UNIVERSITÉ DU QUÉBEC À MONTRÉAL

# ÉVALUATION D'INSTRUMENTS FINANCIERS DE TYPE AMÉRICAIN AVEC DES APPROXIMATIONS PAR CHAÎNES DE MARKOV

THÈSE

PRÉSENTÉE

COMME EXIGENCE PARTIELLE

DU DOCTORAT EN MATHÉMATIQUES

PAR

MARIE-CLAUDE VACHON

FÉVRIER 2025

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

# PRICING AMERICAN-TYPE FINANCIAL INSTRUMENTS WITH MARKOV CHAINS APPROXIMATIONS

THESIS

PRESENTED

AS PARTIAL REQUIREMENT

TO THE DOCTORATE IN MATHEMATICS

BY

MARIE-CLAUDE VACHON

FEBRUARY 2025

## REMERCIEMENTS

Je tiens à exprimer mes plus sincères remerciements à ma directrice, la Dre Anne MacKay, pour sa contribution précieuse à ce projet. Ton sens de la rigueur et tes conseils ont été d'une grande aide tout au long de mes années en recherche. Merci pour le temps que tu as consacré à réviser ce travail et à me guider lorsque j'en avais besoin. Merci pour la liberté que tu m'as accordée pour explorer les sujets qui me passionnent, ainsi que pour ta confiance en mes idées. Enfin, merci pour ton soutien financier qui a été d'une aide précieuse pour mener à bien ce projet.

Un hommage particulier à mon mari et meilleur ami, Roger Ok. Merci de partager avec moi toutes mes folles aventures. Merci de toujours être là pour m'écouter, m'encourager et me soutenir lorsque j'en ai besoin. Merci d'être toi. Je t'aime pour la vie et au-delà. Un remerciement spécial à mes merveilleux enfants, Henri et Juliette. Vous êtes la raison d'être de ce projet. Vous êtes ma lumière, ma source d'inspiration, mon bonheur tous les jours. Je vous aime plus que tout au monde.

Je souhaite exprimer ma profonde gratitude envers ma famille : mes parents, leurs conjoints et mon frère. C'est grâce à vous si je suis devenue la femme que je suis aujourd'hui. Maman, tu es ma meilleure amie et ma confidente. Tu rayonnes de ta belle lumière partout où tu vas. Tu illumines mon chemin. Merci d'être toujours présente pour moi, peu importe les circonstances. Papa, je suis fière de dire que j'ai réussi, même si j'ai souvent douté de moi. Merci pour ton soutien inébranlable et tes encouragements, ils ont été d'une valeur inestimable pour moi. Jean-François, ton enthousiasme et ta positivité sont contagieux. Merci de toujours croire en ta petite sœur.

Ces remerciements ne seraient pas complets sans une pensée pour ma chère tante, Nicole Vachon, qui a révisé plusieurs de mes demandes de bourse au fil des années. Merci d'avoir été présente et d'avoir pris le temps de m'aider peu importe l'heure du jour ou de la nuit. Je t'en serai toujours reconnaissante.

Une pensée spéciale pour le professeur Michel Adès. Merci de partager avec nous votre passion pour les mathématiques et la recherche. Votre dévouement envers vos étudiants est sans limite.

Je tiens aussi à exprimer ma profonde reconnaissance envers le Fonds de recherche du Québec - Nature et technologies pour leur précieuse contribution financière qui a grandement facilité la réalisation de ce

projet. Je remercie également la Faculté des sciences de l'UQAM pour cette belle reconnaissance envers les femmes en science et pour son encouragement envers les femmes ayant un parcours atypique à poursuivre des études supérieures. Merci pour votre engagement envers la diversité et l'égalité des genres dans le domaine des sciences.

Je souhaite adresser une pensée spéciale à mes grandes amies, Iuila Pana et Priscilla Francis. Merci pour votre écoute et vos conseils, vous êtes des amies incroyables. Enfin, je tiens à exprimer ma gratitude envers les clientes, les professeur(e)s et les employées du studio Idolem Yoga à Saint-Eustache. Sans le savoir, chaque jour, vous m'avez inspiré à continuer. Merci du fond du cœur, et namasté.

**CONTENTS**

## LIST OF FIGURES

**LIST OF TABLES**

## ACRONYMS

**CB** convertible bond.

**CBOE** Chicago Board Options Exchange.

**CTMC** continuous-time Markov chain.

**ES** early surrender.

**GMDB** guaranteed minimum death benefit.

**GMMB** guaranteed minimum maturity benefit.

**GMWB** guaranteed minimum withdrawal benefit.

**LSMC** least squares Monte Carlo.

**PDE** partial differential equation.

**TF** Tsiveriotis and Fernandes (1998).

**VA** variable annuity.

**VIX** Chicago Board Options Exchange volatility index.

$\mathbb{R}$ : set of real numbers.

$\mathbb{R}_+$ : set of nonnegative real numbers.

$\mathbb{R}_+^\star$ : set of positive real numbers.

$C^{1,2}((0,T)\times E)$ : collection of all continuous functions $f : (0,T)\times E \to \mathbb{R}$ which have continuous first and the second order partial derivatives, where $E$ denotes an open subset of $\mathbb{R}$.

$X_n \Rightarrow X$ denotes the weak convergence of the sequence of random variables $\{X_n\}_{n=1}^{\infty}$ to the random variable $X$.

$\mathbb{E}_{t,x}[\cdot] = \mathbb{E}[\cdot|X_t = x]$, where $\{X_t\}_{t\geq 0}$ is a stochastic process.

## RÉSUMÉ

Dans cette thèse, nous considérons l'évaluation d'instruments financiers de type américain avec une fonction de gain non bornée, dépendante du temps et discontinue. Ce problème est motivé par l'évaluation de fonds distincts offrant une garantie minimale à l'échéance et de titres de créance tels que les obligations convertibles. Le Chapitre 2 étudie le problème de façon théorique, alors que les Chapitres 3 et 4 présentent des algorithmes efficaces pour l'évaluation de ce type d'instruments financiers. Ces algorithmes sont basés sur une méthode numérique récente qui repose sur les chaînes de Markov à temps continu pour approximer le processus de diffusion sous-jacent.

Plus spécifiquement dans le Chapitre 2, nous effectuons une analyse rigoureuse de la fonction valeur d'un contrat de fonds distinct dans le modèle de Black-Scholes, lorsque les frais et les pénalités de rachat anticipé (ou charges de rachat anticipé) dépendent à la fois du temps et de la valeur du fonds (voir par exemple, Bernard *et al*. (2014a)). Sous l'hypothèse que l'assuré maximise la valeur neutre au risque de son contrat, l'évaluation de fonds distincts équivaut à résoudre un problème d'arrêt optimal similaire à l'évaluation d'option américaine. Cependant, la garantie s'appliquant uniquement à l'échéance du contrat dans un fonds distinct crée une discontinuité dans la fonction de gain à l'échéance, ce qui la distingue des fonctions de gain continues généralement étudiées dans la littérature des options américaines et complique le problème d'arrêt optimal impliqué dans l'évaluation de fonds distinct. En particulier, nous donnons une condition sous laquelle le temps d'arrêt optimal se produit toujours à l'échéance du contrat. En utilisant une représentation alternative de la fonction valeur pour le problème d'optimisation, nous étudions ses propriétés analytiques ainsi que la région de rachat (ou d'exercice) qui en résulte. Nous démontrons que la non-vacuité et la forme de la région de rachat sont entièrement caractérisées par les fonctions de frais et de charge de rachat, fournissant ainsi un outil puissant pour comprendre leur interrelation et comment elles affectent les rachats anticipés et la frontière de rachat optimale. Sous certaines conditions sur ces deux fonctions, nous développons trois représentations de la fonction valeur, deux sont analogues à celle de l'option américaine, alors que l'autre est nouvelle dans la littérature actuarielle et d'évaluation d'options américaines.

Le Chapitre 3 est aussi consacré à l'évaluation de fonds distincts. Dans ce chapitre, une méthode numérique récente basée sur les chaînes de Markov à temps continu est explorée pour l'évaluation de fonds distincts avec une structure de frais générale sous une classe de modèles à volatilité stochastique incluant entre autres les modèles de Heston, Hull-White, Scott, $\alpha$-Hypergeometric, $3/2$, et $4/2$. En particulier, l'impact d'une structure de frais liée à l'indice de volatilité VIX sur la stratégie de rachat optimale d'un contrat de fonds distincts avec une prestation minimale garantie à l'échéance est analysé. En approximant la valeur du fonds par une chaine de Markov à temps continu à deux dimensions, nous développons des algorithmes efficaces. Lorsque le contrat est détenu jusqu'à l'échéance, une expression analytique sous forme matricielle est obtenue pour la valeur du contrat. Nous fournissons également une façon simple et efficace de déterminer la valeur des rachats anticipés à l'aide d'un algorithme récursif et donnons procédure simple pour approximer la surface de rachat optimale.

Le Chapitre 4 étend la méthode numérique explorée dans le Chapitre 3 aux processus de diffusion non homogènes dans le temps pour l'évaluation de titres de créance tels que les obligations, les options sur obligations, les obligations rachetables/rétractables et les obligations convertibles. Plus précisément, des expressions analytiques sous forme matricielle sont obtenues pour approximer le prix d'obligations et d'options sur obligations sous des processus généraux de taux court à une dimension, et un algorithme efficace est

développé pour l'évaluation de dettes rachetables/rétractables. La disponibilité d'une expression analytique pour le prix d'obligations zéro-coupon permet l'ajustement parfait du modèle approximé à la structure à terme actuelle des taux d'intérêt du marché, quelle que soit la complexité du processus de diffusion sous-jacent sélectionné. Nous considérons également l'évaluation d'obligations convertibles sous des processus de diffusion bi-dimensionnels non homogènes dans le temps pour modéliser la dynamique d'une action et des taux courts. Basé sur une approximation à chaine de Markov à temps continu à deux dimensions, un algorithme efficace est développé pour approximer le prix d'obligations convertibles. Lorsque la conversion n'est permise qu'à l'échéance du contrat, une expression analytique sous forme matricielle est obtenue. Des expériences numériques démontrent la précision et l'efficacité de la méthode sur un large éventail de paramètres de modèle et de modèles de taux d'intérêt court.

Mots clés: problèmes d'arrêt optimal, options américaines, méthodes numériques en finance, chaînes de markov à temps continu, volatilité stochastique, taux d'intérêt stochastique, fonds distincts, titres de créance, obligations convertibles.

**ABSTRACT**

In this thesis, we consider the pricing of American-type financial instruments with an unbounded, time-dependent, and discontinuous reward function. This problem is motivated by the pricing of variable annuities with guaranteed minimum maturity benefit and of debt securities such as convertible bonds. Chapter 2 studies the problem from a theoretical perspective, whereas Chapters 3 and 4 present efficient algorithms for the pricing of such financial instruments. These algorithms are based on recently developed numerical techniques that rely on a continuous-time Markov chain (CTMC) to approximate the underlying diffusion process.

Specifically, in chapter 2, we perform a rigorous analytical study of the value of a variable annuity (VA) contract in the Black-Scholes framework, when the fee and the surrender charge are both time and state-dependent (that is, depending on the value of the VA account, see Bernard *et al*. (2014a)). Under the assumption that the policyholder maximizes the risk-neutral value of her contract, the pricing of a variable annuity is equivalent to solving an optimal stopping problem similar to pricing an American option. However, the financial guarantee being applied only at maturity in a VA contract creates a discontinuity in the reward function at maturity, which sets it apart from the continuous reward function of the American option and complicates the optimal stopping problem involved in the pricing of VAs. In particular, we give a condition under which optimal stopping always occurs at maturity. Using an alternative representation for the value function of the optimization problem, we study its analytical properties and the resulting surrender (or exercise) region. We show that the non-emptiness and the shape of the surrender region are fully characterized by the fee and the surrender charge functions, which provides a powerful tool to understand their interrelation and how it affects early surrenders and the optimal surrender boundary. Under certain conditions on these two functions, we develop three representations for the value function; two are analogous to their American option counterpart, and one is new to the actuarial and American option pricing literature.

Chapter 3 is also concerned with the pricing of variable annuities. In this chapter, a recent numerical method based on continuous-time Markov chain approximation is explored for the pricing of VA contracts with general fee structure under a class of time-homogeneous stochastic volatility models which includes the Heston, Hull-White, Scott, $\alpha$-Hypergeometric, $3/2$, and $4/2$ models. In particular, the impact of different Chicago Board Options Exchange volatility index (VIX)-linked fee structures on the optimal surrender strategy of a VA contract with guaranteed minimum maturity benefit (GMMB) is analyzed. Using a two-layer continuous-time Markov chain approximation for the fund value process, efficient algorithms are developed. When the contract is kept until maturity, a closed-form matrix expression is obtained for the value of the VA contract. We also provide a quick and simple way to determine the value of early surrenders via a recursive algorithm and give an easy procedure to approximate the optimal surrender surface.

Chapter 4 extends the numerical methods explored in Chapter 3 to time-inhomogeneous diffusion processes for the pricing of debt securities such as bonds, bond options, callable/putable bonds, and convertible bonds. More precisely, closed-form matrix expressions are obtained to approximate the price of bonds and bond options under general one-dimensional short-rate processes, and a simple efficient algorithm is also developed for the pricing of callable/putable debts. The availability of a closed-form expression for the price of zero-coupon bonds allows for the perfect fit of the approximated model to the current market term structure of interest rate, regardless of the complexity of the underlying diffusion process selected. We further consider the pricing of a convertible bond (CB) under general bi-dimensional time-inhomogeneous

diffusion processes to model equity and short-rate dynamics. Credit risk is also incorporated into the model using the approach of Tsiveriotis and Fernandes (1998). Based on a two-layer CTMC method, an efficient algorithm is developed to approximate the price of convertible bonds. When conversion is only allowed at maturity, a closed-form matrix expression is obtained. Numerical experiments show the accuracy and efficiency of the method across a wide range of model parameters and short-rate models.

Keywords: optimal stopping problems, American options, numerical methods in finance, continuous-time Markov chains, stochastic volatility, stochastic interest rates, variable annuities, debt securities, convertible bonds.

# INTRODUCTION

This thesis is divided into two main parts. The first part, in Chapters 2 and 3, focuses on the pricing of variable annuities, while the second part, in Chapter 4, addresses the pricing of debt securities such as convertible bonds. Chapter 1 establishes the basic framework for valuing these financial securities and highlights the similarities between variable annuity and convertible bond structures. Chapter 2 presents a rigorous analytical study of the pricing formula of a variable annuity contract within the Black-Scholes framework, while Chapter 3 and 4 are concerned with efficient algorithms based on two-layer CTMC approximation for the pricing of these securities. More precisely, Chapter 3 considers the pricing of variable annuities with fee structures tied to the VIX under general stochastic volatility models, and Chapter 4 extends this method to time-inhomogenous bi-dimensional diffusion processes for the pricing of debt securities under general stochastic interest rate models.

Variable annuities are financial products offered by insurance companies mainly used for retirement planning. Typically, policyholders contribute either a single or periodic premium to an investment sub-account (or fund), whose return is linked to a financial portfolio. This makes them comparable to mutual funds, except that variable annuities offer financial guarantees at maturity or death, protecting the investors from market downturns. Various types of protection are available within these products; see, for example, Hardy (2003), Bauer *et al.* (2008), and Feng *et al.* (2022) for details. Common types of investment guarantees include, among others, the guaranteed minimum maturity benefit (GMMB), ensuring a minimum capital amount at contract maturity, and the guaranteed minimum withdrawal benefit (GMWB), providing a minimum level of withdrawals. Typically, these protection riders are financed through a periodic fee generally set as a certain percentage of the fund value. Given their complex structures, variable annuities are subject to different risk factors such as equity, interest rate, and mortality risk. Early surrenders can also lead to liquidity issues. The uncertainty faced by insurers with respect to policyholder surrender behavior is known as surrender risk. Given the significant size of the variable annuity market, appropriately pricing these contracts and understanding optimal surrender behavior is thus essential for risk management purposes, Niittuinperä (2022).

In recent years, the literature has extensively explored the pricing of variable annuities, covering various pricing models and modeling approaches. For a comprehensive review, refer to Feng *et al.* (2022). In certain cases, simplified VA contracts can yield closed-form matrix expressions, as discussed in Hardy (2003)

and Bauer *et al.* (2008). However, incorporating more realistic features like surrender risk to the valuation framework necessitates the use of numerical techniques. Policyholder surrender behavior can significantly affect rider fees and has thus drawn the interest of insurers and researchers. Bacinello *et al.* (2011) propose a general pricing framework under different surrender behavior assumptions (static, dynamic, and mixed), using Monte Carlo simulation. Gao and Ulm (2012) proposes a utility-based approach, whereas Zhu and Welsch (2015) uses novel statistical learning techniques to model withdrawals. Another stream of literature uses the no-arbitrage pricing principle and assumes that policyholders maximize the risk-neutral value of their contract (Grosen and Jørgensen (1997), Milevsky and Salisbury (2001), Bernard *et al.* (2014b), Kang and Ziveyi (2018), among others). It can be argued that it corresponds to the worst-case scenarios for VA providers and provides an upper bound for the prices. Recent studies integrate market frictions like taxes and other expenses into the valuation framework (Bauer *et al.* (2017), Alonso-García *et al.* (2022), Bauer and Moenig (2023)), whereas Moenig and Zhu (2018) and Bernard and Moenig (2019) consider lapse and reentry strategies.

Chapter 2 of this thesis concerns the pricing of variable annuities under rational surrender behavior, meaning that the contract is terminated as soon as it is optimal to do so from a strictly financial perspective. Under this assumption, the pricing of variable annuity corresponds to an optimal stopping problem similar to the pricing of American options. However, the unboundedness and time-dependence of the reward function, combined with its time-discontinuity arising from the difference between the maturity and surrender benefits, sets apart variable annuities from classical American options pricing often studied in the literature (Bassan and Ceci (2002), Bensoussan and Lions (1982), De Angelis and Stabile (2019), Jaillet *et al.* (1990), Krylov (1980), to name a few). The main goal of Chapter 2 is thus to perform a rigorous analytical study of the value function involved in the pricing of variable annuity contracts. The results of this chapter establish a theoretical basis for existing numerical methods already applied in the VA pricing literature. It extends the surrender premium representation of Bernard *et al.* (2014b) to more general fees and surrender charges functions and introduces a new representation of the value function in terms of the current surrender value and an integral term that only takes non-zero value in the continuation region. We called this new representation the "continuation premium". To the author's knowledge, such a representation is new to the actuarial and American option pricing literature, opening the doors for the development of new numerical techniques. Additionally, a detailed theoretical analysis of the shape and the non-emptiness of the surrender region is performed. An explicit condition is derived under which early termination becomes sub-optimal, revealing insights on the interrelation between fees, surrender charges, and surrender risk

and extending the results of Milevsky and Salisbury (2001) to a more general finite time-horizon problem.

Chapters 3 and 4 introduced novel algorithms for pricing variable annuities and financial securities such as convertible debts. The numerical method proposed is based on a two-layer continuous-time Markov chain (CTMC) approximation developed by Cui *et al.* (2018). Building on the ideas of Kushner (1977) and Kushner and DiMasi (1978), the use of CTMC has been developed by Kushner (1990) in control theory for approximating diffusion processes, and subsequently applied to financial mathematics for the pricing of barrier options in one-dimensional Markov models by Mijatović and Pistorius (2013). Then follows the work of Cai *et al.* (2015) for the pricing of Asian options under general one-dimensional Markov processes with jumps, whereas Lo and Skindilias (2014) discussed the calibration of the approximated model and grid designs, and Cai *et al.* (2019) applied the techniques to regimes-switching models. Extension to two-dimensional stochastic volatility models is explored in Chourdakis (2004), and subsequently in Cui *et al.* (2017b), Cui *et al.* (2017c) and Kirkby *et al.* (2017). The idea behind the approach consists of first approximating the variance process by a CTMC, which results in a regime-switching diffusion process. The characteristic function of the equity process is then obtained analytically, and the pricing of the derivative is performed through Fourier techniques. The method applies to a wide class of stochastic volatility models but requires some restrictions on the model parameters for the characteristic function to have an analytical expression. Cui *et al.* (2018) then proposed a two-layer CTMC approximation, which applies to most stochastic volatility models, removing the previous constraint on the diffusion process parameters. The method first approximates the variance process by a CTMC (the first layer). The resulting regime-switching diffusion process is then approximated by a regime-switching CTMC (the second layer) and the technique of Cai *et al.* (2019) is used to map the two-dimensional CTMC process onto a one-dimensional CTMC process on an enlarged state space. The technique is then applied to the pricing of European, Asian, Barrier, Bermudan, and occupation time derivatives under general stochastic volatility models. Ma *et al.* (2020) uses the method for approximating the price of American-type derivatives using an iterative procedure based on the early exercise premium representation of the value function under which the early exercise surface can also be approximated. Ding and Ning (2021) extend the procedure to time-inhomogeneous bi-dimensional diffusion processes, whereas Kirkby *et al.* (2020) explore higher dimension setting. Recently, Kirkby (2023) developed the technique in time-homogeneous stochastic interest rate models and applied CTMC approximations to the pricing of hybrid equity-rate derivatives such as equity swaps.

Various fee designs have been explored in existing literature, Bernard *et al.* (2014a), Bernard and MacKay

(2015), MacKay *et al*. (2017), Bernard and Moenig (2019) and Wang and Zou (2021). Typically, these designs aim to mitigate the exposure of VA providers to different risk factors. Fee structures that are tied to the VIX are inspired directly by industry contracts. Cui (2013) theoretically investigate a particular VIX-linked fee design within the Heston-type stochastic volatility model for the pricing of VA contracts with GMMB rider and highlight its potential in aligning fee incomes with liabilities, whereas Kouritzin and MacKay (2018) analyzed the impact of VIX-linked fees on the sensitivity of the liability of the insurer to volatility risk for VA contracts with GMWB rider. In Chapter 3, we numerically investigate the impact of three different VIX-linked fee structures on the optimal surrender strategy. Fee structures are allowed to be as general as possible, i.e. the fees may depend on the time, the fund value, and the underlying latent variance process. To do so, we use a two-layer CTMC methodology inspired by Cui *et al*. (2018). The approach of Cui *et al*. (2018) is not only theoretically appealing and applies to most stochastic volatility models, but also allows for closed-form matrix expression for the price of European-type derivatives, and simple recursive algorithms can be developed to approximate the price of American-type derivatives. However, for the pricing of long-term financial instruments, such as variable annuities, the method stretches computational resources to unacceptable levels. This chapter addresses this challenge by presenting efficient algorithms for pricing European and American-type derivatives under general two-dimensional models. Additionally, a novel methodology is proposed to approximate the optimal surrender surface. Convergence of the new methodology is also demonstrated theoritically.

Chapter 4 explores CTMC methods for the pricing of debt securities under general stochastic interest rate models. Short-rate models that allow reproducing the term structure of interest rates are time-inhomogeneous, adding to the numerical difficulty of this extension. Indeed, CTMC approximation is better suited for time-homogeneous processes. The reason is that approximating a time-inhomogeneous process by CTMC results in a time-dependent generator (see Ding and Ning (2021)). The theoretical extension of the results of Ding and Ning (2021) to medium/long time horizon derivatives is straightforward but presents numerical challenges since calculations of matrix exponentials are required multiple times to obtain transition probabilities and approximate option prices. Indeed, when the generator is time-dependent, transition probabilities need to be calculated at each time step, adding to the numerical difficulty of valuing medium/long-term derivatives. In Chapter 3, we address this challenge by adapting algorithms developed in Chapter 3 to bi-dimensional time-inhomogeneous processes. Closed-form matrix expressions are obtained for the price of bonds under general time-inhomogeneous short-rate models. Such an extension allows the perfect calibration of the approximated models to the current market term structure of interest rates. Efficient algorithms

are then developed for the pricing of convertible debts under general time-inhomogeneous stochastic interest rate models, and closed-form matrix expressions are obtained when conversion is only allowed at maturity.

# CHAPTER 1

# NOTATION AND SETTING

This thesis covers the pricing of two types of financial instruments: variable annuities and convertible bonds. Although they may seem different in appearance, we observe in this chapter that their structure admits many similarities. Section 1.1 defines their main characteristics and terms, while Section 1.2 describes the market model used in each chapter. Lastly, some basic notions and concepts concerning the pricing of these particular financial instruments are reviewed in Section 1.3.

## 1.1 Financial Instruments

In this section, we describe the main characteristics of variable annuities and convertible bonds. We also introduce the terminology and the notation used in the next chapters.

### 1.1.1 Variable Annuity Contract

In this section, we describe a simplified **variable annuity** (VA) contract offering a **guaranteed minimum maturity benefit** (GMMB) at maturity $T > 0$. At inception of the contract, the policyholder deposits an initial premium in an **investment sub-account** (or **fund**) tracking the financial market. Here, we assume that this sub-account tracks the risky asset with price process $S = \{S_t\}_{0 \leq t \leq T}$. We denote by $F = \{F_t\}_{0 \leq t \leq T}$ the value process of the sub-account.

The guarantees embedded in VA contracts are funded via a fee levied continuously from the investment sub-account at a rate $c = \{c_t\}_{0 \leq t \leq T}$. This **fee rate**, also called *fee process*, can depend on time and the value of the sub-account or other risk factors depending on the context. Generally, we set $c_t := C(t, \cdot)$ where $C$ is called the **fee function**. This is discussed further in the next chapters. The sub-account value is then defined by

$$F_t = S_t e^{-\int_0^t c_u \, \mathrm{d}u}, \quad 0 \leq t \leq T, \tag{1.1}$$

with $F_0 = S_0$.

At maturity, the policyholder receives the maximum between a pre-determined amount $G \in \mathbb{R}_+$ and the value of the investment sub-account. This is defined as the **maturity benefit**. Should she decide to surrender

her contract prior to maturity, the policyholder receives the amount accumulated in the sub-account subject to a penalty charge. This value is called the **surrender benefit** and depends on the **surrender charge**, which is expressed as a certain percentage of the sub-account. If no surrenders occur, the maturity benefit is paid at $T$. We assume that the surrender charge is modeled by a function $g(t, \cdot)$, also called the **surrender charge function**, which can depend on time and other risk factors depending on the context. It is assumed to be non-increasing in time; a later surrender will yield a higher proportion of the account value. In this work, we often use the term surrender charge function to refer to either $g(\cdot, \cdot)$ or $1 - g(\cdot, \cdot)$. Imposing $g(T, \cdot) = 1$ is also common in the actuarial literature. It allows the function $g(\cdot, \cdot)$ to be defined on the closed interval $[0, T]$ and represents the fact that at maturity, the policyholder is entitled to the full amount accumulated in the account. In practice, it is also common to see the surrender charge vanish before maturity. Examples of surrender charges are listed in Palmer (2006).

### 1.1.2 Convertible Bond

**Convertible bonds** (CBs) are hybrid securities that possess features of both debt and equity. They are similar to bonds except that the investor has the right to convert the bond into a predetermined number of shares, called the **conversion ratio** and denoted below by $\eta > 0$, of the issuer during a certain exercise window prior to maturity $T > 0$. This is also referred to as the **conversion value**. At maturity, if conversion is allowed and the bond has not been converted to shares yet, the holder has the right to convert the bond or receive its **face value** $F > 0$. In practice, additional features such as call and put options on the bonds are also generally embedded in CBs.

In this work, we use the term **European-style CB** (resp. **American-style CB**) to refer to a bond under which the conversion option can be exercised at maturity only (resp. at any time prior to maturity).

## 1.2 Market Models

On a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, let $W = \{(W_t^{(1)}, W_t^{(2)})\}_{t \geq 0}$ be a two-dimensional standard Brownian motion, whose augmented filtration is denoted by $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ and let $\mathbb{Q}$ be the risk-neutral probability measure used to price assets.

We consider a financial market consisting of two risk factors, $X = \{X_t\}_{0 \leq t \leq T}$ with $X_t = (X_t^{(1)}, X_t^{(2)})$.

We assume that the dynamics of $X$ under the pricing measure are given by

$$\mathrm{d}X_t^{(i)} = \mu_i(t, X_t)\,\mathrm{d}t + \sum_{j=1}^{2} \sigma_{ij}(t, X_t)\,\mathrm{d}W_t^{(j)}, \quad i \in \{1, 2\}. \tag{1.2}$$

on $[0, T]$ with $T > 0$, the maturity of the financial instrument. Let $\mathcal{S}_X$ be the state space of $X$ and define drift vector by $\mu(t, x) = [\mu_i(t, x)]_{i=1}^2$ with $\mu_i : [0, T] \times \mathcal{S}_X \to \mathbb{R}$ and the diffusion matrix by $\sigma(t, x) = [\sigma_{ij}(t, x)]_{i,j=1}^2$ with $\sigma_{ij} : [0, T] \times \mathcal{S}_X \to \mathbb{R}$. Conditions for (1.2) to be well-defined are discussed in more details in the next chapters.

Each chapter of this thesis focuses on different risk factors relevant to the specific problem being addressed. Chapter 2 examines the VA pricing problem from a theoretical perspective within the Black-Scholes framework, considering only the stock index as a risk factor. In Chapter 3, VIX-linked fee incentives are incorporated into VA pricing, requiring the inclusion of stochastic volatility as a second risk factor. Finally, Chapter 4 addresses the pricing of debt securities, where interest rates play a significant role. Thus, both equity and interest rates are modeled.

### 1.2.1 Black-Scholes Framework

In Chapter 2, a detailed theoretical study of the value of a variable annuity contract under the Black-Scholes setting is performed. Under this market model, the risky asset $\{S_t\}_{0 \le t \le T}$ follows a one-dimensional geometric Brownian motion whose dynamics satisfy

$$\mathrm{d}S_t \;=\; rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t^{(1)}, \tag{1.3}$$

with $r > 0$ the risk-free rate and $\sigma > 0$ the volatility. Hence, using (1.1) and Itô's formula, the dynamics of the fund value process $F$ is given by

$$\mathrm{d}F_t \;=\; (r - c_t)F_t\,\mathrm{d}t + \sigma F_t\,\mathrm{d}W_t^{(1)}, \tag{1.4}$$

where $\{c_t\}_{0 \le t \le T}$ is the fee process.

### 1.2.2 Stochastic Volatility Models

Chapter 3 also studies variable annuities. However, particular fee structures are considered. In this chapter, we numerically study the impact of three different fee structures tied to the VIX on the optimal surrender strategy. Hence, in order to incorporate VIX-linked fee structures into the valuation model, the variance

$\{V_t\}_{0\leq t\leq T}$ of the risky asset process $\{S_t\}_{0\leq t\leq T}$ needs to be modeled as a risk factor. The price of the risky asset is thus described by a two-dimensional process $(S,V)=\{(S_t,V_t)\}_{0\leq t\leq T}$ satisfying

$$\begin{aligned}\mathrm{d}S_t &= rS_t\,\mathrm{d}t+\sigma_S(V_t)S_t\left(\sqrt{1-\rho^2}\,\mathrm{d}W_t^{(1)}+\rho\,\mathrm{d}W_t^{(2)}\right),\\ \mathrm{d}V_t &= \mu_V(V_t)\,\mathrm{d}t+\sigma_V(V_t)\,\mathrm{d}W_t^{(2)},\end{aligned} \tag{1.5}$$

with $S_0=s_0>0$ and $V_0=v_0\in\mathcal{S}_V$ where $\mathcal{S}_V$ denotes the state-space of $V$ (usually $\mathbb{R}$ or $\mathbb{R}_+$ depending on the model), with $r>0$ denoting the risk-free rate and $\rho\in[-1,1]$ the correlation coefficient. The functions $\mu_V:\mathcal{S}_V\to\mathbb{R}$ and $\sigma_S,\,\sigma_V:\mathcal{S}_V\to\mathbb{R}_+$ are assumed to be well-defined such that (1.5) has a unique-in-law weak solution. Other conditions are discussed in Chapter 3.

Hence, using (1.1) and Itô's formula, the dynamics of the fund value process $\{F_t\}_{0\leq t\leq T}$ satisfy

$$\begin{aligned}\mathrm{d}F_t &= (r-c_t)F_t\,\mathrm{d}t+\sigma_S(V_t)F_t\left(\sqrt{1-\rho^2}\,\mathrm{d}W_t^{(1)}+\rho\,\mathrm{d}W_t^{(2)}\right),\\ \mathrm{d}V_t &= \mu_V(V_t)\,\mathrm{d}t+\sigma_V(V_t)\,\mathrm{d}W_t^{(2)},\end{aligned} \tag{1.6}$$

where $\{c_t\}_{0\leq t\leq T}$ is the fee process.

### 1.2.3 Stochastic Interest Rate Models

The last chapter of this thesis concerns the pricing of various debt securities. In particular, we consider the pricing of convertible bonds under general bi-dimensional time-inhomogeneous diffusion processes to model the risky asset $\{S_t\}_{t\geq0}$ and short-rate $\{R_t\}_{t\geq0}$. The price of the risky asset can thus be described by a two-dimensional process $(S,R)=\{(S_t,R_t)\}_{t\geq0}$ satisfying

$$\begin{aligned}\mathrm{d}S_t &= R_tS_t\,\mathrm{d}t+\sigma_S(R_t)S_t\left(\sqrt{1-\rho^2}\,\mathrm{d}W_t^{(1)}+\rho\,\mathrm{d}W_t^{(2)}\right),\\ \mathrm{d}R_t &= \mu_R(t,R_t)\,\mathrm{d}t+\sigma_R(R_t)\,\mathrm{d}W_t^{(2)},\end{aligned} \tag{1.7}$$

with $S_0>0$ and $R_0\in\mathcal{S}_R$, where $\mathcal{S}_R$ denotes the state-space of $R$ (generally $\mathbb{R}$ or $\mathbb{R}_+^\star$ depending on the model), and $\rho\in[-1,1]$ the correlation coefficient. The functions $\mu_R:\mathcal{S}_R\to\mathbb{R}$ and $\sigma_S,\,\sigma_R:\mathcal{S}_R\to\mathbb{R}_+$ are assumed to be well-defined such that (1.7) has a unique-in-law weak solution. Other conditions are discussed in Chapter 4.

## 1.3 Pricing of American-type Financial Instruments

In this thesis, an **American-type financial instrument** (or American-type contract) refers to a financial instrument under which the holder has the right to terminate (or surrender or convert depending on the context) her contract at any time prior to maturity. We describe below their basic concepts and terminology.

### 1.3.1 Reward (or Gain) Function

We consider a particular type of American contract under which the value received upon exercise differs from the maturity benefit. This creates a time discontinuity in the reward (or gain) function, setting them apart from standard American derivatives studied in the literature.

More formally, let $\varphi : [0,T] \times \mathcal{S}_X \to \mathbb{R}_+$ denotes the **reward function** defined by

$$\varphi(t,x) = \begin{cases} f(t,x) & \text{if } t < T, \\ h(x) & \text{if } t = T, \end{cases} \tag{1.8}$$

where $f : [0,T] \times \mathcal{S}_X \to \mathbb{R}_+$ is a $C^{1,2}$ function representing the exercise (or surrender or conversion) value, with $f(T,\cdot)$ being the exercise value at maturity, and $h : \mathcal{S}_X \to \mathbb{R}_+$ is a continuous function representing the maturity benefit. We also suppose that $f(T,x) \leq h(x)$ for all $x \in \mathcal{S}_X$, so that the maturity benefit is always greater or equal to the exercise value at maturity.

**Remark 1.3.1** *Under this assumption, we observe that the reward function can be discontinuous at time $T$ since*

$$\lim_{t \to T^-} \varphi(t,x) = f(T,x) \leq h(x) = \varphi(T,x).$$

In Chapters 2 and 3, variable annuity contracts are considered. The function $f$ represents the surrender value received by the policyholder, while the function $h$ describes the maturity benefit, defined as the maximum between a guaranteed amount $G$ and the value of the investment sub-account at maturity $F_T$.

In Chapter 4, the pricing of convertible debt is considered. Here, we assume that the bond does not pay any coupons and that no other embedded features are included. Under these assumptions, the function $f$ corresponds to the conversion value. At maturity, if the bond has not been converted to shares yet, the holder has the right to convert it or receive its face value $F$. This is modeled by the function $h$ and corresponds to the maturity benefit.

Using this particular notation, we observe that variable annuities and convertible bonds admit a similar structure. However, the pricing of convertible debt can become more complex due to the inclusion of other factors like credit risk, coupons, and additional embedded options such as call and put options. This is discussed further in Chapter 4.

### 1.3.2 Value Function

First, assume that the risk-free rate is constant and denoted by $r > 0$. Under the assumption that the holder of the derivatives maximizes the risk-neutral value of her contract, the time-$t$ value of an American-type derivative with reward function defined in (1.8) is given by

$$v(t,\, x) = \sup_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}_{t,x}[e^{-r(\tau - t)} \varphi(\tau, X_\tau)], \tag{1.9}$$

where $\mathcal{T}_{t,\,T}$ is the set of all stopping times taking value in the interval $[t,\, T]$, and $\mathbb{E}_{t,x}[\cdot]$ is short-hand notation for $\mathbb{E}[\cdot | X_t = x]$. The function $v : [0, T] \times \mathcal{S}_X \to \mathbb{R}_+$ is called the **value function**.

Suppose now that the risk-free rate is stochastic as in Chapter 4 and represented by $X^{(2)}$, the second risk factor in (1.2). Under this assumption, the value function is given by

$$v(t,\, x) = \sup_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}_{t,x}[e^{-\int_t^\tau X_s^{(2)}\, \mathrm{d}s} \varphi(\tau, X_\tau)]. \tag{1.10}$$

Lastly, we say that $\tau_t^x$ is an **optimal stopping time** for (1.10) if $\tau_t^x \in \mathcal{T}_{t,\,T}$ and satisfies

$$v(t,\, x) = \mathbb{E}_{t,x}[e^{-\int_t^{\tau_t^x} X_s^{(2)}\, \mathrm{d}s} \varphi(\tau_t^x, X_{\tau_t^x})]. \tag{1.11}$$

#### 1.3.2.1 Basic Properties of the Value Function

Using the formulation in (1.9) or (1.10), it is straightforward to conclude that

$$v(t,x) \geq f(t,x) \text{ for all } (t,x) \in [0,T) \times \mathcal{S}_X \quad \text{ and } \quad v(T,x) = h(x), \text{ for all } x \in \mathcal{S}_X,$$

since $t \in \mathcal{T}_{t,T}$. The domain of $v$ is typically divided into two regions: the exercise (or surrender or conversion) region $\mathcal{S}$ and the continuation region $\mathcal{C}$, which are defined by

$$\begin{aligned} \mathcal{S} &= \{(t,x) \in [0,T) \times \mathcal{S}_X | v(t,x) = \varphi(t,x)\}, \\ \mathcal{C} &= \{(t,x) \in [0,T) \times \mathcal{S}_X | v(t,x) > \varphi(t,x)\}, \end{aligned}$$

respectively.

#### 1.3.2.2 Other Representations

We now explore common representations of the value function. To facilitate this, we assume a constant risk-free rate denoted by $r > 0$. We also introduce the second-order differential operator $\mathcal{L}$ defined by

$$\mathcal{L} := \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{2}\mu_i(t,x)\frac{\partial}{\partial x_i},$$

where the terms $a_{ij}(t,x)$ are the elements of the matrix $a(t,x) := \sigma(t,x)\sigma^T(t,x)$ given by

$$a_{ij} = \sum_{k=1}^{2}\sigma_{ik}(t,x)\sigma_{jk}(t,x).$$

Free-Boundary Value Problem and Variational Inequalities

Under some conditions on diffusion matrix $\sigma$, the drift vector $\mu$ and the reward function $\varphi$, the relation between partial differential equations and the optimal stopping problem in (1.9) can be established. This result is well-known for standard American option pricing literature when the reward function is continuous or bounded. However, its application to the financial instruments studied in this thesis is not as straightforward. This is discussed further in Chapter 2 in the context of variable annuity pricing under the Black-Scholes setting. More formally, under some regulatory assumptions, it can be shown the value function is a solution of the following **free boundary value problem**

$$\begin{cases} \mathcal{L}v(t,x) + \dfrac{\partial v}{\partial t}(t,x) - rv(t,x) = 0, & (t,x) \in \mathcal{C} \\ v(t,x) > \varphi(t,x), & (t,x) \in \mathcal{C} \\ v(t,\ x) = \varphi(t,x), & (t,x) \in \mathcal{S} \\ v(T,x) = h(x), & x \in \mathcal{S}_X. \end{cases} \tag{1.12}$$

Using the supermartingale properties of the process $\{e^{-rt}v(t,X_t)\}_{0\le t\le T}$, which is also known as the **Snell envelope** of the discounted reward process, we can further demonstrate that

$$\mathcal{L}v(t,x) + \frac{\partial v}{\partial t}(t,x) - rv(t,x) \le 0, \quad (t,x) \in (0,T)\times\mathcal{S}_X. \tag{1.13}$$

Combining (1.12) and (1.13), we can establish the **variational inequality**. This means that the value function $v$ defined in equation (1.9) is a solution to

$$\max\left\{\mathcal{L}v(t,x) + \frac{\partial v}{\partial t}(t,x) - rv(t,x), \varphi(t,x) - v(t,x)\right\} = 0, \tag{1.14}$$

with terminal condition $v(T,x) = h(x)$ (with all the derivatives in the sense of distribution).

Conditions under which (1.12), (1.13) and (1.14) hold are discussed further in Chapter 2 under the Black-Scholes setting.

Early Exercise Premium Representation

Another common characterization of the value function is in terms of the **early exercise premium**, also known as the integral or surrender premium representation in the context of variable annuity pricing. This representation decomposes the value function into two parts: the expected present value of the maturity benefit and a term that only takes a non-zero value when the underlying is in the exercise (or surrender) region. More precisely, under some conditions on the drift and diffusion parameters in (1.2) and the reward function $\varphi$, we can demonstrate that

$$\begin{aligned} v(t,x) = \mathbb{E}_{t,x}\left[e^{-r(T-t)}h(X_T)\right] \\ + \int_t^T \mathbb{E}_{t,x}\left[e^{-r(s-t)}\left(rf(s,\,X_s) - \mathcal{L}f(s,X_s) - f_t(s,\,X_s)\right)\mathbf{1}_{\{(s,X_s)\in\mathcal{S}\}}\right]\,\mathrm{d}s, \end{aligned} \tag{1.15}$$

where $f_t(t,x) = \frac{\partial f}{\partial t}(t,x)$. Note that the term inside the expectation of (1.15) can be calculated explicitly as it solely relies on the exercise value function $f$, which is already determined as part of the model setup in equation (1.8). Establishing (1.15) rigorously in the present context is not trivial. Existing results often require continuity of the reward function. Conditions under which (1.15) is valid for VA pricing are discussed further in Chapter 2 under the Black-Scholes setting.

### 1.3.3 Classical Numerical Techniques

The optimal stopping problem in (1.9) generally does not admit a closed-form analytical solution, unless a trivial optimal stopping time is shown to exist. This is discussed further in Chapter 2. Numerical techniques are thus required to solve the optimization problem. Here we discussed the most commonly used methods, which include finite difference techniques, tree methods, and Monte Carlo simulation.

Methods based on the resolution of partial differential equation (PDE) use (1.12), necessitating certain conditions on $\mu$, $\sigma$, and $\varphi$ to rigorously define the problem. When these are satisfied, finite difference methods are used to approximate the derivative, and the problem is solved iteratively; see, for instance, Hull (2018), Chapter 21.8 for details.

Tree methods and simulation techniques[1], such as the one proposed by Longstaff and Schwartz (2001), are based on a time-discretization of the problem. This is discussed below.

Consider a time partition $0 = t_0 < t_1 < \ldots < t_N = T$ with $t_n = n\Delta$, $n = 0, 1, 2 \ldots, N$ and $\Delta = T/N$ for some $N \in \mathbb{N}$. In this thesis, we use the term **Bermudan-type financial instrument** (or Bermudan contract) to refer to a financial instrument in which the holder has the option to exercise her right prior to maturity on predetermined dates. In the same vein, a financial instrument that can be exercised at maturity only is called a **European-type financial instrument** (or European contract). Note that these terms do not refer to existing contracts, and they are used to simplify explanations. Naturally, as $N \to \infty$, we expect the price of the Bermudan contract to converge to the American-type financial instrument defined in (1.9). This is formalized in Chapter 3.

Define the set $\mathcal{H}_N = \{t_0, t_1, \ldots, t_N\}$. The time-$t$ risk-neutral value of the Bermudan contract with permitted exercise dates $\mathcal{H}_N$ is defined by

$$b_N(t,x) \quad = \quad \sup_{\tau \in \mathcal{T}_N, \tau \geq t} \mathbb{E}_{t,x}[e^{-r(\tau - t)} \varphi(\tau, X_\tau)], \tag{1.16}$$

where $\mathcal{T}_N$ is the set of stopping times taking values in $\mathcal{H}_N$.

We denote the Bermudan contract value process by $B := \{B_n\}_{n=0}^N$, with $B_n := b_N(t_n, X_{t_n})$. Based on the principle of dynamic programming (refer for example to Lamberton (1998), Theorem 10.1.3), it can be shown that the discretized problem admits the following representation:

$$\begin{cases} B_N & = h(X_T) \\ B_n & = \max\left(f(t_n, X_{t_n}), e^{-r\Delta_n} \mathbb{E}_{t_n}[B_{n+1}]\right), \qquad 0 \leq n \leq N-1. \end{cases} \tag{1.17}$$

Therefore, if the process $\{X_{t_n}\}_{n=0}^N$ is approximated correctly, the value of the Bermudan contract can be computed using a backward recursive procedure based on (1.17). The value of the American-type financial

[1] Over the years, several simulation techniques have been developed to value American-type derivatives ranging from random-tree methods (Broadie and Glasserman (1997)) to stochastic mesh techniques (Broadie and Glasserman (2004)) and parametric approximation of early exercise rule (Garcıa (2003)). Other techniques used regression-based methods to approximate the continuation value (Carriere (1996), Longstaff and Schwartz (2001), Tsitsiklis and Van Roy (2001)). The method of Longstaff and Schwartz (2001), also known as least squares Monte Carlo (LSMC), has garnered significant attention in option pricing literature (Clément *et al.* (2002), Stentoft (2004)), with different improvements proposed over time (Fabozzi *et al.* (2017), Auster *et al.* (2022), Wei and Zhu (2022), Xiong *et al.* (2023), Reesor *et al.* (2024), among others). Recent advancements in machine learning and quantum finance have also enabled novel extensions of this technique (Becker *et al.* (2021), Doriguello *et al.* (2021)).

instrument can then be approximated by its Bermudan counterparts using a large number of exercised dates $N$. The main challenge of this technique consists of calculating the **continuation value**, $e^{-r\Delta_n}\mathbb{E}_{t_n}[B_{n+1}]$, at each time step.

When using tree methods, the continuation value can be calculated explicitly since each node is related to subsequent nodes at the previously considered time step. This characteristic makes tree methods very attractive for the pricing of American-type contracts. However, extending these methods to higher-dimensional diffusion processes can be computationally expensive and quickly become inefficient. On the other hand, simulation techniques easily adapt to high-dimension processes, but the calculation of the continuation value requires an additional level of approximation, which makes the methodology usually less efficient than trees. In this work, we provide efficient algorithms for the pricing of American-type financial instruments. The method uses continuous-time Markov chains to approximate the underlying process $X$ and then employs a backward recursive procedure based on (1.17) to approximate the price of the Bermudan contract. The method is very intuitive, similar to tree techniques, and allows for the explicit formulation of the continuation value, which makes the procedure highly efficient. It also adapts to a wide range of diffusion processes in one and two dimensions. Chapter 3 explores this methodology in the context of variable annuity pricing under stochastic volatility models, while Chapter 4 applies the technique to the pricing of debt securities under stochastic interest rate models.

Another common technique uses the exercise premium representation in (1.15) along with a time-discretization of the integral. This has been explored by Kim (1990) and Carr *et al.* (1992) in the context of American call and put options pricing, whereas Bernard *et al.* (2014b) apply the technique to variable annuity pricing. However, establishing equation (1.15) rigorously is non-trivial, and its proof relies on the regularity properties of the value function. This is discussed further in Chapter 2 under the Black-Scholes setting. Extending these results to stochastic volatility or interest rate models presents challenges. This has been done in Lamberton and Terenzi (2019) for standard American-put option pricing under a Heston-type stochastic volatility model, Heston (1993). Additionally, Ma *et al.* (2020) explored this method using continuous-time Markov chain approximation under stochastic volatility models, assuming certain regularity assumptions on the value function. Extending the results developed in Chapter 2 to American-type financial instruments with a discontinuous reward function as described in (1.8) is left as future research.

# CHAPTER 2

# ANALYTICAL STUDY OF A VARIABLE ANNUITY CONTRACT

This chapter is based on the paper "On an Optimal Stopping Problem with a Discontinuous Reward" in collaboration with Dr. Anne MacKay. The research focuses on the pricing of variable annuities within the Black-Scholes setting. Under the assumption that the policyholder maximizes its risk-neutral value, the price of a variable annuity contract corresponds to an optimal stopping problem akin to the pricing of American options. However, the financial guarantee being applied only at maturity in a VA contract creates a discontinuity in the reward function at maturity, which sets it apart from the continuous reward function of the American option often studied in the literature. This unique feature complicates the optimal stopping problem involved in the pricing of variable annuities. The objective of this chapter is thus to perform a rigorous analytical study of the value function associated with variable annuity contracts, considering the challenge posed by the unbounded, time-dependent, and discontinuous reward function.

## 2.1 Introduction

**Variable annuities** (VAs) are structured products sold by insurance companies that are mainly used for retirement planning. An initial premium is deposited in an investment account (or fund) whose return is linked to that of one or more risky assets. They are similar to mutual funds, but they also offer financial guarantees at the end of a pre-determined accumulation period. The embedded guarantees are akin to long-dated options, but they are funded by a periodic fee, generally set as a percentage of the investment account value, rather than being paid upfront. In this chapter, we focus on a GMMB, which provides the policyholder a minimum guaranteed amount at maturity of the contract. Another common feature of variable annuities contracts is that policyholders generally have the right to surrender, or lapse, their contract prior to maturity. When policyholders choose to do so, they receive the value accumulated in the investment account, reduced by a penalty for early surrender (or surrender charge). The uncertainty faced by VA providers with respect to early termination is known as surrender risk. Early surrenders can entail negative consequences such as liquidity issues. Thus, incorporating adequate surrender assumptions in the pricing of VAs is essential for risk management purposes; see Niittuinperä (2022).

Different early surrender modelling approaches have been proposed in the literature (see Bauer *et al.* (2017) and Feng *et al.* (2022) for a review), from the use of utility functions, Gao and Ulm (2012), to modern sta-

tistical techniques, Zhu and Welsch (2015). Early surrenders are also often expressed as a decision taken by policyholders on a strictly rational basis, meaning that the contract is terminated as soon as it is optimal to do so from a financial perspective (see Bernard *et al.* (2014a), Jeon and Kwak (2018), Kang and Ziveyi (2018), Milevsky and Salisbury (2001), and MacKay *et al.* (2023), as discussed in the next chapter, among others). Bauer *et al.* (2017) discusses the impact of other factors on policyholder behavior, such as taxes and expenses. This article gave rise to another stream of literature in which market frictions such as taxation rules are considered (Alonso-García *et al.* (2022), Bauer and Moenig (2023), and Moenig and Bauer (2016)), which helps to explain the discrepancies between market and model fee rates, Moenig and Bauer (2016). Other recent studies include lapse and reentry strategies in their analysis (Bernard and Moenig (2019) and Moenig and Zhu (2018)), whereas Moenig and Zhu (2021) consider a third-party investor to whom the policyholder can sell her contract.

The goal of this chapter is to study the properties of the value function of the VA contract under the assumption that the policyholder maximizes its risk-neutral value, when the fee and the surrender penalty (or charge) are both time and state-dependent (that is, depending on the value of the VA account, see Bernard *et al.* (2014a)). Our setup includes, among others, the constant fee case, the state-dependent fees of Bernard *et al.* (2014a), Delong (2014), and MacKay *et al.* (2017), and the time-dependent fees of Bernard and Moenig (2019) and Kirkby and Aguilar (2023). Under this assumption, valuing a variable annuity contract is equivalent to solving an optimal stopping problem similar to pricing an American option. However, the financial guarantee being applied only at maturity in a VA contract creates a discontinuity in the reward function at maturity, which sets it apart from the continuous reward function of the American option and complicates the optimal stopping problem involved in the pricing of VAs. The impact of the surrender charge and the fee structure on the optimal surrender strategy has been studied in the Black-Scholes framework by Bernard *et al.* (2014a), Bernard and MacKay (2015), Bernard and Moenig (2019), MacKay (2014), MacKay *et al.* (2017), and Moenig and Zhu (2018), and in a more general setting by Kang and Ziveyi (2018), and MacKay *et al.* (2023), as discussed in the next chapter. However, these articles approach the problem from a numerical perspective. To the author's knowledge, it is the first time that such an extensive analytical study of the optimal stopping problem involved in the pricing of a variable annuity contract is performed in the literature.

Recently, Luo and Xing (2021) studied variable annuity contracts under regime-switching volatility models from an optimal stopping perspective. However, they consider a different surrender benefit, which results

in a continuous reward function similar to that of a put option. Chiarolla *et al.* (2022) perform an analysis similar to ours, but for participating policies with minimum rate guarantee and surrender option. Participating policies are akin to variable annuities in that the premium paid by the policyholder tracks a financial portfolio, subject to a minimum rate guarantee. However, the problem they consider is fundamentally different from ours in multiple ways. i) In variable annuity contracts, guarantees are funded via periodic fees set as a percentage of the sub-account value. This creates a discrepancy between the fee amount and the value of the financial guarantee, which becomes an incentive for the policyholder to surrender the policy early (Milevsky and Salisbury (2001)). In participating policies, there is no such ongoing fee; upon termination or at maturity, the policyholder receives the reserve and a given percentage of the surplus, defined as a percentage of the tracking portfolio over the policy reserve (the intrinsic value). ii) In variable annuities, the surrender value differs from that of the maturity payout, which creates a time discontinuity in the reward function at maturity; whereas in participating policies, the intrinsic value is paid at maturity or upon early termination, so that the reward function is independent of time and continuous. iii) In participating policies, the contract is terminated if the underlying portfolio falls below the reserve. There is no such feature in a variable annuity, that is, early termination is at the sole discretion of the policyholder. The resulting optimal stopping problem we study in this chapter is thus very different from the one considered by Chiarolla *et al.* (2022), even if they are motivated by similar insurance products.

The time discontinuity and the unboundedness of the reward function involved in the pricing of variable annuities with GMMB prevent the simple application of results from the American option pricing literature to our problem. For example, to express the value function as the solution to a free-boundary value problem, one needs to establish its continuity. However, continuity of the value function often follows from the continuity of the reward function, which, in our setting, is unbounded and only upper semi-continuous, making the results usually cited in the context of American options inapplicable (see, for instance, Bassan and Ceci (2002), Bensoussan and Lions (1982), De Angelis and Stabile (2019), Jaillet *et al.* (1990), Krylov (1980), and Lamberton (2009)). In this chapter, following the work of Van Moerbeke (1974) and Palczewski and Stettner (2010) (in the context of impulse control problems), we establish continuity of the value function by showing that it admits an alternative representation in terms of a continuous reward function. We can then confirm that it is a solution to a free-boundary value problem and prove further properties, which provides enough regularity to apply a generalized version of Itô's formula to the value process. This allows us to derive various integral representations for the value function, thus extending the results of Bernard *et al.* (2014b) to general time-dependent fee and surrender charge functions and introducing a new decomposition, the

continuation premium representation, to the literature on VA pricing.

The shape and the (non-)emptiness of the exercise region of an American option with a time-homogeneous reward function has first been studied rigorously by Villeneuve (1999). Other authors, such as Jönsson *et al.* (2006), Kotlow (1973), and Villeneuve (2007), also studied the conditions under which the exercise region has a particular shape. These results are however not directly applicable to the problem we study in this chapter, because of the time-dependence and the discontinuity of our reward function. Under the assumption of a general fee and surrender charge structure, we are nonetheless able to identify a condition, expressed as a partial differential inequality, under which the optimal stopping time is always at maturity of the contract, thus mitigating surrender risk. This condition provides a powerful tool to understand the interplay between the fees and surrender charges, and its effect on the optimal stopping strategy. It can also be used to further study and characterize the shape of the optimal exercise, or surrender, region. In particular, we show that the surrender region can be a disconnected set and that the optimal surrender boundary can be discontinuous, thus illustrating a major difference between the pricing of VA contracts and standard American options. Our results extend the work of Milevsky and Salisbury (2001), who consider an infinite horizon, to a more general finite horizon problem.

Our results on the shape of the surrender region also shed light on the link between our original optimal stopping problem and its continuous reward function counterpart. In particular, we obtain conditions under which the two reward functions lead to equivalent optimal stopping problems; that is, we identify cases when they present the same surrender regions and optimal stopping times. While the idea of using an alternate continuous reward process to obtain continuity of the value function is not new to the literature, it is the first time, to the authors' knowledge, that an in-depth comparison of the two problems is presented.

The main contributions of this chapter are listed below.

- We give a condition under which it is never optimal for the policyholder to surrender before maturity of the VA contract.
- We present an alternative representation for the value function of the VA contract in terms of a continuous reward function. As stated above, various results in the literature rely on the continuity of the reward function, in particular those regarding the continuity of the value function. In our chapter, this representation is used in Section 2.4 to perform a rigorous theoretical study of the value function

and to show that the value function of the optimal stopping problem solves a free boundary value problem. This justifies the use of different numerical methods already applied to VA pricing in the actuarial literature.

- We provide two further representations for the value function. For the first one, we decompose the value of the contract into the value of the maturity benefit and an integral term akin to the early exercise premium from the American option literature. The second representation is new to the actuarial and American option pricing literature and expresses the contract value as the sum of the surrender benefit and an integral term coined the continuation premium. These two representations are used in Section 2.4.3 to better understand the value function of a VA contract and in the proofs of Section 2.4.4, which presents characterizations of the (non-) emptiness of the surrender region. The representations could also be used to develop numerical pricing algorithms.

- We obtain conditions under which the original optimal stopping problem and the one with an alternative continuous reward function are equivalent.

This chapter is organized as follows. Section 2.2 presents the optimal stopping problem involved in the pricing of VA contracts with a GMMB. The existence of an optimal stopping time is discussed in Section 2.3. In Section 2.4, we obtain analytical properties of the value function, derive its integral representations and study the shape of the surrender region. We also discuss the equivalence between the original optimal stopping problem and its continuous reward function counterpart. Numerical examples are provided in Section 2.5. Section 2.6 concludes the chapter.

## 2.2 Financial Setting

### 2.2.1 Market Model

On a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, let $W = \{W_t\}_{t\geq 0}$ be a standard Brownian motion whose augmented filtration is denoted by $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$. $\mathbb{Q}$ is the probability measure used to price assets presented below.

We consider a financial market consisting of two primary assets, a risk-free bond $B = \{B_t\}_{t\geq 0}$ and a risky asset $S = \{S_t\}_{t\geq 0}$ whose dynamics under the measure $\mathbb{Q}$ are given by

$$\begin{aligned} \mathrm{d}B_t &= rB_t\,\mathrm{d}t, \\ \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t, \end{aligned} \tag{2.1}$$

where $r$, $\sigma$ are deterministic constants with $\sigma > 0$. It is easy to verify that (2.1) has a unique strong solution given by $B_t = B_0 e^{rt}$ and

$$S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}.$$

Most readers will recognize the so-called Black-Scholes model, presented here directly in terms of its unique risk-neutral measure.

### 2.2.2 Variable Annuity Contract

In this section, we describe a simplified variable annuity contract offering a guaranteed minimum accumulation benefit at maturity $T \in \mathbb{R}_+^\star$. At inception of the contract, the policyholder deposits an initial premium in an investment sub-account (or fund) tracking the financial market. Here, we assume that this sub-account tracks the risky asset with price process $S$. We denote by $F = \{F_t\}_{0\leq t\leq T}$ the value process of the sub-account.

We remark that variable annuity policyholders usually benefit from a return-of-premium guarantee if they die before maturity. Under the assumption that mortality risk is completely diversifiable, it is straightforward to add the death benefit in the analysis of the contract, see for example MacKay *et al.* (2017). For this reason, in this chapter, we ignore mortality risk.

The guarantees embedded in VA contracts are funded via a fee levied continuously from the investment sub-account at a rate $c_t$ defined as

$$c_t := C(t, F_t), \quad 0 \leq t \leq T, \tag{2.2}$$

where $C : [0, T] \times \mathbb{R}_+^\star \to [0, 1]$, so that (2.4) has a unique strong solution. This fee structure is general enough to include state-dependent (Bernard *et al.* (2014b), Delong (2014), MacKay *et al.* (2017)) and time-dependent fees (Bernard and Moenig (2019), Kirkby and Aguilar (2023)). The sub-account value is then defined by

$$F_t = S_t e^{-\int_0^t c_u \,\mathrm{d}u}, \quad 0 \leq t \leq T, \tag{2.3}$$

with $F_0 = S_0$, so that

$$\mathrm{d}F_t = (r - c_t)F_t \,\mathrm{d}t + \sigma F_t \,\mathrm{d}W_t. \tag{2.4}$$

Going forward, we denote by $\{F_s^{t,x}\}_{t\leq s\leq T}$ the solution to (2.4) with starting condition $F_t = x \in \mathbb{R}_+^\star$. When $t = 0$, we simplify the notation and write $F_s^x = F_s^{0,x}$.

At maturity, the policyholder receives the maximum between a pre-determined amount $G \in \mathbb{R}_+^\star$ and the value of the investment sub-account. Given $F_t = x$, the time-$t$ risk-neutral value of the maturity benefit is given by

$$h(t,x) := \mathbb{E}\left[e^{-r(T-t)} \max(G, F_T^{t,x})\right]. \tag{2.5}$$

Should she decide to surrender her contract prior to maturity, the policyholder receives the amount accumulated in the sub-account subject to a penalty charge. We assume that this charge, expressed as a percentage of the sub-account, can depend on time and on the sub-account value. If no surrenders occur, the maturity benefit is paid at $T$.

More formally, let $\varphi : [0,\, T] \times \mathbb{R}_+^\star \to \mathbb{R}_+^\star$ denote the reward (or gain) function defined by

$$\varphi(t,x) = \begin{cases} g(t,x)\, x & \text{if } t < T, \\ \max(G,\, x) & \text{if } t = T, \end{cases} \tag{2.6}$$

where $g : [0,\, T] \times \mathbb{R}_+^\star \to (0,1]$ is in $C^{1,2}$ with bounded first order derivatives on $[0,T] \times \mathbb{R}_+^\star$, non-decreasing in $t$ and satisfies $g(T,x) = 1$ for all $x \in \mathbb{R}_+^\star$. This function represents the amount received by the policyholder upon surrender at $t < T$ or at maturity $T$, given that the sub-account has value $x$.

In practice, $1 - g(\cdot,\cdot)$ is often called the surrender charge. It is non-increasing in time; a later surrender will yield a higher proportion of the account value. In this work, we often use the term **surrender charge function** to refer to either $g(\cdot,\cdot)$ or $1 - g(\cdot,\cdot)$.

Imposing $g(T,\cdot) = 1$ is common in the actuarial literature. It allows the function $g(\cdot,\cdot)$ to be defined on the closed interval $[0,T]$ and represents the fact that at maturity, the policyholder is entitled to the full amount accumulated in the account. In practice, it is also common to see the surrender charge vanish before maturity. Examples of surrender charges are listed in Palmer (2006).

It is the first time, to our knowledge, that state-dependent surrender charges are considered. We will show that state-dependent surrender charges arise naturally when trying to mitigate surrender risk in the presence of a state-dependent fee.

**Remark 2.2.1** *For $x < G$, the function $t \mapsto \varphi(t,x)$ is discontinuous at $T$ since*

$$\lim_{t \to T^-} \varphi(t,x) = g(T,x)x = x < G = \varphi(T,x).$$

Under the assumption that the policyholder maximizes the risk-neutral value of her contract, its time-$t$ value is given by

$$v(t,\, x) = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}[e^{-r(\tau - t)} \varphi(\tau, F_\tau^{t,\,x})] = \sup_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}[e^{-r(\tau - t)} \varphi(\tau, F_\tau^{t,\,x})], \tag{2.7}$$

where $\mathcal{T}_{t,\,T}$ is the set of all stopping times taking value in the interval $[t,\, T]$. Moreover, throughout this chapter, it must be understood that for $s > t$,

$$v(t,\, F_s^{t,x}) = \operatorname*{ess\,sup}_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}[e^{-r(\tau - t)} \varphi(\tau, F_\tau^{t,\,x})],$$

since, in that particular case, the essential supremum does not simplify to an ordinary supremum.

The difference between the value of the full variable annuity contract and the present value of the maturity benefit is the value of the **surrender right**, which we denote by $e : [0, T] \times \mathbb{R}_+^\star \to \mathbb{R}_+$ and define as

$$e(t, x) := v(t, x) - h(t, x).$$

## 2.3 Optimal Stopping Time

In this section, we show that under a simple condition on the fee and the surrender charge function, the optimal stopping problem in (2.7) admits a trivial solution: an optimal stopping time is $T$, the maturity of the contract. In the second part of the present section, we discuss the existence of an optimal stopping time when this condition does not hold.

We say that a stopping time $\tau^\star$ is *optimal* if

$$\begin{aligned} v(t,\, x) &= \sup_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}[e^{-r(\tau - t)} \varphi(\tau, F_\tau^{t,x})] \\ &= \mathbb{E}[e^{-r(\tau^\star - t)} \varphi(\tau^\star, F_{\tau^\star}^{t,x})]. \end{aligned} \tag{2.8}$$

In this case, we also say that $\tau^\star$ is *optimal for* (2.8).

For a general reward function $\varphi$, such an optimal stopping does not necessarily exist. However, under some regularity conditions on the discounted reward process $Z = \{Z_t\}_{0 \le t \le T}$, with $Z_t := e^{-rt} \varphi(t, F_t)$, the existence and the form of an optimal stopping time are well-known from the theory of optimal stopping for random processes in continuous time, see for instance Karatzas and Shreve (1998) (Appendix D, Theorem D.12), Peskir and Shiryaev (2006) (Theorem 2.2), Lamberton (1998) (Section 10.2.1), Lamberton (2009)

(Theorem 2.3.5 and Section 2.3.4), El Karoui (1981), Chapter 2. With the exception of El Karoui (1981), these results rely on the (almost sure) continuity of the reward process, which does not hold for the one involved in variable annuity pricing since its trajectories are discontinuous at time $T$ with positive probability. In fact the trajectories of $Z$ are upper semi-continuous.

### 2.3.1 Surrender Right - Trivial Case

In this section, we derive a simple condition on the fee and surrender charge functions under which it is always optimal for a policyholder maximizing the risk-neutral value of her policy to hold the contract until maturity.

**Lemma 2.1** *Let $g$ and $\varphi$ be defined as in* (2.6)*. If the discounted surrender value process $Y = \{Y_t\}_{0\leq t\leq T}$, with $Y_t := e^{-rt}F_t g(t,\, F_t)$, is a submartingale, then the discounted reward process $Z = \{Z_t\}_{0\leq t\leq T}$, with $Z_t := e^{-rt}\varphi(t, F_t)$, is also a submartingale.*

In the proof below, and in the rest of the chapter when necessary, we use the notation $\mathbb{E}_s[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_s]$ for $s \leq t$.

*Proof. We first observe that for $0 \leq t < T$, $Y_t = Z_t$ so that if $Y$ is a submartingale, then for any $0 \leq s \leq t < T$, $\mathbb{E}_s[Z_t] \geq Z_s$.*

*For $0 \leq s < t = T$, since $g(T, x) = 1$ for all $x \in \mathbb{R}_+^\star$, we have*

$$\mathbb{E}_s[e^{-rT}\varphi(T,\, F_T)] \geq \mathbb{E}_s[e^{-rT}F_T g(T,\, F_T)] \geq e^{-rs}F_s g(s,\, F_s) = e^{-rs}\varphi(s,\, F_s).$$

□

The next result is well-known in optimal stopping theory (see for example Björk (2009), Proposition 21.2) and is reproduced here for completeness. It states that if the discounted reward process is a submartingale, the maturity date of the contract is an optimal stopping time. There is a simple financial interpretation to the above statement. If the discounted reward process $Z$ is a submartingale, then it is non-decreasing on average over time. Hence, it is optimal to hold the contract as long as possible because its value is expected to increase. This is also the reasoning behind the optimal exercise strategy for an American call option on a non-dividend-paying stock.

**Lemma 2.2** *If the discounted reward process $\{Z_t\}_{0\leq t\leq T}$, with $Z_t := e^{-rt}\varphi(t, F_t)$, is a submartingale, then $T$ is an optimal stopping time, that is*

$$v(t,\, x) = \sup_{\tau\in\mathcal{T}_{t,\,T}} \mathbb{E}[e^{-r(\tau-t)}\varphi(\tau, F_\tau^{t,\,x})] = \mathbb{E}[e^{-r(T-t)}\varphi(T, F_T^{t,\,x})].$$

In the following, we denote $g_x(t,x) = \frac{\partial g}{\partial x}(t,x)$, $g_{xx}(t,x) = \frac{\partial^2 g}{\partial x^2}(t,x)$ and $g_t(t,x) = \frac{\partial g}{\partial t}(t,x)$.

**Proposition 2.3** *Let $C$ and $g$ be the fee and the surrender charge functions as defined in* (2.2) *and* (2.6)*, respectively. If*

$$g_t(t,\, x) + (r - C(t,\, x) + \sigma^2)xg_x(t,\, x) + \frac{\sigma^2x^2}{2}g_{xx}(t,\, x) - C(t,x)g(t,\, x) \geq 0, \tag{2.9}$$

*holds for all $(t,\, x) \in [0,T)\times\mathbb{R}_+^\star$, then $T$ is an optimal stopping time for* (2.7)*.*

*Proof. An application of Itô's lemma to the discounted surrender value process $Y = \{Y_t\}_{0\leq t\leq T}$, with $Y_t := e^{-rt}F_tg(t,\, F_t)$, yields*

$$\begin{aligned}\mathrm{d}Y_t = e^{-rt}F_t\Big(g_t(t,\, F_t) + (r - C(t,\, F_t) + \sigma^2)F_tg_x(t,\, F_t) + \frac{\sigma^2F_t^2}{2}g_{xx}(t,\, F_t)\\ - C(t,\, F_t)g(t,\, F_t)\Big)\,\mathrm{d}t + e^{-rt}\sigma F_t\Big(g(t,F_t) + g_x(t,F_t)F_t\Big)\,\mathrm{d}W_t.\end{aligned}$$

*Thus, if*

$$g_t(t,\, x) + (r - C(t,\, x) + \sigma^2)xg_x(t,\, x) + \frac{\sigma^2x^2}{2}g_{xx}(t,\, x) - C(t,x)g(t,\, x) \geq 0,$$

*for all $(t,x)\in[0,T)\times\mathbb{R}_+^\star$, then $Y$ is a submartingale and so is the discounted reward process by Lemma 2.1. The optimal stopping time follows from Lemma 2.2.* □

We now present applications of Proposition 2.3.

**Example 2.3.1** *Let the fee rate be constant, that is $C(t,x) = c > 0$, and let $g(t,x) = e^{-k(T-t)}$ for some $k > 0$. To eliminate the incentive to surrender, it suffices to choose $k$ and $c$ such that* (2.9) *holds; a quick calculation yields $k \geq c$. Therefore, it suffices to set the surrender rate $k$ at or above the fee rate $c$ to eliminate surrender risk. This result is well-known and discussed in Bernard and MacKay (2015), Proposition 3.1 and MacKay (2014), Proposition 4.4.2. In particular, MacKay (2014) shows that in the Black-Scholes setting, $g(t,x) = e^{-c(T-t)}$ is the maximal surrender charge function that can be used to eliminate the surrender incentive when the fee rate is constant.*

**Example 2.3.2** *Let $g(t,x) = 1 - k\left(1 - \frac{t}{T}\right)^3$ for some constant $k > 0$, as in Bacinello and Zoccolan (2019), MacKay (2014) Section 4.3.2., and MacKay* et al. *(2017). As explained in MacKay (2014), this form of surrender charge function mimics the surrender charges in the market, which are usually high in the first years of the contract and drop drastically thereafter. Assume also that the fee function only depends on time, such that $C(t,\cdot) = c(t)$. In this example, we do not specify a particular form for the fee function;* (2.9) *is used to obtain the function $c(t)$ eliminating the surrender incentive. Simple calculations show that* (2.9) *is satisfied if*

$$c(t) \leq \frac{\frac{3k}{T}\left(1 - t/T\right)^2}{1 - k(1 - t/T)^3}. \tag{2.10}$$

*Thus, any function $c(t)$ satisfying this inequality above will eliminate the surrender incentive (from a risk-neutral value maximization perspective). We may simply set $c(t) = \frac{3k/T(1-t/T)^2}{1-k(1-t/T)^3}$. This example highlights the interplay between the fee and the surrender charge structures. Under this setting, it would have been impossible to satisfy* (2.9) *using a constant fee rate, except with $c = 0$, as the function on the right-hand side of* (2.10) *vanishes at maturity. Having a zero fee rate throughout the life of the VA contract is not a feasible solution from the insurer's perspective since the guarantee at maturity must be financed. On the other hand, isolating $k$ in* (2.10) *gives*

$$k \geq \frac{c(t)}{3/T\left(1 - t/T\right)^2 + c(t)(1 - t/T)^3}. \tag{2.11}$$

*We observe that the term on the right-hand side of the inequality goes to infinity as $t$ approaches $T$ if $c(t)$ is kept constant over time. This is obviously not a feasible solution either.*

The examples presented above illustrate the interplay between fees, surrender charges, and surrender incentives. In order to eliminate early surrender incentives, surrender charges and fees must be structured conjointly; that is, a time-dependent fee structure should be paired with a time-dependent surrender charge to satisfy (2.9).

**Remark 2.3.1** *The results of Section 2.3.1 can easily be extended to more general market models and fee structures, such as the VIX-linked fee discussed in Cui* et al. *(2017a), Kouritzin and MacKay (2018), and Chapter 3 of this thesis.*

### 2.3.2 Surrender Right - Non-Trivial Case

In this section, we study the existence of an optimal stopping time for the variable annuity contract when (2.9) is not satisfied. To do so, the definitions below are required.

**Definition 2.3.1** *A process $X = \{X_t\}_{t\geq 0}$ defined on a filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\}$ is said to be optional if it is measurable with respect to the sigma-algebra generated by the right-continuous and adapted processes.*

**Definition 2.3.2** *A right-continuous adapted process $X = \{X_t\}_{t\geq 0}$ defined on a filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\}$ is said to be of class (D) if the family $\{X_\tau, \tau \in \mathcal{T}_{0,\infty}\}$ is uniformly integrable.*

The next theorem is from Theorem 19 of Bassan and Ceci (2002), which summarizes concisely the results of El Karoui (1981). An advised reader will notice differences between the conditions stated in Theorem 19 of Bassan and Ceci (2002) and the ones of Theorem 2.4. This is because Bassan and Ceci (2002) work with bounded reward functions so the class (D) condition of El Karoui (1981) is automatically satisfied in their setting.

**Theorem 2.4 (El Karoui (1981), Theorems 2.28, 2.31 and 2.41, see also Bassan and Ceci (2002), Theorem 19)** *Let the discounted reward process $\{Z_t\}_{t\geq 0}$ defined on some filtered probability space $\{\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P}\}$ be optional, non-negative and of class (D), and let $\{J_t\}_{t\geq 0}$ denote its Snell envelope, that is*

$$J_t = \operatorname*{ess\,sup}_{\tau \geq t} \mathbb{E}[Z_\tau | \mathcal{F}_t].$$

*Then,*

*(i) $J$ is the smallest non-negative supermartingale which dominates $Z$ (El Karoui (1981), Theorem 2.28).*

*(ii) A stopping time $\tau^\star \in \mathcal{T}_{0,\infty}$ is optimal if and only if $J_{\tau^\star} = Z_{\tau^\star}$ a.s. and $\{J_{\tau^\star \wedge t}\}_{t\geq 0}$ is a martingale (El Karoui (1981), Theorem 2.31).*

*(iii) If, in addition, the trajectories of $Z$ are upper semi-continuous, then*

$$\tau^\star = \inf\{t \geq 0 | J_t = Z_t\},$$

*is the smallest optimal stopping time (El Karoui (1981), Theorem 2.41).*

In light of Theorem 2.4, proving the existence of an optimal stopping time is simply a matter of checking that the discounted reward process is positive, optional and of class (D), and has upper semi-continuous trajectories.

**Corollary 2.5** *Let $\varphi$ and $v$ be defined by* (2.6) *and* (2.7)*, respectively. The stopping time given by*

$$\tau_t^x = \inf\left\{t \leq s \leq T, \,\middle|\, v(s, F_s^{t,x}) = \varphi(s, F_s^{t,x})\right\}. \tag{2.12}$$

*is optimal for* (2.7)*.*

*Proof. The discounted reward process $Z = \{Z_t\}_{0\leq t\leq T}$, with $Z_t = e^{-rt}\varphi(t, F_t)$, is positive and adapted with upper semi-continuous trajectories. Moreover, it is dominated by an integrable non-negative random variable, $\sup_{0\leq t\leq T} e^{-rt}\varphi(t, F_t)$*[1]*, so it is of class (D). The result then follows from Theorem 2.4.* □

The next theorem, inspired by Palczewski and Stettner (2010), provides an alternative representation for the optimal stopping problem in (2.7) and will be essential to prove the continuity of the value function.

**Theorem 2.6** *The value function $v$ defined in* (2.7) *can be written as*

$$v(t,x) = \sup_{\tau\in\mathcal{T}_{t,\,T}} \mathbb{E}\left[e^{-r(\tau-t)}\left[g(\tau, F_\tau^{t,x})F_\tau^{t,x} \vee h(\tau,\, F_\tau^{t,x})\right]\right], \tag{2.13}$$

*where*

$$h(s,\, x) = \mathbb{E}\left[e^{-r(T-s)}\max(G, F_T^{s,x})\right], \quad s \in [0,T].$$

*Proof. Without loss of generality, let $t = 0$ and note that for $0 \leq s \leq T$,*

$$h(s, F_s) = \mathbb{E}\left[e^{-r(T-s)}\max(G, F_T)|\mathcal{F}_s\right].$$

[1] To clarify on the integrability of $\sup_{0\leq t\leq T} e^{-rt}\varphi(t, F_t)$, first note that $\mathbb{E}[\sup_{0\leq t\leq T} e^{-rt}\varphi(t, F_t)] \leq G + \mathbb{E}[\sup_{0\leq t\leq T} e^{-rt}F_t]$. Then, using Doob's maximal inequality, we obtain that $\mathbb{E}[\sup_{0\leq t\leq T} e^{-rt}F_t] < \infty$. It follows that $\mathbb{E}[\sup_{0\leq t\leq T} e^{-rt}\varphi(t, F_t)] < \infty$.

*We define the Snell envelope of the discounted continuous reward process* $\{\widetilde{J}_s\}_{0\leq s\leq T}$ *by*

$$\widetilde{J}_s = \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{s,T}} \mathbb{E}\left[e^{-r\tau}\left(g(\tau,F_\tau)F_\tau \vee h(\tau,F_\tau)\right) \mid \mathcal{F}_s\right]$$

*for* $0 \leq s \leq T$*, and recall that the Snell envelope of the discounted original reward process* $\{J_s\}_{0\leq s\leq T}$ *is defined by*

$$J_s = \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{s,T}} \mathbb{E}\left[e^{-r\tau}\varphi(\tau,F_\tau) \mid \mathcal{F}_s\right].$$

*We show that* $\widetilde{J}_s = J_s$ *a.s. for all* $0 \leq s \leq T$*.*

*First observe that since* $g(t,F_t)F_t \vee h(t,F_t) \geq \varphi(t,F_t)$ *for all* $0 \leq t \leq T$*,* $\widetilde{J}_s \geq J_s$ *for all* $0 \leq s \leq T$*.*

*Now we show that* $\widetilde{J}_s \leq J_s$ *for all* $0 \leq s \leq T$*. By definition of* $v(t,F_t)$*,* $g(t,F_t)F_t \vee h(t,F_t) \leq v(t,F_t)$ *for all* $0 \leq t \leq T$*. It follows that*

$$\widetilde{J}_s \leq \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{s,T}} \mathbb{E}\left[e^{-r\tau}v(\tau,F_\tau) \mid \mathcal{F}_s\right] = \operatorname*{ess\,sup}_{\tau\in\mathcal{T}_{s,T}} \mathbb{E}\left[e^{-r\tau}J_\tau \mid \mathcal{F}_s\right].$$

*Since* $\{J_s\}_{0\leq s\leq T}$ *is a supermartingale, by the optimal sampling theorem,* $\mathbb{E}[J_\tau \mid \mathcal{F}_s] \leq J_s$ *for any* $\tau \in \mathcal{T}_{s,T}$*, and thus* $\operatorname{ess\,sup}_{\tau\in\mathcal{T}_{s,T}} \mathbb{E}\left[e^{-r\tau}J_\tau \mid \mathcal{F}_s\right] = J_s$*, which completes the proof.* □

**Remark 2.3.2** *The modified reward function* $g(t,x)x \vee h(t,\,x)$ *is continuous since it is the maximum of two continuous functions. Indeed,* $g(t,x)x$ *is continuous by definition while the continuity of* $h(t,\,x)$ *follows from Theorem 3 of Veretennikov (1981), which is the analog of the Feynman-Kac Theorem (see, for instance, Karatzas and Shreve (1991), Theorem 5.7.6).*

Similarly to Palczewski and Stettner (2010), we use the new representation of the value function in (2.13) to construct another optimal stopping time for the original problem in (2.7). Define

$$\widetilde{\tau}_t^x := \inf\left\{t \leq s \leq T | v(s,F_s^{t,x}) = g(s,F_s^{t,x})F_s^{t,x} \vee h(s,\,F_s^{t,x})\right\}, \tag{2.14}$$

the smallest optimal stopping time for (2.13), as per Theorem 2.4.

**Corollary 2.7** *Let* $h$ *and* $g$ *be defined as in* (2.5) *and* (2.6)*, respectively. The stopping time defined by*

$$\bar{\tau}_t^x = \begin{cases} \widetilde{\tau}_t^x, & g(\widetilde{\tau}_t^x, F_{\widetilde{\tau}_t^x}^{t,x})F_{\widetilde{\tau}_t^x}^{t,x} \geq h(\widetilde{\tau}_t^x, F_{\widetilde{\tau}_t^x}^{t,x}), \\ T, & g(\widetilde{\tau}_t^x, F_{\widetilde{\tau}_t^x}^{t,x})F_{\widetilde{\tau}_t^x}^{t,x} < h(\widetilde{\tau}_t^x, F_{\widetilde{\tau}_t^x}^{t,x}), \end{cases} \tag{2.15}$$

*is optimal for* (2.7)*.*

*Proof. To keep the notation lighter, we write $\bar{\tau} = \bar{\tau}_t^x$ and $\widetilde{\tau} = \widetilde{\tau}_t^x$. By Theorems 2.4 and 2.6, we have*

$$\begin{aligned}
v(t,x) &= \sup_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}\left[e^{-r(\tau-t)}\left[g(\tau, F_\tau^{t,x})F_\tau^{t,x} \vee h(\tau,\, F_\tau^{t,x})\right]\right] \\
&= \mathbb{E}\left[e^{-r(\widetilde{\tau}-t)}\left[g(\widetilde{\tau}, F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x} \vee h(\widetilde{\tau},\, F_{\widetilde{\tau}}^{t,x})\right]\right] \\
&= \mathbb{E}\left[e^{-r(\widetilde{\tau}-t)} g(\widetilde{\tau}, F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x}\mathbf{1}_{\{g(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x} \geq h(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})\}}\right. \\
&\qquad \left. +\, e^{-r(\widetilde{\tau}-t)} h(\widetilde{\tau},\, F_{\widetilde{\tau}}^{t,x})\mathbf{1}_{\{h(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x}) > g(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x}\}}\right] \\
&= \mathbb{E}\left[e^{-r(\widetilde{\tau}-t)} g(\widetilde{\tau}, F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x}\mathbf{1}_{\{g(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x} \geq h(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})\}}\right] \\
&\qquad + \mathbb{E}\left[\mathbb{E}\left[e^{-r(T-t)}\max(G, F_T^{t,x})\mathbf{1}_{\{h(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x}) > g(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x}\}}\middle| \mathcal{F}_{\widetilde{\tau}}\right]\right] \\
&= \mathbb{E}\left[e^{-r(\widetilde{\tau}-t)} g(\widetilde{\tau}, F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x}\mathbf{1}_{\{g(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x} \geq h(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})\}}\right] \\
&\qquad + \mathbb{E}\left[e^{-r(T-t)}\max(G, F_T^{t,x})\mathbf{1}_{\{h(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x}) > g(\widetilde{\tau},F_{\widetilde{\tau}}^{t,x})F_{\widetilde{\tau}}^{t,x}\}}\right] \\
&= \mathbb{E}\left[e^{-r(\bar{\tau}-t)}\varphi(\bar{\tau}, F_{\bar{\tau}}^{t,x})\right].
\end{aligned}$$

*This proves that $\bar{\tau}_t^x$ is optimal for* (2.7). □

It is possible to show that $\bar{\tau}_t^x = \tau_t^x$ a.s., where $\tau_t^x$ is the smallest optimal stopping time for the original problem defined in Corollary 2.5. A comparison of the three stopping time is given in lemma 2.8 below.

**Lemma 2.8** *For $0 \leq t \leq T$, the optimal stopping times defined in* (2.12), (2.14) *and* (2.15) *satisfy $\widetilde{\tau}_t^x \leq \tau_t^x$ a.s. and $\tau_t^x = \bar{\tau}_t^x$ a.s. Moreover, if $\tau_t^x < T$ a.s., then $\widetilde{\tau}_t^x = \tau_t^x$ a.s.*

*Proof. For $t = T$, the result trivially holds with equality. For the rest of the proof, let $t < T$.*

*To show $\widetilde{\tau}_t^x \leq \tau_t^x$, let $F_t = x$ and $v(t,x) > g(t,x)x \vee h(t,x) \geq \varphi(t,x)$. Then, for any $s \in [t,T]$, $g(s, F_s^{t,x})F_s^{t,x} \vee h(s,\, F_s^{t,x}) \geq \varphi(s, F_s^{t,x})$ a.s., so, the process $v(s, F_s)$ must first cross the continuous reward process $g(s,F_s)F_s \vee h(s,F_s)$ to attain the discontinuous reward process $\varphi(s,F_s)$ (since the process $v(s,F_s)$ starts above the two reward processes at $t$). It follows that $\widetilde{\tau}_t^x \leq \tau_t^x$. The other cases are trivial. Indeed, if $v(t,x) = h(t,x)$ then necessarily $h(t,x) \geq xg(t,x)$, since by the continuous reward representation $v(t,x) \geq h(t,x) \vee xg(t,x)$, so that $\widetilde{\tau}_t^x = t$, and $\tau_t^x \geq t$, and the first inequality holds. The last case is when $v(t,x) = xg(t,x)$, which automatically implies $\widetilde{\tau}_t^x = \tau_t^x = t$.*

*To show that $\tau_t^x = \bar{\tau}_t^x$, we fix $\omega \in \Omega$ and consider two cases: $\tau_t^x(\omega) < T$ and $\tau_t^x(\omega) = T$.*

***Case 1*** *– $\tau_t^x(\omega) < T$: By the definition of $\tau_t^x$ and the gain function $\varphi$, we have $v(\tau_t^x(\omega), F^{t,x}_{\tau_t^x(\omega)}(\omega)) = g(\tau_t^x(\omega), F^{t,x}_{\tau_t^x(\omega)}(\omega))F^{t,x}_{\tau_t^x(\omega)}(\omega)$, and note that if $\widetilde{\tau}_t^x(\omega) = \tau_t^x(\omega)$, then it follows that $\widetilde{\tau}_t^x(\omega) = \tau_t^x(\omega) = \bar{\tau}_t^x(\omega)$. Now observe that from the definition of $\tau_t^x(\omega)$, we know that $v(s, F_s^{t,x}(\omega)) > g(s, F_s^{t,x}(\omega))F_s^{t,x}(\omega)$ for all $t \leq s \leq \tau_t^x(\omega)$. Thus, for $\widetilde{\tau}_t^x(\omega)$ to be equal to $\tau_t^x(\omega)$, we need $v(s, F_s^{t,x}(\omega)) > h(s, F_s^{t,x}(\omega))$, for all $t \leq s \leq \tau_t^x(\omega)$. Now, suppose there exists some $s$ in the interval $[t, \tau_t^x(\omega)]$ such that $v(s, F_s^{t,x}(\omega)) = h(s, F_s^{t,x}(\omega))$. In this case, the two stopping times $\tau_t^x(\omega)$ and $\bar{\tau}_t^x(\omega)$ would differ. However, the equality $v(s, F_s^{t,x}(\omega)) = h(s, F_s^{t,x}(\omega))$ implies that the contract is exercised before maturity with probability zero, which contradicts our initial assumption that $\tau_t^x(\omega) < T$. Therefore, we must have $v(s, F_s^{t,x}(\omega)) > h(s, F_s^{t,x}(\omega))$ for all $t \leq s \leq \tau_t^x(\omega)$, which leads to the conclusion that $\widetilde{\tau}_t^x(\omega) = \tau_t^x(\omega) = \bar{\tau}_t^x(\omega)$.*

***Case 2*** *– $\tau_t^x(\omega) = T$: Note that $\tau_t^x(\omega) = T$ implies*

$$v(s, F_s^{t,x}(\omega)) > g(s, F_s^{t,x}(\omega))F_s^{t,x}(\omega)$$

*for all $t \leq s \leq T$. This means that if $\widetilde{\tau}_t^x(\omega) < T$, $v(\widetilde{\tau}_t^x(\omega), F_{\widetilde{\tau}_t^x(\omega)}(\omega)) = h(\widetilde{\tau}_t^x(\omega), F_{\widetilde{\tau}_t^x(\omega)}(\omega))$. Thus, $\bar{\tau}_t^x(\omega) = \tau_t^x(\omega)$.*

*The final equality ($\widetilde{\tau}_t^x = \tau_t^x$ a.s when $\tau_t^x(\omega) < T$a.s.) follows from the arguments in Case 1. This completes the proof.* □

A simple example involving these optimal stopping times is when condition (2.9) of Proposition 2.3 is satisfied.

**Example 2.3.3** *Suppose that the fee and the surrender charge functions defined in* (2.2) *and* (2.6)*, respectively, satisfy* (2.9)*. Hence, by Proposition 2.3, an optimal stopping time for the problem with the discontinuous reward function is $\tau_0^x = T$, so that*

$$v(0, x) = \mathbb{E}[e^{-rT} \max(G, F_T^x)] = h(0, x).$$

*Clearly, an optimal stopping time for the problem with the continuous reward function is $\widetilde{\tau}_0^x = 0$ since $h(0, x) = \mathbb{E}[e^{-rT} \max(G, F_T^x)] = v(0, x)$; whereas $\bar{\tau}_0^x = T$ as per* (2.15)*.*

In the next section, we present a simple condition on the fee and surrender charge functions under which $\widetilde{\tau}_t^x = \tau_t^x$ a.s. This result is particularly interesting since it provides a condition under which the two optimal

stopping problems defined in (2.7) and (2.13) are equivalent. That is, the two problems have the same value function (Theorem 2.6), surrender region (Proposition 2.25), and optimal stopping time (Corollary 2.27).

## 2.4 Analytical study of the value function

In Section 2.4.1 we study the regularity of the value function of the original problem defined in (2.7). In Section 2.4.2, we establish the relationship between $v$, a free-boundary value problem, and a variational inequality, analogously to American option pricing. To do so, we follow the reasoning of Jacka (1991) and Lamberton (1998) in the context of American put option pricing. This allows us to derive, in Section 2.4.3, two other representations for the value function: the surrender premium representation which is analogous to the exercise premium representation in the American option pricing terminology, and the continuation premium representation. This representation is new to the literature on VA and American option pricing and allows us to characterize the (non-)emptiness of the surrender region in Section 2.4.4. Section 2.4.5 presents a condition on the fee and the surrender charge functions under which the optimal stopping problem with discontinuous reward function defined in (2.7) is equivalent to the one with the continuous reward function, in (2.13).

Assumption 2.4.1 concerns the fee function and how it impacts the drift term in the account value process (2.4), which we denote by $\mu : [0, T] \times \mathbb{R}_+^\star \to \mathbb{R}$, with $\mu(t, x) = (r - C(t, x))x$.

**Assumption 2.4.1** *(i) The fee function $C$ is such that $\mu$ is continuous and globally Lipschitz in $x$, that is, there exists $K \geq 0$ such that for all $x, y \in \mathbb{R}_+^\star$, $t \geq 0$,*

$$|\mu(t, x) - \mu(t, y)| \leq K|x - y|.$$

*(ii) For all $x$, the fee function $C(\cdot, x)$ is locally Hölder continuous.*

**Assumption 2.4.2** *The fee and the surrender charge functions only depend on time; that is, we suppose that $C(t, x) = c(t)$ and $g(t, x) = g(t)$ for all $x \in \mathbb{R}_+^\star$.*

**Remark 2.4.1** *Many results in this section require Assumptions 2.4.1 and 2.4.2 to hold. However, note that*

*as soon as Assumption 2.4.2 holds, the drift term in* (2.4) *is globally Lipschitz in* $x$*, so Assumption 2.4.1(i) is automatically satisfied.*

We define the **surrender region** and the **continuation region** by

$$\mathcal{S} = \left\{(t,x) \in [0,T) \times \mathbb{R}_+^\star | v(t,x) = \varphi(t,x)\right\}, \text{ and} \tag{2.16}$$

$$\mathcal{C} = \{(t,x) \in [0,\, T) \times \mathbb{R}_+^\star : v(t,x) > \varphi(t,x)\}, \tag{2.17}$$

respectively. It follows that $\mathcal{C} \cup \mathcal{S} = [0,T) \times \mathbb{R}_+^\star$, since $v(t,x) \geq \varphi(t,x)$ for all $(t,x) \in [0,T] \times \mathbb{R}_+^\star$.

### 2.4.1 Elementary Properties of the Value Function

**Theorem 2.9** *If Assumption 2.4.1 (i) holds, the value function* $v$ *is continuous on* $[0,T] \times \mathbb{R}_+^\star$.

*Proof. The continuity of the reward function implies the continuity of the value function (see Krylov (1980), Theorem 3.1.8, which requires Lipschitz continuity of the drift term in* (2.4)*, hence the need for Assumption 2.4.1(i)). Thus, the assertion follows from the continuous reward representation of the value function in Theorem 2.6 and Remark 2.3.2.* □

Some basic properties of the value function, such as local boundedness, are derived in the next lemma.

**Lemma 2.10** *For every* $(t,\, x) \in E := [0,\, T] \times \mathbb{R}_+^\star$*, the value function* $v$ *satisfies the following properties:*

*(i)* $Ge^{-r(T-t)} \leq v(t,\, x) \leq G + x$;

*(ii)* $v(t,x) \geq \varphi(t,x)$;

*(iii)* $v(T,x) = \max(G,x)$.

The first assertion follows from optional sampling since $\{e^{-r(u-t)}F_u^{t,x}\}_{u=t}^T$ is a supermartingale. Assertions (ii) and (iii) follow easily from (2.7).

**Lemma 2.11** *Under Assumption 2.4.2, the function $x \mapsto v(t,\, x)$ is non-decreasing and convex for all $t \in [0, T]$.*

The proof is a direct consequence of the convexity and the non-decreasing property of $x \mapsto \varphi(t,\, x)$, for each $t \in [0,\, T]$.

**Remark 2.4.2** *The value function in the American option pricing problem is often non-increasing in time. This will always be the case if the underlying asset price process and the reward function are time-homogeneous, which is true for call and put options under the Black-Scholes model. However, in our setting, the reward function is time-dependent and may be increasing in $t$ (because of the surrender charge function $g$). Therefore, the value function is not necessarily monotone in time.*

We now examine the smoothness of the value function. The next two results are inspired by Lamberton (1998), Proposition 10.3.5 in the context of the pricing of an American put option in the Black-Scholes model. We show that the results still hold in our setting despite significant differences in the reward function, namely its time-dependence and the discontinuity at $t = T$.

**Proposition 2.12** *Let Assumption 2.4.2 hold.*

*(i) For every $t \in [0,\, T]$, and for $x,\, y \geq 0$,*

$$|v(t, x) - v(t,\, y)| \leq |x - y|.$$

*(ii) For every $x \in \mathbb{R}_+^\star$, there exist constants $C_1,\, C_2 > 0$ (which may depend on $x$) satisfying,*

$$|v(t, x) - v(s, x)| \leq C_1 |\sqrt{T - t} - \sqrt{T - s}| + C_2 |t - s|,$$

*for $s, t \in [0,\, T]$.*

Assertion (i) follows from the Lipschitz property of $\varphi$, whereas (ii) follows from Proposition 10.3.5 (1) of Lamberton (1998) and the Lipschitz property of the exponential function on a disk. The details of the proof can be found in Section 2.7.

**Remark 2.4.3** *In Proposition 2.12 (i), the assumption that the surrender charge function only depends on time (Assumption 2.4.2) can be relaxed to more general state-dependent surrender charge functions as long as* $x \mapsto xg(t,x)$ *remains Lipschitz-continuous.*

### 2.4.2 Free Boundary Value Problem and Variational Inequality

Using Theorem 2.6 and the previously established properties of the value function, we can apply classical results from optimal stopping theory to establish the relationship between $v$, a free-boundary value problem, and a variational inequality. From the definition of the surrender and continuation regions in (2.16) and (2.17), we can deduce from the continuity of the value function (Theorem 2.9) and the continuity of the reward function on the interval $[0,T) \times \mathbb{R}_+^\star$ that the continuation region is open and the surrender region is closed. The exercise boundary (or surrender boundary) is the boundary $\partial\mathcal{C}$ of $\mathcal{C}$.

Henceforth, the $t$-section of the surrender (resp. continuation) region is denoted by $\mathcal{S}_t$ (resp. $\mathcal{C}_t$). That is, for $t \in [0,\, T)$,

$$\mathcal{S}_t = \{x \in \mathbb{R}_+^\star : v(t,x) = \varphi(t,x)\} \qquad \text{and} \qquad \mathcal{C}_t = \{x \in \mathbb{R}_+^\star : v(t,x) > \varphi(t,x)\}.$$

For $0 \le t \le T$, we define the second-order differential operator $\mathcal{L}_t$ by

$$\mathcal{L}_t := \frac{x^2\sigma^2}{2}\frac{\partial^2}{\partial x^2} + (r - C(t,x))x\frac{\partial}{\partial x},$$

and the function $L : [0,T) \times \mathbb{R}_+^\star \to \mathbb{R}$ by

$$\begin{aligned} L(t,x) &:= \frac{\mathcal{L}_t(xg(t,x))}{x} + g_t(t,x) - rg(t,x) \\ &= g_t(t,\,x) + (r - C(t,\,x) + \sigma^2)xg_x(t,\,x) \\ &\quad + \frac{\sigma^2x^2}{2}g_{xx}(t,\,x) - C(t,x)g(t,\,x). \end{aligned} \tag{2.18}$$

Note that $L$ is the term on the right-hand side of (2.9). When the fee and the surrender charge functions only depend on time (that is, when Assumption 2.4.2 holds), $L$ becomes

$$L(t) = g_t(t) - c(t)g(t).$$

**Theorem 2.13** *Let Assumption 2.4.1 hold. The value function* $v$ *is the unique solution to the boundary value*

*problem*

$$\begin{cases} \mathcal{L}_t f(t,x) + f_t(t,x) - rf(t,x) = 0, & (t,x) \in \mathcal{C} \\ f(t,x) > xg(t,x), & (t,x) \in \mathcal{C} \\ f(t,\, x) = xg(t,x), & (t,x) \in \mathcal{S} \\ f(T,x) = \max(G,x), & x \in \mathbb{R}_+^\star. \end{cases} \tag{2.19}$$

*That is, any solution $v$ to* (2.19) *can be expressed by* (2.7)*. In particular, the function $v \in C^{1,2}$ on $\mathcal{C}$.*

Once the continuity of the value function has been established (Theorem 2.9), the optimal stopping time has been shown to exist (Corollary 2.5), and the martingale property of the Snell envelope of the discounted reward process has been proven (see Theorem 2.4 (ii), the connection between PDE and optimal stopping problems can be established assuming enough regularity of the coefficients in (2.4). A sufficient condition for the results of Friedman (1964) to hold is that the drift term in (2.4) is locally Hölder continuous, which is why Assumption 2.4.1 is needed in Theorem 2.13. The proof makes use of standard PDE results in solving the Dirichlet (or the first initial-boundary value) problem, see for instance Friedman (1964), Theorem 3.4.9.

Thus, the main difficulty in establishing the free-boundary value problem resides in proving the continuity of the value function (Theorem 2.9). Once this has been done, the free-boundary value problems follow using standard arguments. Specifically, the continuity of the value function implies that the continuation region is an open set. From this, the Dirichlet problem (or the first initial-boundary value problem) can be posed in an open subset (typically a ball or a rectangle) of the continuation region. Since the boundary of the defined open subsets is sufficiently regular, standard PDE results guarantee that the Dirichlet problem admits a unique solution. Probabilistic arguments are then used to identify this solution as the value function. For details of the proof in the context of the American put option pricing in the Black-Scholes setting, see Jacka (1991), Proposition 2.6 or Karatzas and Shreve (1998), Theorem 2.7.7, among others. For more general results, one can also refer to Peskir and Shiryaev (2006), Sections 3.7.1 and 4.8.2 or Jacka and Lynn (1992), Proposition 3.1..

Assumption 2.4.1 is not sufficient to provide much information on the shape of the surrender region. However, we show in Section 2.4.4 (Theorem 2.22) that further conditions can be imposed on the fee and the surrender charge functions in order to express the surrender region explicitly in terms of the optimal surrender boundary $b$. Under these assumptions, the free-boundary value problem can be stated more explicitly.

**Corollary 2.14** *Let Assumptions 2.4.1 and 2.4.2 hold. If $L(t) < 0$ for all $t \in [0, T)$, then, the value function $v$ is the unique solution to the boundary value problem*

$$\begin{cases} \mathcal{L}_t f(t,x) + f_t(t,x) - rf(t,x) = 0, & x < b(t),\ t \in [0,\ T), \\ f(t,x) > \varphi(t,x), & x < b(t),\ t \in [0,\ T), \\ f(t,\ x) = \varphi(t,x), & x \geq b(t),\ t \in [0,\ T), \\ f(T,x) = \max(G,x), & x \in \mathbb{R}_+^\star, \end{cases} \tag{2.20}$$

*with $b(t) := \inf(\mathcal{S}_t)$. That is, any solution $v$ to* (2.19) *can be expressed by* (2.7)*. In particular, the function $v$ is $C^{1,2}$ on $\mathcal{C}$.*

Assuming $L(t) < 0$ for all $t \in [0, T)$ in Corollary 2.14 ensures that $\mathcal{S} \neq \emptyset$ (see Proposition 2.23). The proof of this corollary is a direct consequence of Theorem 2.13 and the results of Section 2.4.4 in Theorem 2.22(i).

**Remark 2.4.4** *For each $t \in [0, T)$, the condition $L(t) < 0$ is equivalent to $\frac{\mathrm{d} \ln g(t)}{\mathrm{d}t} < c(t)$.*

**Remark 2.4.5 (Regularity of the value function)** *Fix $t \in [0, T)$ and suppose that $L(t) < 0$ for all $t \in [0, T]$. Corollary 2.14 gives us smoothness of the value function in the continuation region, that is, when $x < b(t)$. On the other hand, when $x > b(t)$, we know that $v(t, x) = g(t)x$, so $v$ is smooth in int$(\mathcal{S})$. What remains is to assess the smoothness of the value function across the boundary (i.e. when $x = b(t)$). When $L(t) \geq 0$ for all $t \in [0, T]$, an optimal stopping time for* (2.7) *is $T$, as per Proposition 2.3. Hence, the optimal stopping problem in* (2.7) *is reduced to the valuation of the present value of the maturity benefit* (2.5)*. The regularity of the value function then follows from the Feynman-Kac representation (see, for instance, Karatzas and Shreve (1991), Theorem 5.7.6), which states that, under some regularity assumptions on the coefficients of* (2.4)*, the value function in* (2.5) *is the unique solution to a Cauchy problem, and such a solution is in $C^{1,2}([0, T) \times \mathbb{R}_+^\star)$.*

Next, we discuss the link between the value function and a variational inequality. In Proposition 2.15 below, we establish a result that is a direct consequence of the process $J = \{J_t\}_{0 \leq t \leq T}$, with $J_t := e^{-rt} v(t,\ F_t)$, being the Snell envelope of the discounted reward process $Z$; thus, $J$ is a supermartingale by Theorem 2.4 (i).

**Proposition 2.15** *Let Assumption 2.4.2 hold. The value function $v$ defined in* (2.7) *satisfies*

$$\mathcal{L}_t v + v_t - rv \leq 0 \tag{2.21}$$

*for all $(t, x) \in (0, T) \times \mathbb{R}_+^\star$ (with all derivatives in the sense of distribution).*

This proposition is analogous to Proposition 10.3.7 of Lamberton (1998) in the context of American put option pricing, and can be extended to the present context without any further complications (see Remark 2.4.6 for a detailed discussion). For an introduction to weak derivatives and distribution theory see Evans (2010), Chapter 5, and Rudin (1991), Chapter 6.

**Remark 2.4.6** *The idea of the proof of Proposition 2.15 rests on the assumption that the value function is regular enough to apply Itô's lemma to $J$ (which is usually not the case since $v$ may not be smooth enough across the boundary, see Remark 2.4.5). $\mathcal{L}_t v + v_t - rv$, being the drift term of $J$, must be non-positive since $J$ is a supermartingale. In the context of American put option pricing, a rigorous proof relying on distribution theory is given in Lamberton (1998), Proposition 10.3.7. Their arguments are similar to the one outlined above, but are generalized to considered derivatives in the sense of distribution. This approach can also be easily adapted to our setting when the fee is time-dependent. The result also holds for more general surrender functions depending on the sub-account value. Thus, in Proposition 2.15, Assumption 2.4.2 can be relaxed to more general surrender charge functions.*

**Remark 2.4.7** *Other authors have obtained results similar to Proposition 2.15 in more general settings, see for instance Bensoussan and Lions (1982) Section 3.4.9 or Jaillet* et al. *(1990), Theorem 3.1 and 3.2. However, the results of Bensoussan and Lions (1982) require the function ($(t, x) \mapsto xg(t, x)$) to be bounded, and the ones of Jaillet* et al. *(1990) consider a continuous and time-independent reward function. Applying the results of Jaillet* et al. *(1990) to the continuous reward representation* (2.13) *by adding a second dimension (time) to the underlying process, as is often done for time-dependent reward functions, is not directly possible because some conditions on the time derivative are not satisfied by the continuous reward function, the first order derivative of $h$ being unbounded. Establishing the results of Proposition 2.15 in a more general market model is thus left as future research.*

Once (2.21) has been established, we can show that the value function solves a variational inequality when Assumptions 2.4.1 and 2.4.2 hold.

**Proposition 2.16** *Let Assumptions 2.4.1 and 2.4.2 hold. The value function $v$ defined in* (2.7) *is a solution to the variational inequality*

$$\max\left\{\mathcal{L}_t v + v_t - rv, \varphi - v\right\} = 0, \tag{2.22}$$

*with the terminal condition $v(T,x) = \max(G,x)$ and all derivatives in the sense of distribution.*

The proof is a direct consequence of Theorem 2.13 and Proposition 2.15.

**Corollary 2.17** *Let Assumption 2.4.2 hold. The partial derivatives (in the sense of distribution) $\partial v/\partial x$, $\partial v/\partial t$ and $\partial^2 v/\partial x^2$ are locally bounded, and $\partial v/\partial x$ is continuous on $[0,T)\times\mathbb{R}_+^\star$.*

*Proof. From Proposition 2.12, we have that the first order derivatives of the value function (in the sense of distribution) in $t$ and in $x$ are locally bounded. Using the convexity of $x \mapsto v(t,x)$ (Lemma 2.11), the local boundedness of the first order derivatives and* (2.21)*, we can show that the second order derivative in $x$ is also locally bounded (see for instance Lamberton (1998), Theorem 10.3.8 or Jaillet* et al. *(1990), Theorem 3.6 for details). Following the argument of Lamberton (1998), Corollary 10.3.10 (or Jaillet* et al. *(1990), Corollary 3.7), we can conclude that the function $\frac{\partial v}{\partial x}$ is continuous on $[0,T)\times\mathbb{R}_+^\star$, see Ladyženskaja* et al. *(1968), Chapter 2, Lemma 3.1. This is also known as the smooth fit condition in the American option terminology.* □

Hence, the main difficulty of this result resides in proving the Lipchitz property of $v$, obtained in Proposition 2.12. The rest of the proof follows using standard arguments; see, for instance, Lamberton (1998), Corollary 10.3.10, Jaillet *et al.* (1990), Corollary 3.7, or Touzi (1999), Lemma 2.1.

### 2.4.3 Surrender and Continuation Premium Representation

In the following, we derive two representations for the value function. The first one is akin to the early exercise premium representation (or integral representation) in the American option terminology.

To the authors' knowledge, the second representation is new to the literature on variable annuities and American options. The particular form of the reward function involved in variable annuity pricing allows the decomposition of the value function in two terms: the current value of the reward process and an integral term which increases only when the sub-account process is in the continuation region. Therefore, we call this second term the **continuation premium**.

**Theorem 2.18** *Let Assumptions 2.4.1 and 2.4.2 hold. Then, the value function can be written as*

$$v(t,x) = h(t,x) + e(t,x), \tag{2.23}$$

*where $h : [0,T]\times\mathbb{R}_+^\star \to \mathbb{R}_+$ is the present value of the maturity benefit, as defined in* (2.5)*, and $e : [0,T] \times \mathbb{R}_+^\star \to \mathbb{R}_+$ denotes the early* ***surrender premium***

$$e(t,x) := \int_t^T \Big(c(s)g(s) - g_t(s)\Big)\mathbb{E}\Big[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{S}\}}\Big]\,\mathrm{d}s, \tag{2.24}$$

*where $g_t = \frac{\partial g}{\partial t}$.*

*Proof. Corollary 2.17 ensures that $v$ is smooth enough to apply Itô's formula for generalized derivatives (see Krylov (1980), Theorem 2.10.1) to $\left\{e^{-rs}v(t,F_s^{t,x})\right\}_{t\le s\le T}$, yielding*

$$\begin{aligned}\mathrm{d}(e^{-rs}v(s,F_s^{t,x})) = e^{-rs}\left(\mathcal{L}_s v(s,F_s^{t,x}) + v_t(s,F_s^{t,x}) - rv(s,F_s^{t,x})\right)\mathrm{d}s\\ + e^{-rs}\sigma F_s^{t,x}v_x(s,F_s^{t,x})\,\mathrm{d}W_s\end{aligned}$$

*for $0 \le s \le T$. Integrating from $t$ to $T$ on both side and multiplying by $e^{rt}$, we get*

$$\begin{aligned}&e^{-r(T-t)}v(T,F_T^{t,x})\\ &\quad= v(t,F_t^{t,x})\\ &\qquad+\int_t^T e^{-r(s-t)}\left(\mathcal{L}_s v(s,F_s^{t,x}) + v_t(s,F_s^{t,x}) - rv(s,F_s^{t,x})\right)\mathrm{d}s\\ &\qquad+\int_t^T e^{-r(s-t)}\sigma F_s^{t,x}v_x(s,\,F_s^{t,x})\,\mathrm{d}W_s.\end{aligned} \tag{2.25}$$

*$\int_t^T e^{-r(s-t)}\sigma F_s^{t,x}v_x(s,\,F_s^{t,x})\,\mathrm{d}W_s$ is a martingale since $|v_x(t,x)| \le 1$ for all $(t,x) \in [0,T)\times\mathbb{R}_+^\star$, as per Proposition 2.12 and Corollary 2.17, and $\mathbb{E}[\int_0^T F_s^2\,\mathrm{d}s] < \infty$. Thus, taking the expectation on both sides of* (2.25) *yields*

$$\begin{aligned}v(t,x) = \mathbb{E}\left[e^{-r(T-t)}v(T,F_T^{t,x})\right]\\ -\int_t^T \mathbb{E}\left[e^{-r(s-t)}\left(\mathcal{L}_s v(s,F_s^{t,x}) + v_t(s,\,F_s^{t,x}) - rv(s,\,F_s^{t,x})\right)\right]\mathrm{d}s.\end{aligned}$$

*Now, $\mathbb{E}\left[e^{-r(T-t)}v(T,F_T^{t,x})\right] = \mathbb{E}\left[e^{-r(T-t)}\max(G,F_T^{t,x})\right] = h(t,\,x)$, and by Corollary 2.14, $\mathcal{L}_t v(s,x) + v_t(s,x) - rv(s,x) = 0$ for all $(s,\,x) \in \mathcal{C}$. Futhermore, for $s < T$, $v(s,\,x) = g(s)x$ when $(s,x) \in \mathcal{S}$. It follows that*

$$\mathcal{L}v(s,F_s^{t,x}) + v_t(s,\,F_s^{t,x}) - rv(s,\,F_s^{t,x}) = F_s^{t,x}\left[g_t(s) - c(s)g(s)\right]\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{S}\}},\quad a.s.,$$

*which concludes the proof.* □

**Remark 2.4.8** *If $L(t) < 0$ for all $t \in [0,T)$, the surrender region has the particular shape $\mathcal{S} = \{(t,x) \in [0,T) \times \mathbb{R}_+^\star | x \geq b(t)\}$ for some $b(t) \geq Ge^{-r(T-t)}$. This will be shown in the next section in Theorem 2.22. Under this assumption, the early surrender premium becomes*

$$e(t,x) := \int_t^T \Big(c(s)g(s) - g_t(s)\Big)\mathbb{E}\Big[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{F_s^{t,x} > b(s)\}}\Big]\,\mathrm{d}s. \tag{2.26}$$

The results presented in Theorem 2.18 are generalizations of Theorem 1 and Equation (11) of Bernard *et al.* (2014b) to time-dependent fee and surrender charge functions.

**Theorem 2.19** *Suppose Assumptions 2.4.1 and 2.4.2 hold. The value function can be written as*

$$v(t,x) = xg(t,x) + f(t,x), \tag{2.27}$$

*where $f : [0,T] \times \mathbb{R}_+^\star \to \mathbb{R}_+$ is the continuation premium given by*

$$\begin{aligned} f(t,x) = \mathbb{E}[e^{-r(T-t)}(G - F_T^{t,x})_+] \\ + \int_t^T \Big(g_t(s) - c(s)g(s)\Big)\mathbb{E}\Big[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{C}\}}\Big]\,\mathrm{d}s. \end{aligned} \tag{2.28}$$

The proof of Theorem 2.19 is presented at the end of the section; it relies on the results below.

**Remark 2.4.9** *If $L(t) < 0$ for all $t \in [0,T)$, the continuation region is given by $\mathcal{C} = \{(t,x) \in [0,T) \times \mathbb{R}_+^\star | x < b(t)\}$ for some $b(t) \geq Ge^{-r(T-t)}$, see Theorem 2.22. Under this assumption, the continuation premium in* (2.28) *becomes*

$$\begin{aligned} f(t,x) := \mathbb{E}\left[e^{-r(T-t)}\left(G - F_T^{t,x}\right)_+\right] \\ + \int_t^T \Big(g_t(s) - c(s)g(s)\Big)\mathbb{E}\Big[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{F_s^{t,x} < b(s)\}}\Big]\,\mathrm{d}s. \end{aligned} \tag{2.29}$$

Theorem 2.19 decomposes the value function in terms of the immediate surrender value plus a term representing the value of holding on to the contract. This term, which we coin the *continuation* premium, is the sum of the financial guarantee at maturity (or the put option) when the contract is held until $T$ and an

integral term equal to the value added by keeping the contract until it is optimal to surrender. The representation of the continuation premium in (2.28) can be particularly helpful to develop numerical methods for approximating the value of a variable annuity contract, since it uses additional information on the shape of the continuation region.

When Assumptions 2.4.1 and 2.4.2 do not hold, it is still possible to decompose the value function in two parts, the surrender benefit and the continuation premium. However, in this case, the continuation premium is written in terms of $\tau_t^x$, which can make it more complicated to compute. This result is given in Lemma 2.20 below and will be used to prove Theorem 2.19.

**Lemma 2.20** *The value function can be written as*

$$v(t,x) = xg(t,x) + f(t,x), \tag{2.30}$$

*where $f : [0,T] \times \mathbb{R}_+^\star \to \mathbb{R}_+$ is the continuation premium given by*

$$\begin{aligned} f(t,x) := \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+ \mathbf{1}_{\{\tau_t^x = T\}}\right] \\ + \int_t^T \mathbb{E}\Big[e^{-r(u-t)} F_u^{t,x} L(u, F_u^{t,x}) \mathbf{1}_{\{u \le \tau_t^x\}}\Big]\, \mathrm{d}u, \end{aligned} \tag{2.31}$$

*with $\tau_t^x$ defined in* (2.12).

*Proof. The result is trivial for $t = T$. For the rest of the proof, fix $(t,x) \in [0,T) \times \mathbb{R}_+^\star$. Recall that $\tau_t^x$ is an optimal stopping time for $v(t,x)$ (see Corollary 2.5). Hence, we have*

$$v(t,x) = \mathbb{E}\left[e^{-r(\tau_t^x - t)} \varphi(\tau_t^x, F_{\tau_t^x}^{t,x})\right].$$

*Notice that the discounted reward process can be decomposed as*

$$e^{-r(s-t)}\varphi(s, F_s^{t,x}) = Y_s^{t,x} + e^{-r(T-t)}(G - F_T^{t,x})_+ \mathbf{1}_{\{s=T\}}$$

*with $Y^{t,x} = \{Y_s^{t,x}\}_{t \le s \le T}$, with $Y_s^{t,x} := e^{-r(s-t)} g(s, F_s^{t,x}) F_s^{t,x}$ and observe that $Y_T^{t,x} = e^{-r(T-t)} F_T^{t,x}$ since $g(T,x) = 1$ for all $x \in \mathbb{R}_+^\star$. The function $xg(t,x)$ is $C^{1,2}$, so we can apply Itô's formula to $Y$, which yields*

$$
\begin{aligned}
&e^{-r(T-t)}\varphi(\tau_t^x, F_{\tau_t^x}^{t,x}) && (2.32)\\
&\quad = Y_{\tau_t^x}^{t,x} + e^{-r(\tau_t^x - t)}(G - F_{\tau_t^x}^{t,x})_+ \mathbf{1}_{\{\tau_t^x = T\}}\\
&\quad = xg(t,x) + e^{-r(T-t)}(G - F_T^{t,x})_+ \mathbf{1}_{\{\tau_t^x = T\}}\\
&\qquad + \int_t^{\tau_t^x} e^{-r(s-t)} F_s^{t,x} L(s, F_s^{t,x})\,\mathrm{d}s && (2.33)\\
&\qquad + \int_t^{\tau_t^x} e^{-r(s-t)} \sigma F_s^{t,x} \left(g(s, F_s^{t,x}) - g_x(s, F_s^{t,x}) F_s^{t,x}\right)\,\mathrm{d}W_s.
\end{aligned}
$$

*The final result is obtained by taking the expectation on both sides and using the zero-mean property of the stochastic integral and Doob's optional sampling theorem. To complete the proof, note that* $f(t,x) \geq 0$ *for all* $(t,x) \in [0,T] \times \mathbb{R}_+^\star$*, since* $v(t,x) \geq \varphi(t,x)$ *as per Lemma 2.10.* □

The proof of Theorem 2.19 also rely on the following Lemma, which is used to remove the dependence of the continuation premium on $\tau_t^x$.

**Lemma 2.21** *Suppose Assumptions 2.4.1 and 2.4.2 hold. For all* $(t,x) \in [0,T] \times \mathbb{R}_+^\star$*,*

$$
\begin{aligned}
&\mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+ \mathbf{1}_{\{\tau_t^x \neq T\}}\right]\\
&\qquad + \mathbb{E}\left[\int_{\tau_t^x}^T \Big(g_t(s) - c(s)g(s)\Big) e^{-r(s-t)} F_s^{t,x} \mathbf{1}_{\{(s,F_s)\in\mathcal{C}\}}\,\mathrm{d}s\right] = 0,
\end{aligned}
\tag{2.34}
$$

*with* $\tau_t^x$ *defined in* (2.12)*.*

*Proof. The proof is trivial for* $t = T$*. For the rest of the proof, fix* $(t,x) \in [0,T) \times \mathbb{R}_+^\star$*. First, note that*

$$
\begin{aligned}
v(t,x) &= \mathbb{E}\left[e^{-r(T-t)}\varphi(T, F_T^{t,x})\right]\\
&\qquad + \mathbb{E}\left[e^{-r(\tau_t^x - t)}\varphi(\tau_t^x, F_{\tau_t^x}) - e^{-r(T-t)}\varphi(T, F_T^{t,x})\right].
\end{aligned}
\tag{2.35}
$$

*Now recall that the discounted reward process admits the following decomposition*

$$
e^{-r(s-t)}\varphi(s, F_s^{t,x}) = Y_s^{t,x} + e^{-r(T-t)}(G - F_T^{t,x})_+ \mathbf{1}_{\{s=T\}},
$$

*with* $Y_s^{t,x} = e^{-r(s-t)} g(s) F_s^{t,x}$ *for* $t \leq s \leq T$*. Hence, applying Itô's lemma to* $Y^{t,x} = \{Y_s^{t,x}\}_{t\leq s\leq T}$*, we*

*obtain*

$$
\begin{aligned}
&e^{-r(T-t)}\varphi(T,F_T^{t,x})\\
&\quad= e^{-r(T-t)}Y_T^{t,x}+e^{-r(T-t)}(G-F_T^{t,x})_+\\
&\quad= e^{-r(\tau_t^x-t)}Y_{\tau_t^x}+e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x=T\}}\\
&\quad+\int_{\tau_t^x}^T e^{-r(s-t)}F_s^{t,x}(g_t(s)-c(s)g(s))\,\mathrm{d}s\\
&\quad+\int_{\tau_t^x}^T e^{-r(s-t)}g(s)F_s^{t,x}\sigma\,\mathrm{d}W_s+e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x\neq T\}}\\
&\quad= e^{-r(\tau_t^x-t)}\varphi(\tau_t^x,F_{\tau_t^x}^{t,x})+\int_{\tau_t^x}^T e^{-r(s-t)}F_s^{t,x}(g_t(s)-c(s)g(s))\,\mathrm{d}s\\
&\quad+\int_{\tau_t^x}^T e^{-r(s-t)}g(s)F_s^{t,x}\sigma\,\mathrm{d}W_s+e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x\neq T\}}.
\end{aligned}
\tag{2.36}
$$

*Replacing* (2.36) *in* (2.35) *yields*

$$
\begin{aligned}
v(t,x)&=\mathbb{E}\left[e^{-r(T-t)}\varphi(T,F_T^{t,x})\right]+\mathbb{E}\left[\int_{\tau_t^x}^T e^{-r(s-t)}F_s^{t,x}(c(s)g(s)-g_t(s))\,\mathrm{d}s\right]\\
&\quad-\mathbb{E}\left[e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x\neq T\}}\right]\\
&=h(t,x)+\mathbb{E}\left[\int_{\tau_t^x}^T e^{-r(s-t)}F_s^{t,x}(c(s)g(s)-g_t(s))\mathbf{1}_{\{(s,F_s^{t,x}\in\mathcal{S}\}}\,\mathrm{d}s\right]\\
&\quad+\mathbb{E}\left[\int_{\tau_t^x}^T e^{-r(s-t)}F_s^{t,x}(c(s)g(s)-g_t(s))\mathbf{1}_{\{(s,F_s^{t,x}\in\mathcal{C}\}}\,\mathrm{d}s\right]\\
&\quad-\mathbb{E}\left[e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x\neq T\}}\right],
\end{aligned}
\tag{2.37}
$$

*where* $h(t,x)$ *is defined as in* (2.5)*. Now note that*

$$
\int_t^{\tau_t^x} e^{-r(s-t)}F_s^{t,x}(c(s)g(s)-g_t(s))\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{S}\}}\,\mathrm{d}s=0\qquad a.s.,
$$

*since for* $s\in[t,\tau_t^x]$, $F_s^{t,x}\in\mathcal{C}_s$*, by definition of* $\tau_t^x$ *being the first entry time of* $F_s^{t,x}$ *in* $\mathcal{S}_s$ *between* $t$ *and* $T$*. Thus,*

$$
\begin{aligned}
&\int_{\tau_t^x}^T e^{-r(s-t)}F_s^{t,x}(c(s)g(s)-g_t(s))\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{S}\}}\,\mathrm{d}s\\
&\qquad=\int_t^T e^{-r(s-t)}F_s^{t,x}(c(s)g(s)-g_t(s))\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{S}\}}\,\mathrm{d}s,\qquad a.s.,
\end{aligned}
$$

*so* (2.37) *becomes*

$$
\begin{aligned}
v(t,x) &= h(t,x) + e(t,x) \\
&\quad + \mathbb{E}\left[\int_{\tau_t^x}^{T} e^{-r(s-t)} F_s^{t,x}(c(s)g(s) - g_t(s))\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{C}\}}\,\mathrm{d}s\right] \\
&\quad - \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x \neq T\}}\right],
\end{aligned} \tag{2.38}
$$

*where* $e(t,x)$ *denotes the surrender premium defined in* (2.24)*. The final result is obtained by comparing* (2.23) *and* (2.38)*.* □

We note that Lemma 2.21 requires assumptions 2.4.1 and 2.4.2 to hold since the proof makes use of the surrender premium representation (2.23) in Theorem 2.18.

Lemmas 2.20 and 2.21 can now be used to prove Theorem 2.19.

*Proof. [Proof of Theorem 2.19] When* $t = T$*, the proof is trivial. For the rest of the proof, fix* $(t,x) \in [0,T)$*. Using the continuation premium representation in* (2.31) *and recalling from the proof of Lemma 2.21 that* $\int_t^{\tau_t^x} e^{-r(s-t)} F_s^{t,x}(c(s)g(s) - g_t(s))\mathbf{1}_{\{(s,F_s^{t,x})\in\mathcal{S}\}}\,\mathrm{d}s = 0$ *a.s., we find that*

$$
\begin{aligned}
v(t,x) &= xg(t) + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x = T\}}\right] + \mathbb{E}\Big[\int_t^{\tau_t^x} e^{-r(s-t)} F_s^{t,x} L(s)\,\mathrm{d}s\Big] \\
&= xg(t) + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+(1 - \mathbf{1}_{\{\tau_t^x \neq T\}})\right] \\
&\qquad + \mathbb{E}\Big[\int_t^{\tau_t^x} e^{-r(s-t)} F_s^{t,x} L(s)\left(\mathbf{1}_{\{F_s^{t,x}\in\mathcal{C}_s\}}) + \mathbf{1}_{\{F_s^{t,x}\in\mathcal{S}_s\}}\right)\,\mathrm{d}s\Big] \\
&= xg(t) + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] - \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\mathbf{1}_{\{\tau_t^x \neq T\}}\right] \\
&\qquad + \mathbb{E}\Big[\int_t^{\tau_t^x} e^{-r(s-t)} F_s^{t,x} L(s)\mathbf{1}_{\{F_s^{t,x}\in\mathcal{C}_s\}}\,\mathrm{d}s\Big] \\
&= xg(t) + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] + \int_t^T L(s)\mathbb{E}\Big[e^{-r(s-t)} F_s^{t,x}\mathbf{1}_{\{F_s^{t,x}\in\mathcal{C}_s\}}\Big]\,\mathrm{d}s,
\end{aligned}
$$

*where the last equality follows from Lemma 2.21.* □

### 2.4.4 Characterization of the Surrender Region

The goal of this section is to study the properties of the surrender and continuation regions. Under certain conditions on the fee and surrender charge functions, we can fully characterize the two regions. The main results of this section characterize specific types of surrender regions and are summarized in Theorem 2.22

below. The rest of the section contains further results that are used in the proof of Theorem 2.22 and that are of interest on their own. The proof of Theorem 2.22 is presented at the end of the section.

Although the shape of the surrender region can vary significantly (see for example Figures 4 and 5 of MacKay *et al.* (2017), when the fee rate is constant and the (time-dependent) surrender charge function satisfies certain conditions, MacKay (2014), Appendix 2.A shows that the surrender region is a connected set. The first part of Theorem 2.22 below indicates that as soon as the surrender region is non-empty and the fee and surrender charge functions depend only on time, then $t$-section of $\mathcal{S}$ has the form $\mathcal{S}_t = [b(t), \infty)$ for some $b(t) \in \mathbb{R}_+^\star \cup \{\infty\}$. This result builds on Jacka (1991), Proposition 2.1 in the context of a bounded and time-homogeneous reward function in a time-homogeneous market model.

The second part of Theorem 2.22 confirms that the optimal stopping time obtained in Proposition 2.3 is unique.

We recall that, under Assumption 2.4.2 , the function $L$, defined in (2.18), becomes $L(t) = g_t(t) - c(t)g(t)$.

**Theorem 2.22** *Let Assumptions 2.4.1 and 2.4.2 hold.*

*(i) If $L(t) < 0$ for all $t \in [0, T)$ then the $t$-sections of $\mathcal{S}$, denoted by $\mathcal{S}_t$, are of the form*

$$\mathcal{S}_t = \{x \in \mathbb{R}_+^\star : (t, x) \in \mathcal{S}\} = [b(t),\, \infty), \tag{2.39}$$

*for some $b(t) \in \mathbb{R}_+^\star \cup \{\infty\}$ satisfying $b(t) \geq Ge^{-r(T-t)}$ and $t \in [0, T)$.*

*(ii) If $L(t) \geq 0$ for all $(t) \in [0, T)$, then $\mathcal{S} = \emptyset$ and $T$ is the unique optimal stopping time for* (2.7).

The proof of Theorem 2.22 is presented at the end of the section, as it builds upon additional results presented below.

**Remark 2.4.10** *The proof of part (ii) of Theorem 2.22 does not require Assumptions 2.4.1 and 2.4.2 to hold. In this more general setting, if $L(t, x) \geq 0$ for all $(t, x) \in [0, T) \times \mathbb{R}_+^\star$, then $\mathcal{S} = \emptyset$ and $T$ is the unique optimal stopping time for* (2.7).

Under the assumptions of Theorem 2.22, $b(t)$ is the smallest sub-account value for which it is optimal to surrender the contract at time $t < T$, and for any fund value greater than $b(t)$, it is also optimal to surrender, so that

$$b(t) = \inf\left\{x \in \mathbb{R}_+^\star \middle| x \in \mathcal{S}_t\right\} = \inf\{\mathcal{S}_t\}, \tag{2.40}$$

with $b(t) = \infty$ if $\mathcal{S}_t = \emptyset$. Under these assumptions, the continuation and the surrender regions can be expressed as

$$\mathcal{C} = \left\{(t,x) \in [0,T) \times \mathbb{R}_+^\star \middle| x < b(t)\right\},$$

and

$$\mathcal{S} = \left\{(t,x) \in [0,T) \times \mathbb{R}_+^\star \middle| x \geq b(t)\right\},$$

respectively. When $b(t) < \infty$ for all $t \in [0,T)$, the surrender boundary splits $[0,T) \times \mathbb{R}_+^\star$ in two regions: the surrender region is at or above the boundary, and the continuation region is below. It follows that the set $\mathcal{S}$ is connected. Henceforth, we say that the surrender region is of "threshold type" if for any $t \in [0,T)$, there exists a $b(t) < \infty$ such that $\mathcal{S}_t = [b(t), \infty)$. Such a geometry for the surrender region can be explained, as in Milevsky and Salisbury (2001), by the fact that when the account value is low, it is financially advantageous for the policyholder to hold on to the contract since there is a significant chance that the guarantee will be triggered at maturity. Since this guarantee is financed by the policyholder via the continuous fee, which reduces the net return on the account, there is a point above which it is no longer profitable to hold the contract and continue paying the fee; this threshold is the optimal surrender boundary.

The rest of this section presents further results on the (non-)emptiness of the surrender region and ends with the proof of Theorem 2.22. Many of the results presented here are inspired by the work of Villeneuve (1999), which we adapted to the time-dependent and discontinuous reward function considered in (2.7). As pointed out by Villeneuve (1999), Remark 2.1, adapting some of their results to a time-dependent payoff is not trivial.

Proposition 2.23 characterizes the non-emptiness of the surrender region. The first part involves relint$(\mathcal{S}_t)$, the relative interior of a $t$-section $\mathcal{S}_t$, which is defined as its interior within the affine hull of $\mathcal{S}_t$. If it is not empty, the relative interior of $\mathcal{S}_t$ forms a vertical line in the plane. If relint$(\mathcal{S}_t) \in$ int$(\mathcal{S})$, it contains some finite $x$ for which $(t,x) \in$ int$(\mathcal{S})$. The second part of Proposition 2.23 (ii) can be seen as a local version of the first part.

**Proposition 2.23** *Let Assumptions 2.4.1 and 2.4.2 hold.*

*(i) For each $t \in [0,T)$, if $\mathcal{S} \neq \emptyset$ and relint$(\mathcal{S}_t) \subset$ int$(\mathcal{S})$, then $L(t) \leq 0$.*

*(ii) If $L(t) < 0$ for all $t \in [0,T)$, then*

- *$\mathcal{S} \neq \emptyset$ and $\mathcal{C} \neq \emptyset$, and*
- *$\cup_{t_1 \leq t < t_2} \mathcal{S}_t \neq \emptyset$ for any $[t_1, t_2) \subset [0,T)$ such that $t_1 < t_2$.*

The proof of Proposition 2.23 is reported in Section 2.7. Proposition 2.24 below characterizes the emptiness of the surrender region.

**Proposition 2.24** *Let Assumptions 2.4.1 and 2.4.2 hold.*

*(i) For each $t \in [0,T)$, if $L(t) > 0$, then $\mathcal{S}_t = \emptyset$ or relint$(\mathcal{S}_t) \not\subset$ int$(\mathcal{S})$.*

*(ii) For any $[t_1, t_2) \subset [0,T)$ such that $t_1 < t_2$, if $L(t) > 0$ for all $(t) \in [t_1, t_2)$, then $\cup_{t_1 \leq t < t_2} \mathcal{S}_t = \emptyset$.*

*(iii) If $\mathcal{S} = \emptyset$, then $g(t) - e^{-\int_t^T c(s)\,\mathrm{d}s} \leq 0$ for all $t \in [0,T)$.*

**Remark 2.4.11** *Some of the assumptions of Proposition 2.24 can be relaxed. The proof of part (i) only requires Assumption 2.4.2. Furthermore, Proposition 2.24(ii) holds in a general setting, without Assumptions 2.4.1 and 2.4.2. In this general setting, it holds that for any $[t_1, t_2) \subset [0,T)$ such that $t_1 < t_2$, if $L(t,x) > 0$ for all $(t,x) \in [t_1, t_2) \times \mathbb{R}_+$, then $\cup_{t_1 \leq t < t_2} \mathcal{S}_t = \emptyset$.*

The proof of Proposition 2.24 is reported in Section 2.7.

**Remark 2.4.12** *The first statements of Propositions 2.23 and 2.24 can apply to surrender charge functions depending on both time and the value of the underlying process. Indeed, the proof relies on the results of Proposition 2.15 which only requires the fee function, and not the surrender charge, to depend exclusively on time, see Remark 2.4.6 for details.*

Proposition 2.23 (ii) and Proposition 2.24 (i), state that the surrender region being empty or not only depends on the sign of $L(t)$. We believe that the results of Proposition 2.23 (ii) can be made more local. That

is, we conjecture that as soon as $L(t) < 0$ for some $t \in [0, T)$, the relative interior of the $t$-section of the surrender region is not empty. This is what we observe in numerical examples, see Section 2.5. Hence, when $L(t)$ changes sign in $[0, T)$, the boundary of the continuation region, $\partial\mathcal{C}$, is discontinuous. This property of the optimal surrender boundary sets the optimal stopping problem involved in variable annuities pricing apart from the pricing of standard American call and put options.

**Remark 2.4.13** *The statements of Proposition 2.23 (ii) and 2.24 (i) do not cover the case where $L(t) = 0$. Heuristically, when $L(t) = 0$ for all $0 \leq t \leq T$, the discounted surrender value process $\{Y_t\}_{0\leq t\leq T}$, with $Y_t = e^{-rt}F_t g(t)$, is a martingale, so one might expect that all stopping times $\tau$ such that $0 \leq \tau \leq T$ are optimal, implying that $\mathcal{S} = [0, T) \times \mathbb{R}^\star_+$, see for instance Björk (2009), Proposition 21.2. However, this is incorrect; due to the time discontinuity of the reward function in* (2.7)*, the discounted reward process $\{Z_t\}_{0\leq t\leq T}$, with $Z_t = e^{-rt}\varphi(t, F_t)$, is a submartingale (see Lemma 2.1). Hence, the policyholder can always profit (on average) from holding on to the contract because of the guaranteed amount at maturity. The optimal stopping time is then $T$ and the surrender region is empty. This illustrates another major difference between the optimal stopping problem in* (2.7) *and the one involved in the pricing of standard American options.*

*The specificity of the VA pricing problem discussed above is also illustrated in the continuation premium representation in* (2.28)*.*

*Indeed, when $L(t) = 0$ for all $t \in [0, T]$, then for any $x \in \mathbb{R}^\star_+$, $f(t, x) = \mathbb{E}[e^{-r(T-t)}(G - F_T^{t,x})_+] > 0$. That is, the continuation premium $f$ is equal to the expected present value of the financial guarantee. Hence, $v(t, x) = \varphi(t, x) + f(t, x) > \varphi(t, x)$ for all $(t, x) \in [0, T) \times \mathbb{R}^\star_+$, which implies that $\mathcal{S} = \emptyset$.*

The results presented in Propositions 2.23 and 2.24 can now be used to prove Theorem 2.22.

*Proof. [Theorem 2.22]*

*To show part (i), let $L(t) < 0$ for all $t \in [0, T)$. It follows from Proposition 2.23 (ii) that $\mathcal{S} \neq \emptyset$, and thus there exists $t \in [0, T)$ for which $\mathcal{S}_t \neq \emptyset$. Fix $t \in [0, T)$. If $\mathcal{S}_t = \emptyset$ then $b(t) = \infty$ and the proof is complete. If $\mathcal{S}_t \neq \emptyset$, define*

$$b(t) = \inf\{x \in \mathbb{R}^\star_+ | v(t, x) = \varphi(t, x)\} = \inf\{\mathcal{S}_t\}.$$

*Since $\mathcal{S}_t$ is non-empty, we know that $b(t) < \infty$. Moreover, since $v$ and $\varphi$ are continuous on $[0,T)$ (see Theorem 2.9), $\mathcal{S}$ is a closed set that is bounded below, so that $b(t) \in \mathcal{S}_t$ and satisfies $v(t,b(t)) = \varphi(t,b(t))$. Thus, for $t < T$,*

$$b(t) \geq g(t)b(t) = v(t,b(t)) \geq \mathbb{E}\left[e^{-r(T-t)}\max(G, F_T^{t,b(t)})\right] \geq G^{-r(T-t)},$$

*since $g$ takes values in $(0,1]$.*

*Next, we show that $\mathcal{S}_t = [b(t), \infty)$. Fix $t \in [0,T)$ and note that if $\mathcal{S}_t = [b(t),\infty)$, then $\mathcal{C}_t = (0, b(t))$. Thus, we need to prove that $x \in \mathcal{C}_t \Rightarrow y \in \mathcal{C}_t$ for any $y < x$, which we do next.*

*Recall that $\tau_t^x = \inf\left\{t \leq s \leq T \,\middle|\, v(s, F_s^{t,x}) = \varphi(s, F_s^{t,x})\right\}$ is optimal for $v(t,x)$. Hence, we have that*

$$\begin{aligned}
0 &\geq v(t,y) - v(t,x) \quad \textit{(since } x \mapsto v(t,x) \textit{ is non-decreasing, see Lemma 2.11)} \\
&\geq \mathbb{E}\left[e^{-r(\tau_t^x - t)}\varphi(\tau_t^x, F_{\tau_t^x}^{t,y}) - e^{-r(\tau_t^x - t)}\varphi(\tau_t^x, F_{\tau_t^x}^{t,x})\right] \\
&= \mathbb{E}\left[e^{-r(\tau_t^x - t)} g(\tau_t^x) e^{(r-\sigma^2/2)(\tau_t^x - t) - \int_t^{\tau_t^x} c(s)\,\mathrm{d}s + \sigma W_{\tau_t^x - t}}(y-x)\mathbf{1}_{\{\tau_t^x < T\}}\right] \\
&\quad + \mathbb{E}\left[e^{-r(\tau_t^x - t)}\left\{\max(G,\, F_{\tau_t^x}^{t,y}) - \max(G,\, F_{\tau_t^x}^{t,x})\right\}\mathbf{1}_{\{\tau_t^x = T\}}\right].
\end{aligned} \tag{2.41}$$

*For all $\omega \in \Omega$ and $\tau \in \mathcal{T}_{t,T}$, we have $F_{\tau(\omega)}^{t,x}(\omega) > F_{\tau(\omega)}^{t,y}(\omega)$, since $y < x$. It follows that $(G - F_\tau^{t,y})_+ - (G - F_\tau^{t,x})_+ \geq 0$, for all $\omega \in \Omega$, so that*

$$\begin{aligned}
&\mathbb{E}\left[e^{-r(\tau_t^x - t)}\left\{\max(G,\, F_{\tau_t^x}^{t,y}) - \max(G,\, F_{\tau_t^x}^{t,x})\right\}\mathbf{1}_{\{\tau_t^x = T\}}\right] \\
&\quad \geq \mathbb{E}\left[e^{-r(\tau_t^x - t)}\left\{F_{\tau_t^x}^{t,y} - F_{\tau_t^x}^{t,x}\right\}\mathbf{1}_{\{\tau_t^x = T\}}\right] \\
&\quad = \mathbb{E}\left[e^{-r(\tau_t^x - t)} g(\tau_t^x) e^{(r-\sigma^2/2)(\tau_t^x - t) - \int_t^{\tau_t^x} c(s)\,\mathrm{d}s + \sigma W_{\tau_t^x - t}}(y-x)\mathbf{1}_{\{\tau_t^x = T\}}\right],
\end{aligned} \tag{2.42}$$

*since $g(T) = 1$. Combining (2.41) and (2.42) yields*

$$\begin{aligned}
0 \geq v(t,y) - v(t,x) &\geq (y-x)\mathbb{E}\left[e^{-r(\tau_t^x - t)} g(\tau_t^x) e^{(r-\sigma^2/2)(\tau_t^x - t) - \int_t^{\tau_t^x} c(s)\,\mathrm{d}s + \sigma W_{\tau_t^x - t}}\right] \\
&= g(t)(y-x)\mathbb{E}\left[\frac{g(\tau_t^x)}{g(t)} e^{-\int_t^{\tau_t^x} c(s)} e^{\sigma W_{\tau_t^x - t} - \frac{\sigma^2}{2}(\tau_t^x - t)\,\mathrm{d}s}\right] \\
&= g(t)(y-x)\mathbb{E}\left[e^{\int_t^{\tau_t^x} \frac{\mathrm{d}\ln g(s)}{\mathrm{d}s} - c(s)\,\mathrm{d}s} e^{\sigma W_{\tau_t^x - t} - \frac{\sigma^2}{2}(\tau_t^x - t)}\right] \\
&> g(t)(y-x)\mathbb{E}\left[e^{\sigma W_{\tau_t^x - t} - \frac{\sigma^2}{2}(\tau_t^x - t)}\right] \\
&= g(t)(y-x).
\end{aligned} \tag{2.43}$$

*Finally, since $x \in \mathcal{C}_t$, we have $v(t,\, x) > g(t)x$, so that $v(t, y) > g(t)y$, which implies $y \in \mathcal{C}_t$.*

*For part (ii), we have from Proposition 2.3 that $v(t, x) = h(t, x)$, for all $(t, x) \in [0, T] \times \mathbb{R}^{\star}_{+}$. Now using Itô's lemma on the discounted surrender value process and the zero-mean property of the stochastic integral, we find that for any $(t, x) \in [0, T) \times \mathbb{R}^{\star}_{+}$,*

$$\begin{aligned} v(t,x) &= \mathbb{E}\left[e^{-r(T-t)} \max(G, F_T^{t,x})\right] \\ &= \mathbb{E}\left[e^{-r(T-t)} g(T, F_T^{t,x}) F_T^{t,x}\right] + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \\ &= g(t,x)x + \mathbb{E}\left[\int_t^T e^{-r(s-t)} F_s^{t,x} L(s, F_s^{t,x})\, \mathrm{d}s\right] + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \\ &> g(t,x)x. \end{aligned}$$

*Hence, $\mathcal{S} = \{(t,x) \in [0,T) \times \mathbb{R}^{\star}_{+} | v(t,x) = \varphi(t,x)\} = \emptyset$, so that $\tau_t^x = \inf\{t \le s \le T | v(s, F_s^{t,x}) = \varphi(s, F_s^{t,x})\} = T$.*

*Now since $\tau_t^x$ is the smallest optimal stopping time for* (2.7)*, as per Theorem 2.4 (iii), we conclude that it is unique (since all other optimal stopping times must be greater than $\tau_t^x$ and smaller than $T$).* □

### 2.4.5 Equivalence of the Optimal Stopping Problems

In this section, we compare the optimal stopping problems with continuous and discontinuous reward functions introduced in Section 2.3, and provide a simple condition under which the two problems lead to the exact same surrender regions and optimal stopping times.

**Proposition 2.25** *Suppose Assumption 2.4.2 holds. If $L(t) < 0$ for all $t \in [0, T)$, then $\mathcal{S} = \widetilde{\mathcal{S}}$.*

The proof of Proposition 2.25 relies on the following lemma.

**Lemma 2.26** *Under the assumptions of Proposition 2.25, $v(t, x) > h(t, x)$ for all $(t, x) \in [0, T) \times \mathbb{R}^{\star}_{+}$, where $h(t, x) = \mathbb{E}\left[e^{-r(T-t)} \max(G, F_T^{t,x})\right]$.*

*Proof. We proceed by contradiction. Suppose that $L(t) < 0$ for all $t \in [0,T)$ and that $v(t,x) = h(t,x)$ for some $(t,x) \in [0,T) \times \mathbb{R}_+^\star$. Using Itô's lemma on the discounted surrender value process and the zero-mean property of the stochastic integral, we find that*

$$\begin{aligned}
v(t,x) &= \mathbb{E}\left[e^{-r(T-t)} \max(G, F_T^{t,x})\right] \\
&= \mathbb{E}\left[e^{-r(T-t)} g(T) F_T^{t,x}\right] + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \\
&= xg(t) + \mathbb{E}\left[\int_t^T e^{-r(s-t)} F_s^{t,x} L(s)\,\mathrm{d}s\right] + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \\
&\leq xg(t) + x \sup_{t\leq s\leq T} L(s) \int_t^T e^{-\int_t^s c(u)\,\mathrm{d}u}\,\mathrm{d}s + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right]
\end{aligned}$$

*Now recall from Lemma 2.10 (ii) that $v(t,x) \geq xg(t)$ for all $(t,x) \in [0,T) \times \mathbb{R}_+^\star$. Hence, it follows that*

$$xg(t) \leq v(t,x) \leq xg(t) + x \sup_{t\leq s\leq T} L(s) \int_t^T e^{-\int_t^s c(u)\,\mathrm{d}u}\,\mathrm{d}s + \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right].$$

*For the last inequality to be satisfied, it must be true that*

$$\mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \geq -x \sup_{t\leq s\leq T} L(s) \int_t^T e^{-\int_t^s c(u)\,\mathrm{d}u}\,\mathrm{d}s.$$

*Using arguments similar to those of the proof of Proposition 2.23(ii), we find a contradiction and therefore conclude that $v(t,x) > h(t,x)$ for all $(t,x) \in [0,T) \times \mathbb{R}_+^\star$.* □

We can now prove Proposition 2.25.

*Proof. [Proposition 2.25] To show that $\mathcal{S} \subseteq \widetilde{\mathcal{S}}$, fix $(t,x) \in \mathcal{S}$, and observe that $\varphi(t,x) = v(t,x) \geq h(t,x)$. Hence, $\varphi(t,x) = \max(\varphi(t,x), h(t,x))$. By definition of $\mathcal{S}$, $v(t,x) = \varphi(t,x) = \max(\varphi(t,x), h(t,x))$, and thus $(t,x) \in \widetilde{\mathcal{S}}$.*

*To show that $\mathcal{S} \supseteq \widetilde{\mathcal{S}}$, we show $\mathcal{C} \subseteq \widetilde{\mathcal{C}}$. Fix $(t,x) \in \mathcal{C}$, so that $v(t,x) > \varphi(t,x)$. Since $L(t) < 0$ for all $t \in [0,T)$, we have by Lemma 2.26 that $v(t,x) > h(t,x)$ for all $(t,x) \in [0,T) \times \mathbb{R}_+^\star$. Thus, $v(t,x) > \max(\varphi(t,x), h(t,x))$, which confirms that $(t,x) \in \widetilde{\mathcal{C}}$.* □

The next corollary establishes the equality of $\tau_t^x$ and $\widetilde{\tau}_t^x$ under a simple condition, complementing the previous results of Lemma 2.8.

**Corollary 2.27** *Suppose Assumption 2.4.2 holds. If $L(t) < 0$ for all $t \in [0,T)$, then $\tau_t^x = \widetilde{\tau}_t^x$.*

*Proof. The stopping times $\tau_t^x$ and $\widetilde{\tau}_t^x$ can be written as*

$$\begin{aligned}\tau_t^x &= \inf\{t \leq s \leq T | (s, F_s^{t,x}) \in \mathcal{S}\},\\ \widetilde{\tau}_t^x &= \inf\{t \leq s \leq T | (s, F_s^{t,x}) \in \widetilde{\mathcal{S}}\}.\end{aligned}$$

*The proof then follows immediately from Proposition 2.25.* □

**Remark 2.4.14** *From the results above, we can conclude that when the fee and surrender charge functions only depend on time (Assumption 2.4.2) and if $L(t) < 0$ for all $t \in [0, T)$, the optimal stopping problem with the continuous reward function* (2.13) *is equivalent to the optimal stopping problem with the discontinuous reward function* (2.7). *That is, the two problems lead to the same value function (Theorem 2.6), surrender region (Proposition 2.25), and optimal stopping time (Corollary 2.27). It also follows that the optimal surrender boundary and the continuation region will be the same for the two problems.*

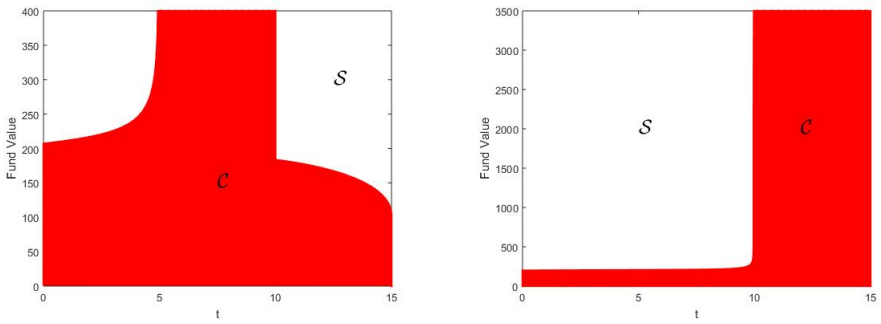
## 2.5 Numerical Examples

In this section, we give two simple examples to illustrate the differences between the optimal stopping problem involved in variable annuities pricing and the one stemming from American call and put option pricing often studied in the literature. When valuing variable annuities, the fee and the surrender charge functions can lead to a disconnected surrender region and a discontinuous optimal stopping boundary. We consider the time-dependent fee functions $c_1$ and $c_2$ defined by

$$\begin{aligned}c_1(t) &= 0.010908 \times \mathbf{1}_{\{t \leq 5\}} + 0.005454 \times \mathbf{1}_{\{5 < t \leq 10\}} + 0.010908 \times \mathbf{1}_{\{10 < t \leq T\}},\\ c_2(t) &= 0.010908 \times \mathbf{1}_{\{t \leq 10\}} + 0.005454 \times \mathbf{1}_{\{10 < t \leq T\}},\end{aligned}$$

with $T = 15$. The surrender function is set to $g(t) = e^{-\kappa(T-t)}$, with $\kappa = 0.0055$. A simple calculation shows that $L(t) > 0$ when $t \in (5, 10]$ for $c_1$ and $t \in (10, T]$ for $c_2$. The numerical example below shows that the results of the preceding sections hold, even though the fee functions $c_1$ and $c_2$ do not satisfy Assumption 2.4.1 (ii), suggesting that the assumptions of Section 2.4 are sufficient conditions and may be relaxed.

Figure 2.1 shows the continuation regions (in red) of the optimal stopping problem with the discontinuous reward function (2.7) when the fee function is modeled by $c_1$ (left panel) and $c_2$ (right panel). We first observe that the surrender regions (in white) are empty between year $5$ and $10$ for the fee function $c_1$, and between year $10$ and $15$ for $c_2$. This corresponds to the time intervals during which $L(t) > 0$, as shown theoretically in Proposition 2.23 (i). Thus, the optimal surrender boundary is discontinuous when the fee is




(a) $c_1$ (b) $c_2$

Figure 2.1: Analysis of the continuation region (in red) of the optimal stopping problem with the discontinuous reward function (2.7). The value of the variable annuity contract is approximated using the continuous-time Markov chain approximation described in the next chapter. Market and VA parameters are $r = 0.03$, $\sigma = 0.2$, $F_0 = G = 100$ and $T = 15$.

given by $c_1$ (left panel). It is already known from the literature of variable annuity that the surrender region can have different shapes depending on the fee function, see for example MacKay *et al.* (2017), Figures 4 and 5. However, such disconnected sets for the surrender region have not been observed in prior numerical work on similar problems, see among others Bernard *et al.* (2014b), MacKay *et al.* (2017), Kang and Ziveyi (2018), and the Chapter 3 of this thesis. Another important feature that is highlighted in the second example (when the fee is modeled by $c_2$) is that we no longer have $\lim_{t\to T} b(t) = G$. This is because the surrender region is empty close to maturity (since $L(t) > 0$). These simple examples illustrate well how different the optimal stopping problem involved in variable annuities pricing differs from the traditional American call and put options often studied in the literature.

## 2.6 Concluding Remarks

In this chapter, we perform a rigorous theoretical analysis of the value function involved in the pricing of a variable annuity contract with guaranteed minimum maturity benefit under the Black-Scholes setting with general fee and surrender charge functions. Because of the time dependence, the discontinuity, and the unboundedness of the reward function, many of the standard results in optimal stopping theory do not apply directly to our problem. We show that the optimal stopping problem in (2.7) admits another representation with a continuous reward function (2.13), which facilitates the study of the regularity of the value function.

In particular, the continuous reward representation allows us to use existing results from American option pricing theory to obtain the continuity of the value function. We also prove its convexity in the state variable, and its Lipschitz property in $x$ and (locally) in $t$. Then, we derive an integral expression for the early surrender premium, generalizing the results of Bernard *et al.* (2014a) to general time-dependent fee and surrender charge functions. This second representation of the value function is also known as the early exercise premium representation in the American option pricing literature. From there, we develop a third representation for the value function in terms of the current surrender value and an integral expression that only takes value in the continuation region. We call this third representation the continuation premium representation.

The continuation premium representation turns out to be very helpful in studying the shape of the surrender region. We show that the (non-)emptiness of the surrender region depends on an explicit condition that is expressed solely in terms of the fee and the surrender charge functions. This result is new to the literature and provides a better understanding of the interaction between the fees and the surrender penalty on surrender incentives. When the surrender region is nonempty, we show that $t$-sections of $\mathcal{S}$ are of the form $\mathcal{S}_t = [b(t), \infty)$, for some $b(t) \in \mathbb{R}_+^\star \cup \{\infty\}$. Investigating the regularity of the optimal surrender boundary theoretically is left as future research.

## 2.7 Appendix - Proofs

### 2.7.1 Proof of Proposition 2.12

The following Lemma is necessary to prove Proposition 2.12.

**Lemma 2.28** *Let $\widetilde{\tau}$ be a stopping time taking values in $[0,1]$ and define $c : \mathbb{R}_+ \mapsto [0,1]$. For $0 \leq s \leq t \leq T$,*

$$\Big| \int_s^{s+\widetilde{\tau}(T-s)} c(u)\,\mathrm{d}u - \int_t^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u \Big| \leq 3|t-s|.$$

*Proof. Since $s \leq t$, $s + \widetilde{\tau}(T-s) \leq t + \widetilde{\tau}(T-t)$. Hence, there are two cases to consider:*

1. *when* $s \le t \le s + \widetilde{\tau}(T-s) \le t + \widetilde{\tau}(T-t)$,

$$\begin{aligned}
&\Big| \int_t^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u - \int_s^{s+\widetilde{\tau}(T-s)} c(u)\,\mathrm{d}u \Big| \\
&\quad = \Big| \int_{s+\widetilde{\tau}(T-s)}^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u - \int_s^t c(u)\,\mathrm{d}u \Big| \le \Big| \int_{s+\widetilde{\tau}(T-s)}^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u \Big| + \Big| \int_s^t c(u)\,\mathrm{d}u \Big| \\
&\quad \le |t + \widetilde{\tau}(T-t) - s - \widetilde{\tau}(T-s)| + |t-s| \le 3|t-s|, \ \textit{(since } c \textit{ is bounded by } 1\textit{);}
\end{aligned}$$

2. *when* $s \le s + \widetilde{\tau}(T-s) \le t \le t + \widetilde{\tau}(T-t)$,

$$\begin{aligned}
&\Big| \int_t^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u - \int_s^{s+\widetilde{\tau}(T-s)} c(u)\,\mathrm{d}u \Big| \\
&\quad \le \Big| \int_t^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u \Big| + \Big| \int_s^{s+\widetilde{\tau}(T-s)} c(u)\,\mathrm{d}u \Big| \\
&\quad \le \Big| \int_{s+\widetilde{\tau}(T-s)}^{t+\widetilde{\tau}(T-t)} c(u)\,\mathrm{d}u \Big| + \Big| \int_s^t c(u)\,\mathrm{d}u \Big| \\
&\quad \le 3|t-s|.
\end{aligned}$$

□

We can now prove Proposition 2.12.

(i) We first note that under Assumption 2.4.2, the function $x \mapsto \varphi(t,\, x)$ satisfies $|\varphi(t,x) - \varphi(t,\, y)| \le |x-y|$ for all $t \in [0,\, T]$. Suppose $x > y$. Since $x \mapsto v(t,x)$ is non-decreasing (Lemma 2.11), we have

$$\begin{aligned}
|v(t,x) - v(t,\, y)| &= v(t,x) - v(t,\, y) \\
&\le \mathbb{E}\left[ e^{-r(\tau_t^x - t)} \left| \varphi(\tau_t^x, F_{\tau_t^x}^{t,x}) - \varphi(\tau_t^x, F_{\tau_t^x}^{t,y}) \right| \right] \\
&\le \mathbb{E}\left[ e^{-r(\tau_t^x - t)} \left| F_{\tau_t^x}^{t,x} - F_{\tau_t^x}^{t,y} \right| \right] \\
&\le \mathbb{E}\left[ e^{-r(\tau_t^x - t)} \left| x e^{(r-\sigma^2/2)(\tau_t^x - t) - \int_0^{\tau_t^x - t} c(t+s)\,\mathrm{d}s + \sigma W_{\tau_t^x - t}} \right.\right. \\
&\qquad \left.\left. - y e^{(r-\sigma^2/2)(\tau_t^x - t) - \int_0^{\tau_t^x - t} c(t+s)\,\mathrm{d}s + \sigma W_{\tau_t^x - t}} \right| \right] \\
&\le |x-y| \mathbb{E}\left[ e^{\sigma W_{\tau_t^x - t} - \sigma^2(\tau_t^x - t)/2} \right] \\
&= |x-y|.
\end{aligned}$$

The last equality follows from an application of Doob's optional sampling theorem.

(ii) Define $M_1 := \sup_{0\le s\le 1}|W_s|$ and let $\widetilde{\tau}\in\mathcal{T}_{0,1}$. Since, $M_1\ge 0$, for any $t\in[0,\,T]$,

$$\begin{aligned}
&\left|\sigma\sqrt{T-t}W_{\widetilde{\tau}(\omega)}(\omega)+(r-\sigma^2/2)\widetilde{\tau}(\omega)(T-t)-\int_0^{\widetilde{\tau}(\omega)(T-t)}c(t+u)\,\mathrm{d}u\right|\\
&\qquad\le \sigma\sqrt{T-t}|W_{\widetilde{\tau}(\omega)}(\omega)|+|r-\sigma^2/2|\widetilde{\tau}(\omega)(T-t)+\int_0^{\widetilde{\tau}(\omega)(T-t)}c(t+u)\,\mathrm{d}u\\
&\qquad\le \sigma\sqrt{T}M_1(\omega)+|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u, \qquad (2.44)
\end{aligned}$$

for all $\omega\in\Omega$.

Moreover, if $|z|\le r$, $|w|\le r$, for some $r>0$, then

$$|e^z-e^w|\le e^r|z-w|. \qquad (2.45)$$

This is known as the Lipschitz property of the exponential function on a disk and follows from $z^{n+1}-w^{n+1}=(z-w)\sum_{k=0}^n z^k w^{n-k}$ for all $n\in\mathbb{N}$.

Now fix $x\in\mathbb{R}_+^\star$ and $0\le s\le t\le T$. Since $v$ is non-monotone in $t$ (see Remark 2.4.2), we consider two cases.

Case 1: $v(s,x)\ge v(t,x)$. Recall from Corollary 2.5 that

$$\tau_s^x=\inf\left\{s\le u\le T\,\middle|\, v(u,F_u^{s,x})=\varphi(u,F_u^{s,x})\right\}$$

is optimal for $v(s,x)$ and let $\widetilde{\tau}=(\tau_s^x-s)/(T-s)$. Then,

$$\begin{aligned}
&|v(s,x)-v(t,\,x)|=v(s,x)-v(t,\,x)\\
&\quad\le \mathbb{E}\left[e^{-r\widetilde{\tau}(T-t)}\left(\varphi\left(\widetilde{\tau}(T-s)+s,F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)-\varphi\left(\widetilde{\tau}(T-t)+t,F^{t,x}_{t+\widetilde{\tau}(T-t)}\right)\right)\right]\\
&\quad\le \mathbb{E}\left[\left(g(\widetilde{\tau}(T-s)+s)F^{s,x}_{s+\widetilde{\tau}(T-s)}-g(\widetilde{\tau}(T-t)+t)F^{t,x}_{t+\widetilde{\tau}(T-t)}\right)\mathbf{1}_{\{\widetilde{\tau}<1\}}\right.\\
&\qquad\qquad\left.+\left(\max(G,F^{s,x}_{s+\widetilde{\tau}(T-s)})-\max(G,F^{t,x}_{t+\widetilde{\tau}(T-t)})\right)\mathbf{1}_{\{\widetilde{\tau}=1\}}\right]\\
&\quad\le \mathbb{E}\left[\left|F^{t,x}_{t+\widetilde{\tau}(T-t)}-F^{s,x}_{s+\widetilde{\tau}(T-s)}\right|\mathbf{1}_{\{\widetilde{\tau}<1\}}+\left|F^{t,x}_{t+\widetilde{\tau}(T-t)}-F^{s,x}_{s+\widetilde{\tau}(T-s)}\right|\mathbf{1}_{\{\widetilde{\tau}=1\}}\right]\\
&\quad= \mathbb{E}\left[\left|F^{t,x}_{t+\widetilde{\tau}(T-t)}-F^{s,x}_{s+\widetilde{\tau}(T-s)}\right|\right]\\
&\quad\le \mathbb{E}\left[xe^{\sigma\sqrt{T}M_1+|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u}\left|(r-\sigma^2/2)\widetilde{\tau}(s-t)\right.\right.\\
&\qquad\left.\left.+\int_s^{s+\widetilde{\tau}(T-s)}c(u)\,\mathrm{d}u-\int_t^{t+\widetilde{\tau}(T-t)}c(u)\,\mathrm{d}u+\sigma\sqrt{T-t}W_{\widetilde{\tau}}-\sigma\sqrt{T-s}W_{\widetilde{\tau}}\right|\right] \text{ (by (2.44) and (2.45))}\\
&\quad\le \mathbb{E}\left[xe^{\sigma\sqrt{T}M_1+|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u}\left\{\left|r-\sigma^2/2\right|\widetilde{\tau}|s-t|\right.\right.\\
&\qquad\left.\left.+3|s-t|+\sigma|W_{\widetilde{\tau}}||\sqrt{T-t}-\sqrt{T-s}|\right\}\right] \text{ (from Lemma 2.28)}
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[xe^{\sigma\sqrt{T}M_1+|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u}|r-\sigma^2/2|\widetilde{\tau}|s-t|\right] \\
&\qquad + 3K_1|s-t| + \mathbb{E}\left[xe^{\sigma\sqrt{T}M_1+|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u}\sigma|W_{\widetilde{\tau}}||\sqrt{T-t}-\sqrt{T-s}|\right] \\
&\leq K_1|r-\sigma^2/2||s-t| + 3K_1|s-t| + K_2\left|\sqrt{T-t}-\sqrt{T-s}\right| \\
&\leq K_1'|s-t| + K_2|\sqrt{T-t}-\sqrt{T-s}|,
\end{aligned}
$$

where the fifth to last equality follows by the scaling property of Brownian motion (see for instance Lamberton (1998), Proposition 10.3.5 (1)), and the constants are defined by $K_1 := xe^{|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u}\mathbb{E}\left[e^{\sigma\sqrt{T}M_1}\right]$, $K_2 := x\sigma e^{|r-\sigma^2/2|T+\int_0^T c(u)\,\mathrm{d}u}\mathbb{E}\left[M_1 e^{\sigma\sqrt{T}M_1}\right]$, and $K_1' := K_1\left(3+|r-\frac{\sigma^2}{2}|\right)$.

Case 2: $v(s,x) \leq v(t,x)$. Let $\widetilde{\tau}$ be defined as above. Note that the surrender function $g$ is Lipschiz, since it is $C^1$ and it has bounded first order derivatives (by hypothesis). Hence, as in the first case, we obtain

$$
\begin{aligned}
&|v(t,x)-v(s,\,x)| = v(t,x)-v(s,\,x) \\
&\quad \leq \mathbb{E}\left[e^{-r\widetilde{\tau}(T-t)}\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right) - e^{-r\widetilde{\tau}(T-s)}\varphi\left(\widetilde{\tau}(T-s)+s, F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\right] \\
&\quad = \mathbb{E}\left[\left(e^{-r\widetilde{\tau}(T-t)} - e^{-r\widetilde{\tau}(T-s)}\right)\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right)\right] \\
&\qquad + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}\left\{\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right) - \varphi\left(\widetilde{\tau}(T-s)+s, F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\right\}\right] \\
&\quad = \mathbb{E}\left[\left|e^{-r\widetilde{\tau}(T-t)} - e^{-r\widetilde{\tau}(T-s)}\right|\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right)\right] \\
&\qquad + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}\left\{\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right) - \varphi\left(\widetilde{\tau}(T-s)+s, F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\right\}\right] \\
&\quad \leq \mathbb{E}\left[e^{rT}|r\widetilde{\tau}(T-t) - r\widetilde{\tau}(T-s)|\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right)\right] \qquad (\text{by (2.45)}) \\
&\qquad + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}\left\{\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right) - \varphi\left(\widetilde{\tau}(T-s)+s, F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\right\}\right] \\
&\quad \leq re^{rT}|t-s| \times \left(G + 4xe^{rT}\right) \qquad (\text{by Lemma 2.10}) \\
&\qquad + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}\left\{\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right) - \varphi\left(\widetilde{\tau}(T-s)+s, F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\right\}\right] \\
&\quad \leq \widetilde{K}_1|t-s| + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}\left\{\varphi\left(\widetilde{\tau}(T-t)+t, F^{t,x}_{t+\widetilde{\tau}(T-t)}\right) - \varphi\left(\widetilde{\tau}(T-s)+s, F^{s,x}_{t+\widetilde{\tau}(T-s)}\right)\right\}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \widetilde{K}_1|t-s| + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}\left(g(\widetilde{\tau}(T-t)+t)F^{t,x}_{t+\widetilde{\tau}(T-t)} - g(\widetilde{\tau}(T-s)+s)F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\mathbf{1}_{\{\widetilde{\tau}<1\}}\right. \\
&\qquad \left. + e^{-r\widetilde{\tau}(T-s)}\left(\max(G, F^{t,x}_{t+\widetilde{\tau}(T-t)}) - \max(G, F^{s,x}_{s+\widetilde{\tau}(T-s)})\right)\mathbf{1}_{\{\widetilde{\tau}=1\}}\right] \\
&= \widetilde{K}_1|t-s| + \mathbb{E}\Bigg[e^{-r\widetilde{\tau}(T-s)}\Big(\left[g(\widetilde{\tau}(T-t)+t) - g(\widetilde{\tau}(T-s)+s)\right]F^{s,x}_{s+\widetilde{\tau}(T-s)}
\end{aligned}
$$

$$
\begin{aligned}
&\quad + g(\widetilde{\tau}(T-t)+t)\left[F^{t,x}_{t+\widetilde{\tau}(T-t)} - F^{s,x}_{s+\widetilde{\tau}(T-s)}\right]\Big)\mathbf{1}_{\{\widetilde{\tau}<1\}} \\
&\quad + e^{-r\widetilde{\tau}(T-s)}\left(\max(G, F^{t,x}_{t+\widetilde{\tau}(T-t)}) - \max(G, F^{s,x}_{s+\widetilde{\tau}(T-s)})\right)\mathbf{1}_{\{\widetilde{\tau}=1\}}\Bigg] \\
&\leq \widetilde{K}_1|t-s| + \mathbb{E}\left[\left[g(\widetilde{\tau}(T-t)+t) - g(\widetilde{\tau}(T-s)+s)\right]e^{-r\widetilde{\tau}(T-s)}F^{s,x}_{s+\widetilde{\tau}(T-s)}\right] \\
&\quad + \mathbb{E}\left[e^{-r\widetilde{\tau}(T-s)}g(\widetilde{\tau}(T-t)+t)\left(F^{t,x}_{t+\widetilde{\tau}(T-t)} - F^{s,x}_{s+\widetilde{\tau}(T-s)}\right)\mathbf{1}_{\{\widetilde{\tau}<1\}}\right. \\
&\quad \left. + e^{-r\widetilde{\tau}(T-s)}\left(\max(G, F^{t,x}_{t+\widetilde{\tau}(T-t)}) - \max(G, F^{s,x}_{s+\widetilde{\tau}(T-s)})\right)\mathbf{1}_{\{\widetilde{\tau}=1\}}\right] \\
&\leq \widetilde{K}_1|t-s| + (G+4x)C|t-s| + \mathbb{E}\left[\left|F^{t,x}_{t+\widetilde{\tau}(T-t)} - F^{s,x}_{s+\widetilde{\tau}(T-s)}\right|\right] \text{ (since } g \text{ is Lipschitz)} \\
&\leq \widetilde{K}_1|t-s| + (G+4x)C|t-s| + K_1'|t-s| + K_2|\sqrt{T-t} - \sqrt{T-s}| \\
&\leq \widetilde{K}_1^*|t-s| + K_2|\sqrt{T-t} - \sqrt{T-s}|,
\end{aligned}
$$

where the second to last inequality follows using the same steps as in Case 1, and with $\widetilde{K}_1 := re^{rT}(G+4xe^{rT})$, $\widetilde{K}_1^* = (G+4x)C + \widetilde{K}_1 + K_1'$, $C$ is the Lipschitz constant of $g$ and $K_1'$ and $K_2$ are the constants defined in Case 1. □

### 2.7.2 Proof of Proposition 2.23

(i) From Theorem 2.13, Remark 2.4.5 and Corollary 2.17, we have that $v \in C^{0,1}([0,T]\times\mathbb{R}_+^\star \setminus \{T\}\times\{G\}) \cap C^{1,2}([0,T]\times\mathbb{R}_+^\star \setminus \partial\mathcal{C})$, so that $v \in C^{1,2}(\text{int}(S))$. Observe that for $x \in \text{relint}(\mathcal{S}_t)$, $v(t,x) = xg(t)$, so that $xL(t) = \mathcal{L}_t v + v_t - rv$. We note that the differential operator can be applied since $x \in int(\mathcal{S})$. Recall from Proposition 2.15 that under Assumption 2.4.2, $\mathcal{L}_t v + v_t - rv \leq 0$ for all $(t,x) \in (0,T)\times\mathbb{R}_+^\star$. The result follows since $x > 0$ in $\text{int}(\mathcal{S})$.

(ii) We first show that if $L(t) < 0$ for all $t \in [0,T)$ then $\mathcal{S} \neq \emptyset$ and $\mathcal{C} \neq \emptyset$. Using the continuation premium representation (2.28), we have

$$
\begin{aligned}
v(t,x) - xg(t) = \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \\
+ \int_t^T L(s)\mathbb{E}\left[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{F_s^{t,x}\in\mathcal{C}_s\}}\right]\,\mathrm{d}s,
\end{aligned}
\tag{2.46}
$$

for all $(t,x) \in [0,T]\times\mathbb{R}_+^\star$. We want to show that- there exists some $(t_1,x_1) \in [0,T)\times\mathbb{R}_+^\star$ satisfying $v(t_1,x_1) = \varphi(t_1,x_1)$ and some $(t_2,x_2) \in [0,T)\times\mathbb{R}_+^\star$ satisfying $v(t_2,x_2) > \varphi(t_2,x_2)$. We proceed by contradiction.

First, suppose that $L(s) < 0$ for all $s \in [0,T)$ and $\mathcal{S} = \emptyset$. Fix $t \in [0,T)$. Since $\mathcal{S} = \emptyset$, we have

$\mathcal{C} = [0,T) \times \mathbb{R}_+^\star$, so that (2.46) becomes

$$\begin{aligned} & v(t,x) - \varphi(t,x) \\ & \quad = \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] + \int_t^T \left(g_t(s) - c(s)g(s)\right) x e^{-\int_t^s c(u)\,\mathrm{d}u}\,\mathrm{d}s \\ & \quad = \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] + x\left(e^{-\int_t^T c(s)\,\mathrm{d}s} - g(t)\right). \end{aligned} \tag{2.47}$$

Now note that $\mathcal{S}_t = \emptyset$ implies $v(t,x) - \varphi(t,x) > 0$ for all $x \in \mathbb{R}_+^\star$. Hence, we deduce that for all $x \in \mathbb{R}_+^\star$,

$$\mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] > x\left(g(t) - e^{-\int_t^T c(s)\,\mathrm{d}s}\right). \tag{2.48}$$

Observe now that $L(s) = g_t(s) - c(s)g(s) < 0 \Leftrightarrow \frac{\mathrm{d}\ln g(s)}{\mathrm{d}s} < c(s)$ for all $s \in [0,T)$. By integrating from $t$ to $T$ on both sides of the inequality and using the fact that $g(T) = 1$, we find that

$$g(t) - e^{-\int_t^T c(s)\,\mathrm{d}s} > 0, \tag{2.49}$$

for all $t \in [0,T)$.

The expectation on the left-hand side of (2.48) is a continuous and monotonically decreasing function of $x$, is equal to $Ge^{-r(T-t)}$ when $x = 0$ and approaches $0$ as $x \to \infty$. The function on the right-hand side is continuous (under Assumption 2.4.1), strictly increasing in $x$ and ranges from 0 to $\infty$. Then, there must exist $y \in \mathbb{R}_+^\star$ such that $\mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] \leq x\left(e^{-\int_t^T c(s)\,\mathrm{d}s} - g(t)\right)$ for all $x \geq y$, contradicting (2.48). We conclude that $\mathcal{S}_t \neq \emptyset$.

For the second part of the proof, we proceed similarly. Suppose that $L(s) < 0$ for all $s \in [0,T)$ and $\mathcal{C} = \emptyset$. Fix $t \in [0,T)$. Since $\mathcal{C} = \emptyset$, (2.46) becomes

$$v(t,x) - \varphi(t,x) = v(t,x) - xg(t) = \mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right].$$

Now note that $\mathcal{C}_t = \emptyset$ implies $v(t,x) - \varphi(t,x) = 0$ for all $x \in \mathbb{R}_+^\star$. However, we know that for all $x \in \mathbb{R}_+^\star$, $\mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] > 0$, so that $\mathcal{C}_t \neq \emptyset$, leading to a contradiction. Thus, we conclude that $\mathcal{C}_t \neq \emptyset$.

Secondly, to show that $\cup_{t_1 \leq t < t_2} \mathcal{S}_t \neq \emptyset$ for any $[t_1,t_2) \subset [0,T)$ such that $t_1 < t_2$, we proceed by contradiction. Fix $0 \leq t_1 < t_2 < T$. Suppose that $L(s) < 0$ for all $s \in [0,T)$ and $\cup_{t_1 \leq t < t_2} S_t = \emptyset$. This implies $\mathcal{C}_t = \mathbb{R}_+^\star$ for all $t \in [t_1,t_2)$, so that $v(t,x) - \varphi(t,x) > 0$ for all $(t,x) \in [t_1,t_2) \times \mathbb{R}_+^\star$. Fix $(t,x) \in [t_1,t_2) \times \mathbb{R}_+^\star$. Using the continuation premium representation (2.28), we have that

$$\begin{aligned} v(t,x) - \varphi(t,x) = \mathbb{E}\left[e^{-r(T-t)}\left(G - F_T^{t,x}\right)_+\right] \\ + \int_t^T L(s)\mathbb{E}\left[e^{-r(s-t)} F_s^{t,x} \mathbf{1}_{\{F_s^{t,x} \in \mathcal{C}_s\}}\right] \mathrm{d}s > 0, \end{aligned}$$

so that,

$$\begin{aligned}\mathbb{E}\left[e^{-r(T-t)}\left(G-F_T^{t,x}\right)_+\right] &> -\int_t^T L(s)\mathbb{E}\left[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{F_s^{t,x}\in\mathcal{C}_s\}}\right]\mathrm{d}s \\ &\geq -\sup_{t\leq s<t_2} L(s)\int_t^{t_2}\mathbb{E}\left[e^{-r(s-t)}F_s^{t,x}\mathbf{1}_{\{F_s^{t,x}\in\mathcal{C}_s\}}\right]\mathrm{d}s \\ &= -x\sup_{t\leq s<t_2} L(s)\int_t^{t_2} e^{-\int_t^s c(u)\,\mathrm{d}u}\,\mathrm{d}s \qquad (2.50)\end{aligned}$$

Note that the function on the right-hand side is positive, equal to 0 when $x = 0$, strictly increasing in $x$ (since $L(t) < 0$ by assumption) and approaches infinity as $x \to \infty$. Then, using arguments similar to those of the proof of Proposition 2.23 (ii), there must exist $y \in \mathbb{R}_+^\star$ such that for all $x \geq y$, $\mathbb{E}\left[e^{-r(T-t)}\left(G-F_T^{t,x}\right)_+\right] \leq -x\sup_{t\leq s<t_2} L(s)\int_t^{t_2} e^{-\int_t^s c(u)\,\mathrm{d}u}\,\mathrm{d}s$, contradicting the strict inequality in (2.50) which must hold for all $x \in \mathbb{R}_+^\star$. We conclude that $\mathcal{S}_t \neq \emptyset$. □

### 2.7.3 Proof of Proposition 2.24

(i) This statement is simply the contrapositive of Proposition 2.23 (i).

(ii) Observe that $\cup_{t_1\leq t<t_2}\mathcal{S}_t = \emptyset$ implies $\cup_{t_1\leq t<t_2}\mathcal{C}_t = [t_1,t_2)\times\mathbb{R}_+^\star$. Thus, we need to show that $v(t,x) > \varphi(t,x)$ for all $(t,x) \in [t_1,t_2)\times\mathbb{R}_+^\star$. Fix $(t,x) \in [t_1,t_2)\times\mathbb{R}_+^\star$ and recall from the proof of Lemma 2.20 that for any $0 < t < s \leq T$

$$e^{-r(s-t)}\varphi(s,F_s^{t,x}) = Y_s^{t,x} + e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{s=T\}},$$

where $Y^{t,x} = \{Y_s^{t,x}\}_{t\leq s\leq T}$, with $Y_s^{t,x} = e^{-r(s-t)}g(s,F_s^{t,x})F_s^{t,x}$, is the discounted surrender value process. Hence, applying Itô's formula to $Y^{t,x}$, we find that

$$\begin{aligned}\mathbb{E}\left[e^{-r(t_2-t)}\varphi(t_2,F_{t_2}^{t,x})\right] &= xg(t,x) + \int_t^{t_2}\mathbb{E}[L(s,F_s^{t,x})e^{-r(s-t)}F_s^{t,x}]\,\mathrm{d}s \\ &\quad + \mathbb{E}\left[e^{-r(T-t)}(G-F_T^{t,x})_+\mathbf{1}_{\{t_2=T\}}\right] \\ &> xg(t,x) = \varphi(t,x),\end{aligned}$$

where the last inequality follows since $L(t,x) > 0$ for all $(t,x) \in [t_1,t_2)\times\mathbb{R}_+^\star$. It follows that

$$v(t,x) \geq \mathbb{E}\left[e^{-r(t_2-t)}\varphi(t_2,F_{t_2}^{t,x})\right] > \varphi(t,x),$$

for all $(t,x) \in [t_1,t_2)\times\mathbb{R}_+^\star$, which completes the proof.

(iii) Since $\mathcal{S} = \emptyset$, $\mathcal{C} = [0,T) \times \mathbb{R}_+^\star$, $v(t,x) - \varphi(t,x) > 0$ for all $(t,x) \in [0,T) \times \mathbb{R}_+^\star$. Then, for all $(t,x) \in [0,T) \times \mathbb{R}_+^\star$, $v(t,x) - \varphi(t,x) > 0$ implies

$$\mathbb{E}\left[e^{-r(T-t)}(G - F_T^{t,x})_+\right] > x\left(g(t) - e^{-\int_t^T c(u)\,\mathrm{d}u}\right). \tag{2.51}$$

As observed in the proof of Proposition 2.23 (ii), the expectation on the left-hand side is a continuous and monotonically decreasing function of $x$, is equal to $Ge^{-r(T-t)}$ when $x = 0$ and decreases to $0$ as $x \to \infty$. Therefore, in order for the inequality to hold for all $(t,x) \in [0,T) \times \mathbb{R}_+^\star$, it must be true that

$$g(t) - e^{-\int_t^T c(u)\,\mathrm{d}u} \leq 0 \tag{2.52}$$

for all $t \in [0,T)$. □

# CHAPTER 3

# VIX-LINKED FEE INCENTIVES IN VARIABLE ANNUITIES

This chapter is based on a collaborative paper authored by Dr. Anne MacKay, Dr. Zhenyu Cui from Stevens Institute of Technology, and myself. The paper has been published in *Quantitative Finance* (see MacKay *et al*. (2023)). In this work, we consider the pricing of variable annuities with general fee structures under a class of stochastic volatility models, which includes the Heston, Hull-White, Scott, $\alpha$-Hypergeometric, $3/2$, and $4/2$ models. In particular, we analyze the impact of different VIX-linked fee structures on the optimal surrender strategy of a VA contract with GMMB using a two-layer continuous-time Markov chain to approximate the fund value process.

## 3.1 Introduction

A variable annuity (or segregated fund in Canada) is a hybrid investment instrument mainly used for retirement planning, which offers a life insurance benefit and a financial guarantee. It allows the policyholder to profit from potential gains resulting from an investment in financial markets, while offering protection against losses. The real options embedded in these products are comparable to exotic options, with the following differences: the benefit may depend on the policyholder's survival (or death), they are long-term investments (generally between 5 and 15 years, or more), and the financial guarantee is funded via a periodic fee (typically set as a percentage of the fund value) as opposed to a premium paid upfront. Different types of protection riders are offered, such as guaranteed minimum maturity benefit (GMMB), guaranteed minimum death benefit (GMDB), and guaranteed minimum withdrawal benefit (GMWB); see (Hardy, 2003) or Bauer *et al*. (2008) for details[1]. This chapter focuses on the GMMB rider, which guarantees the policyholder a minimum amount at the contract's maturity. Considering the significant size of the variable annuity market[2], the management of the risk associated with the guarantees embedded in variable annuities is a major concern for insurance companies, see Niittuinperä (2022). Indeed, variable annuities guarantees entail significant risks given their long-term structure and sensitivity to various financial and demographic risks

[1] There are no consensus among practitioners and scientists for these products' name, and thus, different authors may use different terminologies for the same product.

[2] In the United-States, the total variable annuity sales were $125 billion in 2021, representing an increase of 25% with respect to the total VA sales in 2020. Source: LIMRA Secure Retirement Institute, U.S. Individual Annuities survey `https://www.limra.com/siteassets/newsroom/fact-tank/sales-data/2021/q4/2012-2021-annuity-sales-updated.pdf`.

as well as to policyholders' behavior. For the GMMB rider, this last risk is mostly due to early surrenders.

Surrender risk refers to the uncertainty facing the insurer when a policyholder has the possibility to terminate her contract before its maturity. When she does so, she is entitled to the value accumulated in the variable annuity investment account, subject to a penalty. Kling *et al.* (2014) show that unexpected lapses can represent a significant risk for insurers. For this reason, surrender risk has raised special attention in the literature (Niittuinperä (2022), Chapter 18). Bacinello *et al.* (2011) provide a universal pricing framework for various riders and considers different types of surrender behaviors: static, i.e. the contract is never surrendered; or mixed, i.e. the policyholder acts rationally and surrenders the policy as soon as it is optimal from a risk-neutral valuation perspective. Pricing variable annuities under rational surrender behavior is equivalent to solving an optimal stopping problem and corresponds to the worst-case scenarios for insurers, in the sense that it maximizes the risk-neutral value of the contract from the policyholders' perspective. Under this assumption, Grosen and Jørgensen (1997) study the valuation of interest rate guarantees by assuming that the surrender value will be the same as the benefit value. Milevsky and Salisbury (2001) assume that the policyholder will only get a certain percentage of the fund upon surrender; this hypothesis is more in line with policies seen in practice. Under this assumption, they provide a closed-form analytical solution to the price of a GMDB with surrender in the Black-Scholes framework. In particular, they study the interaction between the surrender charges, the fee rates, and the optimal surrender level. Bernard *et al.* (2014b) study a problem similar to the one of Milevsky and Salisbury (2001), but focus on a GMMB rather than a GMDB. It is well-known that American options with finite maturity generally do not have closed-form solutions. Thus, Bernard *et al.* (2014b) used arbitrage-free techniques in the same vein as Kim (1990) and Carr *et al.* (1992) in the context of American call and put options to derive an analytical expression for the value of the right to surrender, which is analogous to the early exercise premium in American option terminology. In particular, they study the impact of different risk factors influencing the optimal surrender boundary. In that context, Bernard and MacKay (2015) provide a sufficient condition on surrender charges and fees which eliminate surrender incentives for a financially rational policyholder. Recently, Kang and Ziveyi (2018) extended the framework of Bernard *et al.* (2014b) by analyzing how the optimal surrender boundary is affected by changes in different risk factors in a stochastic interest rate and volatility model of the Heston-Hull-White type, whereas Alonso-García *et al.* (2022) incorporate taxes into their analysis.

Other fee structure designs have also been explored by different authors. The idea behind those designs is usually to reduce the insurer's exposure to various risks, such as market volatility and policyholder be-

havior. Bernard *et al.* (2014a) introduce a state-dependent fee structure, where the fee is only paid when the VA account value is below a certain level, and present an analytical formula for the value of a contract with GMMB rider (without early surrender) under this type of fee structure. MacKay *et al.* (2017) study how the fee structure and surrender charges affect the surrender region; they also design surrender charges that eliminate surrender incentives for a financially rational policyholder. Other fee designs have been explored in the literature: Delong (2014) considers a general state-dependent fee structure in a Lévy process driven market, whereas Bernard and Moenig (2019) study lapse-and-reentry in variable annuities with time-dependent fee structure. Finally, in a recent study, Wang and Zou (2021) propose a stochastic control approach to determine the optimal fee structure.

Recently, fee structures that are tied to the Chicago Board Options Exchange (CBOE) volatility index, the VIX[3], have gained attention in the literature, see Cui *et al.* (2017a) and Kouritzin and MacKay (2018). The motivations behind this new fee design come directly from the industry. In 2010, SunAmerica issued two new variable annuities whose fees were tied to the volatility index[4]. More recently, America General Life Insurance Company proposed a fee structure that is linked to the VIX for its Polaris series of variable annuities, see Polaris Platinum O-Series prospectus dated May 3rd, 2021[5]. By allowing the fees to move with the volatility index, the insurer expects to better match the cost of hedging with the premium collected. It also reduces fees for policyholders in low-volatility, rising market environments. The CBOE published two white papers, CBOE (2013b) and CBOE (2013a), illustrating how VIX-linked fee designs can be advantageous to both variable annuity providers and policyholders. Cui *et al.* (2017a) approach the question from a theoretical perspective by analyzing variable annuities without surrender with a fee structure that is tied to the VIX under a Heston-type stochastic volatility model. They provide a closed-form expression for the GMMB rider and observe that such a structure might help realign fee incomes with the value of the financial guarantee. Kouritzin and MacKay (2018) extend the works of Cui *et al.* (2017a) by applying the VIX fee designs to a GMWB rider and by adding jumps to the underlying index value process.

In this work, we allow fee structures to be as general as possible, i.e. the fee structure may depend on

[3] See `https://www.cboe.com/tradable_products/vix/`

[4] See Retirement Income Journal available at `https://retirementincomejournal.com/article/sunamerica-links-\gls{va}-rider-fees-to-volatility-index/`.

[5] See footnote 6 on p.9 of the prospectus (the long-form) available at `https://aig.onlineprospectus.net/AIG/867018103A/index.php?open=POLARIS!5fPLATINUM!5fO-SERIES!5fISP.pdf`.

time, the fund value, and also on the latent variance process, making it possible to link the fee to the VIX. In the constant fee case, it is well-known that the misalignment between the fees and the value of the financial guarantee creates an incentive for the risk-neutral, rational policyholder to surrender her policy early (see Milevsky and Salisbury (2001)). Since VIX-linked fee structures allow for better alignment of the guaranteed value with the corresponding hedging cost, we expect that such fee designs can also help reduce the insurers' exposure to surrender risk. For this reason, we numerically study the impact of three different VIX-linked fee designs on the optimal surrender strategy. To do so, we use a two-layer continuous-time Markov chain (CTMC) approximation for the fund dynamics inspired by Cui *et al.* (2018). Two-layer CTMC approximations have recently been used to price derivatives in stochastic volatility models, see Cui *et al.* (2018), Cui *et al.* (2019) and Ma *et al.* (2021), among others. The methodology proposed by Cui *et al.* (2018) for approximating two-dimensional diffusions is not only theoretically appealing and applies to most stochastic volatility models, but also is simple to implement for pricing European and American options. Their approach is especially efficient for a short/medium time horizon. However, for derivatives with very long maturities, such as those involved in variable annuities pricing, the methodology proposed by Cui *et al.* (2018) stretches the computing resource to unacceptable levels. In this chapter, we adapt their method to long-maturity cases.

The main contributions of this chapter are as follows:

- We extend the work of Cui *et al.* (2018), done in the context of options pricing, by providing novel efficient algorithms to value options with very long maturities, such as variable annuities, under general stochastic volatility models. More precisely, Algorithms 1 and 3 are new to the CTMC literature and allow to accelerate the calculation time considerably. The convergence of the new methodology is also shown theoretically, see Proposition **??**. All the algorithms provided in this chapter apply to a general class of stochastic volatility models and can be used for option pricing under other types of bi-dimensional models.
- We propose a methodology to approximate the optimal surrender surface of a VA contract with a GMMB rider when the underlying index follows a two-dimensional diffusion process. Algorithms 4 and 5, which are based on the Bermudan approximation, can adapt easily to the context of American option pricing to approximate the exercise surface under general stochastic volatility models. This new way of approximating the exercise surface is novel to CTMC literature. The advantage of this method over the current one (see, for instance, Ma *et al.* (2021)) is that the integral representation

of the value function[6] does not need to be derived to obtain the surface approximation.

- The application of CTMC approximation for variable annuity pricing is also new to the literature. Moreover, in this chapter, we analyze the optimal surrender strategy for a VA contract with a GMMB rider under the assumption that the fees are linked to the VIX index. Previously, such an analysis of surrender incentives was performed using a constant fee structure, or in a Black-Scholes framework when the fees are state-dependent. To our knowledge, this is the first time that early surrenders are analyzed conjointly with fees depending on the volatility index under a general class of stochastic volatility models.
- We derive a closed-form analytical expression for the VIX index when the variance process follows a continuous-time Markov chain. This expression can also be used to price VIX derivatives.

The remainder of the chapter is organized as follows. In Section 4.2, we introduce the market model, the VA contract, and the optimal stopping problem involved in the pricing of variable annuities with surrender. A brief introduction to CTMC approximations for a two-dimensional diffusion process is provided in Section 3.3. In Section 3.4, we apply CTMC approximations to VA contract pricing and provide new efficient algorithms. Section 3.5 provides the numerical results and discusses how VIX-linked fees affect surrender incentives. Section 3.6 concludes the chapter.

## 3.2 Financial Setting

### 3.2.1 Market Model

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, where $\mathbb{F}$ is a complete and right-continuous filtration and where $\mathbb{Q}$ denotes the pricing measure for our market, see Remark 3.2.1. We consider a risky asset, whose price can be described by the two-dimensional process $(S, V) = \{(S_t, V_t)\}_{t\geq 0}$ satisfying

$$\begin{aligned} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \sigma_S(V_t)S_t\,\mathrm{d}W_t^{(1)}, \\ \mathrm{d}V_t &= \mu_V(V_t)\,\mathrm{d}t + \sigma_V(V_t)\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{3.1}$$

with $S_0 = s_0 \in \mathbb{R}_+$ and $V_0 = v_0 \in \mathcal{S}_V$ where $\mathcal{S}_V$ denotes the state-space of $V$ (usually $\mathbb{R}$ or $\mathbb{R}_+$ depending on the model, see Table 3.1 for examples), with $r > 0$ denoting the risk-free rate and with $W =$

[6] The integral representation of the value function may be challenging to obtain under general stochastic volatility models unless making some regularity assumptions on the value function as in Ma *et al*. (2021). Indeed, the smoothness of the value function can be difficult to show under such bi-dimensional models; see Terenzi (2019) and Lamberton and Terenzi (2019).

Table 3.1: Examples of stochastic volatility models

| Model Name | Dynamics | Parameters | Cond. for Martingale Measure |
|---|---|---|---|
| Heston (1993) | $\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sqrt{V_t}S_t\,\mathrm{d}W_t^{(1)}$ | $S_0 > 0$, | No additional cond. |
| Heston (1993) | $\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_t^{(2)}$ | $\kappa, \theta, \sigma, V_0 > 0$ | Cui (2013), Proposition 2.5.1 |
| $3/2$ (1997) | $\mathrm{d}S_t = rS_t\,\mathrm{d}t + S_t/\sqrt{V_t}\,\mathrm{d}W_t^{(1)}$ | $S_0 > 0$, | $\rho \leq 0$, |
| Heston (1997) | $\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t - \sigma\sqrt{V_t}\,\mathrm{d}W_t^{(2)}$ | $\kappa, \theta, \sigma, V_0 > 0$ with $\kappa\theta \geq \sigma^2/2$ | Cui (2013), Proposition 2.5.4[7] |
| $4/2$ (2017) | $\mathrm{d}S_t = rS_t\,\mathrm{d}t + S_t\left[a\sqrt{V_t} + b/\sqrt{V_t}\right]\,\mathrm{d}W_t^{(1)}$ | $a, b \in \mathbb{R}$, $S_0 > 0$, | $\sigma^2 \leq 2\kappa\theta + \min(0, 2\rho\sigma b)$ |
| Grasselli (2017) | $\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_t^{(2)}$ | $\kappa, \theta, \sigma, V_0 > 0$, with $\kappa\theta \geq \sigma^2/2$ [8] | Grasselli (2017), Section 2.2 |
| Hull-White (1987) | $\mathrm{d}S_t = rS_t\,\mathrm{d}t + \sqrt{V_t}S_t\,\mathrm{d}W_t^{(1)}$ | $S_0 > 0$, | $\rho \leq 0$, |
| Hull and White (1987) | $\mathrm{d}V_t = \alpha V_t\,\mathrm{d}t + \beta V_t\,\mathrm{d}W_t^{(2)}$ | $\alpha, \beta, V_0 > 0$ | Jourdain (2004), Theorem 1 or Cui (2013), Proposition 2.5.10 |
| Scott (1987) | $\mathrm{d}S_t = rS_t\,\mathrm{d}t + e^{V_t}S_t\,\mathrm{d}W_t^{(1)}$ | $S_0 > 0$, | $\rho \leq 0$, |
| Scott (1987), p.426 | $\mathrm{d}V_t = \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^{(2)}$ | $\kappa, \theta, V_0 \in \mathbb{R}$, $\sigma > 0$ | Jourdain (2004), Theorem 1 |
| $\alpha$-Hypergeometric (2016) | $\mathrm{d}S_t = rS_t + e^{V_t}S_t\,\mathrm{d}W_t^{(1)}$ | $S_0 > 0$, | If $\alpha \geq 2$ or $\alpha < 2$ and either |
| Da Fonseca and Martini (2016) | $\mathrm{d}V_t = (a - be^{\alpha V_t})\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^{(2)}$ | $\alpha, b, \sigma > 0$, $a, V_0 \in \mathbb{R}$ | $\rho \leq 0$, $\alpha > 1$ or $\alpha = 1$ and $b \geq \rho\sigma$ Da Fonseca and Martini (2016), Proposition 6 |

$\{(W_t^{(1)}, W_t^{(2)})\}_{t\geq 0}$ a two-dimensional correlated Brownian motion with cross-variation $[W^{(1)},\, W^{(2)}]_t = \rho t$, where $\rho \in [-1,\, 1]$. For simplicity, $V$ will be referred to as the variance process. We assume that $\mu_V : \mathcal{S}_V \mapsto \mathbb{R}$ is continuous and that $\sigma_S,\, \sigma_V : \mathcal{S}_V \mapsto \mathbb{R}_+$ are continuously differentiable functions with $\sigma_S(\cdot) > 0$, $\sigma_V(\cdot) > 0$ on the state-space $\mathcal{S}_V$ of $V$. Further, we suppose that $\mu_V$, $\sigma_V$ and $\sigma_S$ are defined such that (3.1) has a unique-in-law weak solution.

**Remark 3.2.1** *In* (3.1)*, we start directly with the dynamics under the risk-neutral measure, hence the form of the market price of volatility risk is not necessary in our setting. However, as pointed out by Sin (1998), Jourdain (2004) and Cui (2013), a risk-neutral measure may not always exist under stochastic volatility models; additional conditions must be added to the model parameters in order for* $\{e^{-rt}S_t\}_{t\geq 0}$ *to be a true martingale under* $\mathbb{Q}$*. A list of common SV models is reported in Table 3.1, along with conditions for the martingale property to hold under the risk-neutral measure.*

### 3.2.2 Variable Annuity Contract

A policyholder enters a variable annuity contract by depositing an initial premium $F_0$ into a sub-account, which is then invested in a fund tracking the financial market. For simplicity, we will assume that the sub-account is invested in the risky asset $S$. The policyholder often has the right to surrender the contract, or lapse, prior to maturity. This additional flexibility is often called surrender option (or surrender right) in the literature and significantly complicates the valuation of variable annuity contracts. Below we discuss the risk-neutral valuation approach for a variable annuity, under both the assumption that the policyholder makes use of her surrender right or does not.

To do so, we consider a finite time horizon $T > 0$ and let $F = \{F_t\}_{0 \leq t \leq T}$ denote the variable annuity fund (or sub-account) process. Moreover, we let $C : [0, T] \times \mathbb{R}_+^\star \times \mathcal{S}_V \to \mathbb{R}_+$ denote the fee function and let the continuously compounded fee rate process $\{c_t\}_{0 \leq t \leq T}$ be defined as

$$c_t := C(t, F_t, V_t), \quad 0 \leq t \leq T, \tag{3.2}$$

where $C$ is assumed to be continuous or bounded and such that (3.4) has a unique-in-law weak solution. We allow the fee structure to be as general as possible. This setting includes, among others, state-dependent fee structures (see Bernard *et al.* (2014a), Delong (2014), MacKay *et al.* (2017)), VIX-linked fee structures (see Cui *et al.* (2017a), Kouritzin and MacKay (2018)), and time-dependent fee structures (see Bernard and Moenig (2019)).

We assume that the fees are paid continuously out of the fund at a rate $c_t$, so that the fund value is given by

$$F_t \quad = \quad S_t e^{-\int_0^t c_u \, \mathrm{d}u}, \quad 0 \leq t \leq T, \tag{3.3}$$

with $F_0 = S_0$. Using Itô's lemma, the dynamics of $F$ under the risk-neutral measure are

$$\begin{aligned} \mathrm{d}F_t &= (r - c_t) F_t \, \mathrm{d}t + \sigma_S(V_t) F_t \, \mathrm{d}W_t^{(1)}, \\ \mathrm{d}V_t &= \mu_V(V_t) \, \mathrm{d}t + \sigma_V(V_t) \, \mathrm{d}W_t^{(2)}. \end{aligned} \tag{3.4}$$

For simplicity in this chapter, we assume that interest rates are constant. However, given the long-term maturity of variable annuity contracts, it may be interesting to allow interest rates to be a deterministic

[7] The condition stated in Cui (2013), Proposition 2.5.4 is automatically satisfied by requiring the correlation parameter $\rho$ to be non-positive, as pointed out by Drimus (2012).

[8] If $b \neq 0$, then we must also impose $\kappa\theta \geq \sigma^2/2$ for the model to be well-defined.

function of time. Such an extension, discussed in Appendix 3.8, allows for the exact replication of the term structure of interest rates.

Throughout this chapter, $\mathbb{E}_{t,x,y}[\cdot]$ is short-hand notation for $\mathbb{E}[\cdot|F_t = x, V_t = y]$ and $\mathbb{E}_t[\cdot]$ for $\mathbb{E}[\cdot|\mathcal{F}_t]$, with $x \in \mathbb{R}_+$, $y \in \mathcal{S}_V$ and $t \in [0, T]$. We also use $\mathbb{E}_{x,y}[\cdot]$ to denote $\mathbb{E}_{0,x,y}[\cdot]$.

We focus on a variable annuity with a GMMB whose payoff at maturity $T > 0$ is $\max(G, F_T)$, where $G \in \mathbb{R}_+$ is a predetermined guaranteed amount. Given $(F_t, V_t) = (x, y)$, the time-$t$ risk-neutral value of the variable annuity assuming that it will not be surrendered early is

$$v_e(t, x, y) := \mathbb{E}_{t,x,y}\left[e^{-r(T-t)} \max(G, F_T)\right]. \tag{3.5}$$

On early surrender, the policyholder receives the value of the VA sub-account, reduced by a penalty which, in our setting, can depend on time and on the value of the variance process $V$. When no surrender occurs, the maturity benefit is paid at $T$.

More formally, the VA reward (or gain) function $\varphi : [0, T] \times \mathbb{R}_+ \times \mathcal{S}_V \to \mathbb{R}_+$ is defined by

$$\varphi(t, x, y) = \begin{cases} g(t, y)\, x & \text{if } t < T, \\ \max(G,\, x) & \text{if } t = T, \end{cases} \tag{3.6}$$

where $g : [0, T] \times \mathcal{S}_V \to [0, 1]$ is continuous, non-decreasing in time and satisfies $\lim_{t\to T^-} g(t, y) = 1$ $\forall y \in S_V$. In practice, we usually consider the surrender charge (as a percentage of the account value), $1 - g(\cdot, \cdot)$. A common form for the surrender charge function in the literature is $g(t, y) = e^{-k(T-t)}$ for some constant $k \geq 0$, see for example Shen *et al.* (2016), MacKay *et al.* (2017) and Kang and Ziveyi (2018). It is the first time, to the best of our knowledge, that variance-dependent surrender charges are considered.

**Remark 3.2.2** *For $x < G$, the function $t \mapsto \varphi(t, x, y)$ is discontinuous at $T$ since*

$$\lim_{t\to T^-} \varphi(t, x, y) = g(t, y)x \leq x < G = \varphi(T, x, y).$$

Under the assumption that the policyholder maximizes the risk-neutral value of her VA contract, the time-$t$ value of the variable annuity policy is given by

$$v(t, x, y) = \sup_{\tau \in \mathcal{T}_{t,\,T}} \mathbb{E}_{t,x,y}[e^{-r(\tau-t)}\varphi(\tau, F_\tau, V_\tau)], \tag{3.7}$$

where $\mathcal{T}_{t,\,T}$ is the (admissible) set of all stopping times taking values in the interval $[t,\,T]$.

Similar to the early exercise premium in the American option literature, the value of the right to surrender, denoted by $e : [0, T] \times \mathbb{R}_+ \times \mathcal{S}_V \to \mathbb{R}_+$, is defined by

$$e(t, x, y) := v(t, x, y) - v_e(t, x, y).$$

### 3.3 Continuous-Time Markov Chain Approximation

The CTMC framework outlined in this section has been proposed by Cui *et al.* (2018) for exotic option pricing under stochastic local volatility models. The general idea is to approximate the two-dimensional stock price process by a two dimensional continuous-time Markov chain. This is done by first approximating the variance process by a CTMC, and then by replacing the variance process by its CTMC approximation in the underlying price process. The resulting regime-switching diffusion process is then further approximated by a CTMC, yielding a two-dimensional CTMC process which converges weakly to the original two-dimensional diffusion process, providing that the generator of the CTMC is chosen correctly.

First, we shall briefly recall the basics of continuous-time Markov chains, following sections 6.9 and 6.10 of Grimmet and Stirzaker (2001). The stochastic process $\widetilde{X} = \{\widetilde{X}_t\}_{t\geq 0}$ taking values on some discrete state-space $\mathcal{S}_{\widetilde{X}}$ is called a continuous-time Markov chain if it satisfies the following property (a.k.a. *Markov property*):

$$\mathbb{P}(\widetilde{X}_{t_n} = \widetilde{x}_j | \widetilde{X}_{t_1} = \widetilde{x}_{i_1},\, \ldots,\, \widetilde{X}_{t_{n-1}} = \widetilde{x}_{i_{n-1}}) = \mathbb{P}(\widetilde{X}_{t_n} = \widetilde{x}_j | \widetilde{X}_{t_{n-1}} = \widetilde{x}_{i_{n-1}})$$

for all $\widetilde{x}_j,\, \widetilde{x}_{i_1},\, \ldots,\, \widetilde{x}_{i_{n-1}} \in \mathcal{S}_{\widetilde{X}}$ and any time sequence $t_1 < t_2 < \ldots < t_n$. For $0 \leq s \leq t$, we denote the *transition probability* from state $\widetilde{x}_i \in \mathcal{S}_{\widetilde{X}}$ at time $s$ to state $\widetilde{x}_j \in \mathcal{S}_{\widetilde{X}}$ at time $t$ by $p_{ij}(s,\, t) = \mathbb{P}(\widetilde{X}_t = \widetilde{x}_j | \widetilde{X}_s = \widetilde{x}_i)$. The chain is said to be *homogeneous* if $p_{ij}(s,\, t) = p_{ij}(0,\, t - s)$ for any $i, j,\, s \leq t$. In that case, we use $p_{ij}(t - s)$ to denote $p_{ij}(s,\, t)$.

Going forward, we assume that $\widetilde{X}$ is time-homogeneous and $\mathcal{S}_{\widetilde{X}}$ is finite. The family $\{\mathbf{P}_t := [p_{ij}(t)]_{|\mathcal{S}_{\widetilde{X}}| \times |\mathcal{S}_{\widetilde{X}}|}\}_{t\geq 0}$ of *transition probability matrices* is referred as the *transition semigroup* of the Markov chain.

For an infinitesimal period of length $h > 0$, it can be shown that there exist constants $\{q_{ij}\}_{1\leq i,j\leq |\mathcal{S}_{\widetilde{X}}|}$, also

called *transition rates*, such that

$$p_{ij}(h) = \begin{cases} q_{ij}h + f(h) & \text{if } i \neq j \\ 1 + q_{ij}h + f(h) & \text{if } i = j, \end{cases} \tag{3.8}$$

where $f$ is a function satisfying $\lim_{h\to 0} \frac{f(h)}{h} = 0$. From the above, we can conclude that the transition rates must satisfy

$$\begin{cases} q_{ij} \geq 0, & \text{if } i \neq j, \\ q_{ij} \leq 0, & \text{if } i = j, \end{cases} \tag{3.9}$$

and

$$\sum_{j=1}^{m} q_{ij} = 0, \quad i = 1, 2, \ldots, m. \tag{3.10}$$

The matrix $\mathbf{Q} := [q_{ij}]_{|\mathcal{S}_{\widetilde{X}}|\times|\mathcal{S}_{\widetilde{X}}|}$ is called the *generator* of $\widetilde{X}$. Under some technical conditions[9], it can be shown that the transition probability matrix $\mathbf{P}_t$ has the following matrix exponential representation:

$$\mathbf{P}_t = \exp(\mathbf{Q}t) = \sum_{k=0}^{\infty} \frac{(\mathbf{Q}t)^k}{k!}. \tag{3.11}$$

**Assumption 3.3.1** *The fee function defined in* (3.2) *is time-independent and denoted by* $c$*. That is,* $C(t, x, y) = c(x, y)$ *for all* $0 \leq t \leq T$*. Moreover, we only consider functions* $c$ *that are continuous or bounded.*

Henceforth, we consider that Assumption 3.3.1 holds. That is, we assume that the fee function is time-independent so that the fund process is time-homogeneous. For the CTMC approximation of diffusion processes with time-dependent coefficients, see Ding and Ning (2021).

### 3.3.1 Approximation of the Variance Process $\{V_t\}_{t\geq 0}$

We construct a CTMC $\{V_t^{(m)}\}_{t\geq 0}$ taking values on a finite state-space $\mathcal{S}_V^{(m)} := \{v_1, v_2, \ldots\ v_m\}$, with $v_i \in \mathcal{S}_V$ and $m \in \mathbb{N}$, that converges weakly to $\{V_t\}_{t\geq 0}$ as $m \to \infty$. Weak convergence of $V^{(m)}$ to $V$, is denoted by $V^{(m)} \Rightarrow V$.

Several approaches are available in the literature to construct the finite state-space $\mathcal{S}_V^{(m)}$, from simple uniform to non-uniform grids (see Tavella and Randall (2000), Mijatović and Pistorius (2013) and Lo and

[9] More precisely, the semigroup $\{P_t\}$ must be standard—that is, $p_{ii}(t) \to 1$ and $p_{ij}(t) \to 0$ as $t \downarrow 0$—and uniform—$\sup_i -q_{ii} < \infty$, see Grimmet and Stirzaker (2001), Definitions 6.9.4 and 6.10.3, Theorems 6.10.1, 6.10.5 and 6.10.6 for details.

Skindilias (2014) for examples of non-uniform grids). The specific grid selected for the numerical analysis performed in this chapter is discussed in more details in Section 3.5.

Once the state-space is chosen, the approximating CTMC $\{V_t^{(m)}\}_{t\geq 0}$ is defined via its generator $\mathbf{Q}^{(m)} = [q_{ij}]_{m\times m}$. This generator is constructed so that the first two moments of the transition density of the variance process $\{V_t\}_{t\geq 0}$ and of the approximating CTMC $\{V_t^{(m)}\}_{t\geq 0}$ coincide; these are the so-called *local consistency conditions*, see Kushner (1990) and Lo and Skindilias (2014). More precisely, the elements $q_{ij}$, $1 \leq i, j \leq m$ of the generator of $V^{(m)}$ are chosen so that for a small time increment $h \ll T$,

$$\begin{aligned} \mathbb{E}_t\left[V_{t+h}^{(m)} - V_t^{(m)}\right] &= \mathbb{E}_t\left[V_{t+h} - V_t\right] \simeq \mu_V(V_t)h, \qquad \text{and} \\ \mathbb{E}_t\left[\left(V_{t+h}^{(m)} - V_t^{(m)}\right)^2\right] &= \mathbb{E}_t\left[\left(V_{t+h} - V_t\right)^2\right] \simeq \sigma_V^2(V_t)h, \end{aligned} \tag{3.12}$$

for all $t \geq 0$. To ensure that the local consistency conditions are satisfied, we use the generator proposed by Lo and Skindilias (2014) and given by[10]

$$q_{ij} = \begin{cases} \frac{\sigma_V^2(v_i) - \delta_i \mu_V(v_i)}{\delta_{i-1}(\delta_{i-1}+\delta_i)}, & j = i-1, \\ -q_{i,i-1} - q_{i,i+1}, & j = i, \\ \frac{\sigma_V^2(v_i) + \delta_{i-1}\mu_V(v_i)}{\delta_i(\delta_{i-1}+\delta_i)}, & j = i+1, \\ 0, & j \neq i,\, i-1,\, i+1, \end{cases} \tag{3.13}$$

for $2 \leq i \leq m-1$ and $1 \leq j \leq m$ and where $\delta_i = v_{i+1} - v_i$, $i = 1, 2, \ldots, m-1$. On the borders, we set $q_{12} = \frac{|\mu_V(v_1)|}{\delta_1}$, $q_{11} = -q_{12}$, $q_{m,m-1} = \frac{|\mu_V(v_m)|}{\delta_{m-1}}$, $q_{m,m} = -q_{m,m-1}$; and $0$ elsewhere. The transition rates on the boundaries of the state-space are set so that the absolute instantaneous means are maintained at the endpoints. Other schemes could have also been employed (see Chourdakis (2004), Mijatović and Pistorius (2013)), but we observed that all of these schemes are equivalent numerically.

To obtain a well-defined $\mathbf{Q}^{(m)}$ matrix, the transition rates in (3.13) must also satisfy the conditions in (3.9). Hence, for $2 \leq i \leq m-1$, we must have

$$\begin{cases} \delta_{i-1} \leq \frac{\sigma_V^2(v_i)}{|\mu_V(v_i)|}, & \text{if } \mu_V(v_i) < 0 \\ \delta_i \leq \frac{\sigma_V^2(v_i)}{\mu_V(v_i)}, & \text{if } \mu_V(v_i) > 0. \end{cases} \tag{3.14}$$

If $\mu_V(v_i) = 0$, then no additional condition needs to be added.

[10] An advised reader will notice some differences between the transition rates stated above, and the ones that appear in Lo and Skindilias (2014). However, one can show that the two rate matrices are equivalent with some simple algebra.

**Remark 3.3.1** *A sufficient condition for* (3.14) *to hold is*

$$\max_{1\leq i\leq m-1}\delta_i \leq \min_{1\leq i\leq m-1}\frac{\sigma_V^2(v_i)}{|\mu_V(v_i)|}. \tag{3.15}$$

### 3.3.2 Approximation of the Fund Value Process $\{F_t\}_{0\leq t\leq T}$

The CTMC approximating $\{F_t\}_{t\geq 0}$ is constructed by first replacing the variance process appearing in the drift and diffusion coefficients by their CTMC approximations, and then by further approximating the resulting regime-switching diffusion process by another CTMC. The resulting two-dimensional regime-switching CTMC can then be mapped to a one-dimensional CTMC on an enlarged state-space.

Lemma 3.1 below allows for the removal of the correlation between the Brownian motions in (3.4), which is necessary to construct the CTMC approximation of $\{F_t\}_{t\geq 0}$.

**Lemma 3.1 (Lemma 1 of Cui *et al.* (2018))** *Let $F$ and $V$ be defined as in* (3.4). *Define $\gamma(x) := \int_\cdot^x \frac{\sigma_S(u)}{\sigma_V(u)}\,\mathrm{d}u$ and $X_t := \ln(F_t) - \rho\gamma(V_t)$, $t\in[0,T]$. Then $X$ satisfies*

$$\begin{aligned}\mathrm{d}X_t =& \mu_X(X_t,V_t)\,\mathrm{d}t + \sigma_X(V_t)\,\mathrm{d}W_t^*\\ \mathrm{d}V_t =& \mu_V(V_t)\,\mathrm{d}t + \sigma_V(V_t)\,\mathrm{d}W_t^{(2)},\end{aligned} \tag{3.16}$$

*where $W_t^* := \frac{W_t^{(1)}-\rho W_t^{(2)}}{\sqrt{1-\rho^2}}$ is a standard Brownian motion independent of $W_t^{(2)}$, $\sigma_X(y) := \sqrt{1-\rho^2}\sigma_S(y)$, $\mu_X(x,y) := r - c(e^{x+\rho\gamma(y)},y) - \frac{\sigma_S^2(y)}{2} - \rho\psi(y)$ and*

$$\begin{aligned}\psi(y) &:= \mathcal{L}_v\gamma(y) = \mu_V(y)\gamma'(y) + \frac{1}{2}\sigma_V^2(y)\gamma''(y)\\ &= \mu_V(y)\frac{\sigma_S(y)}{\sigma_V(y)} + \frac{1}{2}\left[\sigma_V(y)\sigma_S'(y) - \sigma_V'(y)\sigma_S(y)\right],\end{aligned}$$

*for $x\in\mathbb{R}, y\in\mathcal{S}_V$.*

The proof relies on the multidimensional Itô formula (see Lemma 1 of Cui *et al.* (2018) for details). Given the CTMC approximation of the process $V^{(m)}$ and its generator $\mathbf{Q}^{(m)}$, the diffusion process in (3.16) can now be approximated by a regime-switching diffusion process $\{X_t^{(m)}\}_{t\geq 0}$:

$$\mathrm{d}X_t^{(m)} = \mu_X(X_t^{(m)},V_t^{(m)})\,\mathrm{d}t + \sigma_X(V_t^{(m)})\,\mathrm{d}W_t^*, \tag{3.17}$$

where regimes are determined by the states of the approximated variance process, $\{v_1,v_2,\ldots,v_m\}$. To construct a regime-switching CTMC $(X_t^{(m,N)},V_t^{(m)})$ approximating the regime-switching diffusion $(X_t^{(m)},V_t^{(m)})$,

we fix a state for the variance process $V_t^{(m)}$ (or equivalently a regime) and construct a CTMC approximation for $X_t^{(m)}$ given that $V_t^{(m)}$ is in that state. This is done using the procedure described in Section 3.3.1 for a one-dimensional diffusion process. The procedure is then repeated for each state in $\mathcal{S}_V^{(m)}$, and the approximating CTMCs are combined with $V^{(m)}$ to obtain the final regime-switching CTMC.

More precisely, let $X_t^{(m,N)}$ be the CTMC approximation of $X_t^{(m)}$ taking values on a finite state-space $\mathcal{S}_X^{(N)} = \{x_1, x_2, \ldots, x_N\}$, $N \in \mathbb{N}$. For each $v_l \in \mathcal{S}_V^{(m)}$, we define the generator $\mathbf{G}_l^{(N)} = [\lambda_{ij}^l]_{N\times N}$ of $X_t^{(m,N)}$ given that the variance process is in state $v_l$ at time $t \geq 0$ by

$$\lambda_{ij}^l = \begin{cases} \frac{\sigma_X^2(v_l) - \delta_i^x \mu_X(x_i, v_l)}{\delta_{i-1}^x(\delta_{i-1}^x + \delta_i^x)} & j = i-1 \\ -\lambda_{i,i-1}^l - \lambda_{i,i+1}^l & j = i \\ \frac{\sigma_X^2(v_l) + \delta_{i-1}^x \mu_X(x_i, v_l)}{\delta_i^x(\delta_{i-1}^x + \delta_i^x)} & j = i+1 \\ 0 & j \neq i,\, i-1,\, i+1, \end{cases} \tag{3.18}$$

for $2 \leq i \leq N-1$ and $1 \leq j \leq N$, where $\delta_i^x = x_{i+1} - x_i$, $i = 1, 2, \ldots, N-1$. On the boundaries, we set $\lambda_{12}^l = \frac{|\mu_X(x_1, v_l)|}{\delta_1^x}$, $\lambda_{11}^l = -\lambda_{12}^l$, $\lambda_{N,N-1}^l = \frac{|\mu_X(x_N, v_l)|}{\delta_{N-1}^x}$, $\lambda_{N,N}^l = -\lambda_{N,N-1}^l$, and $0$ elsewhere.

Using $V^{(m)}$ and the relation presented in Lemma 3.1, the approximated fund process $F^{(m,N)}$, which approximates $F$, is defined by

$$F_t^{(m,N)} := \exp\left\{ X_t^{(m,N)} + \rho\gamma(V_t^{(m)}) \right\}, \quad 0 \leq t \leq T. \tag{3.19}$$

**Remark 3.3.2 (Convergence of the approximation)** *Such a construction of the regime-switching CTMC ensures that the two-dimensional process $(X_t^{(m,N)}, V_t^{(m)})$ converges weakly to $(X_t, V_t)$ as $m, N \to \infty$. The main idea is to show that the generator of $(X_t^{(m,N)}, V_t^{(m)})$ is uniformly close to the infinitesimal generator of $(X_t, V_t)$ as $m, N \to \infty$, to then conclude that $(X_t^{(m,N)}, V_t^{(m)}) \Rightarrow (X_t, V_t)$ using the results of Ethier and Kurtz (2005) which relies on semi-group theory. Moreover, since the function $h : \mathbb{R} \times \mathcal{S}_V \to \mathbb{R}_+$ defined by $h(x, y) = e^{x+\rho\gamma(y)}$ is continuous, we have that $F^{(m,N)} \Rightarrow F$ by the continuous mapping Theorem, see Billingsley (1999) Theorem 2.7. For one-dimensional processes, intuition and detailed explanations of the proof can be found in Mijatović and Pistorius (2013), Section 5 (or in the unabridged version of the paper Mijatović and Pistorius (2009), Section 6); for stochastic volatility models, see Cui* et al. *(2018), Section 2.4.*

The last step is to convert the regime-switching CTMC $(X_t^{(m,N)}, V_t^{(m)})$ into a one-dimensional CTMC process $Y_t^{(m,N)}$ on an enlarged state-space $\mathcal{S}_Y^{(m,N)} := \{1, 2, \ldots, mN\}$. This is done in Theorem 1 of Cai *et al.*

(2019), reproduced below.

**Proposition 3.2 (Theorem 1 of Cai *et al*. (2019))** *Consider a regime-switching CTMC* $(X^{(m,N)}, V^{(m)})$ *taking values in* $\mathcal{S}_X^{(N)} \times \mathcal{S}_V^{(m)}$ *where* $\mathcal{S}_X^{(N)} = \{x_1, x_2, \ldots, x_N\}$ *and* $\mathcal{S}_V^{(m)} = \{v_1, v_2, \ldots, v_m\}$*; and another one-dimensional CTMC,* $\{Y_t^{(m,N)}\}_{0\le t\le T}$*, taking values in* $\mathcal{S}_Y^{(m,N)} := \{1, 2, \ldots, mN\}$ *and its transition rate matrix* $\mathbf{G}^{(m,N)}$ *given by*

$$\begin{pmatrix} q_{11}\mathbf{I}_N + \mathbf{G}_1^{(N)} & q_{12}\mathbf{I}_N & \cdots & q_{1m}\mathbf{I}_N \\ q_{21}\mathbf{I}_N & q_{22}\mathbf{I}_N + \mathbf{G}_2^{(N)} & \cdots & q_{2m}\mathbf{I}_N \\ \vdots & \vdots & \ddots & \vdots \\ q_{m1}\mathbf{I}_N & q_{m2}\mathbf{I}_N & \cdots & q_{mm}\mathbf{I}_N + \mathbf{G}_m^{(N)} \end{pmatrix}, \tag{3.20}$$

*where* $\mathbf{I}_N$ *is the* $N \times N$ *identity matrix,* $\mathbf{G}_l^{(N)} = [\lambda_{ij}^l]_{N\times N}$*,* $l = 1, 2, \ldots, m$ *and* $\mathbf{Q}^{(m)} = [q_{ij}]_{m\times m}$ *are the generators defined in* (3.18) *and* (3.13)*, respectively. Define the function* $\psi : \mathcal{S}_X^{(N)} \times \mathcal{S}_V^{(m)} \to \mathcal{S}_Y^{(m,N)}$ *by* $\psi(x_n, v_l) = (l-1)N + n$ *and its inverse* $\psi^{-1} : \mathcal{S}_Y^{(m,N)} \mapsto \mathcal{S}_X^{(N)} \times \mathcal{S}_V^{(m)}$ *by* $\psi^{-1}(n_y) = (x_n, v_l)$ *for* $n_y \in \mathcal{S}_Y^{(m,N)}$*, where* $n \le N$ *is the unique integer such that* $n_y = (l-1)N + n$ *for some* $l \in \{1, 2, \ldots, m\}$*. Then, we have*

$$\mathbb{E}\left[\Psi(X^{(m,N)}, V^{(m)}) | X_0^{(m,N)} = x_i, V_0^{(m)} = v_k\right] = \mathbb{E}\left[\Psi(\psi^{-1}(Y^{(m,N)})) | Y_0 = (k-1)N + i\right],$$

*for any path-dependent function* $\Psi$ *such that the expectation on the left-hand side is finite.*

## 3.4 Variable Annuity Pricing via CTMC Approximation

In this section, we use the CTMC approximation of the fund value process to price variable annuities under different surrender strategies. We provide a simple way to approximate the optimal surrender surface, which is the extension in three dimensions of the exercise boundary for two-dimensional processes. The present section extends to variable annuity pricing the work of Cui *et al*. (2018) and Cui *et al*. (2019), in which a CTMC approximation is used for option pricing.

Recall that $(X^{(m,N)}, V^{(m)})$ is the regime-switching CTMC approximation of $(X, V)$ (see Lemma 3.1) taking values in a finite state-space $\mathcal{S}_X^{(N)} \times \mathcal{S}_V^{(m)}$ where $\mathcal{S}_X^{(N)} = \{x_1, x_2, \ldots, x_N\}$, $N \in \mathbb{N}$; and $\mathcal{S}_V^{(m)} = \{v_1, v_2, \ldots, v_m\}$, $m \in \mathbb{N}$. We have also defined $F^{(m,N)}$ in terms of $(X^{(m,N)}, V^{(m)})$ in (3.19).

Throughout this section, we denote by $\{\mathbf{e}_{ik}\}_{i,k=1}^{N,m}$ the standard basis in $\mathbb{R}^{mN}$, i.e. $\mathbf{e}_{ik}$ represents a row vector of size $1 \times mN$ having a value of $1$ in the $(k-1)N + i$-th entry and $0$ elsewhere.

### 3.4.1 Variable Annuity without Early Surrenders

Consider an initial premium $F_0 > 0$ and let $V_0 \in \mathcal{S}_V$. The risk-neutral value of a variable annuity contract assuming no early surrenders can be approximated by

$$\begin{aligned} v_e(0, F_0, V_0) &= \mathbb{E}\left[e^{-rT}\max(G, F_T)|F_0, V_0\right] \\ &\approx \mathbb{E}\left[e^{-rT}\max(G, F_T^{(m,N)})|X_0^{(m,N)} = x_i, V_0^{(m)} = v_k\right]. \end{aligned} \tag{3.21}$$

Here, we assume[11] that $x_i \in \mathcal{S}_X^{(N)}$ and $v_k \in \mathcal{S}_V^{(m)}$, with $x_i = \ln(F_0) - \rho\gamma(V_0)$.

**Proposition 3.3** *Let $F_0 > 0$ be the initial premium, with $X_0^{(m,N)} = \ln(F_0) - \rho\gamma(V_0) = x_i \in \mathcal{S}_X^{(N)}$ and $V_0 = v_k \in \mathcal{S}_V$. The risk-neutral value at time $0$ of a variable annuity contract held until maturity $T$ and with guaranteed amount $G > 0$ can be approximated by*

$$\begin{aligned} v_e^{(m,N)}(0, F_0, V_0) &:= \mathbb{E}[e^{-rT}\max(G, F_T^{(m,N)})|F_0^{(m,N)} = F_0, V_0^{(m)} = V_0] \\ &= e^{-rT}\mathbf{e}_{ik}\exp\{T\mathbf{G}^{(m,N)}\}\mathbf{H}, \end{aligned} \tag{3.22}$$

*where $\mathbf{G}^{(m,N)}$ is defined in* (3.20) *and $\mathbf{H}$ is a column vector of size $mN \times 1$ whose $(l-1)N + n$-th entry is given by*

$$h_{(l-1)N+n} = \max\left(G, e^{x_n + \rho\gamma(v_l)}\right), \quad 1 \le l \le m,\ 1 \le n \le N. \tag{3.23}$$

The proof follows by noticing that (3.22) is the matrix representation of the conditional expectation of a (function of a) discrete one-dimensional random variable whose conditional probability mass function is given by the transitional probabilities $p_{(k-1)N+i,j}(T)$, $1 \le j \le mN$, with $\mathbf{P}(T) = [p_{ij}(T)]_{mN\times mN} = \exp\{T\mathbf{G}^{(m,N)}\}$ as per (3.11).

**Remark 3.4.1 (Convergence of VA prices without early surrenders)** *From Remark 3.3.2, we know that $F^{(m,N)} \Rightarrow F$, as $m, N \to \infty$. From the continuous mapping theorem, we also have that $\varphi(T, F_T^{(m,N)}, V_T^{(m)}) \Rightarrow \varphi(T, F_T, V_T)$. For derivatives with a continuous and bounded payoff, convergence of the prices follows directly from the definition of weak convergence, see for example Billingsley (1995), Theorem 25.8, (i) and (ii). However, convergence of the prices for derivatives with unbounded payoffs is*

[11] If $X_0^{(m,N)}$ and $V_0^{(m)}$ are not part of their respective grids, then the two points can be added to the grids, or the option price must be linearly interpolated between grid points, see Remark 3.5.1 for details.

*not as clear. We note that if $F^{(m,N)}$ is a $\mathbb{Q}$-martingale for each $m, N$, then $\{\varphi(T, F_T^{(m,N)}, V_T^{(m)})\}_{m,N=1}^{\infty}$ is uniformly integrable and the convergence of the prices follows. Showing this property is however out of the scope of this chapter. In the one-dimensional setting, convergence of the price approximations is discussed in Zhang and Li (2021). Detailed error and convergence analysis of the two-dimensional CTMC approximation for European call options is performed in Ma* et al. *(2022).*

#### 3.4.1.1 Fast Algorithm

When considering a typical time horizon of $10, 15$ or $20$ years as is often the case in VA pricing, the probability that the fund or the volatility processes reaches high (resp. low) value is higher than when shorter maturities are concerned, and so the grid's upper (resp. lower) bound must be set to a higher (resp. lower) value. This complicates the pricing of VAs compared to financial options (generally written for short or medium time-horizon), since more discretization points $m$ and $N$ are needed in order to capture the distribution of the variance and the fund value process correctly. Theoretically, this is not a problem; however, several numerical issues can occur when implementing the pricing formula numerically. First, the generator matrix $\mathbf{G}^{(m,N)}$ can become very large and thus require a large amount of storage space, which may cause memory problems. Second, calculating the exponential of a large sparse $mN \times mN$ matrix over a long time horizon is time-consuming.

Using the tower property of conditional expectations and an approximation based on the assumption that the variance process remains constant over small time periods (see Proposition 3.4), we propose a new algorithm that speeds up the pricing of the VA contract without early surrenders (see Algorithm 1).

The approximation used in Algorithm 1 follows from Proposition 3.4 presented below.

**Proposition 3.4** *Let $h > 0$ with $h \ll T$ and $0 \leq t \leq T - h$. For any function $\phi$ such that the expectation on the left-hand side of* (3.24) *is finite, we have that*

$$\begin{aligned}
&\mathbb{E}\left[\phi\left(t+h, X_{t+h}^{(m,N)}, V_{t+h}^{(m)}\right) | X_t^{(m,N)} = x_i, V_t^{(m)} = v_k\right] \\
&= \sum_{j=1}^{m} \mathbb{E}\left[\phi(t+h, X_{t+h}^{(m,N)}, V_{t+h}^{(m)}) | V_t^{(m)} = V_{t+h}^{(m)} = v_j, X_t^{(m,N)} = x_i\right] \times \mathbb{P}\left(V_{t+h}^{(m)} = v_j | V_t^{(m)} = v_k\right) + f(h),
\end{aligned} \tag{3.24}$$

*where $f(h)$ is a function satisfying $\lim_{h\to 0} \frac{f(h)}{h} = 0$.*

The proof of Proposition 3.4 is reported in Appendix 3.7.1.

The above proposition allows the separation of the matrices $\{\mathbf{G}_j^{(N)}\}_{j=1}^m$ and $\mathbf{Q}^{(m)}$, so that the new algorithm now requires $m$ times the calculation of the exponential of $N \times N$ matrices and one time the exponential of a $m \times m$ matrix over small time-intervals. Hence, by reducing the size of the matrix in the matrix exponential and the length of the time interval over which the exponential is calculated leads to a significant reduction in computation time and to more effective management of the memory space. Note also that the added cost of computing $m+1$ matrix exponentials (rather than only one) is counterbalanced by the reduced cost of computing the exponential of smaller matrices over a short time interval $h$.

The following notation is used in Algorithm 1 below.

1. We use $M \in \mathbb{N}$ time steps of length $\Delta_M = T/M$.
2. $\mathbf{B} = [b_{jn}]_{j,n=1}^{m,N}$ denotes a matrix of size $m \times N$, containing the value of the VA contract. More precisely, the matrix $\mathbf{B}$ is updated at each time step, so that after the first iteration, $b_{jn} \approx \mathbb{E}\left[e^{-r\Delta_M}\varphi(T, e^{X_T^{(m,N)}+\rho\gamma(V_T^{(m)})}, V_T^{(m)})|X_{T-\Delta_M}^{(m,N)} = x_n, V_{T-\Delta_M}^{(m)} = v_j\right]$; after the second iteration, $b_{jn} \approx \mathbb{E}\left[e^{-2r\Delta_M}\varphi(T, e^{X_T^{(m,N)}+\rho\gamma(V_T^{(m)})}, V_T^{(m)})|X_{T-2\Delta_M}^{(m,N)} = x_n, V_{T-2\Delta_M}^{(m)} = v_j\right]$, and so on.
3. $\mathbf{B}_{*,n} = [b_{jn}]_{j=1}^m$ denotes the $n$-th column of $\mathbf{B}$, $n = 1, 2, \ldots, N$,
4. $\mathbf{B}_{j,*} = [b_{jn}]_{n=1}^N$ denotes the $j$-th row of $\mathbf{B}$, $j = 1, 2, \ldots, m$.
5. The symbol $\top$ indicates the matrix (vector) transpose operation.

**Algorithm 1:** Variable Annuity without Early Surrenders via CTMC Approximation – Fast Algorithm

**Input:** Initialize $\mathbf{Q}^{(m)}$ as in (3.13) and $\mathbf{G}_j^{(N)}$ for $j = 1, 2, \ldots, m$, as in (3.18)

$M \in \mathbb{N}$, the number of time steps,

$\Delta_M \leftarrow T/M$, the size of a time step

Set $\mathbf{B}_{j,*} \leftarrow \left[\varphi(T, e^{x_n + \rho\gamma(v_j)}, v_j)\right]_{n=1}^{N}, j = 1, 2, \ldots, m$

`/* Calculate the transition probability matrices */`

**for** $j = 1, 2, \ldots, m$ **do**

  `/* Adjusted transition probability matrix of` $X^{(m,N)}$ `given` $V^{(m)} = v_j$ `over a period of length` $\Delta_M$ `*/`

  $\mathbf{P}_j^X \leftarrow e^{\mathbf{G}_j^{(N)}\Delta_M} e^{-r\Delta_M}$

`/* Transition probability matrix of` $V^{(m)}$ `over a period of length` $\Delta_M$ `*/`

$\mathbf{P}^V \leftarrow e^{\mathbf{Q}^{(m)}\Delta_M}$

`/* VA valuation */`

**for** $z = M - 1, \ldots, 0$ **do**

  **for** $j = 1, 2, \ldots, m$ **do**

    $\widetilde{\mathbf{H}}_{*,j} \leftarrow \mathbf{P}_j^X \mathbf{B}_{j,*}^\top$

  **for** $n = 1, 2, \ldots, N$ **do**

    $\mathbf{B}_{*,n} \leftarrow \mathbf{P}^V \widetilde{\mathbf{H}}_{n,*}^\top$

**return** $b_{ki}$

The computational gain of using Algorithm 1 over the previous algorithm comes at the cost of a loss of accuracy since the conditional expectations are approximated over small time intervals (refer to Proposition 3.4). Numerical experiments below show that highly accurate results are obtained in seconds when the time step is small, but the algorithm can perform poorly when the time step is not small enough. This new algorithm can be easily adapted to a wide range of payoff functions and extends the work of Cui *et al*. (2018) by accelerating the calculation time considerably. In particular, it allows for efficient pricing of long-maturity derivatives.

**Remark 3.4.2** *Since at each time step, Algorithm 1, or the "Fast Algorithm", takes advantage of the tower property of conditional expectations over short time periods of the same length* $\Delta_M$*, the transition probability matrix can be pre-computed at the beginning of the procedure and stored, accelerating the numerical*

*process greatly.*

### 3.4.2 Variable Annuity with Early Surrenders

We approximate the value of the VA contract (including the right to surrender) by its Bermudan counterpart for a large number of monitoring dates. Bermudan options can be exercised early, but only at predetermined dates $R \subset [0, T]$. Thus, Bermudan options are similar to American options, but the region of the permitted exercise times is a subset of $[0, T]$ containing a finite number of exercise dates, $\{t_0, t_1, \ldots, t_M\}$ with $t_z \in [0, T]$, $z = 0, 1, 2, \ldots, M$ for some $M \in \mathbb{N}$.

In this chapter, we use the term Bermudan (resp. American) contract to refer to a variable annuity under which the policyholder has the right to surrender her contract prior to maturity on predetermined dates (resp. at any time prior to maturity). In the same vein, a variable annuity without surrender rights is also called a European contract. Note that these terms do not refer to existing contracts, and they are used to simplify explanations. Naturally, as $M \to \infty$, we expect the price of the Bermudan contract to converge to the one of a variable annuity with surrender rights as defined in (3.7). The latter is formalized in the following.

Let $\Delta_M = T/M$ for some $M \in \mathbb{N}$ and define the set $\mathcal{H}_M = \{t_0, t_1, \ldots, t_M\}$ where $t_z = z\Delta_M$, $z = 0, 1, \ldots, M$, so that $t_0 = 0$ and $t_M = T$. The time-$t$ risk-neutral value of the Bermudan contract with permitted exercise dates $\mathcal{H}_M$ is defined by

$$b_M(t, x, y) = \sup_{\tau \in \mathcal{T}_{\Delta_M}, \tau \geq t} \mathbb{E}_{t,x,y}[e^{-r(\tau - t)} \varphi(\tau, F_\tau, V_\tau)] \tag{3.25}$$

where $\mathcal{T}_{\Delta_M}$ is the set of stopping times taking values in $\mathcal{H}_M$. Proposition 3.5 below shows that $b_M(t, x, y) \to v(t, x, y)$ as $M \to \infty$.

**Proposition 3.5** *As $M \to \infty$, the value function of the Bermudan variable annuity contract* (3.25) *converges to its American counterpart* (3.7)*, that is,*

$$\lim_{M \to \infty} b_M(t, x, y) = v(t, x, y).$$

The proof can be found in Zhao (2017), Proposition 4.6, under the Barndorff-Nielsen Shephard model. Extending this proof to stochastic volatility models is straightforward. The key idea behind the proof relies on

the fact that the price of a Bermudian option is always less than or equal to the price of an American option, and the price of the Bermudian option increases as the number of exercise dates increases. This results in a sequence of option prices that is both monotone and bounded. The final step is to demonstrate that the limit of this sequence equals the price of the American option.

**Remark 3.4.3** *In the practical context of variables annuities, surrenders are often only allowed at specific times (such as on the policy anniversary dates). In these cases, the Bermudan contract may be more realistic than its American counterpart presented in* (3.7).

We denote the Bermudan contract value process by $B := \{B_z := b_M(t_z, F_{t_z}, V_{t_z})\}_{0\leq z\leq M}$. From the principle of dynamic programming (see for example Lamberton (1998), Theorem 10.1.3), it is well-known that the discretized problem admits the following representation:

$$\begin{cases} B_M &= \varphi(T, F_T, V_T) \\ B_z &= \max\left(\varphi(t_z, F_{t_z}, V_{t_z}), e^{-r\Delta_M}\mathbb{E}_{t_z}[B_{z+1}]\right), \qquad 0 \leq z \leq M-1. \end{cases} \tag{3.26}$$

Using CTMCs, we can define an approximation for the time-$t$ risk-neutral value of the Bermudan contract by

$$b_M^{(m,N)}(t,x,y) \quad = \quad \sup_{\tau\in\mathcal{T}_{\Delta_M},\tau\geq t} \mathbb{E}\left[e^{-r\tau}\varphi(\tau, F_\tau^{(m,N)}, V_\tau^{(m)})|F_t^{(m,N)} = x, V_t^{(m)} = y\right].$$

The approximation of the Bermudan contract value process, denoted by $B^{(m,N)} := \{B_z^{(m,N)} := b_M^{(m,N)}(t_z, F_{t_z}^{(m,N)}, V_{t_z}^{(m)})\}_{0\leq z\leq M}$, is thus given by

$$\begin{cases} B_M^{(m,N)} &= \varphi\left(T, F_T^{(m,N)}, V_T^{(m)}\right) \\ B_z^{(m,N)} &= \max\left(\varphi\left(t_z, F_{t_z}^{(m,N)}, V_{t_z}^{(m)}\right), e^{-r\Delta_M}\mathbb{E}_{t_z}[B_{z+1}^{(m,N)}]\right), \qquad 0 \leq z \leq M-1. \end{cases} \tag{3.27}$$

Or equivalently in terms of the process $Y^{(m,N)}$, we have that

$$\begin{cases} B_M^{(m,N)} &= \varphi(T, \psi^{-1}(Y_T^{(m,N)})) \\ B_z^{(m,N)} &= \max\left(\widetilde{\varphi}(t_z, \psi^{-1}(Y_{t_z}^{(m,N)})), e^{-r\Delta_M}\mathbb{E}_{t_z}[B_{z+1}^{(m,N)}]\right), \qquad 0 \leq z \leq M-1, \end{cases} \tag{3.28}$$

where $\psi^{-1}$ is defined in Proposition 3.2[12]. Finally, we have $b_M(0, F_0, V_0) = B_0 \approx B_0^{(m,N)} = b_M^{(m,N)}(0, F_0, V_0)$. Based on the above, an approximation for the value of the Bermudan contract can be obtained as described in the proposition below.

[12] Recall that $\psi^{-1}(n_y) = (x_n, v_l)$ for $n_y \in \mathcal{S}_Y^{(m,N)}$, where $n \leq N$ is the unique integer such that $n_y = (l-1)N + n$ for some $l \in \{1, 2, \ldots, m\}$.

**Proposition 3.6** *Let $F_0 > 0$, $V_0 \in \mathcal{S}_V$ and $\mathbf{G}^{(m,N)}$ be the generator defined in* (3.20)*. The risk-neutral value of a variable annuity with maturity $T > 0$ and guaranteed amount $G > 0$ can be approximated recursively by*

$$\begin{aligned} \mathbf{B}_M^{(m,N)} &= \mathbf{H}^{(1)}, \\ \mathbf{B}_z^{(m,N)} &= \max\{\mathbf{H}_z^{(2)}, e^{-r\Delta_M} \exp\{\Delta_M \mathbf{G}^{(m,N)}\}\mathbf{B}_{z+1}^{(m,N)}\} \quad 0 \le z \le M-1, \end{aligned} \tag{3.29}$$

*for $M \in \mathbb{N}$ sufficiently large and where the maximum is taken element by element (also known as the parallel maxima). $\mathbf{H}^{(1)}$ and $\mathbf{H}_z^{(2)}$, $z = 0, 1, \ldots, M-1$ are column vectors of size $mN \times 1$ whose $(l-1)N+n$-th entries, $h^{(1)}_{(l-1)N+n}$ and $h^{(2)}_{z,(l-1)N+n}$, are respectively given by*

$$\begin{aligned} h^{(1)}_{(l-1)N+n} &= \max(G, e^{x_n+\rho\gamma(v_l)}), \ \ and \\ h^{(2)}_{z,(l-1)N+n} &= g(t_z, v_l)e^{x_n+\rho\gamma(v_l)}, \end{aligned} \tag{3.30}$$

*$1 \le l \le m$ and $1 \le n \le N$.*

*Specifically, given $X_0^{(m,N)} = x_i = \ln(F_0) - \rho\gamma(V_0)$ and $V_0^{(m)} = V_0 = v_k$, the approximated value of the Bermudan contract is given by*

$$b_M^{(m,N)}(0, F_0, V_0) = \mathbf{e}_{ik}\mathbf{B}_0^{(m,N)}.$$

Hence, based on the last proposition, Algorithm 2 below provides a CTMC approximation for the value of a variable annuity contract (including early surrenders).

**Algorithm 2:** Variable Annuity with Early Surrenders via CTMC Approximation

**Input:** Initialize $\mathbf{G}^{(m,N)}$ as in (3.20), $\mathbf{H}^{(1)}$ and $\mathbf{H}_z^{(2)}$, for $z = 0, 1, \ldots, M-1$, as in (3.30)

$M \in \mathbb{N}$, the number of time steps,

$\Delta_M \leftarrow T/M$, the size of a time step

Set $\mathbf{B}_M^{(m,N)} \leftarrow \mathbf{H}^{(1)}$ and $\mathbf{A}_{\Delta_M} \leftarrow \exp\{\Delta_M \mathbf{G}^{(m,N)}\}e^{-r\Delta_M}$

**for** $z = M-1, M-2, \ldots, 0$ **do**

  $\mathbf{B}_z^{(m,N)} \leftarrow \max\left\{\mathbf{H}_z^{(2)}, \mathbf{A}_{\Delta_M}\mathbf{B}_{z+1}^{(m,N)}\right\}$

$b_M^{(m,N)}(0, F_0, V_0) \leftarrow \mathbf{e}_{ik}\mathbf{B}_0^{(m,N)}$

**return** $b_M^{(m,N)}(0, F_0, V_0)$

**Remark 3.4.4 (Convergence of VA prices with early surrenders)** *Recall, from Remark 3.3.2 that $F^{(m,N)} \Rightarrow F$ as $m, N \to \infty$. The convergence of the price of the Bermudan contract written on $F^{(m,N)}$ to the price of*

*the Bermudan contract written on $F$, that is $B_0^{(m,N)} \to B_0$ as $m, N \to \infty$, follows from Song* et al. *(2013), Theorem 9, and the results of Theorem 2.6 in Chapter 2, on the alternative continuous reward representation of the value function $v$*[13]*. Finally, the convergence of the price of the Bermudan contract to its American counterpart as $M$ goes to infinity follows from Proposition 3.5.*
*To our knowledge, detailed error and convergence analysis for the two-layers CTMC approximation of early-exercise options have not yet been performed in the literature. However, Cui* et al. *(2018) demonstrate the accuracy of the approximation numerically in the context of American put option pricing.*

#### 3.4.2.1 Fast Algorithm

Similar to the no surrender case, the efficiency of Algorithm 2 can be improved significantly by using an approximation based on the assumption that the variance process remains constant over small time periods (see Proposition 3.4).

Let $\boldsymbol{\varphi}(t) := [\varphi(t, e^{x_n + \rho\gamma(v_j)}, v_j)]_{j,n=1}^{m,N}$ be a $m \times N$ matrix representing the payoff at time $t$ for each state in $\mathcal{S}_V^{(m)} \times \mathcal{S}_X^{(N)}$, and $\varphi_{j,*}(t)$ be the $j$-th row of $\boldsymbol{\varphi}(t)$. We denote the matrix (vector) transpose operation by $\top$.
The Fast Algorithm to value VA with surrender rights is given in Algorithm 3. The Fast Algorithms to price VA contracts with and without early surrenders are very similar. The only difference is the additional line 11 in Algorithm 3. In fact, at a given time $t_z$ (that is, we fix one $z$ in the loop line 6 to 11), we can observe that at the end of the inner loop (line 9 and 10), the matrix $\mathbf{B}$ contains the continuation value of the Bermudan contract at $t_z$. Since Bermudan contracts can be surrendered at any time in $\mathcal{H}_M$, we simply need to calculate the maximum between the continuation value and the payoff at $t_z$ to obtain the value of the Bermudan contract at that time (line 11). Therefore, the only difference between the Fast Algorithms for VA pricing with and without early surrenders stems from the fact that the latter contract cannot be surrendered prior to maturity, and thus, only the continuation value needs to be calculated at each time step (that is, line 11 is not used to price VA contracts without early surrenders).

The computational effort in Algorithms 1 and 3 resides in the calculation of the matrix exponentials (line 3 to line 5). Hence, once they are (pre-)computed, one can price variable annuity contracts with and without

[13] Many convergence results, such as the one in Song *et al.* (2013), require the reward function to be bounded. However, as mentioned in Mijatović and Pistorius (2013), Remark 5.4 and Cui *et al.* (2018), Remark 5, the original payoff $\varphi$ can be replaced by the truncated payoff $\varphi \wedge L$ with a constant $L$ sufficiently large without altering the accuracy of the numerical results.

**Algorithm 3:** Variable Annuity with Early Surrenders via CTMC Approximation – Fast Algorithm

**Input:** Initialize $\mathbf{Q}^{(m)}$ as in (3.13) and $\mathbf{G}_j^{(N)}$ for $j = 1, 2, \ldots, m$, as in (3.18)

$M \in \mathbb{N}$, the number of time steps,

$\Delta_M \leftarrow T/M$, the size of a time step

Set $\boldsymbol{\varphi}(t_z) \leftarrow [\varphi(t_z, e^{x_n + \rho\gamma(v_j)}, v_j)]_{j,n=1}^{m,N}$ for $z = 0, 1, \ldots, M$

Set $\mathbf{B}_{j,*} \leftarrow \boldsymbol{\varphi}_{j,*}(t_M)$ for $j = 1, 2, \ldots, m$

/* Calculate the transition probability matrices */

**for** $j = 1, 2, \ldots, m$ **do**

  /* Adjusted transition probability matrix of $X^{(m,N)}$ given $V^{(m)} = v_j$ over a period of length $\Delta_M$ */

  $\mathbf{P}_j^X \leftarrow e^{\mathbf{G}_j^{(N)} \Delta_M} e^{-r\Delta_M}$

/* Transition probability matrix of $V^{(m)}$ over a period of length $\Delta_M$ */

$\mathbf{P}^V \leftarrow e^{\mathbf{Q}^{(m)} \Delta_M}$

/* VA valuation */

**for** $z = M - 1, \ldots, 0$ **do**

  **for** $j = 1, 2, \ldots, m$ **do**

    $\mathbf{E}_{*,j} \leftarrow \mathbf{P}_j^X \mathbf{B}_{j,*}^\top$

  **for** $n = 1, 2, \ldots, N$ **do**

    $\mathbf{B}_{*,n} \leftarrow \mathbf{P}^V \mathbf{E}_{n,*}^\top$

  $\mathbf{B} = \max(\mathbf{B}, \boldsymbol{\varphi}(t_z))$

**return** $b_{ki}$

surrender rights simultaneously at almost no additional cost. This also holds true for any other VA contracts with different guaranteed amounts, that is, a large variety of contracts with different guarantee structures and surrender rights can be priced simultaneously for almost the same computational effort as a single contract. Numerical experiments below demonstrate the efficiency and the accuracy of the Fast Algorithm.

The new algorithm can be easily adapted to a wide range of payoff functions. Thus, it extends the previous work of Cui *et al.* (2018), done in the context of option pricing, by significantly decreasing calculation time and allowing for efficient valuation of long-maturity derivatives.

### 3.4.3 Optimal Surrender Surface

In this section, we provide an algorithm to approximate the optimal surrender strategy for variable annuities with a general fee structure depending on the fund value and the variance process. Policyholder behavior may significantly impact pricing and hedging of variable annuities, Kling *et al.* (2014). Thus, analyzing optimal surrender behavior is crucial for insurers when developing risk management strategies for variable annuities, Bauer *et al.* (2017), Niittuinperä (2022). Optimal surrender strategies have been studied in the literature in different contexts, see for instance MacKay (2014), Bernard *et al.* (2014b), Bernard and MacKay (2015), Shen *et al.* (2016) and Kang and Ziveyi (2018).
The algorithms provided in this section are based on the CTMC Bermudan approximation and can thus be easily adapted to a wide range of payoff functions. In the context of American put option pricing, Ma *et al.* (2021) use the integral representation of the value function to derive a CTMC approximation for the optimal exercise surface. However, the method used by Ma *et al.* (2021) to derive such a representation requires the value function to be smooth enough, which can be difficult to prove for certain payoff functions and under general stochastic volatility models, see Lamberton and Terenzi (2019) for instance. The Bermudan approximation of the optimal surrender surface presented in this section does not require any specific regularity properties on the value function and thus extends the work of Ma *et al.* (2021) to more general payoffs and value functions.

The goal of this section is to approximate the (optimal) surrender surface using the CTMC approximation. To this end, we first introduce additional definitions and notations.

**Definition 3.6.1** *Let $E = [0,T] \times \mathbb{R}_+ \times \mathcal{S}_V$. The **continuation region** $\mathcal{C} \subset E$ is defined as*

$$\mathcal{C} = \{(t,x,y) \in E : v(t,x,y) > \varphi(t,x,y)\},$$

*and the **surrender region** $\mathcal{D} \subseteq E$, as*

$$\mathcal{D} = \{(t,x,y) \in E : v(t,x,y) = \varphi(t,x,y)\}.$$

**Remark 3.4.5** *It follows from Definition 3.6.1 that $E = \mathcal{C} \cup \mathcal{D}$, since $v(t,x,y) \geq \varphi(t,x,y)$ for all $(t,x,y) \in E$.*

If the value function $v$ is continuous, then $\mathcal{C}$ is an open set and $\mathcal{D}$ is closed; and the optimal surrender

surface is the boundary $\partial\mathcal{C}$ of $\mathcal{C}$.

Definition 3.6.1 provides a simple way of approximating the optimal surrender surface via CTMC approximation. To do so, denote by $f_{nl}$ the approximated fund process associated to $(x_n, v_l) \in \mathcal{S}_X^{(N)} \times \mathcal{S}_V^{(m)}$ such that $f_{nl} = e^{x_n+\rho\gamma(v_l)}$ and let $\mathcal{S}_F^{(m,N)} = \{f_{nl}\}_{n,l=1}^{N,m}$ be the state-space of $F^{(m,N)}$. We also denote by $\mathcal{S}_{F,l}^{(m,N)} = \{f_{nl}\}_{n=1}^{N}$ the $l$-section of $\mathcal{S}_F^{(m,N)}$. Moreover, let $\widetilde{\mathcal{H}}_M = \{t_0, t_1, \ldots, t_{m-1}\}$, that is $\widetilde{\mathcal{H}}_M = \mathcal{H}_M \setminus t_M$, and $\widetilde{E}_l = \widetilde{\mathcal{H}}_M \times \mathcal{S}_{F,l}^{(m,N)} \times \{v_l\}$. Using the previous definitions, the $l$-section, $l \in \{1, 2, \ldots, m\}$, of the continuation and the surrender regions can be approximated using the CTMC processes via

$$\mathcal{C}_l^{(m,N)} = \left\{(t_z, f_{nl}, v_l) \in \widetilde{E}_l \middle| b_M^{(m,N)}(t_z, f_{nl}, v_l) > g(t_z, v_l) f_{nl}\right\},$$

and

$$\mathcal{D}_l^{(m,N)} = \left\{(t_z, f_{nl}, v_l) \in \widetilde{E}_l \middle| b_M^{(m,N)}(t_z, f_{nl}, v_l) = g(t_z, v_l) f_{nl}\right\} \cup \{T\} \times \mathcal{S}_{F,l}^{(m,N)} \times \{v_l\},$$

respectively. Hence, the approximated continuation and surrender regions are given by

$$\mathcal{C}^{(m,N)} = \cup_{l=1}^{m} \mathcal{C}_l^{(m,N)}, \quad \text{and} \quad \mathcal{D}^{(m,N)} = \cup_{l=1}^{m} \mathcal{D}_l^{(m,N)}.$$

We use the notation of Proposition 3.6, with $b_{z,(l-1)N+n}$ denoting the $(l-1)N+n$-th entry of $\mathbf{B}_z^{(m,N)}$. For $(t_z, f_{nl}, v_l)$, $0 \le z \le M-1$, $1 \le n \le N$ and $1 \le l \le m$, we set $(t_z, f_{nl}, v_l) \in \mathcal{C}^{(m,N)}$ if $b_{z,(l-1)N+n} > g(t_z, v_l) f_{nl}$, and $(t_z, f_{nl}, v_l) \in \mathcal{D}^{(m,N)}$ otherwise. The approximated optimal surrender surface can then be obtained by analyzing the shape of $\mathcal{C}^{(m,N)}$.

We are now interested in studying the shape of the surrender region. To do so, we fix $t \in [0, T)$ and $y \in \mathcal{S}_V$ and consider the set of points $\mathcal{D}_{t,y} \subseteq \mathbb{R}_+$ for which it is optimal to surrender the contract. More precisely, we define $\mathcal{D}_{t,y}$ by

$$\mathcal{D}_{t,y} = \left\{f \in \mathbb{R}_+ \middle| (t, f, y) \in \mathcal{D}\right\}.$$

Suppose that, for all couples $(t, y) \in [0, T) \times \mathcal{S}_V$, the set $\mathcal{D}_{t,y}$ is of the form $[f^\star(t, y), \infty)$ for some $f^\star(t, y) \in \mathbb{R}_+$. That is, $f^\star(t, y)$ is the smallest fund value for which it is optimal to surrender the contract at time $t$ for a volatility level $y$, and for any fund value greater than $f^\star(t, y)$, it is also optimal to surrender. Mathematically, this may be expressed by

$$f^\star(t, y) := \inf\left\{f \in \mathbb{R}_+ \middle| f \in \mathcal{D}_{t,y}\right\} = \inf\{\mathcal{D}_{t,y}\}. \tag{3.31}$$

Under this assumption, the continuation and the surrender regions can be expressed as

$$\mathcal{C} = \left\{(t, f, y) \in E \middle| f < f^\star(t, y)\right\},$$

and

$$\mathcal{D} = \left\{(t, f, y) \in E \middle| f \geq f^*(t, y)\right\} \cup \{T\} \times \mathbb{R}_+ \times \mathcal{S}_V,$$

respectively.

Hence, under this assumption, the optimal surrender surface $f^*$ splits $E$ in two regions: at or above the surface is the surrender region, and below, the continuation region. That is, the set $\mathcal{D}_{t,y}$ is connected[14]. In this chapter, we say that the surrender region is of "threshold type" if for any $(t, y) \in [0, T) \times \mathcal{S}_V$, the set $\mathcal{D}_{t,y}$ is connected. There is a financial interpretation for such a form for the surrender region. As explained in Milevsky and Salisbury (2001), it is optimal for the policyholder to hold on to the contract when the fund value is low since there is a higher chance that the guarantee will be triggered at maturity.

**Remark 3.4.6** *The surrender region can take any shape; see for examples MacKay* et al. *(2017) Figure 4 and Figure 5. However, for specific fee and surrender charge structures, it can be shown that the surrender region is of threshold type, see for instance MacKay (2014), Appendix 2.A, when the index value process is modelled by a geometric Brownian motion. Other authors take this form for the surrender region as an initial assumption, see for example Kang and Ziveyi (2018). In the context of financial derivative pricing, Jacka (1991), Proposition 2.1, shows that the continuation region of American put options is of threshold type under the Black-Scholes setting whereas Touzi (1999), Section 2, proves it for some stochastic volatility models, and De Angelis and Stabile (2019), Proposition 4.1, in a very general setting.*

When the surrender region is of threshold type, a simple algorithm can be developed to approximate the optimal surrender surface. The idea is based on the definition of $f^\star(t, y)$ in (3.31) above: for each $t_z \in \mathcal{H}_M$ and $v_l \in \mathcal{S}_V^{(m,N)}$, we identify the smallest fund value $f^{(m,N)}(t_z, v_l)$ for which it is optimal to surrender. Algorithm 4 returns the approximated optimal surrender surface $f^{(m,N)}$ (under the assumption that the surrender region is of threshold type) and the approximated value of a variable annuity with early surrenders given $X_0^{(m,N)} = x_i = \ln(F_0) - \rho\gamma(V_0)$ and $V_0^{(m)} = V_0 = v_k$.

We note that the derivation of the optimal surrender surface is not mandatory to obtain the value of the Bermudan contract, as observed from Algorithm 2 or Algorithm 3 in the previous subsection.

Similarly as above, Algorithm 5 is the fast version of Algorithm 4. Recall that $\boldsymbol{\varphi}(t) := [\varphi(t, e^{x_n + \rho\gamma(v_j)})]_{j,n=1}^{m,N}$

[14] A set $X$ is connected if it cannot be divided into two disjoint non-empty open sets.

**Algorithm 4:** Optimal Surrender Surface (of threshold type) via CTMC Approximation

**Input:** Initialize $\mathbf{G}^{(m,N)}$ as in (3.20), $\mathbf{H}^{(1)}$ and $\mathbf{H}_z^{(2)}$, for $z = 0, 1, \ldots, M-1$, as in (3.30)

$M \in \mathbb{N}$, the number of time steps,

$\Delta_M \leftarrow T/M$, the size of a time step

Set $\mathbf{B}_M^{(m,N)} \leftarrow \mathbf{H}^{(1)}$ and $\mathbf{A}_{\Delta_M} \leftarrow \exp\{\Delta_M \mathbf{G}^{(m,N)}\} e^{-r\Delta_M}$

**for** $z = M-1, M-2, \ldots, 0$ **do**

  $\mathbf{B}_z^{(m,N)} \leftarrow \max\left\{\mathbf{H}_z^{(2)}, \mathbf{A}_{\Delta_M}\mathbf{B}_{z+1}^{(m,N)}\right\}$

  **for** $l = 1, 2 \ldots, m$ **do**

    $n \leftarrow 1$

    **while** $\left(b_{z,(l-1)N+n} > g(t_z, v_l)e^{x_n+\rho\gamma(v_l)}\right)$ *and* $(n < N)$ **do**

      $n \leftarrow n+1$

    $f^{(m,N)}(t_z, v_l) \leftarrow e^{x_n+\rho\gamma(v_l)}$

$b_M^{(m,N)}(0, F_0, V_0) \leftarrow \mathbf{e}_{ik}\mathbf{B}_0^{(m,N)}$

**return** $f^{(m,N)}$ and $b_M^{(m,N)}(0, F_0, V_0)$

is a $m \times N$ matrix representing the payoff at time $t$ for each state in $\mathcal{S}_V^{(m)} \times \mathcal{S}_X^{(N)}$, and $\boldsymbol{\varphi}_{j,*}(t)$ (resp $\mathbf{B}_{j,*}$) is the $j$-th row of $\boldsymbol{\varphi}(t)$ (resp. $\mathbf{B}$). We also denote by $b_{jn}$, the $(j,n)$ entry of the matrix $\mathbf{B}$.

**Remark 3.4.7** *Algorithms 2 and 3 do not require the specification of any particular form for the surrender region, which is not the case for many of the numerical procedures presented in the literature (Bernard* et al. *(2014b),Shen* et al. *(2016) and Kang and Ziveyi (2018)). Hence, their scope is more general.*

The accuracy of the approximated surrender boundary is demonstrated numerically in Appendix 3.8.

### 3.4.4 CTMC Approximation of the VIX

In Section 3.5, we analyze numerically the impact of various VIX-linked fee structures on the optimal surrender strategy. Since analytical formulas for the VIX are not always known for all models listed in Table 3.1, we use a CTMC approach to approximate the value of the volatility index. This is the case for the numerical

**Algorithm 5:** Optimal Surrender Surface (of threshold type) via CTMC Approximation – Fast Algorithm

**Input:** Initialize $\mathbf{Q}^{(m)}$ as in (3.13) and $\mathbf{G}_j^{(N)}$ for $j = 1, 2, \ldots, m$, as in (3.18)

$M \in \mathbb{N}$, the number of time steps,

$\Delta_M \leftarrow T/M$, the size of a time step

Set $\boldsymbol{\varphi}(t_z) \leftarrow [\varphi(t_z, e^{x_n + \rho\gamma(v_j)}, v_j)]_{j,n=1}^{m,N}$ for $z = 0, 1, \ldots, M$

Set $\mathbf{B}_{j,*} \leftarrow \boldsymbol{\varphi}_{j,*}(t_M)$ for $j = 1, 2, \ldots, m$

`/* Calculate the transition probability matrices */`

**for** $j = 1, 2, \ldots, m$ **do**

  `/* Ajusted transition probability matrix of` $X^{(m,N)}$ `given` $V^{(m)} = v_j$ `over a period of length` $\Delta_M$ `*/`

  $\mathbf{P}_j^X \leftarrow e^{\mathbf{G}_j^{(N)} \Delta_M} e^{-r\Delta_M}$

`/* Transition probability matrix of` $V^{(m)}$ `over a period of length` $\Delta_M$ `*/`

$\mathbf{P}^V \leftarrow e^{\mathbf{Q}^{(m)} \Delta_M}$

`/* VA valuation */`

**for** $z = M, M-1, \ldots, 0$ **do**

  **for** $j = 1, 2, \ldots, m$ **do**

    $\mathbf{E}_{*,j} \leftarrow \mathbf{P}_j^X \mathbf{B}_{j,*}^\top$

  **for** $n = 1, 2, \ldots, N$ **do**

    $\mathbf{B}_{*,n} \leftarrow \mathbf{P}^V \mathbf{E}_{n,*}^\top$

  $\mathbf{B} = \max(\mathbf{B}, \boldsymbol{\varphi}(t_z))$

  **for** $j = 1, 2 \ldots, m$ **do**

    $n \leftarrow 1$

    **while** $\left(b_{jn} > g(t_z, v_j) e^{x_n + \rho\gamma(v_j)}\right)$ *and* $(n < N)$ **do**

      $n \leftarrow n + 1$

    $f^{(m,N)}(t_z, v_j) \leftarrow e^{x_n + \rho\gamma(v_j)}$

**return** $f^{(m,N)}$ and $b_{ki}$

experiments performed under the 3/2 model whose results are available in Appendix 3.8[15]. In this section, we propose an approximation for the VIX when the variance process is approximated by a CTMC.

Let $\left\{\mathrm{VIX}_t^2\right\}_{t\geq 0}$ be the process representing the square of the VIX, defined by

$$\mathrm{VIX}_t^2 = \mathbb{E}_t\left[\frac{1}{\tau}\int_t^{t+\tau}\sigma_S^2(V_s)\,\mathrm{d}s\right],$$

with $\tau = 30/365$, see Cui *et al.* (2021b) Equation (6) for details.

Recall that $V^{(m)}$ is the CTMC approximation of $V$ taking values on a finite state-space $\mathcal{S}_V^{(m)} := \{v_1, v_2, \ldots, v_m\}$.

When the variance process is a CTMC, an integral expression can be obtained for the value of the volatility index. The CTMC approximation of the VIX, denoted by $\mathrm{VIX}^{(m)} = \{\mathrm{VIX}_t^{(m)}\}_{t\geq 0}$, is given in the proposition below.

**Proposition 3.7** *Given $V_t^{(m)} = v_k$, the square of the VIX index at time $t$ can be approximated by*

$$\begin{aligned}\left(\mathrm{VIX}_t^{(m),k}\right)^2 &:= \mathbb{E}\left[\frac{1}{\tau}\int_t^{t+\tau}\sigma_S^2(V_s^{(m)})\,\mathrm{d}s\Big|V_t^{(m)} = v_k\right]\\ &= \frac{1}{\tau}\int_0^{\tau}\mathbf{e}_k e^{\mathbf{Q}^{(m)}s}\mathbf{H}\,\mathrm{d}s,\end{aligned} \tag{3.32}$$

*where $\tau = 30/365$, $\mathbf{Q}^{(m)}$ is the generator of $V^{(m)}$ defined in* (3.13)*, $\mathbf{e}_k$ is the $k^{th}$ canonical basis vector of $\mathbb{R}^m$ and $\mathbf{H}$ is a $m\times 1$ vector whose $j$-th entry $h_j$ is given by $h_j = \sigma_S^2(v_j)$, $j = 1, 2, \ldots, m$.*

The proof is a direct consequence of Fubini's Theorem and the CTMC representation for conditional expectations.

**Remark 3.4.8** *Since $V^{(m)}$ is time-homogeneous, the approximation does not depend on $t$ and thus, for each $k\in\{1,\ldots,m\}$ it needs to be calculated only once for all $t\geq 0$.*

[15] For the 3/2 model, a closed-form expression for the VIX may be found in Carr and Sun (2007), Theorem 4. However, as pointed out by Drimus (2012), the integral that appears in the analytical formula is difficult to implement and is not suited for fast and accurate numerical methods. For this reason, the CTMC approximation of the VIX is used in the numerical examples under the 3/2 model.

The integral part in (3.32) can be approximated via a quadrature rule. In the numerical example section, this is done by dividing the interval $[0,\, \tau]$ into $n > 0$ equidistant sub-intervals for $z = 0, 1,\, 2, \ldots, n$ with

$$t_z := z\Delta_n, \quad \Delta_n := \tau/n.$$

The approximation then becomes

$$\left(\mathrm{VIX}_t^{(m),k}\right)^2 \approx \frac{\Delta_n}{\tau}\mathbf{e}_k \sum_{z=1}^{n} e^{\mathbf{Q}^{(m)}t_z}\mathbf{H}. \tag{3.33}$$

The approximation in (3.33) can be implemented in a straightforward manner. However, it requires calculating the exponential of a matrix $n$ times, which can be computationally inefficient. By making use of the tower property of conditional expectations, Algorithm 6 speeds up the calculation of (3.33).

We remark that other quadrature rules could be used in for the numerical calculation of (3.32). Using schemes that have a faster rate of convergence may reduce the number of matrix calculations needed.

**Algorithm 6:** Efficient Algorithm for the calculation of the VIX using CTMC approximation

**Input:** Initialize $\mathbf{Q}^{(m)}$ as in (3.13) and $\mathbf{H}$ as in Proposition 3.7

$n \in \mathbb{N}$, the number of time steps,

$\Delta_n \leftarrow \tau/n$, the size of a time step

Set $\mathbf{A}_{\Delta_n} \leftarrow \exp\{\Delta_n \mathbf{Q}^{(m)}\}$, $\mathbf{S} \leftarrow \mathbf{0}_{m\times 1}$ and $\mathbf{E} \leftarrow \mathbf{H}$

**for** $z = n, n-1, \ldots, 1$ **do**

  $\mathbf{E} \leftarrow \mathbf{A}_{\Delta_n}\mathbf{E}$

  $\mathbf{S} \leftarrow \mathbf{S} + \mathbf{E}$

$\mathrm{VIX}_t^{(m),k} \leftarrow \sqrt{\mathbf{e}_k \mathbf{S} \frac{\Delta_n}{\tau}}$

The vector $\mathbf{0}_{m\times 1}$ represents the null column vector of size $m$. Note that Algorithm 6 requires the calculation of a matrix exponential only once at the beginning of the procedure, which makes the algorithm very efficient.

Numerical experiments in the next section also demonstrate the accuracy and the efficiency of Algorithm 6 empirically. This new VIX CTMC approximation can also be used for approximating the price of VIX derivatives. This is left as future research.

## 3.5 Numerical Analysis

In the constant fee case, the misalignment between the fees and the value of the financial guarantee creates an incentive for the policyholder to surrender her policy prematurely (see Milevsky and Salisbury (2001) for details). Indeed, when the fund value is high, the amount of the fees paid is also high, but the put option embedded in a GMMB is out-of-the-money and worth very little since the probability for the guarantee to be triggered at maturity is low. Thus, the policyholder pays high fees for a financial guarantee that has a low value. This is clearly an incentive for a policyholder to surrender her policy prior to maturity. MacKay *et al*. (2017) considered state-dependent fee structures where the fee is paid when the fund value is under a certain level, and showed that this particular type of fee structure reduces insurers' exposure to policyholder behavior under risk-neutral value maximization assumption. Fees that are tied to the S&P volatility index, the VIX, are also studied in the literature, Cui *et al*. (2017a) and Kouritzin and MacKay (2018). Since the volatility is negatively correlated with the stock price (see for instance Rebonato (2004)), we expect VIX-linked fees to be low when the fund value is high and to be higher when the fund value is low. Cui *et al*. (2017a) and Kouritzin and MacKay (2018) showed numerically that linking the fee to the volatility index VIX may help realign revenues with variable annuity liabilities (for VA without surrender rights). Hence, it is reasonable to believe that linking the fees to the VIX may help to reduce insurers' exposure to surrender risk. This will be explored in greater detail in the numerical experiments conducted below.

This section first discusses the market, the VA, and the CTMC parameters used in all numerical experiments performed below. Then, we investigate the efficiency of the Fast Algorithms (Algorithms 1 and 3). In the third subsection, we discuss different structures for the VIX-linked fee, that is, different ways to link the fee rate to the VIX index. Finally, we analyze the impact of different VIX-linked fee structures on the value of VAs and their optimal surrender strategy. We restrict our analysis to the classical Heston model. Using our framework, any model listed in Table 3.1 could have been used.

As supplementary material, we investigate the numerical accuracy of the approximated optimal surrender surface derived in Algorithm 4 (or equivalently 5), and the CTMC approximation of the VIX (Algorithm 6). We also analyze numerically the impact of time-dependent risk-free rates on the variable annuity values and the optimal surrender surface. Finally, we explore numerically the impact of VIX-linked fee structures under the $3/2$ model.

### 3.5.1 Market, VA and CTMC Parameters

We consider a market under regular conditions: low initial variance $v_0$, low long-term variance $\theta$, moderate volatility of volatility $\sigma$, and moderate speed reversion $\kappa$[16]. The initial value of the variance is set to $V_0 = 0.03$, the correlation to $\rho = -0.75$ and the risk-free rate to $r = 0.03$. The selected market parameters are summarized in Table 3.2 and the model dynamic is given in Table 3.1.

Table 3.2: Market parameters

| Parameter | $V_0$ | $\kappa$ | $\theta$ | $\sigma$ | $\rho$ | $r$ |
|---|---|---|---|---|---|---|
| Value | $0.03$ | $2.00$ | $0.04$ | $0.20$ | $-0.75$ | $0.03$ |

The variable annuity parameters are set to $F_0 = S_0 = 100$, $T = 10$ (years), and $G = 100$. We assume that the payoff when the contract is surrendered early is given by $g(t, V_t)F_t$, with

$$g(t, y) = e^{-k(T-t)}, \quad y \in \mathcal{S}_V,$$

and $k = 0.2\%$. The choices for the fee function $c(x, y)$ are discussed in greater detail in the next section.

Note that a numerical analysis under the Heston model with $G = F_0 e^{\widetilde{g}T}$, $\widetilde{g} = 2\%$ (rather than $G = 100$) is also performed in Appendix 3.8. We also investigate the impact of time-dependent risk-free rates on the value of variable annuity contracts and the optimal surrender surface in Section, 3.8.5.

For all numerical examples in this chapter, we use the non-uniform grid proposed by Tavella and Randall (Tavella and Randall (2000), Chapter 5.3). For example, suppose that $\widetilde{X}$ is a one-dimensional diffusion process approximated by a continuous-time Markov chain $\widetilde{X}^{(n)}$ taking values on a finite state-space $\mathcal{S}_{\widetilde{X}} = \{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n\}$, $n \in \mathbb{N}$. The state-space of the approximated process can be determined as follows:

$$\widetilde{x}_i = \widetilde{X}_0 + \widetilde{\alpha} \sinh\left(c_2 \frac{i}{n} + c_1 \left[1 - \frac{i}{n}\right]\right), \quad i = 2, \ldots, n-1, \tag{3.34}$$

[16] Bloomberg provides historical Heston calibrated parameters to market data on a daily basis via its Option Pricing template (OVME). These parameters are often used in practice for over-the-counter option pricing. Bloomberg's Heston calibrated speed reversion parameter is $\kappa = 3.6881$ as of December 31, 2019, $\kappa = 5$ as of March 31, 2020 and $\kappa = 1.1397$ as of September 30, 2020. The parameter selected for our numerical experiments falls approximately in the middle of those of December 2019 and September 2020. In the financial literature, Aït-Sahalia and Kimmel (2007) obtain $\kappa = 5.07$ whereas Garcia *et al.* (2011) obtain $\kappa = 0.173$, and again our values fall between these two values.

where

$$c_1 = \sinh^{-1}\left(\frac{\widetilde{x}_1 - \widetilde{X}_0}{\widetilde{\alpha}}\right), \quad \text{and } c_2 = \sinh^{-1}\left(\frac{\widetilde{x}_n - \widetilde{X}_0}{\widetilde{\alpha}}\right).$$

Here the constant, $\widetilde{\alpha} \geq 0$, controls the degree of non-uniformity of the grid. The choices for the two boundary states and the non-uniformity parameter $\widetilde{\alpha}_v$ (resp. $\widetilde{\alpha}_X$) for $V^{(m)}$ (resp. $X^{(m,N)}$) are discussed in more details below.

**Remark 3.5.1** *When the initial values of the auxiliary and variance processes are not in the grid, they can be inserted (see for example Cui* et al. *(2019), Section 2.3 for details), or the value of the variable annuity must be interpolated between the appropriate grid points.*

Non-uniform schemes have been used frequently in the literature for options pricing via CTMC approximation methods, see Mijatović and Pistorius (2013), Lo and Skindilias (2014), Cai *et al.* (2015), Kirkby *et al.* (2017), Cui *et al.* (2018), Cui *et al.* (2019), Leitao Rodriguez *et al.* (2021) and Ma *et al.* (2021), among others. For a deep analysis of grid designs and how they can affect convergence, the reader is referred to Zhang and Li (2019).

Unless stated otherwise, all numerical experiments are performed using the CTMC parameters listed in Table 3.3. Recall that $m$ is the number of grid points for the variance process whereas $N$ represents the number of grid points of the fund process. $\widetilde{\alpha}_v$ (resp. $\widetilde{\alpha}_X$) is the grid non-uniformity parameter of the variance (resp. the auxiliary) process. The grid's upper and lower bounds are respectively $v_1$ and $v_m$ for the variance process and $x_1$ and $x_N$ for the auxiliary process with $X_0 = \ln(F_0) - \rho\gamma(V_0)$. The values of $V_0$ and $X_0$ are inserted in their respective grid as per Remark 3.5.1. Finally, we use $M = 500 \times 10$ times steps, which corresponds to a computing frequency of approximately twice daily (since there are approximately $250$ trading days per year).

Table 3.3: CTMC parameters

| Parameter | $m$ | $N$ | $v_1$ | $v_m$ | $\widetilde{\alpha}_v$ | $x_1$ | $x_N$ | $\widetilde{\alpha}_X$ | $M$ |
|---|---|---|---|---|---|---|---|---|---|
| Value | $50$ | $2,000$ | $V_0/100$ | $7v_0$ | $0.6571$ | $X_0/10^6$ | $1.95X_0$ | $2/100$ | $5,000$ |

Note that, under the Heston model, good approximations of the transition density of the variance process can be obtained with a small number of grid points, see for example Cui *et al.* (2019) Figure 3.

All the numerical experiments are carried out with Matlab R2015a on a Core i7 desktop with 16GB RAM and speed 2.40 GHz.

### 3.5.2 Efficiency of the Fast Algorithms

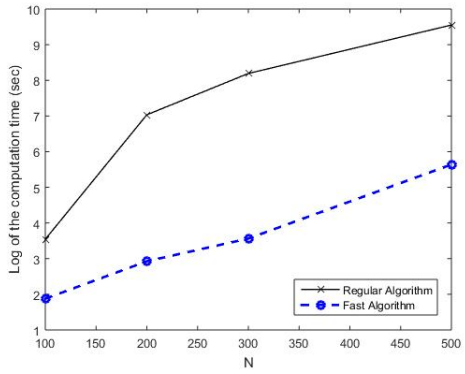
The valuation of options (or variable annuities) using CTMC requires the calculation of a matrix exponential to obtain the transition probability matrix. In Algorithm 2, the two-dimensional process is mapped onto a one-dimensional process resulting in a generator of size $mN \times mN$. Thus, we need to calculate the exponential of a $mN \times mN$ matrix to obtain the transition probability matrix. For large values of $mN$, this procedure might stretch computing resources to unacceptable levels. Algorithms 1 and 3, proposed in Section 3.4, require $m$ times the calculation of the exponential of a $N \times N$ matrix and one time the exponential of a $m \times m$ matrix. As demonstrated below, reducing the size of the matrix in the exponent allows to significantly increase the efficiency of the procedure.

When the size of the exponent in the matrix exponential is greater than $200 \times 200$, we observe that the function *fastExpm* for Matlab, see Mentink-Vigier (2020)[17], which is designed for the fast calculation of matrix exponentials of large sparse matrices, can further speed up the calculation. Combining the function fastExpm and the Fast Algorithms can speed up the code by up to $100$ times (for European and Bermudan contracts). For the European contract, the Fast Algorithm can reduce the computation time by up to $12$ times whereas the function *fastExpm*, by up to $7$ times. For the Bermudan contract, the computation time is reduced by up to $4$ times with the Fast Algorithm and by up to $40$ times with the function *fastExpm*. When $m = 50$ and $N = 100$ the running time is approximately 6 seconds for both the European and the Bermudan contracts, confirming the high efficiency of the new algorithms.

Figure 3.1 illustrates the computation time in seconds of the "Fast Algorithm", Algorithm 1 for the European contract and Algorithm 3 for the Bermudan contract, combined with the function *fastExpm* of Mentink-Vigier (2020) (when the size of the generator in the matrix exponential is greater than $200 \times 200$), and the computation time of the "Regular Algorithm" for the European contract (3.22) and Algorithm 2 for the Bermudan one. The run times in Figure 3.1 are recorded using the market, VA, and CTMC parameters of Subsection 3.5.1, except for the number of grid points for the auxiliary process $X$ which is set to $N = 100, 200, 300$ and $500$, respectively. Moreover, we use a constant fee structure, that is we fix

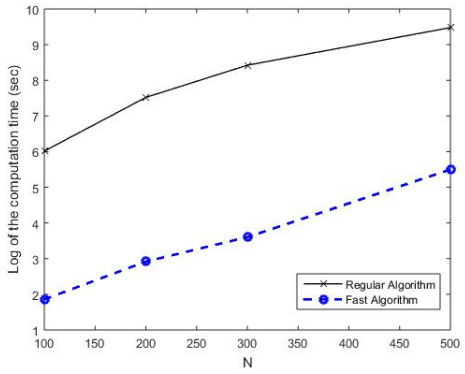
[17] The function *fastExpm* is based on Hogben *et al.* (2011) and Kuprov (2011).

$c(x, y) = 1.5338\%$ for all $(x, y) \in \mathbb{R}_+ \times \mathcal{S}_V$.




(a) European contract

Log of computation time (in seconds)




(b) Bermudan contract

Log of computation time (in seconds)

Figure 3.1: Fast Algorithms: Algorithm 1 for the European contract and Algorithm 3 for the Bermudan contract. Regular Algorithms: Equation (3.22) for the European contract and Algorithm 2 for the Bermudan contract.

We also compare the accuracy of the new algorithms. For comparison, we use the regular algorithms for the European contract and for the Bermudan contract. The absolute difference in the VA prices between the regular and the fast algorithms is around $10^{-3}$ for both European and Bermudan contracts, whereas the relative difference is around $10^{-5}$, confirming the accuracy of the fast algorithms. See also Appendix (3.8) for more details.

**Remark 3.5.2** *When valuing a European contract, the use of the Expokit of Sidje (1998) based on Krylov subspace projection methods accelerates the computation time considerably. Moreover, software packages in Matlab and Fortran can be downloaded for free at* `https://www.maths.uq.edu.au/expokit`. *When* $N = 500$, *we observe that the function expv in the Expokit can accelerate the running time by up to* $7$ *times (compared to the fast algorithm).*

*However, since the function takes advantage of the product of a matrix exponential with a vector, the fast algorithm is more efficient when valuing Bermudan contracts. Indeed, when using the fast algorithm, matrix exponentials are calculated only once at the beginning of the procedure; whereas when using the Expokit, it needs to be calculated at each time step (since we need to calculate the matrix exponential multiplied by a vector in order to make use of the procedure) which slows down the execution considerably. Hence, when valuing a Bermudan contract, Algorithm 3 is up to* $9$ *times faster than the regular algorithm using Expokit*

*procedures of Sidje (1998).*

As mentioned previously, the computational effort in Algorithms 1 and 3 resides in the calculation of the matrix exponentials at the beginning of the two procedures. Hence, once they are (pre-)computed, one can obtain the value of variable annuities with and without surrender rights simultaneously at almost no additional cost. For instance, when $N = 500$, we simultaneously obtain the prices of variable annuities with and without surrender rights in $270$ seconds; whereas the values of the Bermudan and the European contracts can be obtained separately in approximately $250$ seconds for each. This also holds true for any other VA contracts with different guaranteed amounts; that is, a variety of contracts can be priced for almost the same computational effort as a single contract.

### 3.5.3 Fee Structures and Fair Fee Parameters

First, recall from Subsection 3.4.4 that

$$\mathrm{VIX}_t^2 = \mathbb{E}_t\left[\frac{1}{\tau}\int_t^{t+\tau} \sigma_S^2(V_s)\,\mathrm{d}s\right], \qquad \text{with } \tau = 30/365.$$

For all the numerical experiments conducted below, we consider three types of VIX-linked fee structures. As in Cui *et al*. (2017a), we use an uncapped $\mathrm{VIX}^2$-linked fee structure (the "Uncapped $\mathrm{VIX}^2$" fee structure) of the form

$$c_t = \widetilde{c} + \widetilde{m}\mathrm{VIX}_t^2,\ 0 \le t \le T,$$

with $\widetilde{c}, \widetilde{m} \ge 0$.

When the volatility is high (e.g. financial turmoil), the Uncapped $\mathrm{VIX}^2$ fee rate can become excessive (see Table 3.5 for examples). This motivates our choice of imposing a cap to this type of fee. Thus, for the second fee structure, we consider capped $\mathrm{VIX}^2$-linked fees (or simply the "Capped $\mathrm{VIX}^2$" fee structure or Capped fee structure) of the form

$$c_t = \min(\widetilde{c} + \widetilde{m}\mathrm{VIX}_t^2, K),\ 0 \le t \le T,$$

where $\widetilde{c}, \widetilde{m} \ge 0$ and $K > 0$.

Finally, we look at an uncapped fee structure linked to the VIX (rather than the VIX squared) given by

$$c_t = \widetilde{c} + \widetilde{m}\mathrm{VIX}_t,\ 0 \le t \le T,$$

with $\widetilde{c}, \widetilde{m} \geq 0$, also called the "Uncaped VIX" fee structure. Such a structure can help to keep the fee rates reasonable during high volatility periods.

Since the volatility process is Markovian, the volatility index at $t$ will depend only on $V_t$. Therefore, the three fee functions are of the form $c(y)$, $y \in \mathcal{S}_V$.

The fee structure parameters $\widetilde{c}$ and $\widetilde{m}$ are set in such a way that the contract is fair at inception. That is, the initial amount invested by the policyholder, $F_0$, is equal to the expected discounted value of the future benefit (without early surrenders). Those parameters are called the fair parameters and are henceforth denoted by a star: $(\widetilde{c}^*, \widetilde{m}^*)$. We set the fee in this manner to calculate the value added by the right to surrender. To identify the fair fee vector $(\widetilde{c}^*, \widetilde{m}^*)$, we first fix a multiplier $\widetilde{m}^*$, and then solve for the corresponding fair base fee $\widetilde{c}^*$. Note that such a fair base fee $\widetilde{c}^*$ does not exist for all values of $\widetilde{m}^* > 0$. When $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}^2$, the fair parameters are obtained using the exact formula of Cui *et al.* (2017a). For the other fee structures, the fair parameters are calibrated using a CTMC approximation with $N = 100$ (all other CTMC parameters are the same as in Table 3.3), to reduce the computation time. Note that very accurate VA prices are obtained extremely fast with $N = 100$ under the Heston model (see Appendix 3.8 for numerical details). Table 3.4 presents the fair fee vectors $(\widetilde{c}^*, \widetilde{m}^*)$, and Table 3.5 shows examples of fair fee rates produced by each fair fee vector at different volatility levels ($\sqrt{V_t}$).

Table 3.4: Fair fee vectors $(\widetilde{c}^*, \widetilde{m}^*)$

| | | | | |
|---|---|---|---|---|
| | $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$ | | | |
| $\widetilde{m}^*$ | $0.0000$ | $0.1500$ | $0.3000$ | $0.4345$ |
| $\widetilde{c}^*$ | $1.5338\%$ | $1.0036\%$ | $0.4741\%$ | $0.000\%$ |
| | $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 2\%$ | | | |
| $\widetilde{m}^* =$ | $0.0000$ | $0.1500$ | $0.3000$ | $0.4927$ |
| $\widetilde{c}^* =$ | $1.5338\%$ | $1.0112\%$ | $0.5415\%$ | $0.000\%$ |
| | $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$ | | | |
| $\widetilde{m}^* =$ | $0.0000$ | $0.0250$ | $0.0500$ | $0.0836$ |
| $\widetilde{c}^* =$ | $1.5338\%$ | $1.0750\%$ | $0.6164\%$ | $0.000\%$ |

The uncapped $\mathrm{VIX}^2$-linked fee rate can get very high as the volatility increases, reaching levels as high as

Table 3.5: Fair fee rates in %

(a) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$ (%)

| $\sqrt{\mathbf{V_t}}$ (%)$\backslash\widetilde{\mathbf{m}}^*$ | **0.0000** | **0.1500** | **0.3000** | **0.4345** |
|---|---|---|---|---|
| **8.702** | 1.5338 | 1.1550 | 0.7770 | 0.4387 |
| **13.882** | 1.5338 | 1.3168 | 1.1007 | 0.9075 |
| **30.434** | 1.5338 | 2.3314 | 3.1299 | 3.8464 |
| **42.772** | 1.5338 | 3.5808 | 5.6285 | 7.4653 |

(b) $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2(\%), K\}$, $K = 2\%$

| $\sqrt{\mathbf{V_t}}$ (%)$\backslash\widetilde{\mathbf{m}}^*$ | **0.0000** | **0.1500** | **0.3000** | **0.4927** |
|---|---|---|---|---|
| **8.702** | 1.5338 | 1.1627 | 0.8444 | 0.4975 |
| **13.882** | 1.5338 | 1.3245 | 1.1681 | 1.0291 |
| **30.434** | 1.5338 | 2.0000 | 2.0000 | 2.0000 |
| **42.772** | 1.5338 | 2.0000 | 2.0000 | 2.0000 |

(c) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

| $\sqrt{\mathbf{y}}$ (%)$\backslash\widetilde{\mathbf{m}}^*$ | **0.0000** | **0.0250** | **0.0500** | **0.0836** |
|---|---|---|---|---|
| **8.702** | 1.5338 | 1.3262 | 1.1188 | 0.8402 |
| **13.882** | 1.5338 | 1.4363 | 1.3390 | 1.2083 |
| **30.434** | 1.5338 | 1.8188 | 2.1041 | 2.4877 |
| **42.772** | 1.5338 | 2.1113 | 2.6889 | 3.4657 |

$7.4653\%$ when the volatility is $42.772\%$. During the last COVID-19 financial crisis, volatility reached levels as high as $80\%$ in March 2020[18]. This motivates the two other fee structures we propose. We note that the introduction of a cap does not significantly affect the calibrated fair fee parameters. We also observe, from Table 3.5 that the second and the third fee structures allow to keep the fees at reasonable levels during high volatility periods compared to the Uncapped $\mathrm{VIX}^2$ fee structure.

### 3.5.4 Effect of VIX-Linked Fees on Surrender Incentives

Recall that under the Heston (1993) model, the price of the risky asset satisfies

$$\begin{aligned} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \sqrt{V_t}S_t\,\mathrm{d}W_t^{(1)} \\ \mathrm{d}V_t &= \kappa(\theta - V_t)\,\mathrm{d}t + \sigma\sqrt{V_t}\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{3.35}$$

with $S_0$ and $V_0$ are deterministic, and where $W = (W^{(1)}, W^{(2)})^T$ is a bi-dimensional correlated Brownian motion under $\mathbb{Q}$ and such that $[W^{(1)}, W^{(2)}]_t = \rho t$ with $\rho \in [-1, 1]$, the speed of the mean-reversion $\kappa > 0$, the long term variance $\theta > 0$ and the volatility of the variance $\sigma > 0$ (also called the volatility of the volatility).

Moreover, when the market is modeled by (3.35), the VIX has a closed-form expression given by

$$\mathrm{VIX}_t^2 = B + AV_t \tag{3.36}$$

with $A = \frac{1-e^{-\kappa\tau}}{\kappa\tau}$ and $B = \frac{\theta(\kappa\tau-1+e^{-\kappa\tau})}{\kappa\tau}$, see Zhu and Zhang (2007) for details. The three fee structures exposed in the Subsection 3.5.3 can thus be obtained explicitly in terms of the current volatility using (3.36) as follows:

Table 3.6: Fair fee process under the Heston model

| **Fee Structure** | $c_t, 0 \le t \le T$ |
|---|---|
| **Uncapped** $\mathrm{VIX}^2$ | $\widetilde{c}^* + \widetilde{m}^*(A + BV_t)$ |
| **Capped** $\mathrm{VIX}^2$ | $\min\{K, \widetilde{c}^* + \widetilde{m}^*(A + BV_t)\}$ |
| **Uncapped VIX** | $\widetilde{c}^* + \widetilde{m}^*\sqrt{(A + BV_t)}$ |

Now from Lemma 3.1, we find that $\gamma(x) = x/\sigma$. Thus, given a certain fee process $c_t$ (listed in Table 3.6), the

[18] See VIX historical data at `https://www.cboe.com/tradable_products/vix/vix_historical_data/`.

dynamics of the auxiliary process can be derived as

$$\begin{aligned} \mathrm{d}X_t &= \mu_X(X_t, V_t)\,\mathrm{d}t + \sigma_X(V_t)\,\mathrm{d}W_t^*, \\ \mathrm{d}V_t &= \mu_V(V_t)\,\mathrm{d}t + \sigma_V(V_t)\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{3.37}$$

where $\mu_X(X_t, V_t) = r - \frac{\rho\kappa\theta}{\sigma} - c_t + V_t\left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)$, and $\sigma_X(V_t) = \sqrt{(1-\rho^2)V_t}$, $0 \leq t \leq T$.

Using the CTMC technique outlined in Section 3.3 and the market, VA, and CTMC parameters of subsection 3.5.1, we perform the valuation of a variable annuity with and without early surrenders ( "VA with early surrender (ES)" and " VA without ES", respectively). The results are reported in Table 3.7 below[19].

First, we observe that the fair value of the variable annuity without early surrenders is approximately $F_0 = 100$ for all fair fee vectors. This is because fair fee parameters are calibrated such that the value of the VA without surrender rights at time $t = 0$ is equal to the initial premium ($F_0 = 100$). Moreover, under the Uncapped $\mathrm{VIX}^2$-linked structure, fair fee parameters are obtained using the exact pricing formula of Cui *et al.* (2017a), confirming the accuracy of the approximated model. The absolute error is around $10^{-4}$ for all fee vectors for this fee structure. An accuracy of around $10^{-3}$ can be obtained for each value (the value of the VA with and without surrender rights) with fewer grid points with significantly less computational effort. Detailed results with $N = 100$ and $N = 1,000$ are given in Appendix 3.8 with their respective computation time. The value of the surrender right is calculated as the difference between the values of the variable annuity with and without early surrenders.

As $\widetilde{m}^*$ increases, the risk-neutral value of early surrenders remains very close for all fee structures. One might expect the VIX-linked fee structure to reduce the risk-neutral value of the surrender rights since this type of fee structure realigns income and liability (Cui *et al.*, 2017a), but this is not what is observed here. However, VIX-linked fee structures have an impact on optimal surrender strategies, as shown below.

In Figure 3.2, the shape of the approximated optimal surrender surface associated with each of the VIX-linked fee structures is illustrated for different values of the fair multiplier. We observe that the surrender region is of threshold type for all $\widetilde{m}^*$. We also note that, regardless of the value of $\widetilde{m}^*$, the boundary is a concave function of time. It slowly increases to a maximum and decreases rapidly to the guarantee level $G$. This is consistent with earlier findings of Bernard *et al.* (2014b) and Kang and Ziveyi (2018).

[19] Numerical experiments under the Heston model have been performed using Equation (3.22), and Algorithms 2 and 4. Note however that similar results are obtained when using the Fast Algorithms, see Appendix 3.8 for details.

Table 3.7: Variable annuity with and without early surrender (ES).

(a) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$

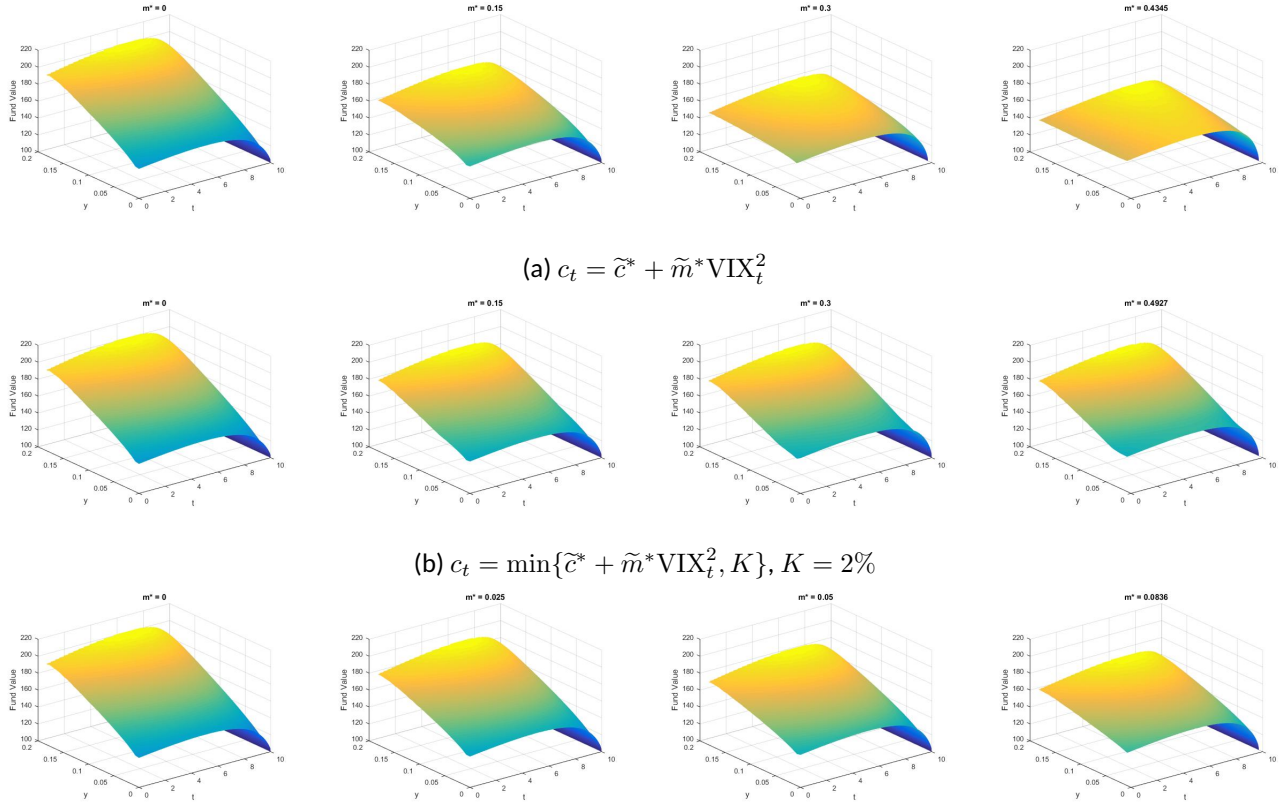
| $\widetilde{\mathbf{m}}^* =$ | $\mathbf{0.0000}$ | $\mathbf{0.1500}$ | $\mathbf{0.3000}$ | $\mathbf{0.4345}$ |
|---|---|---|---|---|
| $\widetilde{\mathbf{c}}^* =$ | $\mathbf{1.5338}\%$ | $\mathbf{1.0036}\%$ | $\mathbf{0.4741}\%$ | $\mathbf{0.000}\%$ |
| **VA without ES** | 100.00090 | 100.00091 | 100.00092 | 100.00093 |
| **VA with ES** | 103.01785 | 103.00823 | 103.00330 | 103.00367 |
| **Value of ES** | 3.01695 | 3.00732 | 3.00238 | 3.00274 |

(b) $c_t = \min\{K, \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2\}$, with $K = 2\%$

| $\widetilde{\mathbf{m}}^* =$ | $\mathbf{0.0000}$ | $\mathbf{0.1500}$ | $\mathbf{0.3000}$ | $\mathbf{0.4927}$ |
|---|---|---|---|---|
| $\mathbf{c}^* =$ | $\mathbf{1.5338}\%$ | $\mathbf{1.0112}\%$ | $\mathbf{0.5415}\%$ | $\mathbf{0.000}\%$ |
| **VA without ES** | 100.00090 | 100.00075 | 100.00070 | 100.00036 |
| **VA with ES** | 103.01785 | 103.00596 | 102.99137 | 102.97560 |
| **Value of ES** | 3.01695 | 3.00521 | 2.99067 | 2.97524 |

(c) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

| $\widetilde{\mathbf{m}}^* =$ | $\mathbf{0.0000}$ | $\mathbf{0.0250}$ | $\mathbf{0.0500}$ | $\mathbf{0.0.0836}$ |
|---|---|---|---|---|
| $\widetilde{\mathbf{c}}^* =$ | $\mathbf{1.5338}\%$ | $\mathbf{1.0750}\%$ | $\mathbf{0.6164}\%$ | $\mathbf{0.000}\%$ |
| **VA without ES** | 100.00090 | 100.00099 | 100.00057 | 100.000167 |
| **VA with ES** | 103.01785 | 103.01080 | 103.00446 | 102.99853 |
| **Value of ES** | 3.01695 | 3.00981 | 3.00389 | 2.99686 |




(a) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$




(b) $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}$, $K = 2\%$

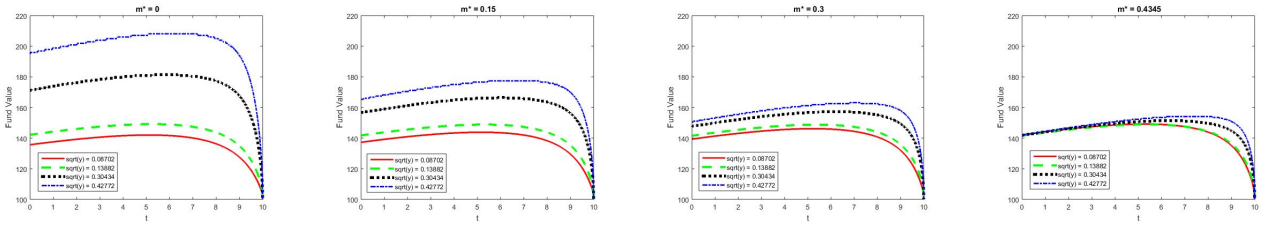



(c) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

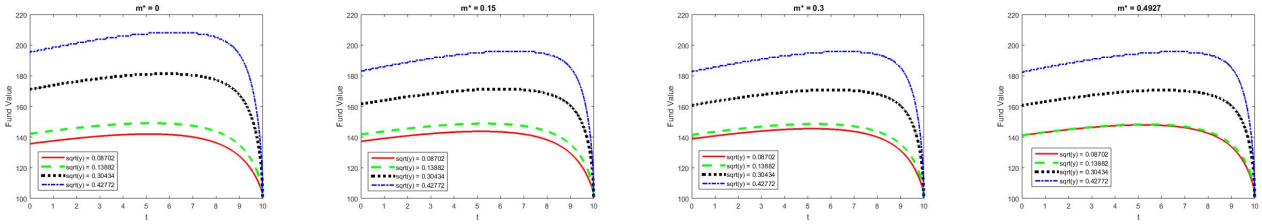
Figure 3.2: Approximated optimal surrender surface of VIX-linked fees VAs for different values of fair multiplier $\widetilde{m}^*$ under the Heston model. The $x$-axis represents the time and the $y$-axis the variance.

For fixed $t \in [0, T]$, we observe, in Figures 3.2 and 3.3, that the $t$ section of $f^{(m,N)}$, the function $y \mapsto f^{(m,N)}(t, y)$, denoted by $f_t^{(m,N)}$, is increasing for all $\widetilde{m}^*$. Hence, when the volatility is high, variable annuities are surrendered at higher fund values than when the volatility is low. This is in line with the findings of Kang and Ziveyi (2018). However, as $\widetilde{m}^*$ increases, the function $f_t^{(m,N)}$ increases at a slower rate (and particularly for the uncapped structures, panel (a) and (c) of Figures 3.2 and 3.3). For instance, fix $\widetilde{m}^* \in \{0, 0.15, 0.3, 0.4345\}$, $y \in \mathcal{S}_V^{(m)}$ and note, from Figure 3.3 (panel (a) or (c)), that the $y$ section of $f^{(m,N)}$, the function $t \mapsto f^{(m,N)}(t, y)$, denoted by $f_y^{(m,N)}$, is pushed upwards as $y$ increases. However, the difference between the low and the high volatility $y$ section is less significant as $\widetilde{m}^*$ grows. This means that the optimal surrender decision for uncapped VIX-linked fees is less sensitive to volatility fluctuations when fees are tied to the volatility index. In the case of the capped structure (panel (b) of Figures 3.2 and 3.3), we also note that approximated optimal surrender surfaces are gradually increasing as volatility grows; however, they now increase at a similar pace for all $\widetilde{m}^*$. This can be interpreted from a financial perspective

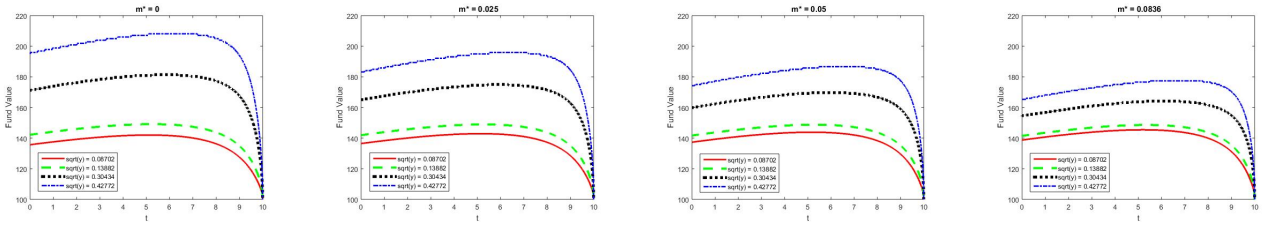
as pointed out below.




(a) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$




(b) $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}$, $K = 2\%$

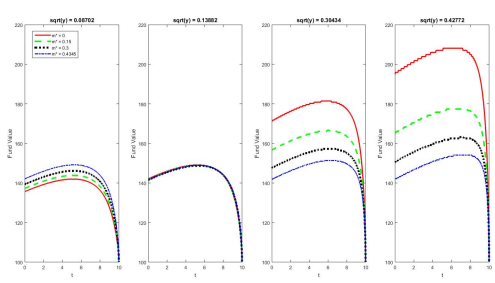



(c) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

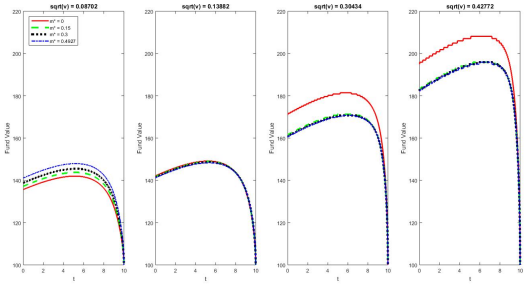
Figure 3.3: The $y$ section of the approximated optimal surrender surface, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.

In Figure 3.4, we fix a volatility level $y$ and we compare the $y$ sections of $f^{(m,N)}$, $f_y^{(m,N)}$, for different fair multipliers $\widetilde{m}^*$. When the volatility is low ($\sqrt{y} = 8.702\%$), see for instance the first graph of Figure 3.4 (a), $f_y^{(m,N)}$ is pushed upwards as $\widetilde{m}^*$ increases. This means that a variable annuity contract with a fully dependent uncapped $\mathrm{VIX}^2$-linked fee structure ($\widetilde{m}^* = 0.4345$) is surrendered at higher fund values than the constant fee one ($\widetilde{m}^* = 0$). Indeed, when the volatility is low, the fee paid under a variable annuity contract with a $\mathrm{VIX}^2$-linked fee structure is also low (see Table 3.5 for examples), making VIX-linked fee contracts more attractive than the constant fee ones. However, as the volatility rises, $\mathrm{VIX}^2$-linked fees also rise and so, the relation between the optimal surrender decision and $\widetilde{m}^*$ reverts. The second graph of Figure 3.4 (a) shows that variable annuity contracts are surrendered at almost all the same fund value levels when the volatility equals to $13.882\%$, regardless of $m^*$. The latter may be explained by the fact that

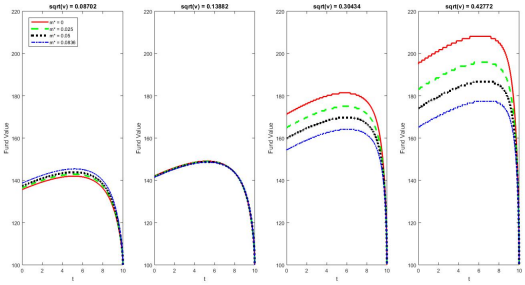
fee rates are all around the same level when $\sqrt{y} = 13.882\%$, that is $\pm 1\%$ as per Table 3.5. However, when the volatility increases (see for instance the third and the last graphs of Figure 3.4 (a)), $\mathrm{VIX}^2$-linked fees are also high (refer again to Table 3.5 for examples), making $\mathrm{VIX}^2$-linked fee variable annuity contracts less attractive than the constant fee ones. And so, when volatility is high, we observe that $f_y^{(m,N)}$ is pushed downward with increasing values of $\widetilde{m}^*$. In other words, for high volatility levels, variable annuity contracts with $\mathrm{VIX}^2$-linked fee structures are surrendered at lower fund values than contracts with constant fee structures. This is financially intuitive, as pointed out by Bernard *et al.* (2014b) under the Black-Scholes setting with a constant fee function, since when the fee gets higher, the policyholder has to pay more for the guarantee, and so, the mismatch between the premium for guarantee and its value is even greater; resulting in earlier exercise time. The analysis above shows that the findings of Bernard *et al.* (2014b) also extend to stochastic fee structures. Similar conclusions can be drawn for the Uncapped VIX-linked fees, Figure 3.4(c).



(a) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$



(b) $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}$, $K = 2\%$



(c) $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

Figure 3.4: The $y$ section of the approximated optimal surrender surface, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.

We observe in Figure 3.4(b) that when the volatility is low, capped $\mathrm{VIX}^2$-linked fee VA contracts are sur-

rendered at lower fund values than constant fee ones, like for the two other fee structures. However, for high enough volatility levels, the cap is reached, and thus, the fee paid under the capped structure is the same for all $m^* > 0$ (see Table 3.5); capped $\mathrm{VIX}^2$-linked fee VA contracts are thus all surrendered at similar fund value level (Figure 3.4 (b) graphs 3 and 4 when $m^* > 0$). This illustrates again the relation that exists between fees and optimal surrender decisions. VA contracts with higher fee rates are surrendered at lower fund values than contracts with lower fee rates.

Similar conclusions can be drawn under the Heston model when $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$ or when the interest rates depend on time; and under the 3/2 stochastic volatility model, see Appendix 3.8 for details.

## 3.6 Concluding Remarks

In this chapter, we provide a framework based on CTMC approximations to analyze the surrender incentives resulting from VIX-linked fees in variable annuities under general stochastic volatility models. Under the assumption that the policyholder maximizes the risk-neutral value of her variable annuity, the pricing of a variable annuity is an optimal stopping problem with a time-discontinuous reward function. Under general fee and surrender charge structures, we develop efficient numerical algorithms based on a two-layers CTMC approximation to price variable annuities with and without early surrenders. We derive a closed-form analytical formula for the value of a variable annuity without surrender rights and provide a quick and simple way of determining early surrenders value via a recursive algorithm. We also present an easy procedure to approximate the optimal surrender surface under the hypothesis that the surrender region is of threshold type. Finally, we observe numerically that VIX-linked fees do not significantly affect the value of early surrenders under the selected set of Heston parameters. However, numerical examples also reveal that the optimal surrender decision is impacted by VIX-linked fee structures. In particular, we observe that the optimal surrender strategy is more stable with respect to volatility changes when the fees are linked to the volatility index.

All algorithms and results of this chapter can easily be adapted to incorporate the term structure of interest rates by modeling the risk-free rate as a deterministic function of time. However, extension to stochastic interest rate models can present some numerical challenges. Indeed, adding a third dimension to the problem necessitates increasing the size of the generator, which can cause numerical challenges specifically for the valuation of long-term derivatives such as variable annuities. Algorithms 1 and 3 of this chapter address the numerical challenge encountered when trying to value long-term derivatives under general

two-dimensional models. The extension to higher dimension models is not straightforward and is left as future research. Moreover, short-rate models that allow to reproduce the term structure of interest rates are time-inhomogeneous, adding to the numerical difficulty of this extension for long-maturity derivatives since matrix exponential of the time-dependent generator now needs to be calculated at each time step to obtain the transition probability matrices. Finally, the CTMC approximation of the VIX presented in Section 3.4.4 can also be used to price exotic path-dependent options on the VIX, which is left as future research.

## 3.7 Appendix - Proofs of Main Results

### 3.7.1 Proof of Proposition 3.4

First, recall that

$$\begin{aligned}&\mathbb{P}(X_{t+h}^{(m,N)}=x_l,V_{t+h}^{(m)}=v_j\big|X_t^{(m,N)}=x_i,V_t^{(m)}=v_k)\\&\quad=\mathbb{P}(Y_{t+h}^{(m,N)}=(j-1)N+l\big|Y_t^{(m,N)}=(k-1)N+i).\end{aligned}$$

By inspection of the matrix $\mathbf{G}^{(m,N)}$ and since $h$ is small, we have by (3.8) that

$$\mathbb{P}(X_{t+h}^{(m,N)}=x_l,V_{t+h}^{(m)}=v_j\big|X_t^{(m,N)}=x_i,V_t^{(m)}=v_k)=\begin{cases}q_{kj}h+f_{ii}^{kj}(h) & \text{if } i=l,j\neq k\\ \lambda_{il}^k h+f_{il}^{kk}(h) & \text{if } i\neq l,j=k\\ 1+(q_{kk}+\lambda_{ii}^k)h+f_{ii}^{kk}(h) & \text{if } i=l,j=k\\ f_{il}^{kj}(h) & \text{if } i\neq l,\ j\neq k.\end{cases}$$

where the functions $\{f_{il}^{kj}\}$ satisfy $\lim_{h\to 0}\frac{f_{il}^{kj}(h)}{h}=0$ for $i,l\in\{1,2,\ldots,N\}$ and $k,j\in\{1,2,\ldots,m\}$.

From the last equality, we observe that the regime-switching CTMC cannot change regime ($V^{(m)}$) and state ($X^{(m,N)}$) simultaneously over small time intervals. It follows that

$$\begin{aligned}&\mathbb{E}\left[\phi\left(t+h,X_{t+h}^{(m,N)},V_{t+h}^{(m)}\right)|X_t^{(m,N)}=x_i,V_t^{(m)}=v_k\right]\\&=\sum_{\substack{j=1\\j\neq k}}^{m}\phi(t+h,x_i,v_j)(q_{kj}h+c_{ii}^{kj}(h))+\sum_{\substack{l=1\\l\neq i}}^{N}\phi(t+h,x_l,v_k)(\lambda_{il}^k h+f_{il}^{kk}(h))\\&\qquad+\phi(t+h,x_i,v_k)\left(1+(q_{kk}+\lambda_{ii}^k)h+f_{ii}^{kk}(h)\right)+\sum_{\substack{j=1\\j\neq k}}^{m}\sum_{\substack{l=1\\l\neq i}}^{N}\phi(t+h,x_l,v_j)f_{il}^{kj}(h)\\&=\phi(t+h,x_i,v_k)+\sum_{j=1}^{m}\phi(t+h,x_i,v_j)q_{kj}h+\sum_{l=1}^{N}\phi(t+h,x_l,v_k)\lambda_{il}^k h+f_N(h),\end{aligned}$$

since for $i, l \in \{1, 2, \ldots, N\}$ and $j, k \in \{1, 2, \ldots, m\}$, $\lim_{h\to 0} \frac{\phi(t+h,x_l,v_j) f_{il}^{kj}(h)}{h} = 0$.

Also, observe that

$$\begin{aligned}
&\mathbb{E}\left[\phi(t+h, X_{t+h}^{(m,N)}, v_j) | V_t^{(m)} = V_{t+h}^{(m)} = v_j, X_t^{(m,N)} = x_i\right] \\
&= \sum_{l=1}^{N} \phi(t+h, x_l, v_j) \mathbb{P}\left(X_{t+h}^{(m,N)} = x_l | V_t^{(m)} = V_{t+h}^{(m)} = v_j, X_t^{(m,N)} = x_i\right) \\
&= \sum_{\substack{l=1 \\ l\neq i}}^{N} \phi(t+h, x_l, v_j)(\lambda_{il}^j h + c_{il}^j(h)) + \phi(t+h, x_i, v_j)(1 + \lambda_{ii}^j h + f_{ii}^j(h)) \\
&= \phi(t+h, x_i, v_j) + \sum_{l=1}^{N} \phi(t+h, x_l, v_j)(\lambda_{il}^j h + f_{il}^j(h)).
\end{aligned}$$

Hence, we have that

$$\begin{aligned}
&\sum_{j=1}^{m} \mathbb{E}\left[\phi(t+h, X_{t+h}^{(m,N)}, v_j) | V_t^{(m)} = V_{t+h}^{(m)} = v_j, X_t^{(m,N)} = x_i\right] \mathbb{P}\left(V_{t+h}^{(m)} = v_j | V_t^{(m)} = v_k\right) \\
&= \sum_{j=1}^{m} \left(\phi(t+h, x_i, v_j) + \sum_{l=1}^{N} \phi(t+h, x_l, v_j)(\lambda_{il}^j h + f_{il}^j(h))\right)\left(q_{kj} h + f_{kj}(h)\right) \\
&\qquad + \phi(t+h, x_i, v_k) + \sum_{l=1}^{N} \phi(t+h, x_l, v_k)(\lambda_{il}^k h + f_{il}^k(h)) \\
= \quad &\phi(t+h, x_i, v_k) + \sum_{j=1}^{m} \phi(t+h, x_i, v_j) q_{kj} h + \sum_{l=1}^{N} \phi(t+h, x_l, v_k) \lambda_{il}^k h + f_m(h).
\end{aligned}$$

The results follows from setting $f(h) = f_m(h) - f_N(h)$. □

## 3.8 Appendix - Supplemental Material

This section provides supplemental material to Chapter 3.

### 3.8.1 Accuracy of the VA price and the Approximated Optimal Surrender Strategy

The purpose of this subsection is to analyze numerically the accuracy of Algorithm 4 to approximate the optimal surrender strategy. In the context of American put options under the Black-Scholes setting, Lamberton (1993) considers a similar approximation for the exercise boundary when the underlying diffusion process is approximated by a Markov chain. In particular, under this setting, Lamberton (1993) proves theoretically the convergence of the approximated exercise boundary (also called the critical price) towards the

exact exercise boundary in their Theorem 2.1.

In order to analyze the accuracy of our methodology against a known benchmark, we use our algorithms in a problem that has been done by others. More precisely, we compare the approach of Algorithm 4 (also called the CTMC Bermudian approximation and denoted below by "CTMC-Berm") for approximating the optimal surrender surface and the value of a variable annuity with early surrenders to the ones obtained using the approach of Bernard *et al*. (2014b), the "Benchmark". The methodology proposed by Bernard *et al*. (2014b) is based on the integral equation method of Kim (1990) under the Black-Scholes model, and cannot be easily extended to more general models. Thus, in order to compare with their method, we assume that the volatility of the index is constant so that

$$\mathrm{d}S_t, = rS_t\,\mathrm{d}t + \sigma S_t\,\mathrm{d}W_t, \quad t \geq 0,$$

where $\sigma > 0$ is the volatility. We also suppose that $c(x, y) = \widetilde{c} > 0$, for all $(x, y) \in \mathbb{R}_+ \times \mathcal{S}_V$, and we use $g(t, y) = 1$ for all $(t, y) \in [0, T] \times \mathcal{S}_V$. Hence, the VA account dynamics is given by

$$\mathrm{d}F_t, = (r - \widetilde{c})F_t\,\mathrm{d}t + \sigma F_t\,\mathrm{d}W_t, \quad 0 \leq t \leq T.$$

We approximate the logarithm of the fund process $\widetilde{X}_t = \ln F_t, 0 \leq t \leq T$ by using the CTMC approximation method. The CTMC approach for one-dimensional diffusion processes works in the exact same way as the one described in Subsection 3.3.1 for the approximation of $V$, see Cui *et al*. (2019) for details. The number of state-space grids and time steps are set to $N = 5,000$ and $M = T \times 500$, respectively. We use the grid of Tavella and Randall (2000), as discussed in Section 3.5.1, with a non-uniformity parameter $\widetilde{\alpha}_{\widetilde{X}} = 5$. The grid upper ($\widetilde{x}_N$) and lower bounds ($\widetilde{x}_1$) of the approximated process $\widetilde{X}$ are set about the mean of $\widetilde{X}_t$ at $t = T/2$ as follows: $\widetilde{x}_N = \bar{\mu}(t) + \gamma\bar{\sigma}(t)$ and $\widetilde{x}_1 = \bar{\mu}(t) - \gamma\bar{\sigma}(t)$ where $\bar{\mu}(t)$ is the conditional mean of $\widetilde{X}_t$ given $\widetilde{X}_0$ and $\bar{\sigma}(t)$, the conditional standard deviation. We use $\gamma = 7.2$. Now note that, under the Black-Scholes setting, we have

$$\bar{\mu}(t) = \widetilde{X}_0 + \left(r - \frac{\sigma^2}{2}\right)t, \quad \text{and} \quad \bar{\sigma}(t) = \sigma\sqrt{t}.$$

As a benchmark, we use the value of the variable annuity with early surrenders and the optimal surrender boundary obtained using the method of Bernard *et al*. (2014b) with $1500$ steps per year. Other parameters are $G = 100$, $r = 0.03$ and $T = 15$. We tested our method for different volatility levels $\sigma \in \{0.1, 0.2, 0.3, 0.4\}$ and initial values of the VA account $F_0 \in \{90, 100\}$. The fair fee $\widetilde{c}^*$ is calculated such that the expected discounted value of the maturity benefit equals the value of the VA account at inception, that is $F_0 = \mathbb{E}\left[e^{-rT}\max(G, F_T)\right]$. We compare to the benchmark the value of the variable annuity

Table 3.8: Relative errors of variable annuities with ES and maximum relative errors of the optimal surrender boundaries.

| | $\sigma$ | $\widetilde{c}^*$ (%) | VA with ES | | | Opt. Surr. Bound. |
|---|---|---|---|---|---|---|
| | | | CTMC-Berm | Bernard et al. (2014) | Rel. Err. | Max Rel. Err. |
| $F_0 = 100$ | 0.1 | 0.1374 | 100.851600 | 100.851748 | 1.47e-6 | 2.60e-3 |
| | 0.2 | 0.9094 | 104.400379 | 100.401287 | 8.70e-6 | 5.20e-3 |
| | 0.3 | 1.9277 | 108.577366 | 108.579001 | 1.51e-5 | 7.78e-3 |
| | 0.4 | 2.9415 | 112.823616 | 112.826112 | 2.21e-5 | 1.04e-2 |
| $F_0 = 90$ | 0.1 | 0.2641 | 91.28494 | 91.285171 | 2.53e-6 | 2.60e-3 |
| | 0.2 | 1.3062 | 94.989758 | 94.990712 | 1.01e-5 | 5.20e-3 |
| | 0.3 | 2.5571 | 99.012022 | 99.013806 | 1.80e-5 | 7.78e-3 |
| | 0.4 | 3.7631 | 103.022547 | 103.025197 | 2.57e-5 | 1.03e-2 |

with early surrenders "VA with ES" and the optimal surrender boundary ("Opt. Surr. Bound."). The results are reported in Table 3.8. The column "Rel. Err." documents the relative errors whereas "Max Re. Err" documents the maximum relative errors[20]. From Table 3.8, we see that the CTMC Bermudian approximation achieves a high level of accuracy across all volatility levels and initial value of VA account $F_0$.

### 3.8.2 Accuracy and Efficiency of the CTMC Approximation for the VIX Index

We assess the accuracy of the CTMC approximation for the volatility index for the 3/2 and Heston models. Under a Heston-type model, the VIX has a closed-form expression given by

$$\mathrm{VIX}_t^2 = B + AV_t \tag{3.38}$$

with $A = \frac{1-e^{-\kappa\tau}}{\kappa\tau}$ and $B = \frac{\theta(\kappa\tau-1+e^{-\kappa\tau})}{\kappa\tau}$, see Zhu and Zhang (2007) for details.

For the 3/2 model, a closed-form expression for the VIX may be found in Carr and Sun (2007), Theorem 4. However, as pointed out by Drimus (2012), the integral that appears in the analytical formula is difficult to implement and is not suited for fast and accurate numerical approximation. For this reason, benchmark

[20] Given a benchmark value $y$ and its approximation $y_{approx}$, the relative error is defined as $|y - y_{approx}|/|y|$ for $y \neq 0$; whereas the maximum relative error is the largest relative error over a sample of benchmark values and their approximations.

results are obtained via Monte Carlo simulation[21] under 3/2 model.

The market and CTMC parameters are set to the ones of Subsection 3.5.1 for the Heston model, except that we set $m = 1,000$ (rather than $m = 50$). For the 3/2 model, we use the parameters reported in Section 3.8.6.1. The results are presented in Table 3.9. The column "$V_t$" (resp. $1/V_t$) displays the initial value of the variance at time $t \geq 0$ for the Heston model (resp. the 3/2 model), the column "CTMC" shows the result of the VIX approximation using Algorithm 6 with $n = 1,000$ time steps, and the column "Benchmark" reports the benchmark value calculated either by using the closed-form formula (3.38) (for the Heston model) or by Monte Carlo simulation (for the 3/2 model). The relative error are documented in column "Rel. err".

Table 3.9: Accuracy of the CTMC- VIX approximation, Algorithm 6.

(a) Heston model

| $\mathbf{V_t}$ | CTMC | Benchmark | Rel. err. |
|---|---|---|---|
| 0.01 | 11.1077% | 11.1068% | 8.10e-05 |
| 0.02 | 14.6829% | 14.6824% | 3.41e-05 |
| 0.04 | 20.0000% | 20.0000% | 0.00e+00 |
| 0.06 | 24.1746% | 24.1749% | 1.24e-05 |
| 0.09 | 29.3433% | 29.3439% | 2.04e-05 |

(b) 3/2 model

| $1/\mathbf{V_t}$ | CTMC | Benchmark | Rel. err. |
|---|---|---|---|
| 0.01 | 11.2202% | 11.2203% | 8.91e-06 |
| 0.02 | 15.7016% | 15.7020% | 2.55e-05 |
| 0.04 | 21.4892% | 21.4901% | 4.19e-05 |
| 0.06 | 25.3096% | 25.3112% | 6.32e-05 |
| 0.09 | 29.2708% | 29.2735%5 | 9.22e-05 |

The results of Table 3.9 confirm the high accuracy of the CTMC-VIX approximation for both models. It is worth mentioning that, when using Algorithm 6, we obtain simultaneously the value of the $\mathrm{VIX}_t^{(m),k}$ for all $v_k \in \mathcal{S}_V^{(m)}$ within less than a 0.1 second for both models. The value of the CTMC approximation given a particular value for $V_t$ is then interpolated between the appropriate grid points. This further increases the efficiency of our algorithm.

### 3.8.3 VA Prices Accuracy and Computation Time under the Heston Model

Table 3.10 shows the value of a variable annuity with and without surrender rights ("VA with ES" and "VA without ES", respectively) under the Heston model with the Uncapped $\mathrm{VIX}^2$-linked fees using $N = 100$, $N = 1,100$ and $M = 252 \times 10$; and using $N = 2,000$ and $M = 500 \times 10$. All other market, VA and CTMC

[21] We simulate 500K paths (plus 500K antithetic variables) using Milstein discretization scheme and we use 5,000-time steps.

parameters are the same as in Subsection 3.5.1.

Table 3.10: Variable annuity with and without early surrenders using CTMC Approximation with $N = 100$ and $1,100$, with $M = 252 \times 10$; and $N = 2,000$ with $M = 500 \times 10$.

| | $\widetilde{\mathbf{m}}^* =$ | $\mathbf{0.0000}$ | $\mathbf{0.1500}$ | $\mathbf{0.3000}$ | $\mathbf{0.4345}$ | |
|---|---|---|---|---|---|---|
| | $\mathbf{c}^* =$ | $\mathbf{1.5338\%}$ | $\mathbf{1.0036\%}$ | $\mathbf{0.4741\%}$ | $\mathbf{0.000\%}$ | **Computation Time (sec.)** |
| | $N = 2,000$ | 100.00090 | 100.00091 | 100.00092 | 100.00093 | $2,500$ |
| **VA without ES** | $N = 1,100$ | 100.00094 | 100.00095 | 100.00096 | 100.00097 | $400$ |
| | $N = 100$ | 100.00750 | 100.00769 | 100.00789 | 100.00807 | $0.85$ |
| | $N = 2,000$ | 103.01785 | 103.00823 | 103.00330 | 103.00367 | $7,100$ |
| **VA with ES** | $N = 1,100$ | 103.01743 | 103.00788 | 103.00300 | 103.00341 | $1,600$ |
| | $N = 100$ | 103.0162 | 103.00676 | 103.00190 | 103.00228 | $54$ |
| | $N = 2,000$ | 3.01695 | 3.00732 | 3.00238 | 3.00274 | N/A |
| **ES value** | $N = 1,100$ | 3.01649 | 3.00693 | 3.00204 | 3.00243 | N/A |
| | $N = 100$ | 3.00862 | 2.99907 | 2.99401 | 2.99421 | N/A |

All the numerical results and computation times in Table 3.10 are performed using Equation 3.22 and Algorithm 2 combine with the Expokit procedures of Sidje (1998), function *expv*. Since fair fees are calibrated at inception using the exact pricing formula of Cui *et al.* (2017a), the benchmark value for the VA without surrender rights is $F_0 = 100$. When $N = 100$, accurate VA prices are obtained extremely fast (within $54$ sec. for the VA with surrender rights and less than a second for the VA without surrender rights). The absolute error is around $10^{-3}$ for both values[22].

Similar results are obtained with the Fast Algorithms (1 and 3). The results are reported in Table 3.11. The values of VA with and without surrender rights are calculated simultaneously, as discussed in Section 3.4.2.1. So, the computation times in column "Computation Time (sec.)" are for both prices.

By comparing the two tables, we note that the Fast Algorithms provide very accurate results extremely fast[23] compared to the original Algorithms.

[22] The benchmark for the VA with early surrenders is the approximated CTMC value obtained using $N = 2,000$ and $M = 500$.

[23] Recall that the computation times in Table 3.11 for the Fast Algorithms are for the value of the VAs with and without ES simultaneously.

Table 3.11: Variable annuity with and without early surrenders (ES) using CTMC Approximation Fast Algorithms with $N = 100$ and $1,100$, with $M = 252 \times 10$; and $N = 2,000$ with $M = 500 \times 10$.

| | $\widetilde{\mathbf{m}}^* =$ | $\mathbf{0.0000}$ | $\mathbf{0.1500}$ | $\mathbf{0.3000}$ | $\mathbf{0.4345}$ | |
|---|---|---|---|---|---|---|
| | $\mathbf{c}^* =$ | $\mathbf{1.5338}\%$ | $\mathbf{1.0036}\%$ | $\mathbf{0.4741}\%$ | $\mathbf{0.000}\%$ | **Computation Time (sec.)** |
| | $N = 2,000$ | 99.99615 | 99.99611 | 99.99607 | 99.99603 | N/A |
| **VA without ES** | $N = 1,100$ | 99.99619 | 99.99615 | 99.99611 | 99.99607 | N/A |
| | $N = 100$ | 100.00276 | 100.00290 | 100.00304 | 100.00317 | N/A |
| | $N = 2,000$ | 103.02361 | 103.01360 | 103.00815 | 103.00789 | $4,405$ |
| **VA with ES** | $N = 1,100$ | 103.02360 | 103.01359 | 103.00814 | 103.00788 | $1,244$ |
| | $N = 100$ | 103.02237 | 103.01251 | 103.00715 | 103.00680 | 9.60 |
| | $N = 2,000$ | 3.02745 | 3.01748 | 3.01208 | 3.01186 | N/A |
| **ES value** | $N = 1,100$ | 3.02741 | 3.01744 | 3.01203 | 3.01181 | N/A |
| | $N = 100$ | 3.01961 | 3.00961 | 3.00411 | 3.00363 | N/A |

### 3.8.4 Other Numerical Analysis of VIX-linked Fee Incentives in the Heston Model

In this appendix, we show the numerical results obtained under the Heston model for the three fee structures (Uncapped $\mathrm{VIX}^2$, Capped $\mathrm{VIX}^2$ and Uncapped VIX), detailed in Subsection 3.5.3, when the guaranteed amount is set to $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$ (rather than $G = F_0$). The conclusions are similar to the ones detailed in Subsection 3.5.4.

Unless stated otherwise, in this section, all market, VA and CTMC parameters are the same as in Subsection 3.5.1 except for $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$ and $M = T \times 252$.

Fair fee parameters $(\widetilde{c}^*, \widetilde{m}^*)$ are presented in Table 3.12, Panel (a), whereas Panel (b) shows the approximated values of early surrenders. When $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$, the fair parameters are obtained using the exact formula of Cui *et al.* (2017a). For the other fee structures, the fair parameters are obtained using CTMC approximation with $N = 100$, $m = 50$, and $M = T \times 252$ to accelerate the computation time.

Figures 3.5, 3.6 and 3.7 display the approximated optimal surrender surfaces under the Heston model for the three fee structures when $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$.

Table 3.12: Fair fee vectors and approximated early surrender values under Heston model when $G = F_0 e^{\widetilde{g}T}$, $\widetilde{g} = 2\%$.

(a) Fair fee vectors $(\widetilde{c}^*, \widetilde{m}^*)$

| $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$ | | | | |
|---|---|---|---|---|
| $\widetilde{m}^*$ | 0.0000 | 0.3000 | 0.8000 | 1.1313 |
| $\widetilde{c}^*$ | 3.82542% | 2.80632% | 1.11545% | 0.0000% |
| $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 4.5\%$ | | | | |
| $\widetilde{m}^*$ | 0.0000 | 0.3000 | 0.8000 | 1.4144 |
| $\widetilde{c}^*$ | 3.82542% | 2.83600% | 1.45530% | 0.0000% |
| $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$ | | | | |
| $\widetilde{m}^*$ | 0.0000 | 0.0750 | 0.1250 | 0.2128 |
| $\widetilde{c}^*$ | 3.82542% | 2.47570% | 1.57660% | 0.00000% |

(b) Early surrender values ("ES" values).

| $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$ | | | | |
|---|---|---|---|---|
| $\widetilde{m}^*$ | 0.0000 | 0.3000 | 0.8000 | 1.1313 |
| ES Value | 5.00216 | 5.01219 | 5.03998 | 5.06681 |
| $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 4.5\%$ | | | | |
| $\widetilde{m}^*$ | 0.0000 | 0.3000 | 0.8000 | 1.4144 |
| ES Value | 5.00216 | 5.00415 | 4.98201 | 4.95968 |
| $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$ | | | | |
| $\widetilde{m}^*$ | 0.0000 | 0.0750 | 0.1250 | 0.2128 |
| ES Value | 5.00216 | 5.00351 | 5.00569 | 5.01254 |

### 3.8.5 Simple Model Extension to Account for the Term Structure of Interest Rates

In the following, we briefly discuss a simple extension to (3.4) to allow interest rates to be a (deterministic) function of time $\widetilde{r} : [0, T] \to [0, 1]$.

Assuming that the fee function only depends on the VIX (or $V_t$) and using the new function for the interest rates, $\widetilde{r}$, we can extend (3.4) as follows

$$\begin{aligned} \mathrm{d}F_t &= (\widetilde{r}(t) - c(V_t))\, F_t\, \mathrm{d}t + \sigma_S(V_t) F_t\, \mathrm{d}W_t^{(1)}, \\ \mathrm{d}V_t &= \mu_V(V_t)\, \mathrm{d}t + \sigma_V(V_t)\, \mathrm{d}W_t^{(2)}. \end{aligned} \tag{3.39}$$

Such an extension allows the exact replication of the term structure of interest rates. This is an important feature given the long-term maturity of variable annuity contracts.

However, the CTMC approximation is better suited for time-homogeneous processes. Recently, Ding and Ning (2021) uses CTMCs to approximate time-inhomogeneous diffusion processes and estimate short-term maturity European and barrier option prices. Theoretical extension to longer-term derivatives is straightforward but can present numerical challenges since matrix exponential of a large matrix now needs to be calculated several times to obtain transition probability matrices. More precisely, the Fast Algorithms (Algorithms 1 and 3) take advantage of pre-computing the transition probabilities at the beginning of the procedure. When the generator is time-dependent, transition probabilities need to be calculated at each time step, adding to the numerical difficulty of valuing long-term derivatives. In the following, we overcome

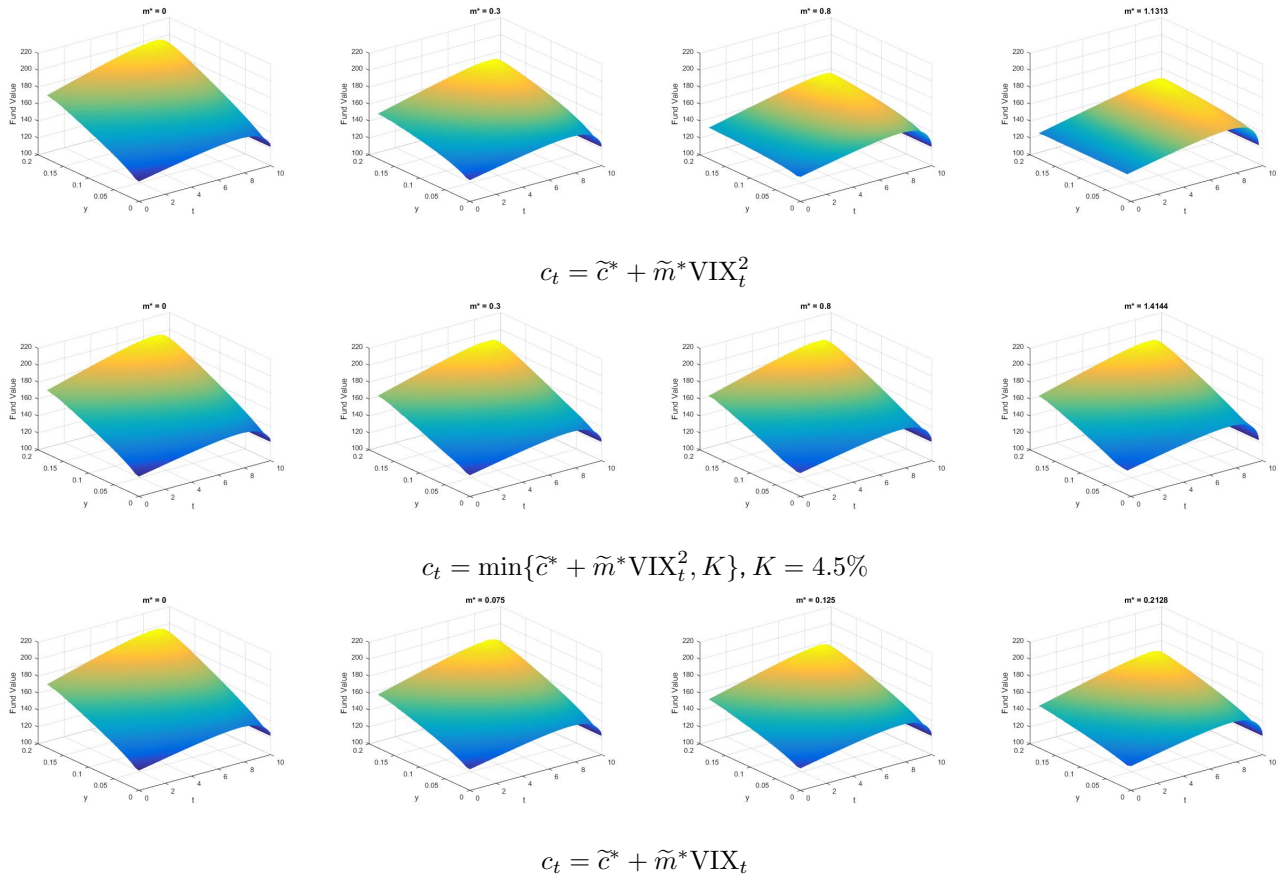



Figure 3.5: Approximated optimal surrender surfaces for different values of fair multiplier $\widetilde{m}^*$ under the Heston model, $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$.

this challenge by working with the discounted fund process $\{F_t\}_{0 \leq t \leq T}$ defined by

$$\widetilde{F}_t = F_t e^{-\int_0^t \widetilde{r}(s)\,\mathrm{d}s}. \tag{3.40}$$

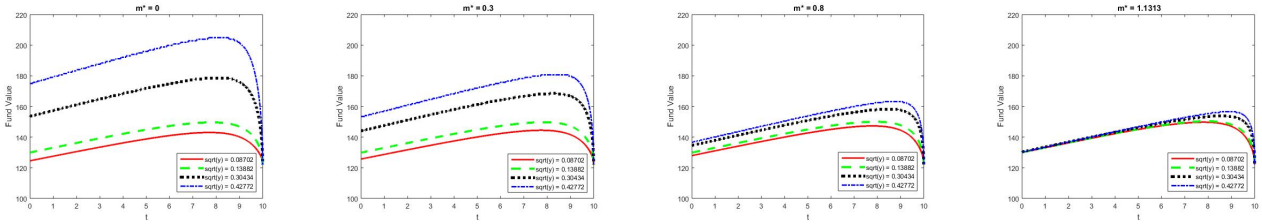
Using Itô's lemma, the dynamics of $\widetilde{F}$ is then given by

$$\mathrm{d}\widetilde{F}_t = -c(V_t)\widetilde{F}_t\,\mathrm{d}t + \sigma_S(V_t)\widetilde{F}_t\,\mathrm{d}W_t^{(1)}. \tag{3.41}$$

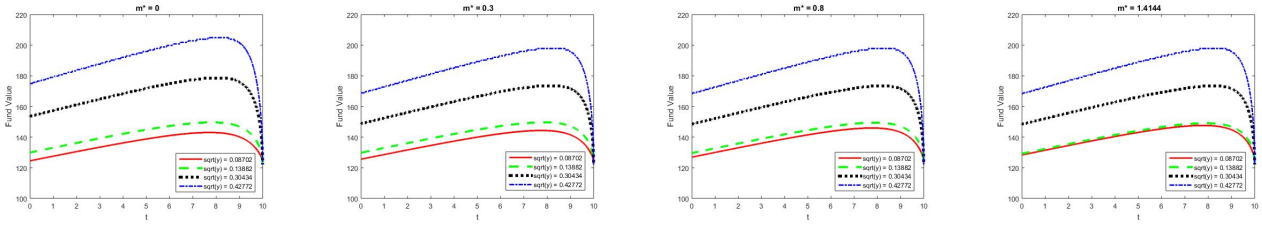
Now note that $\widetilde{F}$ is time-homogeneous and its dynamic is equivalent to (3.4) with $r = 0$. We can thus apply the same procedure as the one described in Section 3.3 to approximate $\widetilde{F}$ by a CTMC. We denote by $\widetilde{F}^{(m,N)}$ the CTMC approximation of $\widetilde{F}$. Then, we can retrieve the value of the approximated fund process $F^{(m,N)}$ by using the relation in (3.40). More precisely, we have that

$$F_t^{(m,N)} := \widetilde{F}_t^{(m,N)} e^{\int_0^t \widetilde{r}(s)\,\mathrm{d}s}, \quad 0 \leq t \leq T. \tag{3.42}$$

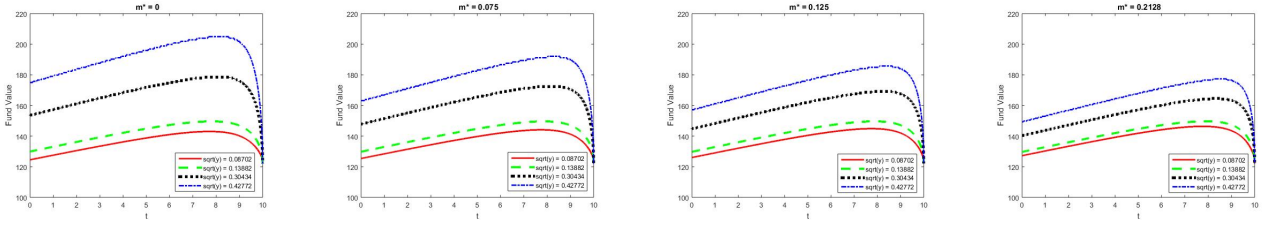
All the results of Section 3.4 can then be adapted with the new definition of the approximated fund process




$$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$$




$$c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 4.5\%$$




$$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$$

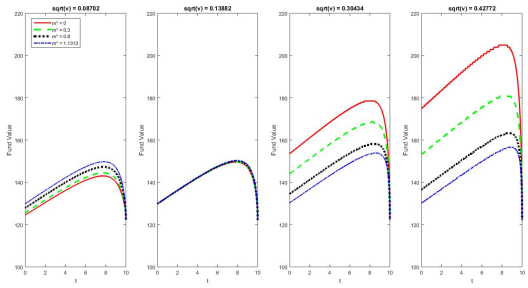
Figure 3.6: The $y$ section of the approximated optimal surrender surface under the Heston model, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$, $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$.

given in (3.42). Note that $D(t,T) := e^{-\int_t^T \widetilde{r}(s)\,\mathrm{d}s}$ corresponds to the discount factor at time $t$ for the maturity $T$, and since the interest rates are assumed to be deterministic, this quantity is known with certainty at $t$. The curve $t \mapsto D(0,t)$ corresponds to the risk-free discount curve and can be retrieved from market data at time $0$. Also, for $t_i < t_j$, we have that

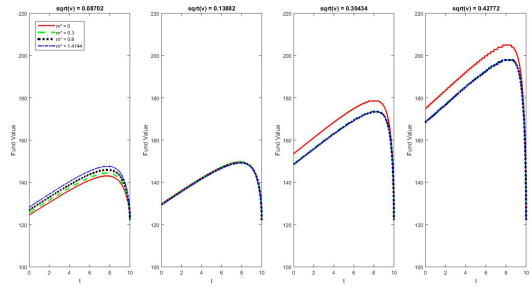
$$D(t_i, t_j) = \frac{D(0,t_j)}{D(0,t_i)}.$$

Finally, all the algorithms in Section 3.4 must be adjusted to account for changing interest rates. So, the cash-flows at time $t_i + \Delta_M$ must be discounted at time $t_i$ using the discount factor $D(t_i, t_i + \Delta_M)$. For example, (3.22) becomes

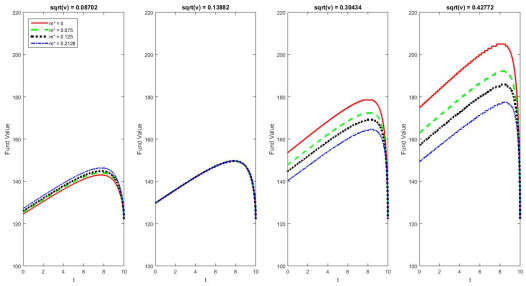
$$\begin{aligned} v_e^{(m,N)}(0,F_0,V_0) &:= \mathbb{E}[e^{-\int_0^T r(s)\,\mathrm{d}s} \max(G, \widetilde{F}_T^{(m,N)} e^{\int_0^T r(s)\,\mathrm{d}s} | \widetilde{F}_0^{(m,N)} = F_0, V_0^{(m)} = V_0] \\ &= D(0,T)\mathbf{e}_{ik} \exp\{T\mathbf{G}^{(m,N)}\}\mathbf{H}, \end{aligned} \tag{3.43}$$



$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$



$c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 4.5\%$



$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

Figure 3.7: The $y$ section of the approximated optimal surrender surface under the Heston model, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$, $G = F_0 e^{\widetilde{g}T}$ with $\widetilde{g} = 2\%$.

and (3.27) becomes

$$\begin{cases} B_M^{(m,N)} & = \varphi\left(T, \widetilde{F}_T^{(m,N)} e^{\int_0^T r(s)\,\mathrm{d}s}, V_T^{(m)}\right), \\ B_z^{(m,N)} & = \max\left(\varphi\left(t_z, \widetilde{F}_{t_z}^{(m,N)} e^{\int_0^{t_z} r(s)\,\mathrm{d}s}, V_{t_z}^{(m)}\right), D(t_z, t_{z+1}) \mathbb{E}_{t_z}\left[B_{z+1}^{(m,N)}\right]\right), \end{cases} \tag{3.44}$$

for $0 \leq z \leq M-1$. We now analyze the impact of time-dependent risk-free rates on optimal surrender strategy and the value of variable annuities. To do so, we use the discount curve reported in Table 3.13.

Table 3.13: Discount curve

| $\mathbf{t}$ | 0.20 | 0.50 | 0.70 | 1.00 | 1.20 | 1.50 | 1.70 | 2.00 | 3.00 | 4.00 | 5.00 | 6.00 | 7.00 | 8.00 | 9.00 | 10.00 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Zero Rate**[24] | 0.0050 | 0.0075 | 0.0100 | 0.0125 | 0.0150 | 0.0175 | 0.0200 | 0.0210 | 0.0220 | 0.0230 | 0.0250 | 0.0270 | 0.0280 | 0.0290 | 0.0295 | 0.0300 |
| $\mathbf{D(0,t)}$ | 0.9963 | 0.9930 | 0.9876 | 0.9822 | 0.9741 | 0.9666 | 0.9589 | 0.9361 | 0.9121 | 0.8825 | 0.8504 | 0.8220 | 0.7929 | 0.7668 | 0.7408 | 0.7111 |

In order to measure the impact of time-dependent risk-free rates on surrender incentives, the discount curve is set in such a way that the 10-year zero rate corresponds to the constant risk-free rate used in Section

[24] The continuously-compounded zero rate, $R(0,t)$, is the interest rate prevailing at time $0$ for the maturity $t$ and it is defined by $R(0,t) := -lnD(0,t)/t$.

3.5. This way fair fees are the same under both models, see Table 3.4. Also, to better capture the effect of interest rates, the short-term zero rates are set to a much lower value than long-term rates. Except for the constant risk-free rate, all other market and CTMC parameters are the same as in the previous analysis, see Tables 3.2 and 3.3. All numerical examples below are performed for a variable annuity with the Uncapped $\mathrm{VIX}^2$ fee structure (see Table 3.6). The values of a variable annuity with and without surrender rights are reported in Table 3.14.

Table 3.14: Variable annuity with and without early surrenders (ES) under Heston model with time-dependent interest rates and $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$, $0 \leq t \leq T$.

| $\widetilde{\mathbf{m}}^* =$ | $\mathbf{0.0000}$ | $\mathbf{0.1500}$ | $\mathbf{0.3000}$ | $\mathbf{0.4345}$ |
|---|---|---|---|---|
| $\widetilde{\mathbf{c}}^* =$ | $\mathbf{1.5338\%}$ | $\mathbf{1.0036\%}$ | $\mathbf{0.4741\%}$ | $\mathbf{0.000\%}$ |
| **VA without ES** | 100.00087 | 100.00088 | 100.00088 | 100.00089 |
| **VA with ES** | 103.01783 | 103.00821 | 103.00328 | 103.00365 |
| **Value of ES** | 3.01697 | 3.00733 | 3.00239 | 3.00276 |

By comparing Table 3.7, Panel (a) and Table 3.14, we note that variable annuity values are almost unchanged with the incorporation of time-dependent risk-free rates. The small difference is probably due to the rounding-off effect of the discount factor at maturity when using the time-dependent discount curve as an input. However, as illustrated below, time-dependent risk-free rates affect the optimal surrender strategy.

First, we observe from Figures 3.8, 3.9 and 3.10, that the conclusions regarding the effects of VIX link fees on the optimal surrender strategy remain the same under both assumptions (constant and time-dependent risk-free rates). However, we also note, from Figure 3.11, that the optimal surrenders occur at lower fund values when the interest rates are time-dependent. So, incorporating the term structure of interest rates does not affect the overall value of the variable annuity, but the optimal surrender behavior is impacted by this new interest rate assumption.

### 3.8.6 Numerical Analysis under 3/2 Model

In this section, we analyze numerically the impact of VIX-linked fee incentives under the 3/2 model.

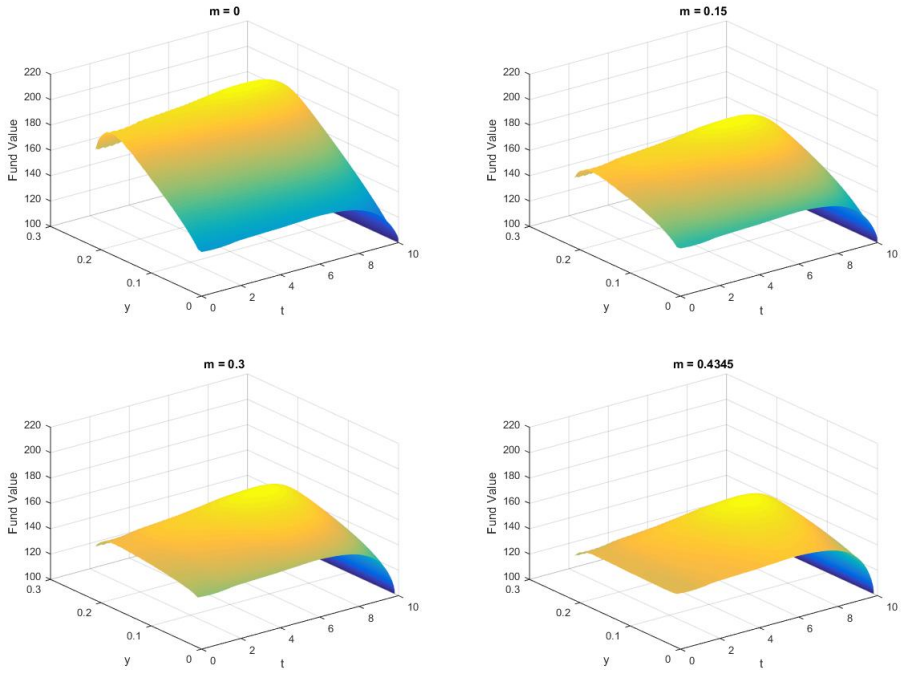



Figure 3.8: Approximated optimal surrender surface of VIX-linked fees VAs for different values of fair multiplier $\widetilde{m}^*$ under the Heston model with time-dependent risk-free rates. The $x$-axis represents the time and the $y$-axis the variance.

#### 3.8.6.1 Market, VA and CTMC Parameters

In order for the results under the 3/2 model to be comparable to the ones obtained under the Heston model, the market parameters ($\theta$, $\kappa$ and $\sigma$ for the 3/2 model) are calibrated to at-the-money call options[25] priced using the Heston model with the market parameters in Table 3.2. The initial value of the variance is set to $V_0 = 0.03$ for the Heston model and $V_0 = 1/0.03$ for the 3/2 model. The correlation $\rho = -0.75$ and the risk-free rate $r = 0.03$ are assumed to be the same under both models. The resulting market parameters are presented in Table 3.15 and the model dynamics is given in Table 3.1.

Table 3.15: Market parameters for the 3/2 model

| Parameter | $V_0$ | $\kappa$ | $\theta$ | $\sigma$ | $\rho$ | $r$ |
|---|---|---|---|---|---|---|
| Value | $1/0.03$ | $5.7488$ | $46.1326$ | $15.4320$ | $-0.75$ | $0.03$ |

[25] We consider 4 options with maturity $T = 0.5,\ 2.5,\ 5$ and $10$. We used $S_0 = K = 100$ and a dividend yield of $1.5338\%$.

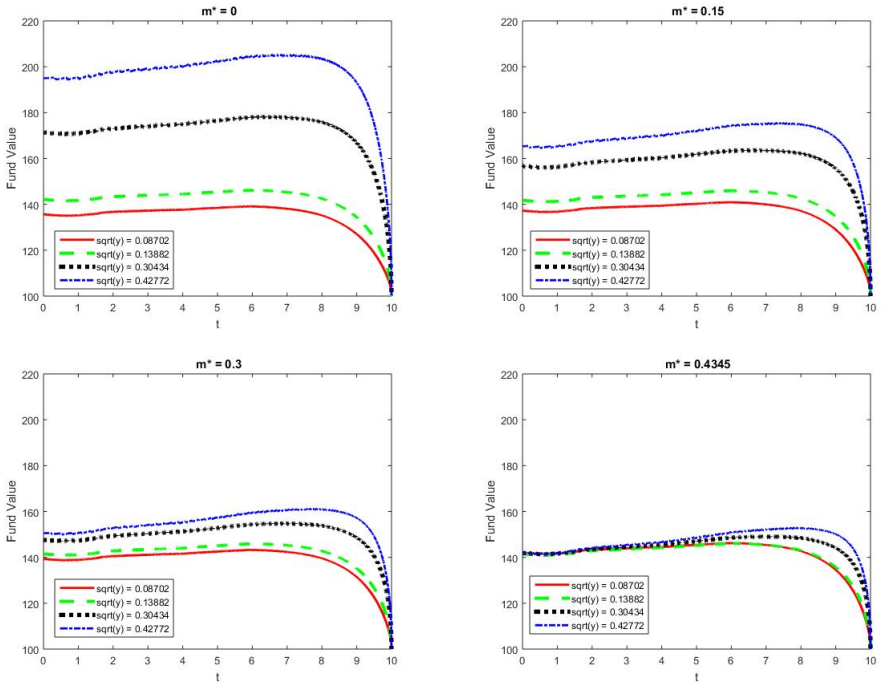



Figure 3.9: The $y$ section of the approximated optimal surrender surface under the Heston model with time-dependent risk-free rates, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.

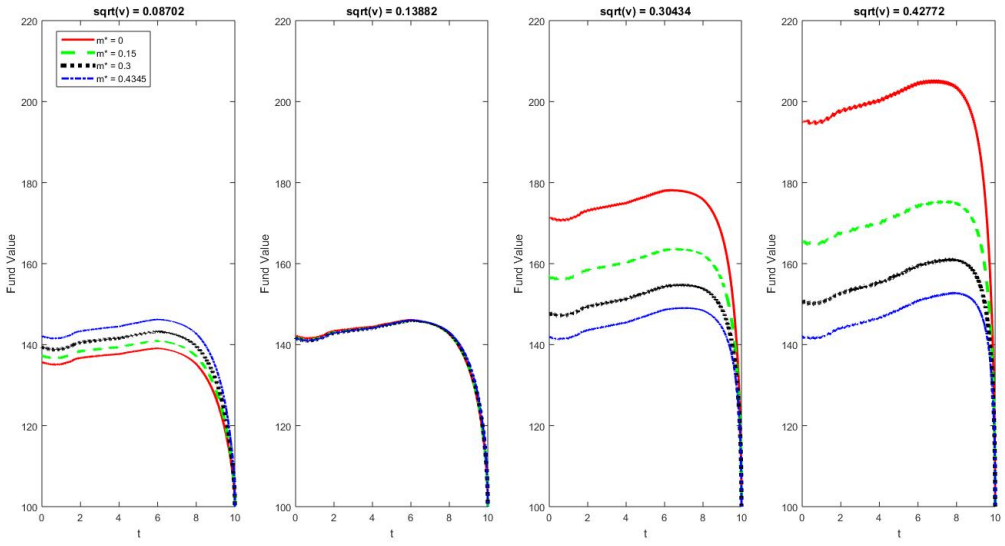



Figure 3.10: The $y$ section of the approximated optimal surrender surface under the Heston model with time-dependent risk-free rates, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.

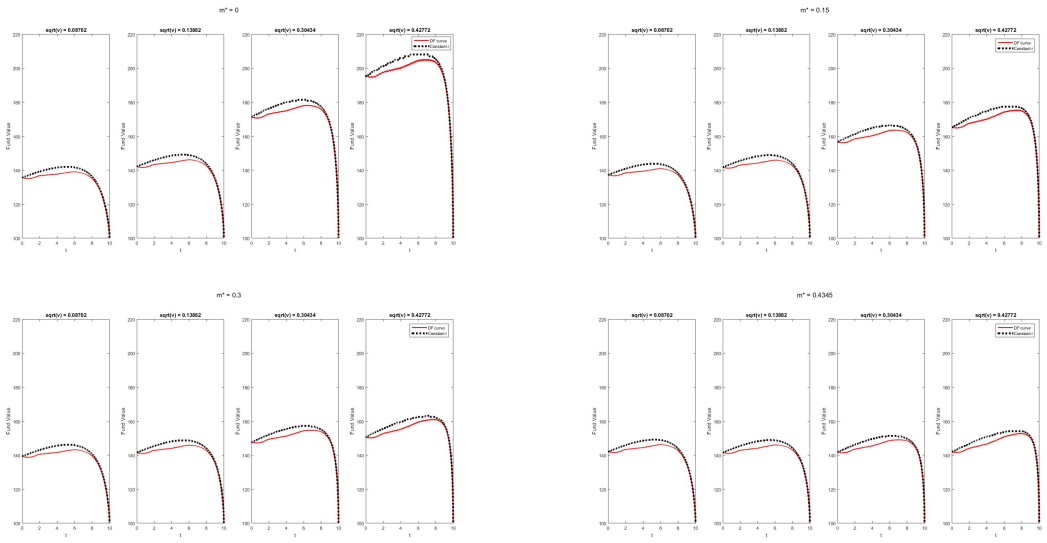


Figure 3.11: Comparison of the $y$ section of the approximated optimal surrender surface under the Heston model with time-dependent risk-free rates and constant risk-free rates at different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.

Unless stated otherwise, all numerical experiments in this section are performed using the CTMC parameters listed in Table 3.16. Note that for the 3/2 model, the VIX is approximated using Algorithm 6 with $n = 1,000$ time steps, whereas in the Heston model, the volatility index is calculated using a closed-form formula.

Table 3.16: CTMC parameters for the 3/2 Model

| Parameter | $m$ | $N$ | $v_1$ | $v_m$ | $\widetilde{\alpha}_v$ | $x_1$ | $x_N$ | $\widetilde{\alpha}_X$ | $M$ | n |
|---|---|---|---|---|---|---|---|---|---|---|
| Value | $1,000$ | $1,000$ | $V_0/200$ | $8v_0$ | $0.5764$ | $-X_0$ | $2X_0$ | $2/100$ | $5,000$ | 1,000 |

### 3.8.6.2 Fee Structures and Fair Fee Parameters

As for the Heston model, we consider three fee structures: the Uncapped $\mathrm{VIX}^2$, the Capped $\mathrm{VIX}^2$ and the Uncapped VIX, see Section 3.5.3 for details. The fair parameters are calibrated using a CTMC approximation with $N = 100$ (all other CTMC parameters are the same as in Table 3.16), to reduce the computation time. Table 3.17 presents the fair fee vectors $(\widetilde{c}^*, \widetilde{m}^*)$ obtained under the 3/2 model. Note how close the fair fee vectors obtained under the 3/2 model (Table 3.17) are from the ones obtained under the Heston model (Table 3.4).

| $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$ | | | | |
|---|---|---|---|---|
| $\widetilde{m}^*$ | $0.0000$ | $0.15000$ | $0.3000$ | $0.4235$ |
| $\widetilde{c}^*$ | $1.5273\%$ | $0.9858\%$ | $0.4449\%$ | $0.0000\%$ |
| $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 2\%$ | | | | |
| $\widetilde{m}^*$ | $0.0000$ | $0.1500$ | $0.3000$ | $0.5791$ |
| $\widetilde{c}^*$ | $1.5273\%$ | $1.0335\%$ | $0.6351\%$ | $0.0000\%$ |
| $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$ | | | | |
| $\widetilde{m}^*$ | $0.0000$ | $0.0250$ | $0.0500$ | $0.0846$ |
| $\widetilde{c}^*$ | $1.5273\%$ | $1.0757\%$ | $0.6241\%$ | $0.0000\%$ |

Table 3.17: Fair fee vectors $(\widetilde{c}^*, \widetilde{m}^*)$ under the 3/2 model

#### 3.8.6.3 Effect of VIX-Linked Fees on Surrender Incentives

Recall from Table 3.1 that under the 3/2 model, the dynamics of the index and the variance processes are given by[26]

$$\begin{aligned} \mathrm{d}S_t &= rS_t\,\mathrm{d}t + \tfrac{1}{\sqrt{V_t}} S_t\,\mathrm{d}W_t^{(1)}, \\ \mathrm{d}V_t &= \kappa(\theta - V_t)\,\mathrm{d}t - \sigma\sqrt{V_t}\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{3.45}$$

where $W = (W^{(1)}, W^{(2)})^T$ is a bi-dimensional correlated Brownian motion under $\mathbb{Q}$ and such that $[W^{(1)},\, W^{(2)}]_t = \rho t$ with $\rho \in [-1, 0]$[27], and $\kappa, \theta, \sigma > 0$ with $2\kappa\theta > \sigma^2$.

Now from Lemma 3.1, we find that $\gamma(x) = -\frac{\ln(x)}{\sigma}$. Thus, given a certain fee process $\{c_t\}_{0\le t\le T}$ (see Subsection 3.5.3 for examples), the dynamics of the auxiliary process is obtained as follows

$$\begin{aligned} \mathrm{d}X_t &= \mu_X(X_t, Y_t)\,\mathrm{d}t + \sigma_X(Y_t)\,\mathrm{d}W_t^*, \\ \mathrm{d}V_t &= \mu_V(V_t)\,\mathrm{d}t + \sigma_V(V_t)\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{3.46}$$

where $\mu_X(X_t, V_t) = r - c_t - \frac{\rho\kappa}{\sigma} + \frac{1}{V_t}\left(\frac{\rho\kappa\theta}{\sigma} - \frac{1}{2} - \frac{\rho\sigma}{2}\right)$, and $\sigma_X(V_t) = \sqrt{\frac{1-\rho^2}{V_t}}$, $0 \le t \le T$.

[26] Note that in this formulation of the 3/2 model, the process $V$ represents the inverse of the variance process. It is also common to see the 3/2 model expressed in terms of its variance, see for instance Drimus (2012).

[27] The parameter $\rho$ is assumed to be non-positive for the martingale measure to exist, see Footnote 7 for details.

When the fee is tied to the VIX, the form of the fee function is not known explicitly at this point under the 3/2 model, since the VIX does not have a closed-form expression (see footnote 15 for details). However, as illustrated in (3.17), when the inverse variance process $V$ is replaced by its CTMC approximation $V^{(m)}$, the auxiliary process $X$ becomes a RS diffusion $X^{(m)}$ with the following dynamics:

$$\mathrm{d}X_t^{(m)} = r - c_t^{(m)} - \tfrac{\rho\kappa}{\theta} + \tfrac{1}{V_t^{(m)}}\left(\tfrac{\rho\kappa\theta}{\sigma} - \tfrac{1}{2} - \tfrac{\rho\sigma}{2}\right)\mathrm{d}t + \sigma_X(V_t^{(m)})\,\mathrm{d}W_t^*, \quad t \geq 0, \tag{3.47}$$

where $c_t^{(m)} = c(X_t^{(m)}, V_t^{(m)})$ is the CTMC approximation of the fee process.

Recall that $\mathrm{VIX}^{(m)}$ is the CTMC approximation of the volatility index; see Proposition 3.7 and Algorithm 6. The three fee structures exposed in Subsection 3.5.3 can then be approximated using $\mathrm{VIX}^{(m)}$ as shown in Table 3.18.

Table 3.18: CTMC approximation of the VIX-linked fee process

| **Fee Structure** | $c_t^{(m)}, 0 \leq t \leq T$ |
|---|---|
| **Uncapped** $\mathrm{VIX}^2$ | $\widetilde{c} + \widetilde{m}(\mathrm{VIX}_t^{(m)})^2$ |
| **Capped** $\mathrm{VIX}^2$ | $\min\{K, \widetilde{c} + \widetilde{m}(\mathrm{VIX}_t^{(m)})^2\}$ |
| **Uncapped VIX** | $\widetilde{c} + \widetilde{m}\mathrm{VIX}_t^{(m)}$ |

The second layer CTMC approximation of Subsection 3.3.2 is then applied to the regime-switching diffusion (3.47) with the approximated fee processes listed in Table 3.18.

Using the CTMC technique outlined in Section 3.3; and the market, the variable annuity, and the CTMC parameters of Subsection 3.8.6.1, we performed the valuation of a variable annuity with and without early surrenders. The results are similar to the ones obtained under the Heston model and are summarized below, confirming again that fees tied to the volatility index do not significantly affect the value of surrender incentives.

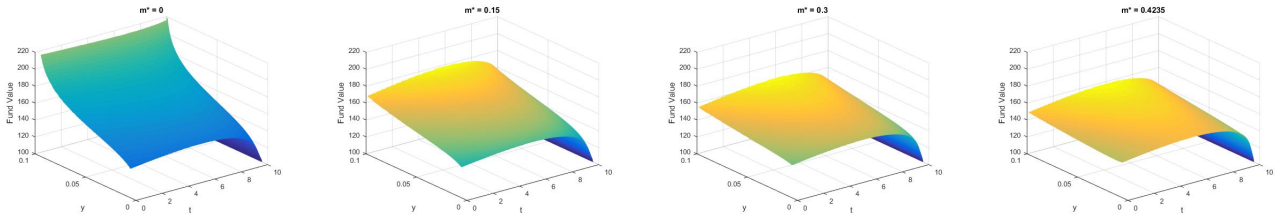
The value of early surrenders ("ES values") under the 3/2 model are presented in Table 3.19.

The approximated optimal surrender surfaces for the three fee structures under the 3/2 model are displayed in Figures 3.12, 3.13 and 3.14. Note that in order for the Figures under the Heston and 3/2 models to be comparable, the $y$-axis under the 3/2 model represents the variance of the fund, that is $\frac{1}{V_t}$, $0 \leq t \leq T$

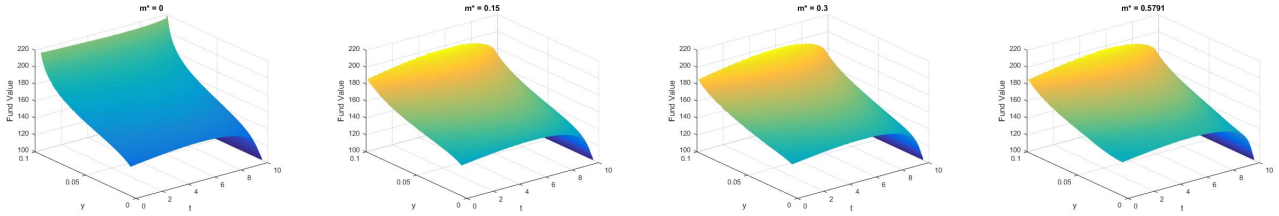
Table 3.19: Approximated early surrender values (ES values) under the 3/2 model.

| | | | | |
|---|---|---|---|---|
| | $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$ | | | |
| $\widetilde{m}^*$ | $0.0000$ | $0.1500$ | $0.3000$ | $0.4235$ |
| ES Value | $2.91799$ | $2.91126$ | $2.90875$ | $2.91078$ |
| | $c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}$, $K = 2\%$ | | | |
| $\widetilde{m}^*$ | $0.0000$ | $0.1500$ | $0.3000$ | $0.5791$ |
| ES Value | $2.91799$ | $2.90594$ | $2.89435$ | $2.88076$ |
| | $c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$ | | | |
| $\widetilde{m}^*$ | $0.0000$ | $0.0250$ | $0.0500$ | $0.0846$ |
| ES Value | $2.9180$ | $2.9120$ | $2.9068$ | $2.9012$ |

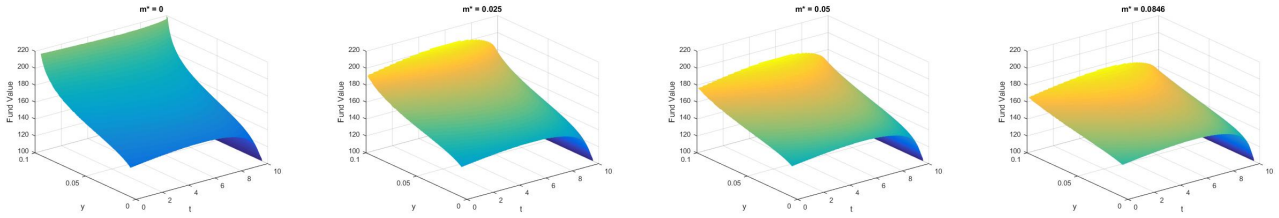
(recall that $V$ is the inverse variance in (3.45)). As noted previously, we observe that the optimal surrender surface is increasing with the volatility. However, when $\widetilde{m}^* = 0$, the surrender surface increases much faster than under the Heston model. We also note that VIX-linked fees help to neutralize the effect of the volatility on the optimal surrender decision, confirming again the relation existing between the fees and the optimal surrender strategies; that is, VA contracts with high fees are surrendered at lower fund values than VA contracts with low fees.




$$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$$

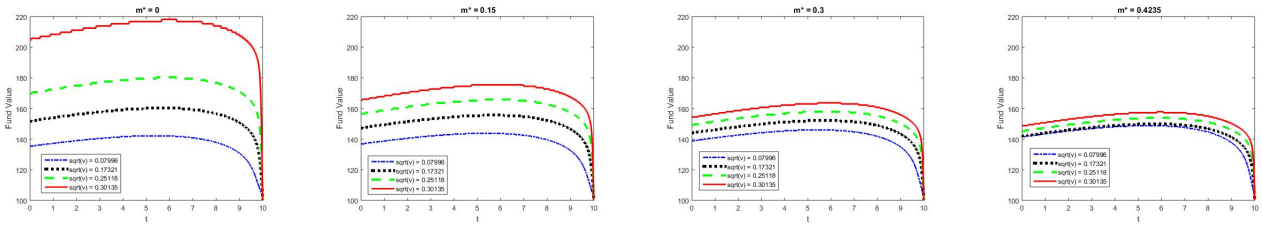



$$c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 2\%$$

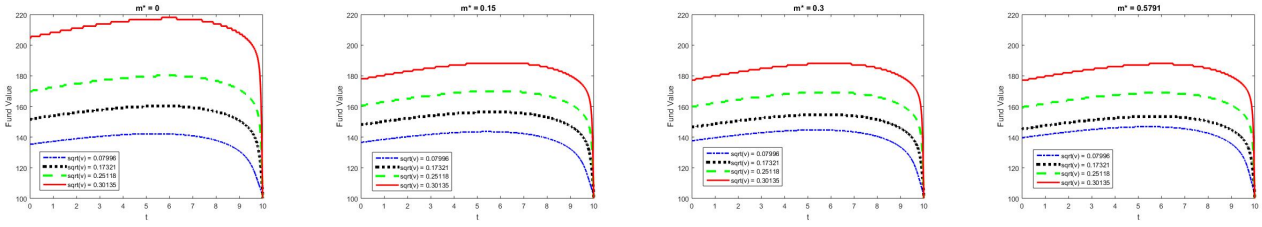



$$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$$

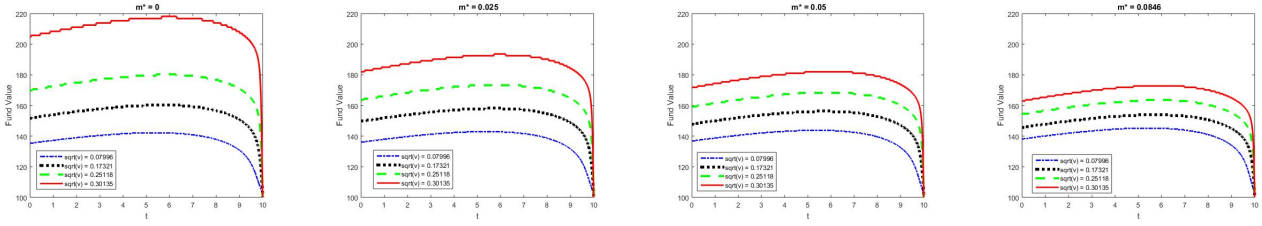
Figure 3.12: Approximated optimal surrender surface for different values of fair multiplier $\widetilde{m}^*$ under the 3/2 model where $x$-axis represents the time and the $y$-axis the variance.




$$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$$

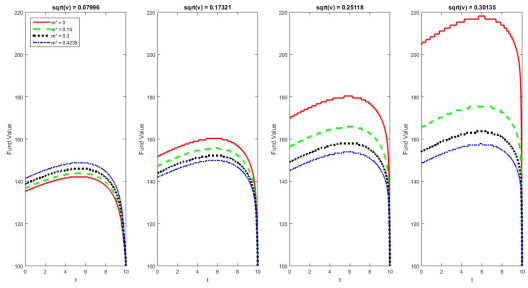



$$c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 2\%$$

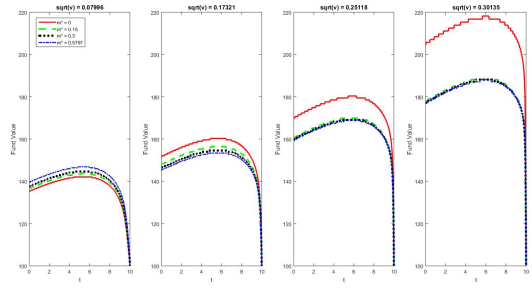



$$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$$

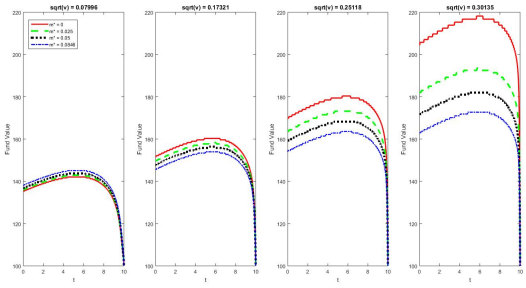
Figure 3.13: The $y$ section of the approximated optimal surrender surface, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.



$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2$



$c_t = \min\{\widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t^2, K\}, K = 2\%$



$c_t = \widetilde{c}^* + \widetilde{m}^* \mathrm{VIX}_t$

Figure 3.14: The $y$ section of the approximated optimal surrender surface, $f_y^{(m,N)}$, for different volatility levels $\sqrt{y}$ and fair multipliers $\widetilde{m}^*$.

# CHAPTER 4

# A UNIFYING APPROACH FOR THE PRICING OF DEBT SECURITIES

This chapter is based on the paper ”A unifying approach for the pricing of debt securities“, co-authored with Dr. Anne MacKay, published in *Quantitative Finance* (see Vachon and MacKay (2024)). It extends the findings from Chapters 2 and 3 to include time-inhomogeneous diffusion processes for pricing debt securities like bonds, bond options, callable/putable bonds, and convertible bonds. In particular, efficient algorithms are developed based on the technique discussed in Chapter 3, which naturally extends to stochastic interest rate models and time-inhomogeneous processes. Additionally, we show the existence of a trivial stopping time at the maturity of the contract for American-style CBs under specific conditions, employing reasoning akin to that presented in Chapter 2.

## 4.1 Introduction

This chapter proposes a unifying framework based on continuous-time Markov chain (CTMC) approximations to price debt securities under general time-inhomogenous short-rate models. Over the last few years, CTMC methods have garnered attention in option pricing literature, see for instance Cui *et al*. (2018), Ding and Ning (2021), and Kirkby (2023), among others. In particular, Cui *et al*. (2018) developed a two-layer CTMC technique to price European, barrier, Bermudian, Asian, and occupation time derivatives under general stochastic volatility models, while Ding and Ning (2021) discussed the extension of the method to time-inhomogeneous processes for the pricing of European and barrier options. Recently, Kirkby (2023) approximated time-homogeneous bi-dimensional diffusion processes to model short-rates and equity for the pricing of hybrid securities such as equity swap and cap via CTMC approximation. The framework outlined below is an extension to the work of Cui *et al*. (2018) to time-inhomogeneous processes for the pricing of debt securities such as bonds, bond options (loan commitments and deposit instruments), callable/putable bonds, and convertible bonds (CBs).

The advantage of CTMC approximations over other numerical techniques resides in their ability to easily adapt to various diffusion processes (homogeneous as well as inhomogeneous processes), and their extension to higher dimension models is also relatively straightforward Cui *et al*. (2018) (for two-dimension), Kirkby *et al*. (2020) (for higher dimension). More importantly, they generally allow for an explicit formulation of expectations (resp. conditional expectations), see, for instance, Cui *et al*. (2019), which facilitate the

pricing of European (resp. American)-type derivatives. In particular, the method allows for a closed-form matrix expression for the price of zero-coupon bonds regardless of the complexity of the short-rate dynamics selected, simplifying the calibration of the approximated model to the current market-term structure for a wide range of short-rate models. Calibration to the current market term structure is essential for practitioners since small deviations in the current short rates can result in significant differences in the value of the derivatives, see Brigo and Mercurio (2006). Moreover, the method generally exhibits a second-order convergence rate; see, for instance, Li and Zhang (2018) and Zhang and Li (2019) for a one-dimensional setting, and Ma *et al*. (2022) for a two-dimensional setting. In this chapter, we develop an easy and efficient algorithm to calibrate the approximated model to the current market term structure of interest rates under general one-dimensional time-inhomogeneous short-rate processes. A closed-form matrix expression is also obtained for the price of bond options (with coupons).

Debt securities often include embedded options such as call and put options, also known as redeemable and retractable bonds, respectively. A callable bond grants the issuer the right to pay back the bond at a predetermined price in the future (the call or the strike price). This type of provision protects the issuer and reduces the value of the bond. Another type of common embedded option in a bond is a put option. A putable bond grants the bondholder the right to sell back the bond to the issuer at a predetermined price (the put or the strike price). These types of options can usually be exercised at any time during a given exercise period (exercise window), which makes them similar to American options with time-dependent strikes. A bond can have both a call and put options embedded simultaneously. Because of the American style of the embedded options, the valuation of callable/putable debt does not admit a closed-form expression; thus, numerical procedures are necessary to price them.

Classical numerical techniques[1] for pricing interest rate derivatives include tree methods, numerical solutions to partial differential equations (PDEs), and Monte Carlo simulation, see for instance Brigo and Mercurio (2006), Section 3.11. Trees are particularly interesting for valuing American-type derivatives because the continuation value can be calculated explicitly at each node of the tree. However, building trees for time-inhomogeneous mean reverting diffusion processes, such as those used for short-rate dynamics, is not always straightforward, and extension to higher dimensions may be difficult. The construction of the

[1] In this chapter, we do not consider advance notices, that is, exercise decisions prior to exercise benefits, which complexify the problem significantly, see for instance Büttler and Waldvogel (1996), D'Halluin *et al*. (2001), Ben-Ameur *et al*. (2007) and Ding *et al*. (2012) for numerical techniques in that particular context.

tree depends on the diffusion model selected (see among others, Ho and Lee (1986), Black *et al.* (1990), Black and Karasinski (1991), Hull and White (1994), Hull and White (1996), Mercurio and Moraleda (2001) and Brigo and Mercurio (2006), Appendix F, for a general formulation), and extra care is required for the transition probabilities to stay positive. Fitting the approximated model to the current market term structure can also be tricky Hull and White (1994), Hull and White (1996). Trees can also make options with path-dependent payoffs, such as Asian options, difficult to value. PDE approaches encounter similar challenges to trees in terms of flexibility in the modelization of the underlying diffusion process. They also make path-dependent payoff valuation challenging, and techniques are generally less intuitive than trees; see, for instance, Duffy (2006) for a general review of the approach for derivatives pricing. In contrast, simulation methods can be implemented for various diffusion models and adapt easily to higher dimensions; however, this type of procedure is usually less efficient computationally than those based on trees or PDEs because the continuation value cannot be calculated explicitly and needs to be approximated. Different techniques are proposed in the literature to price American-type derivatives using simulation methods; see, for instance, Fu *et al.* (2001) for a comparison of these approaches, and Glasserman (2003), Chapter 8, for a review.

In this chapter, a simple and efficient algorithm for pricing callable/putable debt under general one-dimensional time-inhomogeneous short-rate processes is developed. Our methodology overcomes several of the drawbacks of other classical methods. In particular, it is intuitive (similar to tree methods) and adapts to a wide range of diffusion processes (homogeneous as well as inhomogeneous processes). Moreover, because conditional expectations have closed-form matrix expressions, the continuation value can be calculated explicitly, making the approximation extremely efficient and accurate for the pricing of American-type derivatives. Using the methodology of Cui *et al.* (2018), described in Section 4.3.2 of this chapter, the extension of these procedures to two-factor short rate models, see for instance Brigo and Mercurio (2006), Chapter 4, is also straightforward.

Next, the pricing of CBs is considered. CBs are hybrid securities that possess features of both debt and equity. They are similar to bonds except that the investor has the right to convert the bond for a predetermined number of shares, known as the conversion ratio, of the issuing company during a certain exercise window prior to maturity. At maturity, if conversion is allowed and the bond has not been converted to shares, the holder has the right to convert the bond or receive its face value. In practice, additional features such as call and put options are also generally embedded in CBs, so numerical procedures are required to

value these securities.

Over the last 40 years, CB pricing has been studied extensively in the literature. Under the standard Black–Scholes setting, Ingersoll Jr. (1977) proposes a structural approach under which the firm value is the underlying state variable. Then follows the work of Brennan and Schwartz (1977), in which finite difference methods are used to solve a PDE. The structural approach has multiple drawbacks, Batten *et al.* (2014). In particular, since the firm value is not a tradable asset, the calibration of the model may be challenging. Thus, McConnell and Schwartz (1986) propose a reduced-form approach under which the issuing company stock price is the underlying state variable. Working in the Black–Scholes setting with a constant risk-free rate and volatility, McConnell and Schwartz (1986) compensate for the credit risk by adding a constant (the credit spread) over the risk-free rate to discount the cash flows. Since then, multiple authors have incorporated credit risk into the valuation framework adequately. Tsiveriotis and Fernandes (1998) split the bond value into two components: equity and debt. The equity part (when the debt is converted to stock) is discounted at a risk-free rate, whereas the debt part is discounted at a risky rate, where the risky rate can be deduced from the market-observed credit spreads. This approach to model credit risk is still widely used among practitioners for its simplicity and ability to incorporate the main feature of CBs with limited market information, Gushchin and Curien (2008). Following the approach of Jarrow and Turnbull (1995) to model the credit risk, Hung and Wang (2002) and Chambers and Lu (2007) use a binomial tree method and incorporate stochastic risk-free rates in the valuation model, whereas Ayache *et al.* (2003) and Milanov *et al.* (2013) incorporate default risk by modeling the stock price by a jump-diffusion process in a constant risk-free rate environment.

Diverse numerical methods have been proposed to value convertible debt, ranging from the classical tree methods (Hung and Wang (2002),Chambers and Lu (2007), Milanov *et al.* (2013), among others), to finite-difference and finite element approaches (Tsiveriotis and Fernandes (1998), Ayache *et al.* (2003), Barone-Adesi *et al.* (2003), among others) and simulation (Ammann *et al.* (2008), Batten *et al.* (2018)). Recently, Lu and Xu (2017) and Ma *et al.* (2020) developed a two-factor willow-tree approach to price CBs under stochastic interest rates and used the approach of Jarrow and Turnbull (1995) as in Hung and Wang (2002) and Chambers and Lu (2007) to include credit risk. On the other hand, Lin and Zhu (2020) propose a predictor-corrector scheme to solve a PDE under stochastic volatility or interest rate models, whereas Lin and Zhu (2022) use an integral approach under the Black–Scholes setting. Both Lin and Zhu (2020) and Lin and Zhu (2022) ignored credit risk in their valuation framework. This chapter considers the CB pricing problem

under general bi-dimensional time-inhomogeneous diffusion processes, where equity and risk-free rates are the two risk factors. Default/credit risk is also included in the valuation model using the approach of Tsiveriotis and Fernandes (1998). An efficient algorithm to approximate the value of CBs using a two-layer continuous-time Markov chain approximation is developed. When conversion is only permitted at maturity (European-style or European CBs), a closed-form matrix expression is obtained. Numerical experiments reveal that the method is highly efficient and accurate. The advantage of the CTMC approximation over other classical methods resides in its ability to adapt to a wide range of time-inhomogeneous bi-dimensional models while preserving the simplicity of one-dimensional valuation models. The method is intuitive (similar to trees), and it is worth reiterating that it allows for the perfect fit of the current market term structure in a straightforward manner, regardless of the short-rate diffusion process selected.

The CB pricing problem is also studied from a theoretical perspective. In particular, when there is no credit risk, no dividends and other features such as call and put options are ignored, we show that early conversion is sub-optimal such that the problem is reduced to the pricing of a European-style CB. This result also holds for coupon-bearing convertible debt. On the other hand, when credit risk is considered, we show that the value of American-style CBs[2] is bounded from below and above by those of European–style CBs with and without credit risk, respectively.

Finally, numerical experiments demonstrate the high level of accuracy of CTMC methods across a large range of model parameters and short-rate models. The efficiency and numerical convergence of the CTMC methodology in pricing debt securities are also studied empirically, and theoretical convergence is discussed.

The main contributions of this chapter are as follows:

1. This chapter extends the results of Cui *et al*. (2018) to a time-inhomogeneous framework for the pricing of debt securities, such as callable, putable, and convertible bonds.
2. A closed-form matrix expression is obtained to approximate the price of bonds under general time-inhomogeneous short-rate processes. The availability of a closed-form expression to approximate the price of zero-coupon bonds, regardless of the complexity of the short-rate dynamics selected, makes

[2] The term American-style CB (or just CB) is used to refer to a bond under which the conversion option can be exercised at any time prior to maturity.

the method attractive for practitioners since it allows to perfectly calibrate the approximated model to the current market term structure.

3. A closed-form matrix expression is also obtained for the price of bond (with or without coupons) options under general time-inhomogeneous short-rate processes, providing an alternative approximation formula to Kirkby (2023), Proposition 12, which involved an integral to be solved numerically.

4. Efficient procedures are developed to approximate the price of convertible debt under general bi-dimensional time-inhomogeneous diffusion processes, and a closed-form matrix expression is obtained for the price of European-style CBs.

5. The pricing of convertible bonds is also considered from a theoretical perspective. When there is no credit risk and no dividend yield, we show that early conversion is sub-optimal. When credit risk is considered, lower and upper bounds are obtained.

The remainder of the chapter is organized as follows. In Section 4.2, we introduce the market model. A brief introduction to CTMC approximation methods for two-dimensional time-inhomogeneous diffusion processes is provided in Section 4.3. In Section 4.4, CTMC methods are used to approximate the price of bonds, bond options, and debt securities such as callable and putable bonds under general one-dimensional time-inhomogeneous short-rate processes. Section 4.5 discusses the application of CTMC approximation to convertible debt valuation under bi-dimensional time-inhomogeneous diffusion processes. Section 4.6 provides numerical results and shows the high efficiency of CTMC methods over other common numerical techniques. Section 4.7 concludes the chapter.

## 4.2 Financial Setting

### 4.2.1 Market Model

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, where $\mathbb{F}$ denotes a complete and right-continuous filtration and where $\mathbb{Q}$ denotes the pricing measure for our market. We consider a stochastic short-rate process $R$ correlated with the price of a risky asset (or stock) $S$, which can be described by a two-dimensional process $(S, R) = \{(S_t, R_t)\}_{t\geq 0}$ satisfying

$$\begin{aligned} \mathrm{d}S_t &= (R_t - q_t)S_t\,\mathrm{d}t + \sigma_S(R_t)S_t\,\mathrm{d}W_t^{(1)}, \\ \mathrm{d}R_t &= \mu_R(t, R_t)\,\mathrm{d}t + \sigma_R(R_t)\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{4.1}$$

with $S_0 > 0$ and $R_0 \in \mathcal{S}_R$, where $\mathcal{S}_R$ denotes the state-space of $R$ (generally $\mathbb{R}$ or $\mathbb{R}_+$ depending on the model, see Tables 4.1 and 4.2 for details), $q : \mathbb{R}_+ \to [0,1]$ is a continuous function representing a time-deterministic dividend yield, and $W = \{(W_t^{(1)}, W_t^{(2)})\}_{t\geq 0}$ is a two-dimensional correlated Brownian motion with cross-variation $[W^{(1)}, W^{(2)}]_t = \rho t$, with $\rho \in [-1,1]$. We assume that $\mu_R : \mathbb{R}_+ \times \mathcal{S}_R \to \mathbb{R}$ is continuous and that $\sigma_R, \sigma_S : \mathcal{S}_R \to \mathbb{R}_+$ are continuously differentiable with $\sigma_R(\cdot), \sigma_S(\cdot) > 0$ on $\mathcal{S}_R$. Further, we suppose that $\mu_R$, $\sigma_R$ and $\sigma_S$ are defined such that (4.1) has a unique-in-law weak solution.

The function $\sigma_S$ is often set to a constant $\sigma_S(\cdot) = \sigma > 0$ ( Lu and Xu (2017), Ma *et al.* (2020), Kirkby (2023)), such that the stock price follows a geometric Brownian motion with stochastic interest rate. A list of common short-rate models is provided in Tables 4.1, 4.2, and 4.3.

Table 4.1: Example of time-homogeneous short-rate models

| **Time-Homogeneous Models** | | |
|---|---|---|
| **Model** | **Dynamics** | **Parameters** |
| Vasicek, Vasicek (1977) | $\mathrm{d}R_t = \kappa(\theta - R_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$ | $\kappa, \theta, \sigma > 0$, $R_0 \in \mathbb{R}$ |
| Cox–Ingersoll–Ross (CIR), Cox *et al.* (1985) | $\mathrm{d}R_t = \kappa(\theta - R_t)\,\mathrm{d}t + \sigma\sqrt{R_t}\,\mathrm{d}W_t$ | $\kappa, \theta, \sigma, R_0 > 0$, with $2\kappa\theta > \sigma^2$ |
| Dothan, Dothan (1978) | $\mathrm{d}R_t = \kappa R_t\,\mathrm{d}t + \sigma R_t\,\mathrm{d}W_t$ | $\sigma, \kappa \in \mathbb{R}$, $R_0 > 0$ |
| Exponential Vasicek (EV) Brigo and Mercurio (2006), Section 3.2.5 | $\mathrm{d}R_t = R_t(\eta - \alpha \ln R_t)\,\mathrm{d}t + \sigma R_t\,\mathrm{d}W_t$ | $\eta, \alpha, \sigma, R_0 > 0$ |

The time-homogeneous models listed in Table 4.1 are popular because of their analytical tractability. However, they are less used by practitioners because they cannot adequately replicate the term structure of interest rates. Indeed, to be able to capture the discount curve appropriately, models need to have at least one time-dependent parameter. This important feature of interest rate dynamics gives rise to the time-inhomogeneous models; see Table 4.2 for examples. In these models, the yield curve is provided exogenously (as an input to the model). The extended models of Brigo and Mercurio (2006), listed in Table

[3] The short rates are positive under some conditions on function $\theta$.

Table 4.2: Example of time-inhomogeneous short-rate models

| Time-Inhomogeneous Models | | |
|---|---|---|
| **Model** | **Dynamics** | **Parameters** |
| Ho–Lee (HL), Ho and Lee (1986) | $\mathrm{d}R_t = \theta(t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$ | $\sigma > 0$, $R_0 \in \mathbb{R}$ |
| Black–Derman–Toy (BDT), Black *et al.* (1990) | $\mathrm{d}R_t = \theta(t)R_t\,\mathrm{d}t + \sigma R_t\,\mathrm{d}W_t$ | $\sigma > 0$, $R_0 > 0$ |
| Hull–White (HW), Hull and White (1990) | $\mathrm{d}R_t = (\theta(t) - \kappa R_t)\ \mathrm{d}t + \sigma\,\mathrm{d}W_t$ | $\kappa, \sigma > 0$, and $R_0 \in \mathbb{R}$ |
| Black–Karasinski (BK), Black and Karasinski (1991) | $\mathrm{d}R_t = R_t\left(\theta(t) - \kappa \ln R_t\right)\ \mathrm{d}t + \sigma R_t\,\mathrm{d}W_t$ | $\kappa, \sigma, R_0 > 0$ |
| Mercurio and Moraleda (MM), Mercurio and Moraleda (2001) | $\mathrm{d}R_t = R_t\left[\theta(t) - \left(\lambda - \frac{\gamma}{1+\gamma t}\ln R_t\right)\right]\ \mathrm{d}t + \sigma R_t\,\mathrm{d}W_t$ | $\lambda, \gamma \in \mathbb{R}_+$, and $\sigma, R_0 > 0$ |
| Extended CIR (CIR+), Hull and White (1990) | $\mathrm{d}R_t = \theta(t) - \kappa R_t\,\mathrm{d}t + \sigma\sqrt{R_t}\,\mathrm{d}W_t$ | $\kappa, \sigma, R_0 > 0$[3] |

4.3, allow fitting of the initial term structure and reproduce important stylized facts while preserving the analytical tractability of the model through the auxiliary homogeneous process $Y$.

In Tables 4.2 and 4.3, the initial term structure of interest rates is captured through the time deterministic function $\theta$. When using CTMC approximation, an easy recursive procedure is used to find the function $\theta$ that makes the approximated models fit the initial discount curve perfectly. This will be discussed further in Section 4.4.4.

Table 4.3: Extended time-homogeneous models of Brigo and Mercurio (2006), Section 3.8

| **Time-Inhomogeneous Models** | | |
|---|---|---|
| **Model** | **Dynamics** | **Parameters** |
| Extended Vasicek, (EV+) | $\mathrm{d}Y_t = \kappa(\alpha - Y_t)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t$ | $\kappa, \alpha, \sigma > 0,$ |
| Brigo and Mercurio (2006), Section 3.8.4 | $R_t = Y_t + \theta(t)$ | $Y_0, R_0 \in \mathbb{R}$ |
| Extended CIR, (CIR++) | $\mathrm{d}Y_t = \kappa(\alpha - Y_t)\,\mathrm{d}t + \sigma\sqrt{Y_t}\,\mathrm{d}W_t$ | $\kappa, \alpha, \sigma > 0,$ |
| Brigo and Mercurio (2006), Section 3.9 | $R_t = Y_t + \theta(t)$ | $Y_0, R_0 > 0$[3] |
| Extended Exponential Vasicek, (EEV+) | $\mathrm{d}Y_t = Y_t(\eta - \alpha \ln Y_t)\,\mathrm{d}t + \sigma Y_t\,\mathrm{d}W_t$ | $\eta, \alpha, \sigma \in \mathbb{R},$ |
| Brigo and Mercurio (2006), Section 3.8 | $R_t = Y_t + \theta(t)$ | $Y_0, R_0 > 0$[3] |

## 4.3 Continuous-Time Markov Chain Approximation of Nonhomogeneous Processes

The CTMC framework outlined in this section was first proposed by Cui *et al.* (2018) for the pricing of exotic equity options under general stochastic local volatility models. Subsequently, the technique has been extended to time-inhomogeneous processes by Ding and Ning (2021).

### 4.3.1 Approximation of the Short-Rate Process $\{R_t\}_{t\geq 0}$

The objective is to construct a continuous-time Markov chain $\{R_t^{(m)}\}_{t\geq 0}$ taking values on a finite state-space $\mathcal{S}_R^{(m)} := \{r_1, r_2, \dots r_m\}$, with $r_i \in \mathcal{S}_R$, $m \in \mathbb{N}$ and a time-dependent generator $\widetilde{\mathbf{Q}}^{(m)}(t) = [\widetilde{q}_{ij}(t)]_{m\times m}$ that converges weakly to $\{R_t\}_{t\geq 0}$ as $N, m \to \infty$. To denote the weak convergence of $R^{(m)}$ to $R$, we write $R^{(m)} \Rightarrow R$.

The first step is to approximate the state-space of the short-rate process. Several approaches are available in the literature to construct the finite state-space $\mathcal{S}_R^{(m)}$ of $R^{(m)}$, from simple uniform to non-uniform grids (see for instance Tavella and Randall (2000), Mijatović and Pistorius (2013), and Lo and Skindilias (2014) for examples of non-uniform grids). The specific grid choice for this work is discussed in greater detail in Section 4.6.

The next step is to construct the time-dependent generator $\widetilde{\mathbf{Q}}^{(m)}(t)$. For analytical tractability, we suppose further that $\widetilde{\mathbf{Q}}^{(m)}(t)$ is piecewise constant in time[4], that is,

$$\widetilde{\mathbf{Q}}^{(m)}(t) = \sum_{n=1}^{N} \mathbf{Q}_n^{(m)} \mathbf{1}_{[t_{n-1},t_n)}(t). \tag{4.2}$$

for some time partition $0 = t_0 < t_1 < \ldots < t_N = T$ of $[0, T]$, with $T > 0$, $t_n := n\Delta_N$ and $\Delta_N = T/N$, and where $\mathbf{Q}_n^{(m)} = [q_{ij}^{(n)}]_{m\times m}$ denotes the generator on the time interval $[t_{n-1}, t_n)$, whose elements $q_{ij}^{(n)}$, $1 \leq i, j \leq m$, satisfy $q_{ij}^{(n)} \geq 0$ when $i \neq j$, and $q_{ij}^{(n)} \leq 0$ when $i = j$, $n = 1, 2, \ldots N$. Under this assumption, the transition probability matrix $\mathbf{P}(s, t)$ from time $s$ to $t$ has the following matrix representation

$$\mathbf{P}(s,t) = e^{\mathbf{Q}_{i+1}^{(m)}(t_{i+1}-s)} e^{\mathbf{Q}_{i+2}^{(m)}(t_{i+2}-t_{i+1})} \cdots e^{\mathbf{Q}_{j+1}^{(m)}(t-t_j)}, \quad t_i \leq s < t_{i+1},\ t_j \leq t < t_{j+1},\ s < t, \tag{4.3}$$

where

$$\exp\{\mathbf{Q}_n^{(m)} t\} = \sum_{k=0}^{\infty} \frac{(\mathbf{Q}_n^{(m)} t)^k}{k!}, \quad 0 \leq t \leq T, \tag{4.4}$$

see Rindos *et al.* (1995), p.123–124 for details.

Following the work of Lo and Skindilias (2014) and Ding and Ning (2021), the generator $\mathbf{Q}_n^{(m)} = [q_{ij}^{(n)}]_{m\times m}$ on the time interval $[t_{n-1}, t_n)$ is constructed as follows

$$q_{ij}^{(n)} = \begin{cases} \frac{\sigma_R^2(r_i) - \delta_i \mu_R(t_{n-1}, r_i)}{\delta_{i-1}(\delta_{i-1}+\delta_i)}, & j = i-1, \\ -q_{i,i-1}^{(n)} - q_{i,i+1}^{(n)}, & j = i, \\ \frac{\sigma_R^2(r_i) + \delta_{i-1} \mu_R(t_{n-1}, r_i)}{\delta_i(\delta_{i-1}+\delta_i)}, & j = i+1, \\ 0, & j \neq i,\ i-1,\ i+1, \end{cases} \tag{4.5}$$

for $n \in 1, 2, \ldots, N$, $2 \leq i \leq m-1$, $1 \leq j \leq m$, and where $\delta_i = r_{i+1} - r_i$, $i = 1, 2, \ldots, m-1$. On the borders, we set $q_{12}^{(n)} = \frac{|\mu_R(t_{n-1}, r_1)|}{\delta_1}$, $q_{11}^{(n)} = -q_{12}^{(n)}$, $q_{m,m-1}^{(n)} = \frac{|\mu_R(t_{n-1}, r_m)|}{\delta_{m-1}}$, $q_{m,m}^{(n)} = -q_{m,m-1}^{(n)}$, and $0$ elsewhere. At endpoints, other schemes could have also been employed, see Chourdakis (2004) and Mijatović and Pistorius (2013) for examples. However, we note that these schemes are equivalent numerically.

**Remark 4.3.1 (Weak convergence of the approximation)** *Such a construction ensures that the process $R^{(m)}$ converges weakly to $R$ as $N, m \to \infty$, see Mijatović and Pistorius (2009), Section 4 and Corollary 2 for details, and Ding and Ning (2021), Section 2.1.*

[4] This assumption ensures that the transition probability matrix has a simple expression in terms of its generator, see Rindos *et al.* (1995) Equation (8.4) for the general formulation.

### 4.3.2 Approximation of the Stock Process $\{S_t\}_{t\geq 0}$

The idea behind the CTMC approximation of two-dimensional processes is similar to the approximation of one-dimensional processes. The first step is to replace the CTMC approximation of the short-rate process $R^{(m)}$ into an auxiliary process (4.6) obtained from a transformation of (4.1). This results in a regime-switching diffusion process. The regime-switching diffusion process is then approximated by a regime-switching CTMC. The final step and the key to the approximation consists of mapping the two-dimensional regime-switching CTMC onto a one-dimensional CTMC defined on an enlarged state-space. Thus, working with an approximation of two-dimensional processes is similar to working with an approximation of one-dimensional processes.

The next lemma allows the removal of the correlation between the two Brownian motions in (4.1).

**Lemma 4.1 (Cui *et al*. (2018), Lemma 1)** *Let $(S, R)$ be defined as in* (4.1)*. Define $f(r) := \int_\cdot^r \frac{\sigma_S(u)}{\sigma_R(u)}\,\mathrm{d}u$, and $X_t := \ln(S_t) - \rho f(R_t)$ for $t \geq 0$. Then, $X$ satisfies*

$$\begin{aligned} \mathrm{d}X_t &= \mu_X(t, R_t)\,\mathrm{d}t + \sigma_X(R_t)\,\mathrm{d}W_t^\star \\ \mathrm{d}R_t &= \mu_R(t, R_t)\,\mathrm{d}t + \sigma_R(R_t)\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{4.6}$$

*where $W^\star := \frac{W_t^{(1)} - \rho W_t^{(2)}}{\sqrt{1-\rho^2}}$ denotes a standard Brownian motion independent of $W^{(2)}$, $\sigma_X(r) := \sigma_S(r)\sqrt{1-\rho^2}$ and $\mu_X(t,r) := r - q_t - \frac{\sigma_S^2(r)}{2} - \rho\psi(t,r)$, and*

$$\psi(t,r) := \mu_R(t,r)\frac{\sigma_S(r)}{\sigma_R(r)} + \frac{1}{2}\left[\sigma_S'(r)\sigma_R(r) - \sigma_S(r)\sigma_R'(r)\right],$$

*for $r \in \mathcal{S}_R$.*

By replacing the short-rate process in (4.6) by its CTMC approximation $R^{(m)}$, we obtain the following regime-switching diffusion process $\{X_t^{(m)}\}_{t\geq 0}$ satisfying

$$\mathrm{d}X_t^{(m)} = \mu_X(t, R_t^{(m)})\,\mathrm{d}t + \sigma_X(R_t^{(m)})\,\mathrm{d}W_t^\star, \tag{4.7}$$

where the regimes correspond to the state of the approximated short rate process, $\mathcal{S}_R^{(m)} := \{r_1, r_2, \ldots, r_m\}$.

The regime-switching diffusion process $(X^{(m)}, R^{(m)})$ is then approximated by a regime-switching CTMC $(X^{(m,M)}, R^{(m)})$. This is done by fixing a state for the short-rate process $R^{(m)}$ and then constructing a CTMC approximation for $X^{(m)}$ given $R^{(m)}$ is in that state. For this step, the same procedure as that described in Section 4.3.1 can be used. More precisely, let $X^{(m,M)}$ be the CTMC approximation of $X^{(m)}$ taking values on a finite state-space $\mathcal{S}_X := \{x_1, x_2, \ldots, x_M\}$, $M \in \mathbb{N}$.

For each $r_k \in \mathcal{S}_R^{(m)}$, we define the time-dependant generator $\widetilde{\mathbf{\Lambda}}_k^{(N,M)}(t) := \sum_{n=1}^{N} \mathbf{\Lambda}_k^{(n,M)} \mathbf{1}_{[t_{n-1},t_n)}(t)$, where $\mathbf{\Lambda}_k^{(n,M)} = [\lambda_{ij}^{(n,k)}]_{M\times M}$ and

$$\lambda_{ij}^{(n,k)} = \begin{cases} \frac{\sigma_X^2(r_k) - \delta_i^X \mu_X(t_{n-1},r_k)}{\delta_{i-1}^X(\delta_{i-1}^X + \delta_i^X)} & j = i-1 \\ -\lambda_{i,i-1}^{(n,k)} - \lambda_{i,i+1}^{(n,k)} & j = i \\ \frac{\sigma_X^2(r_k) + \delta_{i-1}^X \mu_X(t_{n-1},r_k)}{\delta_i^X(\delta_{i-1}^X + \delta_i^X)} & j = i+1 \\ 0 & j \neq i,\, i-1,\, i+1, \end{cases} \tag{4.8}$$

for $2 \leq i \leq M-1$, $1 \leq j \leq M$, where $\delta_i^X = x_{i+1} - x_i$, $i = 1, 2, \ldots, M-1$. On the boundaries, we set $\lambda_{12}^{(n,k)} = \frac{|\mu_X(t_{n-1},r_k)|}{\delta_1^X}$, $\lambda_{11}^{(n,k)} = -\lambda_{12}^{(n,k)}$, $\lambda_{M,M-1}^{(n,k)} = \frac{|\mu_X(t_{n-1},r_k)|}{\delta_{M-1}^X}$, $\lambda_{M,M}^{(n,k)} = -\lambda_{M,M-1}^{(n,k)}$, and $0$ elsewhere.

Using the regime-switching approximation of $(X, R)$ and the relation between $X$ and $S$ provided in Lemma 4.1, the CTMC approximation of the stock process $S^{(m,M)}$ is defined by

$$S_t^{(m,M)} := \exp\left\{X_t^{(m,M)} + \rho f(R_t^{(m)})\right\}, \quad t \geq 0. \tag{4.9}$$

The final step consists in mapping the two-dimensional regime-switching CTMC onto a one-dimensional CTMC process $Z^{(mM)}$ on an enlarged state-space $\mathcal{S}_Z^{(mM)} := \{1, 2, \ldots, mM\}$. This is done in Proposition 4.3.1, reproduced from Proposition 2 of Ding and Ning (2021).

**Proposition 4.3.1** *[Proposition 2 of Ding and Ning (2021)]*

*Consider a regime-switching CTMC $(X^{(m,M)}, R^{(m)})$ taking values in $\mathcal{S}_X^{(M)} \times \mathcal{S}_R^{(m)}$, where $\mathcal{S}_X^{(M)} = \{x_1, x_2, \ldots, x_M\}$ and $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$, and another one-dimensional CTMC, $\{Z_t^{(mM)}\}_{t\geq 0}$, taking values in $\mathcal{S}_Z^{(mM)} := \{1, 2, \ldots, mN\}$ and its time-dependent generator defined by $\widetilde{\mathbf{G}}^{(mM)}(t) := \sum_{n=1}^{N} \mathbf{G}_n^{(mM)} \mathbf{1}_{[t_{n-1},t_n)}(t)$, where*

$$\mathbf{G}_n^{(mM)} := \begin{pmatrix} q_{11}^{(n)}\mathbf{I}_M + \mathbf{\Lambda}_1^{(n,M)} & q_{12}^{(n)}\mathbf{I}_M & \cdots & q_{1m}^{(n)}\mathbf{I}_M \\ q_{21}^{(n)}\mathbf{I}_M & q_{22}^{(n)}\mathbf{I}_M + \mathbf{\Lambda}_2^{(n,M)} & \cdots & q_{2m}^{(n)}\mathbf{I}_M \\ \vdots & \vdots & \ddots & \vdots \\ q_{m1}^{(n)}\mathbf{I}_M & q_{m2}^{(n)}\mathbf{I}_M & \cdots & q_{mm}^{(n)}\mathbf{I}_M + \mathbf{\Lambda}_m^{(n,M)} \end{pmatrix}, \tag{4.10}$$

*where* $\mathbf{I}_M$ *denotes the* $M \times M$ *identity matrix,* $\mathbf{Q}_n^{(m)} = [q_{ij}^{(n)}]_{m\times m}$ *and* $\mathbf{\Lambda}_k^{(n,M)} = [\lambda_{ij}^{(n,k)}]_{M\times M}$, $k = 1, 2, \ldots, m$, $n = 1, 2, \ldots, N$ *denote the generators defined in* (4.5) *and* (4.8)*, respectively. Define the function* $\psi : \mathcal{S}_X^{(M)} \times \mathcal{S}_R^{(m)} \to \mathcal{S}_Z^{(mM)}$ *by* $\psi(x_l, r_k) = (k-1)M + l$ *and its inverse* $\psi^{-1} : \mathcal{S}_Z^{(mM)} \to \mathcal{S}_X^{(M)} \times \mathcal{S}_R^{(m)}$ *by* $\psi^{-1}(n_z) = (x_l, r_k)$ *for* $n_z \in \mathcal{S}_Z^{(mM)}$*, and* $k = \lceil n_z/M \rceil$, $l = n_z - (k-1)M$*, where* $\lceil x \rceil$ *denotes the largest integer less than* $x$*. Then, we have*

$$\mathbb{E}\left[\Psi(X^{(m,N)}, R^{(m)})\middle| X_0^{(m,N)} = x_i, R_0^{(m)} = r_j\right] = \mathbb{E}\left[\Psi(\psi^{-1}(Z^{(mN)}))\middle| Z_0^{(mN)} = (j-1)M + i\right],$$

*for any path-dependent function* $\Psi$ *such that the expectation on the left-hand side is finite.*

**Remark 4.3.2 (Weak convergence of the approximation)** *Such a construction ensures that* $(X^{(m,M)}, R^{(m)}) \Rightarrow (X, R)$ *and* $S^{(m,M)} \Rightarrow S$ *as* $N, m, M \to \infty$*, see Ding and Ning (2021), Proposition 3, for details.*

## 4.4 Application to the Pricing of Interest Rate Securities

This section provides closed-form matrix expressions for the prices of zero-coupon bonds and European bond options. We also develop an efficient recursive procedure for the pricing of American-type financial instruments such as callable and putable bonds. Calibration to the current term structure of interest rates is also discussed.

Let $0 = t_0 < t_1 < \ldots < t_N = T$ be a time partition of $[0, T]$, where $T > 0$ denotes the maturity of the financial instrument, $t_n = n\Delta_N$, $n = 0, 1, 2, \ldots, N$, and $\Delta_N = T/N$, $N \in \mathbb{N}$. Recall that $R^{(m)}$ is the CTMC approximation of $R$ taking values in a finite state-space $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$, $m \in \mathbb{N}$, and its time dependent generator, $\widetilde{\mathbf{Q}}^{(m)}(t) = \sum_{n=1}^N \mathbf{Q}_n^{(m)} \mathbf{1}_{[t_{n-1}, t_n)}(t)$, is defined in (4.5).

Throughout this section, we denote by $\{\mathbf{e}_k\}_{k=1}^m$ the standard basis in $\mathbb{R}^m$, that is, $\mathbf{e}_k$ represents a row vector of size $1 \times m$ with a value of $1$ in the $k$-th entry and $0$ elsewhere, $\mathbf{1}_{m\times 1}$ denotes an $m \times 1$ unit vector, and

$\mathbf{D}_m := \text{diag}(\boldsymbol{r})$ denotes an $m \times m$ diagonal matrix with the vector $\boldsymbol{r} = (r_1, r_2, \ldots, r_m)$ on its diagonal.

The following results often require Assumption 4.4.1 to hold.

**Assumption 4.4.1** *There exist a $r^\star \in \mathbb{R}$ such that $R_t \geq r^\star$ for all $t \geq 0$.*

This assumption restricts the state-space of the short-rate process for the discount factor to be bounded. This allows the use of some convergence theorems. Note that Vasicek, Ho–Lee, Hull–White, and EV+ models, listed in Tables 4.1, 4.2, or 4.3, do not satisfy this condition. However, numerical results in Section 4.6 and Appendix 4.12.2 show that the theoretical results of this section still hold for the Hull–White and Vasicek models, suggesting that Assumption 4.4.1 can be relaxed under a certain set of parameters. In that case, theoretical results must be shown on a case-by-case basis for each particular model.

### 4.4.1 Zero-Coupon Bond

The CTMC approximation of zero-coupon bond prices has been previously examined in the literature for time-homogeneous short-rate processes; see, for instance, Kirkby (2023), Proposition 3. In this section, we extend these findings to time-inhomogeneous processes. The first result, presented in Lemma 4.2, concerns the Laplace transform of some additive functions, extending Proposition 8 of Cui *et al.* (2018) to time-inhomogeneous processes. This result will be used thereafter to obtain a closed-form matrix expression for the price of zero-coupon bonds.

**Lemma 4.2** *Consider $0 = \widetilde{t}_0 < \widetilde{t}_1 < \ldots < \widetilde{t}_{\widetilde{N}} = T$ a partition of $[0, T]$, with $\widetilde{N} = kN$ for some $k \in \mathbb{N}$, $\Delta_{\widetilde{N}} = T/\widetilde{N}$ and $\widetilde{t}_n = n\Delta_{\widetilde{N}}$, and let $R^{(m)}_{t_i} = R_{t_i} = r_j \in \mathcal{S}^{(m)}_R$, for some $i \in \{0, 1, 2, \ldots, N\}$. It holds that*

$$\mathbb{E}\left[e^{-\sum_{n=ki+1}^{\widetilde{N}} R^{(m)}_{\widetilde{t}_n}\Delta_{\widetilde{N}}} \middle| R^{(m)}_{\widetilde{t}_{ki}} = r_j\right] = \mathbf{e}_j \left(\prod_{n=i+1}^{N} \left(e^{\mathbf{Q}^{(m)}_n \Delta_{\widetilde{N}}} e^{-\mathbf{D}\Delta_{\widetilde{N}}}\right)^k\right) \mathbf{1}_{m\times 1}, \tag{4.11}$$

*and*

$$\mathbb{E}\left[e^{-\sum_{n=ki}^{\widetilde{N}} R^{(m)}_{\widetilde{t}_n}\Delta_{\widetilde{N}}} \middle| R^{(m)}_{\widetilde{t}_{ki}} = r_j\right] = \mathbf{e}_j \left(\prod_{n=i+1}^{N} \left(e^{-\mathbf{D}_m \Delta_{\widetilde{N}}} e^{\mathbf{Q}^{(m)}_n \Delta_{\widetilde{N}}}\right)^k\right) e^{-\mathbf{D}_m \Delta_{\widetilde{N}}} \mathbf{1}_{m\times 1}. \tag{4.12}$$

The proof provided in Appendix 4.8 is intuitive and follows essentially by noticing that (4.11) and (4.12) are

matrix representations of the conditional expectation of a function of a discrete one-dimensional random process whose conditional probabilities are given by (4.3).

**Remark 4.4.1** *Using arguments similar to those of the proof of Lemma 4.2, we can also find that matrices* $e^{\mathbf{D}_m}$ *and* $e^{\mathbf{Q}_n^{(m)}}$ *commute under multiplication in* (4.12). *More precisely, using the notation of Lemma 4.2, we have that*

$$\mathbb{E}\left[e^{-\sum_{n=ki}^{\tilde{N}} R_{\tilde{t}_n}^{(m)} \Delta_{\tilde{N}}} \middle| R_{\tilde{t}_{ki}}^{(m)} = r_j\right] = \mathbf{e}_j \left(\prod_{n=i+1}^{N} \left(e^{-\mathbf{D}_m \Delta_{\tilde{N}}} e^{\mathbf{Q}_n^{(m)} \Delta_{\tilde{N}}}\right)^k\right) e^{-\mathbf{D}_m \Delta_{\tilde{N}}} \times \mathbf{1}_{m\times 1}$$
$$= \mathbf{e}_j e^{-\mathbf{D}_m \Delta_{\tilde{N}}} \left(\prod_{n=i+1}^{N} \left(e^{\mathbf{Q}_n^{(m)} \Delta_{\tilde{N}}} e^{-\mathbf{D}_m \Delta_{\tilde{N}}}\right)^k\right) \times \mathbf{1}_{m\times 1}.$$

Proposition 4.4.1 provides a closed-form matrix expression for the price of a zero-coupon bond under general time-inhomogeneous CTMCs. The result is a natural extension of Proposition 3 of Kirkby (2023) for time-inhomogeneous diffusion processes.

**Proposition 4.4.1** *Let Assumption 4.4.1 hold. Given that* $R_{t_i}^{(m)} = R_{t_i} = r_j \in \mathcal{S}_R^{(m)}$, *for some* $i \in \{0, 1, 2, \ldots, N\}$, *the price at time* $t_i$ *of a zero-coupon bond with maturity* $T \geq t_i$ *can be approximated by*

$$P_j^{(m)}(t_i, T) := \mathbb{E}\left[e^{-\int_{t_i}^{T} R_s^{(m)} \, \mathrm{d}s} \middle| R_{t_i}^{(m)} = r_j\right] = \mathbf{e}_j \left(\prod_{n=i+1}^{N} e^{\left(\mathbf{Q}_n^{(m)} - \mathbf{D}_m\right)\Delta_N}\right) \mathbf{1}_{m\times 1}. \tag{4.13}$$

The proof of Proposition 4.4.1, detailed in Appendix 4.8, relies on Lemma 4.2, the dominated convergence theorem, and the Lie product formula for the limit of matrix exponentials. For time-homogeneous models, an elegant proof can also be found in Kirkby (2023). The proof presented in this chapter employs straightforward and intuitive probabilistic arguments, making it applicable to both time-homogeneous and time-inhomogeneous models. However, it requires Assumption 4.4.1 to be satisfied for the use of dominated convergence, which is also necessary to ensure the convergence of the approximated price to the true price (as in Proposition 3 (iii) of Kirkby (2023)), as discussed in Remark 4.4.2 below.

**Remark 4.4.2 (Convergence of zero-coupon bonds)** *Under the condition of Proposition 4.4.1, the convergence of the approximated bond prices follows from the weak convergence of* $R^{(m)} \Rightarrow R$, *see Remark 4.3.1.*

*Indeed, from there, we have that* $\int_0^t R_s^{(m)}\,\mathrm{d}s \Rightarrow \int_0^t R_s\,\mathrm{d}s$ *from Proposition 4 of Cui* et al. *(2021a). Then, the convergence of the expectation follows directly from the Portmanteau theorem and Assumption 4.4.1 since, for* $x \geq r$*, the function* $e^{-x}$ *is continuous and bounded.*

*Detailed error and convergence analysis for CTMC methods applied to option pricing in one-dimensional settings are discussed in Li and Zhang (2018) and Zhang and Li (2019). Convergence behavior and rates are discussed further in Appendix 4.12.2.1.*

**Remark 4.4.3** *Assumption 4.4.1 ensures that the dominated convergence and Portmanteau theorems can be used in the proof Proposition 4.4.1 and Remark 4.4.2, respectively. However, numerical results in Section 4.6 and Appendix 4.12.2 show these results still hold when* $\mathcal{S}_R = \mathbb{R}$ *under some specific models and set of parameters.*

### 4.4.2 Bond Option

Proposition 4.4.2 provides an explicit closed-form matrix expression to approximate the prices of call and put options on zero-coupon bonds under general time-inhomogeneous short-rate models. To the best of the author's knowledge, this method of approximating the price of zero-coupon bond options is a novel contribution to the literature.

**Proposition 4.4.2** *Let Assumption 4.4.1 hold. Given that* $R_{t_{n_1}}^{(m)} = R_{t_{n_1}} = r_j \in \mathcal{S}_R^{(m)}$*, for some* $n_1 \in \{0, 1, 2, \ldots, N\}$*, the price at* $t_{n_1} \geq 0$ *of a European call (resp. put) option with maturity* $t_{n_2} > t_{n_1}$ *on a zero-coupon bond maturing at time* $T > t_{n_2}$ *with strike* $K > 0$ *can be approximated by*

$$\mathbb{E}\left[e^{-\int_{t_{n_1}}^{t_{n_2}} R_s^{(m)}\,\mathrm{d}s} h\left(P^{(m)}(t_{n_2}, T)\right) \middle| R_{t_{n_1}}^{(m)} = r_j\right] = \mathbf{e}_j \left(\prod_{n=n_1+1}^{n_2} e^{\left(\mathbf{Q}_n^{(m)} - \mathbf{D}_m\right)\Delta_N}\right) \mathbf{H}, \tag{4.14}$$

*where* $h(x) = \max(x - K, 0)$ *(resp.* $h(x) = \max(K - x, 0)$*) denotes the payoff function,* $P^{(m)}(t_{n_2}, T) := \mathbb{E}\left[e^{-\int_{t_{n_2}}^{T} R_s^{(m)}\,\mathrm{d}s} \middle| R_{t_{n_2}}^{(m)}\right]$ *denotes the approximated zero-coupon bond price at* $t_{n_2}$*, and* $\mathbf{H}$ *denotes a column vector of size* $m \times 1$*, whose* $k$*-th entry is given by* $h_k = h\left(P_k^{(m)}(t_{n_2}, T)\right)$*, with* $P_k^{(m)}(t_{n_2}, T)$ *defined in* (4.13)*.*

The proof follows using arguments similar to that of the proof of Proposition 4.4.1.

**Remark 4.4.4 (Convergence of zero-coupon bond options)** *From Remark 4.4.2, we have that* $\int_t^T R_s^{(m)}\,\mathrm{d}s \Rightarrow \int_t^T R^{(m)}\,\mathrm{d}s$ *for all* $t \leq T$*, as* $m, N \to \infty$*, and we conclude that* $P_j^{(m)}(t,T) \to P_j(t,T) := \mathbb{E}\left[e^{-\int_t^T R_s\,\mathrm{d}s}|R_t = r_j\right]$ *for all* $r_j \in \mathcal{S}_R^{(m)}$*. However, this is not sufficient to prove the convergence of zero-coupon bond option prices. Consequently, one first needs to show that* $P^{(m)}(t,T) = \mathbb{E}\left[e^{-\int_t^T R_s^{(m)}\,\mathrm{d}s}|R_t^{(m)}\right] \Rightarrow P(t,T) := \mathbb{E}\left[e^{-\int_t^T R_s\,\mathrm{d}s}|R_t\right]$*, that is, weak convergence of random variables also implies weak convergence of conditional expectations. This has been studied in the context of filtering theory by Goggin (1994), Kouritzin and Zeng (2005), and Crimaldi and Pratelli (2005). However, their results are inapplicable in that particular context since it requires finding a measure under which processes* $e^{-\int_t^T R_s\,\mathrm{d}s}$ *and* $R_t$ *are independent. An alternative way to prove the weak convergence of conditional expectations is to show that the estimation errors converge; see Goggin (1994), Lemma 2.2. Showing this property is, however, out of the scope of this chapter. Numerical experiments in the next section demonstrate the accuracy and efficiency of Proposition 4.4.2 empirically.*

*Detailed error and convergence analysis for CTMC methods applied to option pricing in one-dimensional settings are discussed in Li and Zhang (2018) and Zhang and Li (2019). Extensions to zero-bond option pricing are left for future investigations. Convergence behavior and rates are studied empirically in the next section.*

**Remark 4.4.5 (Extension to coupon-bearing bonds)** *The extension of* (4.14) *to coupon-bearing bonds is straightforward. Indeed, suppose that the underlying bond pays a periodic coupon* $\alpha > 0$ *at time* $t_{n_2+z} < t_{n_2+2z} < \ldots < t_{n_2+\widetilde{N}z} = T$*, with* $z = (N - n_2)/\widetilde{N}$*. Then, it suffices to replace* $P_k^{(m)}(t_{n_2}, T)$ *in Proposition 4.4.2 by*

$$P_k^{(m)}(t_{n_2}, T) + \sum_{n=1}^{\widetilde{N}} \alpha P_k^{(m)}(t_{n_2}, t_{n_2+nz}),$$

*with* $P_k^{(m)}(\cdot,\cdot)$*, defined in* (4.13).

### 4.4.3 Callable/Putable Bond

In the following sections, we develop simple and efficient algorithms for pricing callable and putable debt under general one-dimensional time-inhomogeneous short-rate processes. To the author's knowledge, this type of approximation for callable and putable bond pricing is novel in the literature.

Let $K_t^p, K_t^c \geq 0$ be constants representing the put and call prices (or strike prices) at time $t \leq T$, respectively, and let $F > 0$ be the face value of the bond, and $T > 0$ the maturity of the bond. In the following,

we assume that there is no coupon. However, adjustment to coupon-bearing bonds is straightforward, and it is discussed in greater detail below. Furthermore, we denote by $\mathcal{T}_{t,T}$, the (admissible) set of all stopping times taking values on the interval $[t, T]$.

When the put option can be exercised at any time prior to maturity, the value of a putable debt is equivalent to solving the following optimal stopping problem

$$v_p(t,r) := \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[ e^{-\int_t^\tau R_u \, \mathrm{d}u} \varphi_p(\tau) \middle| R_t = r \right], \tag{4.15}$$

where the reward (or payoff) function $\varphi_p : [0, T] \to \mathbb{R}_+$ is defined by

$$\varphi_p(t) = \begin{cases} K_t^p & \text{if } t < T, \\ \max(F, K_T^p) & \text{if } t = T. \end{cases} \tag{4.16}$$

On the other hand, the value of callable debt is given by

$$v_c(t,r) := \inf_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[ e^{-\int_t^\tau R_u \, \mathrm{d}u} \varphi_c(\tau) \middle| R_t = r \right], \tag{4.17}$$

where the reward function $\varphi_c : [0, T] \to \mathbb{R}_+$ is defined by

$$\varphi_p(t) = \begin{cases} K_t^c & \text{if } t < T, \\ \min(F, K_T^c) & \text{if } t = T. \end{cases} \tag{4.18}$$

Typically, a bond can have both a call and put options embedded; thus, the two problems in (4.16) and (4.17) need to be solved simultaneously. We denote by $v_{cp} : [0, T] \times \mathcal{S}_R \to \mathbb{R}_+$ the value function of the problem when (4.16) and (4.17) are solved together. Often, options are only exercisable during a certain period (exercise period or window). This is discussed further below. Numerical techniques are thus required to solve the problem. Commonly used techniques, such as trees, are based on the Bermudan[5] approximation of $v_{cp}$ and the dynamic programming principle (see, for instance, Lamberton (1998), Theorem 10.1.3). Proposition 4.4.3 relies on the same ideas.

**Proposition 4.4.3** *Let Assumption 4.4.1 hold. The value of a callable and putable bond with maturity $T > 0$ and face value $F > 0$ can be approximated recursively by*

$$\begin{cases} \mathbf{V}_N^{(m)} & = \max\left(\min\left(\mathbf{F}, \mathbf{K}_N^c\right), \mathbf{K}_N^p\right) \\ \mathbf{V}_n^{(m)} & = \max\left(\min\left(\mathbf{K}_n^c, e^{(\mathbf{Q}_{n+1}^{(m)} - \mathbf{D}_m)\Delta_N} \mathbf{V}_{n+1}^{(m)}\right), \mathbf{K}_n^p\right) \quad 0 \le n \le N-1. \end{cases} \tag{4.19}$$

[5] The Bermudan contract refers to a contract under which the embedded options can be exercised on a finite number of predetermined dates, whereas an American contract refers to a contract under which the embedded options can be exercised at any time from the inception to the maturity date.

*for a sufficiently large $N \in \mathbb{N}$, and where $\mathbf{K}_n^a = K_{t_n}^a \mathbf{1}_{m\times 1}$, $a \in \{p, c\}$, $\mathbf{F} = F\mathbf{1}_{m\times 1}$, and the maximum (resp. minimum) is taken element by element (also known as the parallel maxima (resp. minima)). Specifically, given $R_0^{(m)} = R_0 = r_j$, the approximated price of a callable and putable debt is given by*

$$v_{cp}^{(m)}(0, R_0) = \mathbf{e}_j \mathbf{V}_0^{(m)}.$$

The proof of Proposition 4.4.3 follows from the dynamic programming principle and by noting that $\exp\left(\mathbf{Q}_{n+1}^{(m)} - \mathbf{D}_m)\Delta_N\right)\mathbf{V}_{n+1}^{(m)}$ (the continuation value) is the matrix representation of the conditional expectation of a function of a discrete random variable whose conditional probability mass function is given by the transitional probability $p_{ij}(t_n, t_{n+1})$, $1 \leq i, j \leq m$, with $\mathbf{P}(t_n, t_{n+1}) = [p_{ij}(t_n, t_{n+1})]_{m\times m} = \exp\left(\mathbf{Q}_{n+1}^{(m)}\Delta_N\right)$ as per (4.3). The remainder of the proof follows the same reasoning used in the proofs of Lemma 4.2 and Proposition 4.4.1, detailed in Appendix 4.8.

The accuracy of (4.19) in pricing callable and putable debt is demonstrated numerically in Section 4.6. To price a callable only bond, it suffices to set $K_t^p = 0$, and for a putable only bond, one must let $K_t^c \to \infty$, $0 \leq t \leq T$. Different exercise windows can also be incorporated using a similar logic.

**Remark 4.4.6 (Extension to coupon-bearing bonds)** *Proposition 4.4.3 is set up for zero-coupon debt. However, extension to coupon-bearing bonds is straightforward. Indeed, when a coupon $\alpha > 0$ is paid $t_{n+1}$, it just needs to be discounted back at time $t_n$ with the value of the bond at $t_{n+1}$. More precisely,* (4.19) *becomes*

$$\mathbf{V}_n^{(m)} = \max\left(\min\left(\mathbf{K}_n^c, e^{(\mathbf{Q}_{n+1}^{(m)} - \mathbf{D}_m)\Delta_N}\left(\mathbf{V}_{n+1}^{(m)} + \alpha\mathbf{1}_{m\times 1}\right)\right), \mathbf{K}_n^p\right)$$

*for $n \in \{1, 2, \ldots, N-1\}$.*

The results of Lemma 4.2, and Propositions 4.4.1, 4.4.2, and 4.4.3 can be simplified when the short-rate process is time-homogeneous such as for the models listed in Table 4.1, or when it can be expressed as a sum of an auxiliary time-homogeneous process and a deterministic function of time as for the models listed in Table 4.3. This is discussed further in Appendices 4.9 and 4.10.

Extension of these results to two-factor short-rate models can be accomplished using the procedure of Section 4.3, along with Proposition 4.3.1. This is left as future research.

### 4.4.4 Calibration to the Initial Term Structure of Interest Rates

Using the closed-form formula for the price of a zero-coupon bond in (4.13), we can develop an efficient algorithm such that the zero-bond curve[6] of the approximated model fits the market curve.

To do so, we choose a time partition of $[0,T]$, $0 = t_0 < t_1 < \ldots < t_N = T$, with $N \in \mathbb{N}$, $T > 0$, $t_n = n\Delta_N$, $n \leq N$ and $\Delta_N = T/N$. We suppose that there is a time deterministic function $\theta$ that appears in the drift of (4.1) such that $\mu_R(t,r) = \widetilde{\mu}_R(\theta(t),r)$, for $(t,r) \in [0,T] \times \mathcal{S}_R$, as it may often be the case for time-inhomogeneous short-rate models, see the models listed in Table 4.2 for examples. Moreover, we assume that $\theta$ is piecewise constant in time, such that

$$\theta(t) = \sum_{n=1}^{N} \theta_n \mathbf{1}_{[t_{n-1},t_n)}(t). \tag{4.20}$$

for some $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N) \in \mathbb{R}^N$.

Let $t \mapsto P^\star(0,t)$ represent the current market zero-bond curve. The objective is to find the parameters $\boldsymbol{\theta}$ that make the zero-coupon bond prices under the approximated model equal to the market zero-coupon bond prices. Henceforth, we denote these calibrated parameters by $\boldsymbol{\theta}^\star$. Note that matrix $\mathbf{Q}_n$ in (4.5) depends on $\theta_n$ via function $\mu_R$, $n = 1, 2, \ldots, N$. In this subsection, we write $\mathbf{Q}_n^{(m)}(\theta_n)$ for $\mathbf{Q}_n^{(m)}$ to make this relation clearer. By inspecting (4.13), we also note that the zero-coupon bond price at $t_1$ only depends on $\mathbf{Q}_1^{(m)}(\theta_1)$, and the price at $t_2$ depends on $\mathbf{Q}_1^{(m)}(\theta_1)$ and $\mathbf{Q}_2^{(m)}(\theta_2)$; and so on. Thus, the calibrated parameters $\boldsymbol{\theta}^\star$, which make $P_j^{(m)}(0,t_i) = P^\star(0,t_i)$, $i = 1, 2 \ldots, N$, can be obtained recursively starting from $t_1$ to $t_N$. Algorithm 7 provides an efficient recursive procedure to find $\boldsymbol{\theta}^\star$. In Algorithm 7, $\mathbf{I}_{m\times m}$ denotes the identity matrix of size $m \times m$.

When the short-rate process can be modeled as a deterministic shift of a homogeneous process, such as the models listed in Table 4.3, the calibrated parameters $\boldsymbol{\theta}^\star$ have an explicit closed-form expression. This is discussed further in Appendix 4.10.3.

## 4.5 Application to the Pricing of Convertible Bonds

In this section, we develop efficient algorithms for pricing CBs using CTMC approximations. When the conversion feature is only permitted at maturity, a closed-form matrix expression is obtained. In this chapter,

[6] The term zero-bond curve refers to the term structure of discount factors (or zero-coupon bonds).

**Algorithm 7:** Calibration of $\boldsymbol{\theta}$ to the Current Market Term-Structure

**Input:** Let $\mathbf{Q}_n^{(m)}(\theta_n)$ be defined as in (4.5) and $t \mapsto P^\star(0,t)$ be the current market zero-bond curve,

$$n = 1, 2, \ldots, N$$

$N \in \mathbb{N}$, the number of time steps

$\Delta_N \leftarrow T/N$, the size of a time step

Set $t_n = n\Delta_N$, $n = 1, 2, \ldots, N$

Set $\mathbf{D}_m \leftarrow \text{diag}(\boldsymbol{r})$ with $\boldsymbol{r} = (r_1, r_2, \ldots, r_m)$, $r_k \in \mathcal{S}_R^{(m)}$, $k = 1, 2, \ldots, m$

/* Calibration to the current market zero-bond curve $t \mapsto P^\star(0,t)$ */

$\mathbf{A}^\star \leftarrow \mathbf{I}_{m\times m}$ **for** $n = 1, \ldots N$ **do**

  Find $\theta_n^\star$ such that $P^\star(0,t_n) - \mathbf{e}_j \mathbf{A}^\star \times e^{(\mathbf{Q}_n^{(m)}(\theta_n) - \mathbf{D}_m)\Delta_N} \mathbf{1}_{m\times 1} = 0$

  $\mathbf{A}^\star \leftarrow \mathbf{A}^\star \times e^{(\mathbf{Q}_n^{(m)}(\theta_n^\star) - \mathbf{D}_m)\Delta_N}$

**return** $\{\theta_n^\star\}_{n=1}^N$

we use the term European (resp. American)-style CB to refer to a CB under which the investor has the right to convert the bond at maturity only (resp. at any time prior to maturity). The use of CTMC approximation for pricing convertible debt is a novel contribution to the literature.

The frameworks outlined below consider two risk factors: equity and risk-free rate. Default/credit risk is incorporated into the model using the methodology of Tsiveriotis and Fernandes (1998) (TF). Their approach consists of splitting the debt into two components: a cash-only and an equity part. The **cash-only part** consists of coupons and principal payments, whereas the **equity part** consists of equity payments (when the debt is converted to stock). Each part is subject to different credit risks. Indeed, the cash-only part can be seen as a standard bond and is subject to the issuer default risk. Cash-flows are thus discounted at a risky rate. The equity part can be interpreted as an equity derivative and must thus be discounted at the risk-free rate.

In the following, we consider zero-coupon CBs since the extension to coupon-bearing convertible debt is straightforward. Indeed, when conversion can only occur at maturity (European-style), coupons can be added to the price. For American-style CBs, the procedure is similar to callable and putable bonds. That is, when a coupon is paid at time $t_{n+1}$, then it just needs to be discounted back to time $t_n$ with the value of the cash-only part of the bond at $t_{n+1}$, $0 \leq n \leq N-1$. This is discussed further in Remarks 4.5.1 and 4.5.5. Further, we suppose that the risky rate $\{\widetilde{R}_t\}_{0\leq t\leq T}$ is obtained by adding a time-deterministic credit

spread, $c : [0,T] \to [0,1]$, over the risk-free rate, that is, $\widetilde{R}_t = R_t + c_t$, $0 \leq t \leq T$. Finally, the face value of the bonds is denoted by $F > 0$, and $\eta > 0$ represents the conversion ratio.

Recall that $0 = t_0 < t_1 < \ldots < t_N = T$ is a time partition of $[0,T]$, where $T > 0$ denotes the maturity of the financial instrument, $t_n = n\Delta_N$, $n = 0,1,2,\ldots,N$, and $\Delta_N = T/N$, $N \in \mathbb{N}$. $(X^{(m,M)}, R^{(m)})$ denotes the regime-swiching CTMC approximation of $(X,R)$, see Section 4.3.2, taking values of a finite state-space $\mathcal{S}_X^{(M)} \times \mathcal{S}_R^{(m)}$ with $\mathcal{S}_X^{(M)} = \{x_1, x_2, \ldots, x_M\}$ and $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$, $m, M \in \mathbb{N}$. We have also defined $S^{(m,M)}$ in terms of $(X^{(m,M)}, R^{(m)})$ in (4.9), and the generator $\mathbf{G}_n^{(mM)}$ is defined in (4.10), $n = 1,2,\ldots,N$.

Throughout this section, we denote by $\{\mathbf{e}_{kl}\}_{k,l=1}^{m,M}$ the standard basis in $\mathbb{R}^{mM}$, that is, $\mathbf{e}_{kl}$ represents a row vector of size $1 \times mM$ with a value of 1 in the $(k-1)M+l$-th entry and 0 elsewhere. $\mathbf{D}_{mM} := \text{diag}\,(\boldsymbol{d})$ is an $mM \times mM$ diagonal matrix with vector $\boldsymbol{d} = (d_1, d_2, \ldots, d_{mM})$ on its diagonal, where $d_{(k-1)M+l} = r_k \in \mathcal{S}_R^{(m)}$, $k = 1,2,,\ldots,m$, $l = 1,2,,\ldots,M$.

### 4.5.1 European-Style Convertible Bond

Under the approach of Tsiveriotis and Fernandes (1998), the risk-neutral value of a European-style convertible debt, $v_e : [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R \to \mathbb{R}_+$, is given by

$$v_e(t,x,r) = \mathbb{E}\Big[e^{-\int_t^T R_u \,\mathrm{d}u}\eta S_T \mathbf{1}_{\{S_T \geq F/\eta\}} + e^{-\int_t^T R_u + c_u \,\mathrm{d}u} F \mathbf{1}_{\{S_T < F/\eta\}} \big| S_t = x, R_t = r\Big]. \tag{4.21}$$

The cash-only $v_e^{CO} : [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R \to \mathbb{R}_+$ and equity $v_e^E : [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R \to \mathbb{R}_+$ parts of the debt can then be defined as

$$v_e^E(t,x,r) := \mathbb{E}\Big[e^{-\int_t^T R_u \,\mathrm{d}u}\eta S_T \mathbf{1}_{\{S_T \geq F/\eta\}} | S_t = x, R_t = r\Big], \tag{4.22}$$

and

$$v_e^{CO}(t,x,r) := \mathbb{E}\Big[e^{-\int_t^T R_u + c_u \,\mathrm{d}u} F \mathbf{1}_{\{S_T < F/\eta\}} \big| S_t = x, R_t = r\Big], \tag{4.23}$$

respectively.

Under the assumption that the volatility parameter of $S$ in (4.1) is constant, $\sigma_S(r) = \widetilde{\sigma}_S > 0$ for all $r \in \mathcal{S}_R$, and the short-rate process is Gaussian; we can find an explicit expression for (4.21). This is the case for the Vasicek, Ho–Lee, and Hull–White models (see Tables 4.1 and 4.2 for details). This is discussed further in Appendix 4.12.1.

**Proposition 4.5.1** *Let Assumption 4.4.1 hold. Given that $S_0 > 0$, and $X^{(m,M)} = \ln(S_0) - \rho f(R_0) = x_i \in \mathcal{S}_X^{(M)}$, with $R_0 = r_j \in \mathcal{S}_R^{(m)}$, the value of the European-style CB with maturity $T > 0$, face value $F > 0$, and conversion ratio $\eta > 0$ can be approximated by*

$$\begin{aligned} v_e^{(m,M)}(0, S_0, R_0) &:= \mathbb{E}\Big[e^{-\int_0^T R_u^{(m)}\,\mathrm{d}u}\eta S_T^{(m,M)}\mathbf{1}_{\{S_T^{(m,M)} \geq F/\eta\}} \\ &\qquad + e^{-\int_0^T R_u^{(m)} + c_u\,\mathrm{d}u} F\mathbf{1}_{\{S_T^{(m,M)} < F/\eta\}} \big| S_0^{(m,M)} = S_0, R_0^{(m)} = R_0\Big] \\ &= \mathbf{e}_{ji} \prod_{n=1}^{N} e^{\left(\mathbf{G}_n^{(mM)} - \mathbf{D}_{mM}\right)\Delta_N}\mathbf{H}. \end{aligned} \tag{4.24}$$

*where $\mathbf{H}$ denotes a column vector of size $mM \times 1$ whose $(k-1)M + l$-th entry is given by*

$$h_{(k-1)M+l} = \eta e^{x_l + \rho f(r_k)}\mathbf{1}_{\{e^{x_l+\rho f(r_k)} \geq F/\eta\}} + e^{-\int_0^T c_u\,\mathrm{d}u} F\mathbf{1}_{\{e^{x_l+\rho f(r_k)} < F/\eta\}}, \tag{4.25}$$

*for $k = 1, 2, , \ldots, m$, $l = 1, 2, , \ldots, M$.*

The proof follows by noting that (4.24) is the matrix representation of the conditional expectation of a function of a discrete one-dimensional random variable whose conditional probability mass function is given by the transitional probability $p_{kl}(t_n, t_{n+1})$, $1 \leq k, l \leq mM$, with $\mathbf{P}(t_n, t_{n+1}) = [p_{kl}(t_n, t_{n+1})]_{mM \times mM}$ as defined in (4.3), with the generators $\mathbf{Q}_n^{(m)}$ replaced by $\mathbf{G}_n^{(mM)}$. The remainder of the proof follows the same reasoning used in the proofs of Lemma 4.2 and Proposition 4.4.1, detailed in Appendix 4.8.

**Remark 4.5.1 (Extension to coupon-bearing European CBs)** *To extend the valuation to coupon-bearing bonds, it suffices to add the discounted value of the future coupons to the value obtained in* (4.24)*, similar to the approach in Remark 4.4.5. More precisely, assume that a periodic coupon, $\alpha > 0$, is paid at times $t_z < t_{2z} < \ldots < t_{\widetilde{N}z} = T$, with $z = N/\widetilde{N}$. The present value of future coupons is given by*

$$\sum_{n=1}^{\widetilde{N}} \alpha P_j^{(m)}(0, t_{nz}),$$

*with $P_j^{(m)}(\cdot, \cdot)$ defined in* (4.13)*. Adding this to the value of the European-style CB without coupons obtained in* (4.24) *completes the extension.*

**Remark 4.5.2 (Convergence of European-style CBs)** *From Remark 4.3.2, we have that $S^{(m,M)} \Rightarrow S$. The convergence of derivatives with a continuous and bounded payoff function then follows directly from the Portmanteau theorem, see for instance Billingsley (1999), Theorem 2.1. However, for discontinuous and*

*unbounded payoff functions such as that involved in* (4.21)[7]*, the convergence of the prices is not as straightforward. For a continuous and unbounded payoff function* $g$*, Mijatović and Pistorius (2009), Remark 3 and Cui* et al. *(2018), Remark 5, suggest replacing the original payoff function by a truncated payoff* $g \wedge L$*, with a constant* $L > 0$ *chosen to be sufficiently large such that the numerical results are not altered. Kirkby (2023), Proposition 6, shows the convergence of the derivative prices for continuous bounded payoffs, such as equity cap and floor, that is, when the payoff function is bounded from above and below.*

*Detailed error and convergence analysis in the context of European option pricing under two-dimensional stochastic local volatility models are discussed in Ma* et al. *(2022). Extensions to European-style CBs under two-dimensional stochastic interest rate models are left for future research. Numerical experiments in Section 4.6 demonstrate the accuracy and efficiency of the approximation empirically.*

The closed-form matrix expression in (4.24) can be implemented in a straightforward manner. However, as highlighted in Chapter 3, several numerical issues can be encountered when dealing with medium/long time-horizon derivatives because of the size of generator $\mathbf{G}_n^{(mM)}$. Hence, based on Propostion 3.4, a new algorithm that speeds up the pricing of European-style CBs is developed. This fast version of Proposition 4.5.1 is presented in Appendix 4.11.1.

### 4.5.2 Convertible Bond (American-Style)

When the conversion option can be exercised at any time prior to maturity (and call and put features are ignored), the valuation of CBs is equivalent to solving the following optimal stopping problem

$$v(t,x,r) = \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[ e^{-\int_t^\tau R_u + c_u \mathbf{1}_{\{\tau = T, S_T < F/\eta\}} \,\mathrm{d}u} \varphi(\tau, S_\tau) \middle| S_t = x, R_t = r \right], \tag{4.26}$$

where $\mathcal{T}_{t,T}$ denotes the (admissible) set of all stopping times taking values on the interval $[t,T]$, and the reward (or gain) function $\varphi : [0,T] \times \mathbb{R}_+^\star \to \mathbb{R}_+^\star$ is defined by

$$\varphi(t,x) = \begin{cases} \eta x & \text{if } t < T, \\ \max(\eta x, F) & \text{if } t = T. \end{cases} \tag{4.27}$$

[7] The payoff function exhibits a discontinuity in the state variable because of the difference between the risky and the risk-free rates.

**Remark 4.5.3** *When $x < F/\eta$, the reward function is discontinuous at $T$ since*

$$\lim_{t\to T^-} \varphi(t,x) = \eta x < F = \varphi(T,x).$$

Assuming that an optimal stopping time[8] $\tau_t^\star$ exists, the cash-only part $v^{CO} : [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R \to \mathbb{R}_+$ and equity part $v^E : [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R \to \mathbb{R}_+$ of the CB can be defined by

$$v^{CO}(t,x,r) = \mathbb{E}\left[e^{-\int_t^T R_u + c_u \,\mathrm{d}u} F \mathbf{1}_A \middle| S_t = x, R_t = r\right], \text{ and}$$
$$v^{E}(t,x,r) = \mathbb{E}\left[e^{-\int_t^{\tau_t^\star} R_u \,\mathrm{d}u} \eta S_{\tau_t^\star} \mathbf{1}_{A^c} \middle| S_t = x, R_t = r\right],$$

respectively, where $A := \{\tau_t^\star = T, S_T < F/\eta\}$, and $A^c := \{\tau_t^\star < T\} \cup \{\tau_t^\star = T, S_T \geq F/\eta\}$ denotes the complement of $A$. It follows that $v(t,x,r) = v^{CO}(t,x,r) + v^E(t,x,r)$.

When no dividends are paid, and credit risk is assumed to be nil ($q_t = c_t = 0$ for all $t \in [0,T]$), the value of the CB in (4.26), which can be exercised at any time prior to maturity, is equal to the value of the European-style CB (4.21), meaning that an optimal stopping time for (4.26) is at the maturity of the bond. On the other hand, when credit risk is considered, the value of American-style CBs is bounded from below and above by those of European-style CBs with and without credit risk, respectively. This is formalized in the following.

**Proposition 4.5.2** *Assume that $q_t = c_t = 0$ for all $t \in [0,T]$ and $\sigma_S(r) = \widetilde{\sigma}_S > 0$ for all $r \in \mathcal{S}_R$. We have that $v(t,x,r) = v_e(t,x,r)$ for all $(t,x,r) \in [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R$.*

*Proof. We first show that the discounted reward process $\{e^{-\int_0^t R_u \,\mathrm{d}u}\varphi(t,S_t)\}_{0\leq t\leq T}$ is a submartingale. For $0 \leq s \leq t < T$, we have that*

$$\mathbb{E}\left[e^{-\int_0^t R_u \,\mathrm{d}u}\varphi(t,S_t)\middle|\mathcal{F}_s\right] = \mathbb{E}\left[e^{-\int_0^t R_u \,\mathrm{d}u}\eta S_t|\mathcal{F}_s\right] = e^{-\int_0^s R_u \,\mathrm{d}u}\eta S_s = e^{-\int_0^s R_u \,\mathrm{d}u}\varphi(s,S_s),$$

*where the second equality follows from the martingale property of the discounted stock process under the risk-neutral measure (see Remark 4.5.4). On the other hand, if $0 \leq s < t = T$, we have that*

$$\begin{aligned}\mathbb{E}\left[e^{-\int_0^T R_u \,\mathrm{d}u}\varphi(T,S_T)\middle|\mathcal{F}_s\right] &= \mathbb{E}\left[e^{-\int_0^T R_u \,\mathrm{d}u}\max(\eta S_T, F)\middle|\mathcal{F}_s\right] \\ &\geq \mathbb{E}\left[e^{-\int_0^T R_u \,\mathrm{d}u}\eta S_T\middle|\mathcal{F}_s\right] = e^{-\int_0^s R_u \,\mathrm{d}u}\varphi(s,S_s).\end{aligned}$$

[8] An admissible stopping time $\tau_t^\star \in \mathcal{T}_{t,T}$ is said to be optimal for (4.26) if

$$v(t,x,r) = \mathbb{E}\left[e^{-\int_t^{\tau_t^\star} R_u + c_u \mathbf{1}_{\{\tau_t^\star = T, \eta S_T < F\}} \,\mathrm{d}u}\varphi(\tau_t^\star, S_{\tau_t^\star})\middle|S_t = x, R_t = r\right].$$

*The final assertion follows for well-known results in optimal stopping theory, which states that if the discounted reward process is a submartingale, then the maturity of the contract is an optimal stopping time, see Björk (2009), Proposition 21.2.* □

When periodic coupons are paid, the results of Proposition 4.5.2 still hold. This is demonstrated in Corollary 4.4.

Let $\widetilde{v}_e : [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R \to \mathbb{R}_+$ denote the value function of European-style CBs when $c_t = 0$ for all $t \in [0,T]$. Using (4.21), it follows that

$$\begin{aligned} \widetilde{v}_e(t,x,r) &:= \mathbb{E}\Big[e^{-\int_t^T R_u \,\mathrm{d}u} \max\left(\eta S_T, F\right) \Big| S_t = x, R_t = r\Big] \\ &= \mathbb{E}\Big[e^{-\int_t^T R_u \,\mathrm{d}u} \varphi(T, S_T) \Big| S_t = x, R_t = r\Big]. \end{aligned} \tag{4.28}$$

**Corollary 4.3** *Assume that $q_t = 0$ for all $t \in [0,T]$ and $\sigma_S(r) = \widetilde{\sigma}_S > 0$ for all $r \in \mathcal{S}_R$. We have that $v_e(t,x,r) \leq v(t,x,r) \leq \widetilde{v}_e(t,x,r)$ for all $(t,x,r) \in [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R$.*

*Proof. The first inequality follows directly since $T \in \mathcal{T}_{t,T}$. For the second inequality, it suffices to note that*

$$\begin{aligned} v(t,x,r) &= \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[e^{-\int_t^\tau R_u + c_u \mathbf{1}_{\{\tau = T, \S_T < F/\eta\}} \,\mathrm{d}u} \varphi(\tau, S_\tau) \Big| S_t = x, R_t = r\right] \\ &\leq \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E}\left[e^{-\int_t^\tau R_u \,\mathrm{d}u} \varphi(\tau, S_\tau) \Big| S_t = x, R_t = r\right] = \widetilde{v}_e(t,x,r) \end{aligned}$$

*where the last equality follows from Proposition 4.5.2.* □

When periodic coupons are paid, the result of Corollary 4.3 still holds. This is discussed in Corollary 4.8.3.

**Remark 4.5.4** *Condition $\sigma_S(r) = \widetilde{\sigma}_S > 0$ for all $r \in \mathcal{S}_R$ in Propositions 4.5.2 can be relaxed provided that the discounted stock process remains a true martingale under the risk-neutral measure. Indeed, when no additional condition is added, the discounted stock process $\{e^{-\int_0^t R_u \,\mathrm{d}u} S_t\}_{t \geq 0}$ is a local martingale. For $\{e^{-\int_0^t R_u \,\mathrm{d}u} S_t\}_{t \geq 0}$ to be a true martingale, some restrictions must be added to the parameters of the short-rate dynamics. This has been studied in stochastic volatility models for some specific time-homogeneous diffusion processes; see Sin (1998), Jourdain (2004), and Cui (2013). Conditions under which the discounted stock process is a true martingale for the particular short-rate models listed in Tables 4.1, 4.2, and 4.3 are left as future research.*

*Note also that the aforementioned results are applicable only to stocks that do not pay dividends, that is $q_t = 0$ for all $t \geq 0$. When dividends are distributed, the discounted stock process becomes a supermartingale, which makes the arguments in the proof of Proposition 4.5.2 invalid.*

When the conditions of Proposition 4.5.2 are not satisfied, or the debt includes other specific features such as call and/or put options, as is often the case in practice, numerical techniques are required to solve the optimal stopping problem in (4.26). Commonly used methods, such as trees or least-squares Monte Carlo (see Longstaff and Schwartz (2001)), are based on the Bermudan approximation[5] of $v$ and the dynamic programming principle. The same ideas are used in Proposition 4.5.3.

Consequently, we define $\mathbf{H}^{CO}$ (resp. $\mathbf{H}^{E}_{n}$, $0 \leq n \leq N$) as column vectors of size $mM \times 1$ representing the cash-only (resp. equity) part of the reward, whose $(k-1)M + l$-entry are respectively given by

$$h^{CO}_{(k-1)M+l} = F\mathbf{1}_{\{e^{x_l+\rho f(r_k)} < F/\eta\}}, \tag{4.29}$$

and

$$h^{E}_{(k-1)M+l,n} = \begin{cases} \eta e^{x_l+\rho f(r_k)}, & \text{if } 0 \leq n \leq N-1, \\ \eta e^{x_l+\rho f(r_k)}\mathbf{1}_{\{e^{x_l+\rho f(r_k)} \geq F/\eta\}}, & \text{if } n = N, \end{cases} \tag{4.30}$$

for $1 \leq k \leq m$, $1 \leq l \leq M$. Furthermore, let $\mathbf{B}^{CO}_{n}$, $\mathbf{B}^{E}_{n}$, and $\mathbf{B}_n$, $0 \leq n \leq N$, be column vectors of size $mM \times 1$, representing the cash-only part, the equity part, and the total value of the CB at time $t_n$, respectively. We denote by $b_{k,n}$ the $k$-th entry of $\mathbf{B}_n$, $1 \leq k \leq mM$, and define the indicator vector $\mathbf{1}_{\{\mathbf{B}_n = \mathbf{H}^{E}_{n}\}}$, where the $k$-th entry of $\mathbf{1}_{\{\mathbf{B}_n = \mathbf{H}^{E}_{n}\}}$, denoted by $\mathbf{1}_{\{\mathbf{B}_n = \mathbf{H}^{E}_{n}\}}(k)$, is given by $\mathbf{1}_{\{\mathbf{B}_n = \mathbf{H}^{E}_{n}\}}(k) = \mathbf{1}_{\{b_{k,n} = h^{E}_{k,n}\}}$, for each $k \in \mathcal{S}^{mN}_{Z}$. Finally, $\mathbf{1}_{mM\times 1}$ denotes the unit vector of size $mM \times 1$.

**Proposition 4.5.3** *Let Assumption 4.4.1 hold. The value of a CB with maturity $T > 0$, face value $F > 0$, and*

*conversion ratio $\eta > 0$, can be approximated recursively by*

$$
\begin{aligned}
\widehat{\mathbf{B}}_n^{CO\text{'}} &= \begin{cases} \mathbf{H}^{CO}, & \text{if } n = N, \\ e^{-\int_{t_n}^{t_{n+1}} c_u \, \mathrm{d}u} \exp\left\{\left(\mathbf{G}_{n+1}^{(mN)} - \mathbf{D}_{mM}\right)\Delta_N\right\}\mathbf{B}_{n+1}^{CO}, & \text{if } 0 \le n \le N-1, \end{cases} \\
\mathbf{B}_n^{CO} &= \begin{cases} \widehat{\mathbf{B}}_N^{CO}, & \text{if } n = N, \\ \widehat{\mathbf{B}}_n^{CO}\left(\mathbf{1}_{mM\times 1} - \mathbf{1}_{\{\mathbf{B}_n = \mathbf{H}_n^E\}}\right), & \text{if } 0 \le n \le N-1, \end{cases} \\
\widehat{\mathbf{B}}_n^{E} &= \begin{cases} \mathbf{H}_N^E, & \text{if } n = N, \\ \exp\left\{\left(\mathbf{G}_{n+1}^{(mN)} - \mathbf{D}_{mM}\right)\Delta_N\right\}\mathbf{B}_{n+1}^{E}, & \text{if } 0 \le n \le N-1, \end{cases} \\
\mathbf{B}_n^{E} &= \begin{cases} \widehat{\mathbf{B}}_N^{E}, & \text{if } n = N, \\ \mathbf{B}_n - \mathbf{B}_n^{CO}, & \text{if } 0 \le n \le N-1, \end{cases} \\
\mathbf{B}_n &= \begin{cases} \widehat{\mathbf{B}}_N^{CO} + \widehat{\mathbf{B}}_N^{E}, & \text{if } n = N, \\ \max\left(\mathbf{H}_n^E, \widehat{\mathbf{B}}_n^{CO} + \widehat{\mathbf{B}}_n^{E}\right) & \text{if } 0 \le n \le N-1. \end{cases}
\end{aligned}
$$

*for a sufficiently large $N \in \mathbb{N}$, and where the maximum is taken element by element. Specifically, given that $X_0^{(m,M)} = \ln(S_0) - \rho f(R_0) = x_i \in \mathcal{S}_X^{(M)}$ and $R_0^{(m)} = R_0 = r_j \in \mathcal{S}_R^{(m)}$, the value of an American-style CB can be approximated by*

$$v^{(m,M)}(0, S_0, R_0) = \mathbf{e}_{ji}\mathbf{B}_0^{(mM)}.$$

Proposition 4.5.3 is presented as an algorithm in Appendix 4.11. Similar to the European-style CB, the performance of the procedure in Proposition 4.5.3 can be significantly increased using the technique of Chapter 3, Proposition 3.4. This new fast version of the procedure is also reported in Appendix 4.11.

**Remark 4.5.5 (Extension to coupon-bearing bonds)** *Recall that Proposition 4.5.3 is set up for zero-coupon CBs. However, as mentioned previously, adding coupons to the previous procedure is straightforward. Indeed, when a coupon is paid at time $t_{n+1}$, it just needs to be discounted back to time $t_n$ with the cash-only part of the CB. More precisely, when a coupon $\alpha > 0$ is paid at time $t_{n+1}$, then the continuation value of the cash-only part at $t_n$, $\widehat{\mathbf{B}}_n^{CO}$, must be calculated as follow*

$$\widehat{\mathbf{B}}_n^{CO} = e^{-\int_{t_n}^{t_{n+1}} c_u \, \mathrm{d}u} \exp\left\{\left(\mathbf{G}_{n+1}^{(mN)} - \mathbf{D}_{mM}\right)\Delta_N\right\}\left(\mathbf{B}_{n+1}^{CO} + \alpha\mathbf{1}_{mN\times 1}\right),$$

*for $n \in \{0, 1, \ldots, N-1\}$.*

Additional features, such as call and put options, can also be added, similarly as in tree methods (see, for instance, Hung and Wang (2002), Exhibit 2), by modifying the value of the convertible debt, the equity part, and the cash-only part accordingly at each time step.

**Remark 4.5.6 (Convergence of convertible bonds (American-style))** *When there is no credit risk, the convergence follows as in Remark 3.4.4. Indeed, in that particular case, we can rely on the continuous-reward representation of Theorem 2.6, and use the results of Song* et al. *(2013), Theorem 9, to establish the convergence. When credit risk is considered, the convergence is less clear because of the discontinuity in the discounted reward process created by the difference between the risky and risk-free rates.*

*Detailed error and convergence analysis for American-style CBs are left for future research. The accuracy and efficiency of Proposition 4.5.3 in pricing American-style CBs is demonstrated empirically in Section 4.6.*

## 4.6 Numerical Experiments

In this section, numerical experiments are conducted to analyze the performance of the methodology proposed in the previous sections under models listed in Tables 4.2 and 4.3. For the testing, we selected the Hull–White model, which is widely used in practice, and the extended CIR model (CIR++) for its analytical tractability. More precisely, we analyze the accuracy and efficiency[9] of CTMC approximations in valuing different debt securities. Numerical convergence is also investigated. Note that Assumption 4.4.1 is not respected under the Hull–White model.

In Appendix 4.12.2, a similar analysis is performed for the Vasicek and CIR models, two time-homogeneous short-rate processes of Tables 4.1. The accuracy and efficiency of the CTMC methods in approximating zero-bond prices, Proposition 4.4.1, are also investigated, and numerical convergence is analyzed. Additional examples with Dothan, exponential Vasicek, EV+, and EEV+ models are also available upon request. Results under these models are similar to those obtained under the Hull–White and CIR++ short-rate processes documented below.

All the numerical experiments are conducted with Matlab R2015a on a Core i7 desktop with 16GB RAM and a speed of 2.40 GHz. Matrix exponentials are calculated using the function fastExpm for Matlab, see

[9] The term "efficiency" refers to the ratio of the computation time of a procedure to the precision of its numerical result.

Mentink-Vigier (2020), which is designed to accelerate the calculation of large (sparse and full) matrices. Column "CTMC" reports the CTMC approximated value calculated using the results of Sections 4.4 or 4.5. Column "Benchmark" presents the benchmark value, column "Abs. error" documents the absolute error[10], whereas column "Rel. error" provides the relative error. The convergence rate about the number of grid points $m$ is approximated using the following formula:

$$\text{Rate} \approx \frac{\log\left(e_{m_2}/e_{m_1}\right)}{\log\left(m_1/m_2\right)},$$

where $e_m$ is the absolute error using the number of grid points $m$. Throughout this chapter, log refers to the natural logarithm.

In all the following numerical experiments, the model is calibrated to the market risk-free discount curve[11] reported in Table 4.4. The calibration to the market curve is performed using Algorithm 7 for the Hull–White model and Algorithm 8 for the CIR++ model.

Table 4.4: Market zero-bond curve, $t \mapsto P^{\star}(0,t)$.

| $\mathbf{t}$ | 0.26 | 0.47 | 0.72 | 0.97 | 1.22 | 1.47 | 1.72 | 2 | 3 | 4 |
|---|---|---|---|---|---|---|---|---|---|---|
| $\mathbf{P}^{\star}(\mathbf{0},\mathbf{t})$ | 0.986944 | 0.976019 | 0.964123 | 0.953152 | 0.943283 | 0.934357 | 0.926202 | 0.917553 | 0.888740 | 0.861950 |

For the Hull-White model, the state-space of the approximated short-rate process, $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$ with $m \in \mathbb{N}$, is constructed using the non-uniform grid proposed by Tavella and Randall (Tavella and Randall (2000), Chapter 5.3). That is, we first select the grid lower and upper bounds, $r_1, r_m \in \mathcal{S}_R$, and set the other grid points as follows $r_k = R_0 + \widetilde{\alpha}_R \sinh\left(c_2 \frac{k}{m} + c_1\left[1 - \frac{k}{m}\right]\right), \quad k = 2, \ldots, m-1$, where $c_1 = \sinh^{-1}\left(\frac{r_1 - R_0}{\widetilde{\alpha}}\right)$, $c_2 = \sinh^{-1}\left(\frac{r_m - R_0}{\widetilde{\alpha}}\right)$, and $\widetilde{\alpha}_R \geq 0$, controls the degree of non-uniformity of the grid. For the CIR++ models, the same procedure is applied to the auxiliary process $Y^{(m)}$, and we denote by $\widetilde{\alpha}_Y$ the grid non-uniformity parameter. Note that when $R_0$ (or $Y_0$ for the CIR++ model) is not in the grid, it is inserted (see, for instance, Cui *et al.* (2019), Section 2.3 for details). Unless stated otherwise, all model experiments are conducted using the model and the CTMC parameters summarized in Table 4.5.

[10] The *absolute error* is defined as the absolute value of the difference between the CTMC approximated value and the benchmark value.

[11] The market discount curve is obtained from Bloomberg and corresponds to the US Dollar curve 23 as of March 31, 2023, from the Swap Curve Builder (ICVS) page.

Table 4.5: Model and CTMC parameters

| | $R_0$ | $\alpha$ | $\kappa$ | $\sigma$ | $m$ | $r_1$ | $r_m$ | $\widetilde{\alpha}_R$ | $\Delta_N$ |
|---|---|---|---|---|---|---|---|---|---|
| Hull–White | $0.04$ | N/A | $1$ | $0.20$ | $160$ | $-30r_0$ | $25r_0$ | $0.5$ | 1/252 |
| | $Y_0 = R_0$ | $\alpha$ | $\kappa$ | $\sigma$ | $m$ | $y_1$ | $y_m$ | $\widetilde{\alpha}_Y$ | $\Delta_N$ |
| CIR++ | 0.04 | 0.035 | 2 | 0.20 | $160$ | $y_0/100$ | $7y_0$ | $0.5$ | 1/252 |

### 4.6.1 Approximation of Zero-Coupon Bond Option Prices

In this section, we study the accuracy and efficiency of (4.14), as well as the numerical convergence of the approximated zero-coupon bond option prices, under the Hull–White and CIR++ models[12], respectively. Under these particular models, the price of zero-coupon bond options have a closed-form expression, which can be found in Brigo and Mercurio (2006), Section 3, and thus, can serve as a benchmark in our analysis.

We test the accuracy of the approximated option prices for different levels of moneyness and volatilities. The results are summarized in Table 4.6. Column "price-to-strike" shows the price-to-strike ratio, calculated as the actual zero-coupon bond price over the option strike price $K > 0$. The price-to-strike ratio is a measure indicating the degree of moneyness of an option. A ratio above (resp. below) one shows that the call option is in the money (resp. out of the money), whereas a value of one indicates that the option is at the money[13].

We observe that the two models achieve high degrees of accuracy across all parameters and strikes, with absolute errors below $8.11$E-$06$. It is also worth noticing the high precision of the approximation for deep-out-of-the-money options, indicating good approximations of the left tails of the underlying short-rate process. Such a high level of accuracy can be difficult to attain when using other numerical techniques, particularly for out-of-the-money options.

Analogous experiments have been conducted with put options, with similar results for out-of-the-money put options, indicating good approximations of the right tails of the short-rate diffusion process. The ac-

[12] For the CIR++ model, (4.14) can be greatly simplified using the time-homogeneous property of the auxiliary process $Y^{(m)}$, see (4.43) in Appendix 4.10 for details.

[13] The price-to-strike ratios differ between the Hull-White and CIR++ models because the short rate in the CIR++ model cannot become negative, limiting the zero-bond price to $1$.

Table 4.6: Accuracy of the approximated price of zero-coupon bond call options, Proposition 4.4.2 and Corollary 4.9, under the Hull–White and CIR++ models, respectively. Benchmark prices are calculated using closed-form analytical formulas. Except for the number of grid points set to $m = 200$, model and CTMC parameters are as listed in Table 4.5. Zero-coupon bond call option parameters using the notation of Proposition 4.4.2: $t_{n_1} = 0$, $t_{n_2} = 2$, and $T = 4$.

| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
|---|---|---|---|---|
| $\mathbf{0.1}$ | 1.67 | 0.38741911 | 0.38741911 | 4.95E-14 |
| | 1.25 | 0.22924215 | 0.22924215 | 3.43E-10 |
| | 1.00 | 0.07281370 | 0.07281784 | 4.14E-06 |
| | 0.83 | 0.00130678 | 0.00130552 | 1.26E-06 |
| | 0.71 | 0.00000021 | 0.00000023 | 2.49E-08 |
| $\mathbf{0.2}$ | 1.67 | 0.38741912 | 0.38741912 | 4.50E-10 |
| | 1.25 | 0.22939397 | 0.22939439 | 4.21E-07 |
| | 1.00 | 0.08510628 | 0.08509821 | 8.07E-06 |
| | 0.83 | 0.01328299 | 0.01328294 | 4.45E-08 |
| | 0.71 | 0.00083528 | 0.00083714 | 1.86E-06 |
| $\mathbf{0.3}$ | 1.67 | 0.38743466 | 0.38743476 | 9.94E-08 |
| | 1.25 | 0.23168156 | 0.23168115 | 4.13E-07 |
| | 1.00 | 0.10193411 | 0.10192793 | 6.17E-06 |
| | 0.83 | 0.03096382 | 0.03096201 | 1.81E-06 |
| | 0.71 | 0.00681115 | 0.00681259 | 1.44E-06 |
| $\mathbf{0.4}$ | 1.67 | 0.38776664 | 0.38776489 | 1.75E-06 |
| | 1.25 | 0.23777915 | 0.23777104 | 8.11E-06 |
| | 1.00 | 0.12017293 | 0.12017104 | 1.89E-06 |
| | 0.83 | 0.05053076 | 0.05052852 | 2.24E-06 |
| | 0.71 | 0.01840799 | 0.01841005 | 2.05E-06 |

(a) Hull–White model

| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
|---|---|---|---|---|
| $\mathbf{0.1}$ | 1.67 | 0.38741911 | 0.38741911 | 2.46E-11 |
| | 1.25 | 0.22924215 | 0.22924215 | 2.18E-11 |
| | 1.00 | 0.07106519 | 0.07106519 | 1.90E-11 |
| | 0.95 | 0.03153223 | 0.03153224 | 9.24E-09 |
| | 0.92 | 0.00150899 | 0.00150815 | 8.40E-07 |
| $\mathbf{0.2}$ | 1.67 | 0.38741911 | 0.38741911 | 3.58E-10 |
| | 1.25 | 0.22924215 | 0.22924215 | 3.12E-10 |
| | 1.00 | 0.07106519 | 0.07106519 | 2.86E-10 |
| | 0.95 | 0.03153223 | 0.03153224 | 9.24E-09 |
| | 0.92 | 0.00299047 | 0.00299063 | 1.62E-07 |
| $\mathbf{0.3}$ | 1.67 | 0.38741912 | 0.38741911 | 1.61E-10 |
| | 1.25 | 0.22924215 | 0.22924215 | 1.39E-10 |
| | 1.00 | 0.07106791 | 0.07106800 | 8.86E-08 |
| | 0.95 | 0.03153223 | 0.03153224 | 9.24E-09 |
| | 0.92 | 0.00432070 | 0.00431963 | 1.07E-06 |
| $\mathbf{0.4}$ | 1.67 | 0.38741912 | 0.38741911 | 1.55E-10 |
| | 1.25 | 0.22924215 | 0.22924215 | 1.23E-10 |
| | 1.00 | 0.07110811 | 0.07111282 | 4.71E-06 |
| | 0.95 | 0.03153223 | 0.03153224 | 9.24E-09 |
| | 0.92 | 0.00545130 | 0.00544519 | 6.11E-06 |

(b) CIR++ model

curacy of the approximation across different model parameter values has also been tested. The methods exhibit a high level of precision across all parameters. Results are available upon request.

We compare the efficiency of the CTMC methodology (4.14) to the trinomial tree method ("Tree") of Hull and White (1994) and Hull and White (1996), and Monte Carlo simulation ("Sim") using the procedure of Ostrovski (2013). For the Monte Carlo simulation, we used an Euler discretization scheme, with a number of simulations ranging from $10,000$ to $100,000$ and $252$-time steps per year. Figure 4.1 shows the results. The calibration time is not included in the log elapsed time in Figure 4.1, that is, for the tree, the calculation

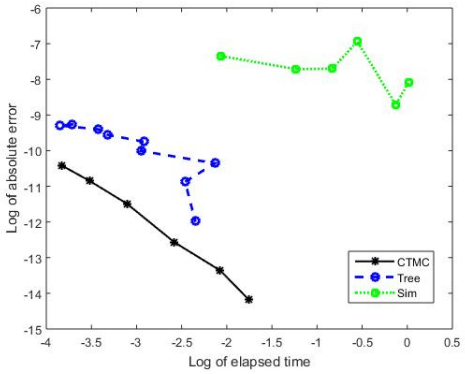
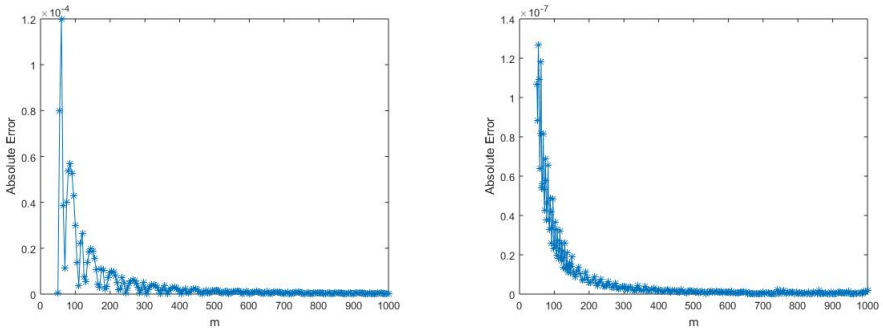



Figure 4.1: Efficiency of the CTMC method (4.14) compared to trees and simulation methods in approximating the price of zero-coupon bond call options under the Hull–White model. Except for the number of grid points $m$, which range from $100$ to $350$, model and CTMC parameters are as listed in Table 4.5. For the tree, we use between 300 and 700 time steps per year. Zero-coupon bond call option parameters using the notation of Proposition 4.4.2: $t_{n_1} = 0$, $t_{n_2} = 2, T = 4$, and $K = 0.9$.

time does not include the construction of the interest rate tree that perfectly fits the market data, and for the CTMC approximation, the calculation time does not include the calibration of the model to the market zero-bond curve using Algorithm 7. Note that the construction of the interest rate tree is generally much faster than the CTMC calibration process with Algorithm 7. When using 252-time steps per year, the tree is built in a fraction of a second, whereas the calibration process with Algorithm 7 takes on average 3.2 seconds with $m = 160$. Figure 4.1 shows the high efficiency of CTMC methods compared to other methods. CTMC approximation clearly outperforms these other techniques in terms of both calculation time and precision. The speed of the approximation for the extended models of Brigo and Mercurio (2006), such as the CIR++ model, is similar to that obtained for the homogeneous models with an average calculation time of less than $0.015$ seconds (excluding calibration).

The convergence patterns of the value of the call option to the analytical price as the number of grid points $m$ increases are displayed in Figure 4.2, whereas Table 4.7 shows the convergence rate. We observe that the approximation achieves superquadratic convergence on average. The CTMC approximation is known to achieve a theoretical quadratic convergence rate (rate = 2) for barrier and European options in one-dimensional diffusion models with particular grid designs; see Zhang and Li (2019) for details. We also note that the two models converge rapidly to their analytical values but exhibit a sawtooth pattern. Such oscillatory behavior has also been observed by Zhang and Li (2019) in the context of double barrier knock-out options pricing within the CTMC approximation framework. In particular, they observe that constructing




(a) Hull–White (b) CIR++

Figure 4.2: Convergence pattern of the approximated zero-coupon bond call prices under Hull–White and CIR++ models. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.5. Zero-coupon bond call option parameters using the notation of Proposition 4.4.2: $t_{n_1} = 0$, $t_{n_2} = 2$, $T = 4$, and $K = 0.9$.

a grid with the strike placed precisely in the middle of two grid points removes oscillations (see Section 4.7 of their paper for details). However, their grid design is not directly applicable to the present context since, in this chapter, the grid represents the state-space of the short-rate process $R^{(m)}$, and the option (and the strike) depends on the zero-bond price $P^{(m)}(\cdot, T)$, whose approximation also depends on the grid design. In the context of tree methods approximation, a detailed study of the oscillatory behavior of European vanilla options has been performed in **?**, whereas Tavella and Randall (2000), Chapter 5, observes that convergence oscillation of the finite difference method can be reduced in a non-uniform grid design when the strike is placed midway between two grid points. Further investigation into how grid design can improve convergence is left for future research. Finally, since Assumption 4.4.1 is not satisfied under the Hull-White model, the preceding experiments show that the results of Section 4.4 can hold under less restrictive conditions for a specific set of parameters[14].

### 4.6.2 Approximation of Callable/Putable Bond Prices

We now examine the accuracy of Proposition 4.4.3 in approximating callable/putable bonds under the Hull–White model. Accordingly, we consider a coupon-bearing bond with semi-annual coupons that mature in $4$ years, $T = 4$. The coupon rate, denoted by $\alpha$, is set to $5\%$ per annum compounded semi-annually. The notional of the debt is set to $F = 100$, and we assume that it can be called at any time between the second

[14] Some testing has also been performed for the EV+ model, and the results are available upon request. These experiments also indicate that Assumption 4.4.1 could potentially be relaxed.

Table 4.7: Approximation of the convergence rate of the price of zero-coupon bond call options, Proposition 4.4.2 and Corollary 4.9, under the Hull–White and CIR++ models, respectively. Benchmark prices are calculated using closed-form analytical formulas. Except for the number of grid points, model and CTMC parameters are as listed in Table 4.5. Zero-coupon bond call option parameters using the notation of Proposition 4.4.2: $t_{n_1} = 0$, $t_{n_2} = 2$, and $T = 4$.

(a) Hull–White model

| m | Abs. error | Rate |
|---|---|---|
| 60 | 1.20E-04 | - |
| 80 | 5.36E-05 | 2.791 |
| 100 | 2.98E-05 | 2.637 |
| 120 | 2.64E-05 | 0.661 |
| 160 | 1.07E-05 | 3.130 |

(b) CIR++ model

| m | Abs. error | Rate |
|---|---|---|
| 60 | 8.19E-08 | - |
| 80 | 4.64E-08 | 1.977 |
| 100 | 2.45E-08 | 2.854 |
| 120 | 1.71E-08 | 1.973 |
| 160 | 8.61E-09 | 2.387 |

and fourth year for no additional cost, that is, $K_t^c = 100$ for $2 \leq t \leq T$, and we let $K_t^c \to \infty$ when $t < 2$ (as exercise is not allowed). Moreover, since there is no put feature, $K^p := K_t^p = 0$ for $0 \leq t \leq T$. Finally, we assume that accrued interest is paid to the bondholder upon redemption[15]. The results are summarized in Table 4.8. The column "CTMC" shows the approximated value of the debt using CTMCs with $m = 160$ as specified in Table 4.5, whereas the column "Benchmark" shows the CTMC approximated value with $m = 350$. The value using the tree method of Hull and White (1994) is reported in the column "Tree".

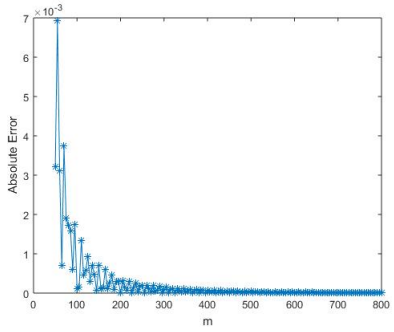



Figure 4.3: Convergence pattern of the approximated price of callable bonds, Proposition 4.4.3, as the number of grid points $m$ increases. Benchmark is calculated using CTMC approximation with $m = 1,000$. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.5. Contract specifications are $F = 100$, $\alpha = 0.05$, $T = 4$, $K_t^c = 100$, and $K^p = 0$, with the call option exercise window starting from $t = 2$ to $T$.

[15] This is a standard assumption in practice, meaning that the call price $K_t^c$ is increased by accrued interest upon redemption.

Table 4.8: Accuracy of Proposition 4.4.3 in approximating the price of callable bonds under the Hull–White model. Benchmark is calculated using CTMC approximation with $m = 350$. Model and CTMC parameters are as listed in Table 4.5. For the tree method, we use $252$-time steps per year. Contract specifications are $F = 100$, $\alpha = 0.05$, $T = 4$, $K_t^c = 100$, and $K^p = 0$, with the call option exercise window starting from $t = 2$ to $T = 4$.

| $\kappa$ | Benchmark | CTMC | Rel. Error | Elapsed Time (sec) | Tree | Rel. Error | Elapsed Time (sec) |
|---|---|---|---|---|---|---|---|
| **0.5** | **91.6418214** | 91.6426071 | 8.57E-06 | 0.0936 | 91.6070243 | 3.80E-04 | 0.0411 |
| **1** | **95.6073132** | 95.6072144 | 1.03E-06 | 0.0897 | 95.5614168 | 4.80E-04 | 0.0301 |
| **2** | **98.5647947** | 98.5648058 | 1.13E-07 | 0.0961 | 98.5108693 | 5.47E-04 | 0.0167 |
| **3** | **99.7010757** | 99.7004280 | 6.50E-06 | 0.0916 | 99.6443817 | 5.69E-04 | 0.0157 |
| $\sigma$ | **Benchmark** | **CTMC** | **Rel. Error** | **Elapsed Time (sec)** | **Tree** | **Rel. Error** | **Elapsed Time (sec)** |
| **0.1** | **98.9838248** | 98.9818310 | 2.01E-05 | 0.0780 | 98.9454895 | 3.87E-04 | 0.0238 |
| **0.2** | **95.6073132** | 95.6072144 | 1.03E-06 | 0.0927 | 95.5614168 | 4.80E-04 | 0.0258 |
| **0.3** | **92.2381361** | 92.2382928 | 1.70E-06 | 0.1011 | 92.1871120 | 5.53E-04 | 0.0244 |
| **0.4** | **88.9338108** | 88.9340669 | 2.88E-06 | 0.1198 | 88.8822619 | 5.80E-04 | 0.0289 |

We observe that the approximated values from the CTMC and tree methods are close to each other, confirming the adequacy of Proposition 4.4.3. Additionally, the relative error of the CTMC approximation is lower than that of the tree method. However, the tree method is shown to be 2 to 5 times faster than the CTMC approximation. The efficiency of the two approaches is shown in the context of zero-coupon bond option pricing in Figure 4.1.

The convergence pattern of the approximated callable price as $m$ grows is displayed in Figure 4.3, whereas Table 4.9 shows the convergence rates. The absolute error decreases rapidly to $0$ but exhibits a sawtooth pattern. As mentioned in Section 4.6.1, previous work in different contexts has shown that grid design can improve convergence and that placing the strike midway between two grid points can reduce or remove oscillatory behavior (see Tavella and Randall (2000) and Zhang and Li (2019) for details). However, the methodologies proposed in these studies are not directly applicable to the present context. Therefore, further investigation into how grid design can improve convergence is left for future research.

Table 4.9: - Approximation of the convergence rate of the price of callable bonds, Proposition 4.4.3, as the number of grid points $m$ increases. Benchmark is calculated using CTMC approximation with $m = 1,000$. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.5. Contract specifications are $F = 100$, $\alpha = 0.05$, $T = 4$, $K_t^c = 100$, and $K^p = 0$, with the call option exercise window starting from $t = 2$ to $T$.

| m | Abs. error | Rate |
|---|---|---|
| 50 | 3.21E-03 | - |
| 80 | 1.73E-03 | 1.31 |
| 120 | 5.81E-04 | 2.70 |
| 140 | 4.75E-04 | 1.31 |
| 175 | 2.47E-04 | 2.93 |

### 4.6.3 Approximation of Convertible Bond Prices

We now investigate the accuracy and efficiency of Proposition 4.5.3 in approximating CB prices under the Black–Scholes–Hull–White model, as well as the numerical convergence of the price estimates. That is, we suppose that the stock price dynamics follow a geometric Brownian motion with stochastic interest rate satisfying

$$\begin{aligned} \mathrm{d}S_t &= (R_t - q_t)S_t\,\mathrm{d}t + \widetilde{\sigma}_S S_t\,\mathrm{d}W_t^{(1)}, \\ \mathrm{d}R_t &= (\theta(t) - \kappa R_t)\,\mathrm{d}t + \widetilde{\sigma}_R\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{4.31}$$

with $\kappa$, $\widetilde{\sigma}_S$, $\widetilde{\sigma}_R$>0, and $[W^{(1)}, W^{(2)}]_t = \rho t$, $\rho \in [-1, 1]$.

From Lemma 4.1, we find that $f(r) = \frac{\widetilde{\sigma}_S}{\widetilde{\sigma}_R} r$. The dynamics of the auxiliary process $X_t = \ln(S_t) - \rho f(R_t)$ can then be derived as

$$\begin{aligned} \mathrm{d}X_t &= \mu_X(t, R_t)\,\mathrm{d}t + \sigma_X(R_t)\,\mathrm{d}W_t^{\star} \\ \mathrm{d}R_t &= (\theta(t) - \kappa R_t)\,\mathrm{d}t + \widetilde{\sigma}_R\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{4.32}$$

with $\mu_X(t, R_t) = R_t - q_t - \frac{\widetilde{\sigma}_S^2}{2} - \rho\frac{\widetilde{\sigma}_S}{\widetilde{\sigma}_R}(\theta(t) - \kappa R_t)$, $\sigma_X = \widetilde{\sigma}_S\sqrt{1-\rho^2}$, and $X_0 = \ln(S_0) - \rho f(R_0)$.

Unless stated otherwise, the model parameters for the short-rate process are the same as those used in previous examples, summarized in Table 4.5. Recall also that function $\theta$ is calibrated to the market zero-bond curve in Table 4.4 using Algorithm 7, as explained in Section 4.4.4. Unless stated otherwise, we suppose that $\widetilde{\sigma}_S = 0.2$, $q_t = 0.02$ for all $t \in [0, T]$, and $\rho = -0.2$. The model parameters are summarized in Table 4.10.

Table 4.10: Model parameters

| Model | $R_0$ | $q_t$ | $\kappa$ | $\widetilde{\sigma}_R$ | $S_0$ | $\widetilde{\sigma}_S$ | $\rho$ |
|---|---|---|---|---|---|---|---|
| Black–Scholes–Hull–White | $0.04$ | 0.02 | $1$ | $0.20$ | $100$ | $0.2$ | $-0.2$ |

The grid used to approximate the short-rate process, $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$, and the auxiliary process $\mathcal{S}_X^{(M)} = \{x_1, x_2, \ldots, x_M\}$, are constructed using the methodology of Tavella and Randall (Tavella and Randall (2000), Chapter 5), as explained at the beginning of this section, with $\widetilde{\alpha}_R$ (resp. $\widetilde{\alpha}_X$) representing the non-uniformity parameter of the grid of $R^{(m)}$ (resp. $X^{(m)}$). Unless otherwise indicated, all numerical experiments are conducted using the CTMC parameters listed in Table 4.11. Note also that the fast versions of Propositions 4.5.1 and 4.5.3 reported in Appendix 4.11 have been used in all numerical examples.

Table 4.11: CTMC parameters

| Model | $m$ | $M$ | $r_1$ | $r_m$ | $\widetilde{\alpha}_R$ | $x_1$ | $x_M$ | $\widetilde{\alpha}_X$ | $\Delta_N$ |
|---|---|---|---|---|---|---|---|---|---|
| Black–Scholes–Hull–White | $160$ | $160$ | $-30R_0$ | $25R_0$ | $0.5$ | $0.64X_0$ | $1.42X_0$ | $2$ | $1/252$ |

For the testing, we consider a convertible bond that pays semi-annual coupons with an annual rate $\alpha = 0.05$ and has a notional value $F = 100$. We suppose that the bond can be converted at any time from inception to maturity ($T = 1$) at a conversion rate $\eta = 1$. Under this set of parameters and assuming that both the dividend yield and credit risk are nil ($q_t = c_t = 0$ for all $t \in [0, T]$), the valuation of American-style CBs simplifies to the valuation of European-style convertible bonds, as stated in Proposition 4.5.2 and generalized to coupon paying bonds in Proposition 4.4. In that particular case, the results of Proposition 4.12.1[16] can thus serve as a benchmark in our analysis. When $q_t, c_t > 0$ for some $t \in [0, T]$, the benchmark is calculated using CTMC approximation with $M = 300$. All other parameters are as stated in Table 4.11. The results are summarized in Table 4.12. We note that the model achieves a high level of accuracy across all model parameters, with an average calculation time of less than 10 seconds. When the short-rate process is time-homogeneous, the matrix exponential can be calculated only once at the beginning of the procedure, which speeds up the procedure significantly. For instance, under the Black–Scholes–Vasicek model, the average calculation time for the CTMC approximated prices is less than 1.7 seconds. Numerical results for that particular model are reported in Appendix 4.12.2.

[16] The expected present value of future coupons should be added to the formula obtained in Proposition 4.12.1.

Table 4.12: Accuracy of the approximation of American-style CB prices, Algorithm 11, under Black–Scholes–Hull–White model. Model and CTMC parameters are as listed in Tables 4.10 and 4.11, respectively. Contract specifications are $F = 100$, $T = 1$, and $\eta = 1$, with an annual coupon rate $\alpha = 0.05$ paid semi-annually.

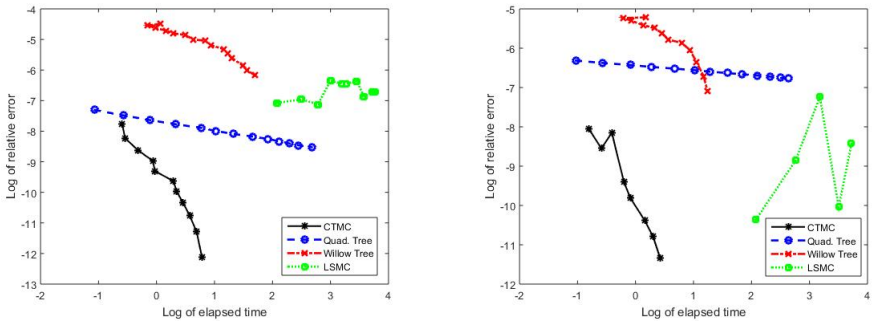
(a) $q_t = c_t = 0$ for all $t \in [0, T]$

| $S_0$ | CTMC | Benchmark | Rel. error |
|---|---|---|---|
| **90** | 105.15488 | 105.15732 | 2.32E-05 |
| **95** | 107.56087 | 107.57350 | 1.17E-04 |
| **100** | 110.48565 | 110.50458 | 1.71E-04 |
| **105** | 113.87792 | 113.89940 | 1.89E-04 |
| **110** | 117.66876 | 117.68977 | 1.79E-04 |
| $\sigma_S$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.1** | 107.36113 | 107.38133 | 1.88E-04 |
| **0.15** | 108.82431 | 108.84084 | 1.52E-04 |
| **0.2** | 110.48565 | 110.50458 | 1.71E-04 |
| **0.3** | 114.04675 | 114.07719 | 2.67E-04 |
| **0.4** | 117.72017 | 117.76496 | 3.80E-04 |
| $\rho$ | **CTMC** | **Benchmark** | **Rel. error** |
| **-0.3** | 110.20950 | 110.22594 | 1.49E-04 |
| **-0.2** | 110.48565 | 110.50458 | 1.71E-04 |
| **0.2** | 111.50772 | 111.53335 | 2.30E-04 |
| **0.3** | 111.74601 | 111.77262 | 2.38E-04 |

(b) $q_t = 0.02$, $c_t = 0.05$ for all $t \in [0, T]$

| $S_0$ | CTMC | Benchmark | Rel. error |
|---|---|---|---|
| **90** | 101.96277 | 101.93530 | 2.69E-04 |
| **95** | 104.84800 | 104.82384 | 2.30E-04 |
| **100** | 108.25070 | 108.21547 | 3.26E-04 |
| **105** | 112.07490 | 112.03748 | 3.34E-04 |
| **110** | 116.24312 | 116.19981 | 3.73E-0 |
| $\sigma_S$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.1** | 105.52046 | 105.48338 | 3.52E-04 |
| **0.15** | 106.75571 | 106.71897 | 3.44E-04 |
| **0.2** | 108.25070 | 108.21547 | 3.26E-04 |
| **0.3** | 111.58183 | 111.55168 | 2.70E-04 |
| **0.4** | 115.08975 | 115.06539 | 2.12E-04 |
| $\rho$ | **CTMC** | **Benchmark** | **Rel. error** |
| **-0.3** | 107.99824 | 107.96134 | 3.42E-04 |
| **-0.2** | 108.25070 | 108.21547 | 3.26E-04 |
| **0.2** | 109.19534 | 109.16259 | 3.00E-04 |
| **0.3** | 109.41008 | 109.38793 | 2.03E-04 |

Figure 4.4 shows the efficiency of the methodology compared to other recently developed numerical approaches. For this testing, we considered zero-coupon CBs ($\alpha = 0$), and we set $\theta(t) = \kappa R_0$ for all $t \in [0, T]$, such that the short-rate process collapses to the Vasicek model with a long-term mean level equal to $R_0$. We compare the CTMC approximated prices to the Willow tree approach ("Willow Tree") of Lu and Xu (2017), the quadrinomial tree ("Quad. Tree") of Battauz and Rotondi (2022), and the LSMC method of Longstaff and Schwartz (2001). All methodologies listed above have been adapted to incorporate credit risk as in the work of Tsiveriotis and Fernandes (1998) for a better comparison. When both credit spread and dividend yield are set to nil, the benchmark is obtained using the closed-form formula derived in Proposition 4.12.1; otherwise, CTMC approximation is used as a benchmark with $M = 1,000$, and all other CTMC parameters are as listed in Table 4.11. Figure 4.4 clearly shows the high efficiency of the CTMC methodologies compared to other methods. CTMC approximation significantly outperforms these other techniques in terms of both

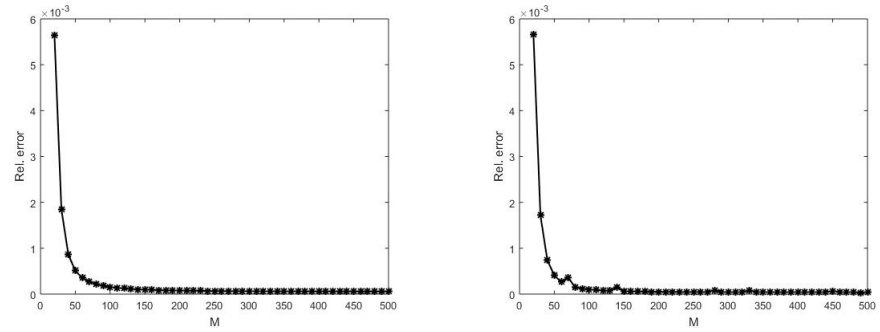



(a) $q_t = c_t = 0$ for all $t \in [0, T]$ (b) $q_t = 0, c_t = 0.05$ for all $t \in [0, T]$

Figure 4.4: Efficiency of the CTMC method in approximating CB prices, Algorithm 11, under the Black—Scholes–Vasicek model. Except for the number of grid points $M$, $\Delta_N = 1/100$ and $\theta(t) = \kappa R_0$ for all $t \in [0, T]$, the model and CTMC parameters are as listed in Tables 4.10 and 4.11, respectively. Contract specifications are $F = 100, T = 1$, $\eta = 1$, and $\alpha = 0$.

precision and calculation time.

The convergence pattern of the approximation as $M$ increases is illustrated in Figure 4.5, whereas Table 4.13 shows the convergence rate.




(a) $q_t = c_t = 0$ for all $t \in [0, T]$ (b) $q_t = 0.02, c_t = 0.05$ for all $t \in [0, T]$

Figure 4.5: Convergence pattern of the approximated CB prices using CTMC method, Algorihtm 11, under Black—Scholes–Hull–White model. Except for the number of grid points $M$ of the auxiliary process, the model and CTMC parameters are as listed in Tables 4.10 and 4.11, respectively. Contract specifications are $F = 100, T = 1$, and $\eta = 1$, with an annual coupon rate $\alpha = 0.05$ paid semi-annually.

Table 4.13: Approximation of the convergence rate of the approximated CB prices using CTMC method, Algorithm 11, under Black–Scholes–Hull-White model. Except for the number of grid points $M$ of the auxiliary process, the model and CTMC parameters are as listed in Tables 4.10 and 4.11, respectively. Contract specifications are $F = 100$, $T = 1$, and $\eta = 1$, with an annual coupon rate $\alpha = 0.05$ paid semi-annually.

(a) $q_t = c_t = 0$ for all $t \in [0, T]$

| m | Rel. error | Rate |
|---|---|---|
| 20 | 5.50E-03 | - |
| 50 | 5.59E-04 | 2.494 |
| 100 | 2.21E-04 | 1.340 |
| 150 | 1.65E-04 | 0.714 |
| 500 | 1.25E-04 | 0.231 |

(b) $q_t = 0.02$, $c_t = 0.05$ for all $t \in [0, T]$

| m | Rel. error | Rate |
|---|---|---|
| 20 | 5.38E-03 | - |
| 50 | 3.43E-04 | 3.004 |
| 100 | 5.17E-05 | 2.731 |
| 150 | 1.30E-05 | 3.396 |
| 500 | 2.45E-06 | 1.390 |

We observe that the approximated prices converge rapidly and smoothly to the benchmark prices, and the approximation achieves superquadratic convergence on average when accounting for both the dividend yield and credit risk.

Appendix 4.12.2 shows similar results under the Black–Scholes–Vasicek model. Analogous numerical analysis has also been performed under the Black–Scholes–CIR model with similar results, which are available upon request. Finally, since Assumption 4.4.1 is not satisfied under the Black-Scholes–Hull–White model, the preceding experiments demonstrate that the results of Section 4.5 can still be valid under less restrictive conditions under a certain set of parameters. Theoretical proof is left as future research.

## 4.7 Conclusion

In this chapter, we provide a general pricing framework based on continuous-time Markov chain approximations to value debt securities under general time-inhomogenous short-rate models. One advantage of CTMC methods over other commonly used numerical techniques is that they allow for a closed-form matrix expression for the price of zero-coupon bonds regardless of the complexity of the short-rate dynamics, making the calibration of the approximated model to the current market-term structure straightforward for a wide range of models. Closed-form matrix expressions are also obtained for the price of bond options and European CBs, and simple and efficient algorithms are developed to approximate the values of callable/putable bonds and CBs (American-style). Numerical results show the high accuracy and great ef-

ficiency of the methodology. Theoretical convergence is also discussed. The problem of CB pricing is also studied from a theoretical perspective. When both credit spread and dividend yield are set to nil and under some conditions on the parameters of the equity process (constant volatility), we show that early conversion is sub-optimal. When default risk is considered, we obtain lower and upper bounds for the value of American-style CBs.

Alternative approaches to credit risk modeling, such as the one of Hung and Wang (2002) and Chambers and Lu (2007) or Milanov *et al*. (2013), can also be explored using CTMCs. The methodologies proposed by Hung and Wang (2002) and Chambers and Lu (2007) can be easily integrated. However, Milanov *et al*. (2013) introduce default risk by incorporating a jump into the equity process. Jump processes in CTMCs have been studied in a one-dimensional setting by Lo and Skindilias (2014), while Kirkby (2023), Appendix B, discusses some extensions to two-dimensional processes. Other developments to the present work can also be considered. For example, building on insights from Brigo and Mercurio (2006), Sections 2.6.1 and 3.3.2., the closed-form expression obtained for the price of bond options in (4.14) and Remark 4.4.5 can be used to approximate other derivatives like caps, floors, and swaptions. The pricing of these derivatives, as well as the associated model calibration, should be studied in greater detail.

## 4.8 Appendix - Proofs

### 4.8.1 Proof of Lemma 4.2

Without loss of generality, suppose that $i = 0$ and that $\widetilde{N} = N$. Since $R^{(m)}$ is a discrete random process whose transitional probabilities $p_{k_1k_2}(t_{n-1}, t_n)$, $1 \leq k_1, k_2 \leq m$, $1 \leq n \leq N$ are given by $\mathbf{P}(t_{n-1}, t_n) = \exp\left\{\mathbf{Q}_n^{(m)}\Delta_N\right\}$ as per (4.3), we have that

$$
\begin{aligned}
&\mathbb{E}\left[e^{-\sum_{n=1}^{N} R_{t_n}^{(m)}\Delta_N}\big|R_0^{(m)} = r_j\right] \\
&\quad= \sum_{k_1,k_2,\ldots,k_N=1}^{m} e^{-\sum_{l=1}^{N} r_{k_l}\Delta_N} p_{jk_1}(t_0,t_1)p_{k_1k_2}(t_1,t_2)\ldots p_{k_{N-1}k_N}(t_{N-1},t_N) \\
&= \sum_{k_1,k_2,\ldots,k_N=1}^{m} p_{jk_1}(t_0,t_1)e^{-r_{k_1}\Delta_N}p_{k_1k_2}(t_1,t_2)e^{-r_{k_2}\Delta_N}\ldots p_{k_{N-1}k_N}(t_{N-1},t_N)e^{-r_{k_N}\Delta_N} \\
&= \mathbf{e}_j\left(\prod_{n=1}^{N}\mathbf{P}(t_{n-1},t_n)e^{-\mathbf{D}_m\Delta_N}\right)\mathbf{1}_{m\times 1} \\
&= \mathbf{e}_j\left(\prod_{n=1}^{N}e^{\mathbf{Q}_n^{(m)}\Delta_N}e^{-\mathbf{D}_m\Delta_N}\right)\mathbf{1}_{m\times 1},
\end{aligned}
$$

where the fourth line is the matrix representation of the third line. The result follows similarly when $\widetilde{N} = kN$. The proof for (4.12) follows using analogous arguments. □

### 4.8.2 Proof of Proposition 4.4.1

Using the notation of Lemma 4.2, we have that

$$\begin{aligned}
&\mathbb{E}\left[e^{-\int_{t_i}^{T} R_s^{(m)}\,\mathrm{d}s}\big|R_{t_i}^{(m)} = r_j\right] \\
&\qquad = \mathbb{E}\left[\lim_{\widetilde{N}\to\infty} e^{-\sum_{n=ki+1}^{\widetilde{N}} R_{\widetilde{t}_n}^{(m)}\Delta_{\widetilde{N}}}\big|R_{\widetilde{t}_{ki}}^{(m)} = r_j\right] \\
&\qquad = \lim_{\widetilde{N}\to\infty} \mathbb{E}\left[e^{-\sum_{n=ki+1}^{\widetilde{N}} R_{\widetilde{t}_n}^{(m)}\Delta_{\widetilde{N}}}\big|R_{\widetilde{t}_{ki}}^{(m)} = r_j\right] \text{ (by the dominated convergence theorem)} \\
&\qquad = \lim_{\widetilde{N}\to\infty} \mathbf{e}_j\left(\prod_{n=i+1}^{N}\left(e^{\mathbf{Q}_n^{(m)}\Delta_{\widetilde{N}}}e^{-\mathbf{D}\Delta_{\widetilde{N}}}\right)^{\frac{\widetilde{N}}{N}}\right)\mathbf{1}_{m\times 1} \text{ (by (4.11))} \\
&\qquad = \mathbf{e}_j\left(\prod_{n=i+1}^{N} e^{(\mathbf{Q}_n^{(m)}-\mathbf{D})\Delta_N}\right)\mathbf{1}_{m\times 1} \text{ (by the Lie product formula).}
\end{aligned}$$

### 4.8.3 Extension of Proposition 4.5.2 and Corollary 4.3

This section shows that the results of Proposition 4.5.2 and Corollary 4.3 still hold when periodic coupons are paid. We suppose that $\widetilde{N} > 0$ periodic coupons $\alpha > 0$ are paid at time $0 < t_z < t_{2z} < \ldots < t_{z\widetilde{N}} = T$, with $z = N/\widetilde{N}$. The time-$t$ risk-neutral value of a European-style CB is given by

$$\begin{aligned}
v_e^{\alpha}(t,x,r) &:= \mathbb{E}\Big[e^{-\int_t^T R_u\,\mathrm{d}u}\eta S_T\mathbf{1}_{\{S_T\geq F/\eta\}} + e^{-\int_t^T R_u+c_u\,\mathrm{d}u}F\mathbf{1}_{\{S_T<F/\eta\}}\big|S_t = x, R_t = r\Big] \\
&\qquad + \alpha\sum_{n=1}^{\widetilde{N}}\mathbb{E}\left[e^{-\int_t^{t_{zn}} R_u+c_u\,\mathrm{d}u}\Big|S_t = x, R_t = r\right]\mathbf{1}_{\{t_{zn}>t\}} \qquad (4.33) \\
&= v_e(t,x,r) + \alpha\sum_{n=1}^{\widetilde{N}} e^{-\int_t^{t_{zn}} c_u\,\mathrm{d}u}P_r(t,t_{zn})\mathbf{1}_{\{t_{zn}>t\}},
\end{aligned}$$

where $P_r(t,t_{zn}) := \mathbb{E}\left[e^{-\int_t^{t_{zn}} R_u\,\mathrm{d}u}\big|R_t = r\right]$ with $P(t,t_{zn}) = 1$ whenever $t_{zn} \leq t$. Hence, the value of European-style CBs with coupons is equal to the value of European-style CBs without coupons to which the present value of future coupons is added. When conversion can occur at any time prior to maturity, the

time-$t$ risk-neutral value of the CB is given by

$$v^{\alpha}(t,x,r) := \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\Bigg[ e^{-\int_t^{\tau} R_u + c_u \mathbf{1}_{\{\tau=T, S_T<F/\eta\}}\,\mathrm{d}u}\varphi(\tau,S_\tau) + \alpha\sum_{n=1}^{\widetilde{N}} e^{-\int_t^{t_{zn}} R_u + c_u\,\mathrm{d}u}\mathbf{1}_{\{t<t_{zn}\leq\tau\}}\Big| S_t = x, R_t = r\Bigg], \tag{4.34}$$

with function $\varphi$ defined in (4.27).

The next corollary extends Proposition 4.5.2 to coupon-bearing bonds. Under the assumption that no dividend yield is paid out and credit risk is nil, the value of coupon-bearing American-type CBs is equivalent to the value of coupon-bearing European-type CBs.

**Corollary 4.4** *Assume $q_t = c_t = 0$ for all $t \in [0,T]$ and $\sigma_S(r) = \widetilde{\sigma}_S > 0$ for all $r \in \mathcal{S}_R$. We have that $v^{\alpha}(t,x,r) = v_e^{\alpha}(t,x,r)$ for all $(t,x,r) \in [0,T]\times\mathbb{R}_+^{\star}\times\mathcal{S}_R$.*

*Proof. Define process $\left\{\widetilde{Z}_s := \alpha\sum_{n=1}^{\widetilde{N}} e^{-\int_t^{t_{zn}} R_u\,\mathrm{d}u}\mathbf{1}_{\{t<t_{zn}\leq s\}}\right\}_{t\leq s\leq T}$, representing the discounted value of future coupons, and note that*

$$\begin{aligned} v^{\alpha}(t,x,r) &\leq \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\left[ e^{-\int_t^{\tau} R(u)\,\mathrm{d}u}\varphi(\tau,S_\tau)\Big| S_t = x, R_t = r\right] \\ &\qquad + \sup_{\tau\in\mathcal{T}_{t,T}} \mathbb{E}\left[\widetilde{Z}_\tau\Big| S_t = x, R_t = r\right] \\ &= v_e(t,x,r) + \mathbb{E}\left[\widetilde{Z}_T\Big| S_t = x, R_t = r\right] \\ &= v_e^{\alpha}(t,x,r), \end{aligned}$$

*where the second equality follows from Proposition 4.5.2 and by noticing that $\widetilde{Z}_s \uparrow \widetilde{Z}_T$. We conclude the proof by observing that $v^{\alpha}(t,x,r) \geq v_e^{\alpha}(t,x,r)$ for all $(t,x,r) \in [0,T]\times\mathbb{R}_+^{\star}\times\mathcal{S}_R$, since $T \in \mathcal{T}_{t,T}$.* □

Let $\widetilde{v}_e^{\alpha} : [0,T]\times\mathbb{R}_+^{\star}\times\mathcal{S}_R \to \mathbb{R}_+$ denote the value of European-style CBs (with periodic coupon $\alpha$) when $c_t = 0$ for all $t \in [0,T]$. Using (4.33), it follows that

$$\begin{aligned} \widetilde{v}_e^{\alpha}(t,x,r) &:= \mathbb{E}\Big[e^{-\int_t^T R_u\,\mathrm{d}u}\max\left(\eta S_T, F\right)\Big| S_t = x, R_t = r\Big] + \alpha\sum_{n=1}^{\widetilde{N}} P_r(t,t_{zn})\mathbf{1}_{\{t_{zn}>t\}} \\ &= \widetilde{v}_e(t,x,r) + \alpha\sum_{n=1}^{\widetilde{N}} P_r(t,t_{zn})\mathbf{1}_{\{t_{zn}>t\}}, \end{aligned}$$

where $\widetilde{v}_e$ is defined in (4.28) and represents the value of European-style CBs when no coupons are paid and when $c_t = 0$ for all $t \in [0, T]$.

The next corollary extends the results of Corollary 4.3 to coupon-bearing bonds. It provides an upper bound for American-type CBs under the assumption that no dividends are distributed.

**Corollary 4.5** *Assume $q_t = 0$ for all $t \geq 0$ and $\sigma_S(r) = \widetilde{\sigma}_S > 0$ for all $r \in \mathcal{S}_R$. We have that $v_e^\alpha(t,x,r) \leq v^\alpha(t,x,r) \leq \widetilde{v}_e^\alpha(t,x,r)$ for all $(t,x,r) \in [0,T] \times \mathbb{R}_+^\star \times \mathcal{S}_R$.*

The proof is akin to that of Corollary 4.3.

## 4.9 Appendix - Time-Homogeneous Models

When the short-rate process is time-homogeneous (see, for instance, models listed in Table 4.1), the generator is time-independent, and the results obtained in Section 4.4 can be simplified. The construction of the generator remains the same as in Section 4.3.1, except that we now have $\mathbf{Q}^{(m)} := \mathbf{Q}_1^{(m)} = \mathbf{Q}_2^{(m)} = \ldots = \mathbf{Q}_N^{(m)}$, which follows because the drift and volatility parameters of $R$ are now time-independent. In this appendix, we summarize the different results obtained previously under the assumption that the short-rate process is time-homogeneous.

Recall that $\{\mathbf{e}_k\}_{k=1}^m$ denotes the standard basis in $\mathbb{R}^m$, that is, $\mathbf{e}_k$ represents a row vector of size $1 \times m$ with a value of $1$ in the $k$-th entry and $0$ elsewhere, $\mathbf{1}_{m\times 1}$ denotes an $m \times 1$ unit vector, and $\mathbf{D}_m := \text{diag}(\boldsymbol{r})$, is an $m \times m$ diagonal matrix with vector $\boldsymbol{r} = (r_1, r_2, \ldots, r_m)$ on its diagonal, $r_k \in \mathcal{S}_R^{(m)}$, $k = 1, 2 \ldots, m$.

### 4.9.1 Zero-Coupon Bond

The first result of this section concerns the price of zero-coupon bonds.

**Corollary 4.6** *Consider a time partition of $[0,T]$, $0 = \widetilde{t}_0 < \widetilde{t}_1 < \ldots < \widetilde{t}_{\widetilde{N}} = T$, with $\widetilde{N} \in \mathbb{N}$, $\Delta_{\widetilde{N}} = T/\widetilde{N}$ and $\widetilde{t}_n = n\Delta_{\widetilde{N}}$. Given that $R^{(m)}_{\widetilde{t}_i} = R_{\widetilde{t}_i} = r_j \in \mathcal{S}_R^{(m)}$, it holds that*

$$\mathbb{E}\left[e^{-\sum_{n=i+1}^{\widetilde{N}} R^{(m)}_{\widetilde{t}_n}\Delta_{\widetilde{N}}} \middle| R^{(m)}_{\widetilde{t}_i} = r_j\right] = \mathbf{e}_j \left(e^{\mathbf{Q}^{(m)}\Delta_{\widetilde{N}}} e^{-\mathbf{D}_m\Delta_{\widetilde{N}}}\right)^{\widetilde{N}-i} \times \mathbf{1}_{m\times 1}, \tag{4.35}$$

*and*

$$\begin{aligned}\mathbb{E}\left[e^{-\sum_{n=i}^{\widetilde{N}} R_{\widetilde{t}_n}^{(m)}\Delta_{\widetilde{N}}}\middle|R_{\widetilde{t}_i}^{(m)}=r_j\right] &= \mathbf{e}_j\left(e^{-\mathbf{D}_m\Delta_{\widetilde{N}}}e^{\mathbf{Q}^{(m)}\Delta_{\widetilde{N}}}\right)^{\widetilde{N}-i}e^{-\mathbf{D}_m\Delta_{\widetilde{N}}}\times\mathbf{1}_{m\times 1}\\ &= \mathbf{e}_j e^{-\mathbf{D}_m\Delta_{\widetilde{N}}}\left(e^{\mathbf{Q}^{(m)}\Delta_{\widetilde{N}}}e^{-\mathbf{D}_m\Delta_{\widetilde{N}}}\right)^{\widetilde{N}-i}\times\mathbf{1}_{m\times 1},\end{aligned} \tag{4.36}$$

*Moreover, if Assumption 4.4.1 holds, the price at time $t\geq 0$ of a zero-coupon bond with maturity $T\geq t$ can be approximated by*

$$P_j^{(m)}(t,T):=\mathbb{E}\left[e^{-\int_t^T R_s^{(m)}\,\mathrm{d}s}\middle|R_t^{(m)}=r_j\right]=\mathbf{e}_j e^{(\mathbf{Q}^{(m)}-\mathbf{D}_m)(T-t)}\mathbf{1}_{m\times 1}, \tag{4.37}$$

*given that $R_t^{(m)}=R_t=r_j\in\mathcal{S}_R^{(m)}$.*

The proof follows by setting $\widetilde{N}=N$ and $\mathbf{Q}^{(m)}=\mathbf{Q}_1^{(m)}=\mathbf{Q}_2^{(m)}=\ldots=\mathbf{Q}_N^{(m)}$ in Lemma 4.2, Remark 4.4.1, and Proposition 4.4.1. (4.37) was previously obtained by Cui *et al.* (2018) Proposition 8 (ii) and Kirkby (2023), Proposition 3, whereas the first equality of (4.36) is provided, in a more general form, in Cui *et al.* (2018) Proposition 8 (i). The proof in Appendix 4.8 differs from these previous proofs and provides a simple and intuitive way of obtaining these results using basic probabilistic arguments.

### 4.9.2 Zero-Coupon Bond Option

The next result concerns the price of European call and put options on zero-coupon bonds.

**Corollary 4.7** *Let Assumption 4.4.1 hold. Given that $R_{t_{n_1}}^{(m)}=r_j\in\mathcal{S}_R^{(m)}$, the price at $t_{n1}\geq 0$ of a European call (resp. put) option with maturity $t_{n_2}>t_{n_1}$ on a zero-coupon bond maturing at time $T>t_{n_2}$ with a strike $K>0$ can be approximated by*

$$\mathbb{E}\left[e^{-\int_{t_{n_1}}^{t_{n_2}} R_s^{(m)}\,\mathrm{d}s}h\left(P^{(m)}(t_{n_2},T)\right)\middle|R_{t_{n_1}}^{(m)}=r_j\right]=\mathbf{e}_j\left(e^{(\mathbf{Q}^{(m)}-\mathbf{D}_m)(t_{n_2}-t_{n_1})}\right)\mathbf{H}, \tag{4.38}$$

*where $h(x)=\max(x-K,0)$ (resp. $h(x)=\max(K-x,0)$) denotes the payoff function, $P^{(m)}(t_{n_2},T):=\mathbb{E}\left[e^{-\int_{t_{n_2}}^{T} R_s^{(m)}\,\mathrm{d}s}\middle|R_{t_{n_2}}^{(m)}\right]$ denotes the approximated zero-coupon bond price at $t_{n_2}$, and $\mathbf{H}$ denotes a column vector of size $m\times 1$ whose $k$-th, $h_k$, entry is given by*

$$h_k=h\left(P_k^{(m)}(t_{n_2},T)\right),$$

*with $P_k^{(m)}(t_{n_2},T)$ defined in* (4.37).

## 4.10 Appendix - Extended Models of Brigo and Mercurio (2006)

In this section, we summarize the results of Section 4.4 under particular time-inhomogeneous short-rate models, for which the short-rate process is obtained by a time-deterministic shift of a time-homogeneous auxiliary process. More precisely, we suppose that the short-rate process can be decomposed as

$$R_t = Y_t + \theta(t), \tag{4.39}$$

for $t \geq 0$, where $\theta$ denotes a continuous deterministic function of time and $Y$ denotes an auxiliary time-homogeneous diffusion process with the following dynamics:

$$\mathrm{d}Y_t = \mu_Y(Y_t)\,\mathrm{d}t + \sigma_Y(Y_t)\,\mathrm{d}W_t, \tag{4.40}$$

where $\mu_Y, \sigma_Y : \mathcal{S}_Y \to \mathbb{R}$ are well-behaved functions such that (4.40) has a unique in-law weak solution with $\mathcal{S}_Y$ the state-space of $Y$. Examples of such diffusion processes are listed in Table 4.3.

In this section, the auxiliary process $Y$ is approximated by a CTMC. We denote by $Y^{(m)}$ the CTMC approximation of $Y$ taking values in a finite state-space $\mathcal{S}_Y^{(m)} = \{y_1, y_2, \ldots, y_m\}$. The time-independent generator of $Y^{(m)}$, denoted by $\mathbf{Q}_Y^{(m)}$, is constructed as in (4.5), with functions $\mu_R$ and $\sigma_R$ replaced by functions $\mu_Y$ and $\sigma_Y$ of (4.40), respectively. Moreover, since the auxiliary process is time-homogenous, the generator of $Y^{(m)}$ does not depend on time, such that $\mathbf{Q}_Y^{(m)} := \mathbf{Q}_1^{(m)} = \mathbf{Q}_2^{(m)} = \ldots = \mathbf{Q}_N^{(m)}$. Using (4.39), the CTMC approximation of the short-rate process $R^{(m)}$ is given by

$$R_t^{(m)} = Y_t^{(m)} + \theta(t),$$

for $t \geq 0$, with $\theta(0) = 0$.

**Remark 4.10.1 (Weak convergence of the approximation)** *The weak convergence of $Y^{(m)}$ to $Y$ follows from Theorem 5.1 of Mijatović and Pistorius (2013). Then, since $R$ is a continuous transformation of $Y$, we conclude that $R^{(m)} \Rightarrow R$ by the continuous mapping Theorem.*

Recall that $\{\mathbf{e}_k\}_{k=1}^m$ denotes the standard basis in $\mathbb{R}^m$, that is, $\mathbf{e}_k$ represents a row vector of size $1 \times m$ with a value of $1$ in the $k$-th entry and $0$ elsewhere, $\mathbf{1}_{m\times 1}$ denotes an $m \times 1$ unit vector, and $\mathbf{D}_Y := \mathrm{diag}(\boldsymbol{y})$ denotes an $m \times m$ diagonal matrix with vector $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$ on its diagonal, with $y_k \in \mathcal{S}_Y^{(m)}$, $k = 1, 2, \ldots, m$.

### 4.10.1 Zero-Coupon Bond

The next corollary shows that the price of zero-coupon bonds inherits the analytical tractability of the homogeneous auxiliary process $Y$.

**Corollary 4.8** *Suppose that there exists $y^\star \in \mathcal{S}_Y$ such that $Y_t \geq y^\star$ for all $t \geq 0$. Then, the price at time $t \geq 0$ of a zero-coupon bond with maturity $T \geq t$ can be approximated by*

$$P_j^{(m)}(t,T) := \mathbb{E}[e^{-\int_t^T R_s^{(m)}\,\mathrm{d}s}\big|R_t^{(m)} = y_j + \theta(t)] = e^{-\int_t^T \theta(s)\,\mathrm{d}s}\widetilde{P}_j^{(m)}(t,T), \tag{4.41}$$

*with*

$$\widetilde{P}_j^{(m)}(t,T) := \mathbb{E}\left[e^{-\int_t^T Y_s^{(m)}\,\mathrm{d}s}\big|Y_t^{(m)} = y_j\right] = \mathbf{e}_j e^{\left(\mathbf{Q}_Y^{(m)} - \mathbf{D}_Y\right)(T-t)}\mathbf{1}_{m\times 1}, \tag{4.42}$$

*where $Y_t = Y_t^{(m)} = y_j \in \mathcal{S}_Y^{(m)}$.*

The proof follows directly from Corollary 4.6.

### 4.10.2 Zero-Coupon Bond Option

The next result concerns the price of European call and put options on zero-coupon bonds. Using the homogeneous property of the auxiliary process, we can obtain a simplified expression for (4.14).

**Corollary 4.9** *Suppose that there exists $y^\star \in \mathcal{S}_Y$ such that $Y_t \geq y^\star$ for all $t \geq 0$. Given that $R_{t_{n_1}}^{(m)} = y_j + \theta(t)$, with $y_j \in \mathcal{S}_Y^{(m)}$, the price at $t_{n_1} \geq 0$ of a European call (resp. put) option with maturity $t_{n_2} > t_{n_1}$ on a zero-coupon bond maturing at time $T > t_{n_2}$ with a strike $K > 0$ can be approximated by*

$$\mathbb{E}\left[e^{-\int_{t_{n_1}}^{t_{n_2}} R_s^{(m)}\,\mathrm{d}s} h\left(P^{(m)}(t_{n_2},T)\right)\Big|R_{t_{n_1}}^{(m)} = y_j + \theta(t)\right]$$
$$= \mathbf{e}_j e^{-\int_{t_{n_1}}^{t_{n_2}} \theta(s)\,\mathrm{d}s}\left(e^{\left(\mathbf{Q}_Y^{(m)} - \mathbf{D}_Y\right)(t_{n_2}-t_{n_1})}\right)\mathbf{H}, \tag{4.43}$$

*where $h(x) = \max(x-K,0)$ (resp. $h(x) = \max(k-x,0)$) denotes the payoff function with $P^{(m)}(t_{n_2},T) := \mathbb{E}\left[e^{-\int_{t_{n_2}}^T R_s^{(m)}\,\mathrm{d}s}\Big|R_{t_{n_2}}^{(m)}\right]$ and $\mathbf{H}$ denotes a column vector of size $m\times 1$ whose $k$-th entry, $h_k$, is given by*

$$h_k = h\left(P_k^{(m)}(t_{n_2},T)\right),$$

*with $P_k^{(m)}(t_{n_2},T)$ defined in (4.41).*

### 4.10.3 Calibration to the Initial Term Structure of Interest Rates

When the short-rate process is of the form $R_t = Y_t + \theta(t)$, as in Table 4.3, the fitting to the term structure of interest rates can be greatly simplified, since the function $\theta$ now appears explicitly in the zero-coupon bond formula (4.41).

Proposition 4.10.1 provides an explicit expression for the function $\theta$ that makes the model zero-coupon bond prices equal to the market prices when the short-rate process is of the form (4.39). As in Section 4.4.4, we assume that the time deterministic function $\theta$ is piecewise constant in time, such that

$$\theta(t) = \sum_{n=1}^{N} \theta_n \mathbf{1}_{[t_{n-1},t_n)}(t),$$

for some $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_N) \in \mathbb{R}^N$. The objective is thus to find parameters $\boldsymbol{\theta}$ such that $P_j^{(m)}(0,t_n) = P^\star(0,t_n)$ for $n = 1, 2, \ldots, N$, where $P^\star$ represents the market zero-coupon bond prices. Those parameters are called the calibrated parameters and are denoted by a star $\boldsymbol{\theta}^\star$.

**Proposition 4.10.1** *Suppose that there exists $y^\star \in \mathcal{S}_Y$ such that $Y_t \geq y^\star$ for all $t \geq 0$. Given $R_0 = R_0^{(m)} = y_j \in \mathcal{S}_Y^{(m)}$, we have that*

$$\theta_n^\star := \begin{cases} -\frac{1}{t_n} \ln\left(\frac{P^\star(0,t_n)}{\widetilde{P}_j^{(m)}(0,t_n)}\right) & \text{if } n = 1 \\ -\frac{1}{t_n - t_{n-1}} \ln\left(\frac{P^\star(0,t_n)}{P^\star(0,t_{n-1})} \frac{\widetilde{P}_j^{(m)}(0,t_{n-1})}{\widetilde{P}_j^{(m)}(0,t_n)}\right) & \text{if } n = 2, 3, \ldots, N, \end{cases} \tag{4.44}$$

*where $\widetilde{P}_j^{(m)}(0,\cdot)$ is defined in* (4.42).

The proof is straightforward from Corollary 4.8.

The calibrated parameters in (4.44) can easily be obtained by calculating the zero-coupon bond price at each time $\{t_1, t_2, \ldots, t_n\}$. This procedure involves calculating matrix exponentials at each time step, which can slow down the execution considerably. However, by taking advantage of the homogeneous property of $Y$, the matrix exponentials can be calculated only once at the beginning of the procedure, which makes it highly efficient. This is illustrated in Algorithm 8. In Algorithm 8, $\widetilde{\mathbf{P}}(t_n) := [\widetilde{P}_j^{(m)}(0,t_n)]_{j=1}^m$ is a column vector of size $m \times 1$, with $\widetilde{P}_j^{(m)}(0,t_n)$ defined in (4.42).

**Algorithm 8:** Calibration of $\boldsymbol{\theta}$ to the Current Market Term-Structure - Extended Models of Brigo and Mercurio (2006)

**Input:** Initialize $\mathbf{Q}_Y^{(m)}$ and let $t \mapsto P^\star(0,t)$ be the current market zero-bond curve

$N \in \mathbb{N}$, the number of time steps

$\Delta_N \leftarrow T/N$, the size of a time step

Set $t_n = n\Delta_N$, $n = 1, 2, \ldots, N$

Set $\mathbf{D}_Y \leftarrow \text{diag}(\boldsymbol{y})$ with $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)$, $y_k \in \mathcal{S}_Y^{(m)}$, $k = 1, 2, \ldots, m$

`/* Adjusted transition probability matrix of` $Y$ `over a period of length` $\Delta_N$ `*/`

$\mathbf{A}_{\Delta_N} \leftarrow e^{(\mathbf{Q}_Y^{(m)} - \mathbf{D}_Y)\Delta_N}$

`/* Calibration to the current market zero-bond curve` $t \mapsto P^\star(0,t)$ `*/`

Set $\widetilde{\mathbf{P}}(t_1) \leftarrow \mathbf{A}_{\Delta_N}\mathbf{1}_{m\times 1}$

Set $\theta_1^\star = -\frac{1}{t_1}\ln\left(\frac{P^\star(0,t_1)}{\mathbf{e}_j\widetilde{\mathbf{P}}(t_1)}\right)$

**for** $n = 2, \ldots, N$ **do**

  $\widetilde{\mathbf{P}}(t_n) \leftarrow \mathbf{A}_{\Delta_N}\widetilde{\mathbf{P}}(t_{n-1})$

  $\theta_n^\star = -\frac{1}{t_n - t_{n-1}}\ln\left(\frac{P^\star(0,t_n)}{P^\star(0,t_{n-1})}\frac{\mathbf{e}_j\widetilde{\mathbf{P}}(t_{n-1})}{\mathbf{e}_j\widetilde{\mathbf{P}}(t_n)}\right)$

**return** $\{\theta_n^\star\}_{n=1}^N$

## 4.11 Appendix - Algorithms

This section presents the results of Section 4.5 into an algorithm format. More precisely, Proposition 4.5.3 is provided in Algorithm 10. Using the results of Proposition 3.4, a fast version of Proposition 4.5.1 for the pricing of European-style CBs is also provided in Algorithm 9, whereas the fast-version of Proposition 4.5.3 for the pricing American-style CBs is provided in Algorithm 11.

### 4.11.1 European-Style Convertible Bond

Using the tower property of conditional expectations and Proposition 3.4 inspired from the work of Chapter 3, we present a new algorithm (Algorithm 9) that speeds up the pricing of European-style CBs. In Chapter 3, we work with stochastic volatility models. However, the reasoning behind the proof is the same for stochastic interest rate models as in the present context.

The following notation is used in Algorithm 9.

1. $\mathbf{B} = [b_{kl}]_{k,l=1}^{m,M}$ denotes a matrix of size $m \times M$, containing the value of the CB.

2. $\mathbf{B}_{*,l} = [b_{kl}]_{k=1}^{m}$ denotes the $l$-th column of $\mathbf{B}$, $l = 1, 2, \ldots, M$,

3. $\mathbf{B}_{k,*} = [b_{kl}]_{l=1}^{M}$ denotes the $k$-th row of $\mathbf{B}$, $k = 1, 2, \ldots, m$.

4. The symbol $\top$ indicates the matrix (vector) transpose operation.

---

**Algorithm 9:** European-style CB – Fast Algorithm

---

**Input:** Initialize $\mathbf{Q}_n^{(m)}$ as in (4.5), and $\mathbf{\Lambda}_k^{(n,M)}$ as in (4.8), for $k = 1, 2, \ldots, m$, $n = 1, 2, \ldots, N$

$N \in \mathbb{N}$, the number of time steps,

$\Delta_N \leftarrow T/N$, the size of a time step

1 For each $k \in \{1, 2, \ldots, m\}$, set

$$\mathbf{B}_{k,*} \leftarrow \left[\eta e^{x_l+\rho f(r_k)} \mathbf{1}_{\{e^{x_l+\rho f(r_k)} \geq F/\eta\}} + e^{-\int_0^T c_u \, \mathrm{d}u} F \mathbf{1}_{\{e^{x_l+\rho f(r_k)} < F/\eta\}}\right]_{l=1}^{M}$$

2 **for** $n = N-1, \ldots, 0$ **do**

3  **for** $k = 1, 2, \ldots, m$ **do**

4  $\mathbf{P}_{n,k}^X \leftarrow e^{\mathbf{\Lambda}_k^{(n+1,M)} \Delta_N} e^{-r_k \Delta_N}$

5  $\mathbf{E}_{*,k} \leftarrow \mathbf{P}_{n,k}^X \mathbf{B}_{k,*}^\top$

6  $\mathbf{P}_n^R \leftarrow e^{\mathbf{Q}_{n+1}^{(m)} \Delta_N}$

7  **for** $l = 1, 2, \ldots, M$ **do**

8  $\mathbf{B}_{*,l} \leftarrow \mathbf{P}_n^R \mathbf{E}_{l,*}^\top$

9 **return** $b_{ji}$

---

**Remark 4.11.1 (Extension to coupon-bearing bonds)** *Algorithm 9 is set up for zero-coupon CBs. However, extension to coupon-bearing bonds is straightforward. More precisely, when a coupon $\alpha > 0$ is paid at time $t_{n+1}$, then the column vector $\mathbf{E}_{*,k}$ at time $t_n$, line 5 of the algorithm, must be modified as follow*

$$\mathbf{E}_{*,k} \leftarrow \mathbf{P}_{n,k}^X \left(\mathbf{B}_{k,*}^\top + \alpha \mathbf{1}_{M\times 1}\right),$$

*for each $k \in \{1, 2, \ldots, m\}$.*

Proposition 3.4 allows the separation of the matrices $\mathbf{\Lambda}_k^{(n,M)}$ and $\mathbf{Q}_n^{(m)}$ at each time step $n \in \{1, 2, \ldots, N\}$. Hence, the matrix exponential of a large sparse matrix $\mathbf{G}_n^{(mM)}$ of size $mM \times mM$ is replaced by $m$ calcu-

lations of the exponential of an $M \times M$ matrix and one calculation of the exponential of an $m \times m$ matrix. Numerical experiments in Section 4.6 show the accuracy and efficiency of the fast algorithm empirically.

### 4.11.2 Convertible Bond (American-style)

Based on Proposition 4.5.3, Algorithm 10 provides the CTMC approximation for the value of a CB given that $X_0^{(m,M)} = \ln(S_0) - \rho f(R_0) = x_i \in \mathcal{S}_X^{(M)}$ and $R_0^{(m)} = R_0 = r_j \in \mathcal{S}_R^{(m)}$. The algorithm is set up for zero-coupon CBs with no additional features, such as call and put options. However, such extensions are straightforward and are discussed further below and in Remark 4.11.2.

Similar to the European-style CB, the performance of Algorithm 10 can be increased by assuming that the short-rate process is constant over small time periods (Proposition 3.4). Let $\widetilde{\mathbf{H}} := [\eta e^{x_l + \rho f(r_k)}]_{k,l=1}^{m,M}$ be an $m \times M$ matrix representing the conversion value. At each time step, matrices $\widetilde{\mathbf{B}}^E$, $\widetilde{\mathbf{B}}^{CO}$, and $\widetilde{\mathbf{B}}$, of size $m \times M$, contain the equity part, cash-only part, and the whole value of the CB, respectively. Furthermore, we denote by $\widetilde{b}_{ij}$ (resp $\widetilde{h}_{ij}$), the $(i,j)$-entry of $\widetilde{\mathbf{B}}$ (resp. $\widetilde{\mathbf{H}}$), and define the matrix indicator $\mathbf{1}_{\{\widetilde{\mathbf{B}}=\widetilde{\mathbf{H}}\}}$, where each element $(i,j)$ of the matrix, denoted $\mathbf{1}_{\{\widetilde{\mathbf{B}}=\widetilde{\mathbf{H}}\}}(i,j)$, is given by $\mathbf{1}_{\{\widetilde{\mathbf{B}}=\widetilde{\mathbf{H}}\}}(i,j) = \mathbf{1}_{\{\widetilde{b}_{ij}=\widetilde{h}_{ij}\}}$, for $1 \leq i \leq m$, $1 \leq j \leq M$. The fast Algorithm to value CBs is provided in Algorithm 11.

**Algorithm 10:** American-style CB

**Input:** Initialize $\mathbf{G}_n^{(mM)}$ as in (4.10) for $n = 1, 2, \ldots, N$, $\mathbf{H}^{CO}$ as in (4.29), and $\mathbf{H}_n^E$ as in (4.30), for

$$n = 0, 1, \ldots, N$$

$N \in \mathbb{N}$, the number of time steps,

$\Delta_N \leftarrow T/N$, the size of a time step

1 Set $\mathbf{D}_{mM} \leftarrow \text{diag}\,(\boldsymbol{d})$ with $\boldsymbol{d} = (d_1, d_2, \ldots, d_{mM})$, and $d_{(k-1)M+l} = r_k \in \mathcal{S}_R^{(m)}$, $k = 1, 2, , \ldots, m$, $l = 1, 2, , \ldots, M$

2 Set $\mathbf{B}_N^{CO} \leftarrow \mathbf{H}^{CO}$, $\mathbf{B}_N^E \leftarrow \mathbf{H}_N^E$, $\mathbf{B}_N \leftarrow \mathbf{B}_N^E + \mathbf{B}_N^{CO}$

3 **for** $n = N-1, N-2, \ldots, 0$ **do**

4 $\quad \mathbf{A}_{n+1} \leftarrow \exp\left\{\Delta_N \left(\mathbf{G}_{n+1}^{(mM)} - \mathbf{D}_{mM}\right)\right\}$,

5 $\quad \mathbf{B}_n^{CO} \leftarrow e^{-\int_{t_n}^{t_{n+1}} c_u \,\mathrm{d}u} \mathbf{A}_{n+1} \mathbf{B}_{n+1}^{CO}$, $\mathbf{B}_n^E \leftarrow \mathbf{A}_{n+1} \mathbf{B}_{n+1}^E$,

6 $\quad \mathbf{B}_n \leftarrow \max\left(\mathbf{H}_n^E, \mathbf{B}_n^E + \mathbf{B}_n^{CO}\right)$

7 $\quad \mathbf{B}_n^{CO} \leftarrow \mathbf{B}_n^{CO}\left(\mathbf{1}_{mM\times 1} - \mathbf{1}_{\{\mathbf{B}_n = \mathbf{H}_n^E\}}\right)$, $\mathbf{B}_n^E = \mathbf{B}_n - \mathbf{B}_n^{CO}$

8 $v^{(m,M)}(0, S_0, R_0) \leftarrow \mathbf{e}_{ji}\mathbf{B}_0$

9 **return** $v^{(m,M)}(0, S_0, R_0)$

**Remark 4.11.2** *The extension of Algorithms 10 and 11 to coupon-bearing bonds is straightforward. When a coupon $\alpha > 0$ is paid at time $t_{n+1}$, the continuation value at time $t_n$ of the cash-only part must be adjusted accordingly. Specifically, line 5 of Algorithm 10 should be updated to*

$$\mathbf{B}_n^{CO} \leftarrow e^{-\int_{t_n}^{t_{n+1}} c_u \,\mathrm{d}u} \mathbf{A}_{n+1}\left(\mathbf{B}_{n+1}^{CO} + \alpha \mathbf{1}_{mM\times 1}\right),$$

*and line 8 of Algorithm 11 should be changed to*

$$\mathbf{E}_{*,k}^{CO} \leftarrow e^{-\int_{t_n}^{t_{n+1}} c_u \,\mathrm{d}u} \mathbf{P}_{n,k}^X \left((\widetilde{\mathbf{B}}_{k,*}^{CO})^\top + \alpha \mathbf{1}_{M\times 1}\right),$$

*for each $k \in \{1, 2, \ldots, M\}$.*

## 4.12 Appendix - Supplemental Material

This document provides supplemental material to Chapter 4.

**Algorithm 11:** American-style CB – Fast Algorithm

**Input:** Initialize $\mathbf{Q}_n^{(m)}$ as in (4.5) and $\mathbf{\Lambda}_k^{(n,M)}$ as in (4.8), for $k = 1, 2, \ldots, m$, $n = 1, 2, \ldots, N$

$N \in \mathbb{N}$, the number of time steps,

$\Delta_N \leftarrow T/N$, the size of a time step

Set $\widetilde{\mathbf{H}} := [\eta e^{x_l + \rho f(r_k)}]_{k,l=1}^{m,M}$

Set $\widetilde{\mathbf{B}}_{k,*}^{E} \leftarrow \left[\eta e^{x_l+\rho f(r_k)} \mathbf{1}_{\{e^{x_l+\rho f(r_k)} \geq F/\eta\}}\right]_{l=1}^{M}$, $k = 1, 2, \ldots, m$

Set $\widetilde{\mathbf{B}}_{k,*}^{CO} \leftarrow \left[F \mathbf{1}_{\{e^{x_l+\rho f(r_k)} < F/\eta\}}\right]_{l=1}^{M}$, $k = 1, 2, \ldots, m$

$\widetilde{\mathbf{B}} \leftarrow \widetilde{\mathbf{B}}^{E} + \widetilde{\mathbf{B}}^{CO}$

**for** $n = N-1, \ldots, 0$ **do**

  **for** $k = 1, 2, \ldots, m$ **do**

    $\mathbf{P}_{n,k}^{X} \leftarrow e^{\mathbf{\Lambda}_k^{(n+1,M)} \Delta_N} e^{-r_k \Delta_N}$

    $\mathbf{E}_{*,k}^{CO} \leftarrow e^{-\int_{t_n}^{t_{n+1}} c_u \, \mathrm{d}u} \mathbf{P}_{n,k}^{X} (\widetilde{\mathbf{B}}_{k,*}^{CO})^{\top}$, $\mathbf{E}_{*,k}^{E} \leftarrow \mathbf{P}_{n,k}^{X} (\widetilde{\mathbf{B}}_{k,*}^{E})^{\top}$

  $\mathbf{P}_n^{R} \leftarrow e^{\mathbf{Q}_{n+1}^{(m)} \Delta_N}$

  **for** $l = 1, 2, \ldots, M$ **do**

    $\widetilde{\mathbf{B}}_{*,l}^{CO} \leftarrow \mathbf{P}_n^{R} (\mathbf{E}_{l,*}^{CO})^{\top}$, $\widetilde{\mathbf{B}}_{*,l}^{E} \leftarrow \mathbf{P}_n^{R} (\mathbf{E}_{l,*}^{E})^{\top}$, $\widetilde{\mathbf{B}} \leftarrow \max\left(\widetilde{\mathbf{H}}, \widetilde{\mathbf{B}}^{E} + \widetilde{\mathbf{B}}^{CO}\right)$

    $\widetilde{\mathbf{B}}^{CO} \leftarrow \widetilde{\mathbf{B}}^{CO}\left(\mathbf{1}_{m \times M} - \mathbf{1}_{\{\widetilde{\mathbf{B}} = \widetilde{\mathbf{H}}\}}\right)$, $\widetilde{\mathbf{B}}^{E} = \widetilde{\mathbf{B}} - \widetilde{\mathbf{B}}^{CO}$

**return** $b_{ji}$

### 4.12.1 Closed-Form Expression for European-style Convertible Bonds under TF approach

In the following, we derive a closed-form analytical formula for European-style CBs (or when the conversion option can only be exercised at maturity). We suppose that credit risk is incorporated into the model using the approach of Tsiveriotis and Fernandes (1998).

Accordingly, we make some simplifying assumptions. We suppose that the dynamics of the stock process are given by

$$\begin{aligned} \mathrm{d}S_t &= (R_t - q_t) S_t \, \mathrm{d}t + \widetilde{\sigma}_S S_t \left(\rho \, \mathrm{d}\widetilde{W}_t^{(2)} + \sqrt{1-\rho^2} \, \mathrm{d}\widetilde{W}_t^{(1)}\right), \\ \mathrm{d}R_t &= (\theta(t) - \kappa R_t) \, \mathrm{d}t + \widetilde{\sigma}_R \, \mathrm{d}\widetilde{W}_t^{(2)}, \end{aligned} \tag{4.45}$$

with $\widetilde{\sigma}_S, \widetilde{\sigma}_R > 0$, $\rho \in [-1, 1]$ and $\widetilde{W} = \{(\widetilde{W}_t^{(1)}, \widetilde{W}_t^{(2)})\}_{t \geq 0}$ is a standard bi-dimensional Brownian motion[17].

[17] The formulation in (4.45) in terms of independent Brownian motion is equivalent to that in (4.1) in terms of correlated Brownian motion. Indeed, define $Z_t = \rho \widetilde{W}_t^{(2)} + \sqrt{1-\rho^2} \widetilde{W}_t^{(1)}$, $t \geq 0$. From Cholesky decomposition, the process $(Z, \widetilde{W}^{(2)})$ is a correlated

Under this assumption, the short-rate process corresponds to the Hull–White model in Table 4.2. When function $\theta(\cdot)$ is constant over time, then the short-rate dynamics collapsed to the Vasicek model[18], whereas when $\kappa = 0$, the Ho–Lee model is obtained. Under these three models, the short-rate process is Gaussian, and the price of the zero-coupon bond can be obtained explicitly by

$$P(t,T) = e^{A(t,T)-B(t,T)R_t}, \tag{4.46}$$

for some time-deterministic functions $A$ and $B$ given in Table 4.14, see Björk (2009), Section 24.4 for details. Finally, note that when $\sigma_R(\cdot) = \kappa = \theta(\cdot) = 0$, then the short-rate process is constant to $R_0$ and (4.45) collapsed to the Black–Scholes model.

Table 4.14: Definition of $A(t,T)$ and $B(t,T)$ for different models of Tables 4.1 and 4.2.

| Model | $A(t,T)$ | $B(t,T)$ |
|---|---|---|
| Vasicek | $\left(\theta - \frac{\widetilde{\sigma}_R^2}{2\kappa^2}\right)[B(t,T)-(T-t)] - \frac{\widetilde{\sigma}_R^2}{4\kappa}B^2(t,T)$ | $\frac{1}{\kappa}[1-e^{-\kappa(T-t)}]$ |
| Ho–Lee | $\int_t^T \theta(s)(s-T)\,\mathrm{d}s + \frac{\widetilde{\sigma}_R^2}{2}\frac{(T-t)^3}{3}$ | $T-t$ |
| Hull–White | $\int_t^T \frac{1}{2}\widetilde{\sigma}_R^2 B(s,T) - \theta(s)B(s,T)\,\mathrm{d}s$ | $\frac{1}{\kappa}[1-e^{-\kappa(T-t)}]$ |

The following proposition provides a closed-form pricing formula for European-style CBs under general stochastic short-rate models (4.45). The pricing of European call options under stochastic interest rate are discussed in Geman *et al.* (1995) (Theorem 2 and Section 3.2), Björk (2009) (Section 26.5), and Brigo and Mercurio (2006) (Appendix B), among others. The general proof uses the change of numéraire techniques developed by Geman *et al.* (1995). An alternative derivation is also provided in Abudy and Izhakian (2013) when the short-rate dynamics is given by the Ornstein–Uhlenbeck process (or, equivalently, the Vasicek model).

**Proposition 4.12.1** *Given $S_t = x > 0$ and $R_t = r \in \mathbb{R}$, the price at time $t$ of a European-style CB with maturity $T > 0$, face value $F > 0$, and conversion ratio $\eta > 0$ is given by*

$$v_e(t,r,x) = \eta x e^{-\int_t^T q_s\,\mathrm{d}s}\Phi(d_1) + e^{-\int_t^T c_u\,\mathrm{d}u}P(t,T)F\Phi(d_2) \tag{4.47}$$

Brownian motion with cross-variation $[Z, \widetilde{W}^{(2)}]_t = \rho t$.

[18] It suffices to set $\theta(t) = \widetilde{\theta}\kappa$, for $0 \le t \le T$, for some constant $\widetilde{\theta} > 0$ to obtain the Vasicek model with a long-term mean parameter equal to $\widetilde{\theta}$.

*where* $\Phi(\cdot)$ *denotes the cumulative distribution of a standard normal distribution,* $d_1 = \frac{\ln\left(\frac{\eta x}{FP(t,T)}\right) - \int_t^T q_s \,\mathrm{d}s + \frac{1}{2}V(t,T)}{\sqrt{V(t,T)}}$*, and* $d_2 = \sqrt{V(t,T)} - d_1$*, with*

$$V(t,T) = \begin{cases} \widetilde{\sigma}_S^2(T-t) + \frac{2\rho\widetilde{\sigma}_S\widetilde{\sigma}_R}{\kappa}\left[(T-t) - B(t,T)\right] + \frac{\widetilde{\sigma}_R^2}{\kappa^2}\left[(T-t) - \frac{\kappa}{2}B^2(t,T) - B(t,T)\right] & \text{if } \kappa \neq 0 \\ \widetilde{\sigma}_S^2(T-t) + \rho\widetilde{\sigma}_S\widetilde{\sigma}_R(T-t)^2 + \frac{\widetilde{\sigma}_R^2}{3}(T-t)^3 & \text{if } \kappa = 0, \end{cases}$$

*and* $P(t,T) = e^{A(t,T)+B(t,T)r}$*, where functions* $A$ *and* $B$ *are defined in Table 4.14.*

*Proof. First, assume first that* $q_t = 0$ *for all* $t \geq 0$ *and recall from* (4.21) *that*

$$v_e(t,x,r) = \mathbb{E}\Big[e^{-\int_t^T R_u \,\mathrm{d}u}\eta S_T \mathbf{1}_{\{S_T \geq F/\eta\}} + e^{-\int_t^T R_u + c_u \,\mathrm{d}u} F \mathbf{1}_{\{S_T < F/\eta\}} \big| S_t = x, R_t = r\Big]. \tag{4.48}$$

*To solve this problem and avoid working with the joint density of* $(\int_t^T R_u \,\mathrm{d}u, S_T)$*, we use the change of numéraire technique discussed in Geman* et al. *(1995), Theorem 2 and Section 3.2 (a). Consequently, we introduce the* $T$*-forward measure, denoted by* $\mathbb{Q}^T$*, which has the zero-coupon bond price process* $P = \{P(t,T)\}_{0 \leq t \leq T}$ *as numéraire. We also introduce the measure* $\mathbb{Q}_S$ *under which the risky asset* $S$ *is the chosen numéraire. Following Geman* et al. *(1995), Theorem 2, we have that*

$$\begin{aligned} v_e(t,x,r) &= P(t,T)\mathbb{E}^{\mathbb{Q}^T}\left[\eta S_T \mathbf{1}_{\{S_T \geq F/\eta\}} \big| S_t = x, R_t = r\right] \\ &\quad + e^{-\int_t^T c_u \,\mathrm{d}u} P(t,T) F \mathbb{Q}^T\left(S_T < F/\eta \big| S_t = x, R_t = r\right) \\ &= \eta x \mathbb{Q}_S\left(S_T \geq F/\eta \big| S_t = x, R_t = r\right) \\ &\quad + e^{-\int_t^T c_u \,\mathrm{d}u} P(t,T) F \mathbb{Q}^T\left(S_T < F/\eta \big| S_t = x, R_t = r\right), \end{aligned} \tag{4.49}$$

*where the first term of the second equality follows from Corollary 2 of Geman* et al. *(1995), and* $\mathbb{E}^{\mathbb{Q}^T}[\cdot]$ *denotes the expectation under the* $T$*-forward measure. To solve* (4.49)*, we need to find the distribution of* $S_T$ *under measures* $\mathbb{Q}^T$ *and* $\mathbb{Q}^S$*, respectively. This is what we do in the following.*

*The first step to obtain the dynamics of* $S$ *under* $\mathbb{Q}^T$ *consists of finding the dynamics of the zero-coupon bond price process* $P$ *under* $\mathbb{Q}$*. Accordingly, recall from* (4.46) *that*

$$P(t,T) = e^{A(t,T) - B(t,T)R_t},$$

*for some time-deterministic functions* $A$ *and* $B$ *defined in Table 4.14, and note that*

$$A_t(t,T) := \frac{\partial A}{\partial t}(t,T) = \theta(t)B(t,T) - \frac{1}{2}\widetilde{\sigma}_R^2 B^2(t,T), \text{ and } B_t(t,T) := \frac{\partial B}{\partial t}(t,T) = \kappa B(t,T) - 1,$$

*see Björk (2009), Section 24.4. Hence, applying Ito's formula to* $\ln P(t,T)$*, we obtain that*

$$\begin{aligned}
\mathrm{d}\ln P(t,T) &= \left(A_t(t,T) - B_t(t,T)R_t\right)\,\mathrm{d}t - B(t,T)\,\mathrm{d}R_t \\
&= \left[\theta(t)B(t,T) - \frac{1}{2}\widetilde{\sigma}_R^2 B^2(t,T) - B_t(t,T)R_t - B(t,T)\left(\theta(t) - \kappa R_t\right)\right]\,\mathrm{d}t - \widetilde{\sigma}_R B(t,T)\,\mathrm{d}\widetilde{W}_t^{(2)} \\
&= \left[R_t - \frac{1}{2}\widetilde{\sigma}_R^2 B^2(t,T)\right]\,\mathrm{d}t - \widetilde{\sigma}_R B(t,T)\,\mathrm{d}\widetilde{W}_t^{(2)},
\end{aligned} \tag{4.50}$$

*so that,*

$$\mathrm{d}P(t,T) = R_t P(t,T)\,\mathrm{d}t - \widetilde{\sigma}_R B(t,T)P(t,T)\,\mathrm{d}\widetilde{W}_t^{(2)}. \tag{4.51}$$

*Now, observe that*

$$\mathbb{Q}^T\left(S_T < F/\eta \middle| S_t = x, R_t = r\right) = \mathbb{Q}^T\left(\frac{S_T}{P(T,T)} < F/\eta \middle| S_t = x, R_t = r\right). \tag{4.52}$$

*From Theorem 1 (i) of Geman* et al. *(1995), we know that process* $f = \{f_t := \frac{S_t}{P(t,T)}\}_{0\leq t\leq T}$ *is a* $\mathbb{Q}^T$*-local martingale (thus, it as no drift). Moreover, since a change of measures only affects the drift of a process, we can deduce the dynamics of process* $f$ *under* $\mathbb{Q}^T$ *from its dynamics under* $\mathbb{Q}$ *by setting the drift term to nil, which we do next. Using* (4.45) , (4.51) *and Itô's formula, we find that*

$$\begin{aligned}
\mathrm{d}f_t &= \frac{1}{P(t,T)}\,\mathrm{d}S_t - \frac{S_t}{P^2(t,T)}\,\mathrm{d}P(t,T) - \frac{1}{P^2(t,T)}\,\mathrm{d}\langle S_t, P(t,T)\rangle + \frac{S_t}{P^3(t,T)}\,\mathrm{d}\langle P(t,T)\rangle \\
&= \left(\rho\widetilde{\sigma}_S\widetilde{\sigma}_R B(t,T) + \widetilde{\sigma}_R^2 B^2(t,T)\right) f_t\,\mathrm{d}t + \widetilde{\sigma}_S\sqrt{1-\rho^2} f_t\,\mathrm{d}\widetilde{W}_t^{(1)} + \left[\rho\widetilde{\sigma}_s + \widetilde{\sigma}_R B(t,T)\right] f_t\,\mathrm{d}\widetilde{W}_t^{(2)}.
\end{aligned} \tag{4.53}$$

*Because* $f$ *must be a local martingale under* $\mathbb{Q}^T$*, we deduce from Girsanov theorem that process* $(\widehat{W}^{(1)}, \widehat{W}^{(2)})$*, defined by*

$$\begin{aligned}
\mathrm{d}\widehat{W}_t^{(1)} &= \mathrm{d}\widetilde{W}_t^{(1)}, \\
\mathrm{d}\widehat{W}_t^{(2)} &= \mathrm{d}\widetilde{W}_t^{(2)} + \widetilde{\sigma}_R B(t,T)\,\mathrm{d}t,
\end{aligned} \tag{4.54}$$

*for* $0 \leq t \leq T$*, is a standard bi-dimensional Brownian motion under* $\mathbb{Q}^T$*, which implies that*

$$\mathrm{d}f_t = \widetilde{\sigma}_S\sqrt{1-\rho^2} f_t\,\mathrm{d}\widehat{W}_t^{(1)} + \left[\rho\widetilde{\sigma}_s + \widetilde{\sigma}_R B(t,T)\right] f_t\,\mathrm{d}\widehat{W}_t^{(2)}. \tag{4.55}$$

*Hence,* $\ln f_T$ *given* $\ln f_t$ *is normally distributed with a mean* $\widehat{\mu} := \ln f_t - \frac{1}{2}V(t,T)$ *(see for instance, Björk (2009), Lemma 4.15), and variance* $\widehat{\sigma}^2 := V(t,T)$*, where*

$$\begin{aligned}
V(t,T) &= \int_t^T \widetilde{\sigma}_S^2(1-\rho^2) + \left[\rho\widetilde{\sigma}_S + \widetilde{\sigma}_R B(s,T)\right]^2\,\mathrm{d}s \\
&= \widetilde{\sigma}_S^2(T-t) + 2\rho\widetilde{\sigma}_S\widetilde{\sigma}_R\int_t^T B(s,T)\,\mathrm{d}s + \widetilde{\sigma}_R^2\int_t^T B^2(s,T)\,\mathrm{d}s.
\end{aligned} \tag{4.56}$$

*The probability* (4.52) *can thus be calculated explicitly using the property of the normal distribution, and we obtain that,*

$$\mathbb{Q}^T\left(f_T < F/\eta \middle| f_t = x/P(t,T)\right) = \Phi\left(\frac{\ln\left(\frac{FP(t,T)}{\eta x}\right) + \frac{1}{2}V(t,T)}{\sqrt{V(t,T)}}\right) = \Phi(d_2). \tag{4.57}$$

*This completes the proof for the second term of* (4.49)*.*

*The challenge now consists of finding an expression for the first term of* (4.49)*. We thus need the distribution of* $S_T$ *under* $\mathbb{Q}_S$*. To this end, we observe that*

$$\begin{aligned}\mathbb{Q}_S\left(S_T \geq F/\eta \middle| S_t = x, R_t = r\right) &= \mathbb{Q}_S\left(1/S_T \leq \eta/F \middle| S_t = x, R_t = r\right)\\ &= \mathbb{Q}_S\left(P(T,T)/S_T \leq \eta/F \middle| S_t = x, R_t = r\right)\\ &= \mathbb{Q}_S\left(1/f_T \leq \eta/F \middle| f_t = x/P(t,T)\right)\end{aligned} \tag{4.58}$$

*Similarly, as above, we use the results of Theorem 1 of Geman* et al. *(1995) to conclude that process* $Y = \{Y_t := P(t,T)/S_t = 1/f_t\}_{0\leq t\leq T}$ *is a local martingale under* $\mathbb{Q}_S$*. Thus, it has no drift, and because a change of measures only affects the drift of a process, we can deduce the dynamics of process* $Y$ *under* $\mathbb{Q}_S$ *from its dynamics under* $\mathbb{Q}^T$*, by setting the drift term to nil. This what we do in the following. Using* (4.55) *and Itô's formula, we find that*

$$\begin{aligned}\mathrm{d}Y_t &= -\frac{1}{f_t^2}\,\mathrm{d}f_t + \frac{1}{f_t^3}\,\mathrm{d}\langle f_t\rangle\\ &= \left[\widetilde{\sigma}_S^2(1-\rho^2) + (\rho\widetilde{\sigma}_S + \widetilde{\sigma}_R B(t,T))^2\right] Y_t\,\mathrm{d}t\\ &\qquad - \widetilde{\sigma}_S\sqrt{1-\rho^2}Y_t\,\mathrm{d}\widehat{W}_t^{(1)} - \left[\rho\widetilde{\sigma}_s + \widetilde{\sigma}_R B(t,T)\right] Y_t\,\mathrm{d}\widehat{W}_t^{(2)}.\end{aligned}$$

*From the local martingale property of* $Y$ *under* $\mathbb{Q}_S$*, we deduce from the Girsanov theorem that process* $(\bar{W}^{(1)}, \bar{W}^{(2)})$*, defined by*

$$\begin{aligned}\mathrm{d}\bar{W}_t^{(1)} &= \mathrm{d}\widehat{W}_t^{(1)} - \widetilde{\sigma}_S\sqrt{1-\rho^2}\,\mathrm{d}t,\\ \mathrm{d}\bar{W}_t^{(2)} &= \mathrm{d}\widehat{W}_t^{(2)} - \rho\widetilde{\sigma}_S + \widetilde{\sigma}_R B(t,T)\,\mathrm{d}t,\end{aligned} \tag{4.59}$$

*for* $0 \leq t \leq T$*, is a standard bi-dimensional Brownian motion under* $\mathbb{Q}_S$*. Thus, the dynamics of* $Y$ *under the new measure are given by*

$$\mathrm{d}Y_t = -\widetilde{\sigma}_S\sqrt{1-\rho^2}Y_t\,\mathrm{d}\bar{W}_t^{(1)} - \left[\rho\widetilde{\sigma}_s + \widetilde{\sigma}_R B(t,T)\right] Y_t\,\mathrm{d}\bar{W}_t^{(2)}.$$

*From there, we conclude that* $\ln Y_T$ *given* $\ln Y_t$ *(or* $\ln f_t$*, since* $\ln Y_t = \ln 1/f_t$*) is normally distributed with mean* $\bar{\mu} := -\ln f_t - \frac{1}{2}V(t,T)$ *and variance* $\bar{\sigma}^2 := \widehat{\sigma}^2 = V(t,T)$*, and the probability* (4.58) *can be calculated explicitly as*

$$\mathbb{Q}_S\left(1/f_T \leq \eta/F \middle| f_t = x/P(t,T)\right) = \Phi\left(\frac{\ln\left(\frac{\eta x}{FP(t,T)}\right) + \frac{1}{2}V(t,T)}{\sqrt{V(t,T)}}\right) = \Phi(d_1). \tag{4.60}$$

*The final assertion then follows from* (4.49),(4.56), (4.57), (4.60)*, and the expression for function* $B$ *in Table 4.14.*

*When* $q_t > 0$ *for some* $t \geq 0$*, the results can be derived using the relationship* $S_t = \widetilde{S}_t e^{-\int_0^t q_s \,\mathrm{d}s}$*,* $t \geq 0$*, where* $\widetilde{S} = \{\widetilde{S}_t\}_{t\geq 0}$ *represents the stock price process when the dividend is assumed to be nil*[19]*.* □

**Remark 4.12.1** *When* $\sigma_R = \kappa = \theta(t) = 0$ *for all* $t \in [0,T]$*, the short-rate is constant to* $R_0$ *over time and the model in* (4.45) *collapsed to the Black–Scholes model. In that case,* $A(t,T) = 0$*,* $B(t,T) = T - t$*, such that* $P(t,T) = e^{-(T-t)R_0}$*, and* $V(t,T) = \sigma_S^2(T-t)$*. The general formula* (4.47) *then becomes*

$$v_e(t,r,x) = \eta x e^{-\int_t^T q_s \,\mathrm{d}s}\Phi(d_1) + e^{-(T-t)R_0}e^{-\int_t^T c_u \,\mathrm{d}u}F\Phi(d_2), \tag{4.61}$$

*with* $d_1 = \frac{\ln\left(\frac{\eta x}{F}\right) + \left(r - \int_t^T q_s \,\mathrm{d}s + \frac{1}{2}\sigma_S^2\right)(T-t)}{\sigma_S\sqrt{T-t}}$*, and* $d_2 = \sigma_S\sqrt{T-t} - d_1$*.*

### 4.12.2 Additional Numerical Experiments

We now investigate the accuracy, convergence, and efficiency of the CTMC method to approximate various debt securities in the Vasicek, CIR, and Dothan models. Note that Assumption 4.4.1 is not respected under the Vasicek model.

Unless stated otherwise, we use the following model and CTMC parameters in all numerical experiments, summarized in Table 4.15.

To construct the state-space of the CTMC, $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$ with $m \in \mathbb{N}$, we use the non-uniform grid proposed by Tavella and Randall (Tavella and Randall (2000), Chapter 5.3), as in Section 4.6.

[19] The value of $\widetilde{S}$ can also be interpreted as the value of the stock price when dividends are continuously reinvested in the stock.

Table 4.15: Model and CTMC parameters

| | $r_0$ | $\kappa$ | $\theta$ | $\sigma$ | $m$ | $r_1$ | $r_m$ | $\widetilde{\alpha}$ | $\Delta_N$ |
|---|---|---|---|---|---|---|---|---|---|
| Vasicek | $0.04$ | $1$ | $0.04$ | $0.20$ | $160$ | $-30R_0$ | $25R_0$ | $0.5$ | 1/252 |
| CIR | $0.04$ | $2$ | $0.035$ | $0.20$ | $160$ | $R_0/100$ | $7R_0$ | $0.5$ | 1/252 |
| Dothan | 0.02 | 0 | N/A | 0.15 | $160$ | $R_0/100$ | $7R_0$ | $0.5$ | 1/252 |

#### 4.12.2.1 Approximation of Zero-Coupon Bond Prices

We now examine the accuracy of the CTMC methods in calculating zero-coupon bond prices, Corollary 4.6. Under the Vasicek and CIR models, the price of zero-coupon bonds has a closed-from expression, which can be found, for instance, in Brigo and Mercurio (2006), Section 3.2. The analytical value of the zero-coupon bond price can thus be used as a benchmark in our experiment. We test the accuracy of the approximated prices across different values of model parameters. The results are summarized in Table 4.16. Column "CTMC" reports the CTMC approximated value using the results of Corollary 4.6.

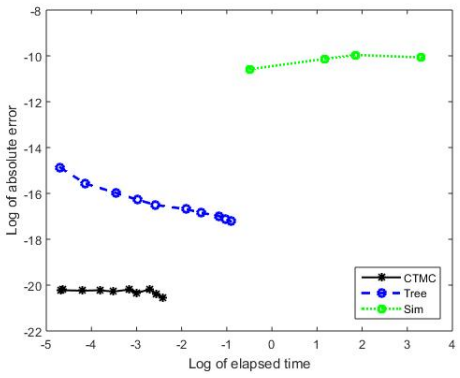



Figure 4.6: Efficiency of the CTMC method in approximating zero-coupon bond prices under the Dothan model. Except for the number of grid points $m$, which range from $100$ to $1000$, model and CTMC parameters are as listed in Table 4.15. Zero-coupon bond price parameters are $t = 0$ and $T = 4$.

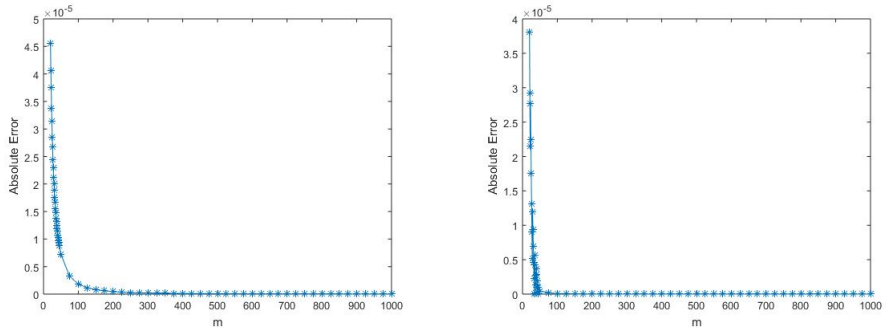
Table 4.16: Accuracy of the zero-coupon bond price approximation (4.37) under Vasicek and CIR models. Model and CTMC parameters are as listed in Table 4.15, and zero-coupon bond parameters are $t = 0$ and $T = 4$.

(a) Vasicek model

| $\kappa$ | CTMC | Benchmark | Abs. error |
|---|---|---|---|
| **0.5** | 0.9625591 | 0.9625609 | 1.77E-06 |
| **1** | 0.8964870 | 0.8964877 | 7.12E-07 |
| **2** | 0.8661056 | 0.8661057 | 9.47E-08 |
| **3** | 0.8587974 | 0.8587974 | 2.27E-08 |
| **4** | 0.8560138 | 0.8560138 | 7.77E-09 |
| $\mathbf{R_0}$ | **CTMC** | **Benchmark** | **Abs. error** |
| **0.02** | 0.9142555 | 0.9142630 | 7.43E-06 |
| **0.03** | 0.9053312 | 0.9053317 | 5.12E-07 |
| **0.04** | 0.8964870 | 0.8964877 | 7.12E-07 |
| **0.05** | 0.8877292 | 0.8877301 | 8.23E-07 |
| $\theta$ | **CTMC** | **Benchmark** | **Abs. error** |
| **0.01** | 0.9814527 | 0.9814529 | 1.93E-07 |
| **0.02** | 0.9522718 | 0.9522722 | 3.77E-07 |
| **0.03** | 0.9239585 | 0.9239590 | 5.50E-07 |
| **0.04** | 0.8964870 | 0.8964877 | 7.12E-07 |
| $\sigma$ | **CTMC** | **Benchmark** | **Abs. error** |
| **0.1** | 0.8630196 | 0.8630198 | 1.70E-07 |
| **0.2** | 0.8964870 | 0.8964877 | 7.12E-07 |
| **0.3** | 0.9551747 | 0.9551765 | 1.77E-06 |
| **0.4** | 1.0438353 | 1.0438513 | 1.60E-05 |

(b) CIR model

| $\kappa$ | CTMC | Benchmark | Abs. error |
|---|---|---|---|
| **0.5** | 0.8656670 | 0.8656663 | 7.09E-07 |
| **1** | 0.8676884 | 0.8676884 | 1.11E-08 |
| **2** | 0.8676884 | 0.8676884 | 1.11E-08 |
| **3** | 0.8681491 | 0.8681491 | 3.96E-10 |
| **4** | 0.8684110 | 0.8684110 | 1.98E-11 |
| $\mathbf{R_0}$ | **CTMC** | **Benchmark** | **Abs. error** |
| **0.02** | 0.8763624 | 0.8763627 | 2.82E-07 |
| **0.03** | 0.8720147 | 0.8720147 | 2.21E-09 |
| **0.04** | 0.8676884 | 0.8676884 | 1.11E-08 |
| **0.05** | 0.8633835 | 0.8633835 | 2.90E-08 |
| $\theta$ | **CTMC** | **Benchmark** | **Abs. error** |
| **0.015** | 0.9303530 | 0.9303518 | 1.20E-06 |
| **0.025** | 0.8984741 | 0.8984740 | 1.82E-07 |
| **0.035** | 0.8676884 | 0.8676884 | 1.11E-08 |
| **0.045** | 0.8379576 | 0.8379576 | 5.63E-10 |
| $\sigma$ | **CTMC** | **Benchmark** | **Abs. error** |
| **0.1** | 0.8673140 | 0.8673140 | 1.93E-10 |
| **0.2** | 0.8676884 | 0.8676884 | 1.11E-08 |
| **0.3** | 0.8683033 | 0.8683025 | 7.69E-07 |
| **0.4** | 0.8691384 | 0.8691428 | 4.36E-06 |




(a) Vasicek (b) CIR

Figure 4.7: Convergence pattern of the approximated zero-coupon bond prices. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.15. Zero-coupon bond price parameters are $t = 0$ and $T = 4$.

Table 4.17: Approximation of the convergence rate of the zero-coupon bond prices. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.15. Zero-coupon bond price parameters are $t = 0$ and $T = 4$.

(a) Vasicek

| m | Abs. error | Rate |
|---|---|---|
| 50 | 7.24E-06 | - |
| 100 | 1.82E-06 | 1.99 |
| 200 | 4.56E-07 | 2.00 |
| 300 | 2.03E-07 | 2.00 |
| 900 | 2.19E-08 | 2.03 |

(b) CIR

| m | Abs. error | Rate |
|---|---|---|
| 50 | 4.57E-07 | - |
| 100 | 3.90E-08 | 3.55 |
| 200 | 4.63E-09 | 3.08 |
| 300 | 1.55E-09 | 2.70 |
| 900 | 9.18E-10 | 0.48 |

We observe that the approximation achieves a high level of precision across all parameters, with an average calculation time of $0.009$ seconds, illustrating the speed of the methodology.

Figures 4.6 demonstrate the efficiency of CTMC methodology in valuing zero-coupon bond prices when the short-rate process follows a geometric Brownian motion (the Dothan model). Thus, alternative methods such as binomial trees ("Tree") or Monte Carlo simulation ("Sim") can be easily implemented. We compare the performance of the CTMC approximation to the Cox–Ross–Rubinstein binomial tree (Cox *et al*. (1979)) and to Monte Carlo simulation. For Monte Carlo simulation, we use an exact scheme with a number of

simulations ranging from 10,000 to 200,000 with the same number of antithetic variables and 500-time steps per year. The benchmark is calculated using the CTMC method with $m = 5,000$. Figure 4.6 shows the high efficiency of the CTMC method compared to these other numerical techniques. Indeed, CTMC approximation clearly outperforms other methods in terms of both calculation time and precision.

Finally, Figure 4.7 shows the convergence pattern of the approximated zero-coupon bond prices as the number of grid points $m$ increases, whereas Table 4.17 shows the convergence rate. For the Vasicek (resp. CIR) model, we note that the approximations achieve quadratic (resp. superquadratic) convergence on average. We also observe that the two models converge smoothly and rapidly to their analytical values. Moreover, since Assumption 4.4.1 is not satisfied under the Vacisek model, the results show that theoretical convergence is possible under less restrictive conditions for a certain set of parameters. Theoretical proof is left as future research.

#### 4.12.2.2 Approximation of the Zero-Coupon Bond Option Prices

In this section, we study the accuracy and the numerical convergence of the zero-coupon bond option prices provided in (4.38), under the Vasicek and CIR models. Under these two models, the price of zero-coupon bond options has a closed-form expression, which can be found in Brigo and Mercurio (2006) Section 3.2., and thus, can serve as a benchmark in our example. We test the accuracy of the approximated option prices for different levels of moneyness and volatilities. The results are summarized in Table 4.18. Column "price-to-strike" shows the price-to-strike ratio, calculated as the actual zero-coupon bond price over the option strike price $K > 0$.

We observe that the approximation achieves a high level of accuracy across all volatilities and strikes. The convergence of the approximated call prices to the analytical formulas is illustrated in Figure 4.8 for the two models, whereas the approximated convergence rates are shown in Table 4.19.

We note that the approximated call prices converge rapidly to their analytical values but exhibit a sawtooth pattern. As mentioned in Section 4.6.1, such oscillatory behavior has been observed in other research (see, for instance, Zhang and Li (2019)). However, the technique proposed by the authors to remove oscillation and improve convergence is not directly applicable in the present context. Further investigation into how grid design can improve convergence is left for future research.

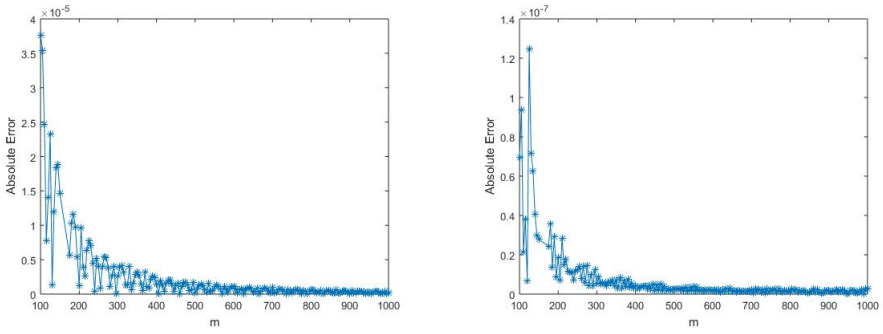
Table 4.18: Accuracy of the zero-coupon bond call option approximation (4.38) under Vasicek and CIR models. Benchmark is calculated using closed-form analytical formulas. Except for the number of grid points set to $m = 1,000$, model and CTMC parameters are as listed in Table 4.15. Zero-coupon bond call option parameters using the notation of Corollary 4.7: $t_{n_1} = 0$, $t_{n_2} = 2$, and $T = 4$.

(a) Vasicek model

| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
|---|---|---|---|---|
| | 1.67 | 0.38319569 | 0.38319569 | 4.29E-09 |
| | 1.25 | 0.22325433 | 0.22325433 | 4.25E-09 |
| $\mathbf{0.1}$ | 1.00 | 0.06581737 | 0.06581712 | 2.56E-07 |
| | 0.83 | 0.01014165 | 0.01014179 | 1.40E-07 |
| | 0.71 | 0.00000011 | 0.00000011 | 6.79E-10 |
| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
| | 1.67 | 0.39232996 | 0.39232997 | 1.76E-08 |
| | 1.25 | 0.22455160 | 0.22455166 | 5.84E-08 |
| $\mathbf{0.2}$ | 1.00 | 0.07590525 | 0.07590491 | 3.39E-07 |
| | 0.83 | 0.01014165 | 0.01014179 | 1.40E-07 |
| | 0.71 | 0.00053736 | 0.00053749 | 1.33E-07 |
| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
| | 1.67 | 0.40772961 | 0.40772968 | 7.30E-08 |
| | 1.25 | 0.22983971 | 0.22983969 | 1.39E-08 |
| $\mathbf{0.3}$ | 1.00 | 0.09108337 | 0.09108360 | 2.27E-07 |
| | 0.83 | 0.01014165 | 0.01014179 | 1.40E-07 |
| | 0.71 | 0.00465377 | 0.00465387 | 1.06E-07 |
| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
| | 1.67 | 0.43035317 | 0.43036293 | 9.76E-06 |
| | 1.25 | 0.24315058 | 0.24315678 | 6.20E-06 |
| $\mathbf{0.4}$ | 1.00 | 0.10988334 | 0.10988798 | 4.64E-06 |
| | 0.83 | 0.01014165 | 0.01014179 | 1.40E-07 |
| | 0.71 | 0.01314694 | 0.01315072 | 3.77E-06 |

(b) CIR model

| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
|---|---|---|---|---|
| | 1.67 | 0.38326833 | 0.38326833 | 5.81E-10 |
| | 1.25 | 0.22191975 | 0.22191976 | 5.04E-10 |
| $\mathbf{0.1}$ | 1.00 | 0.06057118 | 0.06057118 | 4.27E-10 |
| | 0.95 | 0.02020450 | 0.02020450 | 3.55E-09 |
| | 0.92 | 0.00000000 | 0.00000000 | 1.22E-11 |
| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
| | 1.67 | 0.38334989 | 0.38334989 | 3.29E-09 |
| | 1.25 | 0.22190374 | 0.22190374 | 2.86E-09 |
| $\mathbf{0.2}$ | 1.00 | 0.06045761 | 0.06045761 | 2.43E-09 |
| | 0.95 | 0.02020450 | 0.02020450 | 3.55E-09 |
| | 0.92 | 0.00000690 | 0.00000690 | 5.39E-09 |
| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
| | 1.67 | 0.38348301 | 0.38348311 | 1.01E-07 |
| | 1.25 | 0.22187654 | 0.22187663 | 9.48E-08 |
| $\mathbf{0.3}$ | 1.00 | 0.06027987 | 0.06028002 | 1.53E-07 |
| | 0.95 | 0.02020450 | 0.02020450 | 3.55E-09 |
| | 0.92 | 0.00006003 | 0.00006008 | 5.13E-08 |
| $\sigma$ | price-to-strike | CTMC | Benchmark | Abs. error |
| | 1.67 | 0.38366164 | 0.38366831 | 6.68E-06 |
| | 1.25 | 0.22183726 | 0.22184350 | 6.23E-06 |
| $\mathbf{0.4}$ | 1.00 | 0.06011562 | 0.06012477 | 9.15E-06 |
| | 0.95 | 0.02020450 | 0.02020450 | 3.55E-09 |
| | 0.92 | 0.00011258 | 0.00011395 | 1.37E-06 |




(a) Vasicek (b) CIR

Figure 4.8: Convergence pattern of the approximated zero-coupon bond call option prices as the number of grid points $m$ increases. Benchmark is calculated using closed-form analytical formulas. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.15. Zero-coupon bond call option parameters using the notation of Corollary 4.7: $t_{n_1} = 0$, $t_{n_2} = 2$, $T = 4$, $K = 0.9$.

Table 4.19: Approximation of the convergence rate of the zero-coupon bond option prices. Benchmark is calculated using closed-form analytical formulas. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.15. Zero-coupon bond call option parameters using the notation of Corollary 4.7: $t_{n_1} = 0$, $t_{n_2} = 2$, $T = 4$, $K = 0.9$.

(a) Vasicek

| m | Abs. error | Rate |
|---|---|---|
| 100 | 3.77E-05 | - |
| 250 | 4.05E-06 | 2.43 |
| 400 | 1.49E-06 | 2.12 |
| 550 | 1.07E-06 | 1.06 |
| 800 | 5.90E-07 | 1.58 |

(b) CIR

| m | Abs. error | Rate |
|---|---|---|
| 100 | 6.97E-08 | - |
| 250 | 1.25E-08 | 1.88 |
| 400 | 4.09E-09 | 2.37 |
| 550 | 4.00E-09 | 0.07 |
| 800 | 1.12E-09 | 3.40 |

#### 4.12.2.3 Approximation of Callable/Putable Bond Prices

We now examine the accuracy and the convergence of Proposition 4.4.3 in approximating callable/putable bonds under the Vasicek model. Accordingly, we consider a coupon-bearing bond with semi-annual coupons that mature in $4$ years $T = 4$. The coupon rate, denoted below by $\alpha$, is set to $4\%$ per annum compounded semi-annually. The notional of the debt is set to $F = 100$, and we assume that it can be called at any time between the second and the fourth year for no additional cost, that is, $K_t^c = 100$ for $2 \leq t \leq T$ and we let $K_t^c \to \infty$ when $t < 2$ (since exercise is not allowed). Moreover, as there is no put feature, $K^p := K_t^p = 0$ for $0 \leq t \leq T$. Finally, we assume that accrued interest is paid to the bondholder upon redemption. The contract specifications are summarized in Table 4.20.

Table 4.20: Callable bond contract specifications[20]

| $F$ | $\alpha$ | $T$ | $K_t^c$ | $K^p$ |
|---|---|---|---|---|
| $100$ | $0.04$ | $4$ | $100$ | $0$ |

Proposition 4.4.3 is also used to calculate the value of the straight bond (i.e., the value of the coupon-bearing bond with no optionality, when $K_t^c \to \infty$ and $K_t^p = 0$ for all $t \in [0, T]$). The results are summarized in Table 4.21a, whereas those for the callable bond are outlined in Table 4.21b. The value of the optionality is obtained from the difference between the value of the callable and the straight bonds. The call option has a negative value because it is in favor of the issuer and, thus, reduces the value of the bond. For the straight debt, a benchmark can be obtained using a closed-form analytical formula since it can be decomposed in a series of zero-coupon bonds (see Remark 4.4.5). For the callable debt, the benchmark is calculated using CTMC approximation with $m = 2,000$.

[20] $K_t^c$=100 for $2 \leq t \leq T$ and $K_t^c$ is set to a large constant when $t < 2$.

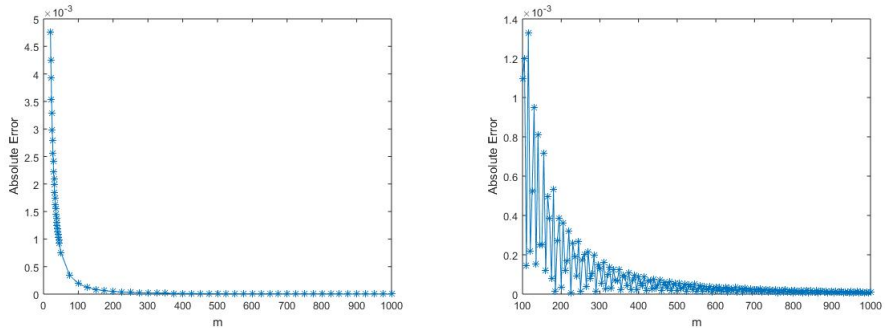
Table 4.21: Accuracy of Proposition 4.4.3 to approximate the values of straight and callable bonds under Vasicek model. Model and CTMC parameters are as listed in Table 4.15. Contract specifications are as listed in Table 4.20, with the call option exercise window starting from $t = 2$ to maturity.

(a) Straight bond

| $\kappa$ | CTMC | Benchmark | Rel. error |
|---|---|---|---|
| **0.5** | 111.5779419 | 111.5781246 | 1.64E-06 |
| **1** | 104.6007798 | 104.6008544 | 7.13E-07 |
| **2** | 101.3606496 | 101.3606597 | 9.94E-08 |
| **3** | 100.5741608 | 100.5741632 | 2.42E-08 |
| **4** | 100.2732264 | 100.2732272 | 8.40E-09 |
| $\mathbf{R_0}$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.02** | 106.6205602 | 106.6213440 | 7.35E-06 |
| **0.03** | 105.6061649 | 105.6062188 | 5.11E-07 |
| **0.04** | 104.6007798 | 104.6008544 | 7.13E-07 |
| **0.05** | 103.6050710 | 103.6051563 | 8.23E-07 |
| $\theta$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.01** | 113.7506391 | 113.7506574 | 1.61E-07 |
| **0.02** | 110.6102083 | 110.6102465 | 3.46E-07 |
| **0.03** | 107.5611515 | 107.5612085 | 5.30E-07 |
| **0.04** | 104.6007798 | 104.6008544 | 7.13E-07 |
| $\sigma$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.1** | 101.0176528 | 101.0176706 | 1.76E-07 |
| **0.2** | 104.6007798 | 104.6008544 | 7.13E-07 |
| **0.3** | 110.8775749 | 110.8777598 | 1.67E-06 |
| **0.4** | 120.3458095 | 120.3474922 | 1.40E-05 |

(b) Callable bond

| $\kappa$ | CTMC | Benchmark | Rel. error |
|---|---|---|---|
| **0.5** | 94.4289102 | 94.4293068 | 4.20E-06 |
| **1** | 95.5617290 | 95.5616095 | 1.25E-06 |
| **2** | 96.9174873 | 96.9173130 | 1.80E-06 |
| **3** | 97.6890278 | 97.6890198 | 8.14E-08 |
| **4** | 98.1735503 | 98.1735203 | 3.06E-07 |
| $\mathbf{R_0}$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.02** | 97.3030355 | 97.3030555 | 2.05E-07 |
| **0.03** | 96.4287924 | 96.4287773 | 1.57E-07 |
| **0.04** | 95.5617290 | 95.5616095 | 1.25E-06 |
| **0.05** | 94.7018444 | 94.7020397 | 2.06E-06 |
| $\theta$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.01** | 100.5832613 | 100.5844116 | 1.14E-05 |
| **0.02** | 98.9373817 | 98.9375213 | 1.41E-06 |
| **0.03** | 97.2623005 | 97.2631690 | 8.93E-06 |
| **0.04** | 95.5617290 | 95.5616095 | 1.25E-06 |
| $\sigma$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.1** | 97.0423712 | 97.0419581 | 4.26E-06 |
| **0.2** | 95.5617290 | 95.5616095 | 1.25E-06 |
| **0.3** | 95.3070805 | 95.3073647 | 2.98E-06 |
| **0.4** | 96.1965169 | 96.1965745 | 6.00E-07 |




(a) Straight bond (b) Callable bond

Figure 4.9: Convergence pattern of the approximated prices of straight and callable bonds as the number of grid points $m$ increases under the Vasicek model. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.15. Contract specifications are as listed in Table 4.20, with the call option exercise window starting from $t = 2$ to maturity.

Table 4.22: Approximation of the convergence rate of straight and callable bond prices under Vasicek model. Except for the number of grid points $m$, model and CTMC parameters are as listed in Table 4.15. Contract specifications are as listed in Table 4.20, with the call option exercise window starting from $t = 2$ to maturity.

(a) Straight bond

| m | Abs. error | Rate |
|---|---|---|
| 100 | 1.91E-04 | - |
| 125 | 1.22E-04 | 2.02 |
| 150 | 8.47E-05 | 1.99 |
| 175 | 6.23E-05 | 1.99 |
| 200 | 4.78E-05 | 1.99 |

(b) Callable bond

| m | Abs. error | Rate |
|---|---|---|
| 100 | 1.10E-03 | - |
| 125 | 5.25E-04 | 3.29 |
| 150 | 2.49E-04 | 4.09 |
| 175 | 7.96E-05 | 7.41 |
| 200 | 3.57E-05 | 6.00 |

Again, we note that the approximation achieves a high level of accuracy in a fraction of a second across all model parameters. The average calculation time is 0.02 seconds.

The convergence pattern of straight and callable bonds is displayed in Figure 4.9, and the approximated convergence rates are shown in Table 4.22. For the straight bond (resp. callable bond), we note that the approximations achieve quadratic (resp. superquadratic) convergence on average. We also observe that the

straight bond converges smoothly to the analytical price. However, the callable debt exhibits a sawtooth pattern. For the two securities, the absolute error decreases rapidly to $0$.

#### 4.12.2.4 Approximation of Convertible Bond Prices

We now investigate the accuracy of Algorithm 11 in approximating CB prices under the Black–Scholes–Vasicek model, along with the numerical convergence of the price estimates.

That is, we suppose that the stock price process follows a geometric Brownian motion with stochastic interest rates satisfying

$$\begin{aligned} \mathrm{d}S_t &= (R_t - q_t)S_t\,\mathrm{d}t + \widetilde{\sigma}_S S_t\,\mathrm{d}W_t^{(1)}, \\ \mathrm{d}R_t &= \kappa(\theta - R_t)\,\mathrm{d}t + \widetilde{\sigma}_R\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{4.62}$$

with $\kappa$, $\theta$, $\widetilde{\sigma}_S$, $\widetilde{\sigma}_R$>0, and $[W^{(1)}, W^{(2)}]_t = \rho t$, $\rho \in [-1, 1]$.

From Lemma 4.1, we find that $f(r) = \frac{\widetilde{\sigma}_S}{\widetilde{\sigma}_R} r$. The dynamics of the auxiliary process $X_t = \ln(S_t) - \rho f(R_t)$ can then be derived as

$$\begin{aligned} \mathrm{d}X_t &= \mu_X(t, R_t)\,\mathrm{d}t + \sigma_X(R_t)\,\mathrm{d}W_t^{\star} \\ \mathrm{d}R_t &= \kappa(\theta - R_t)\,\mathrm{d}t + \widetilde{\sigma}_R\,\mathrm{d}W_t^{(2)}, \end{aligned} \tag{4.63}$$

with $\mu_X(t, R_t) = R_t - q_t - \frac{\widetilde{\sigma}_S^2}{2} - \rho\frac{\widetilde{\sigma}_S}{\widetilde{\sigma}_R}\kappa(\theta - R_t)$, $\sigma_X = \widetilde{\sigma}_S\sqrt{1-\rho^2}$, and $X_0 = \ln(S_0) - \rho f(R_0)$.

Unless stated otherwise, the model parameters for the short-rate process are the same as those used in previous examples, reported in Table 4.15 under Vasicek model. We suppose further that $\widetilde{\sigma}_S = 0.2$, $q_t = 0.02$ for all $t \in [0, T]$, and $\rho = -0.2$. The model parameters are summarized in Table 4.23.

Table 4.23: Model parameters

| Model | $R_0$ | $\kappa$ | $\theta$ | $\widetilde{\sigma}_R$ | $S_0$ | $q_t$ | $\widetilde{\sigma}_S$ | $\rho$ |
|---|---|---|---|---|---|---|---|---|
| Black–Scholes–Vasicek | $0.04$ | $1$ | $0.04$ | $0.20$ | $100$ | $0.02$ | $0.2$ | $-0.2$ |

The grid used to approximate the short-rate process, $\mathcal{S}_R^{(m)} = \{r_1, r_2, \ldots, r_m\}$, and the auxiliary process $\mathcal{S}_X^{(M)} = \{x_1, x_2, \ldots, x_M\}$, are constructed using the methodology of Tavella and Randall (Tavella and Randall (2000), Chapter 5), as explained in Section 4.6, with $\widetilde{\alpha}_R$ (resp. $\widetilde{\alpha}_X$) representing the non-uniformity

parameter of the grid for $R^{(m)}$ (resp. $X^{(m)}$). Unless otherwise indicated, all numerical experiments are conducted using the CTMC parameters listed in Table 4.24.

Table 4.24: CTMC parameters

| | $m$ | $M$ | $r_1$ | $r_m$ | $\widetilde{\alpha}_R$ | $x_1$ | $x_M$ | $\widetilde{\alpha}_X$ | $\Delta_N$ |
|---|---|---|---|---|---|---|---|---|---|
| Black–Scholes–Vasicek | $160$ | $100$ | $-30R_0$ | $25R_0$ | $0.5$ | $0.64X_0$ | $1.42X_0$ | $2$ | $1/100$ |

Table 4.25: CB contract specifications

| $F$ | $\alpha$ | $T$ | $\eta$ |
|---|---|---|---|
| 100 | 0.05 | 1 | 1 |

The contract specifications are summarized in Table 4.25. We consider a convertible bond that pays semi-annual coupons at an annual rate of $\alpha = 0.05$ with a notional $F = 100$. We suppose that the bond can be converted at any time from inception to maturity ($T = 1$) at a conversion rate $\eta = 1$.

Under this set of parameters and when both dividend yield and credit spread are assumed to be nil ($q_t = c_t = 0$ for all $t \in [0, T]$), the valuation of American-style CBs is simplified to that of European-style CBs, see Corollary 4.4. The results of Proposition 4.12.1[21] can thus serve as a benchmark in our analysis. When $q_t, c_t > 0$ for some $t \in [0, T]$, the benchmark is calculated using CTMC approximation with $M = 160$ and $\Delta_N = 1/252$, all other CTMC parameters are as listed in Table 4.24. The results are summarized in Table 4.26.

We note that the model achieves a high level of accuracy across all model parameters. The average calculation time for the CTMC approximated prices is less than 1.70 seconds.

[21] The expected present value of future coupons should be added to the formula obtained in Proposition 4.12.1.

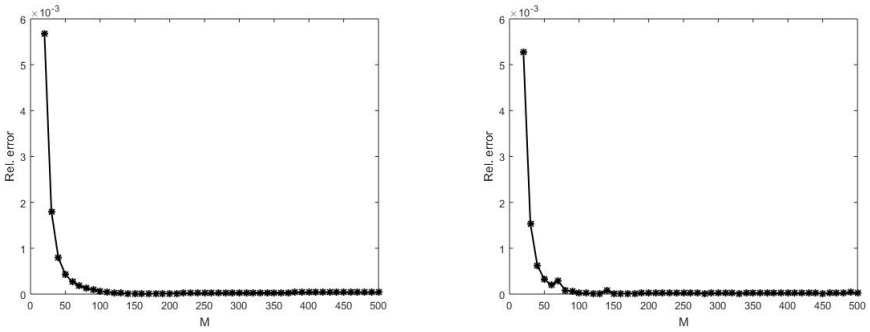
Table 4.26: Accuracy of the CB price approximations, Algorithm 11, under Black–Scholes–Vasicek model. Model, CTMC, and contract parameters are as listed in Tables 4.23 and 4.24 and 4.25, respectively.

(a) $q_t = c_t = 0$ for all $t \in [0, T]$

| $\mathbf{S_0}$ | CTMC | Benchmark | Rel. error |
|---|---|---|---|
| **90** | 105.99171 | 105.99224 | 4.97E-06 |
| **95** | 108.28100 | 108.28568 | 4.33E-05 |
| **100** | 111.08883 | 111.09580 | 6.27E-05 |
| **105** | 114.37154 | 114.37855 | 6.12E-05 |
| **110** | 118.06513 | 118.07046 | 4.51E-05 |
| $\sigma_\mathbf{S}$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.10** | 107.85312 | 107.88135 | 2.62E-04 |
| **0.15** | 109.38108 | 109.39318 | 1.11E-04 |
| **0.20** | 111.08883 | 111.09580 | 6.27E-05 |
| **0.30** | 114.71935 | 114.72313 | 3.30E-05 |
| **0.40** | 118.44863 | 118.45114 | 2.12E-05 |
| $\rho$ | **CTMC** | **Benchmark** | **Rel. error** |
| **-0.3** | 110.80352 | 110.81156 | 7.25E-05 |
| **-0.2** | 111.08883 | 111.09580 | 6.27E-05 |
| **0.2** | 112.14557 | 112.14307 | 2.22E-05 |
| **0.3** | 112.39226 | 112.38624 | 5.36E-05 |

(b) $q_t = 0.02$, $c_t = 0.05$ for all $t \in [0, T]$

| $\mathbf{S_0}$ | CTMC | Benchmark | Rel. error |
|---|---|---|---|
| **90** | 101.80110 | 101.80830 | 7.07E-05 |
| **95** | 104.31639 | 104.32189 | 5.27E-05 |
| **100** | 107.35768 | 107.35983 | 2.00E-05 |
| **105** | 110.84269 | 110.84438 | 1.53E-05 |
| **110** | 114.70970 | 114.70773 | 1.71E-05 |
| $\sigma_\mathbf{S}$ | **CTMC** | **Benchmark** | **Rel. error** |
| **0.10** | 104.28246 | 104.29241 | 9.54E-05 |
| **0.15** | 105.72587 | 105.73050 | 4.38E-05 |
| **0.20** | 107.35768 | 107.35983 | 2.00E-05 |
| **0.30** | 110.84483 | 110.84572 | 7.97E-06 |
| **0.40** | 114.43689 | 114.43768 | 6.86E-06 |
| $\rho$ | **CTMC** | **Benchmark** | **Rel. error** |
| **-0.3** | 107.08461 | 107.08698 | 2.22E-05 |
| **-0.2** | 107.35768 | 107.35983 | 2.00E-05 |
| **0.2** | 108.37107 | 108.36868 | 2.21E-05 |
| **-.3** | 108.60805 | 108.59625 | 1.09E-04 |




(a) $q_t = c_t = 0$ for all $t \in [0, T]$ (b) $q_t = 0.02, c_t = 0.05$ for all $t \in [0, T]$

Figure 4.10: Convergence pattern of the CB price approximations, Algorithm 11, under Black–Scholes–Vasicek model. Except for the number of grid points $M$ of the auxiliary process, the model, CTMC, and contract parameters are as listed in Tables 4.23, 4.24, and 4.25, respectively.

Table 4.27: Approximation of the convergence rate of CB prices, Algorithm 11, under Black–Scholes–Vasicek model. Except for the number of grid points $M$ of the auxiliary process, the model, CTMC, and contract parameters are as listed in Tables 4.23, 4.24, and 4.25, respectively.

(a) $q_t = c_t = 0$ for all $t \in [0, T]$

| m | Rel. error | Rate |
|---|---|---|
| 20 | 5.44E-03 | - |
| 50 | 4.11E-04 | 2.819 |
| 100 | 6.27E-05 | 2.711 |
| 120 | 3.12E-05 | 3.838 |
| 150 | 5.39E-06 | 7.861 |

(b) $q_t = 0.02, c_t = 0.05$ for all $t \in [0, T]$

| m | Rel. error | Rate |
|---|---|---|
| 20 | 5.16E-03 | - |
| 50 | 3.36E-04 | 2.982 |
| 100 | 3.56E-05 | 3.240 |
| 120 | 1.46E-05 | 4.895 |
| 150 | 8.16E-06 | 2.598 |

The convergence patterns of the approximation as $M$ increases are illustrated in Figure 4.10, whereas the approximated convergence rates are shown in Table 4.27. When both credit spread and dividend yield are set to nil ($q_t = c_t = 0$ for all $t \in [0, T]$), the benchmark is calculated using the exact pricing formula of Proposition 4.12.1. When credit risk and dividend yield are considered ($q_t, c_t > 0$ for some $t \in [0, T]$), the benchmark is obtained using CTMC approximation with $M = 1,000$ and $\Delta_N = 1/252$, all other CTMC parameters from table 4.24. Figure 4.10 shows that the approximated prices converge rapidly and smoothly to the benchmark prices.

**INDEX**

**BIBLIOGRAPHY**


Abudy, M. and Izhakian, Y. (2013). Pricing stock options with stochastic interest rate. *International Journal of Portfolio Analysis and Management*, *1*(3), 250–277.

Aït-Sahalia, Y. and Kimmel, R. (2007). Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, *83*(2), 413–452.

Alonso-García, J., Sherris, M., Thirurajah, S. and Ziveyi, J. (2022). *Taxation and policyholder behavior: The case of guaranteed minimum accumulation benefits*. CEPAR Working Paper 2020/17.

Ammann, M., Kind, A. and Wilde, C. (2008). Simulation-based pricing of convertible bonds. *Journal of Empirical Finance*, *15*(2), 310–331.

Auster, J., Mathys, L. and Maeder, F. (2022). JDOI variance reduction method and the pricing of American-style options. *Quantitative Finance*, *22*(4), 639–656.

Ayache, E., Forsyth, P. A. and Vetzal, K. R. (2003). Valuation of Convertible Bonds with Credit Risk. *The Journal of Derivatives*, *11*(1), 9–29.

Bacinello, A. R., Millossovich, P., Olivieri, A. and Pitacco, E. (2011). Variable annuities: A unifying valuation approach. *Insurance: Mathematics and Economics*, *49*(3), 285–297.

Bacinello, A. R. and Zoccolan, I. (2019). Variable annuities with a threshold fee: valuation, numerical implementation and comparative static analysis. *Decisions in Economics and Finance*, *42*, 21–49.

Barone-Adesi, G., Bermudez, A. and Hatgioannides, J. (2003). Two-factor convertible bonds valuation using the method of characteristics/finite elements. *Journal of Economic Dynamics and Control*, *27*(10), 1801–1831.

Bassan, B. and Ceci, C. (2002). Optimal stopping problems with discontinous reward: Regularity of the value function and viscosity solutions. *Stochastics and Stochastic Reports*, *72*(1-2), 55–77.

Battauz, A. and Rotondi, F. (2022). American options and stochastic interest rates. *Computational Management Science*, *19*, 567–604.

Batten, J. A., Khaw, K. L.-H. and Young, M. R. (2014). Convertible bond pricing models. *Journal of Economic Surveys*, *28*(5), 775–803.

Batten, J. A., Khaw, L.-H. K. and Young, M. R. (2018). Pricing convertible bonds. *Journal of Banking & Finance*, *92*, 216–236.

Bauer, D., Gao, J., Moenig, T., Ulm, E. R. and Zhu, N. (2017). Policyholder Exercise Behavior in Life Insurance: The State of Affairs. *North American Actuarial Journal*, *21*(4), 485–501.

Bauer, D., Kling, A. and Russ, J. (2008). A Universal Pricing Framework for Guaranteed Minimum Benefits in Variable Annuities. *ASTIN Bulletin: The Journal of the IAA*, *38*(2), 621–651.

Bauer, D. and Moenig, T. (2023). Cheaper by the bundle: The interaction of frictions and option exercise in variable annuities. *Journal of Risk and Insurance*, *90*(2), 459–486.

Becker, S., Cheridito, P., Jentzen, A. and Welti, T. (2021). Solving high-dimensional optimal stopping problems using deep learning. *European Journal of Applied Mathematics*, *32*(3), 470–514.

Ben-Ameur, H., Breton, M., Karoui, L. and L'Ecuyer, P. (2007). A dynamic programming approach for pricing options embedded in bonds. *Journal of Economic Dynamics and Control*, *31*(7), 2212–2233.

Bensoussan, A. and Lions, J.-L. (1982). *Applications of Variational Inequalities in Stochastic Control*. Noth-Holland Publishing Company.

Bernard, C., Hardy, M. and MacKay, A. (2014a). State-Dependent Fees for Variable Annuity Guarantees. *ASTIN Bulletin: The Journal of the IAA*, *44*(3), 559–585.

Bernard, C. and MacKay, A. (2015). Reducing surrender incentives through fee structure in variable annuities. In K. Glau, M. Scherer, and R. Zagst (dir.), *Innovations in Quantitative Risk Management* 209–223. Springer.

Bernard, C., MacKay, A. and Muehlbeyer, M. (2014b). Optimal surrender policy for variable annuity guarantees. *Insurance: Mathematics and Economics*, *55*, 116–128.

Bernard, C. and Moenig, T. (2019). Where Less Is More: Reducing Variable Annuity Fees to Benefit Policyholder and Insurer. *Journal of Risk and Insurance*, *86*(3), 761–782.

Billingsley, P. (1995). *Probability and Measure* (Third éd.). John Wiley & Sons.

Billingsley, P. (1999). *Convergence of Probability Measures* (2nd éd.). John Wiley & Sons.

Björk, T. (2009). *Arbitrage Theory in Continuous Time* (3rd éd.). Oxford university press.

Black, F., Derman, E. and Toy, W. (1990). A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options. *Financial Analysts Journal*, *46*(1), 33–39.

Black, F. and Karasinski, P. (1991). Bond and Option Pricing when Short Rates are Lognormal. *Financial Analysts Journal*, *47*(4), 52–59.

Brennan, M. J. and Schwartz, E. S. (1977). Convertible Bonds: Valuation and Optimal Strategies for Call and Conversion. *The Journal of Finance*, *32*(5), 1699–1715.

Brigo, D. and Mercurio, F. (2006). *Interest Rate Models – Theory and Practice: With Smile, Inflation and Credit* (2nd éd.). Springer.

Broadie, M. and Glasserman, P. (1997). Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control*, *21*(8-9), 1323–1352.

Broadie, M. and Glasserman, P. (2004). A stochastic mesh method for pricing high-dimensional american options. *Journal of Computational Finance*, *7*(4), 35–72.

Büttler, H.-J. and Waldvogel, J. (1996). Pricing Callable Bonds by Means of Green's Function. *Mathematical Finance*, *6*(1), 53–88.

Cai, N., Kou, S. and Song, Y. (2019). A Unified Framework for Regime-Switching Models. Working paper, Available at SSRN 3310365.

Cai, N., Song, Y. and Kou, S. (2015). A General Framework for Pricing Asian Options Under Markov Processes. *Operations Research*, *63*(3), 540–554.

Carr, P., Jarrow, R. and Myneni, R. (1992). Alternative characterizations of American put options. *Mathematical Finance*, *2*(2), 87–106.

Carr, P. and Sun, J. (2007). A new approach for option pricing under stochastic volatility. *Review of Derivatives Research*, *10*, 87–150.

Carriere, J. F. (1996). Valuation of the early-exercise price for options using simulations and nonparametric regression. *Insurance: Mathematics and Economics*, *19*(1), 19–30.

CBOE (2013a). VIX for Variable Annuities - Part II: A study considering the advantages of tying a Variable Annuity fee to VIX. White Paper.

CBOE (2013b). VIX for Variable Annuities: A study considering the advantages of tying a Variable Annuity fee to VIX. White Paper.

Chambers, D. R. and Lu, Q. (2007). A Tree Model for Pricing Convertible Bonds with Equity, Interest Rate, and Default Risk. *The Journal of Derivatives*, *14*(4), 25–46.

Chiarolla, M. B., De Angelis, T. and Stabile, G. (2022). An analytical study of participating policies with minimum rate guarantee and surrender option. *Finance and Stochastics*, *26*, 173–216.

Chourdakis, K. (2004). Non-Affine Option Pricing. *The Journal of Derivatives*, *11*(3), 10–25.

Clément, E., Lamberton, D. and Protter, P. (2002). An analysis of a least squares regression method for American option pricing. *Finance and Stochastics*, *6*(4), 449–471.

Cox, J. C., Ingersoll Jr., J. E. and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, *53*(2), 385–407.

Cox, J. C., Ross, S. A. and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, *7*(3), 229–263.

Crimaldi, I. and Pratelli, L. (2005). Convergence results for conditional expectations. *Bernoulli*, *11*(4), 737–745.

Cui, Z. (2013). *Martingale Property and Pricing for Time-Homogeneous Diffusion Models in Finance*. (Thèse de doctorat). University of Waterloo.

Cui, Z., Feng, R. and MacKay, A. (2017a). Variable Annuities with VIX-Linked Fee Structure under a Heston-Type Stochastic Volatility Model. *North American Actuarial Journal*, *21*(3), 458–483.

Cui, Z., Kirkby, J. L. and Nguyen, D. (2017b). Equity-linked annuity pricing with cliquet-style guarantees in regime-switching and stochastic volatility models with jumps. *Insurance: Mathematics and Economics*, *74*, 46–62.

Cui, Z., Kirkby, J. L. and Nguyen, D. (2017c). A general framework for discretely sampled realized variance derivatives in stochastic volatility models with jumps. *European Journal of Operational Research*, *262*(1), 381–400.

Cui, Z., Kirkby, J. L. and Nguyen, D. (2018). A General Valuation Framework for SABR and Stochastic Local Volatility Models. *SIAM Journal on Financial Mathematics*, *9*(2), 520–563.

Cui, Z., Kirkby, J. L. and Nguyen, D. (2019). Continuous-Time Markov Chain and Regime Switching Approximations with Applications to Options Pricing. In G. Yin and Q. Zhang (dir.), *Modeling, Stochastic Control, Optimization, and Applications* 115–146. Springer.

Cui, Z., Kirkby, J. L. and Nguyen, D. (2021a). Efficient simulation of generalized SABR and stochastic local volatility models based on Markov chain approximations. *European Journal of Operational Research*, *290*(3), 1046–1062.

Cui, Z., Lee, C. and Liu, M. (2021b). *Valuation of VIX Derivatives through combined Ito-Taylor Expansion and Markov Chain Approximation*. Working paper, Stevens Institute of Technology.

Da Fonseca, J. and Martini, C. (2016). The $\alpha$-hypergeometric stochastic volatility model. *Stochastic Processes and their Applications*, *126*(5), 1472–1502.

De Angelis, T. and Stabile, G. (2019). On Lipschitz Continuous Optimal Stopping Boundaries. *SIAM Journal on Control and Optimization*, *57*(1), 402–436.

Delong, Ł. (2014). Pricing and hedging of variable annuities with state-dependent fees. *Insurance: Mathematics and Economics*, *58*, 24–33.

D'Halluin, Y., Forsyth, P., Vetzal, K. and Labahn, G. (2001). A numerical PDE approach for pricing callable bonds. *Applied Mathematical Finance*, *8*(1), 49–77.

Ding, D., Fu, Q. and So, J. (2012). Pricing callable bonds based on monte carlo simulation techniques. *Technology and Investment*, *3*, 121–125.

Ding, K. and Ning, N. (2021). Markov chain approximation and measure change for time-inhomogeneous stochastic processes. *Applied Mathematics and Computation*, *392*, 125732.

Doriguello, J. F., Luongo, A., Bao, J., Rebentrost, P. and Santha, M. (2021). Quantum algorithm for stochastic optimal stopping problems with applications in finance. *arXiv preprint arXiv:2111.15332*.

Dothan, L. (1978). On the term structure of interest rates. *Journal of Financial Economics*, *6*(1), 59–69.

Drimus, G. G. (2012). Options on realized variance by transform methods: a non-affine stochastic volatility model. *Quantitative Finance*, *12*(11), 1679–1694.

Duffy, D. J. (2006). *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. John Wiley & Sons.

El Karoui, N. (1981). Les Aspects Probabilistes du Contrôle Stochastique. In P. L. Hennequin (dir.), *École d'Été de Probabilités de Saint-Flour IX-1979* 73–238. Springer.

Ethier, S. N. and Kurtz, T. G. (2005). *Markov Processes: Characterization and Convergence*. John Wiley & Sons.

Evans, L. C. (2010). *Partial Differential Equations* (Second éd.). American Mathematical Society.

Fabozzi, F. J., Paletta, T. and Tunaru, R. (2017). An improved least squares monte carlo valuation method based on heteroscedasticity. *European Journal of Operational Research*, *263*(2), 698–706.

Feng, R., Gan, G. and Zhang, N. (2022). Variable annuity pricing, valuation, and risk management: a survey. *Scandinavian Actuarial Journal*, *2022*(10), 867–900.

Friedman, A. (1964). *Partial Differential Equations of Parabolic Type*. Prentice-Hall.

Fu, M. C., Laprise, S. B., Madan, D. B., Su, Y. and Wu, R. (2001). Pricing American options: A Comparison of Monte Carlo Simulation Approaches. *Journal of Computational Finance*, *4*(3), 39–88.

Gao, J. and Ulm, E. R. (2012). Optimal consumption and allocation in variable annuities with Guaranteed Minimum Death Benefits. *Insurance: Mathematics and Economics*, *51*(3), 586–598.

Garcıa, D. (2003). Convergence and biases of monte carlo estimates of american option prices using a parametric exercise rule. *Journal of Economic Dynamics and Control*, *27*(10), 1855–1879.

Garcia, R., Lewis, M.-A., Pastorello, S. and Renault, É. (2011). Estimation of objective and risk-neutral distributions based on moments of integrated volatility. *Journal of Econometrics*, *160*(1), 22–32.

Geman, H., El Karoui, N. and Rochet, J.-C. (1995). Changes of numéraire, changes of probability measure and option pricing. *Journal of Applied Probability*, *32*(2), 443–458.

Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering*, volume 53. Springer.

Goggin, E. M. (1994). Convergence in Distribution of Conditional Expectations. *The Annals of Probability*, 1097–1114.

Grasselli, M. (2017). The 4/2 stochastic volatility model: A unified approach for the Heston and the 3/2 model. *Mathematical Finance*, *27*(4), 1013–1034.

Grimmet, G. R. and Stirzaker, D. R. (2001). *Probability and Random Processes* (Third éd.). Oxford University Press.

Grosen, A. and Jørgensen, P. L. (1997). Valuation of Early Exercisable Interest Rate Guarantees. *The Journal of Risk and Insurance*, *64*(3), 481–503.

Gushchin, V. and Curien, E. (2008). The pricing of convertible bonds within the Tsiveriotis and Fernandes framework with exogenous credit spread: Empirical analysis. *Journal of Derivatives & Hedge Funds*, *14*, 50–65.

Hardy, M. (2003). *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. John Wiley & Sons.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency options. *The Review of Financial Studies*, *6*(2), 327–343.

Heston, S. L. (1997). A Simple New Formula for Options With Stochastic Volatility. Technical report, Available at SSRN 86074.

Ho, T. S. Y. and Lee, S.-B. (1986). Term Structure Movements and Pricing Interest Rate Contingent Claims. *The Journal of Finance*, *41*(5), 1011–1029.

Hogben, H. J., Krzystyniak, M., Charnock, G. T. P., Hore, P. J. and Kuprov, I. (2011). *Spinach*–A software library for simulation of spin dynamics in large spin systems. *Journal of Magnetic Resonance*, *208*(2), 179–194.

Hull, J. and White, A. (1987). The Pricing of Options on Assets with Stochastic Volatilities. *The Journal of Finance*, *42*(2), 281–300.

Hull, J. and White, A. (1990). Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies*, *3*(4), 573–592.

Hull, J. and White, A. (1994). Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models. *The Journal of Derivatives*, *2*(1), 7–16.

Hull, J. and White, A. (1996). Using Hull-White Interest Rate Trees. *The Journal of Derivatives*, *3*(3), 26–36.

Hull, J. C. (2018). *Options, futures, and other derivatives* (10 éd.). Pearson Education.

Hung, M.-W. and Wang, J.-Y. (2002). Pricing Convertible Bonds Subject to Default Risk. *The journal of Derivatives*, *10*(2), 75–87.

Ingersoll Jr., J. E. (1977). A contingent-claims valuation of convertible securities. *Journal of Financial Economics*, *4*(3), 289–321.

Jacka, S. D. (1991). Optimal stopping and the American put. *Mathematical Finance*, *1*(2), 1–14.

Jacka, S. D. and Lynn, J. R. (1992). Finite-horizon optimal stopping, obstacle problems and the shape of the continuation region. *Stochastics and Stochastic Reports*, *39*(1), 25–42.

Jaillet, P., Lamberton, D. and Lapeyre, B. (1990). Variational inequalities and the pricing of american options. *Acta Applicandae Mathematica*, *21*, 263–289.

Jarrow, R. A. and Turnbull, S. M. (1995). Pricing Derivatives on Financial Securities Subject to Credit Risk. *The Journal of Finance*, *50*(1), 53–85.

Jeon, J. and Kwak, M. (2018). Optimal surrender strategies and valuations of path-dependent guarantees in variable annuities. *Insurance: Mathematics and Economics*, *83*, 93–109.

Jönsson, H., Kukush, A. G. and Silvestrov, D. S. (2006). Threshold structure of optimal stopping strategies for American type option. II. *Theory of Probability and Mathematical Statistics*, *72*, 47–58.

Jourdain, B. (2004). *Loss of martingality in asset price models with lognormal stochastic volatility*. Techical report 267, CERMICS.

Kang, B. and Ziveyi, J. (2018). Optimal surrender of guaranteed minimum maturity benefits under stochastic volatility and interest rates. *Insurance: Mathematics and Economics*, *79*, 43–56.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*. Springer-Verlag.

Karatzas, I. and Shreve, S. E. (1998). *Methods of Mathematical Finance*. Springer.

Kim, I. J. (1990). The Analytic Valuation of American Options. *The Review of Financial Studies*, *3*(4), 547–572.

Kirkby, J. L. (2023). Hybrid equity swap, cap, and floor pricing under stochastic interest by Markov chain approximation. *European Journal of Operational Research*, *305*(2), 961–978.

Kirkby, J. L. and Aguilar, J.-P. (2023). Valuation and optimal surrender of variable annuities with guaranteed minimum benefits and periodic fees. *Scandinavian Actuarial Journal*, *2023*(6), 624–654.

Kirkby, J. L., Nguyen, D. and Cui, Z. (2017). A unified approach to bermudan and barrier options under stochastic volatility models with jumps. *Journal of Economic Dynamics and Control*, *80*, 75–100.

Kirkby, J. L., Nguyen, D. H. and Nguyen, D. (2020). A general continuous time Markov chain approximation for multi-asset option pricing with systems of correlated diffusions. *Applied Mathematics and Computation*, *386*, 125472.

Kling, A., Ruez, F. and Ruß, J. (2014). The impact of policyholder behavior on pricing, hedging, and hedge efficiency of withdrawal benefit guarantees in variable annuities. *European Actuarial Journal*, *4*, 281–314.

Kotlow, D. B. (1973). A free boundary problem connected with the optimal stopping problem for diffusion processes. *Transactions of the American Mathematical Society*, *184*, 457–478.

Kouritzin, M. A. and MacKay, A. (2018). VIX-linked fees for GMWBs via explicit solution simulation methods. *Insurance: Mathematics and Economics*, *81*, 1–17.

Kouritzin, M. A. and Zeng, Y. (2005). Weak convergence for a type of conditional expectation: application to the inference for a class of asset price models. *Nonlinear Analysis: Theory, Methods & Applications*, *60*(2), 231–239.

Krylov, N. V. (1980). *Controlled Diffusion Processes*. Springer.

Kuprov, I. (2011). Diagonalization-free implementation of spin relaxation theory for large spin systems. *Journal of Magnetic Resonance*, *209*(1), 31–38.

Kushner, H. J. (1977). *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York.

Kushner, H. J. (1990). Numerical Methods for Stochastic Control Problems in Continuous Time. *SIAM Journal on Control and Optimization*, *28*(5), 999–1048.

Kushner, H. J. and DiMasi, G. B. (1978). Approximations for functionals and optimal control problems on jump diffusion processes. *Journal of Mathematical Analysis and Applications*, *63*(3), 772–800.

Ladyženskaja, O. A., Solonnikov, V. A. and Ural'ceva, N. N. (1968). *Linear and Quasi-linear Equations of Parabolic Type*. American Mathematical Society.

Lamberton, D. (1993). Convergence of the Critical Price In the Approximation of American options. *Mathematical Finance*, *3*(2), 179–190.

Lamberton, D. (1998). American Options. In D. J. Hand and S. D. Jacka (dir.), *Statistic in Finance*. Arnold.

Lamberton, D. (2009). Optimal stopping and American options. Ljubljana Summer School on Financial Mathematics.

Lamberton, D. and Terenzi, G. (2019). Variational Formulation of American Option Prices in the Heston Model. *SIAM Journal on Financial Mathematics*, *10*(1), 261–308.

Leitao Rodriguez, A., Kirkby, J. L. and Ortiz-Gracia, L. (2021). The CTMC-Heston Model: Calibration and Exotic Option Pricing with SWIFT. *Journal of Computational Finance*, *24*(4), 71–114.

Li, L. and Zhang, G. (2018). Error analysis of finite difference and Markov chain approximations for option pricing. *Mathematical Finance*, *28*(3), 877–919.

Lin, S. and Zhu, S.-P. (2020). Numerically pricing convertible bonds under stochastic volatility or stochastic interest rate with an adi-based predictor–corrector scheme. *Computers & Mathematics with Applications*, *79*(5), 1393–1419.

Lin, S. and Zhu, S.-P. (2022). Pricing callable–puttable convertible bonds with an integral equation approach. *Journal of Futures Markets*, *42*(10), 1856–1911.

Lo, C. C. and Skindilias, K. (2014). An improved Markov chain approximation methodology: Derivatives pricing and model calibration. *International Journal of Theoretical and Applied Finance*, *17*(7), 1450047.

Longstaff, F. A. and Schwartz, E. S. (2001). Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies*, *14*(1), 113–147.

Lu, L. and Xu, W. (2017). A Simple and Efficient Two-Factor Willow Tree Method for Convertible Bond Pricing with Stochastic Interest Rate and Default Risk. *The Journal of Derivatives*, *25*(1), 37–54.

Luo, X. and Xing, J. (2021). Optimal Surrender Policy of Guaranteed Minimum Maturity Benefits in Variable Annuities with Regime-Switching Volatility. *Mathematical Problems in Engineering*, *2021*, 1–20.

Ma, C., Xu, W., and Yuan, G. (2020). Valuation model for Chinese convertible bonds with soft call/put provision under the hybrid willow tree. *Quantitative Finance*, *20*(12), 2037–2053.

Ma, J., Yang, W. and Cui, Z. (2021). CTMC intergral equation method for American options under stochastic local volatility models. *Journal of Economic Dynamics and Control*, *128*, 104145.

Ma, J., Yang, W. and Cui, Z. (2022). Convergence analysis for continuous-time Markov chain approximation of stochastic local volatility models: Option pricing and Greeks. *Journal of Computational and Applied Mathematics*, *404*, 113901.

MacKay, A. (2014). *Fee Structure and Surrender Incentives in Variable Annuities*. (Thèse de doctorat). University of Waterloo.

MacKay, A., Augustyniak, M., Bernard, C. and Hardy, M. R. (2017). Risk Management of Policyholder Behavior in Equity-Linked Life Insurance. *Journal of Risk and Insurance*, *84*(2), 661–690.

MacKay, A., Vachon, M.-C. and Cui, Z. (2023). Analysis of VIX-linked fee incentives in variable annuities via continuous-time Markov chain approximation. *Quantitative Finance*, *23*(7-8), 1055–1078.

McConnell, J. J. and Schwartz, E. S. (1986). LYON taming. *The Journal of Finance*, *41*(3), 561–576.

Mentink-Vigier, F. (2020). Fast exponential matrix for Matlab (full/sparse), fastExpm. Available at Github `https://github.com/fmentink/fastExpm`. Accessed: April 27, 2023.

Mercurio, F. and Moraleda, J. M. (2001). A family of humped volatility models. *The European Journal of Finance*, *7*(2), 93–116.

Mijatović, A. and Pistorius, M. (2009). Continuously monitored barrier options under Markov processes (unabridged version with Matlab code). Available at SSRN 1462822.

Mijatović, A. and Pistorius, M. (2013). Continuously monitored barrier options under Markov processes. *Mathematical Finance*, *23*(1), 1–38.

Milanov, K., Kounchev, O., Fabozzi, F. J., Kim, Y. S. and Rachev, S. T. (2013). A Binomial-Tree Model for Convertible Bond Pricing. *The Journal of Fixed Income*, *22*(3), 79–94.

Milevsky, M. A. and Salisbury, T. S. (2001). The Real Option to Lapse a Variable Annuity: Can Surrender Charges Complete the Market? Dans *Conference Proceedings of the 11th Annual International AFIR Colloquium*.

Moenig, T. and Bauer, D. (2016). Revisiting the Risk-Neutral Approach to Optimal Policyholder Behavior: A Study of Withdrawal Guarantees in Variable Annuities. *Review of Finance*, *20*(2), 759–794.

Moenig, T. and Zhu, N. (2018). Lapse-and-Reentry in Variable Annuities. *Journal of Risk and Insurance*, *85*(4), 911–938.

Moenig, T. and Zhu, N. (2021). The Economics of a Secondary Market for Variable Annuities. *North American Actuarial Journal*, *25*(4), 604–630.

Niittuinperä, J. (2022). Chapter 18 – Policyholder Behavior and Management Actions. In *IAA Risk Book*.

Ostrovski, V. (2013). Efficient and Exact Simulation of the Hull-White Model. Available at SSRN 2304848.

Palczewski, J. and Stettner, Ł. (2010). Finite Horizon Optimal Stopping of Time-Discontinuous Functionals with Applications to Impulse Control with Delay. *SIAM Journal on Control and Optimization*, *48*(8), 4874–4909.

Palmer, B. A. (2006). *Equity-Indexed Annuities: Fundamental Concepts and Issues*. Rapport technique, Insurance Information Institute.

Peskir, G. and Shiryaev, A. (2006). *Optimal Stopping and Free-Boundary Problems*. Springer.

Rebonato, R. (2004). *Volatility and Correlation: The Perfect Hedger and the Fox* (Second éd.). John Wiley & Sons.

Reesor, M. R., Stentoft, L. and Zhu, X. (2024). A critical analysis of the Weighted Least Squares Monte Carlo method for pricing American options. *Finance Research Letters*, p. 105379.

Rindos, A., Woolet, S., Viniotis, I. and Trivedi, K. (1995). Exact Methods for the Transient Analysis of Nonhomogeneous Continuous Time Markov Chains. Dans W. J. Stewart (dir.). *Computations with Markov Chains*, 121–133. Springer.

Rudin, W. (1991). *Functional Analysis* (Second éd.). McGraw-Hill.

Scott, L. O. (1987). Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application. *Journal of Financial and Quantitative Analysis*, *22*(4), 419–438.

Shen, Y., Sherris, M. and Ziveyi, J. (2016). Valuation of guaranteed minimum maturity benefits in variable annuities with surrender options. *Insurance: Mathematics and Economics*, *69*, 127–137.

Sidje, R. B. (1998). Expokit: a software package for computing matrix exponentials. *ACM Transactions on Mathematical Software*, *24*(1), 130–156.

Sin, C. A. (1998). Complications with stochastic volatility models. *Advances in Applied Probability*, *30*(1), 256–268.

Song, Q., Yin, G. and Zhang, Q. (2013). Weak Convergence Methods for Approximation of the Evaluation of Path-Dependent Functionals. *SIAM Journal on Control and Optimization*, *51*(5), 4189–4210.

Stentoft, L. (2004). Convergence of the least squares monte carlo approach to american option valuation. *Management Science*, *50*(9), 1193–1203.

Tavella, D. and Randall, C. (2000). Pricing Financial Instruments: The Finite Difference Method. *John Willey & Sons*.

Terenzi, G. (2019). *Option prices in stochastic volatility models*. (Thèse de doctorat). Università di Roma Tor Vergata and Université Paris-Est Marne-La-Vallée.

Touzi, N. (1999). American Options Exercise Boundary When the Volatility Changes Randomly. *Applied Mathematics and Optimization*, *39*, 411–422.

Tsitsiklis, J. N. and Van Roy, B. (2001). Regression methods for pricing complex american-style options. *IEEE Transactions on Neural Networks*, *12*(4), 694–703.

Tsiveriotis, K. and Fernandes, C. (1998). Valuing Convertible Bonds with Credit Risk. *The Journal of Fixed Income*, *8*(2), 95–102.

Vachon, M.-C. and MacKay, A. (2024). A unifying approach for the pricing of debt securities. *Quantitative Finance*, 1–26.

Van Moerbeke, P. (1974). Optimal Stopping and Free Boundary Problems. *The Rocky Mountain Journal of Mathematics*, *4*(3), 539–578.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, *5*(2), 177–188.

Veretennikov, A. J. (1981). On Strong Solutions and Explicit Formulas for Solutions of Stochastic Integral Equations. *Mathematics of the USSR-Sbornik*, *39*(3), 387–403.

Villeneuve, S. (1999). Exercise regions of American options on several assets. *Finance and Stochastics*, *3*, 295–322.

Villeneuve, S. (2007). On Threshold Strategies and the Smooth-Fit Principle for Optimal Stopping Problems. *Journal of Applied Probability*, *44*(1), 181–198.

Wang, G. and Zou, B. (2021). Optimal fee structure of variable annuities. *Insurance: Mathematics and Economics*, *101*, *Part B*, 587–601.

Wei, W. and Zhu, D. (2022). Generic improvements to least squares monte carlo methods with applications to optimal stopping problems. *European Journal of Operational Research*, *298*(3), 1132–1144.

Xiong, L., Luo, J., Vise, H. and White, M. (2023). Distributed Least-Squares Monte Carlo for American Option Pricing. *Risks*, *11*(8), 145.

Zhang, G. and Li, L. (2019). Analysis of Markov Chain Approximation for Option Pricing and Hedging: Grid Design and Convergence Behavior. *Operations Research*, *67*(2), 407–427.

Zhang, G. and Li, L. (2021). *A General Approach for Parisian Stopping Times under Markov Processes*. Working paper, Available at arXiv:2107.06605.

Zhao, Y. (2017). *On properties of the American put option under several models*. (Thèse de doctorat). University of Warwick.

Zhu, Y. and Zhang, J. E. (2007). Variance term structure and VIX futures pricing. *International Journal of Theoretical and Applied Finance*, *10*(1), 111–127.

Zhu, Z. and Welsch, R. E. (2015). Statistical learning for variable annuity policyholder withdrawal behavior. *Applied Stochastic Models in Business and Industry*, *31*(2), 137–147.