Université de Montréal

Simulation and Numerical Procedures for Option Pricing

par

Hatem Ben Ameur

École des Hautes Études Commerciales

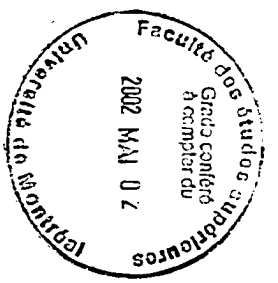
Thèse présentée à la

Faculté des Études Supérieures

en vue de l'obtention du grade de

Philosophiae Doctor (Ph.D.)

en administration



Novembre 2001

Université de Montréal

Cette thèse intitulée
Simulation and Numerical Procedures for Option Pricing

présentée par
Hatem Ben Ameur

et soutenue publiquement, a été évaluée par un jury composé des personnes suivantes:

Désiré Vencatachellum
.....
(président rapporteur de l'École des HEC)

Michèle Breton
.....
(codirectrice de recherche de l'École des HEC)

Pierre L'Écuyer
.....
(codirecteur de recherche de l'Université de Montréal)

Lawrence Kryzanowski
.....
(membre du jury de Concordia University)

Phelim Boyle
.....
(examinateur externe de l'University of Waterloo)

Thèse acceptée le:
12 octobre 2001
.....

Résumé

Cette thèse est composée de trois essais portant sur l'évaluation des options. Le premier propose de mesurer le risque encouru par le signataire d'une option s'il décidait d'une couverture partielle. Le second porte sur l'évaluation des options asiatiques de type américain. Le dernier aborde les clauses de rachat et de remboursement implicites aux obligations. On donne ici leurs résumés.

Dans le premier essai, on considère un signataire d'une option européenne qui décide de couvrir sa position pour certains états de la nature et non pas pour d'autres. Une couverture partielle coûte moins à mettre en oeuvre mais peut aboutir à un défaut. On se propose alors de mesurer le gain et la probabilité de défaut qui en résultent en exprimant l'événement de couverture en fonction du prix du sous-jacent à l'échéance, du maximum des prix et du temps d'atteinte de ce maximum. Dans un premier temps, on se place sous les hypothèses du modèle de Black et Scholes (1973) et on dérive des formules analytiques pour ces indicateurs. Dans un second temps, on considère un modèle à volatilité stochastique et on estime ces indicateurs par simulation de Monte Carlo rendue efficace via les techniques de réduction de variance dites d'induction de corrélation. Les résultats montrent de quelle manière le gain et le risque de défaut sont inversement reliés.

Le second essai porte sur l'évaluation des options asiatiques de type américain sous les hypothèses du modèle de Black et Scholes (1973). Ces options n'admettent pas de solution analytique, d'où le recours à des approches numériques. On pro-

pose une méthode basée sur la programmation dynamique et les éléments finis. Étant donnée la fonction de valeur de l'option à une certaine date, le principe est de l'estimer par une interpolation polynomiale par morçeaux, ce qui permet de valoriser l'option à une date antérieure. En réitérant le même principe depuis l'échéance jusqu'à la date présente, on obtient la valeur de l'option et la stratégie optimale de son exercice. Cette procédure est une alternative viable aux méthodes de différences finies et aux arbres (binomiaux et trinomiaux) surtout si les opportunités d'exercice sont limitées. À l'aide de cette formulation, on établit les propriétés théoriques de la fonction de valeur en fonction du prix et de la moyenne des prix du sous-jacent. Les expériences numériques confirment la convergence, la cohérence et l'efficacité de cette méthode.

Dans le troisième essai, on aborde les options de rachat et de remboursement implicites aux obligations. Ces clauses contractuelles sont de type américain et ne peuvent être valorisées d'une façon analytique. La règle est de recourir à des approches numériques. On formule la question d'évaluation des options de rachat et de remboursement des obligations comme un problème de programmation dynamique dans lequel on modélise le taux sans risque de courte période comme un processus de diffusion tel que proposé par Vasicek (1977). Étant donnée la fonction de valeur de l'obligation rachetable et remboursable à une certaine date, le principe est de l'estimer par une approximation polynomiale par morceaux, ce qui permet de valoriser l'obligation à une date antérieure. En réitérant le même principe depuis

l'échéance jusqu'à la date présente, on évalue ces options et on identifie leurs stratégies optimales d'exercice. À l'aide de la même formulation, on établit les propriétés théoriques de la fonction de valeur d'une obligation rachetable et remboursable. Les expériences numériques indiquent que la méthode est stable, cohérente et efficace.

Abstract

This thesis contains three essays on option pricing. The first discusses a way to measure the risk of the short trader of an option if he decides to hedge partially his position. The second deals with the pricing of American-style Asian options, and the third with call and put options embedded in bonds.

The first essay considers the case of a short trader of an option who decides to partially hedge its position on some states of nature and not on the others. A partial hedge initially costs less than a perfect hedge, but may lead to a default position. It is of interest in that context to estimate the gain and the default risk. Some partial hedging strategies based on the price of the primitive asset at the horizon, its maximum over the trading period, and the time to maximum, are analyzed. Closed-form solutions are derived in the Black and Scholes (1973) model, and efficient Monte Carlo estimators are computed in a stochastic volatility model. The results show how the gain and the default risk inversely change depending on the hedging event.

The second essay deals with the pricing of Asian options based on the arithmetic average, under the Black and Scholes (1973) model. Their evaluation involves estimating an integral (a mathematical expectation) for which no analytical solution is available. Pricing their American-style counterparts, which provide

early exercise opportunities, poses the additional difficulty of solving a dynamic optimization problem to determine the optimal exercise strategy. We develop a numerical method for pricing American-style Asian options based on dynamic programming combined with finite-element piecewise-polynomial approximation of the value function. This method is a viable alternative to Finite Differences algorithms and the lattice-based approach when the early exercise opportunities are limited. Numerical experiments show convergence, consistency, and efficiency. Some theoretical properties of the value function of the option and its optimal exercise strategy are also established.

In the third essay, we consider call and put options embedded in bonds. These options cannot be priced in a closed-form as it is generally the case for American-style financial derivatives. We formulate the problem of pricing the embedded call and put options as a stochastic Dynamic Programming (DP) model. We let the short-term risk-free interest rate move as in Vasicek (1977). We approximate the bond value by a piecewise linear interpolation at each step of the DP procedure, and solve the DP equation in closed-form. This yields both the options' values and their optimal exercice strategies. Then, we use the DP formulation to establish the basic properties of bonds, price their embedded call and put options, and determine their optimal exercise strategies. Numerical investigation shows stability, consistency, and efficiency.

Remerciements

Cette thèse expose les résultats d'un travail élaboré dans une ambiance à la fois sereine et amicale. On dit de cette expérience qu'elle est la plus difficile et la plus risquée de toute une vie. C'est certes vrai mais c'est aussi l'expérience la plus enrichissante et la plus sûre pour le restant d'une vie.

Ce travail n'aurait pu être accompli sans l'aide da ma famille. Je remercie ma femme Sonia et mes filles Beya et Sarah pour leur encouragement continu et inconditionnel. Mon fils Ilyès est encore trop jeune, mais il a surement remarqué mon absence prolongée lors des dernières corrections apportées à cette thèse. En souvenir de ma grande mère Beya qui m'a enseigné mes premiers devoirs. Que Dieu lui accorde sa bénédiction. Je remercie mes parents et mes beaux-parents pour leur soutien durant mon absence. Je remercie mon frère, ma soeur, mon beau-frère et mes belles-soeurs ainsi que les membre de leurs familles pour tout ce qu'ils ont fait pour la mienne.

Je vous remercie, Michèle et Pierre, pour votre encadrement et votre soutien. Vous m'avez accordé toutes les ressources dont j'avais besoin. Votre apport au niveau des questions de recherche, de la méthodologie et de l'élaboration des résultats était primordial. À titre d'exemple, je cite le travail de Michèle au niveau de l'élaboration des propriétés théoriques de la fonction de valeur d'une option

américaine. Je cite aussi le travail de Pierre au niveau de la preuve de convergence de l'algorithme de programmation dynamique (voir Chapitre 2) servant à évaluer une option asiatique de type américain et de l'amélioration de l'efficacité du code informatique correspondant. J'adresse aussi mes remerciements à M. Lawrence Kryzanowski en tant que membre de mon jury de thèse pour ses commentaires pertinents tout le long de mon cursus doctoral.

Je souhaite une bonne continuation pour mes amis d'étude, mes anciens professeurs, mes collègues à l'Université d'Économie et de Gestion de Tunis, les membres du programme doctoral et du département de méthodes quantitatives à l'École des HEC.

J'exprime ma gratitude envers ceux et celles qui m'ont soutenu financièrement durant ma thèse. Je cite le Ministère Tunisien de l'Enseignement Supérieur pour la bourse d'étude qui m'a été accordée. Je remercie notamment M. Afif Hindaoui, recteur de l'Université d'Économie et de Gestion de Tunis, qui a défendu mon dossier de bourse, Mme Michèle Breton qui m'a accordé un appui financier continu à travers ses fonds de recherche, le programme doctoral et la direction de la recherche de l'École des HEC et M. Pierre L'Écuyer pour avoir engagé M. Richard Simard afin d'étendre la portée de mon code informatique. Le travail de M. Simard a été précieux: Il a amélioré l'efficacité de mon code informatique relatif au deuxième essai, implanté plusieurs interpolations de fonctions à 2 variables et proposé

une méthode d'extrapolation, appelée ϵ-algorithm, servant à accélérer la convergence d'une série numérique.

Contents

Derivatives Markets **1**

1 Partial Hedging for Options Based on Extreme Values and Passage Times **3**

1.1 Introduction 3

1.1.1 The Market 3

1.1.2 The No-Arbitrage Property 5

1.1.3 The Completeness Property 7

1.2 Partial Hedging in the Black and Scholes Model 9

1.2.1 Partial Hedging when $A = \{E \leq S(T) \leq a\}$ 11

1.2.2 Partial Hedging when $A = \{E \leq S(T) \leq a, M^S(T) \leq b\}$ 13

1.2.3 Partial Hedging on $\{E \leq S(T) \leq a, M^S(T) \leq b, \theta^S(T) \leq s\}$ 15

1.3 Partial Hedging in a Stochastic Volatility Model 17

1.3.1 A Monte Carlo Experiment 17

1.3.2 Antithetic Variates 22

1.3.3 Control Variates 24

1.3.4 Integrating the Correlation Induction Techniques 27

1.4 Conclusion 27

1.5 References 28

2 Numerical Procedures for Pricing American-style Asian Options... 31

2.1 Introduction 31

2.2 Model and Notation 37

2.2.1 Evolution of the Primitive Asset 37

2.2.2 The Amerasian Contract 38

2.3 The Dynamic Programming Formulation 38

2.4 Characterizing the Value Function and the Exercise Strategy 40

2.4.1 The Value Function v_{n-1} 40

2.4.2 Properties of the Value Function and of the Exercise Frontier 42

2.5 Numerical Solution of the DP Equation 47

2.5.1 A Piecewise Approximation 48

2.5.2 Explicit Integration for Function Evaluation 50

2.5.3 Computational Speed-up and Complexity Analysis 53

2.5.4 Convergence 54

2.5.5 Grid Choice, Refinement, and Convergence Acceleration 58

2.6 Numerical Experiments and Examples 59

2.6.1 Example 1 60

2.6.2 Example 2 63

2.6.3 Example 3 64

2.7 Conclusion 66

2.8 References 67

3 Pricing Call and Put Options Embedded in Bonds 71

3.1 Introduction 71

3.2 The Model 77
3.3 The DP Formulation 83
3.4 Properties of the Bond's Value 85
3.5 Solving the DP Equation 86
3.6 Numerical Experiments 89
3.7 Conclusion 91
3.8 References 92

Derivatives Markets

A financial *derivative* is a contract which provides its holder a future payment that depends on the price of one or several underlying asset(s). The asset underlying a derivative may be a stock, an index, an interest rate, a foreign currency, or a commodity held for investment or consumption. This shows the extreme variety of derivative contracts.

Options are particular derivatives characterized by non-negative payoffs, hence their interpretation as privileges. The most popular are *call* and *put* options. A call (respectively a put) option gives its holder the right to buy (respectively to sell) an asset at or before a known future date for a known price. *Futures* and *swaps* are others popular derivatives. A future is an agreement to buy or to sell an asset at a certain time in the future for a certain price. A swap is an agreement to exchange cash flows in the future according a prearranged formula.

In derivatives markets, there are *hedgers*, *speculators* and *arbitrageurs*. Hedgers trade to cover their positions. Speculators take risk deliberately in attempt to make high returns, while arbitrageurs seek abnormal spreads to make riskless profits. The explosive growth of derivatives markets is explained in large part by a great need for hedging against price fluctuations, and derivatives may be designed to hedge almost all kinds of market risk.

The growth of derivatives markets has improved the brokerage industry. In response, the mathematics and their numerical implementations to evaluate derivatives grew with the same intensity since the early 1970s, and, in conjunction, a new field of research, called *financial engineering*, was initiated.

For complex derivatives, such as path-dependent European options or American-style derivatives, the *efficiency* of the pricing methods becomes a major field of research in financial engineering. Our work belongs to this field. Specifically, we focus on the implementation of some efficient Monte Carlo and Dynamic Programming procedures for option pricing.

Chapter 1
Partial Hedging for Options Based on Extreme Values and Passage Times

1.1 Introduction

1.1.1 The Market

The model described in this section is presented in detail in Karatzas (1996), Chapters 0 and 1. Let $\mathcal{M}$ be a market with $p+1$ traded assets in which trading takes place continuously over the period $[0,T]$. In our presentation, a process $X(t)$ for $t \in [0,T]$ is denoted by $X(\cdot)$. Let $W(\cdot) = (W_1(\cdot), \ldots, W_d(\cdot))'$ be a d-dimensional Brownian motion defined over the probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{F}(\cdot)$ its P-augmented natural filtration. The σ-algebra $\mathcal{F}$ can be chosen as $\mathcal{F}(T)$. The Brownian motions $W_j(\cdot)$, for $j = 1, \ldots, d$, can be interpreted as d sources of systematic risk and the filtration $\mathcal{F}(\cdot)$ as a collection of the increasing sets of information available to investors over time. All the stochastic processes are assumed to be adapted to the filtration $\mathcal{F}(\cdot)$ and verify some additional conditions which guarantee their existence and uniqueness. The adaptability requirement allows dependence on past realizations and precludes anticipation of future realizations. Assume that the market $\mathcal{M}$ is frictionless, that is, there are no taxes, transaction costs, information asymmetries, constraints on short selling or borrowing.

The first asset, called the *bank account*, does not pay dividends. It starts at unity, and moves according to the differential equation

$$dB(t) = B(t)\, r(t)\, dt, \quad \text{for } B(0) = 1 \text{ and } 0 \leq t \leq T, \tag{1.1}$$

where $r(\cdot)$ is the process of risk-free interest-rate. The solution of (1.1) yields the following definition for the *discount factor*

$$\gamma(t) = 1/B(t) = e^{-\int_0^t r(s)ds}, \quad \text{for } 0 \leq t \leq T.$$

If the interest rate process is constant, $r(\cdot) = r$, the discount factor can be written as

$$\gamma(t) = e^{-rt}, \quad \text{for } 0 \leq t \leq T. \tag{1.2}$$

The p remaining assets, called the *primitive assets*, move according to the stochastic differential equation (SDE)

$$\begin{aligned} dS_i(t) &= \mu_i(t)\, S_i(t)\, dt + \sum_{j=1}^{d} \sigma_{ij}(t)\, S_i(t)\, dW_j(t), \\ &\text{for } S_i(0) > 0,\ i = 1, \ldots, p, \text{ and } 0 \leq t \leq T, \end{aligned} \tag{1.3}$$

where $\mu(\cdot) = (\mu_1(\cdot), \ldots, \mu_p(\cdot))'$ is the vector-process of the *appreciation rates* and $\sigma(\cdot) = [\sigma_{ij}(\cdot)]$ is the matrix-process of *volatility*. For simplicity, assume that the primitive assets do not pay dividends.

The Black and Scholes (1973) model assumes a constant interest rate, one source of systematic risk, and one primitive asset. Specifically:

$$dS(t) = \mu S(t)\, dt + \sigma S(t)\, dW(t), \quad \text{for } S(0) > 0 \text{ and } 0 \leq t \leq T, \tag{1.4}$$

where μ and $\sigma > 0$ are assumed to be constants. Another interesting case is a stochastic volatility model similar to those discussed by Hull and White (1987), Johnson and Shanno

(1987), Scott (1987), and Wiggins (1987). Specifically,

$$dS(t) = \mu S(t)dt + \sigma(t)S(t)dW_1(t), \quad \text{for } S(0) > 0 \text{ and } 0 \leq t \leq T, \tag{1.5}$$

where the volatility process $\sigma(\cdot)$ is random and depends on a two-dimensional Brownian motion $W(\cdot) = (W_1(\cdot), W_2(\cdot))'$.

1.1.2 The No-Arbitrage Property

An investor trades continuously on the $p+1$ traded assets by managing a self-financing portfolio which generates the wealth process

$$X(t) = \sum_{i=1}^{p} \pi_i(t) + \left(X(t) - \sum_{i=1}^{p} \pi_i(t) \right), \quad \text{for } 0 \leq t \leq T,$$

where $\pi(\cdot) = (\pi_1(\cdot), \ldots, \pi_p(\cdot))'$, called the *portfolio strategy*, is the vector-process of the dollar amounts invested in the primitive assets. Notice that the components of this portfolio strategy can be positive or negative, depending on the position (long or short) of the investor. The wealth process is denoted by $X^{x,\pi}(\cdot)$ since it depends on the initial wealth $X(0) = x$ and the portfolio strategy $\pi(\cdot)$.

An *arbitrage opportunity* is a strategy $\pi(\cdot)$ such that

$$P(X^{0,\pi}(T) \geq 0) = 1 \text{ and } P(X^{0,\pi}(T) > 0) > 0.$$

A rational investor should take this opportunity when it appears because it costs nothing to hold and may lead to a positive final wealth. A market without arbitrage opportunities is called *arbitrage-free*. There are two sufficient conditions for the arbitrage-free property to be verified. Firstly, there exists a process $\theta : [0,T] \times \Omega \to \mathbb{R}^d$, called the *market-price of*

risk, which verifies

$$\mu(\cdot) - r(\cdot)\,1_p = \sigma(\cdot)\,\theta(\cdot), \tag{1.6}$$

where $1_p = (1, \ldots, 1)' \in \mathbb{R}^p$. Secondly, the exponential process

$$Z(t) = e^{-\frac{1}{2}\int_0^t \|\theta(t)\|^2 dt - \int_0^t \theta(t)' dW(t)}, \quad \text{for } 0 \leq t \leq T, \tag{1.7}$$

is a P-martingale. In practice, it is somewhat hard to check for (1.7). However, a sufficient requirement for (1.7) to hold is the so-called Novikov condition

$$E^P\left[e^{-\frac{1}{2}\int_0^t \|\theta(t)\|^2 dt}\right] < +\infty. \tag{1.8}$$

The market of Black and Scholes (1973) defined in (1.4) is arbitrage-free since $\theta = (\mu - r)/\sigma$ does exist and verifies the Novikov condition (1.8). In the stochastic volatility models introduced in (1.5), the volatility process $\sigma(\cdot)$ is chosen such that the model is arbitrage-free.

Assume that the conditions (1.6) and (1.8) hold and that $p \leq d$. The market is said to be *standard*. The exponential process $Z(\cdot)$ defined in (1.7) allows the construction of a collection of P-equivalent probability measures Q_t, for $t \in [0, T]$, each defined on the corresponding σ-algebra $\mathcal{F}(t)$ by

$$Q_t(A) = E^P[Z(t)\,1_A], \quad \text{for } A \in \mathcal{F}(t).$$

The probability measure Q_T, denoted by Q, is called the *risk-neutral* probability measure. An important feature of Q is that the process

$$B(t) = W(t) + \int_0^t \theta(s)\,ds, \quad \text{for } 0 \leq t \leq T, \tag{1.9}$$

is a Q-Brownian motion. In turn, the equations described in (1.3) can be transformed into

$$dS_i(t) = r(t) S_i(t) dt + \sum_{j=1}^{d} \sigma_{ij}(t) S_i(t) dB_j(t), \quad \text{for } S_i(0) > 0, i = 1, \ldots, p, \text{ and } 0 \leq t \leq T, \tag{1.10}$$

or, equivalently, into

$$d(\gamma(t) S_i(t)) = \gamma(t) \sum_{j=1}^{d} \sigma_{ij}(t) S_i(t) dB_j(t), \quad \text{for } S_i(0) > 0, i = 1, \ldots, p, \text{ and } 0 \leq t \leq T, \tag{1.11}$$

where $B(\cdot) = (B_1(\cdot), \ldots, B_d(\cdot))'$ is a Q-Brownian motion as expressed in (1.9). Equation (1.10) shows that one can ignore the appreciation rate $\mu(\cdot)$ when computing expectations in the form $E^Q[f(S(t), t \in I \subset [0,T])]$, where $f : \mathbb{R}^{|I|} \to \mathbb{R}$. Equation (1.11) says that the discounted price of each primitive asset is a Q-martingale. The notion of efficiency implied by this result is that the "best prediction" of $\gamma(T) S(T)$ is $S(0)$.

1.1.3 The Completeness Property

An *option* is any non-negative and $\mathcal{F}(T)$-measurable random variable Y such that

$$E^Q[\gamma(T) Y] < +\infty.$$

One can think of a contract that gives a payoff of Y at time T. This contract should be interpreted as a privilege since it gives its holder a non-negative amount.

A *call* option pays $C(T) = (S(T) - E)^+$ at time T, where $x^+ = \max(x, 0)$ for $x \in \mathbb{R}$. This contract gives its holder the right to buy the primitive asset at time T at a specified *exercise price* E. A *put* option pays $P(T) = (E - S(T))^+$ at time T. This contract gives its holder the right to sell the primitive asset at time T at a specified exercise price E.

Notice that

$$C(T) - P(T) = S(T) - E. \tag{1.12}$$

A *complete market* is a market on which any contingent claim Y is attainable by a well-selected self-financing portfolio strategy $\pi(\cdot)$:

$$\forall Y, \exists \pi(\cdot) \text{ such that } X^{v,\pi}(T) = Y, \tag{1.13}$$

where

$$v = E^{Q}[\gamma(T) Y]. \tag{1.14}$$

In a complete market, one can start at the initial wealth v defined in (1.14) and find a portfolio strategy $\pi(\cdot)$ as in (1.13), called the *replication portfolio*, such that the final wealth matches with certainty the option payoff. In a standard and complete market, the discounted wealth process $\gamma(\cdot) X^{v,\pi}(\cdot)$ satisfies the martingale property

$$\gamma(s) X^{v,\pi}(s) = E^{Q}[\gamma(t) X^{v,\pi}(t) \mid \mathcal{F}(s)], \quad \text{for } 0 \leq s \leq t \leq T. \tag{1.15}$$

Equation (1.14) can be obtained from (1.15) at $s = 0$ and $t = T$. In fact, the wealth process $X^{v,\pi}(\cdot)$ matches with certainty the option's value during all the trading period $[0, T]$, otherwise arbitrage opportunities will appear. This is why the wealth process is called the *price-process* of the option. In a standard and complete market, the value v defined in (1.14) is the unique rational price of the contingent claim Y. If the market is only standard, there exists in general a whole range of prices

$$I_v = [v_{\min}, v_{\max}],$$

including the value defined in (1.14), that are consistent with the arbitrage-free property.

Assume that the interest rate is constant as in (1.2). Using equations (1.12) and (1.14) and the martingale property of $\gamma(\cdot) S(\cdot)$ described in (1.15), one obtains the put-call parity relationship

$$c - p = S(0) - e^{-rT} E, \tag{1.16}$$

where c and p are respectively the prices of the call option and its corresponding put option. Equation (1.16) shows that the value of a put option can often be deduced from the value of a call option.

A simple criterion exists to check for the completeness property in a standard market. A standard market is complete if and only if

$$p = d, \tag{1.17}$$

and

$$\sigma(t, \omega) \text{ is regular for } t \in [0, T] \text{ and } \omega \in \Omega. \tag{1.18}$$

Roughly speaking, the completeness property is a question of dimension: There must be as many sources of systematic risk as primitive assets. The Black and Scholes model described in (1.4) is complete since it is standard and verifies $p = d = 1$ and $\sigma(t, \omega) = \sigma > 0$ for $t \in [0, T]$ and $\omega \in \Omega$. On the other hand, based on equation (1.17), the stochastic volatility model described in (1.5) is *incomplete*.

1.2 Partial Hedging in the Black and Scholes Model

Assume that the market is standard and complete. The seller of a contingent claim Y can hedge perfectly all risk by starting at the initial wealth v defined in (1.14) and managing

the replication portfolio $\pi(\cdot)$ mentioned in (1.13). Let 1_A be the indicator of an event $A \in \mathcal{F}(T)$. Instead of hedging Y, this investor may want to hedge $Y1_A$. This is less expensive to replicate, since if we define $u = E^Q[\gamma(T)Y1_A]$ and $v = E^Q[\gamma(T)Y]$, then

$$g = v - u > 0. \tag{1.19}$$

The replication takes place only on the hedging event $H = \{Y = 0\} \cup A$. The *gain* is defined as $g = v - u$ and the *default event* is defined as $H^c = \{Y > 0\} \cap A^c$. The *default risk* is

$$P(H^c) = E^P[1_{H^c}]. \tag{1.20}$$

Notice that $P(H^c)$ depends on the appreciation rate $\mu(\cdot)$ which is assumed to be constant in the following.

In the next subsections, some partial hedging strategies are analyzed in terms of the gain and default risk. These strategies account for the price of the primitive asset at the horizon, its maximum during the trading period, and the time at which this maximum occurs. In the first subsection, closed-form solutions are derived for the Black and Scholes model. In the second subsection, efficient Monte Carlo estimators are developed for a stochastic volatility model.

1.2.1 Partial Hedging when $A = \{E \leq S(T) \leq a\}$

Consider a partial replication of the call option $Y = (S(T)-E)^+$ when $A = \{E \leq S(T) \leq a\}$ for some real a greater than E. By equation (1.19), the gain is

$$\begin{aligned} g &= v - u, \\ &= E^Q\left[(S(T)-E)^+\right] - E^Q\left[(S(T)-E)1_{\{E\leq S(T)\leq a\}}\right]. \end{aligned}$$

The cost of the perfect hedge, denoted by v, is the Black and Scholes price which is known in closed-form. The cost of the partial hedge, denoted by u, can also be computed in closed-form as follows. In the Black and Scholes (1973) model defined in (1.4), the price of the primitive asset at the horizon can be written as

$$S(T) = S(0)\exp\left((r-\sigma^2/2)T + \sigma\sqrt{T}Z\right),$$

where Z is a standard normal random variable. The primitive asset price $S(T)$ is then lognormal, and from this we can derive (after some algebraic manipulations) the following expression for u:

$$\begin{aligned} u &= S(0)\left[N(d_1) - N(d_1')\right] - Ee^{-rT}\left[N(d_2) - N(d_2')\right], \qquad (1.21) \\ d_1 &= \left[\ln(S(0)/E) + (r+\sigma^2/2)T\right]/\sigma\sqrt{T}, \\ d_2 &= d_1 - \sigma\sqrt{T}, \\ d_1' &= \left[\ln(S(0)/a) + (r+\sigma^2/2)T\right]/\sigma\sqrt{T}, \\ d_2' &= d_1' - \sigma\sqrt{T}, \end{aligned}$$

where $N(\cdot)$ is the cumulative normal distribution.

Notice the similarity with the Black and Scholes formula since $u \to v$ when $a \to +\infty$. An agent who applies this partial hedging strategy $\pi(\cdot)$ must hold initially $N(d_1) - N(d_1')$ shares of the primitive asset, while a perfect hedge requires $N(d_1)$. The default risk measured under Q can be derived in a similar way. One obtains

$$Q(H^c) = E^Q\left[1_{\{S(T)>a\}}\right] = N(d_2').$$

The default risk $P(H^c)$ is deduced from $Q(H^c)$ by substituting μ for r where μ is the appreciation rate of $S(\cdot)$.

The parameters of the option to be evaluated are: $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma = 0.15$, and $r = 0.05$. The partial hedging parameter is a. A numerical illustration is given in Table 1 whose last column, denoted by ∞, reports the cost v of a full replication. Each cell of this table contains the exact solution computed by numerical integration.

Table 1: Partial Hedging of a Call Option when $A = \{E \leq S(T) \leq a\}$

a	120	125	130	135	150	∞
u	3.9642	4.7197	5.1534	5.3703	5.5199	5.5271
g	1.5630	0.8075	0.3737	0.1569	0.0072	0
$P(H^c)$ for $\mu = 0.05$	0.0622	0.0274	0.0110	0.0041	0.0001	0
$P(H^c)$ for $\mu = 0.10$	0.0967	0.0460	0.0199	0.0080	0.0003	0

When the hedging parameter a increases, the cost u of the partial replication increases and converges to the Black and Scholes price $v = 5.5271$. At the same time, the default risk converges to zero. For example, if the seller decides to hedge the call option only when $S(T) \leq 130$ and not on the others states of nature, he can do so with an initial wealth $u = 5.1534$. This results in a gain of $g = 0.3737$ over the perfect hedge. Nevertheless, the hedger will fall into default with probability $P(H^c) = 0.011$ for $\mu = 0.05$, and $P(H^c) = 0.0199$ for $\mu = 0.1$. Notice that $P(H^c)$ is an increasing function of μ.

1.2.2 Partial Hedging when $A = \{E \leq S(T) \leq a, M^S(T) \leq b\}$

We now consider a hedging event of the form

$$A = \{E \leq S(T) \leq a, M^S(T) \leq b\}, \quad \text{for } E < a < b,$$

where the random variable $M^S(T) = \max\{S(t),\ t \in [0,\ T]\}$ is the maximum attained by the primitive asset price over the trading period. By the Girsanov Theorem [Karatzas and Shreve (1991), Section 3.5], there exists a probability measure $\widetilde{Q}$ under which the process

$$X(\cdot) = \ln(S(\cdot)/S(0))/\sigma,$$

is a $\widetilde{Q}$-Brownian motion. The probability measure $\widetilde{Q}$ is defined by its Radon-Nikodym likelihood ratio

$$dQ/d\widetilde{Q} = \widetilde{Z}(T) = \exp\left((r - \sigma^2/2)X(T)/\sigma - (r - \sigma^2/2)^2 T/2\sigma^2\right).$$

This result has been used judiciously by Conze and Viswanathan (1991) to derive explicit formulas for several lookback options using the risk-neutral evaluation approach. The original results, solutions of a partial differential equation, are derived by Goldman, Sosin, and Gatto (1979).

This change of measure allows the use of the known density function of $(X(T), M^X(T))$ [Karatzas and Shreve (1991), Section 2.8]

$$\varphi(x, y) = 2(2\pi T^3)^{-1/2}(2y - x)e^{-(2y-x)^2/2T}, \quad \text{for } y \geq \max(x, 0),$$

where $M^X(T) = \max\{X(t),\, t \in [0,\, T]\}$ is the maximum attained by the $\widetilde{Q}$-Brownian motion $X(\cdot)$ during $[0, T]$. This result allows for the derivation of the closed-form solutions

$$\begin{aligned} u &= E^Q\left[e^{-rT}(S(T)-E)1_{\{E\le S(T)\le a, M^S(T)\le b\}}\right], \\ &= E^{\tilde{Q}}\left[e^{-rT}\widetilde{Z}(T)(S(0)e^{\sigma X(T)}-E)1_{\left\{\tilde{E}\le X(T)\le \tilde{a}, M^X(T)\le \tilde{b}\right\}}\right], \end{aligned}$$

and

$$\begin{aligned} Q(H^c) &= E^Q[1_{H^c}], \\ &= E^Q\left[1_{\{S(T)>E\}}\right] - E^Q\left[1_{\{E\le S(T)\le a, M^S(T)\le b\}}\right], \\ &= N(d_2) - E^{\tilde{Q}}\left[\widetilde{Z}(T)1_{\left\{\tilde{E}\le X(T)\le \tilde{a}, M^X(T)\le \tilde{b}\right\}}\right], \end{aligned}$$

where $\tilde{z} = \log(z/S(0))/\sigma$ for $z \in \{E, a, b\}$. The default risk $P(H^c)$ is deduced from $Q(H^c)$ by substituting μ for r. These expectations are basically 2-dimensional integrals.

The parameters of the option to be evaluated are: $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma = 0.15$, and $r = 0.05$. The partial hedging parameters are a and $b = a + 3$. Results are shown in Table 2 whose last column, denoted by ∞, reports the cost v of a full replication. Each cell of this table contains the exact solution computed by numerical integration.

Table 2: Partial Hedging of a Call when $A = \{E \le S(T) \le a, M^S(T) \le b\}$

a	120	125	130	135	150	∞
u	3.6810	4.5688	5.0841	5.3418	5.5187	5.5271
g	1.8461	0.9583	0.4430	0.1853	0.0084	0
$P(H^c)$ for $\mu = 0.05$	0.0805	0.0347	0.0137	0.0050	0.0002	0
$P(H^c)$ for $\mu = 0.10$	0.1203	0.0564	0.0242	0.0095	0.0004	0

For example, if the seller decides to hedge the call option only when $S(T) \le 130$ and $M^S(T) \le 133$, and not on the others states of nature, he can do so with an initial wealth $u = 5.0841$. This results in a gain of $g = 0.4430$ over the perfect hedge. Nevertheless,

the hedger will fall into default with probability $P(H^c) = 0.0137$ for $\mu = 0.05$, and $P(H^c) = 0.0242$ for $\mu = 0.1$. In comparison with the results of Table 1, here the cost of any partial hedging strategy is slightly smaller and the default risk is slightly larger.

1.2.3 Partial Hedging on $\{E \leq S(T) \leq a, M^S(T) \leq b, \theta^S(T) \leq s\}$

Consider now a partial hedging strategy on the event

$$A = \{E \leq S(T) \leq a, M^S(T) \leq b, \theta^S(T) \leq s\},$$
$$\text{for } E \leq a \leq b \text{ and } 0 < s < T,$$

where $\theta^S(T) = \inf\{t \in [0,T],\ S(t) = M^S(T)\}$ is the first time when the primitive asset attains its maximum over the trading period. The random variable $\theta^S(T)$ is an example of a random time which is not a stopping time. The same change of measure as in Section 3.1.2 allows one to use the known density function of $(X(T), M^X(T), \theta^X(T))$ [Karatzas and Shreve (1991), Section 2.8]

$$\phi(x,y,\theta^X(T) < s) = \frac{2}{\sqrt{2\pi T^3}}\left[N(-\alpha_+/\beta)(2y-x)e^{-(2y-x)^2/2T} - N(-\alpha_-/\beta)xe^{-x^2/2T}\right],$$
$$\text{for } y \geq \max(x,0) \text{ and } 0 < s < T,$$

where $\alpha_\pm = (y(T-s) \pm (x-y)s)/T$ and $\beta^2 = s(T-s)/T$. From this, we can derive closed-form solutions for the cost of a partial hedge and its associated default risk. The results are

$$u = E^Q\left[e^{-rT}(S(T)-E)1_{\{E\leq S(T)\leq a, M^S(T)\leq b, \theta^S(T)\leq s\}}\right],$$
$$= E^{\tilde{Q}}\left[e^{-rT}\tilde{Z}(T)(S(0)e^{\sigma X(T)} - E)1_{\{\tilde{E}\leq X(T)\leq \tilde{a}, M^X(T)\leq \tilde{b}, \theta^X(T)\leq s\}}\right],$$

and

$$\begin{aligned} Q(H^c) &= E^Q[1_{H^c}], \\ &= E^Q\left[1_{\{S(T)>E\}}\right] - E^Q\left[1_{\left\{E\le S(T)\le a, M^S(T)\le b, \theta^S(T)\le s\right\}}\right], \\ &= N(d_2) - E^{\tilde{Q}}\left[\tilde{Z}(T)1_{\left\{\tilde{E}\le X(T)\le \tilde{a}, M^X(T)\le \tilde{b}, \theta^X(T)\le s\right\}}\right]. \end{aligned}$$

These expectations are basically 3-dimensional integrals transformed into 2-dimensional integrals. The default risk $P(H^c)$ can be deduced from $Q(H^c)$ by substituting μ for r.

The parameters of the option to be evaluated are: $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma = 0.15$, and $r = 0.05$. The partial hedging parameters are a, $b = a + 3$, and $s = 0.48$. Results are shown in Table 3 whose last column, denoted by ∞, reports the cost v of a full replication. In that way, for $a = \infty$, one has $s = 0.5$. Each cell of this table contains the exact solution computed by numerical integration.

Table 3: Partial Hedging a call when $A = \left\{E \le S(T) \le a, M^S(T) \le b, \theta^S(T) \le s\right\}$

a	120	125	130	135	150	∞
u	2.4497	2.9545	3.2201	3.3401	3.4112	5.5271
g	3.0774	2.5727	2.3070	2.1870	2.1159	0
$P(H^c)$ for $\mu = .05$	0.2001	0.1721	0.1605	0.1563	0.1542	0
$P(H^c)$ for $\mu = 0.10$	0.2600	0.2219	0.2047	0.1977	0.1939	0

If the seller decides to hedge the call option only when $S(T) \le 130$, $M^S(T) \le 133$, and $\theta^S(T) \le 0.48$, and not on the others states of nature, he can do so with an initial wealth $u = 3.2201$. This results in a gain of $g = 2.3070$ over the perfect hedge. Nevertheless, the hedger will fall into default with probability $P(H^c) = 0.1605$ for $\mu = 0.05$, and $P(H^c) = 0.2047$ for $\mu = 0.1$. In comparison with the results of Table 1 and Table 2, here the cost of any partial hedging strategy is significantly smaller and the default risk is significantly larger. The reason is that the primitive asset is likely to attain its maximum at

the end of the period: The density function of the time to maximum $\theta^S(T)$ obeys the arcsin law [Karatzas and Shreve (1991), Section 2.8]:

$$h(s) = 2\pi^{-1} \arcsin(\sqrt{s/T}), \quad \text{for } 0 \leq s \leq T.$$

Closed-form solutions can also be derived if the hedging event depends on the final primitive asset price, its first passage time at a certain level, its maximum during the trading period, and the time to maximum.

1.3 Partial Hedging in a Stochastic Volatility Model

1.3.1 A Monte Carlo Experiment

All the random variables introduced in the following sections are assumed to have finite variance. The stochastic volatility model introduced in (1.5) is arbitrage-free but incomplete (see Section 2 for a justification). The dynamic of the primitive asset under Q is

$$dS(t) = rS(t)dt + \sigma_1(t)S(t)dB_1(t), \quad \text{for } S(0) > 0 \text{ and } 0 \leq t \leq T,$$

where the volatility process $\sigma_1(\cdot)$ is a function of a Brownian motion $B(\cdot) = (B_1(\cdot), B_2(\cdot))'$. Several dynamics for volatility have been proposed in the literature (see Detemple and Osakwe (2000) for a general specification). One of these is the following mean-reverting process

$$\begin{aligned} d\sigma_1(t) &= \alpha(\overline{\sigma} - \sigma_1(t))dt + \theta\sigma_1(t)(\rho dB_1(t) + \sqrt{1-\rho^2}dB_2(t)), \\ &\quad \text{for } 0 \leq t \leq T, \end{aligned}$$

where the coefficients α (the reverting rate), $\overline{\sigma}$ (the long-term volatility), θ (the volatility of the volatility), and ρ (the correlation between the innovations) are assumed to be constants. Statistical methods are needed to estimate these coefficients. For simplicity, we assume here that $\rho = 0$ so that

$$d\sigma_1(t) = \alpha(\overline{\sigma} - \sigma_1(t))dt + \theta\sigma_1(t)dB_2(t), \quad \text{for } 0 \leq t \leq T.$$

To make the hedging of contingent claims possible, a second primitive asset $S_2(\cdot)$ is introduced in the market. It is assumed to move under Q according to the stochastic differential equation

$$dS_2(t) = rS_2(t)dt + \sigma_2 S_2(t)dB_2(t), \quad \text{for } S_2(0) > 0 \text{ and } 0 \leq t \leq T,$$

where σ_2 is a positive constant (see equation (1.10) for a justification). The asset $S_2(\cdot)$ could be interpreted as an index of the rest of the economy.

By equation (1.6), the components of the market-price of risk are $\theta_1(t) = (\mu_1 - r)/\sigma_1(t)$ and $\theta_2(t) = (\mu_2 - r)/\sigma_2$ for $0 \leq t \leq T$. By equation (1.9), the dynamics of $S(\cdot)$, $\sigma_1(\cdot)$, and $S_2(\cdot)$ under P are

$$\begin{aligned} dS(t) &= \mu_1 S(t)dt + \sigma_1(t)S(t)dW_1(t), \\ d\sigma_1(t) &= \alpha'(\overline{\sigma}' - \sigma_1(t))dt + \theta\sigma_1(t)dW_2(t), \\ dS_2(t) &= \mu_2 S_2(t)dt + \sigma_2 S_2(t)dW_2(t), \\ &\qquad \text{for } S(0) > 0, \sigma_1(0) > 0, \text{ and } 0 \leq t \leq T, \end{aligned} \tag{1.22}$$

where $\alpha' = \alpha - \theta(\mu_2 - r)/\sigma_2$ and $\overline{\sigma}' = \alpha\overline{\sigma}/\alpha'$.

In the market defined in (1.22), starting at the initial wealth defined in (1.14), any contingent claim Y is attainable by a replication portfolio $\pi(\cdot) = (\pi_1(\cdot), \pi_2(\cdot))'$ as described in (1.13). The price of the call option written on the first primitive asset is $v = E^Q\left[e^{-rT}(S(T) - E)^+\right]$. The partial hedging strategy costs $u = E^Q\left[e^{-rT}(S(T) - E)^+ 1_A\right]$ and default risk is $P(H^c)$. Notice that the appreciation rate μ_1 of the first primitive asset and the parameters of the second primitive asset, μ_2 and σ_2, are needed for estimating $P(H^c)$, but not for v and u.

It is well known that option prices usually do not admit closed-form solutions in this model and that simulation is required. Since the final primitive asset price $S(T)$ cannot be simulated directly, a discrete-time approximation such that the Euler scheme with m periods of length $h = T/m$ can be performed:

$$\begin{aligned} \widehat{S}(kh) - \widehat{S}((k-1)h) &= \widehat{S}((k-1)h)(rh + \widehat{\sigma}_1((k-1)h)\sqrt{h}Z_1(k)), \qquad (1.23) \\ \widehat{\sigma}_1(kh) - \widehat{\sigma}_1((k-1)h) &= \alpha(\overline{\sigma} - \widehat{\sigma}_1((k-1)h)h + \theta\widehat{\sigma}_1((k-1)h)\sqrt{h}Z_2(k), \\ &\qquad \text{for } k = 1, \ldots, m, \end{aligned}$$

where the $\sqrt{h}Z_1(k) = B_1(kh) - B_1((k-1)h)$ and the $\sqrt{h}Z_2(k) = B_2(kh) - B_2((k-1)h)$ are the increments of the Brownian motions $B_1(\cdot)$ and $B_2(\cdot)$. Here $Z_1(1), \ldots, Z_1(m)$, $Z_2(1), \ldots, Z_2(m)$ are independent and identically distributed normal variables.

The error of the Euler approximation when computing an expectation in the form $E\left[f(S(t), t \in I \subset [0,T])\right]$, defined as

$$e(m) = |E[f(\widehat{S}(t), t \in I \subset [0,T])] - E[f(S(t), t \in I \subset [0,T])]|,$$

where $f : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$, is known to be in $O(m^{-1})$. Given a computational budget, a trade-off between the number of time increments m of the Euler approximation and the sample size n of the simulation experiment must be found. Duffie and Glynn (1995) argue that n must increase as $O(m^2)$ so that doubling m necessitates quadrupling n. In the following, $f(\widehat{S}(t), t \in I \subset [0,T])$ is denoted $f(S(t), t \in I \subset [0,T])$.

The Euler approximation is used to simulate n copies of $(S(T), M^S(T), \theta^S(T))$ (we take $n = 4000$), which serve to simulate as many copies of $f(S(T), M^S(T), \theta^S(T))$ where $f : \mathbb{R}^3 \rightarrow \mathbb{R}$. Depending on the function f and the probability measure used, the parameter

$$w = E\left[f(S(T), M^S(T), \theta^S(T))\right]$$

matches v, u, or $P(H^c)$. The crude Monte Carlo estimator of w based on n replications is

$$\widehat{w} = n^{-1} \sum_{i=1}^{n} f((S(T), M^S(T), \theta^S(T))_i),$$

where the $(S(T), M^S(T), \theta^S(T))_i$, for $i = 1, \ldots, n$, are the n copies of $(S(T), M^S(T), \theta^S(T))$. The estimated error of $\widehat{w}$ can be defined as the half-length of the asymptotic 95% confidence interval of w based on the normality assumption,

$$e = 1.96 s_n / \sqrt{n},$$

where s_n is the sample standard error of $f(S(T), M^S(T), \theta^S(T))$.

Through each path, the global maximum $M^S(T)$ is simulated following Beaglehole, Dybvig, and Zhou (1996). The time to maximum is simulated as the midpoint of the subinterval $[(k^* - 1)h, k^*h]$ containing the global maximum.

The parameters of the option to be evaluated are $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma = 0.15$, and $r = 0.05$. The parameters of the volatility are $\alpha = 1.5$, $\overline{\sigma} = 0.15$, and

$\theta = 0.08$. The appreciation rate of the first primitive asset is $\mu_1 = 0.1$, and the parameters of the second primitive asset are $\mu_2 = 0.08$ and $\sigma_2 = 0.12$. Results in Table 4 are obtained at $s = 0.5$, that is, the constraint $\theta^S(T) \leq s$ can be ignored. Each cell of this table contains the Monte Carlo estimate and its estimated error. The partial hedging parameters are a and $b = a + 3$, and the parameters of the simulation are: $m = 60$ and $n = 4000$.

Table 4: Partial Hedging of a Call when $A = \{E \leq S(T) \leq a, M^S(T) \leq b\}$

a	120	125	130	135	150	∞
u	3.46 $\pm$0.16	4.40 $\pm$0.19	4.84 $\pm$0.21	5.09 $\pm$0.22	5.26 $\pm$0.23	5.26 $\pm$0.23
$P(H^c)$	0.120 $\pm$0.010	0.062 $\pm$0.007	0.026 $\pm$0.005	0.011 $\pm$0.003	—	0

Results in Table 4 are similar to those in Table 2, except for the statistical error, which we shall now try to reduce. The relative error of $P(H^c)$, defined as the ratio of the statistical error over the statistical estimation, increases as the parameter a increases, i.e., as the event H^c becomes rarer. This is a typical situation when estimating probability of rare events. At the extreme case $a = 150$, the default event H^c is so rare that we have observed no realization of $f(S(T), M^S(T), \theta^S(T))$ in this region for our 4000 simulation runs. The variance reduction technique, called *Importance Sampling* (see Boyle, Broadie, and Glasserman (1997) and L'Ecuyer (1994) for a discussion), provides a way to handle this type of situation and could be used for large values of a. The idea is to select a change of measure so that the integrand, here $f(S(T), M^S(T), \theta^S(T))$, goes more frequently into the most important regions of the sample space, here H^c. For $a = 150$, one can also see that the simulation could not distinguish between the cost of the partial replication and the cost of the full replication.

In the next subsections, correlation induction techniques are used to reduce the estimated error of the crude Monte Carlo estimators. These variance reduction techniques, namely *Antithetic Variates* and *Control Variates*, induce correlation between estimators in attempt to reduce the variance. The techniques used are discussed, e.g., in Bratley, Fox, and Schrage (1987) and L'Ecuyer (1994).

1.3.2 Antithetic Variates

Let $\widehat{w}_1$ to be an unbiased estimator of w. For simplicity, take $\widehat{w}_1$ as the crude Monte Carlo estimator of w based on one replication. Assume that one can build a second unbiased estimator $\widehat{w}_2$ of w which is negatively correlated with $\widehat{w}_1$. Thus, the unbiased estimator $\widehat{w} = (\widehat{w}_1 + \widehat{w}_2)/2$ of w is expected to have lower variance than each of its components:

$$\mathrm{Var}\left[\widehat{w}\right] = \mathrm{Var}\left[\widehat{w}_1\right]/4 + \mathrm{Var}\left[\widehat{w}_2\right]/4 + \mathrm{Cov}\left[\widehat{w}_1, \widehat{w}_2\right]/2,$$

if $\mathrm{Cov}[\widehat{w}_1,\widehat{w}_2] < 0$ and $\widehat{w}_2$ is well selected. Roughly speaking, if $\widehat{w}_1$ takes high values above its mean w, $\widehat{w}_2$ takes low values below its mean w. Thus, their deviations are mutually compensated in $\widehat{w}$ whence the terminology "Antithetic Variates".

The estimator $\widehat{w}_1$ is often written as a function of some independent and identically basic uniforms $U_1,\ldots,U_q$

$$\widehat{w}_1 = f(U_1, \ldots, U_q),$$

where $f : \mathbb{R}^q \to \mathbb{R}$. If f is monotone, taking

$$\widehat{w}_2 = f(1 - U_1, \ldots, 1 - U_q),$$

ensures the condition $\text{Cov}[\widehat{w}_1, \widehat{w}_2] < 0$ [Bratley, Fox, and Schrage (1987), page 46] and variance reduction. In the case analyzed here, the output $\widehat{w}_1$ is a function of some inputs as shown in (1.23):

$$\widehat{w}_1 = f(Z_1(1), \ldots, Z_1(m), Z_2(1), \ldots, Z_2(m)),$$

where $Z_1(1),\ldots,Z_1(m),Z_2(1),\ldots,Z_2(m)$ are independently and identically distributed normal random variables. By the same argument, taking

$$\widehat{w}_2 = f(-Z_1(1), \ldots, -Z_1(m), Z_2(1), \ldots, Z_2(m)),$$

ensures the condition $\text{Cov}[\widehat{w}_1, \widehat{w}_2] < 0$ and variance reduction. One can focus only on the components where the function is monotone and synchronize between the estimators to induce the attempted negative correlation. For the estimation of u and $P(H^c)$, we observed a variance reduction with the antithetic variates for the first parameter as illustrated by the following example.

The parameters of the option to be evaluated are $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma = 0.15$, and $r = 0.05$. The parameters of the volatility are $\alpha = 1.5$, $\overline{\sigma} = 0.15$, and $\theta = 0.08$. Results are shown in Table 5. Each cell of this table contains the Antithetic Variates estimate and its estimated error. The partial hedging parameters are a and $b = a+3$, and the simulation parameters are $m = 60$ and $n = 4000$. The estimated errors, given in Table 5, show a modest variance reduction in comparison with those of Table 4.

Table 5: Partial Hedging a Call when $A = \{E \leq S(T) \leq a, M^S(T) \leq b\}$

a	120	125	130	135	∞
u	3.68 $\pm$0.08	4.58 $\pm$0.09	5.03 $\pm$0.09	5.29 $\pm$0.10	5.49 $\pm$0.11

1.3.3 Control Variates

Let X be an unbiased estimator of w and let $C = (C_1, \ldots, C_q)'$ be a random vector with a known expected value $\nu = (\nu_1, \ldots, \nu_q)'$, presumably correlated with X. Assume that C is known to the simulator. Think of X as the crude Monte Carlo estimator of w based on one replication when the volatility moves randomly and $C = C_1$ as the synchronous crude Monte Carlo estimator of w when the volatility is constant. The idea behind this technique is to find a vector $\beta = (\beta_1, \ldots, \beta_q)'$ such that the unbiased estimator of w, namely the *controlled estimator*,

$$X_c = X - \beta'(C - \nu),$$

has a lower variance than X. The optimal choice for β, to yield the maximum variance reduction, is

$$\beta^* = \Sigma_C^{-1}\Sigma_{X,C},$$

where Σ_C is the variance matrix of C and $\Sigma_{X,C}$ is the covariance vector between X and the components of C. At β^*, a variance reduction takes place:

$$\mathrm{Var}[X_C] = (1 - R^2_{X,C})\mathrm{Var}[X],$$

where

$$R^2_{X,C} = \Sigma'_{X,C}\Sigma_C^{-1}\Sigma_{X,C}/\mathrm{Var}[X],$$

is the multiple coefficient of correlation between X and the components of C. In the particular case $q = 1$, these results can be written as

$$\beta^* = \mathrm{Cov}[X, C]/\mathrm{Var}[C],$$

and

$$\mathrm{Var}[X_C] = (1 - \mathrm{Corr}[X, C]^2)\mathrm{Var}[X].$$

Roughly speaking, if X increases and takes high values, $\mathrm{Cov}[X, C]\ (C - \nu)/\mathrm{Var}[C]$ necessarily increases. Thus, it controls the excess of X above its mean w via X_c, whence the terminology "Control Variates". In option pricing, the random variable $\gamma(T)S(T)$ is usually taken as a control variable since $S(0) = E^Q[\gamma(T)S(T)]$ is known: The process $\gamma(\cdot)S(\cdot)$ is a Q-martingale as mentioned in (1.15). Several authors, e.g., Clewlow and Carverhill (1994), select a priori $\beta^* = 1$ and report a significant variance reduction. In fact, this choice is not necessarily acceptable but it should work when C is approximately equal to X. In that case, the optimal value for β is expected to be near unity since $\mathrm{Cov}[X, C] \simeq \mathrm{Var}[C]$.

Unfortunately, neither Σ_C nor $\Sigma_{X,C}$ are known in practice and β^* cannot be computed as shown above. An alternative idea is to simulate n copies of (X, C), estimate Σ_C and $\Sigma_{X,C}$ as usual, and define the observations of the controlled estimator as

$$X_{c,i} = X_i - \widehat{\beta}'(C_i - \nu), \quad \text{for } i = 1,...,n,$$

where

$$\widehat{\beta} = \widehat{\Sigma}_C^{-1}\widehat{\Sigma}_{X,C}.$$

The controlled estimator of w is defined as the sample mean of the $X_{c,i}$, for $i = 1,...,n$,

$$\overline{X}_c = \overline{X} - \widehat{\beta}'(\overline{C} - \nu).$$

The sample variance S_c^2 of X_c is defined as usual. Notice that the controlled estimator $\overline{X}_c$ is generally a biased estimator of w since $\widehat{\beta}$ and $\overline{C}$ are a priori correlated. However, Lavenberg and Welch (1981) showed that this bias vanishes when (X, C) is multinormal.

As pointed out by Nelson (1990), under weak conditions, the controlled estimator is convergent as $n \longrightarrow +\infty$:

$$\begin{aligned} \overline{X}_c &\rightarrow w \quad \text{in probability,} \\ S_c^2 &\rightarrow (1 - R_{X,C}^2)\text{Var}[X] \quad \text{in probability,} \\ \sqrt{n}(\overline{X}_c - w)/S_c &\rightarrow N(0,1) \quad \text{in distribution.} \end{aligned}$$

Thus, $\overline{X}_c$ is an asymptotically unbiased estimator of w and has asymptotic smaller variance than the crude Monte Carlo estimator $\overline{X}$. Techniques for reducing the bias of $\overline{X}_c$ for small samples, such as Batching, Jackknifing, and Splitting, are described in Nelson (1990). By splitting optimally into three groups, Avramidis and Wilson (1993) build a controlled estimator which is somewhat more consistent than $\overline{X}_c$, as it converges to w always surely.

The parameters of the option to be evaluated are $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma(0) = 0.15$, and $r = 0.05$. The parameters of the volatility are $\alpha = 1.5$, $\overline{\sigma} = 0.15$, and $\theta = 0.08$. The appreciation rate of the first primitive asset is $\mu_1 = 0.1$, and the parameters of the second primitive asset are $\mu_2 = 0.08$ and $\sigma_2 = 0.12$. Results are shown in Table 6. Each cell of this table contains the Control Variates estimate and its estimated error. The partial hedging parameters are a and $b = a + 3$, and the simulation parameters are $m = 60$ and $n = 4000$. We observe a significant variance reduction resulting from the high correlation between the crude estimators and their associated control variables.

Table 6: Partial Hedging a Call when $A = \{E \leq S(T) \leq a, M^S(T) \leq b\}$

a	120	125	130	135	∞
u	3.63 $\pm$0.04	4.46 $\pm$.04	4.95 $\pm$.04	5.19 $\pm$.02	5.39 $\pm$.01
$P(H^c)$	0.119 $\pm$0.002	0.056 $\pm$0.002	0.023 $\pm$0.002	0.010 $\pm$0.001	0

1.3.4 Integrating the Correlation Induction Techniques

Now, denote X as the antithetic variates estimator of w (based on one replication) when the volatility moves randomly and $C = C_1$ the synchronous antithetic variates estimator of w when the volatility is constant. As pointed out by Avramidis and Wilson (1996), the estimator X can be viewed as an aggregate response and the random variable C as an aggregate control variable.

The parameters of the option to be evaluated are $S(0) = 100$, $E = 100$, $T = 0.5$, $\sigma(0) = 0.15$, and $r = 0.05$. The parameters of the volatility are $\alpha = 1.5$, $\overline{\sigma} = 0.15$, and $\theta = 0.08$. Results are shown in Table 7. Each cell of this table contains the estimate based on the aggregate response and its estimated error. The partial hedging parameters are a and $b = a + 3$, and the simulation parameters are $m = 60$ and $n = 4000$. Additional improvements are realized when integrating the Antithetic Variates and the Control Variates techniques.

Table 7: Hedging Partially a Call when $A = \{E \leq S(T) \leq a, M^S(T) \leq b\}$

a	120	125	130	135	∞
u	3.61 $\pm$.02	4.46 $\pm$.02	4.97 $\pm$.03	5.21 $\pm$.02	5.39 $\pm$.01

1.4 Conclusion

A hedger may find an advantage in partially replicating a contingent claim if the lower cost of a partial hedge more than offsets the added default risk. Several partial replication strategies are possible. In this paper, we measure default probabilities and expected gains for possible strategies using the price of the primitive asset at the horizon, its maximum

over the trading period, and the time to maximum. The results show how the cost of a partial hedge and default risk vary depending on the replication event.

In the Black and Scholes model, the partial hedging strategies are easy to identify and to implement. In the stochastic volatility case, the partial hedging strategy can be approximated based on the sensitivity coefficients of the option (Broadie and Glasserman 1996). In constraint models, hedging contingent claims is not possible and the super-replication cost is excessively high (Cvitanić, Pham, and Touzi 1997). In such models, a partial super-replication may be an interesting solution.

In this paper, we gave examples of how Monte Carlo Monte simulation, a flexible and robust tool, can be used to analyze such strategies.

1.5 References

Avramidis, A.N. and J.R. Wilson, "A Splitting Scheme for Control Variates," *Operations Research Letters*, 14 (1993), 187-198.

Avramidis, A.N. and J.R. Wilson, "Integrated Variance Reduction Strategies for Simulation," *Operations Research*, 44 (1996), 327-346.

Beaglehole, D.R., P.H. Dybvig, and G. Zhou, "Going to Extremes: Correcting Simulation Bias in Exotic Option Valuation," *Financial Analysts Journal*, 53 (1997), 62-68.

Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (1973), 637-654.

Boyle, P.P., M. Broadie, and P. Glasserman, "Monte Carlo Methods for Security Pricing," *Journal of Economic Dynamics and Control*, 21 (1997), 1267-1321.

Bratley, P., B.L. Fox, and L.E. Schrage, *A Guide to Simulation*, Springer-Verlag, New York, 1987.

Broadie, M. and P. Glasserman, "Estimating Security Price Derivatives Using Simulation," *Management Science*, 42 (1996), 269-285.

Clewlow, L. and A. Carverhill, "On the Simulation of Contingent Claims," *Journal of Derivatives*, 2 (1994), 66-74.

Conze, A. and Viswanathan, "Path Dependent Options: The Case of Lookback Options," *Journal of Finance*, 46 (1991), 1893-1906.

Cvitanić, J., H. Pham, and N. Touzi, Super-Replication in Stochastic Volatility Models under Portfolio Constraints, Working Paper, Columbia University, 1997.

Detemple, J. and C.J. Osakwe, The Valuation of Volatility Options, Working Paper, McGill University, 2000.

Duffie, D. and P. Glynn, "Efficient Monte Carlo Simulation of Security Prices," *The Annals of Applied Probability*, 5 (1995), 897-905.

Föllmer, H., Talk at the Isaac Newton Institute for the Mathematical Sciences, Cambridge University, 1995.

Goldman, M.B., H.B. Sosin, and M.A. Gatto, "Path Dependent Options: Buy at the Low, Sell at the High," *Journal of Finance*, 34 (1979), 1111-1127.

Harrison, J.M. and D.M. Kreps, "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory*, 20 (1979), 381-408.

Harrison, J.M. and S. Pliska, "Martingales and Stochastic Integrals in the Theory of Continuous Trading," *Stochastic Processes and their Applications*, 11 (1981), 215-260.

Hull, J. and A. White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42 (1987), 281-299.

Johnson, H. and D. Shanno, "Option Pricing when the Variance is Changing," *Journal of Financial and Quantitative Analysis*, 22 (1987), 143-151.

Karatzas, I., Lectures on the Mathematics of Finance, Centre de Recherches Mathématiques, Université de Montréal, 1996.

Karatzas, I. and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Springer-Verlag, New York, 1991.

Lavenberg, S.S. and P.D. Welch, "A Perspective on the Use of Control Variables to Increase the Efficiency of Monte Carlo Simulations," *Management Science*, 27 (1981), 322-335.

L'Ecuyer, P., "Efficient Improvement via Variance Reduction," *Proceedings of the 1994 Winter Simulation Conference*, IEEE Press, 1994, 122-132.

Nelson, B.L., "Control Variate Remedies," *Operations Research*, 38 (1990), 974-992.

Scott, L.O., "Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application," *Journal of Financial and Quantitative Analysis*, 22 (1987), 419-438.

Wiggins, J.B., "Option Values under Stochastic Volatility, Theory and Empirical Estimates," *Journal of Financial Economics*, 19 (1987), 351-372.

Chapter 2
Numerical Procedures for Pricing American-style Asian Options

2.1 Introduction

In a frictionless market, the no-arbitrage principle allows one to express the value of a derivative as the mathematical expectation of its discounted future payment, with respect to a so-called risk-neutral probability measure. *Options* are particular derivatives characterized by non-negative payoffs. *European-style* options can be exercised at the expiration date only, whereas *American-style* ones offer early exercise opportunities to the holder.

For simple cases, such as for European call and put options written on a stock whose price moves as a geometric Brownian motion (GBM), as studied by Black and Scholes (1973), analytic formulas are available for the fair price of the option. For more complicated derivatives, however, which may involve multiple assets, stochastic time-varying model parameters, possibilities of early exercise, etc., analytic formulas are unavailable. These derivatives are usually priced either via the lattice-based approach, Monte Carlo simulation, or numerical methods (e.g., Boyle, Broadie, and Glasserman 1997, Hull 2000, Wilmott, Dewynne, and Howison 1993).

The payments promised by *Asian options* depend on the average price of the primitive asset. Those based on the arithmetic mean have no known analytic formula, even under the GBM assumption. An Asian option can hedge the risk exposure of a firm that sells or

buys certain types of resources (raw materials, energy, foreign currency, etc.), on a regular basis. The average being less volatile than the price of the underlying asset, these contracts are less expensive than their standard versions. Asian options are heavily traded over-the-counter and, because of the possible lack of depth of these markets, their theoretical values often need to be computed *on-the-fly* for fair negotiations (Vorst 1998).

Asian options come in various flavors. For example, the average can be arithmetic or it can be geometric. One talks of a *plain vanilla* Asian option if the average is computed over the full trading period, and a *backward-starting* if it is computed over a right subinterval of the trading period. This interval usually has a fixed starting point in time.

The Asian option can be *fixed-strike* (if the strike price is a fixed constant) or *floating-strike* (if the strike is itself an average). It is called *flexible* when the payoff is a weighted average, and *equally weighted* when all the weights are equal. The prices are *discretely sampled* if the payoff is the average of a discrete set of the asset price, and *continuously sampled* if the payoff is the integral of the asset price over some time interval, divided by the length of that interval. The options considered in this paper are the most common: *Fixed-strike, equally-weighted, discretely-sampled Asian options based on arithmetic averaging.* Our method could also be adapted to price other kinds of discretely-sampled Asian options.

European-style Asian (named *Eurasian*) options can be exercised at the expiration date only, whereas *American-style* ones (named *Amerasian*) offer earlier exercise opportunities, which may become attractive intuitively when the current asset price is below the current running average (i.e., is pulling down the average) for a call option, and when it is above the running average for a put. Here, we focus on *Amerasian call options*, whose val-

ues are harder to compute than the Eurasian ones, because an optimization problem must be solved at the same time as computing the mathematical expectation giving the option's value.

There is an extensive literature on the pricing of Eurasian options. In the context of the Black and Scholes' model, there is a closed-form analytic solution for the value of discretely-sampled Eurasian options only when they are based on the *geometric average* (Turnbull and Wakeman 1991, Zhang 1995). The idea is that the geometric average of lognormals is a lognormal. Geman and Yor (1993) used Bessel processes and derived exact formulas for the Laplace transform of the value of a continuous-time Eurasian option. For options based on the arithmetic average, solution approaches include *quasi-analytic* approximation methods based on Fourier transforms, Edgeworth and Taylor expansions, and the like (e.g., Bouaziz, Briys, and Crouhy 1994, Carverhill and Clewlow 1990, Curran 1994, Levy 1992, Ritchken, Sankarasubramanian, and Vijh 1993, Turnbull and Wakeman 1991), methods based on partial differential equations (PDEs) and their numerical solution via *finite-difference* (FD) techniques (e.g., Alziary, Décamps, and Koehl 1997, Rogers and Shi 1995, Zvan, Forsyth, and Vetzal 1998), and *Monte Carlo simulation* coupled with variance-reduction techniques (e.g., Glasserman, Heidelberger, and Shahabuddin 1999, Kemna and Vorst 1990, L'Écuyer and Lemieux 1999, Lemieux and L'Écuyer 1998, Lemieux and L'Écuyer 2000).

Techniques for pricing Amerasian options are surveyed by Barraquand and Pudet (1996), Grant, Vora, and Weeks (1997), Zvan, Forsyth, and Vetzal (1998, 1999), and Dempster, Hutton, and Richards (1998). Hull and White (1993) have adapted binomial lattices

(from the binomial model of Cox, Ross, and Rubinstein 1979) to the pricing of Amerasian options. Nevertheless, these models are difficult to implement for path-dependent options, and do not give clear insight on the optimal exercising region (for the American versions). Broadie and Glasserman (1997a) proposed a simulation method based on non recombining trees in the lattice model, which produces two estimators of the option value, one with positive bias and one with negative bias. By taking the union of the confidence intervals corresponding to these two estimators, one obtains a conservative confidence interval for the true value. However, the work and space requirements of their approach increases exponentially with the number of exercise opportunities. Broadie and Glasserman (1997b) then developed a simulation-based stochastic mesh method that accommodates a large number of exercise dates and high-dimensional American options. Their method appears adaptable to Amerasian options, although this is not the route we take here.

Zvan, Forsyth, and Vetzal (1998) have developed stable numerical PDE methods techniques, adapted from the field of computational fluid dynamics, for pricing Amerasian options with continuously sampled prices. Zvan, Forsyth, and Vetzal (1999) have also adapted these PDE methods to Amerasian options with discretely sampled asset prices, and with barriers. The numerical approach introduced in this paper is formulated in discrete time directly.

Pricing American-style options can be formulated as a Markov Decision process, i.e., a stochastic dynamic programming (DP) problem, as pointed out by Barraquand and Martineau (1995) and Broadie and Glasserman (1997b), for example. The DP *value function* expresses the value of an Amerasian option as a function of the current time, current price,

and current average. This value function satisfies a DP recurrence (or Bellman equation), written as an integral equation. Solving this equation yields both the option value and the optimal exercise strategy. For a general overview of stochastic DP, we refer the reader to Bertsekas (1987).

In this paper, we write the DP equation for Amerasian options under the GBM assumption. Using this equation, we prove by induction certain properties of the value function and of the optimal *exercise frontier* (which delimits the region where it is optimal to exercise the option). We then propose a numerical solution approach for the DP equation, based on piecewise polynomial interpolation over rectangular finite elements. This kind of approach has been used in other application contexts, e.g., by Haurie and L'Écuyer (1986), L'Écuyer and Malenfant (1988). In fact, we reformulate the DP equation in a way that simplifies significantly the numerical integration at each step. This is a key ingredient for improving the efficiency of the procedure. Numerical experiments indicate that the method is competitive;i t provides precise results in a reasonable computing time. It could also be easily adapted to price most low-dimensional American-style derivatives such as calls on stocks with dividend paying, puts, lookback options, and options with barriers. The general methodology would also work for other types of models, e.g., for a constant elasticity of variance (CEV) process (Cox 1996, Boyle and Tian 1999). The properties of the value function that we derive in Section 4 are easy to generalize.

The idea of this paper came after reading Grant, Vora, and Weeks (1997). These authors also formulate the problem of pricing an Amerasian option in the dynamic programming framework, but use Monte Carlo simulation to estimate the value function at

each point of some discretization of the state space, and identify a "good" exercise frontier by interpolation. They also propose to restrict the strategy of exercise to a class of sub-optimal rules where the exercise frontier is approximated by two linear segments, at each date of exercise opportunity. They observed on a few numerical examples that restricting the class of strategies in this way did not seem to diminish the value of the option significantly, but they provided no proof that this is true in general. More recently, Longstaff and Schwartz (2001) proposed a general methodology based on Monte Carlo simulation for pricing high-dimensional American-style options, and applied it to Amerasian options.

Here, we suggest replacing simulation at both stages by an exact computation, which is obviously less noisy, and we do not assume a priori a shape of the exercise frontier. For both the simulation approach and our approach, an approximation of the value function must be memorized, so the storage requirement is essentially the same for the two methods.

While finalizing the revision of this paper, we learned about related work by Wu and Fu (2000), who also proved some properties of the exercise frontier for Amerasian options and proposed a different approach, that parameterizes the exercise frontier and optimizes the parameters by a stochastic approximation algorithm combined with a simulation-based perturbation analysis gradient estimation method.

Section 2 presents the model and notation. In Section 3, we develop the DP formulation. In Section 4, we establish certain properties of the value function and of the optimal region of exercise. Our approximation approach is detailed in Section 5. In Section 6, we report on numerical experiments. Section 7 is a conclusion.

2.2 Model and Notation

2.2.1 Evolution of the Primitive Asset

The primitive asset is a stock without dividend whose price moves as a GBM, in a world that satisfies the assumptions of Black and Scholes (1973). Under these assumptions, there is a probability measure Q, called *risk-neutral*, under which the price of the primitive asset $\{S(t), t \in [0,T]\}$ satisfies the stochastic differential equation (SDE)

$$dS(t) = rS(t)dt + \sigma S(t)dW(t), \quad \text{for } 0 \leq t \leq T, \tag{2.24}$$

where $S = S(0) > 0$, r is the risk-free rate, σ is the volatility parameter, T is the maturity date, and $\{W(t), t \in [0,T]\}$ is a standard Brownian motion. The solution of (2.24) is given by

$$S(t'') = S(t')e^{\mu(t''-t')+\sigma[W(t'')-W(t')]}, \quad \text{for } 0 \leq t' \leq t'' \leq T, \tag{2.25}$$

where $\mu = r - \sigma^2/2$. An important feature is that the random variable $S(t'')/S(t')$ is lognormal with parameters $\mu(t''-t')$ and $\sigma\sqrt{t''-t'}$, and independent of the σ-field $\mathcal{F}(t') = \sigma\{S(t), t \in [0,t']\}$, i.e., the trajectory of $S(t)$ up to time t'. This follows from the independent-increments property of the Brownian motion. In addition, the discounted price of the primitive asset is a Q-martingale:

$$\rho(t')S(t') = E[\rho(t'')S(t'') \mid \mathcal{F}(t')], \quad \text{for } 0 \leq t' \leq t'' \leq T, \tag{2.26}$$

where $\{\rho(t) = e^{-rt}, t \in [0,T]\}$ is the discount factor process and E is (all along this paper) the expectation with respect to Q. Details about risk-neutral evaluation can be found in Karatzas and Shreve (1998).

2.2.2 The Amerasian Contract

We consider a model similar to that of Grant, Vora, and Weeks (1997). Let $0 = t_0 \leq t_1 < t_2 < \cdots < t_n = T$ be a fixed sequence of *observation dates*, where T is the *time horizon*, and let m^* be an integer satisfying $1 \leq m^* \leq n$. The increments of time $h = t_i - t_{i-1}$, for $i = 2, \ldots, n$, are assumed to be equally spaced.

The *exercise opportunities* are at dates t_m, for $m^* \leq m \leq n$. If the Amerasian call option is exercised at t_m, its *payoff* (known as the *exercise value*) is $(\overline{S}_m - K)^+ = \max(0, \overline{S}_m - K)$, where K is the *exercise price* and $\overline{S}_m = (S(t_1) + \cdots + S(t_m))/m$ is the arithmetic average of the stock price at the observation dates up to time t_m.

This model is quite flexible. For $n = 1$, we get an European call option and, for $m^* = n > 1$, an Eurasian call option. In fact, we are not really interested in these degenerate cases, but in the case where $m^* < n$.

2.3 The Dynamic Programming Formulation

For $m = 0, \ldots, n$, denote by $v_m(s, \overline{s})$ the value of the option at t_m when $S(t_m) = s$ and $\overline{S}_m = \overline{s}$. This value is a function of the *state variables* $(s, \overline{s})$ and of the time t_m. We take the *state space* as $[0, +\infty)^2$ for convenience, although at each time step, only a subset of this space is reachable. Indeed, one must have $\overline{s} = s > 0$ at t_1 and $\overline{s} > s/m > 0$ at t_m, for $m > 1$.

At time $t_n = T$, if $\overline{S}_n = \overline{s}$, the value of the Amerasian call option coincides with its exercise value

$$v_n(s, \overline{s}) = (\overline{s} - K)^+ .$$

At time t_m, if $\overline{S}_m = \overline{s}$, the exercise value of the option is

$$v_m^e(\overline{s}) = (\overline{s} - K)^+, \quad \text{for } m^* \leq m \leq n. \tag{2.27}$$

From the risk-neutral pricing theory, the *holding value* of the option (its value if not exercised) at time t_m, for $0 \leq m < n$, can be expressed as the conditional expectation of the discounted value of the option at time t_{m+1} :

$$\begin{aligned} & v_m^h(s, \overline{s}) \qquad (2.28) \\ = \; & E\left[\rho v_{m+1}\left(S(t_{m+1}), \overline{S}_{m+1}\right) \mid S(t_m) = s, \overline{S}_m = \overline{s}\right] \\ = \; & \rho E_{m,s,\overline{s}}\left[v_{m+1}\left(S(t_{m+1}), \frac{m\overline{S}_m + S(t_{m+1})}{m+1}\right)\right], \end{aligned}$$

where $\rho = e^{-rh}$ is the discount factor over the period $[t_m, t_{m+1}]$. The value of the option satisfies

$$v_m(s, \overline{s}) = \begin{cases} v_m^h(s, \overline{s}) & \text{if } 0 \leq m < m^* \\ \max\left(v_m^e(\overline{s}), v_m^h(s, \overline{s})\right) & \text{if } m^* \leq m < n \\ v_m^e(\overline{s}) & \text{if } m = n \end{cases} . \tag{2.29}$$

Notice that $v_n(s, \overline{s}) \equiv v_n(\overline{s})$ does not depend on s, and that $v_0(s, \overline{s}) \equiv v_0(s)$ does not depend on $\overline{s}$.

The optimal exercise strategy is defined as follows: In state $(s, \overline{s})$ at time t_m, for $m^* \leq m < n$, exercise the option if $v_m^e(\overline{s}) > v_m^h(s, \overline{s})$, and hold it otherwise. The functions v_m^e, v_m^h, and v_m are defined for all m, via the above recurrence equations.

The natural way of solving (2.28) is via backward iteration: From the known function v_n and using (2.27)–(2.29), compute v_{n-1}, then from v_{n-1} compute v_{n-2}, and so on,

down to v_0. Here, unfortunately, the functions v_m, for $m \leq n-2$, cannot be obtained in closed form (we will give a closed-form expression for v_{n-1} in a moment), so they must be approximated in some way. We propose an approximation method in Section 5. In the next section, we establish some properties of the option value and of its optimal strategy of exercise, which are interesting per se and are also useful for analyzing the approximation technique.

2.4 Characterizing the Value Function and the Exercise Strategy

2.4.1 The Value Function v_{n-1}

Recall that the value function v_n at the horizon $t_n = T$ has the simple form $v_n(s,\overline{s}) = (\overline{s}-K)^+$. We now derive a closed-form expression for the value function at time t_{n-1}, the last observation date before the horizon. We assume that $1 \leq m^* < n$ (otherwise one has $v_{n-1} = v_{n-1}^h$ and the argument simplifies). From (2.28), we have

$$\begin{aligned} v_{n-1}^h(s,\overline{s}) &= \rho E_{n-1,s,\overline{s}}\left[\left(\frac{(n-1)\overline{s}+S(t_n)}{n}-K\right)^+\right] \\ &= \frac{\rho}{n}E_{n-1,s,\overline{s}}\left[\left(S(t_n)-\overline{K}\right)^+\right], \end{aligned}$$

where $\overline{K} = nK-(n-1)\overline{s}$ and $E_{n-1,s,\overline{s}}$ denotes the conditional expectation above.

We first consider the case where $\overline{K} \leq 0$, i.e., $\overline{s} \geq Kn/(n-1)$. The holding value can then be derived from (2.26) as the linear function

$$v_{n-1}^h(s,\overline{s}) = v^{\text{lin}}(s,\overline{s}) = \frac{s}{n}-\frac{\rho}{n}\overline{K} = \frac{s}{n}+\rho\frac{n-1}{n}\overline{s}-\rho K,$$

and the exercise value equals $\overline{s} - K > 0$. One can easily identify the optimal decision (exercise or not) for any given state $(s, \overline{s})$ by comparing this exercise value with the holding value v^{lin}. This yields an explicit expression for the value function. Consider the line defined in the $(s, \overline{s})$ plane by $\overline{s} - K = v^{\text{lin}}(s, \overline{s})$, that is,

$$s - (n - (n-1)\rho)\,\overline{s} + nK(1-\rho) = 0. \tag{2.30}$$

The optimal strategy here is: Exercise the option if and only if $(s, \overline{s})$ is above the line (2.30). This line passes through the point $(K, K)n/(n-1)$ and has a slope of $1/(n-(n-1)\rho) < 1$, so it is optimal to exercise for certain pairs $(s, \overline{s})$ with $s > \overline{s}$, a possibility which was neglected by Grant, Vora, and Weeks (1997). A partial intuition behind this optimal strategy is that for sufficiently large $\overline{s}$ and for $s < \overline{s}$, the average price will most likely decrease in the future (it is pressured down by the current value), so it is best to exercise right away.

We now consider the case $\overline{K} > 0$, i.e., $\overline{s} < Kn/(n-1)$. In this case, the holding value is equivalent to the value of an European call option under the GBM assumption, with strike price $\overline{K}$, initial price s for the primitive asset, maturity horizon $T - t_{n-1} = h$, volatility σ, and risk-free rate r. This value is given by the well-known Black-Scholes formula:

$$v^{\text{BS}}(s, \overline{s}) = \frac{1}{n}\left(\Phi(d_1)s - \rho\overline{K}\Phi(d_1 - \sigma\sqrt{h})\right),$$

where

$$d_1 = \frac{\ln(s/\overline{K}) + (r + \sigma^2/2)h}{\sigma\sqrt{h}}$$

and Φ denotes the standard normal distribution function. If $\overline{s} \leq K$, one must clearly hold the option because the exercise value is 0. For $\overline{s} > K$, the optimal decision is obtained by comparing $v^{\text{BS}}(s, \overline{s})$ with $\overline{s} - K$, similar to what we did for the case where $\overline{K} \leq 0$. We now

have a closed-form expression for v_{n-1}:

$$v_{n-1}(s,\overline{s}) = \begin{cases} \max(\overline{s}-K, v^{\text{lin}}(s,\overline{s})) & \text{if } \overline{s} \geq Kn/(n-1) \\ \max(\overline{s}-K, v^{\text{BS}}(s,\overline{s})) & \text{if } \overline{s} < Kn/(n-1) \end{cases}.$$

We could (in principle) compute an expression for v_{n-2} by placing our expression for v_{n-1} in the DP equations (2.27) and (2.29), although this becomes quite complicated. The functions v_n and v_{n-1} are continuous and non-decreasing with respect to both s and $\overline{s}$. Finally, the optimal exercise region at t_{n-1} is the epigraph of some function φ_{n-1}, i.e., the region where $\overline{s} > \varphi_{n-1}(s)$, where $\varphi_{n-1}(s)$ is defined as the value of $\overline{s}$ such that $v^h_{n-1}(s,\overline{s}) = \overline{s} - K$. In the next subsection, we show that these general properties hold as well at time t_m, for $m < n$.

2.4.2 Properties of the Value Function and of the Exercise Frontier

We now prove certain monotonicity and convexity properties of the value function at each step, and use these properties to show that the optimal strategy of exercise at each step is characterized by a function φ_m whose epigraph partitions the state space in two pieces: At time t_m, for $m^* \leq m < n$, it is optimal to exercise the option if $\overline{s} \geq \varphi_m(s)$, and hold it if $\overline{s} \leq \varphi_m(s)$. We derive these properties when the asset moves as a GBM, but the proofs work as well if the ratio $S(t_{m+1})/S(t_m)$ has a different distribution than the lognormal, provided that it is independent of $\mathcal{F}(t_m)$.

Proposition 1 *For $m < n$, the holding value $v^h_m(s,\overline{s})$ is a continuous, strictly positive, strictly increasing, and convex function of both $s > 0$ and $\overline{s} > 0$. For $m \leq n$, the value function $v_m(s,\overline{s})$ also has the same properties except that it is only non-decreasing in $\overline{s}$.*

Proof. The proof proceeds by backward induction on m. At each step, we define the auxiliary random variable $\tau_{m+1} = S(t_{m+1})/S(t_m)$, which has the lognormal distribution with parameters μh and $\sigma\sqrt{h}$, independently of $\mathcal{F}(t_m)$, as explained in (2.25). A key step in our proof will be to write the holding value $v_m^h(s,\overline{s})$ as a convex combination of a continuous family of well-known functions indexed by τ_{m+1}.

For $m = n-1$, the holding value is

$$\begin{aligned} v_{n-1}^h(s,\overline{s}) &= \rho E_{n-1,s,\overline{s}}\left[v_n\left(\overline{S}_n\right)\right] \\ &= \rho\int_0^{+\infty}\left(\frac{(n-1)\overline{s}+s\tau}{n}-K\right)^+ f(\tau)\,d\tau, \end{aligned}$$

where f is the density of τ_n. For $s > 0$ and $\overline{s} > 0$, this conditional expectation does exist since $v_n\left(\overline{S}_n\right) = \left(\overline{S}_n - K\right)^+$ is dominated by a polynomial of τ_n, a lognormal random variable with finite moments.

The integrand is continuous and bounded by an integrable function of τ over any bounded interval for s and $\overline{s}$. Therefore, the holding value $v_{n-1}^h(s,\overline{s})$ is also continuous by Lebesgue's dominated convergence theorem (Billingsley 1986). The integral is strictly positive because, for instance, the lognormal distribution always gives a strictly positive measure to the event $\{(n-1)\overline{s}+s\tau_n - nK \geq n\}$, on which the integrand is greater than 1.

For $s > 0$, $\overline{s} > 0$, and $\delta > 0$, one has

$$\begin{aligned}
& v_{n-1}^{h}(s, \overline{s}+\delta) - v_{n-1}^{h}(s, \overline{s}) \\
= \; & \rho \int_{0}^{+\infty} \left[\left(\frac{(n-1)(\overline{s}+\delta) + s\tau}{n} - K \right)^{+} - \left(\frac{(n-1)\overline{s} + s\tau}{n} - K \right)^{+} \right] f(\tau)\, d\tau \\
\geq \; & \rho \int_{(nK-(n-1)\overline{s})/s}^{+\infty} \left[\frac{(n-1)(\overline{s}+\delta) + s\tau}{n} - \frac{(n-1)\overline{s} + s\tau}{n} \right] f(\tau)\, d\tau \\
\geq \; & \frac{n-1}{n}\delta > 0.
\end{aligned}$$

The same argument can be used to prove that $v_{n-1}^{h}(s, \overline{s})$ is strictly increasing in s. The convexity of $v_{n-1}^{h}(s, \overline{s})$ follows from the fact that this function is a positively weighted average (a convex combination), of the values of $[((n-1)\overline{s} + s\tau)/n - K]^{+}$, over all positive values of τ, which are (piecewise linear) convex functions of s and $\overline{s}$ for each τ. It is also straightforward to verify directly the definition of convexity for $v_{n-1}^{h}(s, \overline{s})$ (e.g., Rockafellar 1970).

Because the holding value is continuous and strictly positive, the value function

$$v_{n-1}(s, \overline{s}) = \max\left((\overline{s} - K)^{+}, v_{n-1}^{h}(s, \overline{s}) \right)$$

is continuous and strictly positive. It is also convex, non-decreasing in s, and strictly increasing in $\overline{s}$, because it is the maximum of two functions that satisfy these properties. Notice that v_{n-1} is also dominated by a polynomial of a lognormal random variable.

We now assume that the result holds for $m+1$, where $m \leq n-2$, and show that this implies that it holds for m. The holding value at t_m is

$$\begin{aligned}
v_{m}^{h}(s, \overline{s}) &= \rho E_{n-1,s,\overline{s}} \left[v_{m+1}\left(s\tau_{m+1}, \frac{m\overline{s} + s\tau_{m+1}}{m+1} \right) \right] \\
&= \rho \int_{0}^{+\infty} v_{m+1}\left(s\tau, \frac{m\overline{s} + s\tau}{m+1} \right) f(\tau)\, d\tau.
\end{aligned} \tag{2.31}$$

A similar argument, used for $m = n - 1$, ensures the existence of the above integral. The function $v_m^h(s, \bar{s})$ is continuous and strictly positive because the integrand is continuous, strictly positive, and bounded by an integrable function of τ over every bounded interval for s and $\bar{s}$. The other properties can be proved via similar arguments as for the case $m = n-1$. The proof for v_0 is also similar as for $m > 0$. We omit the details. ■

Lemma 2 *For $s > 0$ and $\bar{s}_2 > \bar{s}_1 > 0$, one has*

$$v_m^h(s, \bar{s}_2) - v_m^h(s, \bar{s}_1) \leq \frac{m}{m+1}\rho(\bar{s}_2 - \bar{s}_1) \quad \textit{for } m < n \tag{2.32}$$

and

$$v_m(s, \bar{s}_2) - v_m(s, \bar{s}_1) \leq (\bar{s}_2 - \bar{s}_1) \quad \textit{for } m \leq n. \tag{2.33}$$

Proof. The proof proceeds again by backward induction on m. We will use the property that $b^+ - a^+ \leq b - a$ when $a \leq b$. For $m = n$, one has $v_n(s, \bar{s}_2) - v_n(s, \bar{s}_1) = (\bar{s}_2 - K)^+ - (\bar{s}_1 - K)^+ \leq \bar{s}_2 - \bar{s}_1$ so (2.33) holds for $m = n$. We assume now that (2.33) holds for $m + 1$, where $m < n$, and show that this implies that (2.32) and (2.33) hold for m. From (2.31), one has

$$\begin{aligned}
& v_m^h(s, \bar{s}_2) - v_m^h(s, \bar{s}_1) \\
= \; & \rho \int_0^{+\infty} \left[v_{m+1}\left(s\tau, \frac{m\bar{s}_2 + s\tau}{m+1}\right) - v_{m+1}\left(s\tau, \frac{m\bar{s}_1 + s\tau}{m+1}\right) \right] f(\tau)\, d\tau \\
\leq \; & \rho \int_0^{+\infty} \left[\frac{m\bar{s}_2 + s\tau}{m+1} - \frac{m\bar{s}_1 + s\tau}{m+1} \right] f(\tau)\, d\tau \\
\leq \; & \frac{m}{m+1}\rho(\bar{s}_2 - \bar{s}_1).
\end{aligned}$$

Moreover, $v_m^e(\overline{s}_2) - v_m^e(\overline{s}_1) = (\overline{s}_2 - K)^+ - (\overline{s}_1 - K)^+ \leq \overline{s}_2 - \overline{s}_1$. Now,

$$\begin{aligned} & v_m(s, \overline{s}_2) - v_m(s, \overline{s}_1) \\ = \; & \max\left(v_m^e(\overline{s}_2), v_m^h(s, \overline{s}_2)\right) - \max\left(v_m^e(\overline{s}_1), v_m^h(s, \overline{s}_1)\right) \\ \leq \; & \max\left(v_m^e(\overline{s}_2) - v_m^e(\overline{s}_1), v_m^h(s, \overline{s}_2) - v_m^h(s, \overline{s}_1)\right) \\ \leq \; & \overline{s}_2 - \overline{s}_1. \end{aligned}$$

This completes the proof. ■

Proposition 3 *For $m^* \leq m < n$, there exists a continuous, strictly increasing, and convex function $\varphi_m : (0, +\infty) \to (K, +\infty)$ such that*

$$v_m^h(s, \overline{s}) = \begin{cases} > v_m^e(\overline{s}) & \text{for } \overline{s} < \varphi_m(s) \\ = v_m^e(\overline{s}) & \text{for } \overline{s} = \varphi_m(s) \\ < v_m^e(\overline{s}) & \text{for } \overline{s} > \varphi_m(s). \end{cases} \tag{2.34}$$

Proof. Let $s > 0$ and $m^* \leq m \leq n-1$. We know from Proposition 1 and Lemma 2 that $v_m^h(s, \overline{s})$ is strictly positive and strictly increasing in $\overline{s}$, with a growth rate always less than $\rho m/(m+1) < 1$. On the other hand, $v_m^e(\overline{s}) = (\overline{s} - K)^+$ is 0 for $\overline{s} \leq K$ and increases at rate 1 for $\overline{s} > K$. Therefore, there is a unique value of $\overline{s} > K$, denoted $\varphi_m(s)$, such that (2.34) is satisfied. This proves the existence of the function $\varphi_m(s)$, for all $s > 0$.

To show that $\varphi_m(s)$ is strictly increasing in s, let $s_2 > s_1 > 0$. If we suppose that $\varphi_m(s_1) \geq \varphi_m(s_2)$, we obtain the contradiction

$$\begin{aligned} 0 \; < \; & v_m^h(s_2, \varphi_m(s_2)) - v_m^h(s_1, \varphi_m(s_2)) \\ = \; & v_m^h(s_2, \varphi_m(s_2)) - v_m^h(s_1, \varphi_m(s_1)) + v_m^h(s_1, \varphi_m(s_1)) - v_m^h(s_1, \varphi_m(s_2)) \\ = \; & \varphi_m(s_2) - K - [\varphi_m(s_1) - K] + v_m^h(s_1, \varphi_m(s_1)) - v_m^h(s_1, \varphi_m(s_2)) \\ = \; & v_m^h(s_1, \varphi_m(s_1)) - v_m^h(s_1, \varphi_m(s_2)) - [\varphi_m(s_1) - \varphi_m(s_2)] \leq 0, \end{aligned}$$

where the first inequality is by Proposition 1 and the last inequality by (2.32) in Lemma 2. Therefore, $\varphi_m(s_1) < \varphi_m(s_2)$, i.e., $\varphi_m(s)$ is strictly increasing in s.

Now, consider two states $(s_1, \overline{s}_1)$ and $(s_2, \overline{s}_2)$ where it is optimal to exercise the option at time t_m, i.e., for which $\overline{s}_1 \geq \varphi_m(s_1)$ and $\overline{s}_2 \geq \varphi_m(s_2)$, and let $(s_\lambda, \overline{s}_\lambda) = \lambda(s_1, \overline{s}_1) + (1-\lambda)(s_2, \overline{s}_2)$, for $0 \leq \lambda \leq 1$. By Proposition 1, v_m^h is convex and

$$\begin{aligned} v_m^h(s_\lambda, \overline{s}_\lambda) &\leq \lambda v_m^h(s_1, \overline{s}_1) + (1-\lambda) v_m^h(s_2, \overline{s}_2) \\ &\leq \lambda(\overline{s}_1 - K) + (1-\lambda)(\overline{s}_2 - K) \\ &= \overline{s}_\lambda - K \\ &= v_m^e(\overline{s}). \end{aligned}$$

Therefore, in state $(s_\lambda, \overline{s}_\lambda)$, it is optimal to exercise the option. We have just proved that for any two points lying above the function φ_m, the straight line joining these two points is also above the function φ_m. This implies that φ_m is convex. The convexity then implies the continuity. ■

For $m^* \leq m < n$, we define the optimal *exercise frontier* at time t_m as the graph of φ_m, i.e., the locus of points $(s, \overline{s})$ such that $v_m^h(s, \overline{s}) = v_m^e(\overline{s})$. It is optimal to exercise the Amerasian call option at time t_m if $\overline{s} \geq \varphi_m(s)$, and hold it at least until t_{m+1} if $\overline{s} < \varphi_m(s)$.

2.5 Numerical Solution of the DP Equation

The general idea is to partition the positive quadrant of the plan $(s, \overline{s})$ by a rectangular grid and to approximate the value function at each observation date by a function which is polynomial on each rectangle of the grid, i.e., piecewise polynomial. However, instead

of fitting the approximation to v_m directly, we redefine the value function such that the state variables at time t_m become the stock price at time t_m and the average of prices up to time t_{m-1}, denoted by $(s, \overline{s}')$. This change of variable simplifies the integration when the piecewise-polynomial approximation is incorporated into (2.28), which we compute explicitly. The polynomial functions we tried are linear in s, and linear, quadratic, or cubic in $\overline{s}'$ (Conte and de Boor 1980). Other types of approximations may be used in this context (see, e.g., de Boor 1978), but we found that the technique proposed here gives a good compromise in terms of the amount of work required to achieve a given precision.

2.5.1 A Piecewise Approximation

At time t_m, let

$$\overline{s}' = \begin{cases} (m\overline{s} - s)/(m-1) & \text{if } m > 1 \\ 0 & \text{if } m \leq 1 \end{cases},$$

be the value of $\overline{S}_{m-1}$ if $S(t_m) = s$ and $\overline{S}_m = \overline{s}$, and define

$$w_m(s, \overline{s}') = v_m(s, \overline{s}),$$

where $\overline{s} = ((m-1)\overline{s}' + s)/m$ if $m \geq 1$ and $\overline{s} = 0$ if $m = 0$. The function w_m has the same properties as stated for v_m in Proposition 1, except that w_1 does not depend on $\overline{s}'$. Equation (2.27) can be written in terms of w_m as

$$w_m^e(s, \overline{s}') = v_m^e(s, \overline{s}) = (\overline{s} - K)^+, \quad \text{for } 0 \leq m \leq n, \tag{2.35}$$

equation (2.28) as

$$\begin{aligned} w_m^h(s, \overline{s}') &= v_m^h(s, \overline{s}) \\ &= \rho E_{m,s,\overline{s}}\left[w_{m+1}(s\tau_{m+1}, \overline{s})\right], \quad \text{for } 0 \leq m \leq n, \end{aligned} \tag{2.36}$$

and equation (2.29) as

$$w_m(s,\overline{s}') = \begin{cases} w_m^h(s,\overline{s}') & \text{if } 0 \le m < m^* \\ \max\left(w_m^e(s,\overline{s}'), w_m^h(s,\overline{s}')\right) & \text{if } m^* \le m < n \\ w_m^e(s,\overline{s}') & \text{if } m = n. \end{cases} . \tag{2.37}$$

The idea now is to approximate each value function w_m by a continuous piecewise polynomial function of $(s,\overline{s}')$, of degree 1 (linear) in s and degree d in $\overline{s}'$. We have experimented with $d = 1, 2$, and 3, i.e., (piecewise) bilinear, linear-quadratic, and linear-cubic functions.

We first define a grid by selecting $0 = a_0 < a_1 < \ldots < a_p < a_{p+1} = \infty$ and $0 = b_0 < b_1 < \ldots < b_q < b_{q+1} = \infty$. The grid points $G = \{(a_i, b_j) : 0 \le i \le p \text{ and } 0 \le j \le q\}$ is a partition of the positive quadrant $[0,\infty) \times [0,\infty)$ into the $(p+1)(q+1)$ rectangles $R_{i,j} = \{(s,\overline{s}') : a_i \le s < a_{i+1} \text{ and } b_j \le \overline{s}' < b_{j+1}\}$, for $i = 0,\ldots,p$ and $j = 0,\ldots,q$. In our implementation, we took $a_1 = b_1$ and $a_p = b_q$. Under these conditions, if $(s,\overline{s}') \in G$, then $\overline{s}$ in (2.36) is always in the interval $[b_1, b_q]$, for $m > 0$, which means that we need an approximation of w_{m+1} only over the box $B = [0,\infty) \times [b_1, b_q]$.

Suppose for now that $q-1$ is a multiple of d. We regroup the rectangles covering the box B as follows. For $j \in J^{(d)} = \{1, d+1, 2d+1, \ldots, q-d\}$, let

$$R_{i,j}^{(d)} \begin{cases} R_{0,j} \cup \ldots \cup R_{0,j+d-1} \cup R_{1,j} \cup \ldots R_{1,j+d-1} & \text{for } i = 1 \\ R_{i,j} \cup \ldots \cup R_{i,j+d-1} & \text{for } 1 < i < p-1 \\ R_{p-1,j} \cup \ldots \cup R_{p-1,j+d-1} \cup R_{p,j} \cup \ldots R_{p,j+d-1} & \text{for } i = p-1 \end{cases} .$$

We consider the approximating functions that are linear in s and polynomial of degree d in $\overline{s}'$ over each of the rectangles $R_{i,j}^{(d)}$, and continuous across these rectangles.

To construct these piecewise polynomials, at each step m, for $m = n-2,\ldots,2$, we first compute an approximation of w_m at each point of G, denoted by $\widetilde{w}_m$. This is done via the DP equations (2.35)-(2.37), using an available approximation $\widehat{w}_{m+1}$ of the value function w_{m+1}. Then we interpolate $\widetilde{w}_m$, defined at each point of G, with a piecewise

function linear in s and polynomial of degree d in $\overline{s}'$, for $(s, \overline{s}') \in B$. More specifically, for $m = n-2, \ldots, 2$, over the line segment defined by $s = a_i$ and $\overline{s}' \in [b_j, b_{j+d}]$ for $j \in J^{(d)}$, we approximate w_m by the Newton interpolation polynomial

$$\begin{aligned} \widehat{w}_m(a_i, \overline{s}') &= P^m_{i,j}(\overline{s}') \qquad (2.38) \\ &= \alpha^m_{i,j,0} + \alpha^m_{i,j,1}(\overline{s}' - b_j) + \cdots + \alpha^m_{i,j,d}(\overline{s}' - b_j)\ldots(\overline{s}' - b_{j+d-1}). \end{aligned}$$

The coefficients $\alpha^m_{i,j,l}$ are obtained by interpolating the known values of $\widetilde{w}_m$ at the points $(a_i, b_j), \ldots, (a_i, b_{j+d})$, i.e., by solving the system of linear equations defined by:

$$\widetilde{w}_m(a_i, b_j) = P^m_{i,j}(\overline{s}'), \quad \text{for } \overline{s}' = b_j, \ldots, b_{j+d}. \qquad (2.39)$$

This is done for $i = 1, \ldots, p$, and $j \in J^{(d)}$. Then, over each rectangle $R^{(d)}_{i,j}$ for $i < p$, the approximation $\widehat{w}_m(s, \overline{s}')$ of $w_m(s, \overline{s}')$ is defined by the linear interpolation

$$\widehat{w}_m(s, \overline{s}') = \frac{a_{i+1} - s}{a_{i+1} - a_i} P^m_{i,j}(\overline{s}') + \frac{s - a_i}{a_{i+1} - a_i} P^m_{i+1,j}(\overline{s}'). \qquad (2.40)$$

In the case where $q-1$ is not a multiple of d, if j^* is the largest $j < q$ such that d divides $j-1$, we use a polynomial approximation of degree $q - j^*$ ($< d$) in $\overline{s}'$ over the rectangles $R^{(d)}_{i,j^*}$, redefined by $R^{(d)}_{i,j^*} = R_{i,j^*} \cup \cdots \cup R_{i,q-1}$.

2.5.2 Explicit Integration for Function Evaluation

Now, we examine how to compute $\widetilde{w}_m(a_i, b_j)$ given an available approximation $\widehat{w}_{m+1}$ of w_{m+1}. Observe that $w^h_m(s, \overline{s}')$ in (2.36) is expressed as an expectation with respect to a *single* random variable, τ_{m+1}, and we have chosen our change of variable $(s, \overline{s}) \longrightarrow (s, \overline{s}')$ precisely to obtain this property. More specifically, the holding value $w^h_m(s, \overline{s}')$ is

approximated by

$$\begin{aligned}\widetilde{w}_m^h(s,\overline{s}') &= \rho E_{m,s,\overline{s}}\left[\widehat{w}_{m+1}(s\tau_{m+1},\overline{s})\right] \qquad (2.41)\\ &= \rho E_{m,s,\overline{s}}\left[\sum_{i=1}^{p-1}\left(\frac{a_{i+1}-s\tau_{m+1}}{a_{i+1}-a_i}P_{i,\xi}^{m+1}(\overline{s})+\frac{s\tau_{m+1}-a_i}{a_{i+1}-a_i}P_{i+1,\xi}^{m+1}(\overline{s})\right)I_i\right],\end{aligned}$$

where ξ is the element of $J^{(d)}$ such that $\overline{s}\in[b_\xi,b_{\xi+1}]$ and I_i is the indicator of the event $\{(s\tau_{m+1},\overline{s})\in R_{i,\xi}\}$.

The function $\widetilde{w}_m^h$ is to be evaluated over the points (a_k,b_l) of the grid G. For $s=a_k$, $\overline{s}'=b_l$, and $\overline{s}=c_{k,l}=((m-1)b_l+a_k)/m$, we have by rearranging the terms in (2.41)

$$\widetilde{w}_m^h(a_k,b_l)=\rho\sum_{i=1}^{p}D_{k,i}P_{i,\xi}^{m+1}(c_{k,l}), \qquad (2.42)$$

where ξ is the element of $J^{(d)}$ such that $c_{k,l}\in[b_\xi,b_{\xi+1})$,

$$D_{k,i}=\begin{cases}E\left[\frac{a_2-a_k\tau}{a_2-a_1}I_1\right] & \text{for } i=1,\\ E\left[\frac{a_{i+1}-a_k\tau}{a_{i+1}-a_i}I_i+\frac{a_k\tau-a_{i-1}}{a_i-a_{i-1}}I_{i-1}\right] & \text{for } 1<i<p,\\ E\left[\frac{a_k\tau-a_{p-1}}{a_p-a_{p-1}}I_{i-1}\right] & \text{for } i=p,\end{cases}$$

τ is a lognormal random variable with parameters μh and $\sigma\sqrt{h}$, and I_i is the indicator of the event $\{a_k\tau<a_2\}$ if $i=1$, $\{a_{p-1}\le a_k\tau\}$ if $i=p-1$, and $\{a_i\le a_k\tau<a_{i+1}\}$ if $1<i<p-1$. Defining $A_{k,i}=E[I_i]$ and $B_{k,i}=a_kE[\tau I_i]$, we can rewrite

$$D_{k,i}=\begin{cases}\frac{a_2A_{k,1}-B_{k,1}}{a_2-a_1} & \text{for } i=1,\\ \frac{a_{i+1}A_{k,i}-B_{k,i}}{a_{i+1}-a_i}+\frac{B_{k,i-1}-a_{i-1}A_{k,i-1}}{a_i-a_{i-1}} & \text{for } 1<i<p,\\ \frac{B_{k,p-1}-a_{p-1}A_{k,p-1}}{a_p-a_{p-1}} & \text{for } i=p.\end{cases}$$

Knowing that $E[I(a_k\tau<a_i)]=\Phi(x_{k,i})$ and $E[\tau I(a_k\tau<a_i)]=\Phi\left(x_{k,i}-\sigma\sqrt{h}\right)e^{rh}$ where Φ is the standard normal distribution and $x_{k,i}=[\ln(a_i/a_k)-\mu h]/\left(\sigma\sqrt{h}\right)$, we easily find that

$$A_{k,i}=\begin{cases}\Phi(x_{k,2}) & \text{for } i=1\\ \Phi(x_{k,i+1})-\Phi(x_{k,i}) & \text{for } 1<i<p-1\\ 1-\Phi(x_{k,p-1}) & \text{for } i=p-1\end{cases}$$

and

$$B_{k,i} = \begin{cases} a_k \Phi\left(x_{k,2} - \sigma\sqrt{h}\right) e^{rh} & \text{for } i = 1 \\ a_k \left[\Phi\left(x_{k,i+1} - \sigma\sqrt{h}\right) - \Phi\left(x_{k,i} - \sigma\sqrt{h}\right)\right] e^{rh} & \text{for } 1 < i < p-1 \\ a_k \left[1 - \Phi\left(x_{k,p-1} - \sigma\sqrt{h}\right)\right] e^{rh} & \text{for } i = p-1. \end{cases}$$

The constants $A_{k,i}$, $B_{k,i}$, and $D_{k,i}$ are precomputed before doing the first iteration, and the $D_{k,i}$ are then used to evaluate (2.42) at each step m.

This yields the approximate value function

$$\widetilde{w}_m(a_k, b_l) = \max\left(\widetilde{w}_m^h(a_k, b_l), (c_{k,l} - K)^+\right). \tag{2.43}$$

These values at the grid points are then interpolated to obtain the function $\widehat{w}_m$ as explained earlier. Integration and interpolation stages are repeated successively until $m = 0$, where an approximation of w_0 and of the option value is finally obtained. An important advantage of choosing the same grid G for all m is that the coefficients $D_{k,i}$ can be precomputed once for all. Evaluating $\widetilde{w}_m^h$ at the grid points via (2.42) is then very fast.

It would also be possible to use an adaptive grid, where the grid points change with the observation dates. This could be motivated by the fact that the probability distribution of the state vector changes with time. In that case, the coefficients $D_{k,i}$ would depend on the step m and would have to be recomputed at each observation date.

As it turns out, this procedure evaluates, with no extra cost, the option value and the optimal decision at all observation dates and in all states. This could be used for instance to estimate the sensitivity of the option value with respect to the initial price. Eurasian options can be evaluated via this procedure as well, because they are a special case.

Finally, we explain how the DP procedure works:

Compute $\widetilde{w}_{n-1}(a_k, b_l)$ for all the grid points (a_k, b_l) (see Section 2.4.1);

Compute the coefficients of $\widehat{w}_{n-1}$ by (2.39)-(2.40);

FOR $m = n-1, 0$ by step -1

FOR $k = 1, p$

FOR $l = 1, q$

Compute $\widetilde{w}_m^h(a_k, b_l)$ for all the grid points (a_k, b_l) by (2.42);

Compute $\widetilde{w}_m(a_k, b_l)$ for all the grid points (a_k, b_l) by (2.43);

NEXT l

NEXT k

Compute the coefficients of $\widehat{w}_m$ by (2.39)-(2.40);

NEXT m

2.5.3 Computational Speed-up and Complexity Analysis

Whenever $w_m^e(a_k, b_l) = (c_{k,l} - K)^+ > \widetilde{w}_m^h(a_k, b_l)$ at some point (a_k, b_l) for some $m \geq m^*$, it is optimal to exercise the option at that point and it is optimal to exercise for all $j \geq l$ (see Proposition 3). There is no need to compute $\widetilde{w}_m^h(a_k, b_j)$ for $j \geq l$, and this saves computation effort. We also exploit the fact that whenever $\widetilde{w}_m^h(a_k, b_l)$ is small enough to be negligible (say less than $\epsilon_1 = 10^{-6}$), there is no need to compute $\widetilde{w}_m^h(a_k, b_j)$ for $j < l$, since w_m^h is a strictly increasing function of $\overline{s}'$ (see Proposition 1); it can be replaced by 0.

The time complexity of this algorithm is $O(p^2)$ to precompute the coefficients $D_{k,i}$, plus $O(np^2q)$ to compute the sum in (2.42) for each of the pq grid points at each of the n

steps, plus $O(npq)$ to solve the systems of linear equations giving the coefficients in (2.39). The overall time complexity is thus $O(np^2q)$.

For large p, most of the work is for computing the terms of (2.42) and time can be saved by observing that several of these terms are negligible. When time increment h is small, the largest terms of $\widetilde{w}_m^h(a_k, b_l)$ in (2.42) are typically for i near k; they become negligible when i is far from k. For larger h, the most important terms are for i somewhat larger than k. In our implementation, we initially choose a small constant ϵ_2 (we took $\epsilon_2 = \epsilon_1/p$). When computing (2.42), we first select a starting point j, we add up the terms for $i = j+1, j+2, \ldots$ and stop when whenever a term is less than ϵ_2, then we add the terms $j-1, j-2, \ldots$ and stop whenever a term is less than ϵ_2. To select j, if the term for $i = k$ exceeds $16\epsilon_2$, we take $j = k$, otherwise we try $j = \lfloor (k+p)/2 \rfloor$ and $j = \lfloor (3k+p)/4 \rfloor$. In the (rare) cases where the term for j is still less than $16\epsilon_2$, we sum all the terms for $i = k+1, \ldots, p$.

2.5.4 Convergence

Proving the convergence of the DP algorithm as the grid size becomes finer and finer is not straightforward because the state space is unbounded and the value function increases unboundedly when s and $\overline{s}'$ goes to infinity. However, if $c = \min(a_p, b_q) \to \infty$, then by Lemma 3 of Conze and Viswanathan (1991) and standard large deviations approximation

for the normal distribution, we have

$$\begin{aligned} & Q\left[\max_{0\le t\le T} S(t)/S > c\right] \\ = \; & Q\left[\max_{0\le t\le T} \ln\left(S(t)/S\right) > \ln c\right] \\ = \; & 1-\Phi\left(\frac{\ln(c)-\mu T}{\sigma\sqrt{T}}\right)+\exp\left(\frac{2\mu\ln c}{\sigma^2}\right)\Phi\left(-\frac{\mu T+\ln c}{\sigma\sqrt{T}}\right) \\ = \; & O\left(\frac{1}{\ln c}\exp\left(-\frac{\ln^2 c}{2\sigma^2 T}+O(\ln c)\right)\right) \\ = \; & O\left(\frac{1}{\ln c}c^{-\ln c/(2\sigma^2 T)}\right). \end{aligned}$$

Thus, the probability that the trajectory of $\left\{\left(S(t),\overline{S}(t)\right), 0\le t\le T\right\}$ ever exits the box $B=(0,a_p]\times(0,b_q]$ decreases to 0 at a rate faster than $O(1/P(c))$ for any polynomial P. On the other hand, one can show that the error on the value function can only increase linearly when s or $\overline{s}'$ goes to infinity. As a consequence, the effect of the approximation error outside the box B becomes negligible if the box is large enough. This is the basic idea of the convergence proof detailed in the next proposition. Besides proving convergence, the proposition says that with the bilinear interpolation ($d=1$), if we neglect the effect of the approximation error outside the box B, our DP procedure gives an *upper bound* on the option value.

Proposition 4 *Define* $\delta_a=\sup_{1\le i\le p} a_i-a_{i-1}$ *and* $\delta_b=\sup_{1\le j\le q} b_j-b_{j-1}$ *and assume that each approximation* $\widehat{w}_m$ *is non decreasing. If* $p\to\infty$, $q\to\infty$, $a_p\to\infty$, $b_q\to\infty$, $\delta_a\to 0$, *and* $\delta_b\to 0$, *then for any constant* $c>0$,

$$\sup_{0\le m<n}\sup_{(s,\overline{s}')\in(0,c]^2} \mid \widehat{w}_m(s,\overline{s}')-w_m(s,\overline{s}') \mid \to 0.$$

Moreover, if $d = 1$, $\widehat{w}_m(s, \overline{s}') \geq w_m(s, \overline{s}')$ *for all* m *and all* $(s, \overline{s}') \in (0, c]^2$, *when* a_p *and* b_q *are large enough.*

Proof. In Lemma 2, we showed that the derivative of $v_m(s, \overline{s})$ and $v_m^h(s, \overline{s})$ with respect to $\overline{s}$ never exceeds 1. We now show that their derivatives at step m with respect to s is also bounded by the constant C_m, defined recursively by $C_n = 0$ and

$$\begin{aligned} C_m &= \left(C_{m+1} + \frac{1}{m+1}\right) \rho E[\tau_{m+1}] \\ &= C_{m+1} + \frac{1}{m+1}. \end{aligned}$$

By backward induction on m, it can be shown that for $s_1 \leq s_2$,

$$v_m(s_2, \overline{s}) - v_m(s_1, \overline{s}) \leq (s_2 - s_1) C_m. \tag{2.44}$$

This is clearly true for $m = n$, because $v_n(s, \overline{s})$ does not depend on s. If we assume that (2.44) holds for $m + 1$, then, using Lemma 2,

$$\begin{aligned} & v_m^h(s_2, \overline{s}) - v_m^h(s_1, \overline{s}) \\ = \; & \rho \int_0^\infty \left[v_{m+1}\left(s_2\tau, \frac{m\overline{s} + s_2\tau}{m+1}\right) - v_{m+1}\left(s_1\tau, \frac{m\overline{s} + s_1\tau}{m+1}\right) f(\tau)\, d\tau \right] \\ \leq \; & \rho \int_0^\infty \left[(s_2 - s_1)\tau C_{m+1} + \frac{(s_2 - s_1)\tau}{m+1} \right] f(\tau)\, d\tau \\ = \; & \rho (s_2 - s_1) \left(C_{m+1} + \frac{1}{m+1}\right) E[\tau_{m+1}] \\ = \; & (s_2 - s_1) C_m. \end{aligned}$$

Note that the slope of v_m with respect to s never exceeds that of v_m^h. This implies (2.44) and completes the induction. Moreover, the derivative of w_m does not exceed that of v_m, and similarly for w_m^h, so Lemma 2 and (2.44) hold for these functions as well.

Using these bounds on the slope, and the fact that w_m is increasing with respect to each of its arguments, we obtain that

$$\sup_{1\le k<p,\,1\le l<q} w_m(a_{k+1}, b_{l+1}) - w_m(a_k, b_l) \le C_m\delta_a + \delta_b. \tag{2.45}$$

Define now $\epsilon_n = 0$ and $\epsilon_m = 2\rho\epsilon_{m+1} + C_m\delta_a + \delta_b$, for $m = n-1, \ldots, 0$. Also, define an increasing of square boxes $B_m = (0, c_m]^2$ as follows. Choose $c_0 = c$ arbitrarily, and for $m = 0, \ldots, n-1$, choose $c_{m+1} \ge c_m$ large enough so that for all $(s, \overline{s}') \in B_m$,

$$\begin{aligned} & E_{m,s,\overline{s}}\left[\mid \widehat{w}_{m+1}(s\tau_{m+1}, \overline{s}) - w_{m+1}(s\tau_{m+1}, \overline{s})\, I\left((s\tau_{m+1}, \overline{s}) \notin B_{m+1}\right)\right] \\ \le\ & \epsilon_{m+1}, \end{aligned} \tag{2.46}$$

where I is the indicator function. Such a c_{m+1} exists because both $\widehat{w}_{m+1}$ and w_{m+1} are bounded by a linear function and the probability of exiting the box B_{m+1} decreases faster than the inverse of any (positive) linear function of c_{m+1} when $c_{m+1} \to \infty$, thanks to (??).

We now show, by backward induction on m, that $\mid \widehat{w}_m - w_m \mid$ is bounded by ϵ_m over the box B_m. This is clearly true for $m = n$, because $\widehat{w}_n = w_n$. Now, if we assume that $\mid \widehat{w}_{m+1} - w_{m+1} \mid \le \epsilon_{m+1}$ on B_{m+1}, then we have, for $(s, \overline{s}') \in B_m$,

$$\begin{aligned} & \mid \widetilde{w}_m(s, \overline{s}') - w_m(s, \overline{s}') \mid \\ \le\ & \rho \int_0^\infty \mid \widehat{w}_{m+1}(s\tau, \overline{s}) - w_{m+1}(s\tau, \overline{s}) \mid f(\tau)\, d\tau \\ =\ & \rho E_{m,s,\overline{s}}\left[\mid \widehat{w}_{m+1}(s\tau_{m+1}, \overline{s}) - w_{m+1}(s\tau_{m+1}, \overline{s})\, I\left((s\tau_{m+1}, \overline{s}) \in B_{m+1}\right)\right] + \\ & \rho E_{m,s,\overline{s}}\left[\mid \widehat{w}_{m+1}(s\tau_{m+1}, \overline{s}) - w_{m+1}(s\tau_{m+1}, \overline{s})\, I\left((s\tau_{m+1}, \overline{s}) \notin B_{m+1}\right)\right] \\ \le\ & 2\rho\epsilon_{m+1}. \end{aligned}$$

Then, because $\widehat{w}_m$ is a non-decreasing interpolation of $\widetilde{w}_m$ in the box B_m and using (2.45), we have

$$| \widehat{w}_m(s, \overline{s}') - w_m(s, \overline{s}') | \leq 2\rho\epsilon_{m+1} + C_m\delta_a + \delta_b = \epsilon_m.$$

Under the assumptions of the proposition, $\epsilon_m \to 0$ and this proves the first part.

In the case of the piecewise-linear interpolation ($d = 1$), it is also easily seen that the interpolation always overestimates w_m, because w_m is increasing and convex. That is, by backward induction on m, one can show that $\widehat{w}_m \geq \widetilde{w}_m \geq w_m$. ■

2.5.5 Grid Choice, Refinement, and Convergence Acceleration

In our experiments, we took $p = q$, $a_i = b_i$ for all i, $a_1 = S\exp(\mu t_{n-1} - 5\sigma\sqrt{t_{n-1}})$, $a_{p-1} = S\exp(\mu t_{n-1} + 5\sigma\sqrt{t_{n-1}})$, and $a_p = S\exp(\mu t_{n-1} + 6\sigma\sqrt{t_{n-1}})$, and, for $i = 2, \ldots, p-2$, a_i is the quantile of order $(i-1)/(p-2)$ of the lognormal distribution with parameters μt_{n-1} and $\sigma\sqrt{t_{n-1}}$. Others distributions may be used to generate the points b_j, e.g, the lognormal approximation of the average prices of the stock (Levy 1992).

To assess the discretization error, we applied the algorithm repeatedly for different grid sizes, doubling the value of p ($= q$) each time, until there was no significant change in the option value estimate. We also made a sensitivity analysis with respect to ϵ_1: For $\epsilon_1 < 10^{-6}$ (we took 10^{-7} and 10^{-8}), the results for the option value were the same as those reported here. Also, replacing the numbers 5 and 6 by larger numbers in the definitions of a_1, a_{p-1}, and a_p did not change the results.

The sequence of DP approximations of the option value for successive values of p, where each value of p is twice the previous one, converges to the true option value (with

negligeable error) when $p \to +\infty$. The straightforward way of estimating the option value is then to take the approximation obtained with the finest grid (largest p). But one can do better by using extrapolation methods designed for accelerating the convergence sequences, and which transform a given sequence into another sequence that converges more rapidly under very broad conditions. One of these methods is the *ϵ-algorithm* introduced by Wynn in 1956 (see Brézinski 1978), and which works as follows. Let $s_0^{(1)}, \ldots, s_l^{(1)}$ be the original sequence of length $l+1$, where $l+1$ is odd, and let $s_k^{(0)} = 0$, for all k. For $j = 2, \ldots, l+1$ and $k = 0, \ldots, l-j+1$, define

$$s_k^{(j)} = s_{k+1}^{(j-2)} + \left[s_{k+1}^{(j-1)} - s_k^{(j-1)} \right]^{-1}.$$

The $(l+1)$th and last sequence has a single term, $s_0^{(l+1)}$, which is the final approximation of the convergence point of the original sequence. This algorithm performs surprisingly well in various contexts. In our context, we take it as a heuristic. In this paper, we took $l+1 = 3$. The original sequence is formed by taking the DP approximation of the successive grid sizes $(p/4)^2$, $(p/2)^2$, and p^2. The approximation $s_0^{(3)}$ thus obtained will be compared to the DP approximation obtained with the p^2 grid alone. The CPU time required for computing $s_0^{(3)}$ is approximately the sum of CPU times required by DP for the three grids, and is only marginally more than the CPU time for the p^2 grid alone.

2.6 Numerical Experiments and Examples

We experimented the linear-linear, linear-quadratic, and the linear-cubic interpolations presented in Section 2.5.1. The linear-quadratic interpolation appears to be the most efficient.

This is certainly related to the shape of the value curve: an increasing and convex surface. We report here on the linear-linear ($d = 1$) and the linear-quadratic ($d = 2$) interpolations.

2.6.1 Example 1

For our first example, we take the parameter values $S = 100$, $K = 100$, $T = 1/4$ (years), $\sigma = 0.15$ (per year), $r = 0.05$ (per year), $h = 1/52$ (years), $m^* = 1$, and $n = 13$. We thus have a 13-week contract, with an exercise opportunity at each observation epoch, which is every week. We also consider 3 variants of this example: We first increase the volatility σ from 0.15 to 0.25, we then increase T from $1/4$ to $1/2$ (26 weeks) while keeping $n = 13$, and finally we increase K from 100 to 105, which yields an out-of-the-money option. In each case, we evaluate the Eurasian and Amerasian options with 5 grid sizes. In Tables 1 and 2, for each parameter set and each grid size, we give the DP approximation of the option value obtained with $d = 1$ (above) and $d = 2$ (below). The CPU times reported here are for a 750Mhz PC running the Linux operating system. They are for the first set of parameters, and are approximately the same for the other sets. The programs are written in FORTRAN and were compiled with the GNU g77 compiler.

To assess the precision of the DP procedure, we price the European versions of the contracts in Table 1. The values computed by DP are compared with those computed by efficient Monte Carlo simulation as suggested by Kemna and Vorst (1990). The last column of Table 1 gives their 95% confidence intervals obtained with 10^6 replications.

Table 1: Prices of Eurasian call options

	p					
(K, T, σ)	150	300	600	1200	2400	v_{eu}(sim)
(100, 0.25, 0.15)	2.16581	2.16510	2.16493	2.16489	2.16487	
	2.16547	2.16502	2.16491	2.16488	2.16487	[2.16483, 2.16492]
(100, 0.25, 0.25)	3.36556	3.36439	3.36411	3.36402	3.36402	
	3.36501	3.36426	3.36408	3.36403	3.36402	[3.36393, 3.36416]
(100, 0.50, 0.25)	4.92929	4.92765	4.92726	4.92716	4.92713	
	4.92851	4.92746	4.92721	4.92715	4.92713	[4.92694, 4.92743]
(105, 0.50, 0.25)	2.80811	2.80647	2.80607	2.80597	2.80595	
	2.80734	2.80628	2.80602	2.80596	2.80594	[2.80581, 2.80625]
CPU (h:mn:s)	0:00:01	0:00:10	0:01:16	0:09:47	1:18:25	
	0:00:01	0:00:10	0:01:15	0:09:50	1:23:13	00:00:13

The CPU times are approximately the same with $d = 1$ and 2. With $d = 2$, there are more coefficients to determine in each rectangle than with $d = 1$, but there is only half the number of boxes for the same grid. The DP approximation converges rapidly as the grid size is refined. With the coarse grid ($p = 150$), the error is already less than one quarter of a cent, and the result is obtained in half of a second. The values obtained with the ϵ-algorithm (not shown in the table) converge even faster: With $p = 600$, these values never differ by more than 0.00001 from the values given in the table with $p = 2400$, which can be considered as practically exact.

When we compare the value of the Amerasian option with its Eurasian counterpart, we see that the privilege of early exercise increases the value of the option, as expected. The contract becomes more expensive when the volatility or the maturity date are increased because this gives more chance of achieving a large average. It becomes cheaper when the strike price is increased, i.e., in the case of an out-of-the-money option.

Table 2: Prices of Amerasian call options

	p				
(K, T, σ)	150	300	600	1200	2400
$(100, 0.25, 0.15)$	2.32209	2.32114	2.32091	2.32085	2.32084
	2.32156	2.32101	2.32088	2.32085	2.32084
$(100, 0.25, 0.25)$	3.65221	3.65058	3.65019	3.65009	3.65006
	3.65129	3.65037	3.65013	3.65008	3.65006
$(100, 0.50, 0.25)$	5.33506	5.33273	5.33217	5.33203	5.33200
	5.33375	5.33242	5.33209	5.33201	5.33199
$(105, 0.50, 0.25)$	2.96851	2.96633	2.96580	2.96567	2.96564
	2.96732	2.96605	2.96573	2.96565	2.96563
CPU (h:mn:s)	0:00:01	0:00:06	00:00:48	0:06:04	0:48:37
	0:00:01	0:00:06	00:00:48	0:06:16	0:51:15

To quantify the impact of increasing the number of early exercise opportunities (and observation dates), we performed additional experiments with the same parameter sets as in Table 2, but with different values of n ranging from 1 to 52. For each set of parameters, the top and bottom lines of Table 3 give the value of the Amerasian and the Eurasian call option computed by DP with $p = 600$ and a linear-quadratic interpolation.

We see that increasing n decreases the option value. This may look strange in the case of Amerasian call options since they include much more exercise opportunities. The explanation is that increasing the number of observation dates increases the stability of the average prices, and this offsets the advantage of having more exercise opportunities (for the Amerasian contract). Note that $n = 1$ corresponds to a standard European call. For $n = 2$, it is optimal to exercise at time t_1 only if $S(t_1) = \bar{s} \geq 2K$ (see Section 2.4), which is an extremely rare event with our choice of parameters. This is why the Amerasian and Eurasian options have practically the same value when $n = 2$.

Table 3: Prices of Amerasian and Eurasian call options

	p					
(K, T, σ)	1	2	4	13	26	52
(100, 0.25, 0.15)	3.635	2.842	2.512	2.321	2.289	2.276
	3.635	2.842	2.443	2.165	2.103	2.072
(100, 0.25, 0.25)	5.598	4.395	3.920	3.650	3.608	3.592
	5.598	4.395	3.788	3.364	3.270	3.222
(100, 0.50, 0.25)	8.260	6.463	5.745	5.332	5.266	5.239
	8.260	6.462	5.558	4.927	4.787	4.716
(105, 0.50, 0.25)	5.988	4.245	3.475	2.966	2.858	2.804
	5.988	4.245	3.389	2.806	2.678	2.614

2.6.2 Example 2

Our second example is the one considered by Grant, Vora, and Weeks (1997), referred here by GVW. The time increment is fixed at $h = 1/365$ (one day), the first observation date is at $t_1 = 91/365$ (91 days), and the first exercise opportunity is at $t_{m^\bullet} = 105/365$ (105 days). The other parameters are: $S = 100$, $K = 100$, $T = 120/365$, $\sigma = 0.20$, and $r = 0.09$. Table 4 gives the DP approximation for the Amerasian option with different grid sizes with $d = 2$ (the linear-quadratic approximation). In the second line of each entry (the numbers in parentheses), we give the extrapolation value obtained with the ϵ-algorithm with $l = 3$, based on the DP approximations with $p/4$, $p/2$, and p. For example, the ϵ-algorithm with $p = 300$ is based on the DP approximations obtained with $p = 75$ (not shown here), 150, and 300. The last column, labeled GVW, gives the 95% confidence intervals reported by GVW for the values of some Amerasian call options.

Table 4: Prices of Amerasian call options for the GVW example

	p					
(K, σ)	150	300	600	1200	2400	GVW
$(100, 0.2)$	5.80859	5.80274	5.80059	5.79976	5.79943	5.80 ± 0.02
		(5.799)	(5.799)	(5.799)	(5.799)	
$(105, 0.2)$	3.35848	3.35284	3.35075	3.34994	3.34962	3.35 ± 0.02
		(3.349)	(3.350)	(3.349)	(3.349)	
$(100, 0.3)$	7.97369	7.96367	7.95988	7.95838	7.95776	7.92 ± 0.02
		(7.957)	(7.958)	(7.957)	(7.957)	
$(105, 0.3)$	5.57682	5.56705	5.56329	5.56185	5.56124	5.53 ± 0.02
		(5.560)	(5.561)	(5.561)	(5.561)	
CPU (h:mn:s)	0:00:01	0:00:08	0:01:09	0:08:53	1:08:40	

Our values are similar to those of GVW, except for the two last cases. This could be explained by the fact that GVW systematically underestimate the options' values. Indeed, their exercise strategy is suboptimal and, consequently, their price estimator has a negative bias. Precisely, this negative bias is introduced when they assume that the exercise frontier at each stage is determined by two straight lines. Notice that the ϵ-algorithm gives steady predictions as the grid becomes finer.

2.6.3 Example 3

We tried our method with the example given in Tables 3 and 4 of Zvan, Forsyth, and Vetzal (1999), referred here as ZFV. The parameters of the Amerasian call options are: $n = 250$, $S = K = 100$, $T = 0.25$, $r = 0.1$, and $\sigma = 0.2$ and 0.4. Averaging is between the origin and the maturity of the option. Here, DP is performed with the linear-quadratic interpolation ($d = 2$). CPU times for the DP procedure are for $\sigma = 0.2$; those for $\sigma = 0.4$ are similar.

In Table 5, we report the values of Eurasian call options computed by DP, and compare them with those obtained by ZFV and by efficient Monte Carlo with 10^6 replications (as in Table 1).

Table 5: Prices of Eurasian call options for the ZFV example

	DP					FD		
	σ				CPU	σ		CPU
p	0.2		0.4		h:mn:s	0.2	0.4	h:mn:s
145	2.988	ϵ-alg	5.278	ϵ-alg	0:00:07	2.929	5.160	0:02:42
289	2.942	(2.928)	5.187	(5.159)	0:00:51	2.929	5.161	0:11:50
600	2.932	(2.929)	5.167	(5.161)	0:07:32			
1200	2.930	(2.930)	5.163	(5.162)	0:58:14			
2400	2.930	(2.930)	5.162	(5.162)	7:23:37			

The prices of the Eurasian call options obtained by DP and by FD converge to the same values. The CPU times reported here are not comparable: We used a 750 Mhz computer whereas ZFV used a 200 mhz one, and we probably used a different compiler as well.

Table 6: Prices of Amerasian call options for the ZFV example

	DP					FD		
	σ				CPU	σ		CPU
p	0.2		0.4		h:mn:s	0.2	0.4	h:mn:s
145	3.267	ϵ-alg	5.919	ϵ-alg	0:00:04	3.213	5.828	0:02:48
289	3.217	(3.201)	5.830	(5.799)	0:00:30	3.213	5.825	0:11:54
600	3.206	(3.203)	5.807	(5.801)	0:04:05			
1200	3.204	(3.203)	5.803	(5.801)	0:34:24			
2400	3.203	(3.203)	5.802	(5.801)	4:21:10			

In all cases, the difference between the prices of the Amerasian call options obtained by DP and by FD are less than 0.1% of the initial price. This makes us confident about the DP algorithm, though its convergence is more effective in the case of Bermudan Asian options (see Table 1).

2.7 Conclusion

Numerical methods are required for pricing American-style options: PDEs, the lattice based approach, and the DP formulation are various alternatives. In this paper, we showed how to price an Amerasian call option on a single asset, under the Black and Scholes' assumptions, via DP coupled with a piecewise-polynomial approximation of the value function after an appropriate change of variable. We also used the DP formulation and proved continuity, monotonicity, and convexity properties of the value function and of the optimal exercise frontier. These properties have been used to improve the speed of the DP procedure. This procedure gives an approximation of the value function at each observation date for each state. Thus, it may be used to compute the sensitivity coefficients of the option with no extra cost.

DP shows convergence, consistency, and efficiency. If the exercise opportunities are limited, the DP procedure is particularly efficient because it does not need a time discretization to be implemented.

The computational approach does not rely on the form of the exercise region, and could be adapted for pricing other types of discretely-sampled American-style options for which the relevant information process can be modeled as a Markov process. For the case considered in this paper, the Markov process is $\{(S(t), \overline{S}(t)), \text{for } 0 \leq t \leq T\}$.

Here, we have used piecewise polynomials, with the pieces determined by a rectangular grid that remains the same at all steps. Adapting the grid to the shape of the value function at each step (with the same number of pieces) could provide a better approximation but would bring additional overhead, so it would not necessarily be an improvement.

Perhaps a good compromise would be to readjust the grid every δ steps (say), for some integer δ, and readjust it only in the areas just below the optimal exercise frontier, where the value function is significantly nonlinear.

The GBM assumption could be replaced by the CEV process (Cox 1996). The coefficients $D_{k,i}$ in (2.42) can still be computed in closed-form. Other dynamics for the underlying asset may be assumed, but explicit computation may no longer be feasible.

It may be useful to study, for each case of practical interest, how to exploit the structure of the problem to characterize the value function and the optimal exercise strategy, and to improve the efficiency of the numerical method, as we have done here. When the dimension of the state space is large, e.g., if the payoff depends on several underlying assets, approximating the value function becomes generally much more difficult, (we hit the "curse of dimensionality") and pricing the option then remains a challenging problem. In this case, Monte Carlo or quasi-Monte Carlo simulation may be used to solve the DP equation, but this may alter the efficiency of the DP procedure. Improvements may be realized via variance reduction techniques.

2.8 References

Alziary, B., J.P. Décamps, and P.F. Koehl, "A P.D.E. Approach to Asian Options: Analytical and Numerical Evidence," *Journal of Banking and Finance*, 21 (1997), 613–640.

Barraquand, J. and D. Martineau, "Numerical Valuation of High-Dimensional Multivariate American Securities," Journal of Financial and Quantitative Analysis, 30 (1995), 383–405.

Barraquand, J. and T. Pudet, "Pricing of American Path-Dependent Contingent Claims," *Mathematical Finance*, 6 (1996), 17–51.

Bertsekas, D.P., *Dynamic Programming and Optimal Control*, Volumes I and II, Athena Scientific, Belmont, Mass., 1995.

Billingsley, P., *Probability and Measure*, Second Edition, Wiley, New York, 1986.

Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (1973), 637–654.

Bouaziz, L., E. Briys, and M. Crouhy, "The Pricing of Forward-Starting Asian Options," *Journal of Banking and Finance*, 18 (1994), 823–839.

Boyle, P., M. Broadie, and P. Glasserman, "Monte Carlo methods for Security Pricing," *Journal of Economic Dynamics and Control*, 21 (1997), 1267–1321.

Boyle, P. and Y. Tian, "Pricing Lookback and Barrier Options under the CEV Process," *Journal of Financial and Quantitative Analysis*, 34 (1999), 241–264.

Brézinski, C., *Algorithmes d'Accélération de la Convergence, Étude Numérique*, Éditions Technip, Paris, 1978.

Broadie, M. and P. Glasserman, "Pricing American-Style Securities Using Simulation," *Journal of Economic Dynamics and Control*, 21 (1997a), 1323–1352.

Broadie, M. and P. Glasserman, "A Stochastic Mesh Method for Pricing High-Dimensional Americal Options," 1997b, Manuscript, Columbia University.

Carverhill, A.P. and L.J. Clewlow, "Flexible Convolution," *Risk*, 3 (1990), 25–29.

Conte, S.D. and C. de Boor, *Elementary Numerical Analysis*, Third Edition, McGraw-Hill, New York, 1980.

Conze, A. and Viswanathan, "Path Dependent Options: The Case of Lookback Options," *Journal of Finance*, 46 (1991), 1893–1906.

Cox, J.C., "The Constant Elasticity of Variance Option Pricing Model," *Journal of Portfolio Management*, 22 (1996), 15–17.

Cox, J.C., S.A. Ross, and M. Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (1979), 229–263.

Curran, M., "Valuing Asian and Portfolio Options by Conditioning on the Geometric Mean Price," *Management Science*, 40 (1994), 1705–1711.

de Boor, C., *A Practical Guide to Splines*, number 27 in Applied Mathematical Sciences Series, Springer-Verlag, New York, 1978.

Dempster, M.A.H., J.P. Hutton, and D.G. Richards, "LP Valuation of Exotic American Options Exploiting Structure," *Journal of Computational Finance*, 2 (1998), 61-84.

Geman, H. and M. Yor, "Bessel Processes, Asian Options, and Perpetuities," *Mathematical Finance*, 3 (1993), 349–375.

Glasserman, P., P. Heidelberger, and P. Shahabuddin, "Asymptotically Optimal Importance Sampling and Stratification for Pricing Path Dependent Options," *Mathematical Finance*, 9 (1999), 117–152.

Grant, D., G. Vora, and D. Weeks, "Path-Dependent Options: Extending the Monte Carlo Simulation Approach," *Management Science*, 43 (1997), 1589–1602.

Haurie, A. and P. L'Ecuyer, "Approximation and Bounds in Discrete Event Dynamic Programming," *IEEE Trans. on Automatic Control*, AC-31, 3 (1986), 227–235.

Hull, J., *Options, Futures, and Other Derivative Securities*, Fourth Edition, Prentice-Hall, New Jersey, 2000.

Hull, J. and A. White, "Efficient Procedures for Valuing European and American Path-Dependent Options," Journal of Derivatives, 1 (1993), 21–31.

Karatzas, I. and S.E. Shreve, *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.

Kemna, A. G.Z. and A.C.F. Vorst, "A Pricing Method for Options Based on Average Asset Values," *Journal of Banking and Finance*, 14 (1990), 113–129.

L'Écuyer, P. and C. Lemieux, "Variance Reduction via Lattice Rules," *Management Science*, 9 (2000), 1214-1235.

L'Écuyer, P. and J. Malenfant, "Computing Optimal Checkpointing Strategies for Rollback and Recovery Systems," *IEEE Transactions on Computers*, 37, 4 (1988), 491–496.

Lemieux, C. and P. L'Écuyer, "Efficiency Improvement by Lattice Rules for Pricing Asian Options," In *Proceedings of the 1998 Winter Simulation Conference*, ed. D.J. Medeiros, E.F. Watson, J.S. Carson, and M.S. Manivannan, IEEE Press, 1998, 579–586.

Lemieux, C. and P. L'Écuyer, "Selection Criteria for Lattice Rules and Other Low-Discrepancy Point Sets," *Mathematics and Computers in Simulation* (selected papers from the 1999 IMACS Seminar on Monte Carlo Methods), 1-3 (2001), 139–148.

Levy, E., "Pricing European Average Rate Currency Options," *Journal of International Money and Finance*, 11 (1992), 474–491.

Longstaff, F.A. and E.S. Schwartz, "Valuing American Options by Simulation: A Simple Least-Squares Approach," *Review of Financial Studies*, 14 (2001), 113-147.

Ritchken, P., L. Sankarasubramanian, and A.M. Vijh, "The Valuation of Path-Dependent Contracts on the Average," *Management Science*, 39 (1993), 1202–1213.

Rockafellar, R.T., *Convex Analysis*, Princeton University Press, New Jersey, 1970.

Rogers, L. C.G. and Z. Shi, "The Value of an Asian Option," *Journal of Applied Probability*, 32 (1995), 1077–1088.

Turnbull, S.M. and L.M. Wakeman, "A Quick Algorithm for Pricing European Average Options," *Journal of Financial and Quantitative Analysis*, 26 (1991), 377–389.

Vorst, T.C.F., *Averaging Options*, In the Handbook of Exotic Options, ed. I. Nelken, Irwin Professional Publishing, Chicago, 1998, 175-199.

Wilmott, P., J. Dewynne, and J. Howison, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, United Kingdom, 1993.

Wu, R. and M.C. Fu, "Optimal Exercise Policies and Simulation-Based Valuation for American-Asian Options," 2000, Manuscript, University of Maryland.

Zhang, P.G., "Flexible Arithmetic Asian Options," *Journal of Derivatives*, 2 (1995), 53–63.

Zvan, R., P.A. Forsyth, and K.R. Vetzal, "Robust Numerical methods for PDE Models of Asian Options," *Journal of Computational Finance*, 1 (1998), 39–78.

Zvan, R., P.A. Forsyth, and K.R. Vetzal, "Discrete Asian Barrier Options," *Journal of Computational Finance*, 3 (1999), 41–67.

Chapter 3
Pricing Call and Put Options Embedded in Bonds

3.1 Introduction

A *bond* is a contract which pays to its holder a known amount, called the *face value*, at a known future date, called the *maturity*. A bond may also pay periodically to its holder fixed cash dividends called *coupons*. Otherwise, it is called a *zero-coupon bond*. A bond can be interpreted as a loan with a principal equal to the face value and interest payments equal to the coupons (if any). The borrower is the *issuer* of the bond and the lender, i.e., the holder of the bond, is the *investor*. For a description of fixed income securities in general and of bonds in particular, we refer to Fabozzi (1997).

Several bonds contain one or several options coming in various flavors. Firstly, the issuer of the bond may have the right to purchase back its debt for a known amount, called the *call price*, during a specified period within the bond's life. This is the *call option*. Several government bonds contain a call feature [see Bliss and Ronn (1995) for the history of callable U.S. Treasury bonds from 1917]. Secondly, the investor may have the right to return the bond to the issuer for a known amount, called the *put price*, during a specified period within the bond's life. This is the *put option*. See for example Brennan and Schwartz (1977). In this way, a *savings bond* can be redeemed at any time before its maturity and a *retractable bond* can be redeemed only at a specified date before its maturity. Similarly,

the maturity of an *extensible bond* can be extended to a longer period [Ananthanarayanan and Schwartz (1980) and Longstaff (1990)].

An investor may also have the right to exchange the bond for a given number of an underlying asset during a specified period within the bond's life. In general, this asset is the stock of the bond issuer (a firm in this case). This is called a *conversion option* [Brennan and Schwartz (1980)]. Corporate bonds often contain several embedded options, e.g., the Liquid Yield Option Note (LYON), a product developed by Merrill Lynch Capital Markets in 1985, is a zero-coupon bond that is callable, putable, and convertible.

All the options described above are an integral part of a bond, and cannot be traded alone as is the case for call and put options on stocks (for example). They are said to be *embedded* in the bond. In general, they are of the American-type, so that the bond with its embedded options can be interpreted as an American-style financial derivative with (possibly) a protection period against early exercising.

Other options embedded in bonds do exist. We give here some examples from the Future contract on long-maturity U.S. Treasury bonds traded on the Chicago Board of Trade (CBOT). This Future contract is an agreement to sell or to buy some U.S. Treasury bonds with a total face value of $100,000 at a certain date in the future, called the *delivery date*, for a certain price, called the *delivery price*. This contract contains several embedded options discussed by Boyle (1989) and Cohen (1991) among others. The first, called the *timing option*, gives the seller the right to deliver the underlying Treasury bonds at any time during a specified delivery month. The second, called the *quality option*, gives the seller the right to deliver any U.S. Treasury bond with at least 15 years to the earliest call date or

to maturity at the first delivery date. The third, called the *wild card option*, comes from the time difference between the closing hours of the cash and the Future markets. Thus, from 2:00 PM (Chicago time), when the CBOT closes, the seller has until 8:00 PM to decide whether to deliver or not. Bond markets remain open until 4:00 PM.

There are no analytical formulas for valuing American options, even under very simplified assumptions. Numerical methods, essentially trees and finite-differences (FD), are usually used for pricing using a backward induction framework. Recall that trees are particular discrete-time models and FD are numerical solution methods for Partial Differential Equations (PDE). As an alternative approach, the pricing of Bermudan American financial derivatives can be formulated as a Markov Decision process, i.e., a stochastic *Dynamic Programming* (DP) problem as pointed out by Barraquand and Martineau (1995). Here, the DP function, i.e., the value of the bond with its embedded options, is a function of the current time and of the current interest rate, namely the *state variables*. The set of all the (possible) realizations of the state variables defines the *state space*. This value function verifies a DP recurrence (known as the DP or the *Bellman* equation) via the *risk-neutral* principle of asset pricing. Indeed, the DP equation relates the holding value of the bond at the current time as an expectation, under the so-called *risk-neutral* probability measure, of its future value discounted at the risk-free interest rate. The key point with DP is to solve efficiently the DP equation which yields both the bond value and the optimal exercise strategies of its embedded options. For an overview of stochastic DP, we refer to Bertsekas (1987), and for risk-neutral evaluation, we refer to Karatzas and Shreve (1998).

The problem of pricing options embedded in bonds is related to the term structure of interest rates. In this context, the short term risk-free interest rate is very often used as a Markov process in arbitrage-free markets, but zero-coupon bonds and forward rates are also used. Several approaches are suggested in the literature depending on the nature of the underlying asset(s), the dimension of the state space, and the frequency with which the state variables are observed. For example, a one-factor model may be used for the interest rate, a two-factor model may be used for the interest rate and its random volatility, or for the interest rate and a risky asset (say a stock), and so on. The underlying asset(s) may be observed discretely or continuously.

For discrete-time models, the discrete interest rate is the most widely used, through a binomial tree in the spirit of Cox, Ross, and Rubinstein (1979). These include models by Ho and Lee (1986), Black, Derman, and Toy (1990), and Kalotay, Williams, and Fabozzi (1993). All these models are calibrated to exactly mimic the initial yield curve.

For continuous-time models, the short-term risk-free interest rate is very often modeled as a diffusion process. As pointed out by Chan et al. (1992), most of the alternative dynamics for the interest rate are described by the general stochastic differential equation (SDE)

$$dr_t = (\alpha + \beta r_t)\, dt + \sigma r_t^{\gamma} dB_t, \quad \text{for } 0 \leq t \leq T, \tag{3.47}$$

where α, β, σ, and γ are real parameters and B_t, for $0 \leq t \leq T$, is a standard Brownian motion. Table 1 presents various versions of equation (3.47) used in the literature.

Table 1: Models for the Short-Term Interest Rate

Model	α	β	σ	γ
1. Merton (1973)		0		0
2. Vasicek (1977)				0
3. Brennan-Schwartz (1977)	0	0		1
4. Brennan-Schwartz (1980)				1
5. Marsh-Rosenfeld (1983)	0			
6. Cox-Ingersoll-Ross (1985)				1/2

The model by Merton (1973) is simply a standart Brownian motion with a drift. Vasicek (1977) uses a mean-reverting Ornstein-Uhlenbeck process. This model gives nice distributional results and ensures explicit formulas for zero-coupon bonds and for several European-style interest rate derivatives, but it has the undesirable property of allowing negative interest rates (though with very low probabilities). Several authors take advantage of its properties to price various interest rate derivatives (often in closed-form), e.g., Jamshidian (1989) and Rabinovitch (1989). Brennan and Schwartz (1977, 1980) are pioneers on the modeling of options embedded in bonds. They use the PDE approach and FD algorithms. They let the interest rate move as a geometric Brownian motion (GBM) without a drift to price the call and the put options (Model 3) and as a mean-reverting proportional process to price the conversion option (Model 4). Model 3 has also been used by Dothan (1978) to price bonds in closed-form, and Model 4 by Courtadon (1982) to price several European as well as American options on bonds. Notice that Model 3 is a special case of Model 4 and that the latter includes the GBM process of Black and Scholes (1973). Marsh and Rosenfeld (1983) use the constant elasticity of variance (CEV) process (among others), introduced previously for options pricing [see Cox (1996)]. Cox, Ingersoll, and Ross (CIR) (1985) use the mean-reverting square-root process (MRSR) to handle the interest rate movements. Their basic one-factor formulation is based on an equilib-

rium concept, known to preclude against arbitrage opportunities. It is extendible to several factors, ensures strictly positive interest rates, and gives explicit formulas for zero-coupon bonds and for some European-style interest rate derivatives. Other authors use the MRSR equation to price various interest rate derivatives, e.g., Richard (1978), Ananthanarayanan and Schwartz (1980), and Schaefer and Schwartz (1984). Other dynamics for interest rates can also be found, e.g., Brennan and Schwartz (1979) and Clewlow and Strickland (1997).

The underlying assets may also be zero-coupon bonds, e.g., Briys, Crouhy, and Schöbel (1991), or forward rates, e.g., Heath, Jarrow, and Morton (HJM) (1992). The HJM model is one of the most flexible as it integrates several systematic sources of risk and random coefficients.

Continuous-time models are parsimonious, so that matching all theoretical bond values with their market counterparts gives much more equations than the number of parameters to estimate. A remedy, proposed by Hull and White (1990a), is to augment the model, i.e., add parameters until a calibration becomes possible. This leads to the extended Vasicek model and to the extended CIR model. See the note by Carverhill (1995) and the response by Hull and White (1995) for a discussion about the performance of their "extended" models. Hull and White (1990b, 1993, 1994a, 1994b, 1996) interpret the FD method as a trinomial tree and price several interest rate derivatives within their "extended" models.

In this paper, we formulate the problem of pricing options embedded in bonds as a stochastic DP model, the focus being on the solution of the DP equation. Precisely, we consider call and put options. It is well known that the call option tends to decrease the price of the bond, while the put option has the opposite effect. It is also well known that the

call option is more likely exercised by the bond issuer when the interest rates are low, while the put option is more likely exercised by the investor when the interest rates are high. We use the DP formulation to establish these basic intuitions, to evaluate the impact of these options on the bond price, and to determine their optimal exercise strategies.

We adopt the Vasicek (1977) model for the short-term risk-free interest rate. By the *finite-elements* technique, we approximate the DP function by a piecewise linear interpolation, and, subsequently, we solve the DP equation in closed-form. The PDE approach and FD algorithms could but are not used in this context [Wilmott, Dewynne, and Howison (1997)]. Indeed, the DP formulation and the finite elements technique is a viable alternative to the PDE approach and FD algorithms if the joint distribution of the state variables is known explicitly and the number of exercise opportunities is limited [see Chapter 2].

In Section 2, we present the Vasicek formulation and its distributional properties. In Section 3, we give the DP formulation. In Sections 4 and 5, we derive some theoretical properties of the bond value and solve the DP equation. In Section 6, we give some results. In Section 7, we conclude.

3.2 The Model

Vasicek (1977) introduced a continuous-time, finite horizon, and frictionless market for the short-term risk-free interest rate in which *risk-neutral* evaluation is possible. We present this model directly under the risk-neutral probability measure, denoted by Q, whose existence is guaranteed by the no-arbitrage property.

The risk-free interest rate moves under Q according to the SDE

$$dr_t = \kappa\left(\overline{r} - r_t\right)dt + \sigma dB\left(t\right), \quad \text{for } 0 \leq t \leq T, \tag{3.48}$$

where κ, $\overline{r}$, and σ are real positive constants and $B\left(t\right)$, for $0 \leq t \leq T$, is a standard Brownian motion whose augmented natural filtration is denoted by $\mathcal{F}\left(t\right)$, for $0 \leq t \leq T$. Over time, the interest rate process is pushed towards its *reverting level* $\overline{r}$ at the *reverting rate* κ, and these random reverting cycles are more or less amplified depending on the *volatility* parameter σ. We show later that the standard error of the future interest rate depends on σ.

In this model, a financial derivative can be priced as an expectation under Q of its future payoff discounted at the risk-free rate. This is the fundamental risk-neutral principle of asset pricing. Of course, bonds may be priced in that way. For example, for $0 \leq t' \leq t \leq t'' \leq T$, the rational price at t' of an optionless zero-coupon bond paying 1 at T is

$$\begin{aligned} v_{t'}\left(r\right) &= E\left[e^{-\int_{t'}^{t''} r_t dt} v_{t''}\left(r_{t''}\right) \mid \mathcal{F}\left(t'\right), r_{t'} = r\right] \\ &= E\left[e^{-\int_{t'}^{t''} r_t dt} v_{t''}\left(r_{t''}\right) \mid r_{t'} = r\right], \end{aligned} \tag{3.49}$$

where $v_{t''}$ is the value of the bond at t''. Equation (3.49) at $t' = 0$ and $t'' = T$ gives

$$v_0\left(r\right) = E\left[e^{-\int_0^T r_t dt}\right],$$

where r is the interest rate at time 0.

In the Vasicek (1977) model, it is well known that the random variables $r_{t''}$ and $\int_{t'}^{t''} r_u du$ are normal [see Elliott and Kopp (1999) for example]. In the following, we extend this result to the random vector $\left(r_{t''}, \int_{t'}^{t''} r_u du\right)'$.

Lemma 5 *For f and g two real functions continuously differentiable in $[t', t'']$, for $0 \leq t' \leq t'' \leq T$, and $W(t)$, for $t \in [0, T]$, a standard Brownian motion, one has*

$$\begin{aligned}
&\int_{t'}^{t''} \left(\int_{t'}^{u} f(t)\, g(u)\, dW(t) \right) du \\
&= \int_{t'}^{t''} \left(\int_{t'}^{t''} f(t)\, g(u)\, I(t \in [t', u])\, dW(t) \right) du \\
&= \int_{t'}^{t''} \left(\int_{t'}^{t''} f(t)\, g(u)\, I(u \in [t, t'])\, du \right) dW(t) \\
&= \int_{t'}^{t''} \left(\int_{t}^{t''} f(t)\, g(u)\, du \right) dW(t).
\end{aligned}$$

Proof. We use the integration by parts theorem in stochastic calculus [Øksendal (1995)]. We define the function $h(t) = f(t) \int_t^{t''} g(u)\, du$ and transform the right hand integral as

$$\begin{aligned}
\int_{t'}^{t''} \left(\int_{t}^{t''} f(t)\, g(u)\, du \right) dW(t) &= \int_{t'}^{t''} h(t)\, dW(t) \\
&= h(t'')\, W(t'') - h(t')\, W(t') - \\
&\quad \int_{t'}^{t''} \frac{\partial h}{\partial t}(t)\, W(t)\, dt \\
&= -\int_{t'}^{t''} f(t')\, g(u)\, W(t')\, du \\
&\quad + \int_{t'}^{t''} f(t)\, g(t)\, W(t)\, dt \\
&\quad - \int_{t'}^{t''} \left(\int_{t}^{t''} \frac{\partial f}{\partial t}(t)\, g(u)\, W(t)\, du \right) dt.
\end{aligned}$$

We now use the same theorem to transform $\int_{t'}^{u} f(t)\, g(u)\, dW(t)$ and thereafter the left hand integral as

$$\begin{aligned}\int_{t'}^{t''}\left(\int_{t'}^{u} f(t)\, g(u)\, dW(t)\right) du &= \int_{t'}^{t''} f(u)\, g(u)\, W(u)\, du - \\ &\int_{t'}^{t''} f(t')\, g(u)\, W(t')\, du - \\ &\int_{t'}^{t''}\left(\int_{t'}^{u} \frac{\partial f}{\partial t}(t)\, g(u)\, W(t)\, dt\right) du.\end{aligned}$$

The final result comes from the basic properties of multi-dimensional real integrals. ■

Proposition 6 *Conditioning on the information available at time t', that is, for $r_{t'} = r$, the random vector*

$$\left(r_{t''},\ \int_{t'}^{t''} r_u du\right)',\quad 0 \leq t' \leq t'' \leq T.$$

is a normal vector with mean

$$\mu(r) = \begin{bmatrix} \mu_1(r) = \overline{r} + e^{-\kappa\Delta t}(r - \overline{r}) \\ \mu_2(r) = \overline{r}\Delta t + \frac{1-e^{-\kappa\Delta t}}{\kappa}(r - \overline{r}) \end{bmatrix}$$

and variance

$$\Sigma = \begin{bmatrix} \sigma_1^2 = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\Delta t}\right) & \sigma_{12} = \frac{\sigma^2}{2\kappa^2}\left(1 - 2e^{-\kappa\Delta t} + e^{-2\kappa\Delta t}\right) \\ \sigma_{21} = \sigma_{12} & \sigma_2^2 = \frac{\sigma^2}{2\kappa^3}\left(-3 + 2\kappa\Delta t + 4e^{-\kappa\Delta t} - e^{-2\kappa\Delta t}\right) \end{bmatrix},$$

where $\Delta t = t'' - t'$. It admits the following decomposition

$$\begin{cases} r_{t''} = \mu_1(r) + \sigma_1 Z_1 \\ \int_{t'}^{t''} r_u du = \mu_2(r) + \sigma_2\left[\rho Z_1 + \sqrt{1-\rho^2} Z_2\right] \end{cases},$$

where $(Z_1, Z_2)'$ is a standard normal vector and $\rho = \sigma_{12}/(\sigma_1\sigma_2)$ is the (conditional) coefficient of correlation between $r_{t''}$ and $\int_{t'}^{t''} r_u du$.

Proof. From equation (3.48), one can apply Ito's lemma to the process $\phi(t, r_t) = e^{\kappa t} r_t$, for $t \in [0, T]$, and show that

$$r_u = \overline{r} + e^{-\kappa(u-t')}(r - \overline{r}) + \sigma \int_{t'}^{u} e^{-\kappa(u-t)} dB(t),$$

and consequently that

$$\begin{aligned} \int_{t'}^{t''} r_u du &= \overline{r}\Delta t + \frac{1 - e^{-\kappa\Delta t}}{\kappa}(r - \overline{r}) + \sigma \int_{t'}^{t''} \left(\int_{t'}^{u} e^{-\kappa(u-t)} dB(t) \right) du \\ &= \overline{r}\Delta t + \frac{1 - e^{-\kappa\Delta t}}{\kappa}(r - \overline{r}) + \sigma \int_{t'}^{t''} \left(\int_{t}^{t''} e^{-\kappa(u-t)} du \right) dB(t). \end{aligned}$$

The last equality comes from the Lemma. Conditioning on the information available at time $t' \in [0, T]$, we can decompose each component of the vector $\left(r_{t''}, \int_{t'}^{t''} r_u du \right)'$ into a deterministic part and a random part. The latter part turns out to be a limit of linear combinations of the same standard Brownian motion taken at different points in time. We conclude that the random variables $r_{t''}$ and $\int_{t'}^{t''} r_u du$, conditioned on $F(t')$, are jointly normal.

Now, from basic properties of stochastic integrals [Øksendal (1995)], one can derive the conditional mean and the conditional variance of the vector $\left(r_{t''}, \int_{t'}^{t''} r_u du \right)'$. Its conditional mean is

$$\begin{aligned} & E\left[\left(r_{t''}, \int_{t'}^{t''} r_u du \right)' \mid F(t') \right] \\ = & \left(\overline{r} + e^{-\kappa\Delta t}(r - \overline{r}), \overline{r}\Delta t + \frac{1 - e^{-\kappa\Delta t}}{\kappa}(r - \overline{r}) \right)', \end{aligned}$$

since the centered random vector

$$\left(\int_{t'}^{t''} e^{-\kappa(u-t)} dB(t), \int_{t'}^{t''} \left(\int_{t}^{t''} e^{-\kappa(u-t)} du \right) dB(t) \right)',$$

is independent of $F(t')$. The conditional variance of $r_{t''}$ is

$$\begin{aligned}\text{Var}\left[r_{t''} \mid F(t')\right] &= E\left[\left(\int_{t'}^{t''} \sigma e^{-\kappa(t''-t)} dB(t)\right)^2 \mid F(t')\right] \\ &= \sigma^2 \int_{t'}^{t''} e^{-2\kappa(t''-t)} dt \\ &= \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\Delta t}\right).\end{aligned}$$

The conditional variance of $\int_{t'}^{t''} r_u du$ is

$$\begin{aligned}&\text{Var}\left[\int_{t'}^{t''} r_u du \mid F(t')\right] \\ &= E\left[\left(\int_{t'}^{t''}\left(\int_{t}^{t''} \sigma e^{-\kappa(u-t)} du\right) dB(t)\right)^2 \mid F(t')\right] \\ &= \sigma^2 \int_{t'}^{t''}\left(\int_{t}^{t''} e^{-\kappa(u-t)} du\right)^2 dt \\ &= \frac{\sigma^2}{2\kappa^3}\left(-3 + 2\kappa\Delta t + 4e^{-\kappa\Delta t} - e^{-2\kappa\Delta t}\right).\end{aligned}$$

The conditional covariance between $r_{t''}$ and $\int_{t'}^{t''} r_u du$ is

$$\begin{aligned}&\text{Cov}\left[r_{t''}, \int_{t'}^{t''} r_u du \mid F(t')\right] \\ &= E\left[\int_{t'}^{t''} \sigma e^{-\kappa(t''-t)} dB(t) \int_{t'}^{t''}\left(\int_{t}^{t''} \sigma e^{-\kappa(u-t)} du\right) dB(t) \mid F(t')\right] \\ &= \sigma^2 \int_{t'}^{t''} e^{-\kappa(t''-t)}\left(\int_{t}^{t''} e^{-\kappa(u-t)} du\right) dt \\ &= \frac{\sigma^2}{2\kappa^2}\left(1 - 2e^{-\kappa\Delta t} + e^{-2\kappa\Delta t}\right).\end{aligned}$$

■

3.3 The DP Formulation

In this section, we present the DP function and the DP equation for a zero-coupon bond with its embedded call and put options. We essentially use the risk-neutral principle of asset pricing to assess the DP formulation.

Let $t_0, \ldots, t_M$ be a sequence of dates such that $\Delta t = t_{m+1} - t_m$, for $m = 0, \ldots, M-1$, where $t_0 = 0$ is the origin and $t_M = T$ is the maturity of the bond. We assume that the exercise opportunities of the embedded options are at t_m, for $m = 1, \ldots, M$. In practice, the first increment of time $t_1 - t_0$ and the last one $t_M - t_{M-1}$ may be different from Δt.

Let c_m and p_m be the call and the put prices at t_m, respectively. If the issuer calls back the bond at t_m, he pays a known amount c_m to the investor, and, at the same date, if the investor puts the bond, he receives a known amount p_m from the issuer. Assume that the call and put prices verify $0 < p_m \leq c_m$, as is usual in practice, and $c_M = p_M = 1$. Finally, let $v_m(r)$ be the value of the bond and $v_m^h(r)$ its holding value at time t_m, where r is the interest rate at that time.

At the maturity date $t_M = T$, the value of the bond is

$$v_M(r) = 1, \quad \text{for all } r, \tag{3.50}$$

where 1 is the face value of the bond with its embedded options. Table 2 gives the "payoff" at time t_m by the issuer to the investor under decision pairs.

Table 2: The bond payoff under decision pairs

	Investor	
Issuer	Put	Hold
Call	p_m	c_m
Not to Call	p_m	–

From Table 2, we can specify the optimal strategies of the two agents at each monitoring date. The issuer has a dominating strategy; he will call the bond if $v_m^h(r) > c_m$, otherwise, he is better not to call. In the first case, the investor will put the bond if $p_m > c_m$ (a possibility that we exclude). In the second case, the investor will put the bond if $v_m^h(r) < p_m$. Therefore, for $m = 1, \ldots, M-1$, one has

$$v_m(r) = \begin{cases} c_m & \text{if } v_m^h(r) > c_m \\ v_m^h(r) & \text{if } p_m \leq v_m^h(r) \leq c_m \\ p_m & \text{if } v_m^h(r) < p_m \end{cases}. \tag{3.51}$$

By the risk-neutral principle of asset pricing, we obtain the holding value of the bond $v_m^h(r)$ at time t_m from its value v_{m+1} at time t_{m+1} by

$$\begin{aligned} v_m^h(r) &= E\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} v_{m+1}\left(r_{t_{m+1}}\right) \mid \mathcal{F}(t_m), r_{t_m} = r\right] \\ &= E_{m,r}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} v_{m+1}\left(r_{t_{m+1}}\right)\right]. \end{aligned} \tag{3.52}$$

At the origin, the bond value is a function of the observed interest rate

$$v_0(r) = v_0^h(r), \quad \text{for all } r. \tag{3.53}$$

Equations (3.50-3.53) define the stochastic DP formulation. In particular, equations (3.50), (3.51), and (3.53) define respectively the DP function at maturity, at a given step of the DP procedure, and at the origin. Solving the expectation in (3.52) backwards from the maturity to the origin yields both the initial value of the bond and the optimal exercise strategies of its embedded options. We return to this point later.

3.4 Properties of the Bond's Value

In this section, we study the shape of the value function for a zero-coupon bond with its embedded options. We essentially use the properties of real integrals.

Proposition 7 *For $m = 0, \ldots, M-1$, the holding value of the bond, $v_m^h(r)$, is a strictly positive, continuous, and strictly decreasing function of r. The value function, $v_m(r)$, verifies*

$$v_m(r) = \begin{cases} c_m, & \text{if } r \leq a_m \\ v_m^h(r), & \text{if } a_m < r < b_m \\ p_m, & \text{if } r \geq b_m \end{cases},$$

where a_m and b_m are two thresholds associated with time t_m.

Proof. The proof proceeds by induction on m. For $m = M-1$, equation (3.52) gives

$$\begin{aligned} v_{M-1}^h(r) &= E_{M-1,r}\left[e^{-\int_{t_{M-1}}^{T} r_t dt}\right] \\ &= e^{-\mu_2(r)+\sigma_2^2/2}, \end{aligned}$$

because by conditioning on the information available at time t_{M-1}, the random variable $e^{-\int_{t_{M-1}}^{T} r_t dt}$ is lognormal with parameters $\mu_2(r)$ and σ_2 [see Proposition 6]. Clearly, the holding value of the bond $v_{M-1}^h(r)$ is a strictly positive, continuous, and strictly decreasing function of r, since $\mu_2(r)$ is a continuous and strictly increasing function of r. The thresholds a_{M-1} and b_{M-1} at step $M-1$ do exist since $v_{M-1}^h(r) \to +\infty$ when $r \to -\infty$ and $v_{M-1}^h(r) \to 0$ when $r \to +\infty$ [see equation (3.51)].

Now, we show that the properties under interest are verified at step m once they are verified at step $m+1$. From equation (3.52), one has

$$\begin{aligned} & v_m^h(r) \\ = \; & E_{m,r}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} v_{m+1}\left(r_{t_{m+1}}\right)\right] \\ = \; & \iint_{\mathbb{R}^2} e^{-\mu_2(r)-\sigma_2\left(\rho z_1+\sqrt{1-\rho^2}z_2\right)} v_{m+1}\left(\mu_1(r)+\sigma_1 z_1\right)\phi(z_1)\,\phi(z_2)\,dz_1 dz_2 \\ = \; & e^{-\mu_2(r)+\sigma_2^2(1-\rho^2)/2}\int_{-\infty}^{+\infty} e^{-\sigma_2\rho z_1} v_{m+1}\left(\mu_1(r)+\sigma_1 z_1\right)\phi(z_1)\,dz_1, \end{aligned}$$

where ϕ is the density function of a standard normal random variable.

The holding value is strictly positive and continuous by the Lebesgue's dominated convergence theorem [Billingsley (1995)]. It is a strictly decreasing function of r since $\mu_1(r)$ and $\mu_2(r)$ are strictly increasing functions of r and $v_{m+1}(r)$ is a strictly positive, bounded, bounded away from zero, and non-increasing function of r. Finally, the thresholds a_m and b_m at step m do exist since $v_m^h(r) \to +\infty$ when $r \to -\infty$ and $v_m^h(r) \to 0$ when $r \to +\infty$. ■

3.5 Solving the DP Equation

In this section, we show how to compute the expectation in (3.52) for each m. The idea here is to partition the real axis into a collection of intervals and then to approximate the bond value by a piecewise linear interpolation, so that computation becomes feasible in closed-form.

Let $a_0 = -\infty < a_1 < \ldots < a_p < a_{p+1} = +\infty$ be a set of points and $R_1, \ldots, R_{p+1}$ be a partition of $\mathbb{R}$ into $(p+1)$ intervals such that $R_1 = (-\infty, a_1)$ and

$$R_i = [a_{i-1}, a_i) \quad \text{for } i = 2, \ldots, p+1.$$

Given an approximation $\widetilde{v}_m$ of the bond value v_m at the points a_k and step m, we interpolate this function by a piecewise linear interpolation of the form

$$\widehat{v}_m(a) = \sum_{i=1}^{p+1} (\alpha_i^m + \beta_i^m a) I_i(a), \tag{3.54}$$

where

$$I_i(a) = \begin{cases} 1 & \text{if } a \in R_i \\ 0 & \text{elsewhere} \end{cases}.$$

Its local coefficients α_i^m and β_i^m are obtained by solving the linear equations

$$\widetilde{v}_m(a_i) = \widehat{v}_m(a_i), \quad \text{for } i = 2, \ldots, p, \tag{3.55}$$

and, for $i \in \{1, p+1\}$, they are identical to those of the adjacent interval. Other piecewise polynomial approximations, such as quadratic and cubic splines, could be used in this context [see de Boor (1978) for a general discussion].

Assume now that $\widehat{v}_{m+1}$ is known, and so are its local coefficients at step $m+1$ as in (3.54). The expectation in (3.52) at step m becomes

$$\begin{aligned} \widetilde{v}_m^h(a_k) &= E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} \widehat{v}_{m+1}\left(r_{t_{m+1}}\right)\right] \\ &= \sum_{i=1}^{p+1} \Big(\alpha_i^{m+1} E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} I_i\left(r_{t_{m+1}}\right)\right] + \\ &\quad \beta_i^{m+1} E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} r_{t_{m+1}} I_i\left(r_{t_{m+1}}\right)\right]\Big), \end{aligned} \tag{3.56}$$

where $\widetilde{v}_m^h$ denotes the approximate holding value of the bond. Now, we use the decomposition of the random vector $\left(r_{t_{m+1}}, \int_{t_m}^{t_{m+1}} r_t dt\right)'$ in Proposition 2 to compute explicitly the integrals in (3.56).

Indeed, for $k = 1, \ldots, p$ and $i = 1, \ldots, p+1$, the first integrals can be expressed as

$$\begin{aligned} A_{k,i} &= E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} I_i\left(r_{t_{m+1}}\right)\right] \\ &= e^{-\mu_2(a_k)+\sigma_2^2/2}\left[\Phi\left(a_{k,i}\right) - \Phi\left(a_{k,i-1}\right)\right], \end{aligned}$$

and the second ones as

$$\begin{aligned} B_{k,i} &= E_{m,a_k}\left[e^{-\int_{t_m}^{t_{m+1}} r_t dt} r_{t_{m+1}} I_i\left(r_{t_{m+1}}\right)\right] \\ &= e^{-\mu_2(a_k)+\sigma_2^2/2} \times \\ &\quad \left[\left(\mu_1\left(a_k\right) - \sigma_{12}\right)\left(\Phi\left(a_{k,i}\right) - \Phi\left(a_{k,i-1}\right)\right) - \sigma_1\left(e^{-a_{k,i}^2} - e^{-a_{k,i-1}^2}\right)/\sqrt{2\pi}\right], \end{aligned}$$

where Φ is the cumulative density function of a standard normal random variable and

$$a_{k,j} = \left(a_j - \mu_1\left(a_k\right) + \sigma_{12}\right)/\sigma_1, \quad \text{for } j \in \{i-1, i\}.$$

In this case, the future interest rate may take negative values [see Proposition 1], but with very low probabilities. We take $a_1 = \mu_1(r) - 6\sigma_1$, $a_2 = \mu_1(r) - 4\sigma_1$, $a_{p-1} = \mu_1(r) + 4\sigma_1$, and $a_p = \mu_1(r) + 6\sigma_1$, where r is the interest rate at time 0 and $\Delta t = T$ the maturity of the bond. The integer p is a parameter to be specified. Then, we select the points a_k, for $k = 2, \ldots, p-1$, to be equally spaced within the interval $[a_2, a_{p-1}]$.

Finally, we explain how the DP procedure works:

FOR $m = M, 0$ by step -1

FOR $k = 1, p$

Compute $\widetilde{v}_m^h(a_k)$ for all the points a_k by (3.56);

Compute $\widetilde{v}_m(a_k)$ for all the points a_k by (3.51);

NEXT k

Compute the coefficients of $\widehat{v}_m$ at step m by (3.55);

NEXT m

3.6 Numerical Experiments

The value of an embedded option is obtained by making the difference between the prices of the bond with and without the option. In this section, we price by DP a call and a put options embedded in a zero-coupon bond. We price first some optionless zero-coupon bonds by DP and compare the results with the exact solution of Vasicek (1977).

Consider a zero-coupon bond paying 1 at the maturity T (in years) and let $r_0 = 4.5\%$, $\kappa = 1, \overline{r} = 5\%$, and $\sigma = 0.01$. The number of monitoring dates is denoted by M. Table 3 gives the values of some optionless zero-coupon bonds computed by DP. Its last column indicates the exact solution of Vasicek (1977). CPU times are given in seconds and they are for the last line (the most expensive). In Table 4, all the parameters values are the same as in Table 3 except for the interest rate at the origin, increased to 5.5%.

Table 3: Optionless Zero-Coupon Bonds ($r_0 = 4.5\%$)

	p			
(T, M)	25	50	100	Vasicek
$(1, 2)$	0.9543	0.9543	0.9543	0.9543
$(2, 4)$	0.9088	0.9088	0.9088	0.9088
$(5, 10)$	0.7828	0.7828	0.7828	0.7828
$(10, 20)$	0.6097	0.6097	0.6097	0.6098
CPU (sec)	0.08	0.37	1.01	

Table 4: Optionless Zero-Coupon Bonds ($r_0 = 5.5\%$)

	p			
(T, M)	25	50	100	Vasicek
$(1, 2)$	0.9482	0.9482	0.9482	0.9482
$(2, 4)$	0.9010	0.9010	0.9010	0.9010
$(5, 10)$	0.7750	0.7750	0.7750	0.7751
$(10, 20)$	0.6037	0.6037	0.6037	0.6038
CPU (sec)	0.09	0.28	1.07	

As expected, the value of a zero-coupon bond decreases when the maturity or the interest rate increases. In view of the results, the DP procedure appears to be stable, consistent, and efficient. CPU times are obtained with an old 100 Mhz Silicon Graphics (and an f77 compiler).

Now, consider the 5 years zero-coupon bond in Table 4 for which we add a call and a put features. The call and put prices are specified in the contract and could be selected arbitrarily. Here, they are determined as in Mason et al. (1995) [see the case "Waste Management, Inc."].

Table 5: The Call and Put Prices

m	t_m (in years)	c_m	p_m
1	0.5	0.83070	0.78914
2	1	0.84734	0.80749
3	1.5	0.86452	0.83040
4	2	0.88223	0.85824
5	2.5	0.90051	0.88039
6	3	0.91935	0.90311
7	3.5	0.92641	0.92641
8	4	0.95032	0.95032
9	4.5	0.97484	0.97484
10	5	1	1

In this example, the bond under interest will be exercised for sure after 3.5 years (if not exercised before). The last column of Table 6 reports the options values. They are obtained by making the difference between the price of the bond with its embedded option(s) and the price of its optionless counterpart (0.7751).

Table 6: Pricing the Embedded Call and Put Options

		p				
(T, M)	Option	25	50	100	Vasicek	Option Value
$(5, 10)$	Call	0.7726	0.7727	0.7727	0.7751	0.0024
$(5, 10)$	Put	0.7767	0.7766	0.7766	0.7751	0.0015
$(5, 10)$	Call+Put	0.7751	0.7751	0.7751	0.7751	0
CPU (sec)		0.09	0.3	1.04		

As expected, the call option tends to decrease the bond value and the put option has the opposite effect. In this case, the inclusion of both the call and the put features has negligible effect on the bond value (no effect for the precision given in this example). They tend to compensate each other.

3.7 Conclusion

American-style financial derivatives do not admit closed-form solutions even under very simplified assumptions. In this paper, we adress the problem of pricing the embedded call and put options via a stochastic DP formulation, the focus being on the solution of the DP equation. We let the interest rate move as in Vasicek (1977), we approximate the bond value by a piecewise linear interpolation, and solve the DP equation in closed-form. We use this formulation to price zero-coupon bonds and their embedded options. Results show that the bond value decreases when the call option is included, and increases when the put option is included. Numerical investigation shows stability, consistency, and efficiency. Also, we use the DP formulation to establish some theoretical properties of the bond value with its embedded options. It is a strictly positive, continuous, and non-increasing function of the interest rate. It is equal to the call price for "low" interest rates and to the put price for "high" interest rates.

If the interest rate moves according to the GBM, the CEV, or the CIR process, the market remains arbitrage-free and risk-neutral evaluation is preserved. In this context, the interest rate is Markov and the stochastic DP formulation could be applied.

3.8 References

Ananthanarayanan, A.L. and E.S. Schwartz, "Retractable and Extendible Bonds: The Canadian Experience," *Journal of Finance*, 35 (1980), 31-47.

Barraquand, J. and D. Martineau, "Numerical Valuation of High Dimensional Multivariate American Securities," *Journal of Financial and Quantitative Analysis*, 30 (1995), 383-405.

Bertsekas, D.P., *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, New Jersey, 1987.

Billingsley, P., *Probability and Measure*, Third Edition, John Wiley and Sons, New York, 1995.

Black, F., E. Derman, and W. Toy, "A One-Factor Model of Interest Rates and its Application to Treasury Bond Options," *Financial Analysts Journal*, 46 (1990), 33-40.

Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (1973), 637-654.

Bliss, R.R. and E.I. Ronn, "To Call or not to Call? Optimal Call Policies for Callable U.S. Treasury Bonds," *Economic Review*, 80 (1995), 1-15.

Boyle, P.P., "The Quality Option and the Timing Option in Futures Contracts," *Journal of Finance*, 44 (1989), 101-113.

Brennan, M.J. and E.S. Schwartz, "Savings Bonds, Retractable Bonds, and Callable Bonds," *Journal of Financial Economics*, 5 (1977), 67-88.

Brennan, M.J. and E.S. Schwartz, "A Continuous Time Approach to the Pricing of Bonds," *Journal of Banking and Finance*, 3 (1979), 133-155.

Brennan, M.J. and E.S. Schwartz, "Analyzing Convertible Bonds," *Journal of Financial and Quantitative Analysis*, 15 (1980), 907-929.

Briys, E., M. Crouhy, and R. Schöbel, "The Pricing of Default-Free Interest Rate Cap, Floor, and Collar Agreements," *Journal of Finance*, 46 (1991), 1879-1892.

Carverhill,A., "A Note on the Models of Hull and White for Pricing Options on the Term Structure," *Journal of Fixed Income*, 5 (1995), 89-96.

Chan, K.C., G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, 47 (1992), 1209-1227.

Clewlow,L. and C. Strickland, "Monte Carlo Valuation of Interest Rate Derivatives under Stochastic Volatility," *Journal of Fixed Income*, 7 (1997), 35-45.

Cohen, H., "Evaluating Embedded Options," *Economic Review*, 76 (1991), 9-17.

Courtadon, G., "The Pricing of Options on Default-Free Bonds," *Journal of Financial and Quantitative Analysis*, 17 (1982), 75-100.

Cox, J.C.,."The Constant Elasticity of Variance Option Pricing Model," *Journal of Portfolio Management*, 22 (1996), 15-17.

Cox, J.C., J.E. Ingersoll, and S.A. Ross, "A Theory of the Term Structure of Interest Rates," *Econometrica*, 53 (1985), 385-407.

Cox, J.C., S.A. Ross, and M. Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (1979), 229-263.

de Boor, C., *A Practical Guide to Splines*, Springer-Verlag, New York, 1978.

Dothan, L.U., "On the Term Structure of Interest Rates," *Journal of Financial Economics*, 6 (1978), 59-69.

Elliott, R.J. and P.E. Kopp, *Mathematics of Financial Markets*, Springer-Verlag, New York, 1999.

Fabozzi, F.J., *The Handbook of Fixed Income Securities*, Fifth Edition, McGraw-Hill, New York, 1997.

Heath, D., R. Jarrow, and A. Morton, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation," *Econometrica*, 60 (1992), 77-105.

Ho, T.S.Y. and S.B. Lee, "Term Structure Movements and Pricing Interest Rate Contingent Claims," *Journal of Finance*, 41 (1986), 1011-1029.

Hull, J. and A. White, "Pricing Interest-Rate Derivative Securities," *Review of Financial Studies*, 3 (1990a), 573-592.

Hull, J. and A. White, "Valuing Derivative Securities Using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis*, 25 (1990b), 87-100.

Hull, J. and A. White, "One Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities," *Journal of Financial and Quantitative Analysis*, 28 (1993), 235-254.

Hull, J. and A. White, "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models," *Journal of Derivatives*. 2 (1994a), 7-16.

Hull, J. and A. White, "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models," *Journal of Derivatives*, 2 (1994b), 37-48.

Hull, J. and A. White, "A Note on the Models of Hull and White for Pricing Options on the Term Structure: Response," *Journal of Fixed Income*, 5 (1995), 97-103.

Hull, J. and A. White, "Using Hull-White Interest Rate Trees," *Journal of Derivatives*, 4 (1996), 26-36.

Jamshidian, F., "An Exact Bond Option Formula," *Journal of Finance*, 44 (1989), 205-209.

Kalotay, A.J., G.O. Williams, and F.J. Fabozzi, "A Model for Valuing Bonds and Embedded Options," *Financial Analysts Journal*, 49 (1993), 35-47.

Karatzas, I. and S.E. Shreve, *Methods of Mathematical Finance*, Springer-Verlag, New York, 1998.

Longstaff, F.A., "Pricing Options with Extendible Maturities: Analysis and Applications," *Journal of Finance*, 45 (1990), 935-957.

Marsh, T.A. and E.R. Rosenfeld, "Stochastic Processes for Interest Rates and Equilibrium Bond Prices," *Journal of Finance*, 38 (1983), 635-646.

Mason, S., R. Merton, A. Perold, and P. Tufano, *Cases in Financial Engineering*, Prentice-Hall, New Jersey, 1995.

Merton, R.C., "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4 (1973), 141-183.

Øksendal, B., *Stochastic Differential Equations, An Introduction with Applications*, Fourth Edition, Springer-Verlag, Germany, 1995.

Rabinovitch, R., "Pricing Stock and Bond Options when the Default-Free Rate is Stochastic," *Journal of Financial and Quantitative Analysis*, 24 (1989), 447-457.

Richard, S.F., "An Arbitrage Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, 6 (1978), 33-57.

Schaefer, S.M. and E.S. Schwartz, "A Two-Factor Model of the Term Structure: An Appropriate Analytical Solution," *Journal of Financial and Quantitative Analysis*, 19 (1984), 413-424.

Vasicek, O., "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5 (1977), 177-188.

Wilmott, P., J. Dewynne, and S. Howison, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, United Kingdom, 1993.