Applied Mathematics and Computation 474 (2024) 128698

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Full Length Article

# Pricing European call options with interval-valued volatility and interest rate

Song Wang

*School of Electrical Engineering, Computing & Mathematical Sciences, Curtin University, GPO Box U1987 Perth, WA 6845, Australia*

A R T I C L E I N F O

*MSC:*
65K15
91G20
90C70

*Keywords:*
Interval optimization
Partial-differential inequality constrained optimization
European call option valuation under uncertainties
Variational inequality
Interior penalty method
Finite differences

A B S T R A C T

We propose a novel approach to pricing European call options when both of the volatility of the underlying asset and interest are uncertain. In this approach, we formulate the option pricing problem with uncertain parameters as a partial-differential inequality constrained interval optimization problem. An interior penalty method is then developed for the numerical solution of the finite-dimensional optimization problem arising from the discretization of the continuous pricing problem by a finite difference scheme. A convergence theory for the penalty method is established. An algorithm based on Newton's iterative method is also proposed for solving the penalty equation. Numerical results are presented to demonstrate the effectiveness and usefulness of this approach and the numerical methods.

## 1. Introduction

An options is a financial derivative which gives its hold the right, not obligation, to buy (Call option) or sale (Put Option) a predefined number of shares of an underlying stock at a fixed price, called strike price, on or before a given date (Maturity). If an option can only be exercised on maturity, it is called a European option. Otherwise, it is an American option. Options are used extensively in hedging risks in investments. Options can be traded on a secondary financial market before maturity, and thus how to price an option accurately has been a hot topic for both practitioners and researchers. In [4] Black and Scholes proposed a mathematical model, known as the Black-Scholes model for pricing European options under certain conditions. These conditions include that the short interest rate is constant and the underlying stock price follows a geometric Brownian motion with a constant volatility. Since the publication of this seminal paper by Black and Scholes, valuation of options has attracted a lot of attention from both researchers and practitioners.

Assume the price of a stock, $S$, follows a geometric Brownian motion. Using Ito's lemma and the $\Delta$-hedging strategy, one can easily show that the price of a European call option $V$ satisfies the following Black-Scholes' equation

$$-\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rS\frac{\partial V}{\partial S} + rV = 0 \tag{1}$$

*E-mail address:* Song.Wang@curtin.edu.au.

https://doi.org/10.1016/j.amc.2024.128698
Received 13 December 2023; Received in revised form 12 March 2024; Accepted 19 March 2024

Available online 27 March 2024
0096-3003/

with appropriate boundary and terminal conditions, where $\sigma$ is the volatility of the underlying stock and $r$ the nominal interest rate. This PDE governs the price of a European option when both $\sigma$ and $r$ are given constants.

Many extensions of (1) have been proposed to price options under more realistic economic and market conditions such as transaction costs in trading and stochastic market parameters. In particular, models for pricing options under stochastic volatility $\sigma$ or/and interest rate $r$ in (1) have been established. These models are based on the assumption that the volatility of the underlying stock price and/or interest rate, in addition to the stock price itself, follow their respective Brownian motions. In practice, system and market parameters of any aforementioned models can not be known exactly and need to be calibrated using trading or economic data. Thus, these calibrated data contain uncertainties. Full statistical information for such an uncertain parameter, such as its distribution, can hardly be determined, but partial statistical information on the parameter can be determined easily. Therefore, in practice, we often represent an uncertain parameter approximately by an interval, called an *interval number*. Optimal decision-making for problems with interval coefficients has attracted much attention from researchers and practitioners. In [1], a technique based on the so called 'worst-case scenario' is developed to price a European option when $\sigma$ in (1) is uncertain (non-stochastic), and a computational method for solving this model is developed in [25]. There are some non-PDE based methods for pricing European options when the coefficients are uncertain or fuzzy such as those in [3,16,29], just to name a few. However, to our best knowledge, there are essential no methods for determining optimal solutions of PDE pricing models such as (1) in the open literature when some of the coefficients in (1) are interval numbers.

In this work, we propose a novel approach to determining optimal values of European options when both $\sigma$ and $r$ in (1) are interval numbers, based on some interval number order relations. We will also develop a numerical method for solving the pricing problem and present a mathematical analysis for the numerical method. The rest of this work is organise as follows.

In the next section, we first give a brief account of interval number order relations. We then formulate the option pricing problem with uncertain $\sigma$ and $r$ as a constrained interval optimization problem and show that the feasible set of solutions to this continuous optimization problem is non-empty. In Section 3, we discretize the interval optimization problem using a finite-difference scheme to form a finite-dimensional constrained interval optimization problem. In Section 4, we propose to approximate the KKT conditions of the discrete interval optimization problem by a penalty equation, and establish a convergence theory for the solutions from the penalty equation. An algorithm is also proposed for numerically solving the penalty equation. In Section 5, we present some numerical results to demonstrate the usefulness and effectiveness of this approach.

## 2. Preliminaries and problem formulation

### 2.1. Interval numbers and orders

When a real number $a$ is uncertain, we often represent it as an interval number, that is, $a = [a_L, a_R]$, where $a_L$ and $a_R$ are two real numbers such that $a_L \le a_R$. This interval number can alternatively be written as $a = \langle a_C, a_W \rangle$, where $a_C = (a_L + a_R)/2$ and $a_W = (a_R - a_L)/2$. Definitions for extensions of arithmetic operations on scalars to interval numbers have been discussed and used in many existing works such as [2,5,6,8,11,15,19,20,24]. In this work, we use the following definitions of extensions of arithmetic operations ([15]).

**Definition 2.1.** Let $a = [a_L, a_R] = \langle a_C, a_W \rangle$ and $b = [b_L, b_R] = \langle b_C, b_W \rangle$ be two interval numbers. Then, the arithmetic operations addition (+), multiplication ($\cdot$) and division (/) are defined respectively as

- $a + b = [a_L + b_L, a_R + b_R] = \langle a_C + b_C, a_W + b_W \rangle$,
- $a \cdot b = [\min\{a_L b_L, a_R b_L, a_L b_R, a_R b_R\}, \max\{a_L b_L, a_R b_L, a_L b_R, a_R b_R\}]$,
- $a/b = [\min\{a_L/b_L, a_R/b_L, a_L/b_R, a_R/b_R\}, \max\{a_L/b_L, a_R/b_L, a_L/b_R, a_R/b_R\}]$, when $0 \in [b_L, b_R]$, and
- $ka = \langle ka_C, |k| a_W \rangle = \begin{cases} [ka_L, ka_R] & k \ge 0 \\ [ka_R, ka_L] & k < 0 \end{cases}$ for any real number $k$.

The order relationships on interval numbers defined below are important in an minimization problem.

**Definition 2.2.** Let $a = \langle a_C, a_W \rangle = (a_L, a_R)$ and $b = \langle b_C, b_W \rangle = (b_L, b_R)$. The interval orders $\preceq_{CW}$, $\prec_{CW}$, $\preceq_{CR}$ and $\prec_{CR}$ for $a$ and $b$ are defined as follows.

- $a \preceq_{CW} b$ iff $a_C \le b_C$ and $a_W \le b_W$,
- $a \prec_{CW} b$ iff $a \preceq_{CW} b$ and $a \ne b$,
- $a \preceq_{LR} b$ iff $a_L \le b_L$ and $a_R \le b_R$,
- $a \prec_{LR} b$ iff $a \preceq_{CR} b$ and $a \ne b$.

These order relations are partial orders on the set of all interval numbers. We may also define the following order relation

- $a \preceq_{CR} b$ iff $a_C \le b_C$ and $a_R \le b_R$,
- $a \prec_{CR} b$ iff $a \preceq_{CR} b$ and $a \ne b$.

It has been shown in [15] that $a \preceq_{CR} (\prec_{CR}) b$ if and only if $a \preceq_{CW} (\prec_{CW}) b$ or $a \preceq_{LR} (\prec_{LR}) b$. Thus, we will use this order relation in the rest of this work. For notational simplicity we will simply write $\preceq_{CR}$ as $\preceq$.

We comment that $\preceq_{CR}$ is a partial order. However, we will show that under this partial order relationship, the solution set of our optimization-based pricing problem is non-empty. Also, it is straightforward to show our results to be presented in the rest of this paper hold true when $\preceq_{CR}$ is replaced by the full order relationship proposed in [11].

### 2.2. The option pricing problem

Assume the unit price $S(t)$ of a traded risky asset follows a geometric Brownian motion at any time $t$, i.e., $S$ satisfies the following stochastic differential equation:

$$dS(t) = \mu S(t)dt + \sigma S(t)dW,$$

where $\mu$ is the drift rate of $S(t)$, $\sigma$ is the volatility and $dW$ denotes the increment of a Wiener process $W$. An investor would like to sell a European call option on this risky asset with strike price $K > 0$ and maturity $T > 0$, assuming the risk-free return rate is $r > 0$. A natural question is how to determine the fair price $V$ of this option when both $\sigma$ and $r$ are interval numbers.

When $\sigma$ and $r$ follow their respective Brownian motions, their dynamics can be modelled by a stochastic differential equation. In this case, $V$ is governed by a multi-factor Black-Scholes equation [10,14,28]. However, for our case, full statistical properties of $\sigma$ and $r$ are unknown and these parameters can only be characterized as interval numbers. In [1], a pricing model for the case that $\sigma$ is an interval number is developed based on the 'worse-case scenario'. In our present work, we shall use the interval number optimization technique and the $\Delta$-hedging strategy to price a European call option when both $\sigma$ and $r$ are interval numbers. More specifically, we consider the hedging strategy for the investor issuing one European call whose value at $t$ is $V(S(t),t)$. It is known that the investor can use the $\Delta$-hedging strategy to neutralize the risk in selling the option, i.e. the investor needs to hold $\frac{\partial V}{\partial S}$ number of the underlying shares at time $t$, so that the portfolio

$$\Pi(S,t) = \frac{\partial V(S(t),t)}{\partial S} S(t) - V(S(t),t) \tag{2}$$

becomes risk-free. Thus, the investment gain in $\Pi$ in an infinitesimal $[t, t+dt)$ is a combination of $dS$ and $dV$, i.e., $d\Pi = \frac{\partial V}{\partial S} dS - dV$. Applying Ito's lemma to $\Pi(S(t),t)$, we can represent $d\Pi$ as

$$d\Pi = -\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right) dt. \tag{3}$$

In the conventional case that $\sigma$ and $r$ are non-interval numbers, we expect this portfolio $\Pi$ is also risk-free, i.e.,

$$d\Pi = r\Pi dt = r\left(\frac{\partial V}{\partial S} S - V\right) dt, \tag{4}$$

which, when combined with (3), yields (1). However, in the case that both $\sigma$ and $r$ are interval numbers, (1) cannot usually hold true. This is because that, from (3), we see that $d\Pi$ is an interval number depending on $\sigma$ and the right-hand side of (4) is an interval number depending on $r$. There usually does not exist a function $V$ such that the two interval numbers are equal to each other. Thus, it is generally not possible to determine $V$ so that the conventional Black-Scholes equation (1) holds true when $\sigma$ and $r$ are interval numbers.

To remedy this difficulty, we propose a new technique to determine the option price. Note that the coefficient of $dt$ on the right-hand side of (3) represents the growth rate in the value of the portfolio $\Pi$ at $t$. Since this growth rate as a interval number can not be exactly equal to the (uncertain) risk-free growth rate $r\Pi$ on the right-hand side of (4), we propose to use to following optimization problem to determine $V$ at any time $t$: find option price $V(S,t)$ satisfying

$$\text{minimize} \int_0^{S_{\max}} \Pi^2(S,t)dS \tag{5}$$

$$\text{subject to } r(t)\left(\frac{\partial V}{\partial S} S - V\right) \preceq -\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2(t) \frac{\partial^2 V}{\partial S^2}\right) \tag{6}$$

for $S \in (0, S_{\max})$ almost everywhere (*a.e.*), where $S_{\max} \gg K$ is a positive number, $\Pi$ is the portfolio defined in (2), and $\preceq$ represents one of the order relations defined in Definition 2.2. The explanation of this optimization problem is that we look for $V$ which, at time $t$ and when $dt \to 0^+$, minimizes the $L_2$-norm squared of the portfolio $\Pi$ in a computable range $(0, S_{\max})$ of the underlying stock price $S$, subject to the constraint that the gain rate of $\Pi$ (as an interval number) is not smaller than that of the risk-free portfolio. The minimization at $t$ can also be regarded as the worst-case scenario of the portfolio while its return is at lease equal to that of the risk-free portfolio. This is reasonable because the interval number $r$ should be the minimum return rate an investor expects.

The following theorem shows that (5)–(6) is equivalent to the conventional Black-Scholes partial differential equation model when both $\sigma$ and $r$ are real numbers.

**Theorem 1.** *When $\sigma > 0$ and $r > 0$ become real numbers, the solution to the optimization problem defined in (5)–(6) is determined by the Black-Scholes equation (1).*

**Proof.** When $\sigma > 0$ and $r > 0$ are real numbers, the order relation $\preceq$ becomes $\leq$, and thus (6) becomes a conventional constraint that the return rate of $\Pi$ at $t$ is bounded below by $r$. For any $t > 0$, consider the interval $[t-\Delta t, t]$, where $\Delta t > 0$ is a sufficiently small increment in time. We write $\Pi(S,t) = \Pi(S,t-\Delta t) + \Delta\Pi(S,t)$, where $\Pi(S,t-\Delta t) \geq 0$ is the value of the portfolio at $t-\Delta t$. Integrating $\Pi^2(S,t)$ from 0 to $S_{\max}$ and using this equality and (3) with $dt$ replaced with $\Delta t$, we have

$$\int_0^{S_{\max}} \Pi^2(S,t)dS = \int_0^{S_{\max}} \left[\Pi(S,t-\Delta t) - \left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)_{t-\Delta t} \Delta t\right]^2 dS$$
$$\geq \int_0^{S_{\max}} \left[\Pi(S,t-\Delta t) + \left(r\left(\frac{\partial V}{\partial S}S - V\right)\right)_{t-\Delta t} \Delta t\right]^2 dS$$

for any feasible $V$ by (6). Therefore, when

$$-\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}\right)_{t-\Delta t} = \left[r\left(\frac{\partial V}{\partial S}S - V\right)\right]_{t-\Delta t},$$

the cost in (5) is minimized. Finally, when $\Delta t \to 0+$, the above equality becomes (1). □

From the definition of $\Pi$ in (2), we see that the cost function in (5) can be rewritten as

$$\int_0^{S_{\max}} \Pi^2(S,t)dS = \int_0^{S_{\max}} \left(\left(\frac{\partial V}{\partial S}\right)^2 S^2 - 2\frac{\partial V}{\partial S}SV + V^2\right) dS. \tag{7}$$

Using integration by parts, we have

$$\int_0^{S_{\max}} \frac{\partial V}{\partial S}SV = V^2(S_{\max})S_{\max} - \int_0^{S_{\max}} \left(V^2 + \frac{\partial V}{\partial S}SV\right) dS,$$

from which we obtain

$$\int_0^{S_{\max}} \frac{\partial V}{\partial S}SV\,dS = \frac{1}{2}V^2(S_{\max},t)S_{\max} - \frac{1}{2}\int_0^{S_{\max}} V^2 dS,$$

where $V(S_{\max},t)$ is the boundary condition of $V$ at $(S_{\max},t)$, which is predetermined in computation. Thus, combining the above equality and (7) gives

$$\int_0^{S_{\max}} \Pi^2(S,t)dS = \int_0^{S_{\max}} \left[\left(\frac{\partial V}{\partial S}\right)^2 S^2 + 2V^2\right] dS - V^2(S_{\max},t)S_{\max}. \tag{8}$$

Since $V(S_{\max},t)$ is a prescribed boundary condition, using (8), we may rewrite (5)–(6) as the following equivalent minimization problem:

$$\text{minimize} \int_0^{S_{\max}} \left[\left(\frac{\partial V}{\partial S}\right)^2 S^2 + 2V^2\right] dS \tag{9}$$

$$\text{subject to } r(t)\left(\frac{\partial V}{\partial S}S - V\right) \preceq -\left(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(t)S^2\frac{\partial^2 V}{\partial S^2}\right) \tag{10}$$

for $(S,t) \in (0,S_{\max}) \times [0,T)$ *a.e.*.

We comment that in this continuous optimization problems, we only require that the constraint is satisfied almost everywhere in $(0,S_{\max}) \times [0,T)$. This is because a subset of $(0,S_{\max}) \times [0,T)$ with a zero measure on which (10) is not satisfied by $V$ does not affect the value of the cost function in (9).

In (9)–(10), we have assumed that $\sigma$ is a local volatility and $r$ is also a function of $t$. The boundary and payoff conditions for this call option are

$$V(0,t) = 0, \quad t \in [0,T), \tag{11}$$

$$V(S_{\max},t) = S_{\max} - K\exp\left(-\int_t^T r(\tau)d\tau\right), \quad t \in [0,T), \tag{12}$$

$$V(S,T) = \max\{S-K,0\}, \quad S \in (0,S_{\max}). \tag{13}$$

We also represent the interval numbers $\sigma(t)$ and $r(t)$ as $\sigma(t) = [\sigma_L(t), \sigma_R(t)]$ and $r(t) = [r_L(t), r_R(t)]$ for $t \in [0,T)$, where $\sigma_L, \sigma_R, r_L$ and $r_R$ are given functions of $t$ satisfying $\sigma_L(t) < \sigma_R(t)$ and $r_L(t) < r_R(t)$ for all $t \in [0,T)$.

We remark that $V(S_{\max}, t)$ in (12) is also an interval number because of $r$. In computation, we need to fix this boundary condition by replacing $r(\tau)$ in (12) with $(r(\tau)_L + r_R(\tau))/2$.

We now show in the following theorem that the set of feasible solutions to (9)–(10) is non-empty in the weak sense, i.e., there exists at least one function that satisfies (10) almost everywhere ((*a.e.*) on $(0, S_{\max}) \times (0,T)$ and the boundary and terminal conditions (11)–(12).

**Theorem 2.** *The set of feasible solutions to (9)–(13) is non-empty.*

**Proof.** Consider the function $U(S,t) := \max\left\{0, S - K\exp\left(-\int_t^T r(\tau)d\tau\right)\right\}$. We shall show $U(S,t)$ satisfies (10) for $(S,t) \in (0, S_{\max}) \times [0,T)$ *a.e.* and the payoff and boundary conditions (11)–(13). It is trivial to verify that $U(S,t)$ satisfies the payoff and boundary conditions (11)–(12), and we only prove the former.

We now show $U(S,t)$ satisfies (10) for $(S,t) \in (0, S_{\max}) \times (0,T)$ *a.e.*. From its definition, we have that, when $S < K\exp\left(-\int_t^T r(\tau)d\tau\right)$, $U(S,t) = 0$, and thus (10) holds true in the sub-domain $\{(S,t) : S < K\exp\left(-\int_t^T r(\tau)d\tau\right), 0 < t < T\}$.

When $S > K\exp\left(-\int_t^T r(\tau)d\tau\right)$, from the definition of $U$ we have

$$\frac{\partial U}{\partial S} = 1, \quad \frac{\partial U}{\partial t} = -r(t)K\exp\left(-\int_t^T r(\tau)d\tau\right), \quad \frac{\partial^2 U}{\partial S^2} = 0.$$

Therefore, for any $t \in (0,T)$, we have, in this case,

$$r(t)\left(\frac{\partial U}{\partial S}S - U\right) = r(t)K\exp\left(-\int_t^T r(\tau)d\tau\right) = -\left(\frac{\partial U}{\partial t} + \frac{1}{2}\sigma^2(t)S^2\frac{\partial^2 U}{\partial S^2}\right)$$

Combining the above two cases we see that (10) is satisfied by $U$ with '$\preceq$' replaced with '$=$' in

$$\left\{(S,t) \in (0, S_{\max}) \times (0,T) : S \neq K\exp\left(-\int_t^T r(\tau)d\tau\right)\right\}. \quad \square$$

In Theorem 2, we have shown that the feasible solution set of the pricing problem (9)–(10) is non-empty under the partial order interval relation $\preceq$. As mentioned before, Theorem 2 also holds true when $\preceq$ is replaced by the full interval order relation proposed in [11]. In the rest of this work, we shall use $\preceq$ in our discussion, as we are able to make an optimal decision for the pricing problem even this partial interval order relationship is used.

## 3. Discretization of the infinite-dimensional optimization problem

We now discretize the optimization problem (9)–(10) by a finite difference scheme based on the upwind finite difference schemes in [17,23]. Let $M$ and $N$ be two positive integers. We partition $(0, S_{\max}) \times (0,T)$ into a set of sub-domains using the break points $(S_i, t_n)$ for $i = 0,1,...,M$ and $n = 0,1,...,N$ satisfying

$$0 = S_0 < S_1 < \ldots < S_M = S_{\max}, \quad T = t_0 > t_1 > \ldots > t_N = 0.$$

Let $h_i := S_{i+1} - S_i$ for $i = 0,1,...,M-1$ and $\Delta t_n = t_n - t_{n-1}$ for $n = 1,2,...,N$. (Note $\Delta t_n < 0$ for all feasible $n$.) For any $i \in \{1,2,...,M-1\}$ and $n \in \{1,2,...,N\}$, we approximate the derivatives in (10) at $(S_i, t_n)$ respectively by the following finite differences used in [17]:

$$\frac{\partial V}{\partial t} \approx \frac{V_i^{n-1} - V_i^n}{-\Delta t_n}, \quad r(t)\left(S\frac{\partial V}{\partial S} - V\right) \approx r^n\left(S_i\frac{V_{i+1}^n - V_i^n}{h_i} - V_i^n\right),$$

$$\frac{1}{2}\sigma^2(t)S^2\frac{\partial^2 V}{\partial S^2} \approx (\sigma^n)^2(-p_{i,i-1}V_{i-1}^n - p_{i,i}V_i^n - p_{i,i+1}V_i^{n+1}),$$

where $V_j^k$ denotes an approximation to $V(S_j, t_k)$ for all feasible $j$ and $k$, $\sigma^n = \sigma(t_n)$ and $r^n = r(t_n)$ are interval numbers at $t_n$, and

$$p_{i,i-1} = \frac{-S_i^2}{h_{i-1}(h_{i-1} + h_i)}, \quad p_{i,i+1} = \frac{-S_i^2}{h_i(h_{i-1} + h_i)}, \quad p_{i,i} = -p_{i,1} - p_{i,3} = \frac{S_i^2}{h_{i-1}h_i}. \quad (14)$$

Replacing the derivatives in (10) with the approximations defined above gives

$$r^n\left(S_i\frac{V_{i+1}^n - V_i^n}{h_i} - V_i^n\right)$$
$$\preceq -\frac{V_i^{n-1}-V_i^n}{|\Delta t_n|} + (\sigma^n)^2(p_{i,i-1}V_{i-1}^n + p_{i,i}V_i^n + p_{i,i+1}V_{i+1}^n)$$

for $i = 1, 2, ..., M-1$ and $n = 1, 2, ..., N$ with the payoff and boundary conditions (11)–(13). Discretizing (9) using the mesh defined and rewriting the above inequality in a matrix form, we have the following finite-dimensional optimization problem approximating (9)–(10):

$$\min\ 2\sum_{i=0}^{M}(V_i^n)^2 h_{i-1/2} + \sum_{i=0}^{M-1} h_i S_{i+1/2}^2\frac{(V_{i+1}^n - V_i^n)^2}{h_i^2} \tag{15}$$

$$\text{subj. to } r^n\left(-Q\hat{V}^n + b_2^n\right) \preceq \frac{1}{|\Delta t_n|}\hat{V}^n + (\sigma^n)^2\left(P\hat{V}^n - b_1^n\right) - \frac{1}{|\Delta t_n|}\hat{V}^{n-1}, \tag{16}$$

where $h_{i-1/2} = (h_{i-1}+h_i)/2$ for $i = 0, 1, ..., M$ with $h_0 = 0 = h_M$, $S_{i+1/2} = (S_i + S_{i+1})/2$, $\hat{V}^k = (V_1^k, V_2^k, ..., V_{M-1}^k)^\top$ for $k = n-1, n$, $P$ is a tri-diagonal matrix of order $(M-1)\times(M-1)$ with non-zero entries defined in (14), $Q$ is a bi-diagonal $(M-1)\times(M-1)$ matrix with $q_{i,i} = \frac{S_i}{h_i} + 1$ and $q_{i,i+1} = -\frac{S_i}{h_i}$, and $b_1^n$ and $b_2^n$ are $(M-1)\times 1$ matrices defined by

$$b_1^n = -(p_{1,0}V_0^n, ..., p_{M-1,M}V_M^n)^\top \text{ and } b_2^n = -(0, ..., 0, q_{M-1,M}V_M^n)^\top.$$

For any $n = 1, 2..., N$, we define

$$\begin{cases} r_*^n = \arg\min_{r_L(t_n)\le r\le r_R(t_n)} r\left(-Q\hat{V}^n + b_2^n\right), \\ r^{*n} = \arg\max_{r_L(t_n)\le r\le r_R(t_n)} r\left(-Q\hat{V}^n + b_2^n\right), \end{cases} \tag{17}$$

$$\begin{cases} \sigma_*^n = \arg\min_{\sigma_L(t_n)\le\sigma\le\sigma_R(t_n)} \sigma^2\left(P\hat{V}^n - b_1^n\right), \\ \sigma^{*n} = \arg\max_{\sigma_L(t_n)\le\sigma\le\sigma_R(t_n)} \sigma^2\left(P\hat{V}^n - b_1^n\right). \end{cases} \tag{18}$$

Note that the above optimization processes are performed element-by-element, and thus $r_*^n, r^{*n}, \sigma_*^n, \sigma^{*n} \in \mathbb{R}^{(M-1)}$. From the above it is easy to see that $r_{*,i}^n = \frac{1}{2}[\text{sgn}((-Q\hat{V}^n + b_2^n)_i) + 1]r_L(t_n) - \frac{1}{2}[\text{sgn}((-Q\hat{V}^n + b_2^n)_i) - 1)]r_R(t_n)$, and $r_i^{*n}$, $\sigma_{*,i}^n$ and $\sigma_i^{*n}$ can be evaluated in a similarly way for $i = 1, 2, ..., M-1$, where sgn() denotes the sign function.

We comment that (15)–(16) is to minimize the sum of squared weighted $\ell_2$-norms of $\hat{V}^n$ and the forward differences of $\hat{V}^n$. From the definition of $S_{i+\frac{1}{2}}$ and $h_{i-\frac{1}{2}}$ we see that the values of the weights of the two sums in (15) may differ to each other by a few orders of magnitudes. This unbalance in the numerical values of the two terms will cause difficulties in numerical solution of (15)–(16), as, in computation, (15)–(16) is to effectively minimize the second sum in (15). To remedy this difficulty in computation, we need to scale the two sums in (15). More specifically, we replace the cost function in (15) with $2\sum_{i=0}^{M}(V_i^n)^2 h_{i-1/2} + \alpha\sum_{i=0}^{M-1} S_{i+1/2}^2\frac{(V_{i+1}^n - V_i^n)^2}{h_i}$, where $\alpha \in (0, 1]$ is a chosen constant. Thus, in what follows, we shall use this scaling parameter $\alpha$ in this optimization problem.

Note $V_0^n$ and $V_M^n$ are given boundary conditions. When all $h_i$'s are sufficiently small, we may omit the first and last terms in both of the sums in (15) involving the boundary conditions. Thus, from the definition of $\preceq$ (i.e., $\preceq_{CR}$) and using the optimal solutions in (17)–(18), we rewrite (15)–(16) as the following equivalent problem.

$$\min\ 2\sum_{i=1}^{M-1} h_{i-\frac{1}{2}}(V_i^n)^2 + \alpha\sum_{i=1}^{M-2}\frac{S_{i+\frac{1}{2}}^2}{h_i}(V_{i+1}^n - V_i^n)^2 \tag{19}$$

$$\text{subj. to } c^n(\sigma_A^n, r_A^n) - A^n(\sigma_A^n, r_A^n)\hat{V}^n \le 0, \tag{20}$$

$$c^n(\sigma^{*n}, r^{*n}) - A^n(\sigma^{*n}, r^{*n})\hat{V}^n \le 0 \tag{21}$$

for $n = 1, 2, ..., N$, where $\sigma_A^n = \frac{1}{2}(\sigma_*^n + \sigma^{*n})$, $r_A^n = \frac{1}{2}(r_*^n + r^{*n})$,

$$A^n(\sigma, r) := \frac{1}{|\Delta t_n|}I + (\sigma .* \sigma) .* P + r .* Q, \tag{22}$$

$$c^n(\sigma, r) := \frac{1}{|\Delta t_n|}\hat{V}^{n-1} + (\sigma .* \sigma) .* b_1^n + r .* b_2^n \tag{23}$$

for $\sigma, r \in \mathbb{R}^{(M-1)}$, where $.*$ denotes the element-by-element product operator (used in Matlab). The payoff (terminal) condition is $\hat{V}^0 = (V(S_1, T), ..., V(S_{M-1}, T))^\top$ with $V(S, T)$ given in (13).

We comment that the constraints (20) and (21) are nonlinear, as $\sigma_A^n$, $\sigma^{*n}$, $r_A^n$ and $r^{*n}$ are functions of the decision variable $\hat{V}^n$.

For the system matrix $A^n(\sigma, r)$ in (22), we have the following theorem.

**Theorem 3.** *When $|\Delta t|$ is sufficiently small, the matrix $A^n(\sigma, r)$ defined in (22) is a positive-definite $M$-matrix.*

The proof of this theorem can be found in [17].

## 4. Solution of the finite-dimensional optimization problem

We now consider the solution of (19)–(21). Since the 2nd sum in (19) is a quadratic form, we rewrite it as the following symmetric form:

$$\sum_{i=1}^{M-2} \frac{S_{i+\frac{1}{2}}^2}{h_i}(V_{i+1}^n - V_i^n)^2 = \frac{1}{2}(\hat{V}^n)^\top E \hat{V}^n, \tag{24}$$

where $E$ is an $(M-1)\times(M-1)$ tri-diagonal symmetric matrix with non-zero entries given below:

$$\begin{cases} e_{i,i-1} = -\dfrac{S_{i-\frac{1}{2}}^2}{h_{i-1}} \text{ for } i = 2,3,\dots,M-1, \; e_{i,i+1} = -\dfrac{S_{i+\frac{1}{2}}^2}{h_i} \text{ for } i = 1,2,\dots,M-2, \\ e_{i,i} = |e_{i,i-1}| + |e_{i,i+1}| \text{ for } i = 1,2,\dots,M-1 \text{ with } e_{0,1} = 0 = e_{M-1,M}. \end{cases} \tag{25}$$

The KKT conditions for (19)–(21) are listed below:

$$[H + \alpha E]\hat{V}^n - (A^n(\sigma_A^n, r_A^n))^\top \lambda_1 - (A^n(\sigma^{*n}, r^{*n}))^\top \lambda_2 = 0, \tag{26}$$

$$\begin{cases} c^n(\sigma_A^n, r_A^n) - A^n(\sigma_A^n, r_A^n)\hat{V}^n \le 0, \\ -\lambda_1 \le 0, \\ \lambda_1^\top \left(c^n(\sigma_A^n, r_A^n) - A^n(\sigma_A^n, r_A^n)\hat{V}^n\right) = 0, \end{cases} \tag{27}$$

$$\begin{cases} c^n(\sigma^{*n}, r^{*n}) - A^n(\sigma^{*n}, r^{*n})\hat{V}^n \le 0, \\ -\lambda_2 \le 0, \\ \lambda_2^\top \left(c^n(\sigma^{*n}, r^{*n}) - A^n(\sigma^{*n}, r^{*n})\hat{V}^n\right) = 0, \end{cases} \tag{28}$$

where $\lambda_1 \in \mathbb{R}^{M-1}$ and $\lambda_2 \in \mathbb{R}^{M-1}$ are multipliers and $H = 4\mathrm{diag}(h_{\frac{1}{2}}, h_{1+\frac{1}{2}}, \dots, h_{(M-1)-\frac{1}{2}})$ is an $(M-1)\times(M-1)$ diagonal matrix. For the matrix $H + \alpha E$ we have the following theorem.

**Theorem 4.** *The matrix $H + \alpha E$ with $\alpha > 0$ is a positive-definite $M$-matrix.*

**Proof.** We first prove positive-definiteness of $H + \alpha E$. For any $u \in \mathbb{R}^{M-1}$, from the definitions $H$ and $E$ in (24), we have

$$u^\top (H + \alpha E)u = \sum_{i=1}^{M-1} h_{i-\frac{1}{2}} u_i^2 + \alpha \sum_{i=1}^{M-2} \frac{S_{i+\frac{1}{2}}^2}{h_i}(u_{i+1} - u_i)^2 \ge \left(\min_{1\le i\le M-1} h_{i-\frac{1}{2}}\right) \|u\|_2^2,$$

where $\|\cdot\|_2$ denotes the $\ell_2$-norm on $\mathbb{R}^{M-1}$. Thus, it is positive-definite.

To prove that $H + \alpha E$ is an $M$-matrix, we note that from (25), we see that $E$ is a diagonally dominant with $e_{i,i} > 0$ and $e_{i,j} \le 0$ when $j \ne i$. Therefore, $H + \alpha E$ satisfies the properties that it is strictly diagonally dominant with positive diagonal and non-positive off-diagonal entries. It is clear that $E$ is irreducible. Thus, $H + \alpha E$ is irreducibly diagonally dominant, and so it is am $M$-matrix (see, eg, [22]). □

### 4.1. The interior penalty method

Penalty methods have been shown to be very efficient and effective for solving both infinite- and finite-dimensional complementarity problems (see, for example, [7,12,13,26,27]). In this subsection, we shall use the idea proposed in [26,27] and construct a system of nonlinear equations to approximate (26)–(28). This idea is based on an interior method for constrained nonlinear optimization problems ([9]).

Note that (26)–(28) is a mixed complementarity problem containing two sub complementarity problems. If we let $x = \hat{V}^n$ and $y = (\lambda_1^\top, \lambda_2^\top)^\top \in \mathbb{R}^{2\times(M-1)}$, we can rewrite (26)–(28) in the following matrix form.

$$(H + \alpha E)x - F^{n\top} y := (H + \alpha E)x - \left[ A^n(\sigma_A^n, r_A^n)^\top \quad A^n(\sigma^{*n}, r^{*n})^\top \right] y = 0, \tag{29}$$

$$\begin{cases} F^n x - f^n := \begin{bmatrix} A^n(\sigma_A^n, r_A^n) \\ A^n(\sigma^{*n}, r^{*n}) \end{bmatrix} x - \begin{bmatrix} c^n(\sigma_A^n, r_A^n) \\ c^n(\sigma^{*n}, r^{*n}) \end{bmatrix} \geq 0, \\ y \geq 0, \\ y^\top (F^n x - f^n) = 0, \end{cases} \tag{30}$$

where the definitions of the matrices introduced in (29)–(30) are self-explanatory.

Let $\mathcal{K} := \left\{ \begin{bmatrix} u \\ v \end{bmatrix} \in \mathbb{R}^{M-1} \times \mathbb{R}^{2(M-1)} : v > 0 \right\}$. Following the ideas in [26,27], we propose to approximate (29)–(30) with the following problem. Find $(x_\mu^\top, y_\mu^\top)^\top \in \mathcal{K}$ satisfying

$$G^n \begin{bmatrix} x_\mu \\ y_\mu \end{bmatrix} - \begin{bmatrix} 0 \\ \mu./y_\mu \end{bmatrix} := \begin{bmatrix} H + \alpha E & -(F^n)^\top \\ F^n & 0 \end{bmatrix} \begin{bmatrix} x_\mu \\ y_\mu \end{bmatrix} - \begin{bmatrix} 0 \\ \mu./y_\mu \end{bmatrix} = \begin{bmatrix} 0 \\ f^n \end{bmatrix}, \tag{31}$$

where $\mu$ is a (small) positive constant, called a penalty constant, $x_\mu$ and $y_\mu$ are approximations to $x$ and $y$ respectively and ./ denotes the element-by-element division operator used in Matlab programming language, i.e. $\mu./y_\mu = (\mu/y_{\mu,1}, \mu/y_{\mu,2}, ..., \mu/y_{\mu,M-1})^\top$. In (31), we used 0 to represent zero matrices of an appropriate size determined from its context.

For the penalty equation (31), we have the following theorem.

**Theorem 5.** *Eq. (31) has the following properties when $|\Delta t_n|$ is sufficiently small.*

1. *For any $x \in \mathbb{R}^{M-1}$ and $y \in \mathbb{R}^{2(M-1)}$, the system matrix of (31) satisfies*

$$\begin{bmatrix} x^\top & y^\top \end{bmatrix} G^n \begin{bmatrix} x \\ y \end{bmatrix} \geq C\|x\|_2^2 \tag{32}$$

   *for some positive constant $C$.*
2. *Let $(x_\mu^\top, y_\mu^\top)^\top \in \mathcal{K}$ be a solution to (31) for any $\mu$ satisfying $0 < \mu \leq \mu_0$, where $\mu_0$ is a given constant. Then, there exists a constant $L > 0$, independent of $\mu$, such that $\|x_\mu\|_2 + \|y_\mu\|_2 \leq L$.*

**Proof.** We firs prove Item 1. It is easy seen that

$$\begin{bmatrix} x^\top & y^\top \end{bmatrix} G^n \begin{bmatrix} x \\ y \end{bmatrix} = x^\top (H + \alpha E)x - x^\top (F^n)^\top y + y^\top F^n x = x^\top (H + \alpha E)x \geq C\|x\|_2^2$$

by Theorem 4.

Let $L > 0$ be a *generic* positive constant, independent of $\mu$. We rewrite $y_\mu$ as $y_\mu = \begin{bmatrix} y_\mu^{(1)} \\ y_\mu^{(2)} \end{bmatrix}$, where $y_\mu^{(1)}, y_\mu^{(2)} \in \mathbb{R}^{M-1}$. For any $i = 1, 2, ..., M-1$, from the definition of $F^n$ we see that left-multiplying (31) by $(0, ..., y_{\mu,i}^{(1)}, 0, ..., y_{\mu,i}^{(2)}, 0, ...)$ gives

$$y_{\mu,i}^{(1)}[A_i^n(\sigma_A^n, r_A^n)x_\mu - c^n(\sigma_A^n, r_A^n)] = \mu, \quad y_{\mu,i}^{(2)}[A_i^n(\sigma^{*n}, r^{*n})x_\mu - c^n(\sigma^{*n}, r^{*n})] = \mu \tag{33}$$

for $i = 1, 2.., M-1$, where $B_i$ denotes the $i$th row of a matrix $B$. When $\mu \to 0+$, we consider the following two cases.

**Case 1.** If $A_i^n(\sigma_A^n, r_A^n)x_\mu - c^n(\sigma_A^n, r_A^n) \to 0$, it is necessary that $y_{\mu,i}^{(2)} \to 0+$ as both $A_i^n(\sigma_A^n, r_A^n)x_\mu - c^n(\sigma_A^n, r_A^n)$ and $A_i^n(\sigma^{*n}, r^{*n})x_\mu - c^n(\sigma^{*n}, r^{*n})$ cannot approach 0 at the same time, since $(\sigma_A^n, r_A^n) \neq (\sigma^{*n}, r^{*n})$.

**Case 2.** Symmetrically, if $A_i^n(\sigma^{*n}, r^{*n})x_\mu - c^n(\sigma^{*n}, r^{*n}) \to 0$, then $y_{\mu,i}^{(1)} \to 0+$.

Thus, combining these two cases above we have the complementary condition that $y_{\mu,i}^{(1)} y_{\mu,i}^{(2)} \to 0+$ for $i = 1, 2, ..., M-1$, as $\mu \to 0+$. Using this condition, we have that, for $i = 1, 2, ..., M-1$, either $y_{\mu,i}^{(1)} \leq L$ or $y_{\mu,i}^{(2)} \leq L$. This implies there are $(M-1)$ components in $y_\mu$, each of which is bounded above by $L$. We write these $(M-1)$ components as $\bar{y}_\mu \in \mathbb{R}^{M-1}$ which satisfies $\|\bar{y}_\mu\|_2 \leq \sqrt{M-1}L$. We denote the rest of components in $y_\mu$ as $\hat{y}_\mu \in \mathbb{R}^{M-1}$. Therefore, from the definition of $F^n$, we see that the first block of equations in (31) can be written in the following form:

$$\begin{aligned} (H + \alpha E)x_\mu &- A^{n\top}(\sigma_A^n, r_A^n)y_\mu^{(1)} - A^{n\top}(\sigma^{*n}, r^{*n})y_\mu^{(2)} \\ &= (H + \alpha E)x_\mu - \bar{A}^{n\top}\bar{y}_\mu - \hat{A}^{n\top}\hat{y}_\mu = 0, \end{aligned} \tag{34}$$

where $\bar{A}^n$ and $\hat{A}^n$ are $(M-1) \times (M-1)$ matrices containing rows of $A^n(\sigma_A^n, r_A^n)$ and $A^n(\sigma^{*n}, r^{*n})$ corresponding respectively to the rows of $\bar{y}_\mu$ and $\hat{y}_\mu$ in $y_\mu$.

From Theorem 3, we see that both $A^n(\sigma_A^n, r_A^n)$ and $A^n(\sigma^{*n}, r^{*n})$ are positive-definite $M$-matrices when $|\Delta t_n|$ is sufficiently small. Since $\hat{A}^n$ is formed by a combination of $(M-1)$ different rows from $A^n(\sigma_A^n, r_A^n)$ and $A^n(\sigma^{*n}, r^{*n})$, $\hat{A}^n$ should also be an $M$-matrix since it is still irreducibly diagonally dominant. Note $\hat{A}^n$ also contains the diagonal matrix $\frac{1}{|\Delta t_n|}I$ from the definition of $A^n$ in (22), it should be positive-definite as well when $|\Delta t_n|$ is sufficiently small. Thus, we conclude that $\hat{A}^n$ is a positive-definite $M$-matrix. Now, left-multiplying (34) with $\hat{y}_\mu^\top$, re-arranging the resulting equation and using the positive-definiteness of $\hat{A}^n$, we have

$$C\|\hat{y}_\mu\|_2^2 \le L\hat{y}_\mu^\top \hat{A}^{n\top}\hat{y}_\mu = \hat{y}_\mu^\top\left[(H+\alpha E)x_\mu - \bar{A}^{n\top}\bar{y}_\mu\right] \le L\|\hat{y}_\mu\|_2\left(\|x_\mu\|_2+1\right).$$

Thus, combining $\|\bar{y}_\mu\|_2 \le \sqrt{M-1}L$ and the above estimate, we obtain $\|y_\mu\|_2 \le L\left(\|x_\mu\|_2+1\right)$.

Left-multiplying (31) with $(x_\mu^\top, y_\mu^\top)$ gives

$$\begin{bmatrix} x_\mu^\top & y_\mu^\top \end{bmatrix} G^n \begin{bmatrix} x_\mu \\ y_\mu \end{bmatrix} - \mu y_\mu./y_\mu = y_\mu^\top f^n. \tag{35}$$

Since $y_\mu./y_\mu = 2(M-1)$, using (32) and the estimate for $\|y_\mu\|_2$ obtained above, we have from (35)

$$C\|x_\mu\|_2^2 \le 2\mu(M-1) + \|f^n\|_2\|y_\mu\|_2 \le L(1+\|x_\mu\|_2).$$

It is trivial to show that the above inequality implies $\|x_\mu\|_2^2 \le L$. (Recall $L>0$ is a generic constant, independent of $\mu$.) □

Using Theorem 5, we have the following result.

**Theorem 6.** *Let the conditions in Theorem 5 be fulfilled. The penalty equation (31) has a unique solution in $\mathcal{K}$.*

The proof of this theorem is essentially a rewriting of that of Theorem 2.3 in [26]. Thus, we omit this proof and refer the reader to [26].

### 4.2. Convergence

The following theorem establishes an upper bound as a function of $\mu$ for the $\ell_2$-norm of the difference between the solutions to (31) and (29)–(30), respectively.

**Theorem 7.** *Let $(x^\top, y^\top)^\top$ be the solution to (29)–(30) and $(x_\mu^\top, y_\mu^\top)^\top \in \mathcal{K}$ the solution to (31). Then, there exists a constant $L>0$, independent of $\mu$, such that*

$$\|x_\mu - x\|_2 + \|y_\mu - y\|_2 \le L\sqrt{\mu}. \tag{36}$$

**Proof.** We first show that $(x_\mu^\top, y_\mu^\top)^\top$ satisfies (30) when $\mu \to 0+$. Since $y_\mu > 0$, from the 2nd block of equations in (31) we see tat $F^n x_\mu > f^n$, and thus the inequalities in (30) are satisfied by $z_\mu$. From (33), we have that

$$y_\mu^\top(F^n x_\mu - f^n) = 2(M-1)\mu, \tag{37}$$

and so, when $\mu \to 0+$, $y_\mu^\top(F^n x_\mu - f^n) = 0$, which is the equality in (30).

Let $L$ denote generic positive constant, independent of $\mu$. For any $u \in \mathbb{R}^{M-1}$, left-multiplying (29) by $(u-x)^\top$ yields

$$(u-x)^\top[(H+\alpha E)x - F^{n\top}y] = 0. \tag{38}$$

Similarly, for any $v \in \mathbb{R}^{2(M-1)}$ satisfying $v>0$, we have

$$(v-y)^\top(F^n x - f^n) = v^\top(F^n x - f^n) - y^\top(F^n x - f^n) \ge 0,$$

since $F^n x - f^n \ge 0$, $y^\top(F^n x - v^\top f^n) = 0$ by (30) and $v>0$. Combining the above inequality with (38) gives

$$\left(\begin{bmatrix} u^\top & v^\top \end{bmatrix} - \begin{bmatrix} x^\top & y^\top \end{bmatrix}\right)\left(G^n\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 0 \\ f^n \end{bmatrix}\right) \ge 0. \tag{39}$$

Now, left-multiplying (31) with $[(x_\mu - x)^\top \ (y_\mu - y)^\top]$, we have

$$\begin{bmatrix} (x_\mu - x)^\top & (y_\mu - y)^\top \end{bmatrix}\left(G^n\begin{bmatrix} x_\mu \\ y_\mu \end{bmatrix} - \begin{bmatrix} 0 \\ f^n \end{bmatrix}\right) - (y_\mu^\top - y^\top)\mu./y_\mu = 0. \tag{40}$$

Replacing $u$ and $v$ in (39) with $x_\mu$ and $y_\mu$ respectively and taking both sides of the resulting inequality from the corresponding sides of (40) gives

$$\begin{bmatrix} (x_\mu - x)^\top & (y_\mu - y)^\top \end{bmatrix} G^n \begin{bmatrix} x_\mu - x \\ y_\mu - y \end{bmatrix} - \mu y_\mu./y_\mu + \mu y./y_\mu \le 0.$$

Since $y \le 0$ and $y_\mu < 0$, $\mu y./y_\mu > 0$. Also, $y_\mu./y_\mu = 2(M-1)$. Thus, using (32), we obtain from the above inequality $C\|x_\mu - x\|_2^2 \le 2(M-1)\mu$, yielding

$$\|x_\mu - x\|_2 \le L\sqrt{\mu}. \tag{41}$$

We now prove the convergence of $y_\mu$. As in the proof of Theorem 5, we introduce the partition $y_\mu = ((y_\mu^{(1)})^\top, (y_\mu^{(2)})^\top)^\top$. Similarly, we write $y$ as $y = ((y^{(1)})^\top, (y^{(2)})^\top)^\top$. Taking both sides of the 1st block of equations in (31) from the corresponding sides of (29) yields

$$(H + \alpha E)(x - x_\mu) + A^{n\top}(\sigma_A^n, r_A^n)(y_\mu^{(1)} - y^{(1)}) + A^{n\top}(\sigma^{*n}, r^{*n})(y_\mu^{(2)} - y^{(2)}) = 0.$$

From the proof of Item 2 of Theorem 5, particularly (33), we see that $y_{\mu,i}^1 y_{\mu,i}^2 \le \mathcal{O}(\mu)$ for $i = 1, ..., M-1$. Similarly, using the complementarity condition in (30) we can also easily show that $y^1$ and $y^2$ satisfy the complementary slackness condition $y_i^{(1)} y_i^{(2)} = 0$ for all feasible $i$. Furthermore, it is easy to see that the combined complementary slackness condition is $(y_i^1 + y_{\mu,i}^1)(y_i^2 + y_{\mu,i}^2) \le \mathcal{O}(\mu)$. Therefore, if we use the matrices $\bar{A}^n$ and $\hat{A}^n$ introduced in the proof of Item 2 of Theorem 5, we may rewrite the above equation in the following form

$$(H + \alpha E)(x - x_\mu) + \bar{A}^{n\top}(\bar{y}_\mu - \bar{y}) + \hat{A}^{n\top}(\hat{y}_\mu - \hat{y}) = 0,$$

where $\bar{y}_\mu - \bar{y} \in \mathbb{R}^{(M-1)}$ contains $(M-1)$ entries and $\hat{y}_\mu - \hat{y} \in \mathbb{R}^{(M-1)}$ contains the remaining ones from $y_\mu - y$ satisfying $\hat{y}_{\mu,i} + \hat{y}_i \le \mathcal{O}(\mu)$. Left-multiplying the above equality with $(\bar{y}_\mu - \bar{y})^\top$ and reorganising the resulting equality, we have

$$\begin{aligned}(\bar{y}_\mu - \bar{y})^\top \bar{A}^{n\top}(\bar{y}_\mu - \bar{y}) &= (\bar{y}_\mu - \bar{y})^\top \left[(H + \alpha E)(x - x_\mu) - \hat{A}^{n\top}(\hat{y}_\mu - \hat{y})\right] \\ &\le L\|\bar{y}_\mu - \bar{y}\|_2 \left(\|x - x_\mu\|_2 + \|\hat{y}_\mu - \hat{y}\|_2\right). \end{aligned} \tag{42}$$

(Recall $L > 0$ is a generic positive constant, independent of $\mu$.) In the proof of Theorem 5, we showed that $\bar{A}^n$ is a positive-definite $M$-matrix. Also, $\|\hat{y}_\mu - \hat{y}\|_2 \le (M-1)^{1/2}\mathcal{O}(\mu)$ since $\hat{y}_{\mu,i} + \hat{y}_i \le \mathcal{O}(\mu)$ for all feasible $i$. Thus, using the positive-definiteness of $\bar{A}^n$, this estimate and (41), we have from (42)

$$\|\bar{y}_\mu - \bar{y}\|_2 \le L(\sqrt{\mu} + (M-1)^{1/2}\mu) \le L\sqrt{\mu},$$

when $\mu > 0$ is sufficiently small. Combining the above estimate and (41) gives (36). □

### 4.3. Algorithm

We rewrite (31) as the following a nonlinear system:

$$W^n(x_\mu, y_\mu) := G^n \begin{bmatrix} x_\mu \\ y_\mu \end{bmatrix} - \begin{bmatrix} 0 \\ \mu./y_\mu \end{bmatrix} - \begin{bmatrix} 0 \\ f^n \end{bmatrix} = 0 \tag{43}$$

and consider the numerical solution of (43) in $\mathcal{K}$. It is easy to see that $W$ a smooth function in $\mathcal{K}$. Thus, in this subsection, we propose a algorithm based on the conventional Newton's iterative method for nonlinear system of equations. Note the computed option price $x_\mu$ should be non-negative. Thus, we introduce a sufficiently small positive parameter $\delta$ and look for a solution to (43) in $\hat{\mathcal{K}} = \{(u^\top, v^\top)^\top \in \mathcal{K} : u_i \ge 0, v_j \ge \delta, i = 1, ..., M-1, j = 1, ..., 2(M-1)\}$. Also, since the Jacobian matrix of $P$, i.e. $G^n + \mu./(y_\mu.*y_\mu)$, may be ill-conditioned in some of the Newton iterations, we introduce a (small) positive regularity parameter $\rho$, similar to that in Lavenberg-Marquardt algorithm [18,21], and replace the Jacobian matrix with $G^n + \mu./(y_\mu.*y_\mu) + \rho I$. The algorithm is given below.

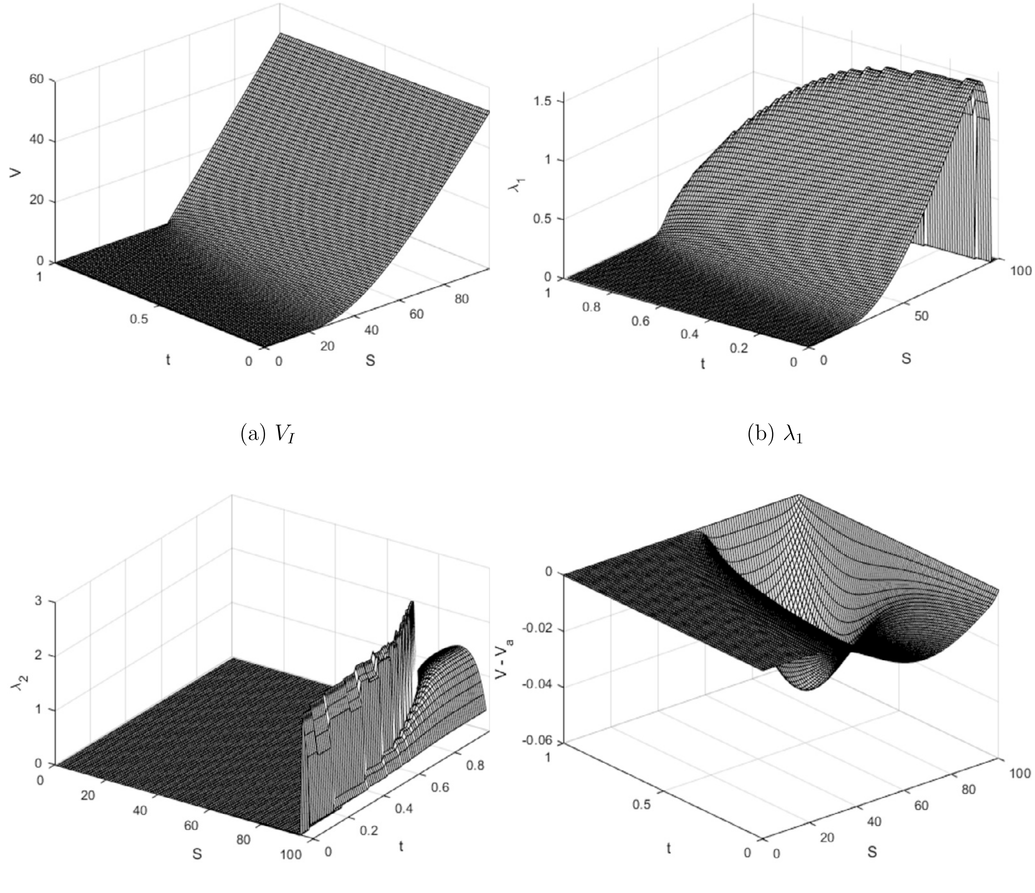
Choose parameters $\mu, \alpha, \delta, \rho, \varepsilon \in (0,1)$ with $\varepsilon$ the tolerance in the Newton's iterative scheme. Calculate the payoff and boundary conditions $x_i^0$ for $i = 1, 2, ..., M-1$, $x_0^j$ and $x_M^j$ for $j = 1, 2, ..., N$. Let $y_i^0 = 1$ for $i = 1, 2, ..., 2(M-1)$.

**for** $n = 1$ *to* $N$ **do**
- $\hat{x}^0 \leftarrow x^{n-1}$, $\hat{y}^0 \leftarrow y^{n-1}$,
- **for** $k = 1, 2, ...$ **do**
  - Evaluate $\sigma_A$, $\sigma^*$, $r_A$ and $r^*$ using (17)–(18) .
  - Calculate $A^n(\sigma_A.r_A)$, $A^n(\sigma^*, r^*)$, $c^n(\sigma_A, r_A)$ and $c^n(\sigma^*, r^*)$ using (22)–(23).
  - Solve $\left[G^n + \mu./(y^{k-1}.*y^{k-1}) + \rho I\right]\begin{bmatrix} p_x \\ p_y \end{bmatrix} = -W(\hat{x}^{k-1}, \hat{y}^{k-1})$ for $\begin{bmatrix} p_x \\ p_y \end{bmatrix}$.
  - $\hat{x}^k \leftarrow \max\{\hat{x}^{k-1} + p_x, 0\}$ and $\hat{y}^k \leftarrow \max\{\hat{y}^{k-1} + p_y, \delta\}$ .
  - Evaluate $Err = \max\left\{\left\|\begin{bmatrix} p_x \\ p_y \end{bmatrix}\right\|_2, \left|\hat{x}^{k\top}(F^n\hat{x}^k - f^n) - 2(M-1)\mu\right|\right\}$, where $F^n$ and $f^n$ are defined in (30).
  - **if** $Err < \varepsilon$ **then**
    - Break from loop $k$
  - **end**
- **end**
- $(x^n, y^n) \leftarrow (\hat{x}^k, \hat{y}^k)$ .

**end**

**Algorithm 1:** Solution of the option pricing problem.

We comment that in the stopping criterion, we require that both the $\ell_2$-norm of the correction of the current Newton iterate and the complementarity condition (37) to be smaller than the tolerance $\varepsilon$.




(a) $V_I$ (b) $\lambda_1$

(c) $\lambda_2$ (d) $V_I - V_A$

**Fig. 1.** Computed $V$, $\lambda_1$, $\lambda_2$ and $V - V_A$ for Test 1.

## 5. Numerical experiments

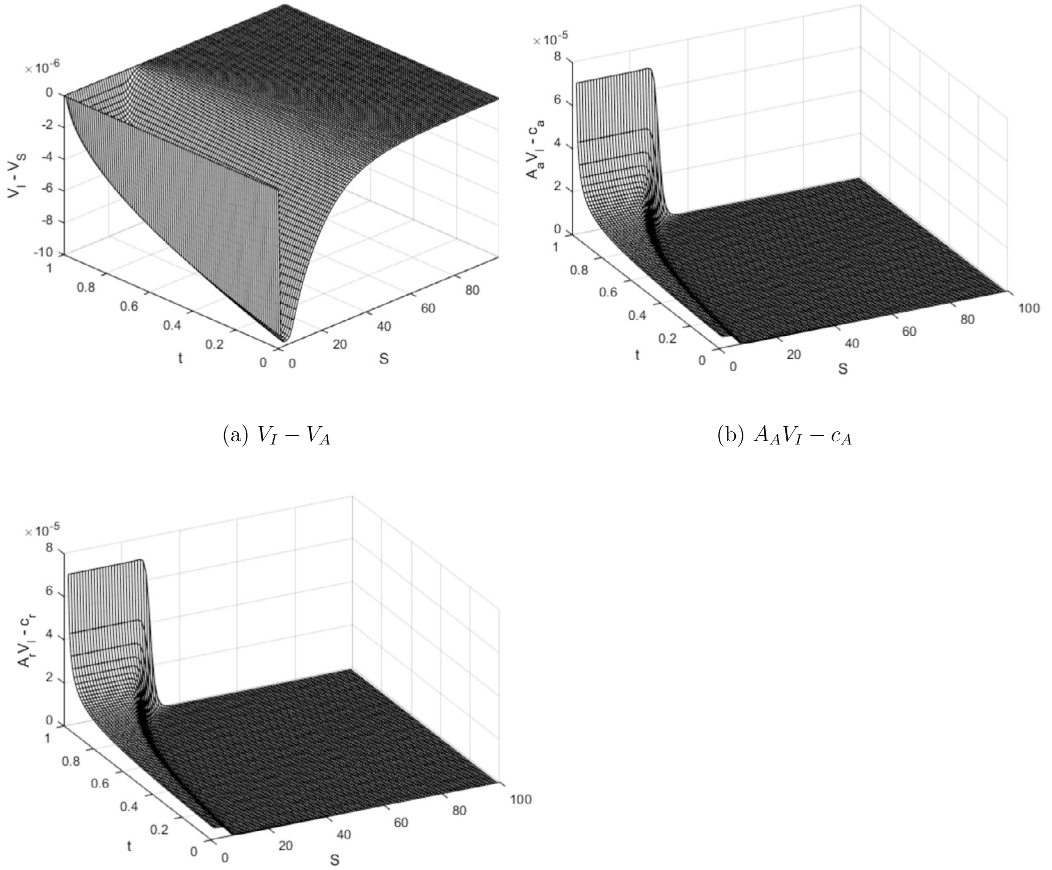
In this section, we numerically price some model European options using the above pricing model and numerical methods. In all of the numerical experiments in this section, the parameter in Algorithm 1 are chosen to be $\mu = 10^{-12} = \delta$, $\alpha = 1/S_{\max}^2$, $\rho = 10^{-9}$ and $\epsilon = 10^{-8}$. All numerical experiments have been carried out under Matlab programming environment.

**Test 1.** European call option with $K = 50$, $S_{\max} = 100$ and $T = 1$. The volatility and interest rate are chosen to be constant interval numbers $\sigma = [0.3, 0.6]$ and $r = [0.03, 0.05]$, respectively.

To solve this problem, we choose the uniform mesh for $(0, S_{\max}) \times (0, 1)$ with mesh sizes $h_i = 1/M$ and $|\Delta t_n| = 1/N$ for all feasible $i$ and $n$, where $M = 100 = N$. The penalty equation (31) (or (43)) is solved using Algorithm 1. The computed option value $V_I$ and the multipliers $\lambda_1$ and $\lambda_2$ are depicted in Fig. 1. From the plots of $\lambda_1$ and $\lambda_2$ we see that the constraint (20) is active and (21) is in-active in a large portion of the solution region. We also solve this pricing problem as a conventional European option pricing one with non-interval coefficients $\sigma = 0.45$ and $r = 0.04$ and the solution is denoted as $V_A$. In Fig. 1(d), we plot the difference $V_I - V_A$, from which we see that the two solutions are different, though they are close to each other.

We now use this test problem to numerically verify the result in Theorem 1. We assume $\sigma = [0.45, 0.45]$ and $r = [0.04, 0.04]$, and other parameters are the same as in Test 1. This problem is solved by Algorithm 1, and the numerical solution is denoted as $V_I$. The difference $V_I - V_A$, $A^n(\sigma_A, r_A)V_I - c_A$ and $A^n(\sigma_R, r_R)V_I - c_R$ for $n = 0, 1, ..., N$ are depicted in Fig. 2, from which we see that when the interval numbers reduce to a real ones, the solution $V_I$ becomes that to the conventional Black-Scholes equation, $V_A$, with constant $\sigma$ and $r$ (up to an approximation error). Also, the inequality constraints in (20)–(21) become the same equality, up to an approximation error.

**Test 2.** European call option with $K = 50$, $S_{\max} = 100$, $T = 1$. The volatility is a time-dependent interval number $\sigma = [0.5 - 0.2(1 + 0.2\text{rand}), 0.5 - 0.2(1 + 0.2\text{rand})]$, where 'rand' is the random number generator in Matlab. The interest rate is chosen to be time-dependent interval number defined by $r(t) = [0.05 + 0.02\sin(5t) - 0.01(1 + 0.2\text{rand}), 0.05 + 0.02\sin(5t) + 0.01(1 + 0.2\text{rand})]$.




(a) $V_I - V_A$

(b) $A_A V_I - c_A$

(c) $A_R V_I - c_R$

**Fig. 2.** Computed $V_I - V_A$, $A_A V_I - c_A$ and $A_R V_I - c_R$ for Test 1.

To solve this test problem, we use the same uniform mesh used for Test 1, and on this mesh, we generate sequences of interval numbers $[\sigma_L(t_n), \sigma_R(t_n)]$ and $[r_L(t_n), r_R(t_n)]$ for $n = 0, 1, \ldots, N$ using the functions defined above. The discrete optimization problem is solved using Algorithm 1, and the computed option price, and the multipliers are denoted as $V_I$, $\lambda_1$ and $\lambda_2$ respectively. For comparison, we also solve this problem using non-interval $\sigma_A(t_n) = 0.5(\sigma_L(t_n) + \sigma_R(t_n))$ and $r_A(t_n) = 0.5(r_L(t_n) + r_R(t_n))$ for $n = 0, 1, \ldots, N$, and the numerical solution is denoted as $V_A$. The computed $V_I$, $\lambda_1$, $\lambda_2$ and the difference $V_I - V_A$ are plotted in Fig. 3. Comparing Fig. 3 with Fig. 1 we see that both problems have similar behaviours, i.e. the constraints (20)–(21) are inactive when $S$ is small and active as $S$ increases. From Fig. 3(d) we also see, as in Test 1, $V_I$ and $V_A$ are different, but close to each other.

## 6. Conclusions

In this work, we formulate the valuation of European call options with uncertain volatility and interest rate as an infinite-dimensional interval optimization problem. We propose an interior penalty method with a penalty constant $\mu \in (0, 1)$ for the numerical solution of the discretized optimization problem and prove that the solution to the penalty equation converges to that of the KKT conditions of the discretized optimization problem at the rate $\mathcal{O}(\mu^{1/2})$. An algorithm based on Newton's iterative method is proposed for solving the nonlinear penalty equation. Numerical results are presented to show that the solutions from this novel optimization-based pricing model is close, but not equal, to the option price obtained by taking the averages of uncertain volatility and interest rate. Numerical results also show that, when the interval volatility and interest rate become real numbers, i.e., the widths of the intervals become zero, the price of an option from this new model become, up to an approximation error, that from the conventional Black-Scholes' model.

## Data availability

No data was used for the research described in the article.

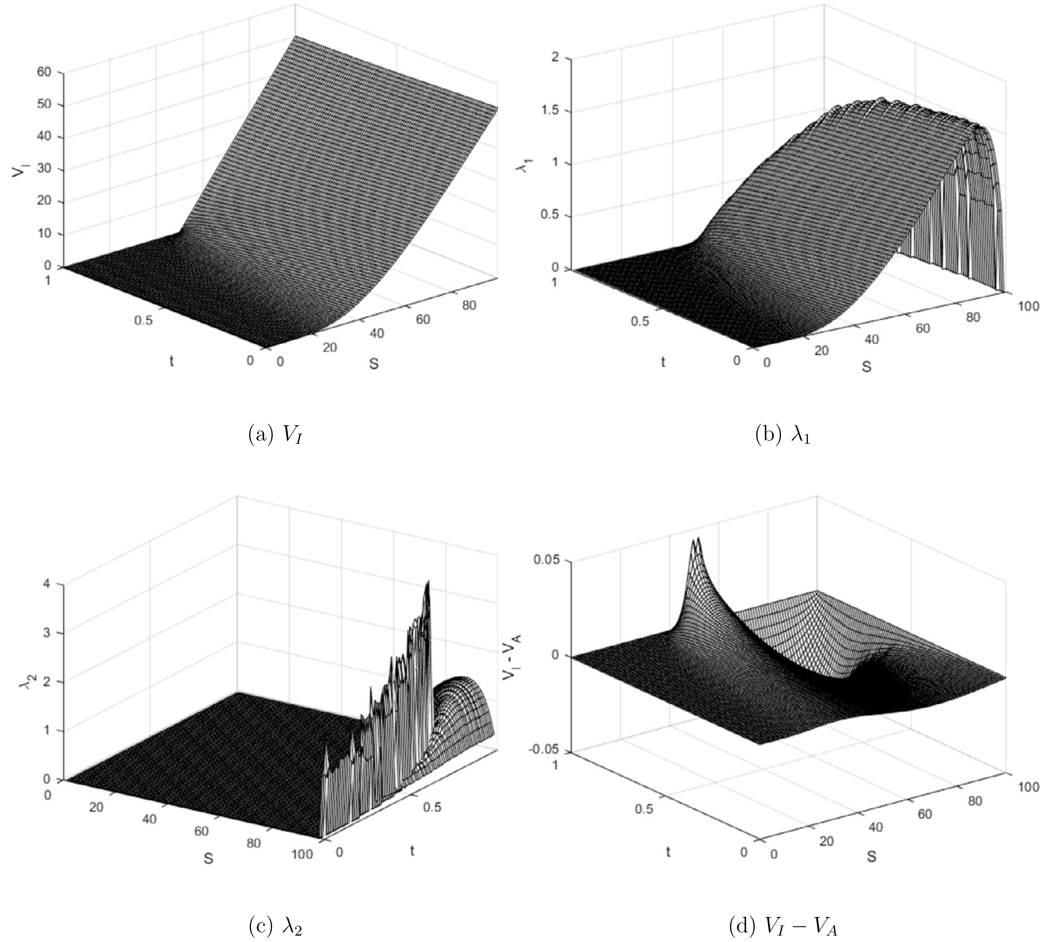



**Fig. 3.** Computed $V_I$, $\lambda_1$, $\lambda_2$ and $V_I - V_A$ for Test 2.

## Acknowledgement

This work is partially supported by the Australian Research Council (DP190103361).

## References

[1] M. Avellaneda, A. Levy, A. Paras, Pricing and hedging derivatives securities in markets with uncertain volatilities, Appl. Math. Finance 2 (1995) 73–88.
[2] A.K. Bhurjee, G. Panda, Efficient solution of interval optimization problem, Math. Methods Oper. Res. 76 (2012) 273–288.
[3] L. Bian, Z. Li, Fuzzy simulation of European option pricing using sub-fractional Brownian motion, Chaos Solitons Fractals 153 (2021) 111442.
[4] F.S. Black, M.S. Scholes, The pricing of options and corporate liabilities, J. Polit. Econ. 81 (1973) 637–654.
[5] A. Bryla, On solving an optimization problem with interval coefficients, in: S. Butenko, P.P.M. Pardolas, V. Shylo (Eds.), Optimization Methods and Applications, in: Springer Optimization & Its Application, Springer, 2017.
[6] S. Chanas, D. Kuchta, Multiobjective programming in optimization of interval objective functions – a generalized approach, Eur. J. Oper. Res. 94 (1996) 594–598.
[7] W. Chen, S. Wang, A penalty method for a fractional order parabolic variational inequality governing American put option valuation, Comput. Math. Appl. 67 (2014) 77–90.
[8] J. Cheng, Z. Liu, M. Tang, J. Tan, Robust optimization of uncertain structures based on normalized violation degree if interval constraint, Comput. Struct. 182 (2017) 41–54.
[9] A. Forsgren, P.E. Gill, M.H. Wright, Interior methods for nonlinear optimization, SIAM Rev. 44 (2002) 525–597.
[10] M. Garman, A general theory of asset valuation under diffusion state processes, Working paper No. 50, University of California, Berkeley, 1976.
[11] B. Hu, S. Wang, A novel approach in uncertain programming part I: new arithmetic and order relations for interval numbers, J. Ind. Manag. Optim. 2 (2006) 351–371.
[12] C.C. Huang, S. Wang, A power penalty approach to a nonlinear complementarity problem, Oper. Res. Lett. 38 (2010) 72–76.
[13] C.C. Huang, S. Wang, A penalty method for a mixed nonlinear complementarity problem, Nonlinear Anal. 75 (2012) 588–597.
[14] J.C. Hull, A. White, The pricing of options on assets with stochastic volatilities, J. Finance 42 (1987) 281–300.
[15] H. Ishibuchi, H. Tanaka, Multiobjective programming in optimization of interval objective function, Eur. J. Oper. Res. 48 (1990) 219–225.
[16] C.-F. Lee, G.-H. Tzeng, S.-Y. Wang, A new application of fuzzy set theory to the Black-Scholes option pricing model, Expert Syst. Appl. 29 (2005) 330–342.
[17] D.C. Lesmana, S. Wang, An upwind finite difference method for a nonlinear Black-Scholes equation governing European option valuation under transaction costs, Appl. Math. Comput. 219 (2013) 8811–8828.

[18] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Q. Appl. Math. 2 (1944) 164–168.
[19] S.-T. Liu, R.-T. Wang, A numerical solution method to interval quadratic programming, Appl. Math. Comput. 189 (2007) 1274–1281.
[20] X. Liu, Z. Zhang, L. Yin, A multi-objective optimization method for uncertain structure based on nonlinear interval number programming method, Mech. Based Des. Struct. Mach. 45 (2016) 25–42.
[21] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM J. Appl. Math. 11 (1963) 413–441.
[22] R.S. Varga, Matrix Iterative Analysis, Prentice-Hall, Engelwood Cliffs, NJ, 1962.
[23] C. Vazquez, Am upwind numerical approach for an American and European option pricing model, Appl. Math. Comput. 97 (1998) 273–286.
[24] S. Wang, Numerical solution of an obstacle problem with interval coefficients, Numer. Algebra Control Optim. 10 (2020) 22–38.
[25] K. Zhang, S. Wang, A computational scheme for uncertain volatility model in option pricing, Appl. Numer. Math. 59 (2009) 1754–1767.
[26] S. Wang, An interior penalty method for a large-scale finite-dimensional nonlinear double obstacle problem, Appl. Math. Model. 58 (2018) 217–228.
[27] S. Wang, K. Zhang, An interior penalty method for a finite-dimensional linear complementarity problem in financial engineering, Optim. Lett. 12 (2018) 1161–1178.
[28] P. Wilmott, J. Dewynne, J. Howison, Option Pricing: Mathematical Models and Computation, Oxford Financial Press, Oxford, 1993.
[29] H.-C. Wu, Using fuzzy sets theory and Black–Scholes formula to generate pricing boundaries of European options, Appl. Math. Comput. 185 (2007) 136–146.