ON A SEMILINEAR BLACK AND SCHOLES PARTIAL DIFFERENTIAL EQUATION FOR VALUING AMERICAN OPTIONS. PART II: APPROXIMATE SOLUTIONS AND CONVERGENCE

FRED E. BENTH, KENNETH H. KARLSEN, AND KRISTIN REIKVAM

ABSTRACT. In [7], we proved that the American (call/put) option valuation problem can be stated in terms of one single semilinear Black and Scholes partial differential equation set in a fixed domain. The semilinear Black and Scholes equation constitutes a starting point for designing and analyzing a variety of "easy to implement" numerical schemes for computing the value of an American option. To demonstrate this feature, we propose and analyze an upwind finite difference scheme of "predictor–corrector type" for the semilinear Black and Scholes equation. We prove that the approximate solutions generated by the predictor–corrector scheme respect the early exercise constraint and that they converge uniformly to the the American option value. A numerical example is also presented. Besides the predictor–corrector schemes, other methods for constructing approximate solution sequences are discussed and analyzed as well.

CONTENTS

1. INTRODUCTION

Let $T > 0$ be fixed and $t < T$. Suppose that the price dynamics of a dividend paying stock X_s is governed by a geometric Brownian motion (under the unique equivalent martingale measure Q), i.e., it evolves according to the stochastic differential equation

$$dX(s) = (r - d)X(s)\,ds + \sigma X(s)\,dW(s), \qquad s \in (t, T],$$

where $d \geq 0$ is the constant dividend yield for the stock, $r \geq 0$ is the risk-free interest rate, $\sigma > 0$ is the volatility, and $\{W(s) \mid s \in [0, T]\}$ is a standard Brownian motion. Starting at time t with initial condition $X(t) = x$, it is well known that the arbitrage-free value of an American option with expiration at time T is given by

$$(1.1) \qquad V(t, x) = \sup_{t \leq \tau \leq T} \mathbb{E}^{t,x}\left[e^{-r(\tau - t)} g(X(\tau))\right],$$

where the supremum is taken over all $\mathcal{F}_t$ stopping times $\tau \in [t, T]$, $\mathbb{E}^{t,x}$ denotes expectation under the equivalent martingale measure conditioned on $X(t) = x$, and $g : \mathbb{R} \to \mathbb{R}$ is the payoff function. Herein we will focus on call and put options, i.e. options with payoff,

$$(1.2) \qquad g(x) = \begin{cases} (x - K)^+, & \text{call option}, \\ (K - x)^+, & \text{put option}, \end{cases}$$

where $K > 0$ is the strike price. In this paper, we use Duffie [10], Karatzas and Shreve [17], Musiela and Rutkowksi [19], and Myneni [20] as general references on the American option valuation problem.

In the literature one can find two main approaches for determining the function V in (1.1): (i) The *free boundary problem* formulation. (ii) The *quasi-variational inequality* formulation. It is well known that there is no explicit formula for V, as opposed to the value of an European option for which an analytical formula exists. Consequently, with both approaches one has to resort to numerical schemes for finding V. However, the two approaches lend themselves to different numerical schemes. We refer to [10, 17, 19, 20] (and also the references therein) for mathematical and numerical aspects of the free boundary problem and quasi-variational inequality formulations as well as their advantages and disadvantages.

In [7], we presented a new approach for determining the value of an American option. The function V in (1.1) is the value function of an optimal stopping problem for which the dynamic programming principle holds [24]. Using the dynamic programming principle, we proved in [7] that V uniquely solves (in a viscosity solution sense) the following *semilinear Black and Scholes equation* set in a fixed domain:

$$(1.3) \qquad \partial_t v + (r - d)x\partial_x v + \frac{1}{2}\sigma^2 x^2 \partial_x^2 v - rv = -q(x, v), \qquad (t, x) \in Q_T,$$

where Q_T denotes the time-space cylinder $[0, T) \times [0, \infty)$. The *nonlinear reaction term* $q : \mathbb{R} \times \mathbb{R} \to [0, \infty)$ takes the form

$$(1.4) \qquad q(x, v) = c(x)H(g(x) - v),$$

with $c : \mathbb{R} \to [0, \infty)$ being the "cash flow" function

$$(1.5) \qquad c(x) = \begin{cases} (dx - rK)^+, & \text{call option,} \\ (rK - dx)^+, & \text{put option,} \end{cases}$$

and $H : \mathbb{R} \to \{0, 1\}$ the Heaviside function

$$(1.6) \qquad H(\xi) = \begin{cases} 0, & \xi < 0, \\ 1, & \xi \geq 0. \end{cases}$$

We augment (1.3) with the terminal condition

$$(1.7) \qquad v(T, x) = g(x), \qquad x \in [0, \infty),$$

where g is the payoff function in (1.2). One should note that in (1.3) there is no free boundary that need to be computed as part of the solution nor are there "side constraints" that need to be satisfied by the solution (as opposed to the quasi-variational formulation). We refer to [7] for the motivation behind the semilinear Black and Scholes equation (1.3) and its rigorous derivation from (1.1) via the dynamic programming principle. We also refer to [7] for an overview of relevant literature. Here we only mention that (1.3) can be viewed as an infinitesimal (partial differential equation) version of the well known *early exercise premium representation of the American option* (i.e., the separation of the American option price into the corresponding European option price plus an early exercise premium) and that an initial motivation for our work in [7] was [18].

It was pointed out in [7] that the semilinear Black and Scholes equation does not make sense as it stands in (1.3) if classical (i.e., $C^{1,2}$) solutions — or more generally continuous viscosity solutions [9, 11] — are sought. This is related to the fact that the nonlinearity $v \mapsto q(x, v)$ in (1.3) is *discontinuous*. Guided by the dynamic programming principle, we suggested in [7] a suitable definition of a *viscosity solution* for (1.3). We recall this definition in Section 2 of the present paper. It was proved in [7] that the function V defined in (1.1) is the unique such viscosity solution of (1.3) satisfying the terminal condition (1.7). In other words, the terminal value problem (1.3)–(1.7) constitutes an alternative formulation of the American (call/put) option valuation problem.

From a practical point of view, the advantage with the semilinear Black and Scholes equation (1.3) is that it suggests a natural recipe for constructing "easy to implement" numerical schemes for valuing American options. In fact, any European option solver can be turned into an American option solver via (1.3). In Section 4, we device an upwind finite difference scheme of "predictor–corrector type" for (1.3). From (1.1), it is seen that the American option value V always satisfies the *early exercise constraint*

$$(1.8) \qquad V \geq g \text{ on } \overline{Q_T}.$$

The suggested predictor–corrector scheme automatically satisfies a discrete analog of (1.8). Using the the Barles–Ishii–Perthame weak viscosity limit method [9, 11], we give an easy L^∞_{loc} (i.e., uniform on compacta) convergence proof for the predictor–corrector scheme. A numerical example is presented in Section 6.

We must stress that the particular numerical scheme studied herein is chosen just for its simplicity in terms of presentation, mathematical analysis, and implementation. But, at the same time, this simple choice illustrates the basic advantages of using the semilinear Black and Scholes equation (1.3) as the governing partial differential equation for American option valuation. In practical applications, however, one should use numerical schemes that result from a more sophisticated discretization of (1.3) than the one used herein. We leave this as a challenging topic for future work.

In addition to the numerical scheme, we also analyze various sequences of "semi–discrete approximate solutions", which are obtained by solving "approximate semilinear Black and Scholes equations" in which the discontinuous right-hand side q in (1.3) has been replaced by some continuous approximation. This is the topic of Section 3. Related to this, let us mention that some of the so-called *penalty schemes* found in recent computational finance literature (see, e.g., [12, 21, 25]) can be interpreted as numerical schemes obtained by discretizing an approximate semilinear Black and Scholes equation. A consequence of this point of view is that one can adopt the techniques developed herein to give easy convergence proofs for penalty schemes. We make a further remark about this in Section 4.

To demonstrate that the mathematical framework herein can be used to analyze other numerical schemes for valuing American options as well, we give in Section 5 a rather elementary convergence proof for a numerical scheme based on the classical Brennan and Schwartz algorithm [8] and the finite difference approach.

Finally, we mention that future work will be devoted to extending the semilinear Black and Scholes equation and its mathematical/numerical theory to the multi-asset setting.

2. Viscosity solutions

In this section, we recall from [7] the definition of a viscosity solution for the semilinear Black and Scholes equation. For a general introduction to the viscosity solution theory, we refer to Crandall, Ishii, and Lions [9] and Fleming and Soner [11]. We also recall from [9, 11] how to perform weak limit operations with viscosity solutions. This will be needed later when we prove convergence of various sequences of approximate solutions.

Before stating the definition of a viscosity solution, we need to introduce some notation. We start with the following spaces of semicontinuous functions:

$$USC\left(\overline{Q_T}\right) = \left\{v : \overline{Q_T} \to \mathbb{R} \cup \{-\infty\} \,\middle|\, v \text{ is upper semicontinuous}\right\},$$

$$LSC\left(\overline{Q_T}\right) = \left\{v : \overline{Q_T} \to \mathbb{R} \cup \{+\infty\} \,\middle|\, v \text{ is lower semicontinuous}\right\}.$$

Next, $C^{\alpha,\beta}\left(\overline{Q_T}\right)$ denotes the space of functions on $\overline{Q_T}$ that are $\alpha \geq 0$ times continuously differentiable in t and $\beta \geq 0$ times continuously differentiable in x. The space of continuous functions on Q_T is denoted by $C\left(\overline{Q_T}\right)$, i.e., $C\left(\overline{Q_T}\right) = C^{0,0}\left(\overline{Q_T}\right)$.

In what follows, we will need the nonlinear functions $q^*, q_* : \mathbb{R} \times \mathbb{R} \to [0, \infty)$ defined by

$$(2.1) \qquad q^*(x,v) = c(x)H^*(g(x) - v), \qquad q_*(x,v) = c(x)H_*(g(x) - v),$$

where $H^*, H_* : \mathbb{R} \times \mathbb{R} \to [0, \infty)$ are defined as

$$H^*(\xi) = \begin{cases} 0, & \xi < 0, \\ 1, & \xi \geq 0, \end{cases} \qquad H_*(\xi) = \begin{cases} 0, & \xi \leq 0, \\ 1, & \xi > 0. \end{cases}$$

Remark. Observe that q^* and q_* are respectively upper and lower semicontinuous on $\mathbb{R} \times \mathbb{R}$. Moreover, we have $q_* \leq q^*$ and $q^* \equiv q$. In fact, q^* and q_* are respectively the upper and lower semicontinuous envelopes of the nonlinear function q defined in (1.6) (see [7]).

The fact that the mapping $v \mapsto q(x, v)$ is discontinuous makes it a non-trivial matter to decide what one should mean by a viscosity solution or even a classical $C^{1,2}$ solution of (1.3). Using the dynamic programming principle in optimal stopping theory, we show in [7] that the notion of viscosity solution found in Definition 2.1 below is the natural one; that is, it identifies the American option value as the unique viscosity solution of the terminal value problem (1.3)–(1.7). Remarkably, it turns out that this definition of a viscosity solution is an adaption to our setting of the one used by Ishii [13] for first order Hamilton-Jacobi equations with discontinuous Hamiltonians.

Definition 2.1. (i) A locally bounded function $v \in USC\left(\overline{Q_T}\right)$ is a *viscosity subsolution* of (1.3) if and only if $\forall \phi \in C^{1,2}\left(\overline{Q_T}\right)$ we have:

(2.2) $\quad \begin{cases} \text{for each } (t, x) \in Q_T \text{ being a local maximizer of } v - \phi, \\ \partial_t \phi(t, x) + (r - d)x\partial_x \phi(t, x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \phi(t, x) - rv(t, x) + q^*(x, v(t, x)) \geq 0. \end{cases}$

If, in addition, $v|_{t=T} \leq g$ on $[0, \infty)$, then v is a viscosity subsolution of (1.3)–(1.7).
(ii) A locally bounded function $v \in LSC\left(\overline{Q_T}\right)$ is a *(viscosity) supersolution* of (1.3) if and only if $\forall \phi \in C^{1,2}\left(\overline{Q_T}\right)$ we have:

(2.3) $\quad \begin{cases} \text{for each } (t, x) \in Q_T \text{ being a local minimizer of } v - \phi, \\ \partial_t \phi(t, x) + (r - d)x\partial_x \phi(t, x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \phi(t, x) - rv(t, x) + q_*(x, v(t, x)) \leq 0. \end{cases}$

If, in addition, $v|_{t=T} \geq g$ on $[0, \infty)$, then v is a viscosity supersolution of (1.3)–(1.7).
(iii) A function $v \in C\left(\overline{Q_T}\right)$ is a *viscosity solution* of (1.3) if and only if it is simultaneously a viscosity sub– and supersolution of (1.3). If, in addition, $v|_{t=T} = g$ on $[0, \infty)$, then v is a *viscosity solution* of the terminal value problem (1.3)–(1.7).

Remark. For convenience, we adopt the terms *subsolution* and *supersolution* instead of viscosity subsolution and viscosity supersolution. Furthermore, it is well known (see e.g. [9, 11]) that we can replace "local" by "strict local" or "global" by "strict global". We can also assume that the extremum of $v - \phi$ has the value zero. There are equivalent formulations of sub– and supersolutions based on so-called semijets (or semidifferentials). These formulations were used in [7], but they will not be needed in this paper.

Remark. Note that if v is a subsolution (supersolution) for $x > 0$, then it is automatically a subsolution (supersolution) for $x \geq 0$. We refer to [7] for a proof of this result.

The remaining part of this section is devoted to weak (half-relaxed) limit operations with viscosity solutions [9, 11]. This will set the scene for the convergence proofs presented in the subsequent sections. Regarding limit operations with viscosity solutions, the following comparison principle will be of fundamental importance:

Theorem 2.1 ([7]). *Let v_{sub} and v_{sup} be respectively a sub– and supersolution of (1.3)–(1.7). Suppose there exists a finite constant C such that*

$$(2.4) \qquad v_{\mathrm{sub}}(t,x), -v_{\mathrm{sup}}(t,x) \le C\,(1+x)\,, \qquad (t,x) \in \overline{Q_T},$$

Then we have

$$v_{\mathrm{sub}} \le v_{\mathrm{sup}} \;\; on \;\; \overline{Q_T}.$$

Later we shall repeatedly be faced with the problem of passing to the limit as $\varepsilon \downarrow 0$ in various sequences $v_\varepsilon : \overline{Q_T} \to \mathbb{R}$ of approximate solutions. A natural procedure for doing so is to prove that v_ε is uniformly bounded (in ε) as well as equicontinuous on compacta. The Ascoli-Arzela compactness theorem then gives L^∞_{loc} (i.e., uniform on compacta) convergence along a subsequence of v_ε to a locally bounded and continuous function v. However, a merit of the viscosity solution theory is that one can dispense with the equicontinuity estimate provided there is a comparison principle in the class of semicontinuous sub– and supersolutions (Theorem 2.1). This limiting procedure is known as the Barles–Ishii–Perthame weak limit method [3, 14]. For completeness, we recall here the definitions of the upper and lower weak (or half-relaxed) limits of v_ε.

Definition 2.2. *Suppose v_ε is locally uniformly bounded. (i) The upper weak limit of v_ε, denoted by $\overline{v}$, is defined as*

$$\overline{v}(t,x) = \limsup_{\substack{\overline{Q_T} \ni (s,y) \to (t,x) \\ \varepsilon \downarrow 0}} v_\varepsilon(s,y)$$

$$= \lim_{\delta \downarrow 0} \sup \Big\{ v_\varepsilon(s,y) \,\Big|\, (s,y) \in \overline{Q_T}, |t-s|, |x-y| \le \delta, 0 < \varepsilon \le \delta \Big\}.$$

(ii) The lower weak limit of v_ε, denoted by $\underline{v}$, is defined as

$$\underline{v}(t,x) = \liminf_{\substack{\overline{Q_T} \ni (s,y) \to (t,x) \\ \varepsilon \downarrow 0}} v_\varepsilon(s,y)$$

$$= \lim_{\delta \downarrow 0} \inf \Big\{ v_\varepsilon(s,y) \,\Big|\, (s,y) \in \overline{Q_T}, |t-s|, |x-y| \le \delta, 0 < \varepsilon \le \delta \Big\}.$$

Since v_ε is locally uniformly bounded, the weak limits $\overline{v}$ and $\underline{v}$ are well-defined (finite). Some properties (to be used later) of the weak limits are given in the next lemma (whose proof can found in, e.g., [9, 11]).

Lemma 2.1. *(i) The upper weak limit $\overline{v}$ belongs to $USC\left(\overline{Q_T}\right)$ and the lower weak limit $\underline{v}$ belongs to $LSC\left(\overline{Q_T}\right)$. (ii) If $\overline{v} = \underline{v} = v$ on a compact subset of $\overline{Q_T}$, then v is continuous and $v_\varepsilon \to v$ in L^∞ (i.e., uniformly) on this set as $\varepsilon \downarrow 0$. (iii) Let $v_\varepsilon \in USC\left(\overline{Q_T}\right) \left(LSC\left(\overline{Q_T}\right)\right)$ be locally uniformly (in ε) bounded. Let $(t,x) \in \overline{Q_T}$ be a strict local maximizer of $\overline{v} - \phi$*

(minimizer for $\underline{v} - \phi$), $\phi \in C^{1,2}\left(\overline{Q_T}\right)$. Then there exist subsequences, which we do not relabel, $(t_\varepsilon, x_\varepsilon) \to (t, x)$ and $v_\varepsilon(t_\varepsilon, x_\varepsilon) \to \overline{v}(t, x)$ $(\underline{v}(t, x))$ as $\varepsilon \downarrow 0$ such that $(t_\varepsilon, x_\varepsilon)$ is a local maximizer (minimizer) of $v_\varepsilon - \phi$ for each $\varepsilon > 0$.

Regarding the general use of the weak limit method as a tool for proving convergence of approximate solutions of fully nonlinear degenerate second order partial differential equations, we refer [4, 9, 11] and the references therein. For some concrete applications to partial differential equations arising in finance theory, see [1, 2] and the references therein.

3. Semi–discrete approximations

In this section we present and analyze several examples of "semi–discrete" approximate solutions of the semilinear Black and Scholes equation (1.3). These approximations can be used to construct "fully discrete" numerical schemes for computing the value of an American option (see, e.g., [12, 21, 25] as well as Section 5 herein). We also introduce the basic techniques for analyzing approximate solution sequences associated with the semilinear Black and Scholes equation.

Let us start with the classical penalization technique [5, 6], which considers the following equation for each $\varepsilon > 0$:

$$(3.1) \qquad \partial_t v_\varepsilon + (r - d)x\partial_x v_\varepsilon + \frac{1}{2}\sigma^2 x^2 \partial_x^2 v_\varepsilon - r v_\varepsilon = -\frac{1}{\varepsilon}\left(g(x) - v_\varepsilon\right)^+, \qquad (t, x) \in Q_T,$$

with terminal data

$$(3.2) \qquad\qquad v_\varepsilon(T, x) = g(x), \qquad x \in [0, \infty).$$

From standard viscosity solution theory [9, 11], we know that there exists a unique viscosity solution v_ε to (3.1)–(3.2) satisfying $0 \leq v_\varepsilon \leq C(1 + x)$ on $\overline{Q_T}$, where the constant C is independent of ε. Regarding the penalization method, we have the following theorem:

Theorem 3.1. *Let v be the unique viscosity solution of (1.3)–(1.7). For each $\varepsilon > 0$, let v_ε be the unique viscosity solution of (3.1)–(3.2). Then $v_\varepsilon \to v$ in $L^\infty_{\text{loc}}\left(\overline{Q_T}\right)$ as $\varepsilon \downarrow 0$.*

Proof. Let $\overline{v}$ and $\underline{v}$ be the upper and lower weak limits of v_ε (see Definition 2.2). From Lemma 2.1, $\overline{v} \in USC\left(\overline{Q_T}\right)$ and $\underline{v} \in LSC\left(\overline{Q_T}\right)$. Obviously, $0 \leq \overline{v}, \underline{v} \leq C(1 + x)$ on $\overline{Q_T}$.

We prove first that $\underline{v}$ is a supersolution of (1.3)–(1.7), starting with the terminal condition (1.7). To get a contradiction, suppose there exist $y \in [0, \infty)$ and $\delta > 0$ such that $\underline{v}(T, y) \leq g(y) - \delta$. Pick sequences $(t_\varepsilon, x_\varepsilon) \to (T, y)$ and $v_\varepsilon(t_\varepsilon, x_\varepsilon) \to \underline{v}(T, y)$ as $\varepsilon \downarrow 0$. Because of (3.2), $t_\varepsilon < T \ \forall \varepsilon \leq \varepsilon_0$ for some $\varepsilon_0 > 0$. We next pick a function $\tilde{g} \in C^2\left([0, \infty)\right)$ such that $\tilde{g} \leq g$ on $[0, \infty)$, $\tilde{g}(y) = g(y) - \frac{\delta}{2}$, and $\tilde{g} = \text{Const}$ on $[\tilde{K}, \infty)$ with $\tilde{K} > y$. Define the function $G = -C(T - t) + \tilde{g}$ for a constant $C > 0$. Observe that $G < g$ on Q_T. By an appropriate choice of C, G turns out to be a subsolution of (3.1)–(3.2):

$$\partial_t G(t, x) + (r - d)x\partial_x G(t, x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 G(t, x) - r G(t, x) + \frac{1}{\varepsilon}\left(g(x) - G\right)^+$$

$$\geq C + (r - d)x\partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x),$$

By choosing

$$C \geq - \min_{x \in [0,\infty)} \left\{ (r-d)x \partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) \right\},$$

we see that G becomes a subsolution. Observe that the minimum is finite since $\tilde{g} = \text{Const}$ on $[\tilde{K}, \infty)$. The comparison principle for (3.1)–(3.2) (see [9, 11]) implies $v_\varepsilon \geq G$ on $\overline{Q_T}$ for any $\varepsilon \in (0, \varepsilon_0]$. Sending $\varepsilon \downarrow 0$, we get $\underline{v} \geq G$ on $\overline{Q_T}$ and, in particular, $\underline{v}(T, y) \geq \tilde{g}(y) = g(y) - \frac{\delta}{2}$, which is a contradiction. This concludes the proof of the terminal condition $\underline{v}|_{t=T} \geq g$ on $[0, \infty)$. Next we prove that $\underline{v}$ is a supersolution of (1.3). Let $(t, x) \in Q_T$ be a strict local minimizer of $\underline{v} - \phi$, $\phi \in C^{1,2}\left(\overline{Q_T}\right)$. First, we claim that

(3.3) $$\underline{v}(t, x) \geq g(x).$$

To get a contradiction, suppose $\underline{v}(t, x) \leq g(x) - \delta$ for some $\delta > 0$. Applying Lemma 2.1, there exist sequences $(t_\varepsilon, x_\varepsilon) \to (t, x)$ and $v_\varepsilon(t_\varepsilon, x_\varepsilon) \to \underline{v}(t, x)$ as $\varepsilon \downarrow 0$ such that $(t_\varepsilon, x_\varepsilon)$ is a local minimizer of $v_\varepsilon - \phi$ for each ε. Obviously, there exists an $\varepsilon(\delta) > 0$ such that

$$v_\varepsilon(t_\varepsilon, x_\varepsilon) \leq g(x_\varepsilon) - \frac{\delta}{2}, \qquad \forall \varepsilon \leq \varepsilon(\delta).$$

In view of this and since v_ε is a supersolution of (3.1), we have (for ε small enough)

$$\partial_t \phi(t_\varepsilon, x_\varepsilon) + (r-d)x_\varepsilon \partial_x \phi(t_\varepsilon, x_\varepsilon) + \frac{1}{2}\sigma^2 x_\varepsilon^2 \partial_x^2 \phi(t_\varepsilon, x_\varepsilon) - r v_\varepsilon(t_\varepsilon, x_\varepsilon)$$

$$\leq -\frac{1}{\varepsilon}\left(g(x_\varepsilon) - v_\varepsilon(t_\varepsilon, x_\varepsilon)\right) \leq -\frac{\delta}{2\varepsilon}.$$

Sending $\varepsilon \downarrow 0$ in this inequality gives a contradiction since the left-hand side converges to a finite number while the right-hand side converges to $-\infty$. This proves (3.3), which in turn implies $q_*\left(x, \underline{v}(t, x)\right) = 0$. To conclude the proof of the supersolution property of $\underline{v}$ at (t, x), it only remains to send $\varepsilon \downarrow 0$ in the inequality

$$\partial_t \phi(t_\varepsilon, x_\varepsilon) + (r-d)x_\varepsilon \partial_x \phi(t_\varepsilon, x_\varepsilon) + \frac{1}{2}\sigma^2 x_\varepsilon^2 \partial_x^2 \phi(t_\varepsilon, x_\varepsilon) - r v_\varepsilon(t_\varepsilon, x_\varepsilon) \leq 0.$$

In [7], it is proved that g is a subsolution of (1.3)–(1.7). We have just proved that $\underline{v}$ is a supersolution of (1.3)–(1.7). Consequently, Theorem 2.1 tells us that

(3.4) $$\overline{v}, \underline{v} \geq g \text{ on } \overline{Q_T}.$$

We prove next that $\overline{v}$ is a subsolution of (1.3)–(1.7), starting with the terminal condition (1.7). To get a contradiction, suppose there exist $y \in [0, \infty)$ and $\delta > 0$ such that $\overline{v}(T, y) \geq g(y) + \delta$. Pick sequences $(t_\varepsilon, x_\varepsilon) \to (T, y)$ and $v_\varepsilon(t_\varepsilon, x_\varepsilon) \to \overline{v}(T, y)$ as $\varepsilon \downarrow 0$. Because of (3.2), $t_\varepsilon < T \, \forall \varepsilon \leq \varepsilon_0$ for some $\varepsilon_0 > 0$. We next pick a function $\tilde{g} \in C^2\left([0, \infty)\right)$ such that $\tilde{g} \geq g$ on $[0, \infty)$, $\tilde{g}(y) = g(y) + \frac{\delta}{2}$, and $\tilde{g} = g$ on $[\tilde{K}, \infty)$ with $\tilde{K} > \max(y, K)$. Define the function $G = C(T-t) + \tilde{g}$ and note that $G \geq g$ on $\overline{Q_T}$. Following the same calculation as in the subsolution case, we see that by choosing

$$C \geq \max_{x \in [0,\infty)} \left\{ (r-d)x \partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) \right\},$$

G becomes a supersolution of (3.1)–(3.2). Note that the maximum is finite: For $x > \tilde{K}$ we have $\tilde{g} = g$. Thus, for a call option $\tilde{g}(x) = x - K$ for $x > \tilde{K}$, and the expression inside the curly brackets is equal to $-dx + rK \leq -d\tilde{K} + rk$. For a put option, on the other hand, $\tilde{g}(x) = 0$ for $x > \tilde{K}$. The comparison principle for (3.1)–(3.2) (see [9, 11]) then implies $v_\varepsilon \leq G$ on $\overline{Q_T}$ for any $\varepsilon \in (0, \varepsilon_0]$. Sending $\varepsilon \downarrow 0$, we get $\overline{v} \leq G$ on $\overline{Q_T}$ and thus $\overline{v}(T, y) \leq \tilde{g}(y) = g(y) + \frac{\delta}{2}$, which is a contradiction. Hence $\overline{v}|_{t=T} \leq g$ on $[0, \infty)$. Let us now prove that $\overline{v}$ is a subsolution of (1.3). Let $(t, x) \in Q_T$ be a strict local maximizer of $\overline{v} - \phi$, $\phi \in C^{1,2}(\overline{Q_T})$. We use again Lemma 2.1 to find sequences $(t_\varepsilon, x_\varepsilon) \to (t, x)$ and $v_\varepsilon(t_\varepsilon, x_\varepsilon) \to \overline{v}(t, x)$ as $\varepsilon \downarrow 0$ such that $(t_\varepsilon, x_\varepsilon)$ is a local maximizer of $v_\varepsilon - \phi$ for each ε. If $\overline{v}(t, x) = g(x)$, then we use (3.4) and argue as follows: Since $\overline{v}(t, x) = g(x)$ and $\overline{v} \geq g$ on $\overline{Q_T}$, we conclude that $g - \phi$ has a local maximum at (t, x). Since g is a subsolution of (1.3) (see [7]), (2.2) follows. Finally, if $\overline{v}(t, x) > g(x)$, then it is clear that $v_\varepsilon(t_\varepsilon, x_\varepsilon) > g(x_\varepsilon)$ for any ε sufficiently small. Since v_ε is a subsolution of (3.1), we have (for any ε small enough)

$$\partial_t \phi(t_\varepsilon, x_\varepsilon) + (r - d)x_\varepsilon \partial_x \phi(t_\varepsilon, x_\varepsilon) + \frac{1}{2}\sigma^2 x_\varepsilon^2 \partial_x^2 \phi(t_\varepsilon, x_\varepsilon) - r v_\varepsilon(t_\varepsilon, x_\varepsilon) \geq 0.$$

Since $\overline{v}(t, x) > g(x)$, $q_*(x, \overline{v}(t, x)) = 0$. Sending $\varepsilon \downarrow 0$ in the above inequality thus yields the subsolution property of $\overline{v}$ at (t, x).

Theorem 2.1 gives that $\overline{v} \leq \underline{v}$, and thus $\overline{v} = \underline{v} = v$. Lemma 2.1 then concludes the proof of the theorem. $\square$

Let $q_\varepsilon : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ designate a reasonable approximation to q. The term "reasonable" will be made precise through conditions (3.12) and (3.13) in Lemma 3.1 below. Another way to construct semi–discrete approximate solutions to the semilinear Black and Scholes equation (1.3) is then to consider the following equation for each $\varepsilon > 0$:

$$(3.5) \qquad \partial_t v_\varepsilon + (r - d)x\partial_x v_\varepsilon + \frac{1}{2}\sigma^2 x^2 \partial_x^2 v_\varepsilon - r v_\varepsilon + q_\varepsilon(x, v_\varepsilon) = 0, \qquad (x, t) \in Q_T,$$

which is augmented with terminal data

$$(3.6) \qquad v_\varepsilon(T, x) = g(x), \qquad x \in [0, \infty).$$

Let us look at some choices of q_ε of the form

$$(3.7) \qquad q_\varepsilon(x, v) = c(x)H_\varepsilon(g(x) - v), \qquad \varepsilon > 0,$$

where $H_\varepsilon : \mathbb{R} \to \mathbb{R}$ is an approximation to the Heaviside function (1.6). One example may be the "symmetric approximation"

$$(3.8) \qquad H_\varepsilon(\xi) = \begin{cases} 0, & \xi < -\varepsilon, \\ \frac{\xi + \varepsilon}{2\varepsilon}, & -\varepsilon \leq \xi < \varepsilon, \\ 1 & \xi \geq \varepsilon. \end{cases}$$

Similarly, we have "approximation from above"

$$
(3.9) \qquad H_\varepsilon(\xi) = \begin{cases} 0, & \xi < -\varepsilon, \\ \frac{\xi+\varepsilon}{\varepsilon}, & -\varepsilon \le \xi < 0, \\ 1 & \xi \ge 0, \end{cases}
$$

as well as "approximation from below"

$$
(3.10) \qquad H_\varepsilon(\xi) = \begin{cases} 0, & \xi < 0, \\ \frac{\xi}{\varepsilon}, & 0 \le \xi < \varepsilon, \\ 1 & \xi \ge \varepsilon. \end{cases}
$$

Our final example is (there are of course infinitely many more and they do not need to be smooth as is the case with the examples here)

$$
(3.11) \qquad H_\varepsilon(\xi) = \frac{\varepsilon}{\varepsilon - \xi^-}.
$$

From standard viscosity solution theory [9, 11], we know that for any reasonable q_ε there exists a unique viscosity solution v_ε of (3.1)–(3.2) satisfying $0 \le v_\varepsilon \le C(1+x)$ on $\overline{Q_T}$, where the constant C is independent of ε. The next lemma shows that viscosity solutions (in the sense of Definition 2.1) are stable with respect to weak limits.

Lemma 3.1. *For each $\varepsilon > 0$, let q_ε be a locally uniformly bounded function such that the comparison principle holds for (3.5)–(3.6), and $v_\varepsilon \in USC\left(\overline{Q_T}\right)$ (LSC $\left(\overline{Q_T}\right)$) be a locally uniformly bounded subsolution (supersolution) of (3.5)–(3.6). . Suppose there exist $\varepsilon_0 > 0$ and finite constants $\underline{C}, \overline{C} \ge 0$ such that*

$$
(3.12) \qquad \begin{aligned} q_\varepsilon(x,v) &\le c(x) + \overline{C} \quad \text{whenever } g(x) - v < 0 \text{ and } \varepsilon \le \varepsilon_0, \\ q_\varepsilon(x,v) &\ge c(x) - \underline{C} \quad \text{whenever } g(x) - v > 0 \text{ and } \varepsilon \le \varepsilon_0, \end{aligned}
$$

where the "cash flow" function c is defined in (1.5). Let $\overline{q}$ and $\underline{q}$ be respectively the upper and lower weak limits of q_ε. If

$$
(3.13) \qquad \overline{q}(x,v) \le q^*(x,v) \quad \forall(x,v) \in \mathbb{R} \times \mathbb{R}, \qquad \underline{q}(x,v) \ge q_*(x,v) \quad \forall(x,v) \in \mathbb{R} \times \mathbb{R},
$$

where q^*, q_* are defined in (2.1), then the upper weak limit $\overline{v}$ (lower weak limit $\underline{v}$) of v_ε is a subsolution (supersolution) of (1.3)–(1.7).

Proof. From Lemma 2.1, $\overline{v}$ is upper semicontinuous and $\underline{v}$ is lower semicontinuous. Also, there holds that $0 \le \overline{v}, \underline{v} \le C(1+x)$ on $\overline{Q_T}$. Let us now prove that $\overline{v}$ satisfies the terminal condition. To this end, choose a function $\tilde{g} \in C^2\left([0,\infty)\right)$ such that $\tilde{g} \ge g$ on $[0,\infty)$, $\|\tilde{g} - g\|_{L^\infty([0,\infty))} \le \delta$ for some $\delta > 0$, and $\tilde{g} = g$ on $[\tilde{K}, \infty)$ with $\tilde{K} > K$. Define the function $G = C(T-t) + \tilde{g}$ for a constant $C > 0$, and note that $G > g$ on Q_T. Observe that, for

any $\varepsilon \leq \varepsilon_0$ and $(t, x) \in Q_T$,

$$\partial_t G(t, x) + (r - d)x\partial_x G(t, x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 G(t, x) - rG(t, x) + q_\varepsilon(x, G(t, x))$$

(3.14)
$$\leq -C + (r - d)x\partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) + q_\varepsilon(x, G(t, x))$$

$$\leq -C + (r - d)x\partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) + c(x) + \overline{C},$$

where we have used (3.12) to derive the second inequality. Now choose

$$C \geq \overline{C} + \max_{x \in [0,\infty)} \left\{ (r - d)x\partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) + c(x) \right\}.$$

Note that the maximum is finite since

$$(r - d)x\partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) = -c(x), \qquad x > \tilde{K}.$$

Plugging this C into (3.14), we conclude that G is a supersolution of (3.5) and obviously also of (3.5)–(3.6) (at least when $\varepsilon \leq \varepsilon_0$). The comparison principle for (3.5)–(3.6) (see [9, 11]) then implies $v_\varepsilon \leq G$ on $\overline{Q_T}$ for any $\varepsilon \leq \varepsilon_0$. Hence $\overline{v} \leq G$ on $[0, \infty)$ and, in particular, there holds that $\overline{v}(T, x) \leq \tilde{g}(x) \leq g(x) + \delta$ for $x \in [0, \infty)$. Since $\delta > 0$ was arbitrary, we conclude that $\overline{v}|_{t=T} \leq g$ on $[0, \infty)$. Similarly, we can prove that $\underline{v}$ satisfies the terminal condition. This time we choose a function $\tilde{g} \in C^2([0, \infty))$ such that $\tilde{g} \leq g$ on $[0, \infty)$, $\|\tilde{g} - g\|_{L^\infty([0,\infty))} \leq \delta$, and $\tilde{g} = g$ on $[\tilde{K}, \infty)$ with $\tilde{K} > K$. Define the function $G = -C(T - t) + \tilde{g}$ for a constant $C > 0$, and note that $G < g$ on Q_T. In view of (3.12), it is not difficult to see that if we choose C so that

$$C \geq \underline{C} - \min_{x \in [0,\infty)} \left\{ (r - d)x\partial_x \tilde{g}(x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \tilde{g}(x) - r\tilde{g}(x) + c(x) \right\},$$

then G becomes a subsolution of (3.1)–(3.2) whenever $\varepsilon \leq \varepsilon_0$. The comparison principle for (3.5)–(3.6) (see [9, 11]) then implies $v_\varepsilon \geq G$ on $\overline{Q_T}$ (for small enough ε). We now end the proof as we did for $\overline{v}$ and obtain $\underline{v}|_{t=T} \geq g$ on $[0, \infty)$.

Next we prove that $\overline{v}, \underline{v}$ are respectively sub– and supersolutions of (1.3). We present here only the subsolution proof (the supersolution proof is similar). Let $(t, x) \in Q_T$ be strict local maximizer of $\overline{v} - \phi$, $\phi \in C^{1,2}(\overline{Q_T})$. Applying Lemma 2.1, there exist sequences $(t_\varepsilon, x_\varepsilon) \to (t, x)$ and $v_\varepsilon(t_\varepsilon, x_\varepsilon) \to \overline{v}(t, x)$ as $\varepsilon \downarrow 0$ such that each $(t_\varepsilon, x_\varepsilon)$ is a local maximizer of $v_\varepsilon - \phi$, and hence

(3.15)
$$\partial_t \phi(t_\varepsilon, x_\varepsilon) + (r - d)x_\varepsilon \partial_x \phi(t_\varepsilon, x_\varepsilon) + \frac{1}{2}\sigma^2 x_\varepsilon^2 \partial_x^2 \phi(t_\varepsilon, x_\varepsilon) - rv_\varepsilon(t_\varepsilon, x_\varepsilon)$$
$$+ q_\varepsilon(x_\varepsilon, v_\varepsilon(t_\varepsilon, x_\varepsilon)) \geq 0,$$

or, after rearranging,

$$q_\varepsilon(x_\varepsilon, v_\varepsilon(t_\varepsilon, x_\varepsilon)) \geq -\partial_t \phi(t_\varepsilon, x_\varepsilon) - (r - d)x_\varepsilon \partial_x \phi(t_\varepsilon, x_\varepsilon)$$
(3.16)
$$-\frac{1}{2}\sigma^2 x_\varepsilon^2 \partial_x^2 \phi(t_\varepsilon, x_\varepsilon) + rv_\varepsilon(t_\varepsilon, x_\varepsilon).$$

Obviously, by (3.13) and the definition of $\overline{q}$, we have

$$q^* \left(x, \overline{v}(t, x)\right) \geq \overline{q} \left(x, \overline{v}(t, x)\right) \geq \limsup_{\varepsilon \downarrow 0} q_\varepsilon(x_\varepsilon, v_\varepsilon(t_\varepsilon, x_\varepsilon)).$$

Hence, taking lim sup on both sides in (3.16) and observing that lim sup coincides with lim on the right-hand side, we have the subsolution property of $\overline{v}$ at (t, x):

$$\partial_t \phi(t, x) + (r - d)x\partial_x \phi(t, x) + \frac{1}{2}\sigma^2 x^2 \partial_x^2 \phi(t, x) - r\overline{v}(t, x) + q^* \left(x, \overline{v}(t, x)\right) \geq 0.$$

$\square$

A consequence of the previous lemma is the following theorem:

Theorem 3.2. *For each $\varepsilon > 0$, suppose q_ε is a locally uniformly bounded function such that the comparison principle holds for (3.5)–(3.6), (3.12) and (3.13) hold, and v_ε is a locally uniformly bounded viscosity solution of (3.5)–(3.6). Then $v_\varepsilon \to v$ as $\varepsilon \downarrow 0$, where v is the unique viscosity solution of (1.3)–(1.7). The convergence takes place in $L^\infty_{\mathrm{loc}}\left(\overline{Q_T}\right)$.*

Proof. By definition, $\underline{v} \leq \overline{v}$ on $\overline{Q_T}$. From Lemma 3.1, $\overline{v}$ and $\underline{v}$ are respectively sub– and supersolutions of (1.3)–(1.7) and $\overline{v}$ is a supersolution of (1.3)–(1.7). Theorem 2.1 yields therefore $\overline{v} \leq \underline{v}$ on $\overline{Q_T}$, and hence $\underline{v} = \overline{v} = v$ is the viscosity solution of (1.3). The uniform convergence follows from Lemma 2.1. $\square$

Regarding the choices of q_ε discussed above, we have the following convergence theorem:

Theorem 3.3. *For each $\varepsilon > 0$, let v_ε be the unique viscosity solution of (3.5)–(3.6) with q_ε defined via (3.7) and one of the choices (3.8), (3.9), (3.10), or (3.11). As $\varepsilon \downarrow 0$, v_ε converges in $L^\infty_{\mathrm{loc}}\left(\overline{Q_T}\right)$ to the unique viscosity solution v of (1.3)–(1.7).*

Proof. In what follows, let q_ε be any one of the choices mentioned in the theorem. Standard viscosity solution theory [9, 11] provides us with the existence of a unique viscosity solution of (3.5)–(3.6) satisfying $0 \leq v_\varepsilon \leq C\left(1 + x\right)$ on $\overline{Q_T}$, where C is independent of ε. Moreover, the comparison principle holds. It is easy to check that q_ε satisfies (3.12) and that the upper and lower weak limits $\overline{q}$ of and $\underline{q}$ of q_ε coincide with q^* and q_* respectively. Hence an application of Theorem 3.2 concludes the proof. $\square$

4. THE PREDICTOR–CORRECTOR SCHEME

The semilinear Black and Scholes equation (1.3) provides a natural recipe for turning any numerical scheme for the European option valuation problem into a numerical scheme for the American option valuation problem. We have chosen to illustrate this feature by devising a very simple explicit upwind finite difference scheme of "predictor–corrector type" for (1.3). It will become apparent, however, that everything presented in this section can be extended to other (more sophisticated) numerical schemes for (1.3).

To implement a numerical scheme on a computer, we must *truncate* the infinite domain $[0, \infty)$ to a finite domain $[0, L)$, where $L < \infty$ is fixed, and then provide a reasonable boundary condition at $x = L$ (see (4.7) below). The truncation technique is a classical one in numerical finance (see [2] for a theoretical investigation of it). The choice of boundary

condition at $x = L$ will not affect the theoretical convergence analysis given later. The reason is that we will let $L \uparrow \infty$ as $\Delta x \downarrow 0$ in the convergence analysis.

Let $\Delta x > 0$ and $\Delta t > 0$ be the spatial and temporal discretization parameters, respectively. The spatial domain $[0, L]$ is then discretized into grid cells

$$I_j = \left[x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}} \right), \qquad j = 1, \ldots, J_L - 1,$$

where $x_\ell = \ell \Delta x$ for $\ell = 0, \frac{1}{2}, 1, \ldots, J_L - 1, J_L - \frac{1}{2}, J_L$. Furthermore, we set $I_0 = \left[0, x_{\frac{1}{2}} \right)$ and $I_{J_L} = \left[x_{J_L - \frac{1}{2}}, x_{J_L} \right]$. Choose the integer J_L such that $J_L \Delta x = L$. For the convergence analysis, we let

(4.1) $$J_L \Delta x = L \uparrow \infty \text{ as } \Delta x \downarrow 0.$$

Similarly, the time interval $[0, T]$ is discretized into time strips

$$I^n = [t^n, t^{n+1}), \qquad n = N - 1, \ldots, 0,$$

where $t^n = n\Delta t$ for $n = 0, \ldots, N$. The integer N is chosen such that $N\Delta t = T$. We denote by R_j^n the rectangle $I^n \times I_j$. For $j = 0, \ldots, J_L$ and $n = N, N - 1, \ldots, 0$, V_j^n denotes the predictor-corrector approximation (yet to be defined) associated with the point (t^n, x_j).

We extend the difference solution $\{V_j^n\}$ to all of $Q_T^L = [0, T] \times [0, L]$ by setting

(4.2) $$v_\Delta(t, x) = \begin{cases} V_j^n, & (t, x) \in R_j^n, \ j = 0, \ldots, J_L, \ n = N - 1, \ldots, 0, \\ V_j^N, & t = T, \ x \in I_j, \ j = 0, \ldots, J_L, \end{cases}$$

where Δ is used as short-hand notation for Δx. Let us now introduce the explicit predictor–corrector scheme. To this end, set $\lambda = \frac{\Delta t}{\Delta x}$ and $\mu = \frac{\Delta t}{(\Delta x)^2}$. To simplify the presentation, we use Δ_+ and Δ_- to designate the difference operators in the x direction:

$$\Delta_+ V_j^n = V_{j+1}^n - V_j^n, \qquad \Delta_- V_j^n = V_j^n - V_{j-1}^n.$$

For the same reason, we introduce the *upwind numerical flux function* $F : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by

$$F(a, b) = \begin{cases} b, & \text{when } r - d \geq 0, \\ a, & \text{when } r - d < 0. \end{cases}$$

The suggested numerical scheme for (1.3)–(1.7), which uses the upwind numerical flux function F for the convection part and centered differencing for the parabolic part, takes the following *predictor–corrector form* for $j = 0, \ldots, J_L - 1$ and $n = N - 1, \ldots, 0$:

Predictor step:

(4.3) $$V_j^{n+\frac{1}{2}} = V_j^{n+1} + \lambda(r - d)x_j \Delta_- F\left(V_j^{n+1}, V_{j+1}^{n+1} \right) + \mu \frac{1}{2}\sigma^2 x_j^2 \Delta_- \Delta_+ V_j^{n+1} - \Delta t r V_j^{n+1}.$$

Corrector step:

(4.4) $$V_j^n = V_j^{n+\frac{1}{2}} + \Delta t c(x_j) H\left(g(x_j) - V_j^{n+\frac{1}{2}} \right),$$

We start the backward iteration (4.3)–(4.4) by setting

(4.5) $$V_j^N = g(x_j), \qquad j = 0, \ldots, J_L.$$

We will impose the following Dirichlet condition at $x = 0$:

$$(4.6) \qquad\qquad V_0^{n+\frac{1}{2}} = V_0^n = g(0), \qquad n = N - 1, \ldots, 0.$$

Based on the asymptotic behavior of the American option value (1.1) as $x \uparrow \infty$, we will impose the following Dirichlet condition at $x = L$:

$$(4.7) \qquad\qquad V_{J_L}^{n+\frac{1}{2}} = V_{J_L}^n = g(L), \qquad n = N - 1, \ldots, 0.$$

Note that when $V_j^{n+\frac{1}{2}} \leq g(x_j)$, the updating formula (4.4) reduces to

$$(4.8)$$
$$V_j^n = V_j^{n+1} + \lambda(r - d)x_j \Delta_- F\left(V_j^{n+1}, V_{j+1}^{n+1}\right) + \mu\frac{1}{2}\sigma^2 x_j^2 \Delta_- \Delta_+ V_j^{n+1} - \Delta t r V_j^{n+1} + \Delta t c(x_j).$$

Otherwise, (4.4) reduces to

$$(4.9) \qquad V_j^n = V_j^{n+1} + \lambda(r - d)x_j \Delta_- F\left(V_j^{n+1}, V_{j+1}^{n+1}\right) + \mu\frac{1}{2}\sigma^2 x_j^2 \Delta_- \Delta_+ V_j^{n+1} - \Delta t r V_j^{n+1}.$$

For the call (put) option we have $c(x_j) = dx_j - rK$ $(rK - dx_j)$ if $x_j > \frac{r}{d}K$ $(< \frac{r}{d}K)$, and the updating formula (4.8) is possibly in effect. Otherwise, $c(x_j) = 0$, and the updating formula (4.9) is effective.

We assume that the following parabolic CFL condition holds:

$$(4.10) \qquad\qquad \lambda|r - d|L + \mu\sigma^2 L^2 + \Delta t r \leq 1.$$

Note that when $\Delta(= \Delta x) \downarrow 0$, then also $\Delta t \downarrow 0$ by this condition.

The following lemma shows that the approximate solution v_Δ satisfies the early exercise constraint (1.8) and possesses (at most) linear growth as $x \uparrow \infty$, which implies that the finite difference scheme (4.4) is L_{loc}^∞ stable.

Lemma 4.1. *Suppose the parabolic CFL condition* (4.10) *holds. Then the predictor–corrector solution* $\left\{V_j^n\right\}$ *defined by* (4.3)–(4.4)–(4.5)–(4.6)–(4.7) *satisfies*

$$(4.11) \qquad\qquad V_j^n \geq g(x_j), \qquad j = 0, \ldots, J_L, \ n = N, N - 1, \ldots, 0.$$

Consequently, the approximate solution v_Δ *defined by* (4.2) *and* $\left\{V_j^n\right\}$ *satisfies*

$$(4.12) \qquad\qquad v_\Delta(t, x) \geq g(x) - O(\Delta x), \qquad (t, x) \in \overline{Q_T^L}.$$

Furthermore, there exist finite constants C_1 *and* C_2, *independent of* Δ, *such that*

$$(4.13) \qquad\qquad V_j^n \leq C_1 + C_2 x_j, \qquad j = 0, \ldots, J_L, \ n = N, N - 1, \ldots, 0.$$

For the call option, $C_1 = 0$ *and* $C_2 = 1$. *For the put option,* $C_1 = K$ *and* $C_2 = 0$. *Consequently,*

$$(4.14) \qquad\qquad v_\Delta(t, x) \leq C_1 + C_2 x + O(\Delta x), \qquad (t, x) \in \overline{Q_T^L}.$$

Proof. The proof is inductive. Observe first that statements (4.11) and (4.13) hold for $(j = 0, n = N, N - 1, \ldots, 0)$, $(j = J_L, n = N, N - 1, \ldots, 0)$, and $(j = 0, \ldots, J_L, n = N)$. For $(j = 1, \ldots, J_L - 1, n = N - 1, \ldots, 0)$, we assume that a statement holds at time level

$n + 1$ and then seek to prove that it holds at time level n, starting with statement (4.11). First note that if $V_j^{n+\frac{1}{2}} > g(x_j)$, then

$$V_j^n = V_j^{n+\frac{1}{2}} > g(x_j),$$

and we are finished. In what follows, we therefore assume $V_j^{n+\frac{1}{2}} \leq g(x_j)$, so that

$$(4.15) \qquad V_j^n = V_j^{n+\frac{1}{2}} + \Delta t c(x_j).$$

Introducing the function $\mathcal{S}$ defined by

$$(4.16) \qquad \begin{aligned} &\mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) \\ &= V_j^{n+1} + \lambda(r-d)x_j\Delta_- F\left(V_j^{n+1}, V_{j+1}^{n+1}\right) + \mu\frac{1}{2}\sigma^2 x_j^2 \Delta_-\Delta_+ V_j^{n+1} - \Delta t r V_j^{n+1}, \end{aligned}$$

we can write (4.3) as

$$\mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^{n+\frac{1}{2}} = 0.$$

Under the parabolic CFL condition (4.10), it straightforward to check that

$$\partial\mathcal{S}/\partial V_{j-1}^{n+1},\ \partial\mathcal{S}/\partial V_j^{n+1},\ \partial\mathcal{S}/\partial V_{j+1}^{n+1} \geq 0,$$

which implies that the finite difference scheme (4.21) is *monotone*. Set $g_j = g(x_j)$. Since by assumption $V_j^{n+1} \geq g_j\ \forall j$, we then have

$$V_j^{n+\frac{1}{2}} = \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) \geq \mathcal{S}\left(x_j, g_{j-1}, g_j, g_{j+1}\right)\ \forall j.$$

Hence

$$(4.17) \qquad \begin{aligned} V_j^{n+\frac{1}{2}} &\geq g_j + \lambda(r-d)x_j\Delta_- F\left(g_j, g_{j+1}\right) + \mu\frac{1}{2}\sigma^2 x_j^2 \Delta_-\Delta_+ g_j - \Delta t r g_j, \\ &\geq g_j + \lambda(r-d)x_j\Delta_- F\left(g_j, g_{j+1}\right) - \Delta t r g_j\ \forall j. \end{aligned}$$

where we have used the convexity of g to derive the last inequality. To be concrete in what follows, we assume $r - d < 0$, so that (4.17) reads

$$V_j^{n+\frac{1}{2}} \geq g_j - \lambda(d-r)x_j\left(g_j - g_{j-1}\right) - \Delta t r g_j.$$

The case $r - d \geq 0$ can be treated similarly. We have two cases to consider: (i) $x_j \leq K$. Then for the call option

$$V_j^{n+\frac{1}{2}} \geq g_j - 0 \geq g_j - \Delta t c(x_j),$$

and for the put option

$$\begin{aligned} V_j^{n+\frac{1}{2}} &\geq g_j + \Delta t(d-r)x_j - \Delta t r(K - x_j) \\ &= g_j - \Delta t(rK - dx_j) \geq g_j - \Delta t c(x_j). \end{aligned}$$

(ii) $x_j > K$. Then for the call option

$$\begin{aligned} V_j^{n+\frac{1}{2}} &\geq g_j - \Delta t(d-r)x_j - \Delta t r(x_j - K) \\ &= g_j - \Delta t(dx_j - rK) \geq g_j - \Delta t c(x_j). \end{aligned}$$

and for the put option

$$V_j^{n+\frac{1}{2}} \geq g_j - 0 \geq g_j - \Delta t c(x_j).$$

Summing up,

(4.18) $$V_j^{n+\frac{1}{2}} \geq g_j - \Delta t c(x_j) \; \forall j.$$

Plugging this lower bound for $V_j^{n+\frac{1}{2}}$ into (4.15) gives the desired result (4.11).

From (4.11) and the definition of v_Δ (see (4.2)), we get (4.12).

Finally, we prove that (4.13) and (4.14) hold, starting with (4.13) and the put option. Assume that $V_j^{n+1} \leq K \; \forall j$. It then follows from from the monotonicity of $\mathcal{S}$ that

$$V_j^{n+\frac{1}{2}} \leq K(1 - \Delta t r),$$

and hence, via (4.4),

$$V_j^n \leq K(1 - \Delta t r) + \Delta t c(x_j) \leq K \; \forall j.$$

For the call option, we assume $V_j^{n+1} \leq x_j \; \forall j$. Again from the monotonicity of $\mathcal{S}$, we get

$$V_j^{n+\frac{1}{2}} \leq x_j(1 - \Delta t r) - \Delta t(d - r)x_j = x_j - \Delta t d x_j,$$

and therefore

$$V_j^n \leq x_j - \Delta t d x_j + c(x_j) \leq x_j \; \forall j.$$

This concludes the proof of (4.13), which also implies (4.14). $\qquad\square$

Theorem 4.1. *Suppose the parabolic CFL condition (4.10) holds. Denote by v the unique viscosity solution of (1.3)–(1.7). Let v_Δ be the approximate solution defined by (4.2) and the predictor–corrector scheme (4.3)–(4.4)–(4.5)–(4.6)–(4.7). Then*

$$v_\Delta \to v \; \text{in} \; L_{\text{loc}}^\infty \left(\overline{Q_T}\right) \; \text{as} \; \Delta \downarrow 0.$$

Proof. Let $\overline{v}$ and $\underline{v}$ be respectively the upper and lower weak limits of v_Δ. We have $\overline{v} \in USC\left(\overline{Q_T}\right)$, $\underline{v} \in LSC(Q_T)$, and $\underline{v} \leq \overline{v}$ on $\overline{Q_T}$. In view of Lemma 4.1, $\overline{v}, \underline{v}$ are finite at each point $(t, x) \in \overline{Q_T}$. More precisely, $0 \leq \overline{v}, \underline{v} \leq C(1 + x)$ on $\overline{Q_T}$, for some constant $C > 0$. For the moment, suppose that we can prove that $\overline{v}$ and $\underline{v}$ are respectively sub– and supersolutions of (1.3)–(1.7). An application of Theorem 2.1 then gives

(4.19) $$\overline{v} \leq \underline{v} \; \text{on} \; \overline{Q_T}.$$

We are thus finished since this implies $\overline{v} = \underline{v}$ on $\overline{Q_T}$ and, via Lemma 2.1, the approximate solution v_Δ converges as $\Delta \downarrow 0$ to the unique viscosity solution $v = \overline{v} = \underline{v}$ of (1.3)–(1.7).

We will first prove that $\overline{v}$ is subsolution of (1.3). Pick a local maximizer (t, x) of $\overline{v} - \phi$ for some $\phi \in C^{1,2}(\overline{Q_T})$. Without loss of generality, we assume $\overline{v}(t, x) = \phi(t, x)$ and that the maximizer is strict. Moreover, we can assume $x > 0$. In view of (4.1), there exists $\Delta x_0 > 0$ such that (t, x) is an interior point of Q_T^L for any $\Delta x < \Delta x_0$. In what follows, we assume $\Delta x < \Delta x_0$. In view of Lemma 4.1, we obviously have

(4.20) $$\overline{v}, \underline{v} \geq g \; \text{on} \; \overline{Q_T}.$$

First we consider the case $\overline{v}(t,x) = g(x)$. Since $\overline{v}(t,x) = g(x)$ and $\overline{v} \geq g$ on $\overline{Q_T}$, we conclude that $g - \phi$ has a local maximum at (t,x). In [7], we proved that g is a subsolution of (1.3). Hence (2.2) holds. Next we consider the case $\overline{v}(t,x) \geq g(x) + \delta$ for some $\delta > 0$. Then $q^*(x, \overline{v}(t,x)) = 0$. Applying Lemma 2.1, there exist sequences $(t_\Delta, x_\Delta) \to (t,x)$ and $v_\Delta(t_\Delta, x_\Delta) \to \overline{v}(t,x)$ as $\Delta \downarrow 0$ such that (t_Δ, x_Δ) is a local maximizer of $v_\Delta - \phi$ for each Δx. For any Δx sufficiently small, we have $v_\Delta(t_\Delta, x_\Delta) \geq g(x_\Delta) + \frac{\delta}{2}$. Without loss of generality, we can assume that there are integers (j,n) such that $(t_\Delta, x_\Delta) = (t^n, x_j)$ and hence $v_\Delta(t_\Delta, x_\Delta) = V_j^n$. Note that

$$V_j^n = V_j^{n+\frac{1}{2}} + \Delta t c(x_j) H\left(g(x_j) - V_j^{n+\frac{1}{2}}\right),$$

where $V_j^{n+\frac{1}{2}}$ is defined in (4.3). From this it follows that

$$V_j^{n+\frac{1}{2}} = V_j^n - \Delta t c(x_j) H\left(g(x_j) - V_j^{n+\frac{1}{2}}\right) \geq g(x_j) + \frac{\delta}{4},$$

by choosing Δ sufficiently small. This implies that $V_j^n = V_j^{n+\frac{1}{2}}$, so that (4.4) reduces to

$$(4.21) \quad V_j^n = V_j^{n+1} + \lambda(r-d)x_j \Delta_- F\left(V_j^{n+1}, V_{j+1}^{n+1}\right) + \mu \frac{1}{2}\sigma^2 x_j^2 \Delta_- \Delta_+ V_j^{n+1} - \Delta t r V_j^{n+1}.$$

Using the monotone function $\mathcal{S}$ defined in (4.16), we can write (4.21) as

$$(4.22) \qquad\qquad \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n = 0.$$

The fact that $v_\Delta - \phi$ has a maximum at (t_Δ, x_Δ) implies

$$\phi_\cdot^{n+1} \geq V_\cdot^{n+1} - (v_\Delta(t_\Delta, x_\Delta) - \phi(t_\Delta, x_\Delta)).$$

We have introduced the notation $\phi_j^n = \phi(t^n, x_j)$. By the monotonicity of $\mathcal{S}$ and (4.22), we therefore obtain

$$(4.23) \quad \begin{aligned} &\mathcal{S}\left(x_j, \phi_{j-1}^{n+1}, \phi_j^{n+1}, \phi_{j+1}^{n+1}\right) - \phi_j^n \\ &\geq \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n + \Delta t r\left(v_\Delta(t_\Delta, x_\Delta) - \phi(t_\Delta, x_\Delta)\right) \\ &= \Delta t r\left(v_\Delta(t_\Delta, x_\Delta) - \phi(t_\Delta, x_\Delta)\right). \end{aligned}$$

Dividing (4.23) by Δt and Taylor expanding around $(t_\Delta, x_\Delta) = (t^n, x_j)$ yields

$$\partial_t \phi(t_\Delta, x_\Delta) + (r-d)x_\Delta \partial_x \phi(t_\Delta, x_\Delta) + \frac{1}{2}\sigma^2 x_\Delta^2 \partial_x^2 \phi(t_\Delta, x_\Delta) - r v_\Delta(t_\Delta, x_\Delta) \geq -O(\Delta),$$

where "$O(\Delta)$" means "$\leq C\left(\Delta t + \Delta x\right)$" for some constant $C \geq 0$ being independent of Δx and Δt. Sending $\Delta \downarrow 0$ in this inequality finalizes the proof of the subsolution property. This shows that the finite difference scheme (4.21) is *consistent*.

We now prove that $\underline{v}$ is a supersolution of (1.3). Pick a local minimizer (t,x) of $\overline{v} - \phi$ with $\phi \in C^{1,2}(\overline{Q_T})$. Without loss of generality, we can assume $\overline{v}(t,x) = \phi(t,x)$ and that the minimizer is strict. As before, we can find sequences $(t_\Delta, x_\Delta) \to (t,x)$ and $v_\Delta(t_\Delta, x_\Delta) \to \overline{v}(t,x)$ as $\Delta \downarrow 0$ such that (t_Δ, x_Δ) is a local minimizer of $v_\Delta - \phi$ for each Δ.

We can also assume that each point (t_Δ, x_Δ) coincides with a grid point. Since $\underline{v} \geq g$ on $\overline{Q_T}$, we have $q_*(x, \underline{v}(t, x)) = 0$. Note that (4.4) reads

$$(4.24) \qquad \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n = -\Delta t c(x_j) H\left(g(x_j) - V_j^{n+\frac{1}{2}}\right) \leq 0.$$

In view of this, arguing analogously to what we did in the subsolution proof yields

$$\partial_t \phi(t_\Delta, x_\Delta) + (r-d)x_\Delta \partial_x \phi(t_\Delta, x_\Delta) + \frac{1}{2}\sigma^2 x_\Delta^2 \partial_x^2 \phi(t_\Delta, x_\Delta) - rv_\Delta(t_\Delta, x_\Delta) \leq O(\Delta).$$

Sending $\Delta \downarrow 0$ in this inequality finalizes the proof of the supersolution property.

Finally, let us prove that $\overline{v}$ and $\underline{v}$ satisfy the terminal condition, i.e., we have $\overline{v}|_{t=T} \leq g$ and $\underline{v}|_{t=T} \geq g$ on $[0, \infty)$. In view of (4.20), it is sufficient to prove that

$$(4.25) \qquad\qquad\qquad \overline{v}|_{t=T} \leq g \text{ on } [0, \infty),$$

which actually implies $\overline{v}|_{t=T} = \underline{v}_{t=T} = g$ on $[0, \infty)$. Pick a function $\tilde{g} \in C^2\left([0, \infty)\right)$ such that $\tilde{g} \geq g$ on $[0, \infty)$, $\|\tilde{g} - g\|_{L^\infty([0,\infty))} \leq \delta$ for some $\delta > 0$, and $\tilde{g} = g$ on $[\tilde{K}, \infty)$ with $\tilde{K} > \max(K, \frac{r}{d}K)$. Define the function $G = C(T - t) + \tilde{g}$ for a constant $C > 0$, and note that $G > g$ on Q_T. Let $J_{\tilde{K}}$ be the integer such that $\tilde{K} \in I_{J_{\tilde{K}}}$. Set $G_j^n = G(t^n, x_j)$, $\tilde{g}_j = \tilde{g}(x_j)$, and

$$\tilde{C} = \max_{j=0,\ldots,J_{\tilde{K}}} \left\{ (r-d)x_j \frac{1}{\Delta x}\Delta_- F\left(\tilde{g}_j, \tilde{g}_{j+1}\right) + \frac{1}{2}\sigma^2 x_j^2 \frac{1}{(\Delta x)^2}\Delta_-\Delta_+ \tilde{g}_j - r\tilde{g}_j \right\}.$$

Note that $\tilde{C}$ is finite and independent of Δx, since $\tilde{g} \in C^2$. Set $\tilde{c} = \max_{x \in \left[0, \tilde{K}+1\right]} c(x)$. For $j \leq J_{\tilde{K}}$, we have

$$\mathcal{S}\left(x_j, G_{j-1}^{n+1}, G_j^{n+1}, G_{j+1}^{n+1}\right) - G_j^n \leq \Delta t\left[-C + \tilde{C} - c(x_j) + \tilde{c}\right] \leq -\Delta t c(x_j),$$

provided we choose $C \geq \tilde{C} + \tilde{c}$. For $j > J_{\tilde{K}}$, it is easy to check that for any $C \geq 0$ we have

$$\mathcal{S}\left(x_j, G_{j-1}^{n+1}, G_j^{n+1}, G_{j+1}^{n+1}\right) - G_j^n = \Delta t\left[-C - (dx - rK)\right] \leq -\Delta t c(x_j)$$

for the call option and

$$\mathcal{S}\left(x_j, G_{j-1}^{n+1}, G_j^{n+1}, G_{j+1}^{n+1}\right) - G_j^n = \Delta t\left[-C\right] \leq -\Delta t c(x_j)$$

for the put option. Hence, by choosing $C \geq \tilde{C} + \tilde{c}$, we have

$$(4.26) \qquad\qquad \mathcal{S}\left(x_j, G_{j-1}^{n+1}, G_j^{n+1}, G_{j+1}^{n+1}\right) - G_j^n \leq -\Delta t c(x_j),$$

for $j = 1, \ldots, J_L - 1$ and $n = N - 1, \ldots, 0$. From the definition of the predictor–corrector scheme (4.3)–(4.4), it follows that

$$(4.27) \qquad\qquad \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n \geq -\Delta t c(x_j),$$

for $j = 1, \ldots, J_L - 1$ and $n = N - 1, \ldots, 0$. We now claim that that

$$(4.28) \qquad\qquad V_j^n \leq G_j^n, \qquad j = 0, \ldots, J_L, \; n = N, N-1, \ldots, 0.$$

Obviously, by (4.5), (4.6), and (4.7), this is obviously true for $(j = 0, \ldots, J_L, n = N)$, $(j = 0, n = N, N - 1, \ldots, 0)$, and $(j = J_L, n = N, N - 1, \ldots, 0)$. To prove (4.28) for the

remaining indices, we proceed by induction on n. Suppose $V_j^{n+1} \leq G_j^{n+1}$. Then, using (4.27), (4.26), and that $\mathcal{S}$ is a monotone mapping, we reach the desired result

$$V_j^n \leq \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) + \Delta t c(x_j) \leq \mathcal{S}\left(x_j, G_{j-1}^{n+1}, G_j^{n+1}, G_{j+1}^{n+1}\right) + \Delta t c(x_j) = G_j^n.$$

From (4.28), it immediately follows that

$$v_\Delta(t, x) \leq G(t, x) + O(\Delta x), \qquad (t, x) \in \overline{Q_T^L}.$$

and therefore $\overline{v} \leq G$ on $\overline{Q_T}$. In particular, $\overline{v}(T, x) \leq \tilde{g}(x) \leq g(x) + \delta$. Since $\delta > 0$ was arbitrary, the proof of (4.25) is finished. $\qquad\square$

Remark. Although this section shows that one can directly discretize the semilinear Black and Scholes equation (1.3) to obtain a "good" numerical scheme for computing the value of an American option, there is some numerical literature [12, 21, 25] on so-called *penalty schemes*. We claim that penalty schemes can be viewed as numerical schemes obtained by discretizing an approximation to the semilinear Black and Scholes equation (1.3), namely (3.1) with a suitable choice of ε [12, 25] or (3.5) with a suitable choice of q_ε. For example, the choice of q_ε based on (3.7) and (3.11) is related to the penalty scheme used in [21]. The authors of [12, 21, 25] do not provide rigorous convergence proofs for their penalty schemes. At a semi–discrete level, we have already provided rather general convergence theorems (see Theorems 3.2, 3.3, and 3.1). Regarding convergence proofs for numerical schemes based on discretizing (3.1) or (3.5), we only mention that the convergence arguments presented herein can be modified so as to apply to such numerical schemes.

From a computational point of view, the explicit predictor–corrector scheme (4.4) is not particularly useful because of the severe time step restriction imposed by the parabolic CFL condition (4.10). However, (4.10) can be easily avoided by replacing the explicit predictor step (4.3) by an implicit one. We end this section by briefly discussing an implicit predictor–corrector scheme. To simplify matters, we consider only the put option and set $d = 0$. We then consider the following "CFL condition free" predictor–corrector scheme:

$$(4.29) \qquad \begin{cases} V_j^{n+\frac{1}{2}} = V_j^{n+1} + \lambda r x_j \Delta_- V_{j+1}^{n+\frac{1}{2}} + \mu \dfrac{1}{2}\sigma^2 x_j^2 \Delta_- \Delta_+ V_j^{n+\frac{1}{2}} - \Delta t r V_j^{n+\frac{1}{2}}, \\[2mm] V_j^n = V_j^{n+\frac{1}{2}} + \Delta t r K H\left((K - x_j)^+ - V_j^{n+\frac{1}{2}}\right). \end{cases}$$

The terminal and boundary conditions for (4.29) are as before (see (4.5), (4.6), and (4.7)). Note that in (4.29) we do not have to solve a nonlinear algebraic system at each time step but merely a tridiagonal linear system. This makes (4.29) an efficient numerical scheme.

We have the following convergence theorem for (4.29):

Theorem 4.2. *The statements in Lemma 4.1 and Theorem 4.1 hold true for the approximate solution $\{V_j^n\}$ generated by the implicit predictor–corrector scheme* (4.29).

Proof. The proof follows very closely the proofs of Lemma 4.1 and Theorem 4.1 for the explicit predictor–corrector scheme, and hence we will omit most of it. Since we do not have any CFL condition relating the convergence to zero of Δt with that of Δx, we need in this case to assume that Δt goes to zero as a function of Δx when $\Delta x \downarrow 0$. To illustrate

a typical difference between the methods of proof for the explicit and implicit schemes, we will, however, prove that the early exercise constraint (4.11) holds. To this end, let us write the predictor step in (4.29) as

$$(4.30) \qquad \frac{V_j^{n+1} - V_j^{n+\frac{1}{2}}}{\Delta t} + \mathcal{S}\left(x_j, V_{j-1}^{n+\frac{1}{2}}, V_j^{n+\frac{1}{2}}, V_{j+1}^{n+\frac{1}{2}}\right) = 0,$$

where

$$\mathcal{S}\left(x_j, V_{j-1}^{n+\frac{1}{2}}, V_j^{n+\frac{1}{2}}, V_{j+1}^{n+\frac{1}{2}}\right) = rx_j \frac{1}{\Delta x} \Delta_- V_{j+1}^{n+\frac{1}{2}} + \frac{1}{2}\sigma^2 x_j^2 \frac{1}{(\Delta x)^2} \Delta_- \Delta_+ V_j^{n+\frac{1}{2}} - rV_j^{n+\frac{1}{2}}.$$

Observe that

$$\partial \mathcal{S}/\partial V_{j-1}^{n+\frac{1}{2}}, \partial \mathcal{S}/\partial V_{j+1}^{n+\frac{1}{2}} \geq 0.$$

For any j, set $g_j = g(x_j)$ and $W_j^{n+p} = V_j^{n+p} - g_j$ with $p = 1, \frac{1}{2}$. Observe that $W_j^{n+1} \geq 0$ for all j. The scheme for $W_j^{n+\frac{1}{2}}$ reads

$$\frac{W_j^{n+1} - W_j^{n+\frac{1}{2}}}{\Delta t} + \mathcal{S}\left(x_j, W_{j-1}^{n+\frac{1}{2}}, W_j^{n+\frac{1}{2}}, W_{j+1}^{n+\frac{1}{2}}\right) + \mathcal{S}\left(x_j, g_{j-1}, g_j, g_{j+1}\right) = 0.$$

Let $W_\ell^{n+\frac{1}{2}} = \min_j W_j^{n+\frac{1}{2}}$ for some ℓ and assume $W_\ell^{n+\frac{1}{2}} < 0$ (otherwise there is nothing to prove). Then

$$\frac{W_\ell^{n+1} - W_\ell^{n+\frac{1}{2}}}{\Delta t} + \mathcal{S}\left(x_\ell, W_\ell^{n+\frac{1}{2}}, W_\ell^{n+\frac{1}{2}}, W_\ell^{n+\frac{1}{2}}\right) + \mathcal{S}\left(x_\ell, g_{\ell-1}, g_\ell, g_{\ell+1}\right) \leq 0,$$

which implies that

$$(1 + \Delta t r) W_\ell^{n+\frac{1}{2}} \geq \Delta t \mathcal{S}\left(x_j, g_{j-1}, g_j, g_{j+1}\right) \geq -\Delta t r K,$$

where the last inequality follows as in the proof of (4.11) for the explicit predictor–corrector scheme. In particular, $W_j^{n+\frac{1}{2}} \geq -\Delta t r K$ for all j. Plugging this into (4.29) yields (4.11), and hence (4.12).

Similarly, we can prove that (4.13) and (4.12) hold.

Equipped with (4.29) and the monotonicity of $\mathcal{S}$ in two of the variables, the proof that Theorem 4.1 holds for (4.29) is more or less identical to the proof for the explicit predictor–corrector scheme. $\qquad \square$

5. The Brennan and Schwartz algorithm

By applying a predictor–corrector discretization to (3.1), we can device a numerical scheme that is identical to the one analyzed in Section 4 except that the corrector step (4.4) is replaced by the following one:

$$V_j^n = V_j^{n+\frac{1}{2}} + \frac{\Delta t}{\varepsilon}\left(g(x_j) - V_j^{n+\frac{1}{2}}\right)^+,$$

where ε has to be chosen. We are here interested in the particular choice $\varepsilon = \Delta t$, which results in the following explicit scheme

$$(5.1) \quad \begin{cases} V_j^{n+\frac{1}{2}} = V_j^{n+1} + \lambda(r-d)x_j\Delta_-F\left(V_j^{n+1}, V_{j+1}^{n+1}\right) + \mu\frac{1}{2}\sigma^2 x_j^2\Delta_-\Delta_+V_j^{n+1} - \Delta tr V_j^{n+1}, \\ \\ V_j^n = \max\left(V_j^{n+\frac{1}{2}}, g(x_j)\right) \end{cases}$$

The terminal and boundary conditions for (5.1) are as in Section 4 (see (4.5), (4.6), and (4.7)). We call (5.1) the *Brennan and Schwartz algorithm*, referring to the classical scheme studied by Brennan and Schwartz in [8]. The theoretical justification of the Brennan and Schwartz algorithm is a delicate issue that has been treated by Jaillet, Lamberton, and Lapeyre [15, 16] using the theory of quasi-variational inequalities due to Bensoussan and Lions [5, 6]. Viscosity solution theory for quasi-variational inequalities can be found in [22, 23]. The purpose of this section is to show that we can use the mathematical framework for the semilinear Black and Scholes equation to give an elementary convergence proof for the Brennan and Schwartz scheme defined in (5.1).

We have the following theorem:

Theorem 5.1. *Suppose the parabolic CFL condition (4.10) holds. Then the statements in Lemma 4.1 and Theorem 4.1 hold true for the approximate solution $\left\{V_j^n\right\}$ generated by the Brennan and Schwarz algorithm (5.1).*

Proof. The proof follows the lines of the proofs of Lemma 4.1 and Theorem 4.1, and hence we will be rather brief here. We trivially have $V_j^n \geq g_j$ for all j, n, so that (4.11) and hence (4.12) hold. The upper bounds (4.13) and (4.12) can be proved as before. Let us now prove that the upper weak limit $\overline{v}$ is a subsolution of (1.3)–(1.7), starting with the terminal condition (1.7). Since $V_j^{n+1} \geq g_j \;\forall j$, we know from (4.18) that

$$-\left(g(x_j) - V_j^{n+\frac{1}{2}}\right)^+ \geq -\Delta tc(x_j) \;\forall j.$$

Plugging this into (5.1), we get

$$\mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n \geq -\Delta tc(x_j) \;\forall j,$$

where $\mathcal{S}$ is defined in (4.16). Now doing exactly as we did in the proof of Theorem 4.1, we end up with (4.25). It remains to prove that $\overline{v}$ is a subsolution of (1.3). Pick a strict local maximizer (t, x) of $\overline{v} - \phi$ for some $\phi \in C^{1,2}(Q_T)$ with $x > 0$. The case $\overline{v}(t, x) = g(x)$ can treated as in the proof of Theorem 4.1. Let us therefore assume $\overline{v}(t, x) > g(x) + \delta$ for some $\delta > 0$. Then $q^*(x, \overline{v}(t, x)) = 0$. Applying Lemma 2.1, there exist sequences $(t_\Delta, x_\Delta) \to (t, x)$ and $v_\Delta(t_\Delta, x_\Delta) \to \overline{v}(t, x)$ as $\Delta \downarrow 0$ such that (t_Δ, x_Δ) is a local maximizer of $v_\Delta - \phi$ for each Δx. For any Δx sufficiently small, we have $v_\Delta(t_\Delta, x_\Delta) > g(x_\Delta)$. Pick integers (j, n) such that $(t_\Delta, x_\Delta) = (t^n, x_j)$ and $v_\Delta(t_\Delta, x_\Delta) = V_j^n$. Using the monotone function $\mathcal{S}$ defined in (4.16), we can write (5.1) as

$$\mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n = 0.$$

From here on, we can conclude exactly as in the proof of Theorem 4.1.

Let us now prove that the lower weak limit $\overline{v}$ is a supersolution of (1.3)–(1.7). In view of (4.20), we know already that $\underline{v}$ satisfies the terminal condition. Replacing (4.24) by

$$(5.2) \qquad \mathcal{S}\left(x_j, V_{j-1}^{n+1}, V_j^{n+1}, V_{j+1}^{n+1}\right) - V_j^n = -\left(g(x_j) - V_j^{n+\frac{1}{2}}\right)^+ \leq 0,$$

the proof that $\overline{v}$ is a supersolution of (1.3) goes exactly as in the proof of Theorem 4.1. $\qquad\square$

6. A numerical example

In this section, we test the predictor–corrector scheme defined and analyzed in Section 4. In the computer program that have been implemented, we specify the spatial discretization parameter Δx and then choose Δt according to

$$(6.1) \qquad \Delta t \leq \frac{(\Delta x)^2}{\Delta x|r - d|L + \sigma^2 L^2 + (\Delta x)^2 r},$$

so that the parabolic CFL condition (4.10) holds. We present here only numerical results for the put option with $d = 0$. Furthermore, we use the following parameters:

$$r = 0.1, \quad \sigma = 0.2, \quad K = 1, \quad T = 1.0, \quad L = 4, \quad \Delta x = 0.00533, \quad \Delta t = 0.0043.$$

The choice $\Delta x = 0.00533$ corresponds to 75 grid points and $\Delta t = 0.0043$ is chosen according to (6.1). For comparison, we use an "exact solution" computed by the predictor–corrector scheme on a very fine grid. The exact and numerical solutions are displayed in Figure 1. The predictor–corrector scheme gives a fairly good approximation to the exact solution. The largest difference is seen in the early exercise region (which roughly speaking corresponds to "$x > 0.8 - 0.9$"). However, this difference is largely due to the use of a first order upwind discretization of the convection term in the Black and Scholes operator. By using a higher order discretization of the convection term, this difference can be greatly reduced (we will come back to this in future work).

As mentioned before, instead of discretizing the semilinear Black and Scholes equation (1.3), one can discretize the approximate semilinear Black and Scholes equation (3.5). To be concrete, consider (3.5) with q_ε defined via (3.7) and $H_\varepsilon(\xi) = \frac{\varepsilon}{\varepsilon - \xi}$. An explicit finite difference discretization (3.5) with this choice of q_ε produces the following *penalty scheme*:

$$(6.2) \qquad \begin{cases} V_j^n = V_j^{n+1} + \lambda(r - d)x_j \Delta_- V_{j+1}^{n+1} + \mu\dfrac{1}{2}\sigma^2 x_j^2 \Delta_- \Delta_+ V_j^{n+1} \\[2mm] \qquad - \Delta t r V_j^{n+1} + \Delta t r K \dfrac{\varepsilon}{\varepsilon + V_j^{n+1} - (K - x_j)}, \end{cases}$$

where $\varepsilon > 0$ is a parameter that has to be specified. The terminal and boundary conditions for (6.2) are as in Section 4. We note that the penalty scheme (6.2) coincides with the one used in [21]. In [21], it was proved that (6.2) satisfied a discrete analog of the early exercise constraint (1.8) provided the following strengthened CFL condition holds:

$$\Delta t \leq \frac{(\Delta x)^2}{\Delta x r L + \sigma^2 L^2 + r(\Delta x)^2 + \frac{rK}{\varepsilon}(\Delta x)^2}.$$
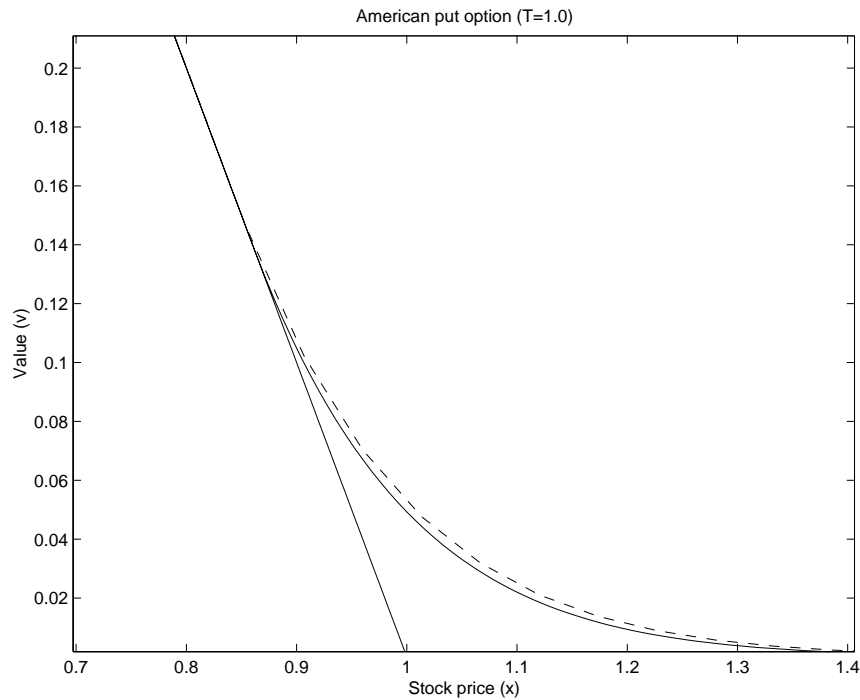
FIGURE 1. The price of the American put option with expiration time $T = 1.0$. The exact solution (along with the payoff function) is shown as solid line and the predictor–corrector solution is shown as a dashed line

Adopting the method of proof herein, one can also prove under this CFL condition that the approximate solutions generated by the penalty scheme (6.2) converge to unique viscosity solution of (1.3)–(1.7) (i.e., the American put option value).

For the numerical example, the penalty scheme used the same discretization parameters as the predictor–corrector scheme. This forces us to choose $\varepsilon = 0.0009$. The exact and numerical solutions are displayed in Figure 2. The penalty solution is comparable to the predictor–corrector solution. However, the "zoom–in plots" in Figure 3 reveals that the penalty solution consistently lies above the predictor–corrector solution. This is not surprising since the reaction term in (6.2) is non-zero (but small) also in the optimal exercise region. Although we did not gain anything here by using a continuous approximation q_ε instead of the discontinuous nonlinear reaction term q in (1.4), it may be an advantage to use q_ε instead of q in a numerical scheme in which the nonlinear reaction term is going to be discretized implicitly. But this remains to be investigated.

REFERENCES

[1] G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance theory. In *Numerical methods in finance*, pages 1–21. Cambridge Univ. Press, Cambridge, 1997.
[2] G. Barles, C. Daher, and M. Romano. Convergence of numerical schemes for parabolic equations arising in finance theory. *Math. Models Methods Appl. Sci.*, 5(1):125–143, 1995.
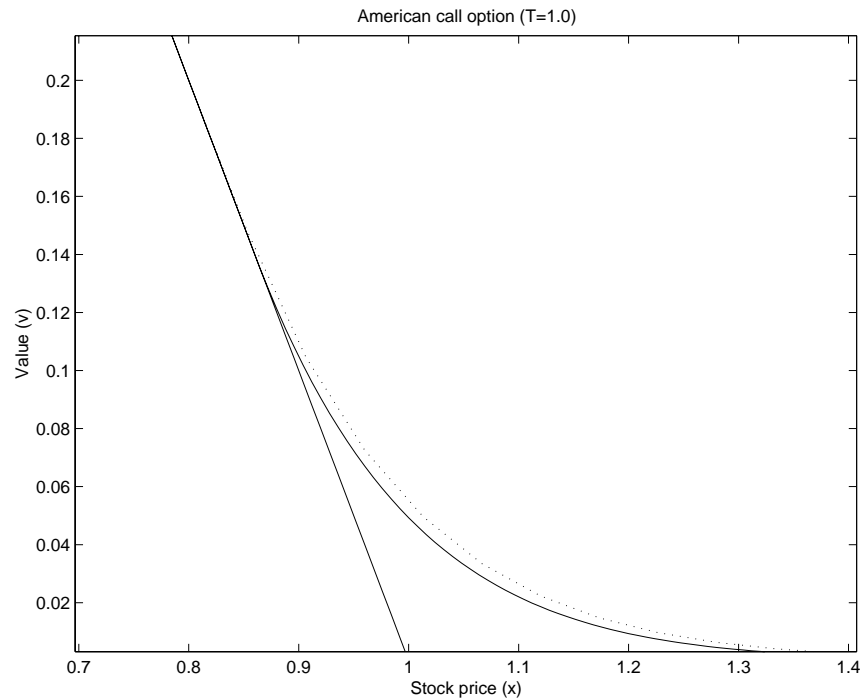
FIGURE 2. The price of the American put option with expiration time $T = 1.0$. The exact solution (along with the payoff function) is shown as solid line and the penalty solution is shown as dotted line.

[3] G. Barles and B. Perthame. Comparison principle for Dirichlet-type Hamilton-Jacobi equations and singular perturbations of degenerated elliptic equations. *Appl. Math. Optim.*, 21(1):21–44, 1990.

[4] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptotic Anal.*, 4(3):271–283, 1991.

[5] A. Bensoussan and J.-L. Lions. *Applications of variational inequalities in stochastic control.* North-Holland Publishing Co., Amsterdam, 1982. Translated from French.

[6] A. Bensoussan and J.-L. Lions. *Impulse control and quasivariational inequalities.* Gauthier-Villars, Montrouge, 1984.

[7] F. E. Benth, K. H. Karlsen, and K. Reikvam. On a semilinear Black and Scholes partial differential equation for valuing American options. Part I: viscosity solutions and well-posedness. Preprint, 2001.

[8] M. J. Brennan and E. S. Schwartz. The valuation of American put options. *J. Finance*, 32:449–462, 1977.

[9] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[10] D. Duffie. *Dynamic asset pricing theory.* Princeton University Press, Princeton, 1996.

[11] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions.* Springer-Verlag, New York, 1993.

[12] P. A. Forsyth and K. R. Vetzal. Quadratic convergence of a penalty method for valuing american options. Preprint, 2001.

[13] H. Ishii. Hamilton-Jacobi equations with discontinuous Hamiltonians on arbitrary open sets. *Bull. Fac. Sci. Engrg. Chuo Univ.*, 28:33–77, 1985.
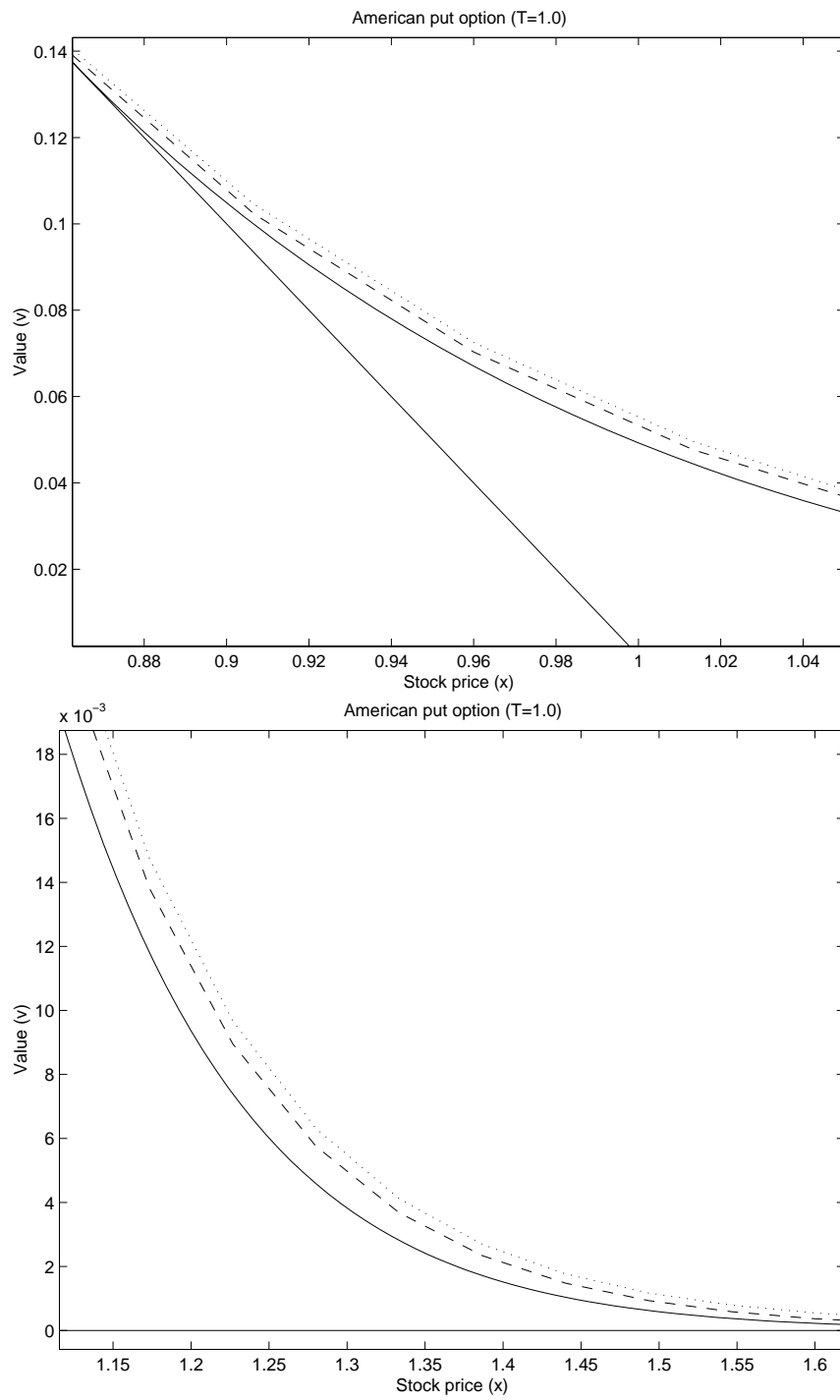
FIGURE 3. "Zoom–in plots" of the exact solution (solid line), the predictor–corrector solution (dashed line), and the penalty solution (dotted line).

[14] H. Ishii. A boundary value problem of the Dirichlet type for Hamilton-Jacobi equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 16(1):105–135, 1989.

[15] P. Jaillet, D. Lamberton, and B. Lapeyre. Inéquations variationnelles et théorie des options. *C. R. Acad. Sci. Paris Sér. I Math.*, 307(19):961–965, 1988.

[16] P. Jaillet, D. Lamberton, and B. Lapeyre. Variational inequalities and the pricing of American options. *Acta Appl. Math.*, 21(3):263–289, 1990.

[17] I. Karatzas and S. E. Shreve. *Methods of mathematical finance*. Springer-Verlag, New York, 1998.

[18] V. A. Kholodnyi. A nonlinear partial differential equation for American options in the entire domain of the state variable. In *Proceedings of the Second World Congress of Nonlinear Analysts, Part 8 (Athens, 1996)*, volume 30, pages 5059–5070, 1997.

[19] M. Musiela and M. Rutkowski. *Martingale methods in financial modelling*. Springer-Verlag, Berlin, 1997.

[20] R. Myneni. The pricing of the American option. *Ann. Appl. Probab.*, 2(1):1–23, 1992.

[21] B. F. Nielsen, O. Skavhaug, and A. Tveito. Penalty and front-fixing methods for the numerical solution of American option problems. *J. Comp. Finance*. To appear.

[22] B. Øksendal and K. Reikvam. Viscosity solutions of optimal stopping problems. *Stochastics Stochastics Rep.*, 62(3-4):285–301, 1998.

[23] H. Pham. Optimal stopping of controlled jump diffusion processes: a viscosity solution approach. *J. Math. Systems Estim. Control*, 8(1):27 pp. (electronic), 1998.

[24] A. N. Shiryayev. *Optimal stopping rules*. Springer-Verlag, New York, 1978.

[25] R. Zvan, P. A. Forsyth, and K. R. Vetzal. Penalty methods for American options with stochastic volatility. *J. Comput. Appl. Math.*, 91(2):199–218, 1998.

(Fred E. Benth)
DEPARTMENT OF MATHEMATICAL SCIENCES
NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY.
N–7491 TRONDHEIM, NORWAY
   E-mail address: benth@math.ntnu.no
   URL: http://www.math.ntnu.no/~benth/

(Kenneth H. Karlsen)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BERGEN
JOHS. BRUNSGT. 12
N–5008 BERGEN, NORWAY
   E-mail address: kennethk@mi.uib.no
   URL: http://www.mi.uib.no/~kennethk/

(Kristin Reikvam)
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF OSLO
P.O. BOX 1053, BLINDERN
N–0316 OSLO, NORWAY
   E-mail address: kre@math.uio.no
   URL: http://www.math.uio.no/~kre/