Computers and Mathematics with Applications 62 (2011) 1–26

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

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

A predictor–corrector scheme based on the ADI method for pricing American puts with stochastic volatility

Song-Ping Zhu *, Wen-Ting Chen

School of Mathematics and Applied Statistics, University of Wollongong NSW 2522, Australia

ARTICLE INFO

Article history:
Received 30 June 2010
Received in revised form 21 October 2010
Accepted 28 March 2011

Keywords:
ADI scheme
Stability analysis
American put options
Heston model

ABSTRACT

In this paper, we introduce a new numerical scheme, based on the ADI (alternating direction implicit) method, to price American put options with a stochastic volatility model. Upon applying a front-fixing transformation to transform the unknown free boundary into a known and fixed boundary in the transformed space, a predictor–corrector finite difference scheme is then developed to solve for the optimal exercise price and the option values simultaneously. Based on the local von Neumann stability analysis, a stability requirement is theoretically obtained first and then tested numerically. It is shown that the instability introduced by the predictor can be damped, to some extent, by the ADI method that is used in the corrector. The results of various numerical experiments show that this new approach is fast and accurate, and can be easily extended to other types of financial derivatives with an American-style exercise.

Another key contribution of this paper is the proposition of a set of appropriate boundary conditions, particularly in the volatility direction, upon realizing that appropriate boundary conditions in the volatility direction for stochastic volatility models appear to be controversial in the literature. A sound justification is also provided for the proposed boundary conditions mathematically as well as financially.

1. Introduction

It is well known that one of the most important topics in quantitative finance research is the valuation of option derivatives. Empirical evidence suggests that the Black–Scholes model, which is a breakthrough in the financial area, is inadequate to describe asset returns and the behavior of the option markets [1]. This is because their assumption on the log-normality of the value of the underlying asset has somewhat oversimplified the real process of the asset price. One possible remedy is to assume that the volatility of the asset price also follows a stochastic process [2–5]. In this paper, we will use the stochastic model introduced by Heston for pricing American options [4]. In this model, it assumes that the variance (the square of the underlying price volatility) follows a random process known in financial literature as the Cox–Ingersoll–Ross (CIR) process and in mathematical statistics as the Feller process [3,6]. Empirical studies suggest that this non-negative and mean-reverting process is indeed more consistent with what has been observed in real markets [7–9]. For example, Adrian and Victor [1] showed that the time-dependent probability distribution of the changes of the stock index generated in the Heston model agrees well with the Dow-Jones data after the calibration of the parameters in this model.

How to rationally price an option remains one of the major challenges in today's finance industry. This is even for so for pricing American options as the challenge stems from the nonlinearity originated from the inherent characteristics that an American option can be exercised at any time during its lifespan and thus the additional right of being able to exercise the option early, in comparison with a European option, casts the problem into a free boundary problem, which is far more difficult to deal with, even under the traditional Black–Scholes framework. In this area, there have been predominantly two kinds of approaches, numerical methods and analytical approximations, for the valuation of American

* Corresponding author. Fax: +61 2 42214845.
E-mail addresses: spz@uow.edu.au (S.-P. Zhu), wc904@uow.edu.au (W.-T. Chen).

0898-1221/$ – see front matter
doi:10.1016/j.camwa.2011.03.101

options under the Black–Scholes framework. While the former typically includes the finite difference method [10,11], the finite element method [12], the binomial method [13] and the Monte Carlo simulation method [14], the latter includes the Richardson extrapolation approximation [15], the Laplace transform method [16], the algebraic equation method [17,18] and the integral-equation method [19–22]. Furthermore, recently, an analytical solution for American puts on a non-dividend underlying asset was even found out by Zhu [23]. It must be pointed out that all these approaches cannot be easily extended to the Heston model, primarily due to the fact that, under stochastic volatility, the optimal exercise price depends, in addition to time, on the dynamics of volatility. In other words, the introduction of a second stochastic process has considerably complicated the solution process in pricing American options.

In the last decade, several numerical approaches based on the finite difference method (FDM) are introduced to solve the free boundary problem associated with the valuation of American options under the Heston model. For instance, Clarke and Parrott [24], used a special version of a projected full approximation scheme with multigrid to solve the American option pricing problem. One advantage of such a multilevel method is that the number of iterations required to solve a linear complementarity problem is essentially independent of the grid size. However, their method is rather complicated because of the use of a special projected linear Gauss–Seidel smoother. Ikonen and Toivanen [25] calculated the option values by using the operator splitting method, in which an auxiliary variable is used to improve the accuracy. However, their method still requires a relatively large amount of computational storage space. Zvan et al. [26] applied the penalty method to the American option pricing problem. Their method is simpler than the one used by Clark and Parrott, but still needs a relatively large amount of computational resources to produce an accurate result.

Since most of the numerical methods in the literature are either too complicated to implement or with very low computational efficiency, it is desirable to have alternative ways to deal with the valuation of American options with stochastic volatility. In this paper, we propose an approach based on a predictor–corrector framework, which is commonly used to numerically solve nonlinear partial differential equations (PDEs). The idea behind the predictor–corrector method is to use a suitable combination of an explicit and an implicit technique to obtain a method with better convergence characteristics. Previously, this scheme has only been applied to the pricing problem under the Black–Scholes model, such as Zhu and Zhang reported in [11]. Though their method is efficient and accurate, it cannot be applied directly to the stochastic models. The purpose of this article is to introduce a new predictor–corrector scheme, which is not only suitable for the Heston model, but also for other stochastic models. In our new approach, we adopt the so-called front-fixing transformation [27] to let the unknown boundary be included in the governing equation as a nonlinear term in exchange for a fixed boundary. To tackle the nonlinear nature of American option pricing problem, which is explicitly exposed in the transformed equation, we use a predictor–corrector finite difference scheme at each time step to convert the nonlinear PDE to two linearized difference equations associated with the prediction and correction phase, respectively. The prediction phase, constructed by an explicit Euler scheme, is used to calculate the optimal exercise price, whereas the correction phase, designed by the alternating direction implicit, or ADI, method, continues to do the calculation of the option price together with the correction of the optimal exercise price. The ADI scheme used in the corrector is efficient in computing multi-dimensional problems. Moreover, it is also suggested that the good convergence property of the ADI scheme can somehow, damp the instability that might be introduced by the predictor. With the perfect combination of the explicit Euler scheme and the ADI method, the originally nonlinear problem has been successfully converted to a set of linear algebraic equations, which can be solved efficiently. In comparison with the numerical methods in the literature, the advantage of the current scheme is obvious. For example, first, our method requires almost the same storage space as a one-dimensional problem does and it will not increase even when the method is applied to option pricing problems on multi-assets. This is, however, not the case for the numerical methods proposed in [25], as a substantially larger amount of the storage space is required, which will also increase as the number of the assets increases. Second, in addition to the option values, the present method captures the entire optimal exercise boundary as part of the solution procedure, whereas in [25], the optimal exercise price cannot be obtained simultaneously, and needs to be solved with some extra effort. Finally, our method requires no iterations, and can be easily extended to the valuation of American options under other models.

It is usually easy to design a numerical scheme to solve a PDE system, but much harder to provide a theoretical threshold for the stability and the convergence of the scheme. It is probably even worse to theoretically define a suitable stability criterion for the predictor–corrector method, since it is a hybrid finite difference method. For this reason, the issue of stability requirement was not even attempted in [11] for the Black–Scholes case. One could naturally imagine that with the complexity of the Heston model, it would have made a theoretical stability analysis much less achievable. Based on the local von Neumann stability analysis, combining with the "frozen" coefficient technique, which is commonly used for the stability analysis of the variable-coefficient problem [28], we have managed to not only verify that the ADI method used for the European puts under the Heston model is unconditionally stable, but also give a proper stability requirement for the predictor–corrector approach.

In the subsequent sections, we will present this new approach together with the numerical results for American put options under the Heston model. The paper is organized as follows: in Section 2, we introduce the PDE system that the price of an American put must satisfy under the Heston model, with our emphasis being placed on properly closing the system with appropriate boundary conditions, which appear to be controversial in the literature. In Section 3, we present our predictor–corrector approach in detail as well as the implementation of the ADI scheme. In Section 4, numerical examples and some analyses are presented to demonstrate the convergence and accuracy of the current scheme. Concluding remarks are given in the last section.

2. Pricing American options under the Heston model

Although the Heston model has been studied by a number of authors [7–9], we still describe it in reasonable detail, in this section, for the sake of completeness of the paper and easiness of reference for the readers. However, our emphasis will be placed on the discussion of appropriate boundary conditions in the volatility direction, which contributes greatly to the proper close of the PDE system to be solved numerically in the later sections.

2.1. The Heston model

Heston [4], assumed that the underlying S_t, as a function of time, follows the stochastic differential equation (SDE) of a geometric Brownian motion in the Itô form:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dw_1, \tag{2.1}$$

where μ is the drift rate, w_1 is a standard Brownian motion, and $\sqrt{v_t}$ is the standard deviation of the stock returns $\frac{dS_t}{S_t}$. Furthermore, he assumed that the variance v_t (the square of the volatility) is governed by the following mean-reverting SDE

$$dv_t = \kappa(\eta - v_t)dt + \sigma\sqrt{v_t}dw_2. \tag{2.2}$$

Here, η is the long time mean of v_t, κ is the rate of relaxation to this mean, and σ is the volatility of the volatility. w_2 is also a standard Brownian motion, and it is related to w_1 with a correlation factor $\rho \in [-1, 1]$. Eq. (2.2) is known in the financial literature as the Cox–Ingersoll–Ross (CIR) process and in mathematical statistics as the Feller process [3,6]. Various studies [7–9] suggest that it is consistent with the real market. Most remarkably, Heston found a closed-form exact solution for the price of European-style options. Unfortunately, the approach he adopted could not be easily extended to the case of American options; no analytical solution for American options under the Heston Model has been discovered yet.

2.2. The PDE for the value of American puts and the corresponding boundary conditions

Let $U(S, v, t)$ denote the value of an American put option, with S being the price of the underlying asset, v being the variance and t being the time. For simplicity, we assume that the underlying pays no dividend. Under the Heston Model, it can be easily shown that under the risk-neutral argument, the value of a put option U should satisfy the following bivariate PDE:

$$\frac{1}{2}vS^2\frac{\partial^2 U}{\partial S^2} + \rho\sigma vS\frac{\partial^2 U}{\partial S\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2} + rS\frac{\partial U}{\partial S} + \left[\kappa(\eta - v) - \lambda\sigma\sqrt{v}\right]\frac{\partial U}{\partial v} - rU + \frac{\partial U}{\partial t} = 0, \tag{2.3}$$

where λ is the market price of risk, r is the risk-free interest rate. In this paper, for simplicity, we set λ to zero, and the extensions to the case that λ is non-zero should be straightforward. The terminal condition to Eq. (2.3) is given by the payoff function. For an American put option, it reads: $U(S, v, T) = \max(K - S, 0)$, where K is the strike price. Since at the expiry date, the optimal exercise price of an American put on a non-dividend underlying asset is equal to the strike price, this condition can be simplified as [23]:

$$U(S, v, T) = 0, \quad \text{for } S > S_f(v, T) = K,\ v > 0. \tag{2.4}$$

For the valuation of American puts, a set of appropriate boundary conditions is also needed together with the terminal condition (2.4) to solve Eq. (2.3). It is obvious that the boundary conditions along the S direction are easy to justify. They are just the same as those in the Black–Scholes model. The value of a put should satisfy the far-field boundary condition,

$$\lim_{S\to\infty} U(S, v, t) = 0,$$

which means that when the price of the underlying becomes extremely large, a put option becomes worthless. On the other hand, just as in the Black–Scholes model, there is a critical asset price, below or equal to which it is optimal to exercise the put option. It can be shown, under the no-arbitrage argument, that the boundary conditions at the optimal exercise boundary $S = S_f$ are [29]:

$$U(S_f, v, t) = K - S_f,$$

$$\frac{\partial U}{\partial S}(S_f, v, t) = -1.$$

It should be noted that the above two conditions look very similar to the case with constant volatility. However, the main difference between the constant volatility model and the stochastic volatility model lies in the fact that in the latter case, such as in the Heston model, the optimal exercise price S_f depends, in addition to time, on the dynamics of the volatility. In other words, S_f is a function of both v and t.

The boundary conditions along the v direction remain unclear in the literature. Even for the European case, it is still controversial whether or not Heston's analytical formula [4] does indeed satisfy the given boundary conditions along the v direction. While some (e.g. [30]) chose the boundary conditions along the v direction, for European puts, by taking the limit of Black–Scholes' formula with respect to σ, others (e.g. [31]) chose to neglect the boundary conditions along the v direction altogether for American puts. From the financial point of view, there is no reason why the boundary conditions in the v direction should be different when the moving boundary for the case of American puts is associated with the S direction only. There were also some authors [24,25,32] who argued that boundary conditions along the v direction are still necessary in their numerical approaches for solving the American option pricing problem under the Heston model and chose the two boundary conditions:

$$\lim_{v\to 0} U(S, v, t) = \max(K - S, 0), \tag{2.5}$$

$$\lim_{v\to \infty} \frac{\partial U}{\partial v}(S, v, t) = 0. \tag{2.6}$$

However, they did not explain exactly why the option price should be equal to the payoff function at $v = 0$. As different boundary conditions imposed will undoubtedly affect the value of an option, the controversy of what boundary conditions should be imposed in the v direction has clearly jeopardized the uniqueness of the solution and thus needs to be properly investigated. In the following, we will discuss what special price the option should be along the boundary of the v direction.

Whether or not boundary conditions are needed for $v = 0$ and if they are needed, what would be their appropriate form, should be discussed from both mathematical and financial points of view.

Mathematically, $v = 0$ is a so-called "degenerate" boundary of Eq. (2.3), because its characteristic form vanishes as v approaches zero [33]. From the mathematical point of view, boundary conditions along degenerate boundaries are not needed at all if the Fichera function is nonnegative, but should be imposed otherwise [33]. For the Heston model, it can be shown that the Fichera function along $v = 0$ equals $\kappa\eta - \frac{\sigma^2}{2}$. Therefore, if $\kappa\eta \geq \frac{\sigma^2}{2}$, the pricing system without boundary constraints at $v = 0$ is already closed, and thus there is no need to prescribe any condition along the boundary $v = 0$ at all for this case. On the other hand, if $\kappa\eta < \frac{\sigma^2}{2}$, appropriate boundary conditions at $v = 0$ are still "needed" to ensure the uniqueness of the solution. However, the Fichera function does not reveal what the specific boundary conditions should be prescribed for this case. We believe that one now has to use a financial argument to set up an appropriate boundary condition for this case.

Under the risk-neutral argument, when $v \to 0$, the leading order term of the solution of the SDE (2.1) is $S = e^{rt}S_0$, i.e., the underlying becomes virtually riskless, and its price should appreciate at a deterministic rate r when $v \to 0$. Therefore, if $S < K$, the put option should be immediately exercised as there is no reason to hold the option any more if one knows that the underlying will *definitely* increase for sure. In other words, the underlying price range $[0, K)$ belongs to the "exercise" region, i.e., $[0, K) \subseteq [0, \lim_{v\to 0} S_f(v, t)]$, and thus $\lim_{v\to 0} S_f(v, t) \geq K$. On the other hand, if $S > K$, it is obvious that the value of the put option becomes zero, and therefore it is better to hold the option as the option may still have some time value before its expiration date is reached. That is to say, the underlying price range (K, ∞) belongs to the "continuous" region, i.e., $(K, \infty) \subseteq (\lim_{v\to 0} S_f(v, t), \infty)$, and thus $\lim_{v\to 0} S_f(v, t) \leq K$. A combination of the above two statements leads to the conclusion that

$$\lim_{v\to 0} S_f(v, t) = K. \tag{2.7}$$

Having established that $S_f(v, t)$ must be equal K when $v \to 0$, it is then trivial to show

$$\lim_{v\to 0} U(S, v, t) = 0 \tag{2.8}$$

from the definition of S_f associated with put options.

Two points need to be remarked. Firstly, Eq. (2.8) is a simplified version of Eq. (2.5); it can be deduced from Eq. (2.5) by considering the definition of S_f, as well as the fact that $\lim_{v\to 0} S_f(v, t) = K$. Therefore, we believe those [24,32] who proposed to use Eq. (2.5) as the boundary condition at $v \to 0$ are correct, although we argue that it is better to use the simplified version Eq. (2.8). Second, this boundary condition based on a financial argument is irrespective of the ratio $\frac{2\kappa\eta}{\sigma^2}$ being greater than 1 or not. Naturally, one may wonder the consistency of the two arguments for the case of $\kappa\eta \geq \frac{\sigma^2}{2}$, for which the mathematical argument shows that there is no need for any boundary condition at $v = 0$, whereas the financial argument suggests that the option value vanishes there. In other words, if one also adopts the boundary condition Eq. (2.8) for the case $\kappa\eta \geq \frac{\sigma^2}{2}$, will this value be consistent with the inherent value of the PDE system when $v \to 0$, and thus would not be an inappropriate "boundary condition" that will certainly "deteriorate" the well-posedness of the pricing problem? Note that the PDE system here refers to (2.16) but without boundary condition at $v = 0$, which is already closed for this case. Fortunately, we have managed to prove the consistency of Eq. (2.8) with the asymptotic behavior of the unknown function $U(S, v, t)$ as $v \to 0$ when $\kappa\eta \geq \frac{\sigma^2}{2}$. For briefness, we leave such a proof in Appendix A.

Combining both arguments, we can now confidently conclude that Eq. (2.8) should be used as the boundary condition at $v \to 0$ for pricing an American put option, regardless of what the ratio $\frac{2\kappa\eta}{\sigma^2}$ is.

We now show what the boundary condition should be imposed as $v \to \infty$. One needs to understand how the volatility impacts on the option price, since v in the Heston model is nothing but the square of the volatility. Roughly speaking, the volatility of the underlying asset is a measure of the uncertainty of the future underlying price movements. As a result, if volatility increases, the probability of the underlying asset price varying within a large range would increase too, resulting in higher option prices for both puts and calls, since the holder of the option will have a chance to cash in when the movement of the underlying is in his favor and do nothing when it is not. Mathematically, this is equivalent to saying that the option price $U(S, v, t)$ is a monotonic increasing function with respect to v.

Like the case that the boundary conditions for European puts and American puts, in the S direction, are different at the lower boundary of the computational domain, those associated with European puts and American puts in the v direction are of different forms as well; they need to be discussed separately.

First, we consider the case of European puts. Market observations show that when the volatility is extremely large, the option price U is independent of volatility changes [24], i.e.,

$$\lim_{v\to\infty} \frac{\partial U}{\partial v}(S, v, t) = 0. \tag{2.9}$$

In fact, this has been used by many authors as their boundary condition for the option price at the large end of v (e.g. [24, 25]). However, we believe that under the same argument we should be able to deduct an even simpler boundary condition that not only reflects the fact that the vega approaches zero when $v \to \infty$, but also considerably facilitates the computation. This is achieved by realizing the fact that when $v \to \infty$, the second-order partial derivative of $U(S, v, t)$ with respect to v should vanish too if the vega approaches zero, i.e.,

$$\lim_{v\to\infty} \frac{\partial^2 U}{\partial v^2}(S, v, t) = 0. \tag{2.10}$$

Then, utilizing Eqs. (2.9) and (2.10), the Heston operator

$$\mathcal{L}_{\mathcal{H}} = \frac{1}{2}vS^2\frac{\partial^2}{\partial S^2} + \rho\sigma vS\frac{\partial^2}{\partial S\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2} + rS\frac{\partial}{\partial S} + \kappa[\eta - v]\frac{\partial}{\partial v} - r + \frac{\partial}{\partial t},$$

degenerates to be the Black–Scholes operator, with σ^2 being substituted by v,

$$\mathcal{L}_{\mathcal{BS}} = \frac{1}{2}vS^2\frac{\partial^2}{\partial S^2} + rS\frac{\partial}{\partial S} - r + \frac{\partial}{\partial t}.$$

This implies that we can use the option value calculated from the Black–Scholes formula as the boundary value of $U(S, v, t)$ when $\sigma \to +\infty$ under the Heston model. In other words, through this argument, we have successfully converted the Neumann boundary condition Eq. (2.9) into a Dirichlet boundary condition

$$\lim_{v\to\infty} U(S, v, t) = Ke^{-r(T-t)}, \tag{2.11}$$

the implementation of which would require far less computational effort. It should be remarked that Eq. (2.11) is a special case of Eq. (2.9), which means that the solution obtained by satisfying Eq. (2.11) would automatically satisfy Eq. (2.9), but not vice versa.

Now, we consider the case of American puts. To better articulate the establishment of appropriate boundary condition for this case, we form and prove the following proposition.

Proposition 1. *When v approaches infinity, the value of an American put option reaches the strike price K asymptotically, i.e., $\lim_{v\to\infty} U_A(S, v, t) = K$.*

Proof. First, we consider the effect of the expiration date on the option prices. It is a well-known fact that the value of American puts is an increasing function of the time to expiry, simply because the longer the tenor of an option is, the more "right" the holder has in terms of exercising the option [34]. Mathematically, this is equivalent to saying that the option price $U_A(S, v, t)$ is a monotonic decreasing function of t, i.e.,

$$\lim_{v\to\infty} U_A(S, v, T) \le \lim_{v\to\infty} U_A(S, v, t) \quad \forall t \in [0, T]. \tag{2.12}$$

Moreover, an American put is always worth at least as much as its European counterpart [34], i.e.,

$$\lim_{v\to\infty} U_E(S, v, T) \le \lim_{v\to\infty} U_A(S, v, T). \tag{2.13}$$

From Eq. (2.11), we have

$$\lim_{v\to\infty} U_E(S, v, T) = Ke^{-r(T-t)}|_{t=T} = K,$$

which, combined with the inequalities (2.12) and (2.13), yields,

$$K \le \lim_{v\to\infty} U_A(S, v, t). \tag{2.14}$$

On the other hand, it is also well known that no matter what the underlying value becomes, the price of an American put option can never be worth more than its strike price [34], i.e.,

$$\lim_{v\to\infty} U_A(S, v, t) \le K. \tag{2.15}$$

Clearly, from Eqs. (2.14) and (2.15) one must reach the conclusion

$$\lim_{v\to\infty} U_A(S, v, t) = K. \quad \square$$

This can also be understood from the financial point of view. From the definition of a put option, its value is bounded up by the strike price K [34]. Therefore, if one can show that K is also the least upper bound, one must then conclude that option price must reach K when $v \to \infty$, based on another financial intuition that any option value monotonically increases with volatility. The "least upper boundness" of K can be easily established by arguing that if there was another upper bound that is less than K, since the infinite volatility implies that the underlying may take any value within the range of $S \in [0, \infty)$ with a probability one, the option value can always be greater than this "upper bound". A contradiction is thus reached.

In summary, the properly-closed PDE system for pricing American put options under the Heston model can be written as:

$$\begin{cases} \dfrac{1}{2}vS^2\dfrac{\partial^2 U}{\partial S^2} + \rho\sigma vS\dfrac{\partial^2 U}{\partial S\partial v} + \dfrac{1}{2}\sigma^2 v\dfrac{\partial^2 U}{\partial v^2} + rS\dfrac{\partial U}{\partial S} + \kappa(\eta - v)\dfrac{\partial U}{\partial v} - rU + \dfrac{\partial U}{\partial t} = 0, \\ U(S, v, T) = 0, \\ \lim\limits_{S\to\infty} U(S, v, t) = 0, \\ U(S_f(v, t), v, t) = K - S_f(v, t), \\ \dfrac{\partial U}{\partial S}\Big(S_f(v, t), v, t\Big) = -1, \\ \lim\limits_{v\to 0} U(S, v, t) = 0, \\ \lim\limits_{v\to\infty} U(S, v, t) = K, \end{cases} \tag{2.16}$$

for $S \in [S_f(v, t), \infty)$, $v \in [0, \infty)$, and $t \in [0, T]$. There are several remarks before we introduce an efficient and accurate numerical scheme to solve this system in the next section. First of all, the two newly-introduced boundary conditions in the v direction have coincidentally well manifested the monotonicity of the option price with respect to v as well as its boundedness [34]:

$$\max(K - S, 0) \le U(S, v, t) \le K.$$

Or, the option price is expected to monotonically increase from its lower bound $\max(K - S, 0)$ to its upper bound K when v varies from 0 to ∞ (one should not forget that when $v = 0$, the solution domain of the above differential system is $S \in [K, \infty)$ as we have shown before already). Second, while the last boundary condition is a special case of that adopted by some authors (e.g. [24,25,32]), the fact that this is now a Dirichlet boundary condition rather than a Neumann boundary condition as used in (e.g. [24,25,32]) has considerably facilitated the numerical solution procedure to be shown in the next section. In other words, the solution satisfying the differential system (2.16) must also satisfy those obtained with the last boundary condition replaced by the Neumann boundary condition (2.9), but not vice versa. Finally, one should notice that there exists a singularity at the corner where $S \to \infty, v \to \infty$, in addition to the well-known one along the moving boundary $S = S_f$. However, unlike the latter, the former does not cause any computational difficulties as it is a simple and removable singularity that commonly exists in diffusion problems with different boundary values prescribed on two adjacent boundaries.

3. Numerical method based on the ADI scheme

Upon establishing a closed differential system (2.16) for the price of American puts under the Heston model, we propose a new predictor–corrector approach based on the ADI method, in this section, to solve this system in two phases within a time step: a prediction phase in which a rough value of the optimal exercise price S_f is calculated, and a correction phase in which the option value U as well as the final value of S_f is determined.

3.1. Coordinate transformation

In order to solve the PDE system (2.16) effectively, we introduce a new variable as the time to expiration: $\tau = T - t$. It should be noted that the backward problem (2.16) has been changed to an initial value problem (3.1) after the introduction of τ.

If we want to solve PDE system (2.16) directly using FDM, some kinds of iterative methods may be adopted because of the existence of the free boundary. To avoid iterations, we first adopt the Landau transform [27], i.e.,

$$x = \ln\left(\frac{S}{S_f}\right)$$

to shift the moving boundary conditions to fixed boundary conditions before applying our predictor–corrector method.

After some rather simple algebraic manipulations, the PDE system (2.16) can be written as: for $x \in [0, \infty)$, $v \in [0, \infty)$, $\tau \in [0, T]$

$$\begin{cases} \mathbb{L}U = 0, \\ U(x, v, 0) = 0, \\ \lim_{x\to\infty} U(x, v, \tau) = 0, \\ U(0, v, \tau) = K - S_f(v, \tau), \\ \frac{\partial U}{\partial x}\left(0, v, \tau\right) = -S_f(v, \tau), \\ \lim_{v\to 0} U(x, v, \tau) = 0, \\ \lim_{v\to\infty} U(x, v, \tau) = K, \end{cases} \tag{3.1}$$

where

$$\mathbb{L} = \left[\frac{1}{2}v + \frac{1}{2}\frac{\sigma^2 v}{S_f^2}\left(\frac{\partial S_f}{\partial v}\right)^2 - \frac{\rho\sigma v}{S_f}\frac{\partial S_f}{\partial v}\right]\frac{\partial^2}{\partial x^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2} + \left(\rho\sigma v - \frac{\sigma^2 v}{S_f}\frac{\partial S_f}{\partial v}\right)\frac{\partial^2}{\partial x\partial v}$$
$$+ \left[-\frac{1}{2}v + \frac{1}{2}\frac{\sigma^2 v}{S_f^2}\left(\frac{\partial S_f}{\partial v}\right)^2 - \frac{1}{2}\frac{\sigma^2 v}{S_f}\frac{\partial^2 S_f}{\partial v^2} + r - \kappa(\eta - v)\frac{1}{S_f}\frac{\partial S_f}{\partial v} + \frac{1}{S_f}\frac{\partial S_f}{\partial \tau}\right]\frac{\partial}{\partial x} + \kappa(\eta - v)\frac{\partial}{\partial v} - r - \frac{\partial}{\partial \tau}.$$

In order to simplify the notation of $\mathbb{L}$, we introduce three new notations

$$\xi = \frac{1}{S_f}\frac{\partial S_f}{\partial v}, \qquad \beta = \frac{1}{S_f}\frac{\partial^2 S_f}{\partial v^2}, \qquad \lambda = \frac{1}{S_f}\frac{\partial S_f}{\partial \tau},$$

so that $\mathbb{L}$ can be written as:

$$\mathbb{L} = a(v)\frac{\partial^2}{\partial x^2} + b(v)\frac{\partial^2}{\partial v^2} + c(v)\frac{\partial^2}{\partial x\partial v} + (d(v) + \lambda)\frac{\partial}{\partial x} + e(v)\frac{\partial}{\partial v} - r - \frac{\partial}{\partial \tau},$$

where

$$a(v) = \frac{1}{2}v + \frac{1}{2}\sigma^2\xi^2 v - \rho\sigma v\xi, \qquad b(v) = \frac{1}{2}\sigma^2 v, \qquad c(v) = \rho\sigma v - \sigma^2 v\xi,$$
$$d(v) = -\frac{1}{2}v + \frac{1}{2}\xi^2\sigma^2 v - \frac{1}{2}\sigma^2 v\beta + r - \kappa(\eta - v)\xi, \qquad e(v) = \kappa(\eta - v).$$

One should notice that after the Landau transform, the nonlinear nature of the problem is explicitly exposed in the transformed equation. It consists the optimal exercise price, which is also part of the solution. Before our new predictor–corrector approach can be applied, we should discretize the highly nonlinear PDE system (3.1), which will be demonstrated in detail in the next subsection.

3.2. Discretization of the PDE system

Now the option pricing problem is defined on an unbounded domain

$$\left\{(x, v, \tau)|x \ge 0, v \ge 0, \tau \in [0, T]\right\}.$$

In order to use finite difference approximation for spacial variables, we need to truncate the semi-infinite domain into a finite domain:

$$\left\{(x, v, \tau) \in [0, x_{\max}] \times [0, v_{\max}] \times [0, T]\right\}. \tag{3.2}$$

Theoretically, $x_{\max}$ and $v_{\max}$ should be sufficiently large to eliminate the boundary effect. However, based on Willmott et al.'s estimate [29] that the upper bound of the asset price $S_{\max}$ is typically three or four times the strike price, it is reasonable for us to set $x_{\max} = \ln 5$. On the other hand, the volatility value is usually very small. The highest value of the volatility that has ever been recorded on Chicago Board Options Exchange (CBOE) is only 0.85 [35]. Thus, it is quite reasonable to set $v_{\max} = 1$, and this has also been the case in many previous studies, (e.g. [25]).

The discretization is performed by placing a set of uniformly distributed grids in the computation domain (3.2). With the number of steps in the x, v and τ directions being denoted by N_x, N_v and N_τ, respectively, the step sizes are correspondingly $\Delta x = \frac{X_{\max}}{N_x}$, $\Delta v = \frac{V_{\max}}{N_v}$ and $\Delta\tau = \frac{T}{N_\tau}$. The value of the unknown function U at a grid point is thus denoted by

$$U_{i,j}^n \approx U(x_i, v_j, \tau_n) = U(i\Delta x, j\Delta v, n\Delta\tau),$$

where $i = 0, \ldots, N_x$; $j = 0, \ldots, N_v$; $n = 0, \ldots, N_\tau$.

The discretization of the PDE system (3.1) needs to be conducted both in the interior domain

$$\Omega = \left\{(i\Delta x, j\Delta v)|i = 1\cdots N_x - 1, j = 1\cdots N_v - 1\right\},$$

and along the boundary $\partial\Omega = \partial_x\Omega \bigcup \partial_v\Omega \bigcup \partial_{xv}\Omega$, in which

$$\partial_x\Omega = \left\{(i\Delta x, j\Delta v)|i = 0, N_x, j = 1\cdots N_v - 1\right\},$$
$$\partial_v\Omega = \left\{(i\Delta x, j\Delta v)|i = 1\cdots N_x - 1, j = 0, N_v\right\},$$
$$\partial_{xv}\Omega = \left\{(i\Delta x, j\Delta v)|i = 0, N_x, j = 0, N_v\right\}.$$

For those grid points that belong to Ω, we use the standard second-order central difference scheme to approximate the first-order and second-order non-cross spatial derivatives. The cross-derivative term, on the other hand, is discretized as:

$$(\delta_{xv}U)_{i,j} = \frac{(\delta_x U)_{i,j+1} - (\delta_x U)_{i,j-1}}{2\Delta v},$$

where

$$(\delta_x U)_{i,j+1} = \frac{U_{i+1,j+1} - U_{i-1,j+1}}{2\Delta x}, \qquad (\delta_x U)_{i,j-1} = \frac{U_{i+1,j-1} - U_{i-1,j-1}}{2\Delta x}.$$

Next, we consider the discretization of the grid points that belong to $\partial\Omega$. The treatment of the Dirichlet boundary conditions is quite standard. However, it is a bit more difficult to deal with the Neumann boundary condition. In the literature, there are usually two different approaches for such kinds of boundary conditions with second-order accuracy [28]. The first one is to introduce a fictitious grid point $U^n_{-1,j}$, and approximate the Neumann boundary condition at $x = 0$ with the central difference,

$$\frac{\partial U^n_{0,j}}{\partial x} = \frac{U^n_{1,j} - U^n_{-1,j}}{2\Delta x}.$$

Then, together with the assumption that the governing equation in the PDE system (3.1) is also satisfied at the boundary point $U^n_{0,j}$, we obtain two different algebraic equations, from which $U^n_{-1,j}$ can be eliminated. The details of the algebraic manipulations are given in [11]. In essence, this method is to use the grid point $U^n_{1,j}$ and some extra information from the PDE, to approximate the derivative $\frac{\partial U^n_{0,j}}{\partial x}$.

The second approach is to use the so-called one-sided difference, which is, in essence, a form of extrapolation that determines the value of the unknown function on the boundary in terms of its values at the interior grid points [28]. From the Taylor series, we obtain the approximation:

$$U^n_{1,j} \approx U^n_{0,j} + \Delta x\frac{\partial U^n_{0,j}}{\partial x} + \frac{1}{2}(\Delta x)^2\frac{\partial^2 U^n_{0,j}}{\partial x^2} + \mathcal{O}((\Delta x)^3), \tag{3.3}$$

$$U^n_{2,j} \approx U^n_{0,j} + 2\Delta x\frac{\partial U^n_{0,j}}{\partial x} + 2(\Delta x)^2\frac{\partial^2 U^n_{0,j}}{\partial x^2} + \mathcal{O}((\Delta x)^3). \tag{3.4}$$

By eliminating the second-order derivatives in Eqs. (3.3) and (3.4), we obtain

$$\frac{\partial U^n_{0,j}}{\partial x} = \frac{4U^n_{1,j} - U^n_{2,j} - 3U^n_{0,j}}{2\Delta x} + \mathcal{O}((\Delta x)^3).$$

In short, this method is to use the interior grid values $U^n_{1,j}$, $U^n_{2,j}$ and the known boundary value $U^n_{0,j}$ to approximate the derivative $\frac{\partial U^n_{0,j}}{\partial x}$.

In our work, we adopt the second approach to approximate the derivative $\frac{\partial U^n_{0,j}}{\partial x}$. The specific reason will be stated in the next section. For readers' convenience, we summarize the finite difference equations written on a grid point (i, j, n) as:

$$\begin{cases} \dfrac{\partial U^n_{i,j}}{\partial \tau} = a_j\delta_{xx}U^n_{i,j} + b_j\delta_{vv}U^n_{i,j} + (d_j + \lambda_j)\delta_x U^n_{i,j} + c_j\delta_{xv}U^n_{i,j} + e_j\delta_v U^n_{i,j} - rU^n_{i,j}, \\ U^0_{i,j} = 0, \\ U^n_{0,j} = K - S^n_f(j), \\ \dfrac{4U^n_{1,j} - U^n_{2,j} - 3U^n_{0,j}}{2\Delta x} = -S^n_f(j), \\ U^n_{i,0} = 0, \\ U^n_{i,nv} = K, \end{cases} \tag{3.5}$$

where

$$\xi_j^n = \frac{1}{S_f^n(j)} \frac{S_f^n(j+1) - S_f^n(j-1)}{2\Delta v}, \qquad \beta_j^n = \frac{1}{S_f^n(j)} \frac{S_f^n(j+1) - 2S_f^n(j) + S_f^n(j-1)}{(\Delta v)^2},$$

$$\lambda_j = \frac{1}{S_f^n(j)} \frac{\partial S_f^n(j)}{\partial \tau}, \qquad a_j = \frac{1}{2} v_j + \frac{1}{2}\sigma^2 v_j (\xi_j^n)^2 - \rho\sigma v_j \xi_j, \qquad b_j = \frac{1}{2}\sigma^2 v_j,$$

$$c_j = \rho\sigma v_j - \sigma^2 v_j \xi_j^n, \qquad d_j = r - \frac{1}{2} v_j + \frac{1}{2}(\xi_j^n)^2 \sigma^2 v_j - \frac{1}{2}\sigma^2 v_j \beta_j^n + \kappa[\eta - v_j]\xi_j^n,$$

$$e_j = \kappa[\eta - v_j].$$

One should notice that in the PDE system (3.5), the time derivative remains un-discretized. The discretization of the time derivative is completed in the next subsection when we design a fast and efficient numerical scheme to solve this highly nonlinear system through a linearization.

3.3. Linearization of the nonlinear PDE system

It can be seen from the PDE system (3.5) that, if the optimal exercise price S_f^n were known at the beginning of the $(n+1)$th time step, this system would become a linear one, and the option price at the $(n+1)$th step could be worked out directly. Based on this point, we propose a new predictor–corrector approach to solve the highly nonlinear PDE system (3.5) in two phases within a time step.

The first phase is to work out an estimated value of the optimal exercise price at the $(n+1)$th time step, denoted as $\widetilde{S}_f^{n+1}$. For simplicity, we omit the subscript in the v direction. Recall that there are two ways to approximate the first-order derivative at $x = 0$, i.e.,

$$\frac{\partial U_0^{n+1}}{\partial x} = \frac{U_1^{n+1} - U_{-1}^{n+1}}{2\Delta x} + \mathcal{O}((\Delta x)^3) = -S_f^{n+1}, \tag{3.6}$$

$$\frac{\partial U_0^{n+1}}{\partial x} = \frac{4U_1^{n+1} - U_2^{n+1} - 3U_0^{n+1}}{2\Delta x} + \mathcal{O}((\Delta x)^3) = -S_f^{n+1}. \tag{3.7}$$

If Eq. (3.6) is adopted to approximate $\frac{\partial U_0^{n+1}}{\partial x}$, one needs to assume that the governing equation contained in (3.1) is also satisfied at $x = 0$, which will lead to a complicated nonlinear differential equation of $S_f(v, \tau)$ with respect to v, due to the complexity of the Heston operator. However, if the second approximation formula Eq. (3.7) is adopted, we only need to deal with a much simpler process that will lead to an explicit exposure of $S_f(v, \tau)$, while maintaining the same order of accuracy. From the Dirichlet boundary condition at $x = 0$, we know that

$$U_0^{n+1} = K - S_f^{n+1}. \tag{3.8}$$

Substituting Eq. (3.8) into (3.7), we can obtain a relation among U_1^{n+1}, U_2^{n+1} and S_f^{n+1} at the $(n+1)$th time step as:

$$S_f^{n+1} = \frac{3K + U_2^{n+1} - 4U_1^{n+1}}{3 + 2\Delta x}. \tag{3.9}$$

Then, by applying the explicit Euler scheme to the time derivative $\frac{\partial U_{i,j}^n}{\partial \tau}$ and the implicit Euler scheme to the time derivative $\frac{\partial S_f^n(j)}{\partial \tau}$, respectively, in the governing equation contained in (3.5), we obtain

$$U_{i,j}^{n+1} = U_{i,j}^n + \Delta\tau \left[a_j \delta_{xx} U_{i,j}^n + b_j \delta_{vv} U_{i,j}^n + \left(d_j + \frac{1}{S_f^n(j)} \frac{S_f^{n+1}(j) - S_f^n(j)}{\Delta\tau} \right) \delta_x U_{i,j}^n + c_j \delta_{xv} U_{i,j}^n + e_j \delta_v U_{i,j}^n \right] - r\Delta\tau U_{i,j}^n, \quad i = 1, 2. \tag{3.10}$$

Consequently, we can obtain a predicted value of the optimal exercise price at the new time step:

$$\tilde{S}_f^{n+1}(j) = \frac{3K + \mathcal{D}(U_{2,j}^n - 4U_{1,j}^n)}{3 + 2\Delta x - \frac{\delta_x \left(U_{2,j}^n - 4U_{1,j}^n \right)}{S_f^n(j)}}.$$

Here, the operator $\mathcal{D}$ is defined as:

$$\mathcal{D} = I + \Delta\tau \left[a_j \delta_{xx} + b_j \delta_{vv} + \left(d_j - \frac{1}{\Delta\tau} \right) \delta_x + c_j \delta_{xv} + e_j \delta_v \right] - r\Delta\tau I.$$

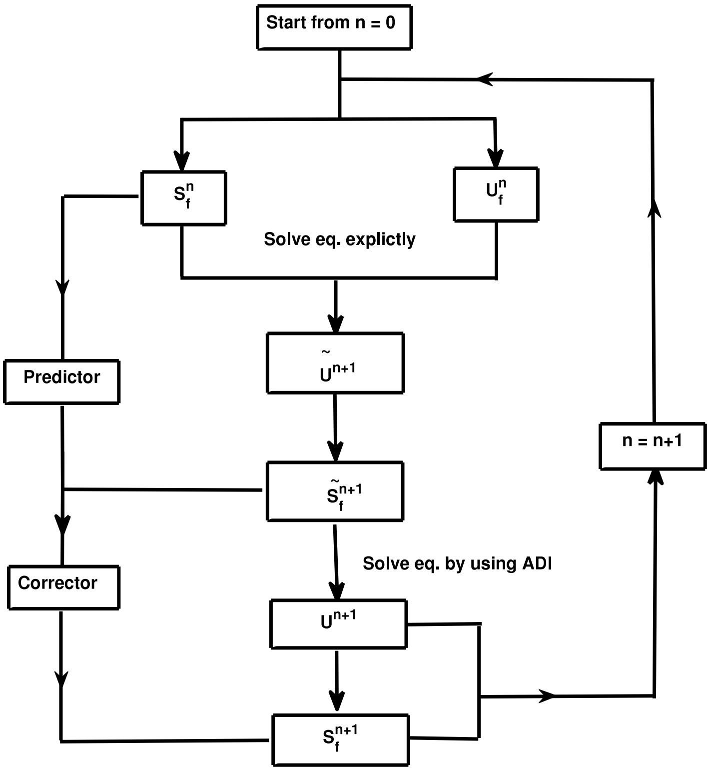


Fig. 1. Schematic flow chart of the scheme.

It should be noted that with the calculated $\tilde{S}_f^{n+1}$, we can also obtain a predicted boundary value $\tilde{U}_0^{n+1}$ from Eq. (3.8), which completes the first phase of prediction.

The second phase is to calculate U^{n+1} at all grid points by using the estimated $\tilde{S}_f^{n+1}$ and $\tilde{U}_0^{n+1}$. Then, by using the newly-obtained U_1^{n+1} and U_2^{n+1}, we can obtain the corrected values of S_f^{n+1} and U_0^{n+1}. Unlike the first phase in which the explicit Euler scheme is used to construct the predicted values, the whole process of calculating U^{n+1} from the linearized PDE system (3.5) is based on the ADI scheme, which will be described in detail in the next subsection. Then, repeat this prediction–correction process until the expiration date is reached. The schematic flow chart of the scheme is provided in Fig. 1 to summarize what has been described above.

3.4. The ADI scheme

The ADI scheme is a very powerful tool that is especially useful for solving parabolic equations on rectangular domains. It can be also applied to equations of other types or on more general domains [28]. Generally, the ADI scheme is a way of reducing a two-dimensional problem to a succession of many one-dimensional problems. The efficiency of the ADI method lies in the fact that those reduced one-dimensional problems usually possess a good structure, that is, their final matrix is tridiagonal and can thus be efficiently dealt with. On the other hand, the ADI method requires less storage space, because it solves the two directions alternatively by fixing the variable in one direction at a time step and thus it needs almost the same storage space as that required to solve a one-dimensional problem. Since our problem now is a two dimensional one, excluding time, it is better to use the ADI method. In the following, we will illustrate how this method can be applied to our case.

According to what we have discussed before, our problem at hand now is to solve the PDE system (3.5). The time derivative of U needs to be discretized before the ADI method can be used. While the details of the derivation are provided in Appendix A, the finite difference equation, to which the ADI method is applied, is of the form:

$$(I-\theta A_1)(I-\theta A_2)U^{n+1}=[I+A_0+(1-\theta)A_1+A_2]U^n-(I-\theta A_1)\theta A_2 U^n, \tag{3.11}$$

where the definitions of linear operators A_0, A_1 and A_2 are also left in Appendix B.

In terms of the ADI scheme, the simplest one is the Douglas-Rachford (DR) method [36], which is of the first order in time. Other methods, such as the Craig–Sneyed (CS) method [37], the Hundsdorfer and Verwier (HV) method [38], are much more complicated, but are of more than first-order accuracy in the time direction. We chose the DR method to calculate U^{n+1} defined in Eq. (3.11). One of the most important reasons is that in the predictor, we have used the explicit Euler scheme, which is a first-order scheme, to construct the predicted value of S_f^{n+1}. Theoretically speaking, an increase of the order of accuracy in the correction phase beyond that in the prediction phase is futile, as far as the overall order of accuracy of the entire procedure is concerned. Moreover, it is far more efficient to use the DR method, because there are only two steps in the DR method, while others have at least four.

The DR method involves two steps, in which the original operator in Eq. (3.11) is split into two that are applied in two spatial directions respectively. First, we compute an intermediate value, Y, from

$$(I-\theta A_1)Y = [I + A_0 + (1-\theta)A_1 + A_2]U^n, \tag{3.12}$$

with the nodal values in the v direction fixed. The corresponding matrix form for calculating Y can be simply written as

$$AY_j = P_j + \mathbf{bnd}\, x_j,$$

with the details of A, Y_j, P_j and $\mathbf{bnd}\, x_j$ being defined in Appendix C. Note that the matrix A is tridiagonal, so the Thomas algorithm [28] can be used to accelerate the computational speed. In short, the way to implement the first step is to use a loop on j, and within each loop the Thomas algorithm is used to solve for the values of $Y_{i,j}$, for $i = 0 \cdots N_x$.

Having computed Y, the second stage of computation is to compute U^{n+1} from

$$(I-\theta A_2)U^{n+1} = Y - \theta A_2 U^n, \tag{3.13}$$

by fixing the variable in the x direction. The corresponding matrix form is:

$$BU_i^{n+1} = Q_i + \mathbf{bnd}\, v_i,$$

where B, U_i^{n+1}, Q and $\mathbf{bnd}\, v_i$ are also defined in Appendix C. Similar to Eq. (3.12), this is a system of $N_x - 1$ tridiagonal systems of equations, one tridiagonal system for each value of i. Again, we use a loop on i, within which the Thomas algorithm is used to solve for the values of $U_{i,j}^{n+1}$, for $j = 0 \cdots N_v$.

One can easily show that solving Eqs. (3.12) and (3.13) in an alternative way is equivalent to solving the original Eq. (3.11).

Before presenting the results of the numerical implementation of this two-stage solution procedure in the next section, it should be remarked that boundary values of the intermediate variable Y need to be produced first before the interior Y values can be computed from Eq. (3.12). The calculation of Y values on the two boundaries needs to be treated differently. The calculation on the right boundary is straightforward, since the U values there are always equal to zero. That is, we can simply set

$$Y_{N_x} = 0. \tag{3.14}$$

On the other hand, the calculation of Y values on the left boundary $x = 0$ is a bit more complicated, as U is time dependent there. Eq. (3.13) is utilized because the estimated U^{n+1} values on the boundary have already been obtained in the prediction phase, although its interior values have not been calculated yet. That is, we can calculate Y_0 from

$$\begin{aligned} Y_0 &= (I-\theta A_2)\tilde{U}_0^{n+1} + \theta A_2 U_0^n, \\ &= (I-\theta A_2)(K - \tilde{S}_f^{n+1}) + \theta A_2 (K - S_f^n), \end{aligned}$$

where $\tilde{S}_f^{n+1}$ is the predicted optimal exercise price obtained at the beginning of the $(n+1)$th time step. This completes the ADI scheme and thus the entire numerical procedure, that we have used to calculate the results presented in the next section.

4. Examples and discussions

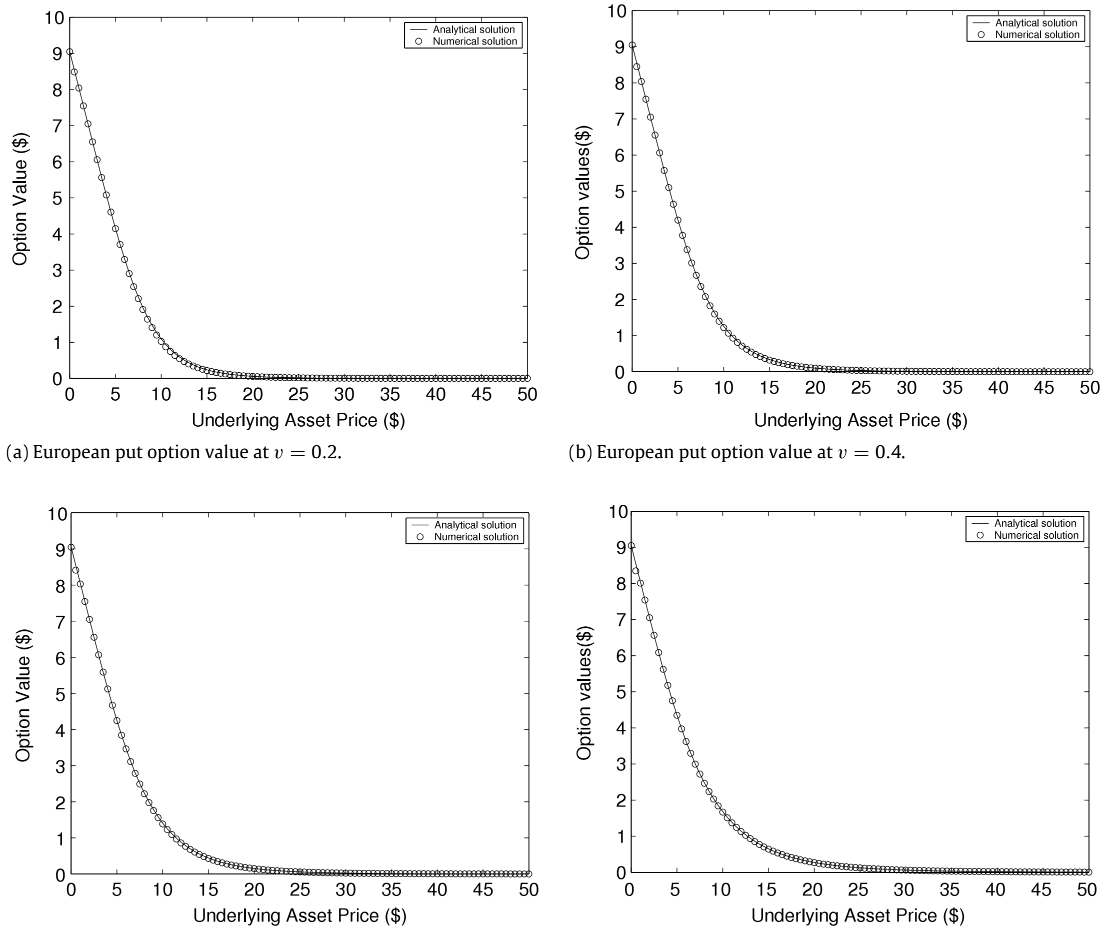
Even if the ADI method used for the corrector is unconditionally stable, our predictor–corrector finite difference scheme is only conditionally stable since the explicit Euler scheme for the predictor is conditionally stable. The stability condition for the current approach remained undiscussed in the literature even when it was applied to the Black–Scholes case [11]. The complicity of the Heston model has no doubt added the degree of difficulty in verifying the conditional stability of the current approach. However, based on the local von Neumann stability analysis [28], we were able to theoretically obtain the stability requirement for the current approach applied to the Heston model.

In this section, we shall present some numerical results as well as some useful discussions, which reveal the essence of the current scheme. The section is organized into three subsections, according to three important issues that should be addressed.

In the first subsection, we calculate the European puts to test the reliability of the chosen ADI method. In the second subsection, we discuss the convergence of the current approach. The stability condition is obtained by using von Neumann stability analysis. In the last subsection, we study the accuracy and efficiency of the current scheme.

4.1. Valuing European put options using the ADI scheme

Since the corrector plays a crucial role in the current approach, it is important to test first whether the chosen DR method can be used to construct a good corrector. An efficient way to illustrate the reliability of the DR method is to calculate the value of a European put, and compare it with the existing Heston formula [4].



(a) European put option value at $v = 0.2$.

(b) European put option value at $v = 0.4$.

(c) European put option value at $v = 0.6$.

(d) European put option value at $v = 1$.

Fig. 2. European put option prices with different variance values.

The option's parameters are set as follows: reversion rate $\kappa = 5$, reversion level $\eta = 0.16$, volatility factor of volatility $\sigma = 0.9$, risk-free interest rate $r = 0.1$, correlation factor $\rho = 0.1$, time to expiration $T - t = 1$ (year), strike price $K = \$10.0$. The computation domain is truncated as:

$$[0, S_{\max}] \times [0, v_{\max}] \times [0, T] = [0, 200] \times [0, 5] \times [0, 1].$$

In this numerical test, we apply the Crank–Nicolson scheme, i.e., $\theta = \frac{1}{2}$, to the time derivative. The grid numbers in the x direction and v direction are chosen to be 400 and 150, respectively. Furthermore, a relatively large grid size in the time direction is used, i.e., $\Delta\tau = \frac{1}{20}$. Fig. 2 shows the comparison of the option values calculated by the DR method and those obtained by using Heston's analytical formula. The option prices presented in Fig. 2 are plotted against the underlying asset with different fixed variance values. The excellent agreement of the two results confirms that the DR method is accurate.

Having gained confidence on the chosen DR method as a good corrector, we then apply the predictor–corrector scheme to American put options under the Heston model, which will be demonstrated in the following subsection.

4.2. Discussion on convergence

In this subsection, we study the convergence of the proposed numerical scheme. As has been pointed out previously, our problem in hand is originally a nonlinear one before the linearization process is invoked. Therefore, the convergence of the linearized system and the overall convergence of the numerical solution to that of the original nonlinear PDE should be both discussed.

4.2.1. Stability analysis

For the convergence of the linearized system, the Lax Equivalence Theorem [28] can be applied, which states that the convergence of a numerical scheme is equivalent to its consistency and stability. The proof of the consistency is trivial and is thus omitted here. The proof of stability, however, is not so trivial because our method is a hybrid finite difference scheme, and it is applied to a problem with variable coefficients. Intuitively, even if the ADI scheme for the corrector is unconditionally stable (see Theorem 2), the current approach should be only conditionally stable since the explicit Euler scheme is used in the predictor. However, it can be anticipated that the unconditional stability in the correction phase would be weakened to some extent. This is because the main ingredient of the predictor–corrector method is the corrector, which serves as a feedback mechanism. The feedback can damp the instability that has been introduced by the predictor, and thus relax the stability requirement of the explicit Euler scheme used in the predictor. It is extremely difficult to calculate exactly "how much" the corrector can influence the predictor, and in the literature, no one has discussed this issue in detail. Based on the local von Neumann stability analysis, we have managed to show the conditional stability for the predictor–corrector scheme. In the following work, the stability requirements of the explicit Euler scheme Eq. (3.10) and the ADI scheme Eqs. (3.12)–(3.13) will be first discussed. Finally, we obtain the conditional stability of the predictor–corrector approach.

The von Neumann stability analysis is usually restricted to problems with constant coefficients. However, the stability conditions obtained for constant coefficient schemes can be used to give stability conditions for the same scheme applied to equations with variable coefficients, mainly due to the fact that, in essence, instability is a local phenomenon with the high frequency modes being the most unstable ones that would result in the collapse of a numerical scheme [28]. The general procedure is to consider the so-called frozen coefficient problem, which is the corresponding problem with constant coefficients obtained by fixing the coefficients at their values attained at each grid point in the computational domain. If each frozen coefficient problem is stable, then the variable coefficient problem is also stable [28]. For simplicity, in the following analysis, we do not consider the overall effect of the boundary conditions between subdomains.

When using the frozen coefficient technique, for a typical node (i, j, n), we "freeze" all the coefficients in the governing equation contained in Eq. (3.5) across the whole computational domain, as if they were all constants. In short, we first consider the stability condition of the DR method applied to a two-dimensional convection–diffusion equation with constant coefficients, i.e.,

$$\frac{\partial U}{\partial \tau} = a\frac{\partial^2 U}{\partial x^2} + b\frac{\partial^2 U}{\partial v^2} + c\frac{\partial^2 U}{\partial x \partial v} + d\frac{\partial U}{\partial x} + e\frac{\partial U}{\partial v} - rU, \tag{4.1}$$

where a, b and r are of positively values. The standard procedure of von Neumann stability analysis is to express $U^n_{k,m}$ in Eqs. (3.12)–(3.13) by $g^n e^{ik\varphi} e^{im\phi}$, and $Y_{k,m}$ by $\tilde{g} g^n e^{ik\varphi} e^{im\phi}$ [28], where g and $\tilde{g}$ are the amplification factors of Eqs. (3.13) and (3.12), respectively, with $\varphi, \phi \in [-\pi, \pi]$. As a result, Eqs. (3.12)–(3.13) are transformed to:

$$\tilde{g}(1-\theta z_1) = 1 + z_0 + (1-\theta) z_1 + z_2;$$
$$g(1-\theta z_2) = \tilde{g} - \theta z_2,$$

which result in the amplification factor:

$$g = 1 + \frac{z_0 + z_1 + z_2}{(1-\theta z_1)(1-\theta z_2)},$$

where

$$z_1 = -\frac{4a\Delta\tau}{\Delta x^2}\sin^2\frac{\varphi}{2} - \frac{r\Delta\tau}{2} + i\frac{d\Delta\tau}{\Delta x}\sin\varphi,$$
$$z_2 = -\frac{4b\Delta\tau}{\Delta v^2}\sin^2\frac{\phi}{2} - \frac{r\Delta\tau}{2} + i\frac{e\Delta\tau}{\Delta v}\sin\phi,$$
$$z_0 = -\frac{c\Delta\tau}{\Delta x \Delta v}\sin\varphi\sin\phi,$$

after some algebraic manipulations.

Proposition 2. (i) *If the coefficients a, b, c, d, e, c, r are chosen such that*

$$|z_0| \le 2\sqrt{\Re(z_1)\Re(z_2)}, \tag{4.2}$$

then for $\theta \ge \frac{1}{2}$, the DR method (Eqs. (3.12)–(3.13)) is unconditionally stable, i.e., $|g| \le 1$.[1]

(ii) *Assuming that the coefficients satisfy $c^2 \le 4ab$, the fully explicit scheme (Eqs. (3.12)–(3.13) with $\theta = 0$), is stable if and only if*

$$\Delta\tau \le \frac{1}{\frac{2a}{\Delta x^2} + \frac{2b}{\Delta v^2}}. \tag{4.3}$$

1 Refer to [39].

Proof. (i) First, we define two vectors as $v_j = \begin{pmatrix} \sqrt{-2\Re(z_j)} \\ \frac{|1+\theta z_j|}{\sqrt{2\theta}} \end{pmatrix}$, where $j = 1, 2$. It is trivial to show $\|v_j\|_2 = \frac{|1-\theta z_j|}{\sqrt{2\theta}}$. Thus, we have

$$\begin{aligned} \frac{|1-\theta z_1|}{\sqrt{2\theta}} \frac{|1-\theta z_2|}{\sqrt{2\theta}} &= \|v_1\|_2 \, \|v_2\|_2, \\ &\geq v_1 \cdot v_2, \\ &= 2\sqrt{\Re(z_1)\Re(z_2)} + \frac{|1+\theta z_1|}{\sqrt{2\theta}} \frac{|1+\theta z_2|}{\sqrt{2\theta}}, \\ &\geq |z_0| + \left| \frac{(1-\theta z_1)}{\sqrt{2\theta}} \frac{(1-\theta z_2)}{\sqrt{2\theta}} + z_1 + z_2 \right|. \end{aligned} \tag{4.4}$$

Without loss of generality, we assume that $|1-\theta z_1||1-\theta z_2| \neq 0$. Dividing both sides of Eq. (4.4) with $|1-\theta z_1|\,|1-\theta z_2|$, we have

$$\begin{aligned} \frac{1}{2\theta} &\geq \left| \frac{z_0}{(1-\theta z_1)(1-\theta z_2)} \right| + \left| \frac{1}{2\theta} + \frac{z_1+z_2}{(1-\theta z_1)(1-\theta z_2)} \right|, \\ &\geq \left| \frac{1}{2\theta} + \frac{z_0+z_1+z_2}{(1-\theta z_1)(1-\theta z_2)} \right|. \end{aligned} \tag{4.5}$$

On the other hand, when $\theta \geq \frac{1}{2}$, we have

$$\begin{aligned} g &= \left| 1 + \frac{z_0+z_1+z_2}{(1-\theta z_1)(1-\theta z_2)} \right|, \\ &= \left| 1 - \frac{1}{2\theta} + \frac{1}{2\theta} + \frac{z_0+z_1+z_2}{(1-\theta z_1)(1-\theta z_2)} \right|, \\ &\leq 1 - \frac{1}{2\theta} + \left| \frac{1}{2\theta} + \frac{z_0+z_1+z_2}{(1-\theta z_1)(1-\theta z_2)} \right|, \end{aligned}$$

which, combined with Eq. (4.5), yields $|g| \leq 1$. Therefore, if

$$|z_0| \leq 2\sqrt{\Re(z_1)\Re(z_2)},$$

the ADI scheme (for $\theta \geq \frac{1}{2}$) is unconditionally stable.

(ii) When $\theta = 0$, the DR method turns out to be a fully explicit scheme. In this case, it is efficient to first assume that $\frac{\Delta\tau}{(\Delta S)^2} = \mu_1$, $\frac{\Delta\tau}{(\Delta v)^2} = \mu_2$, where μ_1 and μ_2 are fixed constants. Thus, we have

$$\begin{aligned} \lambda_1 &= \frac{\Delta\tau}{\Delta x} = \sqrt{\Delta t \mu_1} \sim \mathcal{O}(\sqrt{\Delta\tau}), \\ \lambda_2 &= \frac{\Delta\tau}{\Delta v} = \sqrt{\Delta t \mu_2} \sim \mathcal{O}(\sqrt{\Delta\tau}), \end{aligned}$$

which indicate that, in comparison with those $\mathcal{O}(1)$ terms, the lower-order derivative terms only give an $\mathcal{O}(\Delta\tau)$ contribution to $|g|^2$, and thus do not affect the stability. In fact, ignoring lower-order derivative terms when determining stability is a commonly used technique in the stability analysis [28].

Now, we need to prove that, with $\Delta\tau \leq \frac{1}{\frac{2a}{\Delta x^2} + \frac{2b}{\Delta v^2}}$,

$$\left(1 - \frac{c\Delta\tau}{\Delta x \Delta v} \sin\varphi \sin\phi - \frac{4a\Delta\tau}{\Delta x^2} \sin^2\frac{\varphi}{2} - \frac{4b\Delta\tau}{\Delta v^2} \sin^2\frac{\phi}{2} \right)^2 \leq 1, \quad (\varphi, \phi) \in [-\pi, \pi] \times [-\pi, \pi],$$

which is equivalent to showing that

$$0 \leq \frac{c\Delta\tau}{\Delta x \Delta v} \sin\varphi \sin\phi + \frac{4a\Delta\tau}{\Delta x^2} \sin^2\frac{\varphi}{2} + \frac{4b\Delta\tau}{\Delta v^2} \sin^2\frac{\phi}{2} \leq 2, \quad (\varphi, \phi) \in [-\pi, \pi] \times [-\pi, \pi], \tag{4.6}$$

and vice versa. This is achieved by using the following procedure.

Define a function

$$P(\varphi, \phi) = \tilde{c} \sin\varphi \sin\phi + 2\tilde{a} \sin^2\frac{\varphi}{2} + 2\tilde{b} \sin^2\frac{\phi}{2},$$

where

$$\tilde{c} = c\sqrt{\mu_1\mu_2}, \qquad \tilde{a} = 2a\mu_1, \qquad \tilde{b} = 2b\mu_2.$$

Since $P(\varphi, \phi)$ is continuous over the domain $[-\pi, \pi] \times [-\pi, \pi]$, the maximum and minimum values can be both achieved in this domain. As a result, Eq. (4.6) is satisfied if and only if

$$0 \le P_{\min} \le P_{\max} \le 2. \tag{4.7}$$

Suppose (φ, ϕ) is an interior point of the domain of the function P, and it is also a local maximum or minimum point, we have $P_\varphi(\varphi, \phi) = P_\phi(\varphi, \phi) = 0$, i.e.,

$$\begin{cases} \tilde{c}\cos\varphi \sin\phi + \tilde{a}\sin\varphi = 0, \\ \tilde{c}\cos\phi \sin\varphi + \tilde{b}\sin\phi = 0. \end{cases} \tag{4.8}$$

It is straightforward to show that Eq. (4.8) is equivalent to

$$\begin{cases} ((\tilde{c}^2 + \tilde{a})^2)\sin^2\varphi = \dfrac{\tilde{c}^4 - \tilde{a}^2\tilde{b}^2}{\tilde{c}^2}, \\ ((\tilde{c}^2 + \tilde{b})^2)\sin^2\phi = \dfrac{\tilde{c}^4 - \tilde{a}^2\tilde{b}^2}{\tilde{c}^2}. \end{cases} \tag{4.9}$$

Clearly, the origin $(\varphi, \phi) = (0, 0)$ is the only solution of Eq. (4.9) because of the constraint $c^2 \le 4ab$, which is equivalent to $\tilde{c}^2 \le \tilde{a}\tilde{b}$. Thus, in $(-\pi, \pi) \times (-\pi, \pi)$, the only possible local maximum or minimum point of P is $(\varphi, \phi) = (0, 0)$. Comparing the function values at (0, 0), and along the boundary $\varphi = \pm\pi$, or $\phi = \pm\pi$, it is clear that

$$P_{\max} = 4a\mu_1 + 4b\mu_2,$$
$$P_{\min} = 0.$$

Consequently, if $\Delta\tau \le \frac{1}{\frac{2a}{\Delta x^2} + \frac{2b}{\Delta v^2}}$, then Eq. (4.7) is satisfied, and vice versa. Therefore, under the assumption that the coefficients of Eq. (4.1) satisfy $c^2 \le 4ab$, the sufficient and necessary condition for the stability of the explicit scheme (Eqs. (3.12)–(3.13) with $\theta = 0$) is $\Delta\tau \le \frac{1}{\frac{2a}{\Delta x^2} + \frac{2b}{\Delta v^2}}$. This completes the proof. □

By means of Proposition 2, the following theorems can be proved.

Theorem 1. *The DR method that is applied to the valuation of European puts is unconditionally stable* $\left(\theta \ge \frac{1}{2}\right)$.

Proof. In the case of European puts, z_0, $\Re(z_1)$, and $\Re(z_2)$ are defined as follows:

$$\Re(z_1) = -\frac{2vx^2\Delta\tau}{\Delta x^2}\sin^2\frac{1}{2}\varphi - \frac{r\Delta\tau}{2},$$
$$\Re(z_2) = -\frac{2\sigma^2 v\Delta\tau}{\Delta v^2}\sin^2\frac{1}{2}\phi - \frac{r\Delta\tau}{2},$$
$$|z_0| = \left|\frac{\rho\sigma vx\Delta\tau}{\Delta x\Delta v}\sin\varphi\sin\phi\right|,$$

where x and v are defined over the whole computational domain. We have,

$$\begin{aligned} 4\Re(z_1)\Re(z_2) - |z_0|^2 &= 4\left(\frac{2vx^2\Delta\tau}{\Delta x^2}\sin^2\frac{1}{2}\varphi + \frac{r\Delta\tau}{2}\right)\left(\frac{2\sigma^2 v\Delta\tau}{\Delta v^2}\sin^2\frac{1}{2}\phi + \frac{r\Delta\tau}{2}\right) - \left(\frac{\rho\sigma vx\Delta\tau}{\Delta x\Delta v}\sin\varphi\sin\phi\right)^2, \\ &\ge \left(\frac{4v\sigma x\Delta\tau}{\Delta x\Delta v}\sin\frac{1}{2}\varphi\sin\frac{1}{2}\phi\right)^2\left(1 - \rho^2\cos^2\frac{1}{2}\varphi\cos^2\frac{1}{2}\phi\right), \\ &\ge 0, \end{aligned}$$

which means that for any fixed point (x, v) across the whole computational domain, Eq. (4.2) is always satisfied. According to the first part of the proposition and the frozen coefficient technique, the ADI scheme is unconditionally stable. This completes the proof. □

Theorem 2. *Assuming that the optimal exercise price is known in advance, and all the variable coefficients are both bounded and sufficiently smooth, the DR method that is applied to the corrector (Eqs. (3.12)–(3.13)) is unconditionally stable* $\left(\theta \ge \frac{1}{2}\right)$.

Proof. In this case, z_0, $\Re(z_1)$, and $\Re(z_2)$ are defined as follows:

$$\Re(z_1) = -\frac{(2v + 2\sigma^2 v\xi^2 - 4\rho\sigma v\xi)\Delta\tau}{\Delta x^2}\sin^2\frac{1}{2}\varphi - \frac{r\Delta\tau}{2},$$
$$\Re(z_2) = -\frac{2\sigma^2 v\Delta\tau}{\Delta v^2}\sin^2\frac{1}{2}\phi - \frac{r\Delta\tau}{2},$$

$$|z_0| = \left| \frac{(\rho\sigma v - \sigma^2 v\xi)\Delta\tau}{\Delta x \Delta v} \sin\varphi \sin\phi \right|.$$

We have,

$$\begin{aligned}
4\Re(z_1)\Re(z_2) - |z_0|^2 &= 4\left(\frac{(2v + 2\sigma^2 v\xi^2 - 4\rho\sigma v\xi)\Delta\tau}{\Delta x^2} \sin^2\frac{1}{2}\varphi + \frac{r\Delta\tau}{2}\right)\left(\frac{2\sigma^2 v\Delta\tau}{\Delta v^2}\sin^2\frac{1}{2}\phi + \frac{r\Delta\tau}{2}\right) \\
&\quad - \left(\frac{(\rho\sigma v - \sigma^2 v\xi)\Delta\tau}{\Delta x\Delta v}\sin\varphi\sin\phi\right)^2, \\
&\geq \left(\frac{4v\sigma\Delta\tau}{\Delta x\Delta v}\sin\frac{1}{2}\varphi\sin\frac{1}{2}\phi\right)^2\left[(1+\sigma^2\xi^2 - 2\rho\sigma\xi) - (\rho - \sigma\xi)^2\cos^2\frac{1}{2}\varphi\cos^2\frac{1}{2}\phi\right], \\
&\geq \left(\frac{4v\sigma\Delta\tau}{\Delta x\Delta v}\sin\frac{1}{2}\varphi\sin\frac{1}{2}\phi\right)^2 [(1+\sigma^2\xi^2 - 2\rho\sigma\xi) - (\rho-\sigma\xi)^2], \\
&= \left(\frac{4v\sigma\Delta\tau}{\Delta x\Delta v}\sin\frac{1}{2}\varphi\sin\frac{1}{2}\phi\right)^2 (1-\rho^2), \\
&\geq 0,
\end{aligned}$$

which means that for any fixed point (x, v) across the whole computational domain, Eq. (4.2) is always satisfied. According to the first part of the proposition and the frozen coefficient technique, the ADI scheme that is applied to the corrector is unconditionally stable. This completes the proof. □

Theorem 3. *Assuming that the optimal exercise price is known in advance, and all the variable coefficients are both bounded and sufficiently smooth, the fully explicit scheme that is applied to the predictor Eq.* (3.10) *($\theta = 0$) is stable if and only if*

$$\Delta\tau \leq \frac{2}{a_1 + a_2},$$

where

$$a_1 = \frac{2v_{\max} - 4\rho\sigma v_{\max}\xi_{\min} + 2\sigma^2 v_{\max}\xi_{\min}^2}{(\Delta x)^2},$$

$$a_2 = \frac{2\sigma^2 v_{\max}}{(\Delta v)^2}.$$

(Here, we further assume that the correlation factor $\rho \geq 0$.)

Proof. According to the second part of Proposition 2, and by using the frozen coefficient technique, we only need to show that, with

$$\Delta\tau \leq \frac{2}{a_1 + a_2},$$

the constraints $c \leq 4ab$, and $\Delta\tau \leq \frac{1}{\frac{2a}{\Delta x^2} + \frac{2b}{\Delta v^2}}$ are both satisfied at each point across the whole computational domain. In this case, a, b, c are defined as:

$$a = \left(\frac{1}{2}v + \frac{1}{2}\sigma^2 v\xi^2 - \rho\sigma v\xi\right), \qquad b = \frac{1}{2}\sigma^2 v,$$

$$c = \rho\sigma v - \sigma^2 v\xi.$$

Since

$$\begin{aligned}
4ab - c^2 &= (v + \sigma^2 v\xi^2 - 2\rho\sigma v\xi)\sigma^2 v - (\rho\sigma v - \sigma^2 v\xi)^2, \\
&= (v\sigma)^2(1-\rho^2), \\
&\geq 0,
\end{aligned}$$

all the frozen coefficient problems are stable if and only if

$$\Delta\tau \leq \frac{1}{\frac{2a_{\max}}{\Delta x^2} + \frac{2b_{\max}}{\Delta v^2}}.$$

Note that the optimal exercise price $S_f(v, \tau)$ is a monotonic decreasing function of v. On the other hand, it is assumed that all the coefficients are both bounded. Therefore, there exists a $\xi_{\min}$, such that $\xi_{\min} \leq \xi_m \leq 0$. As a result, $a_{\max} = a_1$, $b_{\max} = a_2$, this completes the proof. □

Based on Theorem 3, one can easily find out that the stability condition of the fully explicit scheme is closely related to the value of σ. More steps in the time direction are needed when σ is large.

Through previous analyses, we know that with the optimal exercise price known in advance, the corrector is unconditionally stable. Thus, if the predictor constructed by the explicit Euler scheme is stable, the whole process is stable. In other words, the amplification factor of the predictor is also the one for the entire predictor–corrector scheme. The main difference between our predictor and the single explicit Euler scheme lies in the fact that we use the corrected value, which is produced by the ADI method, to continue the prediction of the next time step. As a result, the good stable property of the ADI scheme has somehow relaxed the overall amplification factor. It would be ideal to show the above statement numerically. This is achieved after we prove Theorem 4 as follows.

Theorem 4. *If g_1 is the amplification factor of the explicit Euler scheme used in the predictor, and g_2 is the one of the ADI scheme used in the corrector, then the predictor–corrector method is stable if and only if $|g_1 g_2| \le 1 + M\Delta\tau$.*

Proof. We use tildes to denote the value that is obtained after the predictor. In the following proof, we have omitted the subscript in the v direction or x direction for simplicity.

Since

$$\tilde{S}_f^{n+1} = \frac{3K + \tilde{U}_2^{n+1} - 4\tilde{U}_1^{n+1}}{3 + 2\Delta x},$$

we can easily deduce:

$$|\tilde{S}_f^{n+1} - \tilde{S}_f^0| \le \frac{|\tilde{U}_2^{n+1} - \tilde{U}_2^0| + 4|\tilde{U}_1^{n+1} - \tilde{U}_1^0|}{3 + 2\Delta x}.$$

Thus, the predictor is stable if and only if the process of computing $\tilde{U}^{n+1}$ is stable.

From Eq. (3.10), we know that

$$\tilde{U}^{n+1} = (I + A_0^* + A_1^* + A_2^*)U^n \tag{4.10}$$

where a star denotes the parameter $\theta = 0$. Here, U^n is solved from

$$(I - \theta A_1)Y = [I + A_0 + (1-\theta)A_1 + A_2]\tilde{U}^{n-1}, \tag{4.11}$$

$$(I - \theta A_2)U^n = Y - \theta A_2 \tilde{U}^{n-1}, \tag{4.12}$$

where $\tilde{U}^{n-1}$ is obtained after the predictor of the nth time step. Again, we take all the variable coefficients in these two stages to be "frozen" at constant values, so that the von Neumann analysis can be applied. By using Fourier transform on Eqs. (4.10) and (4.11)–(4.12), we obtain:

$$\begin{aligned}
\tilde{\tilde{U}}_{k,m}^{n+1} &= (1 + z_0^* + z_1^* + z_2^*)\hat{U}_{k,m}^n, \\
&= g_1 \hat{U}_{k,m}^n, \\
\hat{U}_{k,m}^n &= 1 + \frac{z_0 + z_1 + z_2}{(1-\theta z_1)(1-\theta z_2)} \tilde{\tilde{U}}_{k,m}^{n-1}, \\
&= g_2 \tilde{\tilde{U}}_{k,m}^{n-1},
\end{aligned}$$

where $\tilde{\tilde{U}}_{k,m}^{n+1}$ and $\hat{U}_{k,m}^n$ are the Fourier transform of $\tilde{U}_{k,m}^{n+1}$ and $U_{k,m}^n$, respectively. Thus,

$$\tilde{\tilde{U}}_{k,m}^{n+1} = (g_1 g_2)\tilde{\tilde{U}}_{k,m}^{n-1},$$

which shows that the overall amplification factor g for the predictor–corrector approach is nothing but $g = g_1 g_2$.

Therefore, our method is stable if and only if

$$|g_1(x_i, v_j, \tau_n; \Delta\tau) g_2(x_i, v_j, \tau_{n-1}; \Delta\tau)| \le 1 + M\Delta\tau, \tag{4.13}$$

where M is a constant that is independent of φ, ϕ, as well as all the step sizes, and i, j, n are defined over the whole computational domain. This completes the proof. □

A further analysis of inequality (4.13) shows:

$$|g_1(x_i, v_j, \tau_n; \Delta\tau)| \le \frac{1 + M\Delta\tau}{|g_2(x_i, v_j, \tau_{n-1}; \Delta\tau)|}. \tag{4.14}$$

Since the corrector is unconditionally stable, i.e.,

$$|g_2| \le 1,$$

Table 1
Comparison of the stability requirement. Model parameters are $\kappa = 2.5$, $\eta = 0.16$, $r = 0.1$, $\rho = 0.1$, $T = 0.25$ (year), $K = \$10.0$.

σ	Predictor–corrector	Explicit
0.01	(25, 32, 13)	(25, 32, 250)
	(50, 64, 130)	(50, 64, 1000)
0.1	(25, 32, 39)	(25, 32, 250)
	(50, 64, 140)	(50, 64, 1000)
0.45	(25, 32, 100)	(25, 32, 400)
	(50, 64, 1100)	(50, 64, 1600)

Table 2
Comparison of the computed option prices with the reference solutions. Model parameters are $\kappa = 2.5$, $\eta = 0.16$, $\sigma = 0.45$, $r = 0.1$, $\rho = 0.1$, $T = 0.25$ (year), $K = \$10.0$.

Volatility value	(N_x, N_v, N_τ)	Asset values				
		8	9	10	11	12
$v = 0.0625$	(25, 32, 200)	2.0000	1.0682	0.4920	0.1950	0.0760
	(50, 64, 1600)	2.0000	1.0794	0.4828	0.1851	0.0634
	(100, 100, 6000)	2.0000	1.0774	0.4789	0.1796	0.0622
	Approximation 1 [40]		1.072	0.475	0.174	
	Approximation 2 [40]		1.077	0.478	0.178	
	Ref. [32]		1.077	0.479	0.178	
	Ref. [14]		1.075	0.478	0.177	
$v = 0.25$	(25, 32, 200)	2.9877	1.3669	0.8521	0.4787	0.2913
	(50, 64, 1600)	2.0904	1.3645	0.8410	0.4921	0.2756
	(100, 100, 6000)	2.0903	1.3644	0.8382	0.4884	0.2685
	Approximation 1 [40]		1.365	0.838	0.488	
	Approximation 2 [40]		1.362	0.836	0.487	
	Ref. [32]		1.364	0.837	0.487	
	Ref. [14]		1.363	0.837	0.488	

it is straightforward that

$$\Delta\tau_{\text{explicit}} \le \Delta\tau_{\text{predictor–corrector}}.$$

Therefore, comparing with the explicit Euler scheme, the stability requirement of our current scheme is less restrictive. This allows us to choose a larger time step than the fully explicit Euler scheme, and thus considerably enhance the computational efficiency.

This theorem leads to the conclusion that $\Delta\tau_{\text{explicit}} \le \Delta\tau_{\text{predictor–corrector}}$, because the corrector is unconditionally stable, or $g_1 < 1$.

Ideally, the above theoretical statement has been verified by the following numerical experiment (see Table 1). It shows the comparison of the smallest number of time steps that each method requires to calculate a convergent solution. One should notice that in Table 1, the "Predictor–corrector" refers to the current method with parameter $\theta = \frac{1}{2}$, while the "Explicit" refers to the fully explicit scheme, which is a special case of our scheme. It should be also noted that the third number in the triplets (N_x, N_v, N_τ) is the smallest number that each method needs to produce a convergent solution. It can be clearly seen that our current method requires a lower number of time steps than the fully explicit one. Moreover, our method is far more stable when σ is small. It confirms that the stability requirement of the predictor–corrector method has been influenced by the corrector. In other words, a better corrector can improve the computational efficiency of the whole scheme significantly.

Having demonstrated the convergence of the linearized system, we need to show that the numerical solution does converge to that of the original nonlinear PDE. Since there is no analytical solution for American puts under the Heston model, the only reasonable approach is to compare our numerical solution with other published results. This would be demonstrated by the following numerical examples.

4.2.2. Computed option prices and optimal exercise prices

We calculate two sets of American put options with different parameters using the current method based on the Crank–Nicolson scheme, i.e., $\theta = \frac{1}{2}$. These prices are presented in Tables 2 and 3 for the asset values $S = 8, 9, 10, 11, 12$, and for variance values $v = 0.0625, 0.25$. We have used different discretization grids in order to study the accuracy of the numerical solutions. The prices reported in [14,24,26,32,40,41] are also shown in these tables for comparison. It can be seen

Table 3
Comparison of the computed option prices with the reference solutions. Model parameters are $\kappa = 5, \eta = 0.16, \sigma = 0.9, r = 0.1, \rho = 0.1, T = 0.25$ (year), $K = \$10.0$.

Volatility value	(N_x, N_v, N_τ)	Asset price				
		8	9	10	11	12
$v = 0.0625$	(25, 32, 500)	2.0000	1.0752	0.5100	0.2200	0.0943
	(50, 64, 5000)	2.0000	1.0908	0.5073	0.2133	0.0837
	(100, 100, 50 000)	2.0000	1.0987	0.5082	0.2106	0.0861
	PSOR	2.0000	1.1075	0.5190	0.2129	0.0818
	Ref. [41]	2.000	1.107	0.517	0.212	0.0815
	Ref. [24]	2.000	1.1080	0.5316	0.2261	0.0907
	Ref. [26]	2.000	1.1076	0.5202	0.2138	0.0821
$v = 0.25$	(25, 32, 500)	2.0701	1.3366	0.8131	0.4654	0.2645
	(50, 64, 5000)	2.0787	1.3335	0.7999	0.4540	0.2474
	(100, 100, 50 000)	2.0781	1.3337	0.7965	0.4496	0.2441
	PSOR	2.0785	1.3336	0.7956	0.4481	0.2427
	Ref. [41]	2.079	1.334	0.796	0.449	0.243
	Ref. [24]	2.0733	1.3290	0.7992	0.4536	0.2502
	Ref. [26]	2.0784	1.3337	0.7961	0.4883	0.2428

that even with the most coarse grid, the error is only about 10^{-2}. Furthermore, the prices obtained with the finest grid are fairly close to the ones in [14,24,26,32,40,41], and the error is about 10^{-4}. This confirms that our numerical solution does converge to that of the original nonlinear PDE.

Depicted in Fig. 3(a) and (b) are the option price $U(S, v, \tau)$ as a function of S with different parameters. Clearly, the option price is a decreasing function of the asset values. Moreover, the "smooth pasty" conditions across the free boundary, which are usually difficult to implement numerically, are also satisfied well. In Fig. 4, we show the option price $U(S, v, \tau)$ as a function of S with fixed variance $v = 0.25$ at three instants: $\tau = T - t = 0.5$ (year), $\tau = T - t = 0.25$ (year) and $\tau = T - t = 0.1$ (year). Clearly, as it gets closer to the expiration of the option, i.e., $\tau = 0$, the option price becomes closer to the payoff function $\max(K - S, 0)$. Moreover, the optimal exercise price $S_f(v, \tau)$ as a function of time to expiration with different fixed variance values is shown in Fig. 3(c) and (d). As expected, the optimal exercise price is a monotonically decreasing function with both τ and v.

4.2.3. Convergence rate

Another important issue for a numerical scheme is its convergence rate. Theoretically speaking, we should have a first-order convergence in the τ direction and a second-order convergence in both x and v direction. This is because the explicit Euler scheme is used to construct the predictor, and the central finite difference is adopted for the space variables.

It is demonstrated that the optimal exercise price is far more difficult to calculate accurately than the option price [16]. Furthermore, once S_f is determined accurately, the calculation of the option price becomes straightforward. Consequently, it suffices to focus on the calculation of S_f to study the order of convergence for the current scheme.

Theoretically speaking, to obtain the order of convergence in one direction, we should examine the ratios of the consecutive errors of S_f with the grid spacing along this direction being successively decreased, while the grid spacings along other directions being fixed to be sufficiently small. However, as demonstrated earlier, there is no analytical solution available for the American puts with stochastic volatility. One of the standard ways to demonstrate the rate of convergence is to calculate a reference solution based on very fine grids, and use it as if it were the exact solution. We have conducted such an experiment, with the reference solution being constructed with the number of grids defined as $(N_x, N_v, N_\tau) = (400, 400, 100\,000)$, under the parameters settings: $\kappa = 1.5, \eta = 0.16, \sigma = 0.1, r = 0.1, \rho = 0.1, T = 0.1$ (year), $K = \$10.0$.

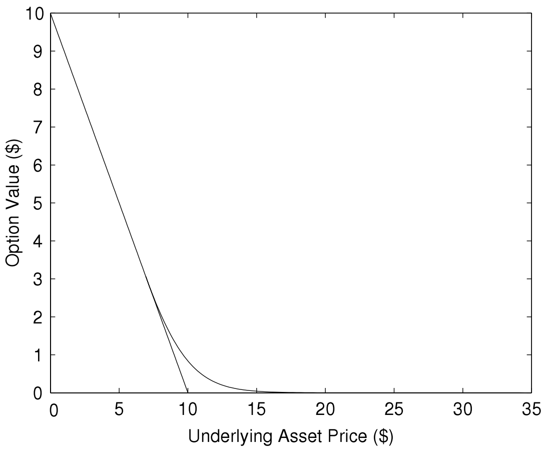
To obtain the order of convergence along the τ direction, we fix the spacial grid sizes to be $\Delta x = \frac{X_{\max}}{250}, v = \frac{V_{\max}}{250}$, and vary the number of time steps from 4000 to 5000. The errors reported in the following tables are the L_2-norm difference between the computed numerical values and the reference solution. Moreover, the Experimental Order of Convergence (EOC) appearing in the $(i + 1)$th row of Tables 4–6 is defined as,

$$\mathrm{EOC}_{i+1} = \frac{\ln \mathrm{error}_{i+1} - \ln \mathrm{error}_i}{\ln N_{\tau,i} - \ln N_{\tau,i+1}}.$$

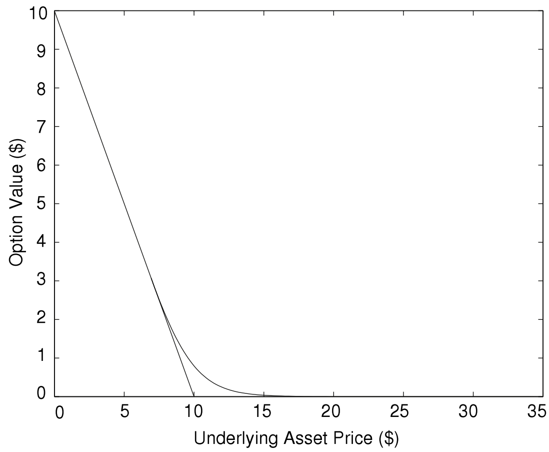
According to Table 4, when the grid sizes in both x and v directions are fixed and kept to be quite small, the EOCs reported are close to 1, indicating that our scheme is first-order convergent in the time direction.

Similarly, when we fix the time step size to $\Delta\tau = \frac{T}{100\,000}$, and the grid size in the x (or v) direction to be $\frac{X_{\max}}{400}$ (or $\frac{V_{\max}}{400}$), and increase the grid number in the v direction (or in the x direction), we find that the EOCs approach 2, as shown in Tables 5 and 6, respectively. This indicates that a second-order convergence is achieved in both x and v directions.

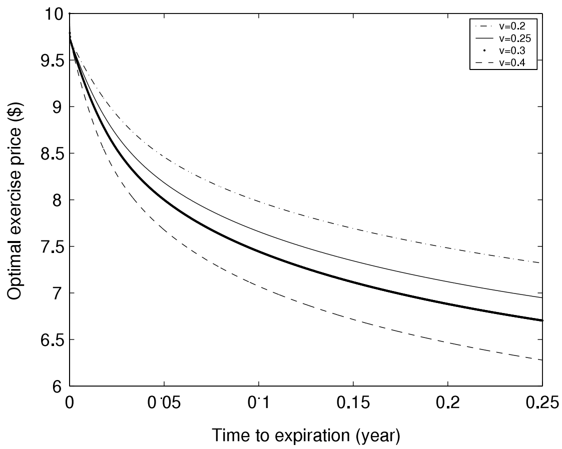
To better investigate the convergence rate of the current scheme, we also calculated the EOCs with the time and spatial steps adjusted to each other according to the expected order of error $\mathcal{O}(\Delta\tau) + \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta v^2)$. Specifically, we choose



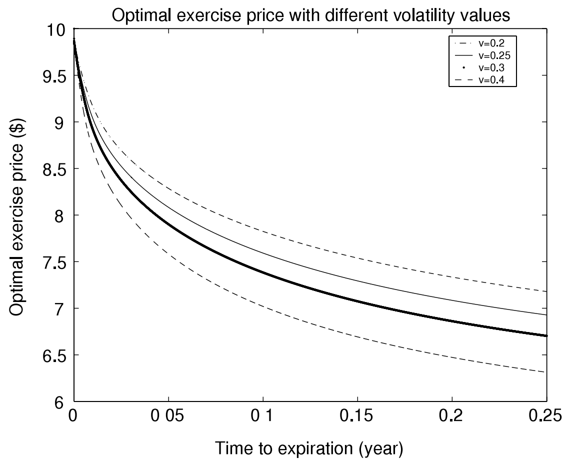
(a) American put option value at $v = 0.25$ with $\kappa = 2.5$, $\sigma = 0.45$..



(b) American put option value at $v = 0.25$ with $\kappa = 5$, $\sigma = 0.9$.



(c) Optimal exercise prices with different volatility values. Model parameters are $\kappa = 2.5$, $\sigma = 0.45$.



(d) Optimal exercise prices with different volatility values. Model parameters are $\kappa = 5$, $\sigma = 0.9$.

Fig. 3. American puts and the optimal exercise prices calculated with $\eta = 0.16$, $r = 0.1$, $\rho = 0.1$, $T = 0.25$ (year), $K = \$10.0$, $\theta = \frac{1}{2}$.

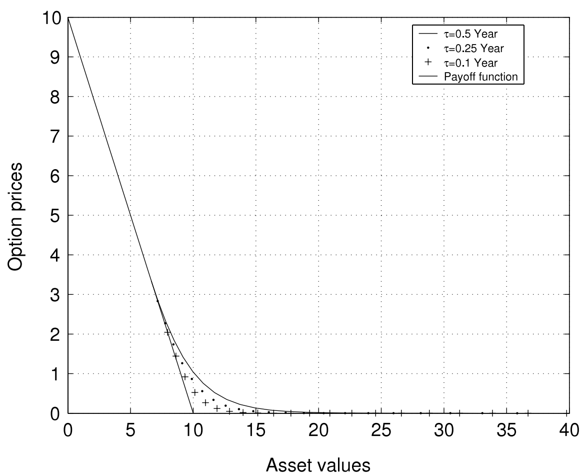


Fig. 4. Option prices at different times to expiration. Model parameters are $\kappa = 2.5$, $\eta = 0.16$, $\sigma = 0.45$, $r = 0.1$, $\rho = 0.1$, $K = \$10.0$.

varying grid sizes, i.e., $\Delta\tau_i = h_i^2\Delta\tau_1$, $\Delta x_i = h_i\Delta x_1$, $\Delta v_i = h_i\Delta v_1$, where h_i is the rate of the grid spacings used in the i-th line of Table 7 to those appearing in the first line of the same table. We could anticipate that if the theoretical order of

Table 4
EOC in the time direction, calculated with $\Delta x = \frac{X_{\max}}{250}$, $\Delta v = \frac{V_{\max}}{250}$.

No. of time steps	Difference	Ratio
4000	4.8699E−4	
4500	4.3639E−4	0.9314
5000	3.9212E−4	1.0153
5500	3.7315E−4	1.0679

Table 5
EOC in the x direction, calculated with $\Delta\tau = \frac{T}{100\,000}$, $\Delta v = \frac{V_{\max}}{400}$.

No. of steps in x-direction	Difference	Ratio
20	0.2092	
30	0.0891	2.1052
40	0.0469	2.2305
80	0.0097	2.2711

Table 6
EOC in the v direction, calculated with $\Delta\tau = \frac{T}{100\,000}$, $\Delta x = \frac{X_{\max}}{400}$.

No. of steps in v-direction	Difference	Ratio
10	0.0547	
20	0.0152	1.8475
40	0.0046	1.7244
80	0.0013	1.8231

Table 7
EOC with varying grid sizes, with $\Delta x_1 = \frac{X_{\max}}{10}$, $\Delta v_1 = \frac{V_{\max}}{10}$, and $\Delta\tau_1 = \frac{T_{\max}}{5}$.

h	Difference	Ratio
1	0.9265	
0.5	0.2783	1.7352
0.25	0.0580	2.2625
0.1	0.0070	2.3077

convergence is achieved in all directions, the EOC, which is now defined as

$$\mathrm{EOC}_{i+1} = \frac{\ln \mathrm{error}_{i+1} - \ln \mathrm{error}_i}{\ln h_{i+1} - \ln h_i},$$

should approach 2. From Table 7, we find that the EOCs are approximately equal to 2, which confirms again that the current scheme is indeed first-order convergent in the time direction, and second-order convergent in both x and v directions.

4.3. Discussion on accuracy and efficiency

In the financial industry, an ideal numerical scheme is the one with both high efficiency and accuracy. However, generally, the computational efficiency is always inversely proportional to accuracy. Thus, whether or not one can achieve a high efficiency at the expense of losing certain degrees of accuracy should be considered as a key criterion for choosing good numerical method.

In this subsection, we study numerically the relationship between the efficiency and the accuracy of our current scheme. We shall demonstrate that with an acceptable accuracy, our scheme can achieve a relatively high speed. Once again, we shall demonstrate this issue by showing the results of the calculation of the optimal exercise price only.

The tested example is chosen with parameter values: $\kappa = 1.5$, $\eta = 0.16$, $\sigma = 0.2$, $r = 0.1$, $\rho = 0.1$, $T - t = 0.25$ (year), and $K = \$10.0$.

Unfortunately, there is no exact or analytical solution for the valuation of American put options with stochastic volatility. Thus, we computed the reference values using the predictor–corrector method with a very fine grid defined by $N_x = 200$, $N_v = 400$, and $N_\tau = 50\,000$. These values are used to verify the accuracy of the computed S_f values based on some coarse grid.

Table 8 shows the relationship of the efficiency and accuracy of our scheme. Here, the computational efficiency is measured by the total CPU time in seconds consumed for each run, while the accuracy is measured by the relative error

Table 8
Report on CPU time vs. relative error.

(N_x, N_v, N_τ)	CPU time (s)	Error (%)
(13, 25, 25)	0.5470	6.67
(13, 50, 25)	0.8440	6.43
(26, 50, 80)	3.4370	2.26
(52, 50, 250)	19.5420	1.09
(104, 100, 2000)	529.4920	0.89

over the whole computational domain, which is defined as:

$$\text{Error} = \frac{\|S_f - \tilde{S_f}\|_\infty}{\|\tilde{S_f}\|_\infty}$$

where S_f and $\tilde{S_f}$ denote the computed values and the reference values, respectively, and $\|\cdot\|_\infty$ denotes the infinite norm. All the experiments here were performed within Matlab6.1 on an Intel Pentium 4, 3 GHz machine.

As clearly shown in Table 8, the accuracy is inversely varying with the efficiency. Furthermore, a better resolution in the asset values or the variance values implies more computational time. It can also be observed that our numerical results converge quickly to the reference values, and even with the most coarse grid, the accuracy is still acceptable by the industry standard. On the other hand, based on previous sections, it is suggested that the current method requires a lower number of time step intervals when σ is small. One should note that σ is usually very small in the real market. Therefore, our scheme is indeed a good one with a high computational efficiency and a satisfactory accuracy, and is suited for practical use.

5. Conclusion

In this paper, we first propose a set of appropriate boundary conditions along the v direction for American puts under the Heston model, and make the valuation system closed. Then, under this closed system, a new predictor–corrector scheme based on the ADI method is demonstrated and tested. The novelty of the current scheme, in comparison with some other finite difference methods in the literature, lies in the fact that it requires no embedded iterations, and can capture the whole optimal exercise boundary as part of the solution procedure. Based on the local von Neumann analysis, combined with the "frozen" coefficient technique, a conditional stability requirement is also obtained for the predictor–corrector method. It is suggested that the good convergence property of the corrector can influence the whole procedure significantly, which has also been confirmed by some of our numerical results. This approach can be easily extended to other stochastic volatility models and to non-zero choices of the market price of risk function, as long as a front-fixing transform exists. Various numerical examples suggest that the proposed scheme is both accurate and efficient, and is also suitable for practical use.

Acknowledgments

The authors gratefully acknowledge a useful discussion on the issue of the appropriateness of the boundary conditions along the $v = 0$ boundary with Prof. Daniel Ševčovič, Comenius University, Slovakia, who visited the University of Wollongong in March of 2010. The addition of Theorem 5 was suggested by Prof. Daniel Ševčovič and proved by the authors.

The authors would also like to thank an anonymous referee who raised this point in his reviewer's report. The suggested revision for the discussion based on the Fichera function has certainly improved this part of the paper.

Appendix A

In this appendix we consider the consistency between Eq. (2.8) and a newly constructed PDE system, which is the same as the PDE system (2.16), but without the boundary constraint at $v = 0$. As demonstrated in Section 2.2, this new PDE system is already closed when $\kappa\eta \geq \frac{\sigma^2}{2}$. Note that in the following theorem, "the PDE system" always refers to this newly-constructed PDE system, unless stated elsewhere.

Theorem 5. *For $\kappa\eta \geq \frac{\sigma^2}{2}$, the proposed boundary condition (2.8) for $v = 0$, is consistent with the inherent value of the PDE system as $v \to 0$.*

Proof. From a perturbation point of view, this statement is equivalent to showing that Eq. (2.8) will not create any singularity or boundary layer as $v \to 0$. Mathematically, we need to prove that the leading order term as $v \to 0$ deduced from the PDE system is exactly the same as Eq. (2.8).

The asymptotic behavior of the PDE system can be examined with the perturbation technique. First, we rewrite v as $v = \epsilon V$, with ϵ being a small parameter and V being an $\mathcal{O}(1)$ variable, to reflex the fact that v is an extremely small number in the limit process of v approaching zero (cf. [42]). Now, substituting $v = \epsilon V$ into the governing equation of the PDE system, and assuming that its solution can be written in powers of ϵ, i.e., $U(S, V, t) = \sum_{n=0}^{\infty} U_n(S, V, t)\epsilon^n$, we obtain the leading order term as

$$U_0(S, V, t) = C_1(S, t) + C_2(S, t)\frac{V^{1-\alpha}}{\alpha}$$

as $v \to 0$, where $\alpha = \frac{2\kappa\eta}{\sigma^2}$. Clearly, when $\alpha > 1$, i.e., $\kappa\eta \geq \frac{\sigma^2}{2}$, U_0 has an exponential growth as $V \to 0$, unless C_2 is set to zero. With C_2 being set to zero to maintain the boundedness of the American put value as the solution of the PDE system, U_0 is now a function of S and t only. Since U_0 also needs to satisfy the initial and boundary conditions along the t and S directions, respectively, one has no choice but letting $C_1(S, t) = \max(K - S, 0)$. Using Eq. (2.7), the boundary condition (2.8) can be easily deduced. Consequently, the consistency between Eq. (2.8) and the inherent value of the PDE system as $v \to 0$ for the case of $\kappa\eta \geq \frac{\sigma^2}{2}$ is established. □

The proof of this theorem leads to the final establishment of the PDE system (2.16) for pricing American put options under the Heston model, with a key boundary condition (2.8) being imposed regardless of what the ratio $\frac{2\kappa\eta}{\sigma^2}$ is.

Appendix B

Once the space discretization is performed, the governing equation in the PDE system (3.1) becomes:

$$\frac{\partial U_{i,j}^n}{\partial \tau} = a_j\delta_{xx}U_{i,j}^n + b_j\delta_{vv}U_{i,j}^n + (d_j + \lambda_j)\delta_x U_{i,j}^n + c_j\delta_{xv}U_{i,j}^n + e_j\delta_v U_{i,j}^n - rU_{i,j}^n.$$

By applying the first-order fully implicit Euler scheme to the time derivative $\frac{\partial U_{i,j}^n}{\partial \tau}$, we can obtain

$$\frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta\tau} = a_j\delta_{xx}U_{i,j}^{n+1} + b_j\delta_{vv}U_{i,j}^{n+1} + (d_j + \lambda_j)\delta_x U_{i,j}^{n+1} + c_j\delta_{xv}U_{i,j}^{n+1} + e_j\delta_v U_{i,j}^{n+1} - rU_{i,j}^{n+1}.$$

In order to demonstrate the ADI method clearly, we use linear operators A_0, A_1 and A_2 to denote the mixed derivative, the spatial derivatives in the x direction and the spatial derivatives in the v direction respectively, i.e.,

$$A_0U_{i,j}^n = \Delta\tau \cdot c_j\delta_{xv}U_{i,j}^n,$$
$$A_1U_{i,j}^n = \Delta\tau \cdot \left(a_j\delta_{xx}U_{i,j}^n + (d_j + \lambda_j)\delta_x U_{i,j}^n - \frac{r}{2}U_{i,j}^n\right),$$
$$A_2U_{i,j}^n = \Delta\tau \cdot \left(b_j\delta_{vv}U_{i,j}^n + e_j\delta_v U_{i,j}^n - \frac{r}{2}U_{i,j}^n\right).$$

Thus, the governing equation in the PDE system (3.1) can be written shortly as:

$$[I - (A_0 + A_1 + A_2)]U^{n+1} = U^n + \mathcal{O}((\Delta\tau)^2).$$

Similarly, the first-order explicit Euler scheme reads:

$$[I + (A_0 + A_1 + A_2)]U^n = U^{n+1} + \mathcal{O}((\Delta\tau)^2).$$

Thus, the weighted average of the fully implicit scheme and explicit scheme can be written as

$$[I - \theta(A_0 + A_1 + A_2)]U^{n+1} = [I + (1 - \theta)(A_0 + A_1 + A_2)]U^n + \mathcal{O}((\Delta\tau)^3). \tag{B.1}$$

One should note that when θ is equal to zero or one, Eq. (B.1) goes back to fully explicit Euler scheme or fully implicit one. When $\theta = \frac{1}{2}$, it is equivalent to apply the Crank–Nicolson scheme to the time derivative $\frac{\partial U^n}{\partial \tau}$.

Adding $\theta^2 A_1A_2U^{n+1}$ to both sides of Eq. (B.1), we have:

$$\begin{aligned}[I - \theta A_0 - \theta A_1 - \theta A_2 + \theta^2 A_1A_2]U^{n+1} &= [I + (1 - \theta)A_0 + (1 - \theta)A_1 + (1 - \theta)A_2 + \theta^2 A_1A_2] \\ &\quad \times U^n + \theta^2 A_1A_2(U^{n+1} - U^n) + \mathcal{O}((\Delta\tau)^3).\end{aligned} \tag{B.2}$$

As $U^{n+1} - U^n \sim \mathcal{O}(\Delta\tau)$, and $A_1A_2 \sim \mathcal{O}((\Delta\tau)^2)$, we can merge the term $\theta^2 A_1A_2(U^{n+1} - U^n)$ into the error term. Thus, Eq. (B.2) becomes:

$$(I - \theta A_1)(I - \theta A_2)U^{n+1} - \theta A_0U^{n+1} = [I + (1 - \theta)A_0 + (1 - \theta)A_1 + (1 - \theta)A_2 + \theta^2 A_1A_2]U^n + \mathcal{O}((\Delta\tau)^3). \tag{B.3}$$

However, it is still difficult to solve Eq. (B.3) alternatively in two directions because of the existence of θA_0U^{n+1}. One possible measure is to treat the mixed derivative fully explicit at the expense of losing one order of accuracy in the time direction, i.e.,

$$\begin{aligned}(I - \theta A_1)(I - \theta A_2)U^{n+1} &= [I + A_0 + (1 - \theta)A_1 + (1 - \theta)A_2 + \theta^2 A_1A_2]U^n + \theta A_0(U^{n+1} - U^n) + \mathcal{O}((\Delta\tau)^3), \\ &= [I + A_0 + (1 - \theta)A_1 + (1 - \theta)A_2 + \theta^2 A_1A_2]U^n + \mathcal{O}((\Delta\tau)^2),\end{aligned}$$

which is equivalent to:

$$(I - \theta A_1)(I - \theta A_2)U^{n+1} = [I + A_0 + (1 - \theta)A_1 + A_2]U^n - (I - \theta A_1)\theta A_2U^n. \tag{B.4}$$

Here, the linear operators A_0 and A_1 in Eq. (B.4) are time-dependent. We value them at the $(n+\theta)$th time step, i.e.,

$$\begin{aligned} A_0 &= \Delta\tau c_j \delta_{xv} \\ &= \Delta\tau\Big(\rho\sigma v - \sigma^2 v\xi\Big), \\ A_1 &= \Delta\tau\left[a_j\delta_{xx} + (d_j+\lambda_j)\delta_x - \frac{r}{2}U^n_{i,j}\right] \\ &= \Delta\tau\left[\left(\frac{v}{2} + \frac{\sigma^2 v}{2}\xi^2 - \rho\sigma v\xi\right)\delta_{xx} + \left(-\frac{v}{2} + \frac{\sigma^2 v}{2}\xi^2 - \frac{\sigma^2 v}{2}\beta + r - \kappa(\eta - v)\xi + \lambda\right)\delta_x - \frac{1}{2}rI\right], \end{aligned}$$

where

$$\begin{aligned} \xi &= \frac{\delta_v(\theta S_f^{n+1} + (1-\theta)S_f^n)}{\theta S_f^{n+1} + (1-\theta)S_f^n}, \\ \beta &= \frac{\delta_{vv}(\theta S_f^{n+1} + (1-\theta)S_f^n)}{\theta S_f^{n+1} + (1-\theta)S_f^n}, \\ \lambda &= \frac{S_f^{n+1} - S_f^n}{(\theta S_f^{n+1} + (1-\theta)S_f^n)\Delta\tau}. \end{aligned}$$

Appendix C

The matrix form for calculating Y corresponding to Eq. (3.12) is:

$$AY_j = P_j + \mathbf{bnd}\, x_j,$$

with

$$\begin{aligned} Y_j &= (Y_{1,j}, Y_{2,j}, \ldots, Y_{N_x-1,j})^T, \\ P_j &= (p_{1,j}, p_{2,j}, \ldots, p_{N_x-1,j})^T, \\ \mathbf{bnd}\, x_j &= \begin{pmatrix} \theta\left(\dfrac{a_j\Delta\tau}{(\Delta x)^2} - \dfrac{d'_j\Delta\tau}{2\Delta x}\right)Y_{0,j} \\ \vdots \\ \theta\left(\dfrac{a_j\Delta\tau}{(\Delta x)^2} + \dfrac{d'_j\Delta\tau}{2\Delta x}\right)Y_{N_x,j} \end{pmatrix}, \end{aligned}$$

$$\begin{aligned} p_{ij} &= U^n_{i,j} + \frac{c_j\Delta\tau}{4\Delta x\Delta v}\Big(U^n_{i+1,j+1} + U^n_{i-1,j-1} - U^n_{i+1,j-1} - U^n_{i-1,j+1}\Big) + (1-\theta)\left(\frac{a_j\Delta\tau}{(\Delta x)^2} - \frac{d'_j\Delta\tau}{2\Delta x}\right)U^n_{i-1,j} \\ &\quad + (1-\theta)\left(-\frac{a_j\Delta\tau}{(\Delta x)^2} - \frac{r\Delta\tau}{2}\right)U^n_{i,j} + (1-\theta)\left(\frac{a_j\Delta\tau}{(\Delta x)^2} + \frac{d'_j\Delta\tau}{2\Delta x}\right)U^n_{i+1,j} + \left(\frac{b_j\Delta\tau}{(\Delta v)^2} - \frac{e_j\Delta\tau}{2\Delta v}\right)U^n_{i,j-1} \\ &\quad + \left(-\frac{b_j\Delta\tau}{(\Delta v)^2} - \frac{r\Delta\tau}{2}\right)U^n_{i,j} + \left(\frac{b_j\Delta\tau}{(\Delta v)^2} + \frac{e_j\Delta\tau}{2\Delta v}\right)U^n_{i,j+1}, \end{aligned}$$

$$A = \begin{pmatrix} 1+\theta\left(\dfrac{2a_j\Delta\tau}{\Delta x^2} + \dfrac{r\Delta\tau}{2}\right) & -\theta\left(\dfrac{a_j\Delta\tau}{\Delta x^2} + \dfrac{d'_j\Delta\tau}{2\Delta x}\right) & & \\ -\theta\left(\dfrac{a_j\Delta\tau}{\Delta x^2} - \dfrac{d'_j\Delta\tau}{2\Delta x}\right) & 1+\theta\left(\dfrac{2a_j\Delta\tau}{\Delta x^2} + \dfrac{r\Delta\tau}{2}\right) & & \\ & \ddots & \ddots & \\ & \ddots & \ddots & \\ & & & 1+\theta\left(\dfrac{2a_j\Delta\tau}{\Delta x^2} + \dfrac{r\Delta\tau}{2}\right) \end{pmatrix},$$

where $d'_j = (d_j + \lambda)$, and $\mathbf{bnd}\, x_j$ stands for the vector that contains the boundary value of the intermediate variable Y in the x direction.

The corresponding matrix form for computing U^{n+1} from Eq. (3.13) is:

$$BU_i^{n+1} = Q_i + \mathbf{bnd}\, v_i,$$

with

$$U_i^{n+1} = (U_{i,1}^{n+1}, U_{i,2}^{n+1} \cdots U_{i,N_v-1}^{n+1})^T,$$

$$Q_i = (q_{i,1}, q_{i,2} \cdots q_{i,N_v-1})^T,$$

$$\mathbf{bnd}\, v_j = \begin{pmatrix} \theta\left(\dfrac{b_j \Delta\tau}{\Delta v^2} - \dfrac{e_j \Delta\tau}{2\Delta v}\right) U_{i,0}^{n+1} \\ \vdots \\ \theta\left(\dfrac{b_j \Delta\tau}{\Delta v^2} + \dfrac{e_j \Delta\tau}{2\Delta v}\right) U_{i,N_v}^{n+1} \end{pmatrix},$$

$$q_{i,j} = Y_{i,j} - \theta\left(\frac{b_j \Delta\tau}{\Delta v^2} - \frac{e_j \Delta\tau}{2\Delta v}\right) U_{i,j-1}^n - \theta\left(-\frac{b_j \Delta\tau}{\Delta v^2} - \frac{r\Delta\tau}{2}\right) U_{i,j}^n - \theta\left(\frac{b_j \Delta\tau}{\Delta v^2} + \frac{e_j \Delta\tau}{2\Delta v}\right) U_{i,j+1}^n,$$

$$B = \begin{pmatrix} 1+\theta\left(\dfrac{2b_j\Delta\tau}{\Delta v^2} + \dfrac{r\Delta\tau}{2}\right) & -\theta\left(\dfrac{b_j \Delta\tau}{\Delta v^2} + \dfrac{e_j \Delta\tau}{2\Delta v}\right) & & & \\ -\theta\left(\dfrac{b_j \Delta\tau}{\Delta v^2} - \dfrac{e_j \Delta\tau}{2\Delta v}\right) & 1+\theta\left(\dfrac{2b_j\Delta\tau}{(\Delta v)^2} + \dfrac{r\Delta\tau}{2}\right) & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1+\theta\left(\dfrac{2b_j\Delta\tau}{\Delta v^2} + \dfrac{r\Delta\tau}{2}\right) \end{pmatrix},$$

where **bnd** v_j stands for the vector that includes the boundary value of U^{n+1} in the v direction.

References

[1] A.D. Adrian, M.Y. Victor, Probability distribution of returns in the Heston model with stochastic volatility, Quantitative Finance 2 (2002) 443–453.
[2] C.A. Ball, A. Roma, Stochastic volatility option pricing, Journal of Financial and Quantitative Analysis 29 (1994) 589–687.
[3] J.-P. Fouque, G. Papanicolaou, K.R. Sircar, Derivatives in Financial Markets with Stochastic Volatility, Cambridge University Press, Cambridge, 2000.
[4] S. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, The Reviews of Financial Studies 6 (1993) 327–343.
[5] J.C. Hull, A. White, The pricing of options on assets with stochastic volatilities, The Journal of Finance 42 (1987) 281–300.
[6] W. Feller, Two singular diffusion problems, Annals of Mathematics 54 (1951) 173–182.
[7] G. Bakshi, C. Cao, Z. Chen, Empirical performance of alternative option pricing models, The Journal of Finance 52 (5) (1997) 2003–2049.
[8] J. Pan, The jump-risk premia implicit in options: evidence from an integrated time-series study, The Journal of Financial Economics 63 (2002) 3–50.
[9] R.G. Tompkins, Stock index futures markets: stochastic volatility and smiles, Journal of Futures Markets 21 (2001) 43–78.
[10] L. Wu, Y.K. Kwok, A front-fixing finite difference method for the valuation of American options, The Journal of Financial Engineering 6 (2) (1997) 83–97.
[11] S.-P. Zhu, J. Zhang, A new predictor–corrector scheme for valuating American puts, Applied Mathematics and Computation 217 (9) (2011) 4439–4452.
[12] W. Allegretto, Y.-P. Lin, H. Yang, A fast and highly accurate numerical methods for the valuation of American options, Discrete and Continuous Dynamical Systems. Series B. Applications and Algorithm 8 (2002) 127–136.
[13] J. Cox, S. Ross, M. Rubinstein, Option pricing—a simplified approach, Journal of Financial Economics 7 (1979) 229–263.
[14] F. Longstaff, E. Schwartz, Valuing American options by simulation: a simple least-squares approach, The Reviews of Financial Studies 14 (2001) 113–147.
[15] R. Geske, H. Johnson, The American put option valued analytically, The Journal of Finance 39 (1984) 1511–1524.
[16] S.-P. Zhu, A new analytical-approximation formula for the optimal exercise boundary of American put options, International Journal of Theoretical and Applied Finance 9 (7) (2006) 1141–1177.
[17] D.S. Bunch, H. Johnson, The American put option and its critical stock price, The Journal of Finance 5 (2000) 2333–2356.
[18] S.-P. Zhu, Z.-W. He, Calculating the early exercise boundary of American put options with an approximation, International Journal of Theoretical and Applied Finance 10 (7) (2007) 1203–1227.
[19] J. Huang, M. Subrahmanyam, G. Yu, Pricing and hedging American options: a recursive integration method, The Reviews of Financial Studies 9 (1996) 277–300.
[20] D. Ševčovič, Analysis of the free boundary for the pricing of an American call option, European Journal of Applied Mathematics 12 (2001) 25–37.
[21] D. Ševčovič, An iterative algorithm for evaluating approximations to the optimal exercise boundary for a nonlinear Black–Scholes equation, Canadian Applied Mathematics Quarterly 15 (1) (2007) 77–97.
[22] R. Stamicar, D. Ševčovič, J. Chadam, The early exercise boundary for the American put near expiry: numerical approximation, Canadian Applied Mathematics Quarterly 7 (4) (1999) 427–444.
[23] S.-P. Zhu, An exact and explicit solution for the valuation of American put options, Quantitative Finance 6 (3) (2006) 229–242.

[24] N. Clarke, K. Parrott, Multigrid for American option pricing with stochastic volatility, Applied Mathematical Finance 6 (3) (1999) 177–195.
[25] S. Ikonen, J. Toivanen, Effcient numerical methods for pricing American options under stochastic volatility, Numerical Methods for Partial Differential Equations 24 (1) (2008) 104–126.
[26] R. Zvan, P.A. Forsyth, K.R. Vetzal, Penalty methods for American options with stochastic volatility, Journal of Computational and Applied Mathematics 91 (1998) 199–218.
[27] H.G. Landau, Heat conduction in a melting solid, Quarterly of Applied Mathematics 8 (1950) 81.
[28] C.J. Strikwerda, Finite Difference Schemes and Partial Differential Equations, Chapman & Hall, New York, 1989.
[29] P. Wilmott, J. Dewynne, S. Howison, Option Pricing, first ed., Oxford Financial Press, 1993
[30] S.-S. Lin, Finite difference schemes for Heston's model, A Graduate Project, 2008.
[31] Z. Zhe, K.-G. Lim, A non-lattice pricing model of American options under stochastic volatility, Journal of Futures Markets 26 (5) (2006) 417–448.
[32] K. Ito, J. Toivanen, Lagrange multiplier approach with optimized finite difference stencils for pricing American options under stochastic volatility, SIAM Journal on Scientific Computing 31 (2009) 2646–2664.
[33] G. Fichera, Existential analysis of the solutions of mixed boundary value problems, related to second order elliptic equation and systems of equations, selfadjoint, Annali della Scuola Normale Superiore, III 1 (1–4) (1949) 75–100.
[34] J.C. Hull, Options, Futures, and Other Derivatives, the third ed., Prentice Hall Press, 1997.
[35] CBOE's volatility index, 2009. Available at: http://www.cboe.com/micro/vix/volatility-qrg.pdf.
[36] J. Douglas, H.H. Rachford, On the numerical solution of heat conduction problems in two and three space variables, Transactions of the American Mathematical Society 82 (1956) 421–439.
[37] I.J.D. Craig, A.D. Sneyd, An alternating-direction implicit scheme for parabolic equations with mixed derivatives, Computers and Mathematics with Applications 16 (4) (1988) 341–350.
[38] W. Hundsdorfer, Accuracy and stability of splitting with stabilizing corrections, Applied Numerical Mathematics 42 (2002) 213–233.
[39] K.J. in't Hout, B.D. Welfert, Stability of ADI schemes applied to convection–diffusion equations with mixed derivative terms, Applied Numerical Mathematics 57 (2006) 19–35.
[40] A. Medvedev, O. Scaillet, Pricing American options under stochastic volatility and stochastic interest rates, Journal of Financial Economics 98 (1) (2010) 145–159.
[41] C.W. Oosterlee, On multigrid for linear complementarity probelms with application to American-style options, Electronic Transactions on Numerical Analysis 15 (2003) 165–185.
[42] J.D. Evans, R. Kuske, J.B. Keller, American options on assets with dividends near expiry, Mathematical Finance 12 (3) (2002) 219–237.