

Chen, Wenting; Zhu, Song-Ping

## Article

# On the asymptotic behavior of the optimal exercise price near expiry of an American put option under stochastic volatility

Journal of Risk and Financial Management

## Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* Chen, Wenting; Zhu, Song-Ping (2022) : On the asymptotic behavior of the optimal exercise price near expiry of an American put option under stochastic volatility, Journal of Risk and Financial Management, ISSN 1911-8074, MDPI, Basel, Vol. 15, Iss. 5, pp. 1-19, <https://doi.org/10.3390/jrfm15050189>

This Version is available at:

<https://hdl.handle.net/10419/274711>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>



Article

# On the Asymptotic Behavior of the Optimal Exercise Price Near Expiry of an American Put Option under Stochastic Volatility

Wenting Chen <sup>1</sup> and Song-Ping Zhu <sup>2,\*</sup>

<sup>1</sup> School of Business, Jiangnan University, Wuxi 214126, China; wtchen@jiangnan.edu.cn

<sup>2</sup> School of Mathematics and Applied Statistics, University of Wollongong, Wollongong City 2500, Australia

\* Correspondence: spz@uow.edu.au

**Abstract:** The behavior of the optimal exercise price of American puts near expiry has been well studied under the Black–Scholes model as a result of a series of publications. However, the behavior of the optimal exercise price under a stochastic volatility model, such as the Heston model, has not been reported at all. Adopting the method of matched asymptotic expansions, this paper addresses the asymptotic behavior of American put options on a dividend-paying underlying with stochastic volatility near expiry. Through our analyses, we are able to show that the option price will be quite different from that evaluated under the Black–Scholes model, while the leading-order term of the optimal exercise price remains almost the same as the constant volatility case if the spot volatility is given the same value as the constant volatility in the Black–Scholes model. Results from numerical experiments also suggest that our analytical formulae derived from the asymptotic analysis are quite reasonable approximations for options with remaining times to expiry in the order of days or weeks.

**Keywords:** singular perturbation; matched asymptotic expansions; American put options; optimal exercise price; the Heston model



**Citation:** Chen, Wenting, and Song-Ping Zhu. 2022. On the Asymptotic Behavior of the Optimal Exercise Price Near Expiry of an American Put Option under Stochastic Volatility. *Journal of Risk and Financial Management* 15: 189. <https://doi.org/10.3390/jrfm15050189>

Academic Editor: Zbigniew Palmowski

Received: 3 March 2022

Accepted: 13 April 2022

Published: 19 April 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

It is well known now that the Black–Scholes model, which is a breakthrough in the financial area, is inadequate to describe asset returns and behaviors of the option markets (see [Dragulescu and Yakovenko 2002](#)). This is because the Black–Scholes model assumes that the underlying asset satisfies the lognormal distribution, the right tail of which has quite often displayed a mismatch with empirically observed data. One possible remedy is to assume that the volatility of the asset price also follows a stochastic process (see [Ball and Roma 1994](#); [Fouque et al. 2000](#); [Heston 1993](#); [Hull and White 1987](#)), rather than a constant, as assumed in the original Black–Scholes' framework. In this paper, we use the stochastic model introduced by [Heston \(1993\)](#) to price American options. In this model, we assume that the variance (the square of the underlying price volatility) follows a random process known in the financial literature as the Cox–Ingersoll–Ross (CIR) process and in mathematical statistics as the Feller process (see [Feller 1951](#); [Fouque et al. 2000](#)). Empirical studies suggest that this non-negative and mean-reverting process is indeed more consistent with the observations of real markets (see [Bakshi et al. 1997](#); [Pan 2002](#); [Tompkins 2001](#)). For example, [Dragulescu and Yakovenko \(2002\)](#) 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 index after a careful calibration process to determine the appropriate parameters to be used in the model. Supported by much empirical evidence in the literature already, the Heston model has recently received much research attention in traditional finance, as well as quantitative finance, particularly in the latter (see [Altmayer and Neuenkirch 2016](#); [Malham et al. 2021](#); [Mickel and Neuenkirch 2021](#) and the references therein). In this paper, we shall complement the existing literature with a study on the asymptotic behavior of optimal exercise price near expiry of an American put option under a typical stochastic volatility model.

As is well known now, the main difficulty for pricing American options stems from the inherent nonlinear nature of the option contract itself, i.e., the additional right written in the contract for the holder to exercise the option at any time prior to the expiry date. In the context of partial differential equation (PDE) approaches, this is reflected in the fact that the corresponding PDE is associated with an unknown moving boundary, and the problem thus becomes a moving boundary problem. As a result, no useful analytical methods are hitherto available for pricing American options under the Heston model, and thus, numerical methods are preferred by market practitioners. However, it is usually difficult to retain accuracy in approximating the optimal exercise price by means of the numerical methods, and the inaccuracy becomes more intolerable when the time is closer to expiry. This is because within the short tenor, the velocity, with which the optimal exercise price reaches its final value, is extremely fast and is thus difficult to approximate by numerical methods. For example, when using the predictor–corrector finite difference scheme (see [Zhu and Chen 2011](#)), a very fine discretization must be adopted near expiry, which is both expensive and limited in accuracy. Therefore, it is quite reasonable to infer that, under the Heston model, the optimal exercise price is also singular at expiry, which is true under the Black–Scholes’ framework (see [Evans et al. 2002](#)). On the other hand, it is quite useful to determine the asymptotic behavior of the optimal exercise price near expiry, since this asymptotic solution can be used as a complement to the numerical approaches to calculate the option values and the optimal exercise price for other larger times away from the expiry.

In the literature, some analyses of the asymptotic behavior of the optimal exercise price near expiry have already been carried out. For instance, [Barles et al. \(1995\)](#) derived the first term for the optimal exercise price of American puts on a non-dividend-paying underlying by constructing a subsolution, as well as a supersolution. The examples in their paper considered options with maturities less than one year, in which case, the inaccuracy of the approximation was not significant. [Chen and Chadam \(2007\)](#) analyzed the same problem as Barles et al. did and provided four approximations for the optimal exercise prices of American puts with different maturities by the method of integral equations. [Evans et al. \(2002\)](#) considered the asymptotic behavior of the optimal exercise price near expiry on an asset with a constant dividend yield rate. The approximation was derived both with the utilization of the method of integral equations and the method of matched asymptotic expansions and was expected to be useful for time scales in the order of days and weeks from expiry. Moreover, [Chen and Zhu \(2009\)](#) extended Evans et al.’s results to the local volatility model, in which the volatility is assumed to be dependent on both time and the underlying asset price. [Yang \(2019\)](#) considered the short-maturity asymptotic behavior of the optimal exercise price of American lookback options under the random walks of the order-two model. On the other hand, it is a non-trivial task to derive asymptotic behaviors for multi-dimensional problems, and thus, this has received much research interest. For example, [Qin et al. \(2019\)](#) derived the asymptotic behavior of the optimal exercise strategy for a small number of executive stock options. It is also not easy, as will be shown in this paper, to determine the asymptotic behavior of the optimal exercise price near expiry under the Heston model, since in this case, the optimal exercise price depends, in addition to time, on the dynamics of volatility as well. In other words, the introduction of a second stochastic process has produced a number of new phenomena, which have in turn made the problem much more complicated and totally different from the case with a single stochastic process being used to describe the behavior of the underlying asset only while the volatility is assumed to be a constant.

The aim of this paper is to present an explicit analytical approximation for the optimal exercise price near expiry under the Heston model by means of the method of matched asymptotic expansions. The approximation could help readers understand clearly the near-expiry behavior of the optimal exercise price, which is indeed useful for both theoretical and practical purposes. It turns out that, even with stochastic volatility taken into consideration, the convergence rate for the calculation of the optimal exercise price is almost the same as that for the constant volatility case.

The paper is organized as follows. In Section 2, we introduce the PDE system that the price of an American put option must satisfy under the Heston model. In Section 3, we deduce the asymptotic behavior of the optimal exercise price near expiry by using the singular perturbation method. In Section 4, we compare our approximation with the numerical results calculated by the predictor–corrector finite difference method (see [Zhu and Chen 2011](#)), to illustrate the reliability of our asymptotic solution. Concluding remarks are given in Section 5.

## 2. American Puts under the Heston Model

In this section, the Heston model and the PDE system for American puts underneath will be briefly reviewed. There are two main reasons why we choose the Heston model for the current work. Firstly, various studies (see [Bakshi et al. 1997](#); [Pan 2002](#); [Tompkins 2001](#)) suggest that the Heston model is consistent with the real market. Secondly, this model is analytical achievable only for European options. In the case of American options, no analytical solution for American options under the Heston model has been discovered yet.

The hypotheses used in the current paper are the same as those of the Heston model. In the Heston model, the underlying  $S_t$ , as a function of time, is assumed to follow the stochastic differential equation (SDE) of a geometric Brownian motion (GBM) in the Itô form:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_t^1, \tag{1}$$

where  $\mu$  is the drift rate,  $W_t^1$  is a standard Brownian motion, and  $\sqrt{v_t}$  is the standard deviation of the stock returns  $\frac{dS_t}{S_t}$ . Furthermore, the variance  $v_t$  (the square of the volatility) is assumed to be governed by the following mean-reverting SDE:

$$dv_t = \kappa(\eta - v_t)dt + \sigma\sqrt{v_t}dW_t^2. \tag{2}$$

Here,  $\eta$  is the long-term mean of  $v_t$ ,  $\kappa$  is the rate of relaxation to this mean, and  $\sigma$  is the volatility of volatility.  $W_t^2$  is also a standard Brownian motion, and it is related to  $W_t^1$  with a correlation factor  $\rho \in [-1, 1]$ . (2) known in the financial literature as the CIR process and in mathematical statistics as the Feller process (see [Feller 1951](#); [Fouque et al. 2000](#)).

Let  $P_A(S, v, t)$  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. Then, under the proposed processes (1)–(2), it is shown that the valuation of an American put option can be formulated as a free boundary problem (see [Zhu 2006](#)), in which the boundary location itself is part of the solution of the problem. Specifically,  $P_A$  satisfies

$$\left\{ \begin{array}{l} \frac{1}{2}vS^2 \frac{\partial^2 P_A}{\partial S^2} + \rho\sigma vS \frac{\partial^2 P_A}{\partial S \partial v} + \frac{1}{2}\sigma^2 v \frac{\partial^2 P_A}{\partial v^2} + (r - D)S \frac{\partial P_A}{\partial S} \\ + \kappa(\eta - v) \frac{\partial P_A}{\partial v} - rP_A + \frac{\partial P_A}{\partial t} = 0, \\ P_A(S, v, T_E) = \max(K - S, 0), \\ \lim_{S \rightarrow \infty} P_A(S, v, t) = 0, \\ P_A(S_f(v, t), v, t) = K - S_f(v, t), \\ \frac{\partial P_A}{\partial S}(S_f(v, t), v, t) = -1, \\ \lim_{v \rightarrow 0} P_A(S, v, t) = \max(K - S, 0), \\ \lim_{v \rightarrow \infty} P_A(S, v, t) = K. \end{array} \right. \tag{3}$$

This PDE system is defined on  $S \in [S_f(v, t), +\infty]$ ,  $v \in [0, +\infty]$  and  $t \in [0, T_E]$ . We remark that in the above PDE system,  $K$  is the strike price,  $D$  is the dividend yield,  $T_E$  is the

expiry date, and  $r$  is the risk-free interest rate. Unlike the Black–Scholes’ case, the unknown optimal exercise price  $S_f$  now depends on both the time and the volatility.

For simplicity, in our work, we assume that  $r$  is greater than  $D$ . A similar analysis can also be carried out for the case of  $r < D$ , and is thus not included in the current paper. It should be remarked that, when  $r > D$ , the riskless growing argument proposed in [Zhu and Chen \(2011\)](#) also holds, and thus, the boundary conditions established therein can still be used here. Moreover, it is also straightforward to show that (see [Zhu and Chen 2011](#)):

$$S_f(v, T_E) = K, S_f(0, t) = K,$$

which financially simply states that at the expiration date or when the spot volatility is zero, the optimal exercise price is equal to the strike price. However, just as a similar case of the Stephan problem (see [Zhu 2006](#)), the “velocity”, with which the optimal exercise price reaches its final value, is extremely fast and is difficult to approximate by numerical approaches. Therefore, in the next section, we shall construct the asymptotic behavior of the optimal exercise price near expiry by using the method of matched asymptotic expansions.

### 3. Matched Asymptotic Expansions for the Optimal Exercise Price near Expiry

To make the analysis convenient, we shall first non-dimensionalize all variables by using the new variables

$$S = Ke^x, \quad P = \frac{P_A e^{q\tau}}{K} + e^{q\tau}(e^x - 1), \quad S_f = Ke^{x_f}, \quad \tau = \frac{\sigma^2}{2}(T_E - t),$$

where the parameters  $q$  and  $d$  are defined as

$$q = \frac{2r}{\sigma^2}, \quad d = \frac{2D}{\sigma^2},$$

respectively. Then, (3) can be written in the dimensionless form:

$$\left\{ \begin{array}{l} \frac{\partial P}{\partial \tau} = \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial x^2} + (q - v - \frac{v}{\sigma^2}) \frac{\partial P}{\partial x} + v \frac{\partial^2 P}{\partial v^2} + \frac{2\kappa}{\sigma^2} (\eta - v) \frac{\partial P}{\partial v} \\ \quad + \frac{2\rho v}{\sigma} \frac{\partial^2 P}{\partial x \partial v} + e^{q\tau}(de^x - q), \\ P(x, v, 0) = \max(e^x - 1, 0), \\ P(x_f, v, \tau) = 0, \\ \frac{\partial P}{\partial x}(x_f, v, \tau) = 0, \\ \lim_{x \rightarrow \infty} P(x, v, \tau) = e^{q\tau}(e^x - 1), \\ \lim_{v \rightarrow 0} P(x, v, \tau) = e^{q\tau} \max(1 - e^x, 0) + e^{q\tau}(e^x - 1), \\ \lim_{v \rightarrow \infty} P(x, v, \tau) = e^{q\tau+x}, \end{array} \right. \tag{4}$$

together with two more conditions for the optimal exercise price:

$$x_f(v, 0) = x_f(0, \tau) = 0. \tag{5}$$

One should notice that, although the governing differential equation itself in (4) is linear in terms of the unknown function  $P$ , it is the unknown boundary that has made this PDE system highly nonlinear. The nonlinearity of the problem will be clearly manifested once a Landau transform is used to convert the moving boundary problem into a fixed boundary problem, as demonstrated by [Zhu and Chen \(2011\)](#). On the other hand, the high nonlinearity, as well as the introduction of another new variable  $v$  has resulted in the analytical methods being less achievable than the Black–Scholes’ case (see [Evans et al.](#)

2002). Consequently, we shall use the method of matched asymptotic expansions, which is an ideal tool to deal with the nonlinear problems, to construct an approximation of  $x_f(v, \tau)$  for the PDE System (4). Hereafter, we only consider the options with a short tenor, i.e.,

$$\tau = \epsilon T, \tag{6}$$

where  $T = \mathcal{O}(1)$  and  $\epsilon$  is a small positive parameter. By substituting (6) into (4), we obtain the PDE system for  $P(x, v, T)$ :

$$\left\{ \begin{array}{l} \frac{\partial P}{\partial T} = \epsilon \left[ \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial x^2} + \left( q - d - \frac{v}{\sigma^2} \right) \frac{\partial P}{\partial x} + v \frac{\partial^2 P}{\partial v^2} \right. \\ \quad \left. + \frac{2\kappa}{\sigma^2} (\eta - v) \frac{\partial P}{\partial v} + \frac{2\rho v}{\sigma} \frac{\partial^2 P}{\partial x \partial v} \right] + \epsilon e^{\epsilon q T} (de^x - q), \\ P(x, v, 0) = \max(e^x - 1, 0), \\ P(x_f, v, T) = 0, \\ \frac{\partial P}{\partial x}(x_f, v, T) = 0, \\ \lim_{x \rightarrow \infty} P(x, v, T) = e^{\epsilon q T} (e^x - 1), \\ \lim_{v \rightarrow 0} P(x, v, T) = e^{q\epsilon T} \max(1 - e^x, 0) + e^{q\epsilon T} (e^x - 1), \\ \lim_{v \rightarrow \infty} P(x, v, T) = e^{q\epsilon T + x}. \end{array} \right. \tag{7}$$

Unlike the constant volatility case discussed in Evans et al. (2002), the PDE System (7) needs to be dealt with with care; there are several different regions, or the so-called “boundary layers”, in which either the unknown function  $P$  or its partial derivatives change rapidly. The increased complexity, as will be shown later, is a result of the introduction of stochastic volatility and, consequently, a new dimension of the PDE system. The “boundary layer” is a phrase commonly used in physics and fluid mechanics to describe the layer of fluid in the immediate vicinity of a boundary surface (see Friedman 2008), and it has been adopted in most of singular perturbation analyses as well. Therefore, we shall also use it in this paper to derive the specific boundary layer structure for the PDE System (7).

The procedure usually begins with the assumption that the solution can be expanded in powers of  $\epsilon$ , i.e.,

$$P_{outer} = P_0 + \epsilon P_1 + \mathcal{O}(\epsilon^2), \tag{8}$$

in which the subscript stands for the outer solution in the region outside of boundary layer where a singular expansion is required to deal with rapid changes of the value of  $P$  or its derivatives. According to the fundamental theory of the method of matched asymptotic expansions Smith (2009), only a regular expansion in the form of (8) is needed for the “outer region”; the correct order for the entire problem will be automatically determined once the “inner solution” and “outer solution” are forced to match. After some simple calculations, we obtain the solution as

$$P_{outer}(x, T) = e^x - 1 + qT\epsilon(e^x - 1) + \mathcal{O}(\epsilon^2). \tag{9}$$

It should be noted that the above solution is valid for the domain  $x > 0, 0 \leq v < \infty$ , and fails to satisfy the boundary condition at  $v = \infty$ . Therefore, there is a boundary layer near  $v = \infty$ , with the layer thickness in the order of  $\epsilon$ . It can be inferred that, in order to satisfy the boundary condition at  $v = \infty$ , the exact solution of (7), compared with  $P_{outer}$ , changes rapidly only within the  $\mathcal{O}(\epsilon)$  neighborhood of  $v = \infty$ . On the other hand, in the traded market, the volatility value is usually very small, and the highest value of the volatility that has ever been recorded on the Chicago Board Options Exchange (CBOE) is only 0.85 (see CBOE 2022). Therefore, with these two points in mind, it is perfectly

justifiable to not analyze this boundary layer at all; our solution will be a reasonably good approximation for any large enough, but finite  $v$  values.

On the other hand, it is clear that as  $x \rightarrow 0^+$ , too many terms on the right-hand side of the governing equation contained in (7) have been dropped, and thus, we need to perform an asymptotic analysis in the vicinity of  $x = 0$ . By using the stretched variable:

$$X = \frac{x}{\epsilon^\alpha}, \tag{10}$$

and substituting (10) into (7), we obtain

$$\frac{\partial P}{\partial T} = \epsilon^{1-2\alpha} \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial X^2} + \epsilon^{1-\alpha} \left( \rho - v - \frac{v}{\sigma^2} \right) \frac{\partial P}{\partial X} + \epsilon v \frac{\partial^2 P}{\partial v^2} + \epsilon \frac{2\kappa}{\sigma^2} (\eta - v) \frac{\partial P}{\partial v} + \epsilon^{1-\alpha} \frac{2\rho v}{\sigma} \frac{\partial^2 P}{\partial X \partial v} + \epsilon e^{\epsilon q T} (d e^{\epsilon^\alpha X} - q).$$

The significant degeneration of the above operator arises if  $\alpha = \frac{1}{2}$ , and thus, the boundary layer near  $x = 0$  is with the thickness of  $\mathcal{O}(\sqrt{\epsilon})$ . Assuming that  $P(X, v, T)$  has a regular expansion in this region, we write

$$P = \sqrt{\epsilon} P_0 + \epsilon P_1 + \epsilon^{\frac{3}{2}} P_2 + \mathcal{O}(\epsilon^2), \tag{11}$$

where the explicit analytical expressions of  $P_0, P_1$ , and  $P_2$  are derived in Appendix A. The solution derived in this boundary layer can be referred as the inner solution with respect to (with respect to hereafter)  $P_{outer}$ .

It can be easily shown that the first-order derivative of  $P_0$  with respect to  $v$  becomes unbounded at the corner  $(X, v) = (0, 0)$ . Consequently, the derivatives of  $P_0$  with respect to  $v$  cannot be ignored when deriving the PDE system for  $P_0$  around that corner. Another local analysis is again needed. By setting  $V = \frac{v}{\epsilon^\beta}$  and investigating the significant degeneration of the corresponding operator, it is obvious that  $\beta = 1$  is a well-balanced choice. Assuming that  $Z_1(X, V, T)$  (the solution at the corner) can be expanded in powers of  $\epsilon$ , we obtain

$$Z_1 = \sqrt{\epsilon} Z_{11} + \epsilon Z_{12} + \epsilon^{\frac{3}{2}} Z_{13} + \mathcal{O}(\epsilon^2),$$

where the explicit analytical expressions of  $Z_{11}, Z_{12}$ , and  $Z_{13}$  are derived in Appendix B. Furthermore, it can be easily shown that  $Z_{11}$  is continuous, but not differentiable with respect to  $X$  at the corner  $(X, v) = (0, 0)$ . Thus, the following stretched variables are needed:

$$X_1 = \frac{x}{\epsilon^\mu}, \quad V_1 = \frac{v}{\epsilon^\nu},$$

where  $\mu = \nu = 1$  are determined by investigating the significant degeneration of the corresponding operator again.

The above analysis has clearly demonstrated the boundary layer structure of our problem, as shown in Figure 1, where the Roman numbers *I, II, III*, and *IV* stand for four different regions in which local analysis needs to be carried out consecutively. In particular, Region *I* represents the valid domain of  $P_{outer}$ , while Region *II* shows the  $\mathcal{O}(\sqrt{\epsilon})$  boundary layer near  $x = 0$ . Regions *III* and *IV* denote the corner boundary layers, and they are defined as

$$\begin{aligned} (x, v) &\in [-\sqrt{\epsilon}, \sqrt{\epsilon}] \times [0, \epsilon], \\ (x, v) &\in [-\epsilon, \epsilon] \times [0, \epsilon], \end{aligned}$$

respectively. Moreover, near  $x_f$ , there might be another boundary layer, which will be discussed later. Note again that the  $\mathcal{O}(\epsilon)$  boundary layer near  $v = \infty$  has been ignored.

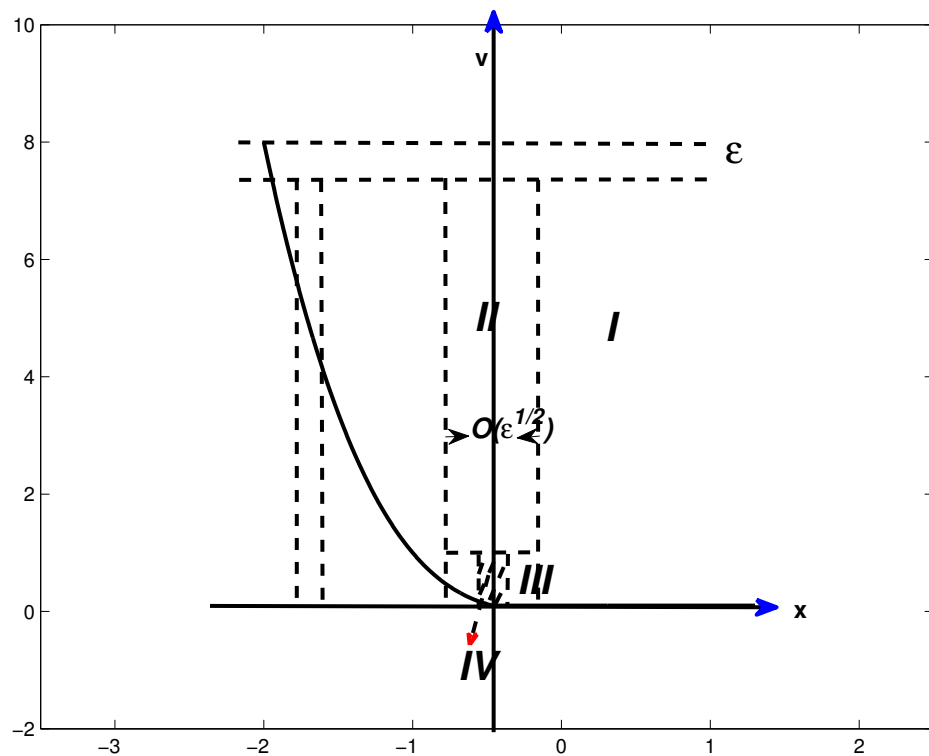


Figure 1. Boundary layer structure.

It should be remarked here that the introduction of a second stochastic process for  $v$  has indeed made the analysis much more complicated and totally different from that of the Black–Scholes’ case, as shown in Evans et al. (2002). The complexity, as well as the difference has been clearly manifested when analyzing the solution within the domain  $x \in U(0, \sqrt{\epsilon}), 0 \leq v < \infty$ . Here, the notation  $U(a, \delta)$  denotes the neighborhood of a point  $a$  with radius  $\delta > 0$ , i.e.,

$$U(a, \delta) = \{x \mid 0 \leq |x - a| < \delta\}.$$

In the classical Black–Scholes’ case, the analysis usually stops once the inner solution within the  $\mathcal{O}(\sqrt{\epsilon})$  boundary layer near  $x = 0$  has been found. Under the Heston model, however, the “inner” region  $x \in U(0, \sqrt{\epsilon})$  (relative to the “outer” Region I) needs to be further divided into another set of “inner” and “outer” regions because there exists another boundary layer in the  $v$  direction as a result of Region II still containing the singular point at the origin. This new inner region is denoted as Region III. Similarly, the same process needs to be repeated, which yields Region IV (the dashed area of Figure 1) as a boundary layer in  $x$  again. It is envisaged that this “cascade” phenomenon of sub-dividing regions is a particular feature associated with the Heston model when it is used to price American options. However, since we are only interested in an approximate solution, based on the method of matched asymptotic expansions, of  $\mathcal{O}(\epsilon)$ , it suffices to stop the sub-division at Region IV, as shall be discussed later.

With the division of the original solution domain into the above four regions, we can now prove the following lemma for the asymptotic behavior  $x_f(v, \tau)$  when  $v \notin U(0, \epsilon)$ . Note that hereafter, we shall use a new definition  $\mathcal{O}_s(\cdot)$ , i.e.,  $f(\epsilon) = \mathcal{O}_s(g(\epsilon))$ , if  $f(\epsilon) = \mathcal{O}(g(\epsilon))$ , and  $f(\epsilon) \neq o(g(\epsilon))$  for  $\epsilon \rightarrow 0$ . As shall be demonstrated, this new definition enables us to give a sharp estimate of the target functions.



**Lemma 1.** (i) For  $v = \mathcal{O}_s(1)$ , we have  $x_f(v, \tau) \notin U(0, \sqrt{\epsilon})$ . (ii) For  $v = \mathcal{O}_s(\epsilon^\beta)$ , where  $0 < \beta < 1$ , we have  $x_f(v, \tau) \in U(0, \sqrt{\epsilon})$ , and the leading-order term of  $x_f$ , say  $x_0$ , should be bounded as  $v \rightarrow 0$  and satisfy  $\lim_{v \rightarrow 0} \frac{x_0(v, \tau)}{2\sqrt{a\tau}} = -\infty$ , where  $a = \frac{v}{\sigma^2}$ .

**Proof.** (i) The first part of the lemma is proven by means of the method of proof by contradiction. Assuming that  $x_f(v, \tau) \in U(0, \sqrt{\epsilon})$ , we have  $\lim_{\epsilon \rightarrow 0} \frac{x_f(v, \tau)}{\sqrt{\epsilon}} = X_0$ , where  $X_0$  should have finite values for any  $v = \mathcal{O}_s(1)$  and  $T = \mathcal{O}(1)$ . Therefore, we can rescale  $x_f(v, \tau)$  and expand it in terms of  $\sqrt{\epsilon}$ , i.e.,

$$X_f = \frac{x_f}{\sqrt{\epsilon}} = X_0 + \sqrt{\epsilon}X_1 + \mathcal{O}(\epsilon).$$

In order to satisfy the moving boundary conditions, the leading-order term should at least satisfy

$$P_0(X_0, v, T) = \frac{\partial P_0}{\partial X}(X_0, v, T) = 0,$$

which yields

$$\frac{vT}{\sigma\sqrt{\pi}}e^{-\frac{\sigma X_0^2}{4vT}} + \frac{X_0}{2} \operatorname{erfc}\left(-\frac{\sigma X_0}{2\sqrt{vT}}\right) = 0, \tag{12}$$

$$\operatorname{erfc}\left(\frac{-\sigma X_0}{2\sqrt{vT}}\right) = 0. \tag{13}$$

However, when  $v = \mathcal{O}_s(1)$ , the only solution of (12) and (13) is  $X_0 = -\infty$ , and this is in contrast with our assumption that  $x_f(v, \tau) \in U(0, \sqrt{\epsilon})$ . Therefore, when  $v = \mathcal{O}_s(1)$ , the location of the free boundary should be outside the  $\mathcal{O}(\sqrt{\epsilon})$  layer near  $x = 0$ , and thus,  $\lim_{\epsilon \rightarrow 0} \frac{x_f(v, \tau)}{\sqrt{\epsilon}} = -\infty$ . This completes the proof.

(ii) It is straightforward to show that if  $\lim_{v \rightarrow 0} \frac{x_0(v, \tau)}{2\sqrt{a\tau}} = -\infty$ , (12) and (13) are always satisfied. On the other hand, when  $v = \mathcal{O}_s(\epsilon^\beta)$ , the leading-order term of  $x_f(v, \tau)$ , i.e.,  $x_0$ , can be bounded and satisfy  $\lim_{v \rightarrow 0} \frac{x_0(v, \tau)}{2\sqrt{a\tau}} = -\infty$  at the same time. Therefore, when  $v = \mathcal{O}_s(\epsilon^\beta)$ , where  $0 < \beta < 1$ ,  $x_f$  should be located inside  $U(0, \sqrt{\epsilon})$ . This completes the proof.  $\square$

Since the location of the free boundary  $x_f(v, \tau)$  differs with respect to  $v$ , it is much more convenient to discuss the asymptotic behavior of  $x_f(v, \tau)$  in different ranges of  $v$  separately.

**Case I.**  $v = \mathcal{O}_s(1)$

In this case,  $x_f(v, \tau)$  is located outside  $U(0, \sqrt{\epsilon})$ , which indicates that there is another boundary layer near  $x_f(v, \tau)$ . Now, we perform the local analysis in the vicinity of  $x_f$  by using the stretched variable:

$$Z = \frac{x - x_f(v, \tau)}{\epsilon}, \tag{14}$$

where  $Z = \mathcal{O}(1)$ . Substituting (14) into the governing equation contained in (7), we obtain

$$\left\{ \begin{aligned} \epsilon \frac{\partial P}{\partial T} - \frac{x_f}{\partial T} \frac{\partial P}{\partial Z} &= \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial Z^2} + \epsilon \left( q - d - \frac{v}{\sigma^2} \right) \frac{\partial P}{\partial Z} + v \left[ \epsilon^2 \frac{\partial^2 P}{\partial v^2} - 2\epsilon \frac{\partial^2 P}{\partial v \partial Z} \frac{\partial x_f}{\partial v} \right. \\ &+ \left. \frac{\partial^2 P}{\partial Z^2} \left( \frac{\partial x_f}{\partial v} \right)^2 - \epsilon \frac{\partial P}{\partial Z} \frac{\partial^2 x_f}{\partial v^2} \right] + \frac{2\kappa}{\sigma^2} (\eta - v) \left( \epsilon^2 \frac{\partial P}{\partial v} - \epsilon \frac{\partial P}{\partial Z} \frac{\partial x_f}{\partial v} \right) \\ &+ 2 \frac{\rho v}{\sigma} \left( \epsilon \frac{\partial^2 P}{\partial v \partial Z} - \frac{\partial^2 P}{\partial Z^2} \frac{\partial x_f}{\partial v} \right) + \epsilon^2 e^{\epsilon q T} (d e^{\epsilon z + x_f} - q), \\ P(0, v, T) &= 0, \\ \frac{\partial P}{\partial Z}(0, v, T) &= 0. \end{aligned} \right. \tag{15}$$

Again, an expansion in regular powers of  $\epsilon$  gives the solution of (15) as

$$P = \mathcal{O}(\epsilon^2). \tag{16}$$

In order to match the solution in Region II with the solution near  $x_f$ , we found the asymptotic behaviors of  $h_0, h_1$ , and  $h_2$  as  $\xi \rightarrow -\infty$ :

$$\begin{aligned} h_0(\xi) &= \frac{\sqrt{a}}{2\sqrt{\pi}} \frac{e^{-\xi_1^2}}{\xi_1^2} + \mathcal{O}\left(\frac{e^{-\xi_1^2}}{\xi_1^4}\right), \\ h_1(\xi) &= d - q + \mathcal{O}(\xi_1 e^{-\xi_1^2}), \\ h_2(\xi) &= 2d\sqrt{a}\xi_1 + \mathcal{O}(\xi_1^4 e^{-\xi_1^2}), \end{aligned}$$

where  $\xi_1 = \frac{\xi}{\sqrt{a}}$ .

Now, if we follow Evans et al. (2002) and match the  $P_\tau$  values of (11) and (16) by taking the limit  $x \rightarrow x_f$ , we obtain the following transcendental equations:

$$\frac{\sqrt{a}}{2\sqrt{\pi\tau}} e^{-\frac{x_f^2}{4a\tau}} + d - q = 0, \quad \text{for } d < q, \tag{17}$$

$$\frac{\sqrt{a}}{2\sqrt{\pi\tau}} e^{-\frac{x_f^2}{4a\tau}} + d\sqrt{a}x_f = 0, \quad \text{for } d = q, \tag{18}$$

which lead to the solutions of the form

$$x_f(v, \tau) = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|} \quad \text{for } d < q, \tag{19}$$

$$x_f(v, \tau) = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{1}{4\sqrt{\pi}d\tau} \right|}, \quad \text{for } d = q, \tag{20}$$

respectively, after higher-order terms are ignored.

It should be remarked that matching  $P_\tau$  values is not the unique choice. Evans et al. (2002) did it in this way without any detailed explanation. In fact, this approach may be at odds with the conventional method of matched asymptotic expansions in which  $P$ , instead of  $P_\tau$ , values should be matched in the “inner” and “outer” regions. Clearly, if a solution is obtained with  $P$  values being matched, so should the  $P_\tau$  values, provided that the  $P$  function is of sufficient smoothness. However, the converse is not true. One naturally wonders whether or not the two different approaches would lead to the same conclusion, once all the higher-order terms are ignored. Without a great deal of additional effort, it can be shown that matching the  $P$  values of (11) and (16) by taking the limit  $x \rightarrow x_f$  leads to

$$\frac{2a^{\frac{3}{2}} \tau^{\frac{3}{2}} e^{-\frac{x_f^2}{4a\tau}}}{\sqrt{\pi} x_f^{*2}} + (d - q)\tau + d\tau x_f^* = 0. \tag{21}$$

In comparison with (17) and (18), it is much more difficult to go through some additional order analysis to simplify (21), in order to analytically obtain the asymptotic behavior of  $x_f^*$  like those presented in (19) and (20). However, it can still be shown that  $x_f^*$  lies between  $x_f$  and  $-2\sqrt{a\tau}$  (see Lemma 2 below), where  $x_f$  is solved from matching with  $P_\tau$ . Hence, as  $\tau$  approaches zero, the difference between  $x_f$  and  $x_f^*$  will become smaller. It should be noted that we only need to consider the negative root of (21) because of the physical restriction that  $x_f \leq 0$ .

**Lemma 2.** For small  $\tau$ , (21) has only one negative solution  $x_f^*$ , and moreover, it satisfies

$$x_f < x_f^* < -2\sqrt{a\tau},$$

where

$$x_f = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|}, \quad \text{for } d < q \tag{22}$$

$$x_f = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{1}{4\sqrt{\pi d\tau}} \right|}, \quad \text{for } d = q. \tag{23}$$

**Proof.** Denote

$$f(x) = \frac{\sqrt{a\tau}e^{-x^2}}{2\sqrt{\pi}x^2} + (d-q)\tau + 2d\sqrt{a\tau}^{\frac{3}{2}}x. \tag{24}$$

Comparing (24) with (21), it is obvious that  $x = \frac{x_f^*}{2\sqrt{a\tau}}$ . Then, taking the first-order derivative of  $f(x)$  with respect to  $x$ , we obtain:

$$f'(x) = -e^{-x^2} \frac{\sqrt{a\tau}}{\sqrt{\pi}} \left( \frac{1}{x} + \frac{1}{x^3} \right) + 2d\sqrt{a\tau}^{\frac{3}{2}}. \tag{25}$$

Since the parameters  $a$ ,  $\tau$ , and  $d$  are all greater than zero, we have  $f'(x) > 0$  for any  $x < 0$ . Therefore,  $f(x)$  is monotonically increasing for negative  $x$  values.

On the other hand, it is straightforward to show that when  $d < q$ ,

$$f(-1) = \frac{\sqrt{a\tau}e^{-1}}{2\sqrt{\pi}} + (d-q)\tau - 2d\sqrt{a\tau}^{\frac{3}{2}},$$

$$f\left(-\sqrt{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|}\right) = \tau(q-d) \left( \frac{1}{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|} - 1 \right) - 2d\sqrt{a\tau}^{\frac{3}{2}} \sqrt{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|}.$$

Furthermore, for reasonably small  $\tau$ , we have

$$\frac{\sqrt{a\tau}e^{-1}}{2\sqrt{\pi}} > (q-d)\sqrt{\tau} + 2d\sqrt{a\tau},$$

$$\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right| > 1.$$

Therefore,  $f(-1) > 0$  and  $f\left(-\sqrt{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|}\right) < 0$ . Based on the monotonicity of  $f(x)$ , the only negative root of  $f$  should lie between  $-\sqrt{\left| \ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}} \right|}$  and  $-1$ .

Similarly, we can show that when  $q = d$ , it lies between  $-\sqrt{|\ln \frac{1}{4\sqrt{\pi d\tau}}|}$  and  $-1$ . Therefore, for small  $\tau$ , (21) has only one negative solution  $x_f^*$ , and moreover, it satisfies

$$x_f < x_f^* < -2\sqrt{a\tau}. \tag{26}$$

This completes the proof.  $\square$

On the other hand, Equation (21) can also be numerically solved, and numerical evidence indeed suggests that the difference between  $x_f^*$  and  $x_f$  is negligible. This probably explains why Evans et al. (2002) chose to match the  $P_\tau$  values in their analysis.

By substituting (19)–(20) into (11), we find that the free boundary conditions can be satisfied in the following asymptotic sense, i.e.,

$$P(x_f, v, \tau) - P_{exact}(\tilde{x}_f, v, \tau) = \mathcal{O}(\tau) \quad \text{for } d < q, \tag{27}$$

$$P(x_f, v, \tau) - P_{exact}(\tilde{x}_f, v, \tau) = \mathcal{O}\left(\tau^{\frac{3}{2}}\sqrt{\ln \frac{1}{\tau}}\right) \quad \text{for } d = q,$$

$$\frac{\partial P}{\partial x}(x_f, v, \tau) - \frac{\partial P_{exact}}{\partial x}(\tilde{x}_f, v, \tau) = \mathcal{O}(\tau^{\frac{3}{2}}) \quad \text{for } d \leq q, \tag{28}$$

where  $\tilde{x}_f$  stands for the exact free boundary.

**Case II.**  $v = \mathcal{O}(\epsilon^\beta), 0 < \beta < 1$

According to the second part of Lemma 1,  $x_f(v, \tau)$  should be located inside  $U(0, \sqrt{\epsilon})$ . Now, we assume that (11) can satisfy the conditions across the free boundary. By taking the limit  $\epsilon \rightarrow 0$ , we deduce the asymptotic behavior of  $h_0, h_1$  and  $h_2$ , respectively, as:

$$h_0(\xi) = \frac{\sqrt{a}}{2\sqrt{\pi}} \frac{e^{-\xi_1^2}}{\xi_1^2} + \mathcal{O}\left(\sqrt{a} \frac{e^{-\xi_1^2}}{\xi_1^4}\right), \tag{29}$$

$$h_1(\xi) = d - q + \mathcal{O}(\xi_1 e^{-\xi_1^2}), \tag{30}$$

$$h_2(\xi) = 2d\sqrt{a}\xi_1 + \mathcal{O}\left(\frac{\xi_1^4 e^{-\xi_1^2}}{\sqrt{a}}\right),$$

where  $\xi_1 = \frac{\xi}{\sqrt{a}}$ . It is not strange that the asymptotic behaviors of  $h_0, h_1$ , and  $h_2$  as  $\epsilon \rightarrow 0$  are quite similar to those derived in Case I, since when taking the limit  $\epsilon \rightarrow 0$ , it is equivalent to  $v \rightarrow 0$ , which is also equal to  $\xi_1 \rightarrow -\infty$  and  $a \rightarrow 0$ .

On the other hand, upon applying the free boundary conditions on (11), the leading-order term of  $x_f$  should at least satisfy

$$P_\tau(x_f, v, \tau) = \mathcal{O}(\tau),$$

which is the same as what we have obtained in Case I. It is now quite trivial to show that the leading-order term of  $x_f$  is the same as the one derived in Case I, and the expansion satisfies the free boundary conditions in almost the same asymptotic sense, i.e.,

$$P(x_f, v, \tau) - P_{exact}(\tilde{x}_f, v, \tau) = \mathcal{O}(\tau) \quad \text{for } d < q, \tag{31}$$

$$P(x_f, v, \tau) - P_{exact}(\tilde{x}_f, v, \tau) = \mathcal{O}\left(\tau^{\frac{3}{2} + \frac{\beta}{2}}\sqrt{\ln \frac{1}{\tau}}\right) \quad \text{for } d = q.$$

$$\frac{\partial P}{\partial x}(x_f, v, \tau) - \frac{\partial P_{exact}}{\partial x}(\tilde{x}_f, v, \tau) = \mathcal{O}\left(\tau^{\frac{3}{2}}\right) \quad \text{for } d \leq q. \tag{32}$$

**Case III.**  $v = \mathcal{O}(\epsilon)$

In this case, if we follow the procedure described previously, we should first determine whether  $x_f(v, \tau) \in U(0, \epsilon)$  or not, which is equivalent to analyzing whether the solution in Region IV can satisfy the free boundary conditions or not. It is straightforward to show that the governing equation for the solution in Region IV is

$$\frac{\partial Z_2}{\partial T} = \frac{V}{\sigma^2} \frac{\partial^2 Z_2}{\partial X_1^2} + (q - v) \frac{\partial Z_2}{\partial X_1} + \frac{2\rho V}{\sigma} \frac{\partial^2 Z_2}{\partial X_1 \partial V} + V \frac{\partial^2 Z_2}{\partial V^2} + \frac{2\kappa\eta}{\sigma^2} \frac{\partial Z_2}{\partial V} + d - q, \quad (33)$$

which contains all the derivatives the original equation has. Now, we are in an unfortunate situation of having to solve almost the full problem to be able to determine what is going on in Region IV. If we could solve this problem, one might wonder why it was necessary to bother with an approximation in the first place. There is no need to argue with this sentiment, and this is indeed one of the situations where the perturbation methods show some of their limitations. However, there are several remarks that should be made. Firstly, though this layer problem cannot be solved in a closed form, we may still be able to extract some useful information about the solution. Secondly, this problem can be further dealt with if we again apply the method of matched asymptotic expansions to this corner, i.e., by rescaling  $T_1 = \frac{T}{\epsilon}$ , and following almost the same procedure as demonstrated previously. Fortunately, there is no need to go through such a cumbersome analysis again. In order to demonstrate the reasons in a clear way, we adopt a new notation  $\bar{x}_f$  for the actual optimal exercise boundary in this case.

As it turns out, when  $v = \mathcal{O}(\epsilon)$ , the location of the optimal exercise boundary should be located either inside  $U(0, \epsilon)$  or outside. However, it is claimed that no matter what is the truth, the approximation of  $x_f$  derived before for  $d < q$  can still be used here. Firstly, if  $\bar{x}_f$  is located outside  $U(0, \epsilon)$ , then (33) is defined on

$$\Omega = \{-\infty < X_1 < +\infty, 0 < V < +\infty, 0 < T < +\infty\}.$$

Taking the corresponding boundary conditions into consideration, it can be identified that  $Z_2 = \epsilon Z_{21} + \mathcal{O}(\epsilon^2)$ , with  $Z_{21}$  twice continuously differentiable with respect to  $X_1$  or  $V$  on  $\Omega$ . The stretched  $x_f$ , which is derived in the previous two cases, reads

$$X_f(V, T) = \frac{x_f}{\epsilon} = \mathcal{O}(1),$$

which means that  $X_f$  is finite for any  $0 < V < +\infty$  and  $0 < T < +\infty$ . Based on the continuous property of both  $Z_{21}$  and its first-order derivative with respect to  $X_1$ , it is clear that if we adopt the approximation of  $x_f$  derived previously as the free boundary here, then the option price  $Z_2$  satisfies the free boundary conditions in the  $\mathcal{O}(\tau)$  sense, which is almost the same as the previous two cases; see (27) and (28) and (31) and (32).

Secondly, if  $\bar{x}_f$  is located inside  $U(0, \epsilon)$ , it is meaningless to derive its actual form, since  $x_f = \mathcal{O}(\epsilon)$  is already a good approximation for  $\bar{x}_f$ .

Based on Case I to Case III, it can be concluded that when  $D < r$ , (19) can be used as an approximation for small  $\tau$  and  $0 \leq v < \infty$ . Written in original variables, we obtain the leading-order term of the optimal exercise price as:

$$S_f(v, t) = K - K\sqrt{v(T_E - t)} \sqrt{\ln \frac{v}{8\pi(T_E - t)(r - D)^2}}, \quad \text{for } D < r. \quad (34)$$

For  $D = r$ , (20) is valid for small  $\tau$  and  $\epsilon \leq v < \infty$ . Therefore,

$$S_f(v, t) = K - K\sqrt{2v(T_E - t)} \sqrt{\ln \frac{1}{4\sqrt{\pi}(T_E - t)D}}, \quad \text{for } D = r. \quad (35)$$

It is quite interesting to note that the leading-order terms of the optimal exercise price are similar to the ones with constant volatility (see Evans et al. 2002), with  $\sigma^2$  being

substituted by  $v$ . One possible reason is that the moving boundary only occurs along the  $S$  direction, and no critical points appear along the  $v$  direction. Furthermore, in Case I and Case II, the impact of  $v$  is less significant, so that  $v$  becomes a parameter rather than a variable; (see (A1)–(A3)). However, the option prices are much more complicated and totally different than those with constant volatility; (see (A5), (A11), and (A15)).

#### 4. Numerical Results

In this section, we compare the estimated optimal exercise prices with those obtained from numerical simulations based on a predictor–corrector finite difference scheme (see Zhu and Chen 2011). We expect our approximation to be useful for options with tenor in the order of days and weeks.

Depicted in Figure 2 are the optimal exercise prices as a function of the time to expiry  $T - t$  with different fixed values of  $v$ . Here, the reference results refer to those obtained by using the predictor–corrector scheme on extremely fine grids, while the estimated values are calculated directly from (34). We point out that it is reasonable to use the numerical results produced by the predictor-correct scheme on extremely fine grids as benchmark solutions, because the convergence of this method was already tested by Zhu and Chen (2011). From Figure 2, it is clear that, for reasonably short maturities, our approximation agrees well with those numerical results, as one would have expected. To show this in a more quantitative base, we further calculate the accuracy of our approximation, which is shown in Table 1. Note that model parameters for both Figure 2 and Table 1. are  $r = 0.1$ ,  $\sigma = 0.2$ ,  $\rho = 0.1$ ,  $\eta = 0.16$ , and  $\kappa = 1.5$  and the strike price  $K = \$10.0$ . There is no preference in choosing the parameters, and they are all picked randomly. Furthermore, since the optimal exercise price does not depend on the spot price, it is unnecessary to provide the value of  $S$ . In Table 1, the “rel-err” refers to the relative error for fixed values of  $v$ , which is defined as

$$rel - err = \frac{|S_f(v, \tau) - \tilde{S}_f(v, \tau)|}{|S_f(v, \tau)|},$$

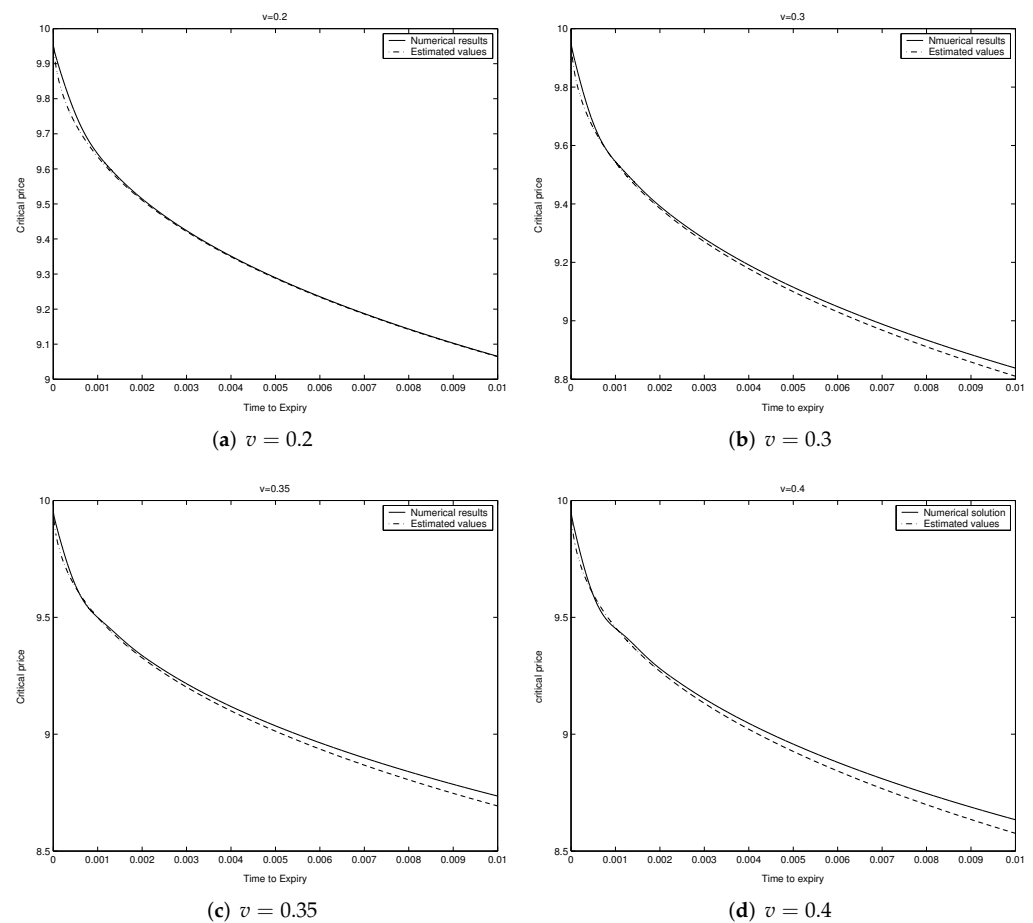
and the “overall-err” refers to the overall error across the whole computational domain, defined as

$$overall - err = \frac{\|S_f(:, \tau) - \tilde{S}_f(:, \tau)\|_2}{\|S_f(:, \tau)\|_2},$$

and for all  $v$ . Here,  $\|\cdot\|_2$  denotes the  $L_2$ -norm for vectors, and  $S_f$  and  $\tilde{S}_f$  stand for the numerical solutions and the estimated values, respectively. From Table 1, it can be seen that for the maturities in the order of days and weeks, the overall errors are all very small, which demonstrates that our estimation is quite reasonable.

Table 1. Report on the relative error and the overall error.

Time to Expiry (Year)	Rel-Err at $v = 0.05$	Rel-Err at $v = 0.1$	Rel-Err at $v = 0.2$	Rel-Err at $v = 0.25$
0.002	$3.445 \times 10^{-2}\%$	$2.2791 \times 10^{-2}\%$	$4.1927 \times 10^{-2}\%$	$6.1511 \times 10^{-2}\%$
0.005	0.16%	0.11%	$1.727 \times 10^{-2}\%$	$9.3112 \times 10^{-2}\%$
0.01	0.36%	0.26%	$1.0267 \times 10^{-2}\%$	0.16%
Time to Expiry (Year)	Rel-Err at $v = 0.5$	Rel-Err at $v = 0.6$	Rel-Err at $v = 0.7$	Overall-Err
0.002	0.18%	0.26%	0.47%	0.45%
0.005	0.54%	0.73%	0.94%	0.77%
0.01	1.05%	1.43%	1.84%	1.49%



**Figure 2.** Comparison of the numerical results and the estimated values.

### 5. Conclusions

In this paper, we considered the near-expiry behavior of the optimal exercise price for American put options on a dividend-paying underlying with stochastic volatility. Based on the method of matched asymptotic expansions, explicit analytical expressions of the optimal exercise price and the values of American puts were derived. It turns out that the option prices are quite different from the corresponding Black–Scholes’ case, while the leading-order term of the optimal exercise price remains almost the same as the constant volatility case if the spot volatility given the same value as the constant volatility appearing in the Black–Scholes model. Numerical experiments suggest that the current approximation is reasonably accurate, and thus very useful to price American options close to expiry, say for example, with tenor in the order of days and weeks. Therefore, our asymptotic analysis complements purely numerical analyses, for which dealing with the singularity of the optimal exercise price of American puts near expiry always poses a conflict of balance between numerical accuracy and efficiency. It has of course also enhanced our understanding of the asymptotic behavior of the optimal exercise price of American puts near expiry, which is an important aspect of this very fundamental problem in modern quantitative finance.

**Author Contributions:** Conceptualization, S.-P.Z. and W.C.; methodology, S.-P.Z.; software, W.C.; validation, S.-P.Z. and W.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** The Ministry of Education of Humanities and Social Science Project of China No. 21YJAZH005; the Research Base of Humanities and Social Sciences outside Jiangsu Universities Research Center of Southern Jiangsu Capital Market (2017ZSJD020).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** This work is supported by the Ministry of Education of Humanities and Social Science Project of China No. 21YJAZH005 and the Research Base of Humanities and Social Sciences outside Jiangsu Universities Research Center of Southern Jiangsu Capital Market (2017ZSJD020).

**Conflicts of Interest:** The authors declare no conflict of interest.

### Appendix A

The PDE systems for  $P_0$ ,  $P_1$ , and  $P_2$  are:

$$\left\{ \begin{array}{l} \frac{\partial P_0}{\partial T} = \frac{v}{\sigma^2} \frac{\partial^2 P_0}{\partial X^2}, \\ P_0(X, v, 0) = \max(X, 0), \\ \lim_{X \rightarrow \infty} P_0(X, v, T) = X, \\ \lim_{X \rightarrow -\infty} P_0(X, v, T) = 0, \end{array} \right. \tag{A1}$$

$$\left\{ \begin{array}{l} \frac{\partial P_1}{\partial T} = \frac{v}{\sigma^2} \frac{\partial^2 P_1}{\partial X^2} + (q - d - \frac{v}{\sigma^2}) \frac{\partial P_0}{\partial X} + \frac{2\rho v}{\sigma} \frac{\partial^2 P_0}{\partial X \partial v} + d - q, \\ P_1(X, v, 0) = \max(\frac{1}{2} X^2, 0), \\ \lim_{X \rightarrow \infty} P_1(X, v, T) = \frac{1}{2} X^2, \\ \lim_{X \rightarrow -\infty} \frac{\partial P_1}{\partial X}(X, v, T) = 0, \end{array} \right. \tag{A2}$$

$$\left\{ \begin{array}{l} \frac{\partial P_2}{\partial T} = \frac{v}{\sigma^2} \frac{\partial^2 P_2}{\partial X^2} + (q - d - \frac{v}{\sigma^2}) \frac{\partial P_1}{\partial X} + \frac{2\rho v}{\sigma} \frac{\partial^2 P_1}{\partial X \partial v} \\ \quad + v \frac{\partial^2 P_0}{\partial v^2} + \frac{2\kappa}{\sigma^2} (\eta - v) \frac{\partial P_0}{\partial v} + dX, \\ P_2(X, v, 0) = \max(\frac{1}{6} X^3, 0), \\ \lim_{X \rightarrow \infty} P_2(X, v, T) = \frac{1}{6} X^3 + qXT, \\ \lim_{X \rightarrow -\infty} \frac{\partial^2 P_2}{\partial X^2}(X, v, T) = 0. \end{array} \right. \tag{A3}$$

To find the solution of the PDE System (A1), we assume it can be written as

$$P_0(X, v, T) = \sqrt{T} h_0(\xi; a), \tag{A4}$$

where  $\xi = \frac{X}{2\sqrt{T}}$  and  $a = \frac{v}{\sigma^2}$ . By substituting (A4) into (A1), we obtain the ODE system for  $h_0(\xi; a)$  as

$$\left\{ \begin{array}{l} ah_0''(\xi; a) + 2\xi h_0'(\xi; a) - 2h_0(\xi; a) = 0, \\ \lim_{\xi \rightarrow \infty} h_0(\xi; a) = 2\xi, \\ \lim_{\xi \rightarrow -\infty} h_0(\xi; a) = 0. \end{array} \right.$$

The analytical solution of this ODE system can be readily found as

$$h_0(\xi; a) = \frac{\sqrt{a}}{\sqrt{\pi}} e^{-\frac{\xi^2}{a}} + \xi \operatorname{erfc}\left(-\frac{\xi}{\sqrt{a}}\right). \tag{A5}$$



Similarly, by assuming that the solution of PDE System (A2) is in the form of

$$P_1(X, v, T) = Th_1(\xi; a) = Ta\tilde{h}_1(\xi_1),$$

where  $\xi_1 = \frac{\xi}{\sqrt{a}}$ , we have

$$\left\{ \begin{aligned} \tilde{h}_1''(\xi_1) + 2\xi_1\tilde{h}_1'(\xi_1) - 4\tilde{h}_1(\xi_1) &= 2\frac{(a-q+d)}{a} \operatorname{erfc}(-\xi_1) \\ &+ \frac{4\rho}{\sigma\sqrt{\pi a}}\xi_1 e^{-\xi_1^2} + 4\frac{(q-d)}{a}, \\ \lim_{\xi_1 \rightarrow \infty} \tilde{h}_1(\xi_1) &= 2\xi_1^2, \\ \lim_{\xi_1 \rightarrow -\infty} \tilde{h}_1'(\xi_1) &= 0. \end{aligned} \right. \tag{A6}$$

Suppose that the solution  $\tilde{h}_1(\xi_1)$  has the following structure:

$$h_1(\xi) = f(\xi_1)e^{-\xi_1^2} + g(\xi_1) \int_{-\infty}^{\xi_1} e^{-t^2} dt + m(\xi_1), \tag{A7}$$

where  $f(\xi_1)$ ,  $g(\xi_1)$ , and  $m(\xi_1)$  are polynomials in  $\xi_1$ . By substituting (A7) into (A6), we obtain

$$\begin{aligned} (f'' - 2\xi_1 f' + 2g' - 6f)e^{-\xi_1^2} + (g'' + 2\xi_1 g' - 4g) \int_{-\infty}^{\xi_1} e^{-t^2} dt + m'' + 2\xi_1 m' - 4m \\ = \frac{4(a-q+d)}{a\sqrt{\pi}} \int_{-\infty}^{\xi_1} e^{-t^2} dt + \frac{4\rho}{\sigma a\sqrt{\pi}}\xi_1 e^{-\xi_1^2} + 4\frac{(q-d)}{a}, \end{aligned}$$

which, combined with the boundary conditions at  $\xi_1 = \pm\infty$ , yields

$$\left\{ \begin{aligned} m'' + 2\xi_1 m' - 4m &= \frac{4(q-d)}{a}, \\ \lim_{\xi_1 \rightarrow -\infty} m' &= 0, \end{aligned} \right. \tag{A8}$$

$$\left\{ \begin{aligned} g'' + 2\xi_1 g' - 4g &= \frac{4(a-q+d)}{a\sqrt{\pi}}, \\ \lim_{\xi_1 \rightarrow \infty} \sqrt{\pi}g + m &= 2\xi_1^2, \end{aligned} \right. \tag{A9}$$

$$f'' - 2\xi_1 f' + 2g' - 6f = \frac{4\rho}{\sigma\sqrt{\pi a}}\xi_1. \tag{A10}$$

The polynomial solutions of (A8)–(A10) can be readily found as:

$$\begin{aligned} m(\xi_1) &= \frac{d-q}{a}, \\ g(\xi_1) &= \frac{2}{\sqrt{\pi}}\xi_1^2 + \frac{q-d}{\sqrt{a\pi}}, \\ f(\xi_1) &= \frac{1}{\sqrt{\pi}}\left(1 - \frac{\rho}{2\sigma a}\right)\xi_1. \end{aligned}$$

Therefore,

$$h_1(\xi; a) = \frac{1}{\sqrt{\pi}}\left(a - \frac{\rho}{2\sigma}\right)\frac{\xi}{\sqrt{a}}e^{-\frac{\xi^2}{a}} + \left(\frac{2}{\sqrt{\pi}}\xi^2 + \frac{q-d}{\sqrt{\pi}}\right) \int_{-\infty}^{\frac{\xi}{\sqrt{a}}} e^{-t^2} dt + d - q. \tag{A11}$$

By means of the above solution technique, it is not hard to find the solution of (A3), though the expression is quite complicated. By assuming

$$P_2(X, v, T) = T^{\frac{3}{2}}h_2(\xi; a) = T^{\frac{3}{2}}a^{\frac{3}{2}}\tilde{h}_2(\xi_1),$$

and substituting it into (A3), we obtain the ODE system for  $\tilde{h}_2(\xi_1)$  as

$$\left\{ \begin{aligned} \tilde{h}_2''(\xi_1) + 2\xi_1\tilde{h}_2'(\xi_1) - 6\tilde{h}_2(\xi_1) &= (C\xi_1^4 + A\xi_1^2 + B)e^{-\xi_1^2} + \frac{8(a+d-q)}{a\sqrt{\pi}} \int_{-\infty}^{\xi_1} e^{-t^2} dt - \frac{8d}{a}\xi_1, \\ \lim_{\xi_1 \rightarrow \infty} \tilde{h}_2(\xi_1) &= \frac{4}{3}\xi_1^3 + \frac{2q}{a}\xi_1, \\ \lim_{\xi_1 \rightarrow -\infty} \tilde{h}_2''(\xi) &= 0, \end{aligned} \right.$$

where

$$\begin{aligned} A &= \frac{2(a-q+d)\rho}{a^2\sigma\sqrt{\pi}} - \frac{\rho}{4\sigma\sqrt{\pi}a^2} \left( -\frac{32\rho}{\sigma} - 16a + 16q - 16d \right) - \frac{2}{\sigma^2\sqrt{\pi}a^2}, \\ B &= \frac{2(a-q+d)}{a^2} \left( -\frac{\rho}{2\sigma\sqrt{\pi}} + \frac{a}{\sqrt{\pi}} + \frac{q-d}{\sqrt{\pi}} \right) - \frac{\rho}{4\sigma\sqrt{\pi}a^2} \left( 8a + \frac{4\rho}{\sigma} - 8q + 8d \right) \\ &\quad + \frac{1}{\sigma^2\sqrt{\pi}a^2} - \frac{4\kappa}{\sigma^2\sqrt{\pi}a^2} \left( \frac{\eta}{\sigma^2} - a \right), \\ C &= -\frac{4\rho^2}{\sigma^2\sqrt{\pi}a^2}. \end{aligned}$$

Suppose  $\tilde{h}_2(\xi_1)$  can be written as

$$\tilde{h}_2(\xi_1) = f(\xi_1)e^{-\xi_1^2} + g(\xi_1) \int_{-\infty}^{\xi_1} e^{-t^2} dt + m(\xi_1),$$

where  $f(\xi_1)$ ,  $g(\xi_1)$ , and  $m(\xi_1)$  are polynomials in  $\xi_1$ . By using the same procedure as in deriving  $h_1$ , the systems for  $f(\xi_1)$ ,  $g(\xi_1)$ , and  $m(\xi_1)$  can be easily found as

$$\left\{ \begin{aligned} m'' + 2\xi_1m' - 6m &= -8\frac{d}{a}\xi_1, \\ \lim_{\xi_1 \rightarrow -\infty} m'' &= 0, \end{aligned} \right. \tag{A12}$$

$$\left\{ \begin{aligned} g'' + 2\xi_1g' - 6g &= \frac{8(a+d-q)}{a\sqrt{\pi}}\xi_1, \\ \lim_{\xi_1 \rightarrow \infty} \sqrt{\pi}g + m &= \frac{4\xi_1^2}{3} + 2\frac{q}{a}\xi_1, \end{aligned} \right. \tag{A13}$$

$$f'' - 2\xi_1f' - 8f + 2g' = C\xi_1^4 + A\xi_1^2 + B. \tag{A14}$$

The polynomial solutions of (A12) and (A13) are:

$$\begin{aligned} m(\xi_1) &= 2\frac{d}{a}\xi_1, \\ g(\xi_1) &= \frac{4\xi_1^3}{3\sqrt{\pi}} + \frac{2(q-d)\xi_1}{a\sqrt{\pi}}, \\ f(\xi_1) &= -\frac{C}{16}\xi_1^4 - \left( \frac{C}{16} + \frac{\tilde{A}}{12} \right)\xi_1^2 - \frac{C}{64} - \frac{\tilde{A}}{48} - \frac{\tilde{B}}{8}, \end{aligned}$$

where

$$\begin{aligned} \tilde{A} &= A - \frac{8}{\sqrt{\pi}}, \\ \tilde{B} &= B - \frac{4(q-d)}{a\sqrt{\pi}}. \end{aligned}$$

Therefore,

$$\begin{aligned} h_2(\xi; a) &= a^{\frac{3}{2}} \left[ -\frac{C}{16} \frac{\xi^4}{a^2} - \left( \frac{C}{16} + \frac{\tilde{A}}{12} \right) \frac{\xi^2}{a} - \frac{C}{64} - \frac{\tilde{A}}{48} - \frac{B}{8} \right] e^{-\frac{\xi^2}{a}} \\ &\quad + a^{\frac{3}{2}} \left( \frac{4\xi^3}{3\sqrt{\pi}a^{\frac{3}{2}}} + \frac{2(q-d)\xi}{\sqrt{\pi}a^{\frac{3}{2}}} \right) \int_{-\infty}^{\frac{\xi}{\sqrt{a}}} e^{-t^2} dt + 2d\xi. \end{aligned} \tag{A15}$$

### Appendix B

The PDE systems for  $Z_{11}$ ,  $Z_{12}$ , and  $Z_{13}$  are:

$$\left\{ \begin{aligned} \frac{\partial Z_{11}}{\partial T} &= V \frac{\partial^2 Z_{11}}{\partial V^2} + \frac{2\kappa}{\eta} \frac{\partial Z_{11}}{\partial V}, \\ Z_{11}(X, V, 0) &= \max(X, 0), \\ \lim_{V \rightarrow \infty} Z_{11}(X, v, T) &= \max(X, 0), \\ \lim_{V \rightarrow 0} P_0(X, v, T) &= X, \quad \text{for } X > 0, \end{aligned} \right. \tag{A16}$$

$$\left\{ \begin{aligned} \frac{\partial Z_{12}}{\partial T} &= V \frac{\partial^2 Z_{12}}{\partial V^2} + \frac{2\kappa}{\eta} \frac{\partial Z_{12}}{\partial V} + (q-d) \frac{\partial Z_{11}}{\partial X} + \frac{2\rho V}{\sigma} \frac{\partial^2 Z_{11}}{\partial X \partial V} + d - q, \\ Z_{12}(X, v, 0) &= \max\left(\frac{1}{2}X^2, 0\right), \\ \lim_{V \rightarrow \infty} Z_{12}(X, v, T) &= \frac{1}{2}X^2, \quad \text{for } X > 0, \\ &\quad (d-q)T, \quad \text{for } X < 0, \\ \lim_{V \rightarrow 0} Z_{12} &= \frac{1}{2}X^2, \quad \text{for } X > 0, \end{aligned} \right. \tag{A17}$$

$$\left\{ \begin{aligned} \frac{\partial Z_{13}}{\partial T} &= V \frac{\partial^2 Z_{13}}{\partial V^2} + \frac{2\kappa}{\eta} \frac{\partial Z_{13}}{\partial V} + (q-d) \frac{\partial Z_{12}}{\partial X} + \frac{2\rho V}{\sigma} \frac{\partial^2 Z_{12}}{\partial X \partial V} \\ &\quad + \frac{V}{\sigma^2} \frac{\partial^2 Z_{11}}{\partial X^2} - \frac{2\kappa V}{\sigma^2} \frac{\partial Z_{11}}{\partial V} + dX, \\ Z_{13}(X, v, 0) &= \max\left(\frac{1}{6}X^3, 0\right), \\ \lim_{V \rightarrow \infty} Z_{13}(X, v, T) &= qTX + \frac{1}{6}X^3, \quad \text{for } X > 0, \\ &\quad dTX, \quad \text{for } X < 0, \\ \lim_{V \rightarrow 0} Z_{13}(X, v, T) &= qTX + \frac{1}{6}X^3, \quad \text{for } X > 0. \end{aligned} \right. \tag{A18}$$

It is straightforward to derive the solutions of the above PDE systems, i.e.,

$$Z_{11}(X, V, T) = \begin{cases} X, & \text{for } X > 0, \\ 0, & \text{for } X < 0. \end{cases}$$

$$Z_{12}(X, V, T) = \begin{cases} \frac{X^2}{2}, & \text{for } X > 0, \\ (d - q)T, & \text{for } X < 0. \end{cases}$$

$$Z_{13}(X, V, T) = \begin{cases} dTX + \frac{1}{6}X^3, & \text{for } X > 0, \\ dTX, & \text{for } X < 0. \end{cases}$$

## References

- Altmyer, Martin, and Andreas Neuenkirch. 2016. Discretizing the Heston model: An analysis of the weak convergence rate. *IMA Journal of Numerical Analysis* 37: 1930–60.
- Bakshi, Gurdip, Charles Cao, and Zhiwu Chen. 1997. Empirical performance of alternative option pricing models. *Journal of Finance* 52: 2003–49. [CrossRef]
- Ball, Clifford A., and Antonio Roma. 1994. Stochastic volatility option pricing. *Journal of Financial and Quantitative Analysis* 29: 589–687. [CrossRef]
- Barles, Guy, Julien Burdeau, Marc Romano, and Nicolas Samsøen. 1995. Critical stock price near expiration. *Mathematical Finance* 5: 77–95. [CrossRef]
- CBOE. 2022. CBOE's Volatility Index. Available online: [https://www.gurufocus.com/economic\\_indicators/234/cboe-volatility-index-the-vix](https://www.gurufocus.com/economic_indicators/234/cboe-volatility-index-the-vix) (accessed on 1 March 2022).
- Chen, Wen-Ting, and Song-Ping Zhu. 2009. Optimal exercise price of American options near expiry. *The ANZIAM Journal* 51: 145–61. [CrossRef]
- Chen, Xinfu, and John Chadam. 2007. A mathematical analysis of the optimal exercise boundary for American put options. *SIAM Journal on Mathematical Analysis* 38: 1613–41. [CrossRef]
- Dragulescu, Adrian A., and Victor M. Yakovenko. 2002. Probability distribution of returns in the Heston model with stochastic volatility. *Quantitative Finance* 2: 443–53. [CrossRef]
- Evans, Jonathan D., Rachel Kuske, and Joseph B. Keller. 2002. American options on assets with dividends near expiry. *Mathematical Finance* 12: 219–37. [CrossRef]
- Feller, William. 1951. Two singular diffusion problems. *Annals of Mathematics* 54: 173–82. [CrossRef]
- Fouque, Jean-Pierre, George Papanicolaou, and K. Ronnie Sircar. 2000. *Derivatives in Financial Markets with Stochastic Volatility*. Cambridge: Cambridge University Press.
- Friedman, Avner. 2008. *Partial Differential Equations*. New York: Dover Publications.
- Heston, Steven L. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6: 327–43. [CrossRef]
- Hull, John, and Alan White. 1987. The pricing of options on assets with stochastic volatilities. *Journal of Finance* 42: 281–300. [CrossRef]
- Malham, Simon JA, Jiaqi Shen, and Anke Wiese. 2021. Series expansions and direct inversions for the Heston model. *SIAM Journal on Financial Mathematics* 12: 487–549. [CrossRef]
- Mickel, Annalena, and Andreas Neuenkirch. 2021. The weak convergence rate of two semi-exact discretization schemes for the Heston model. *Risks* 9: 23. [CrossRef]
- Pan, Jun. 2002. The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study. *Journal of Financial Economics* 63: 3–50. [CrossRef]
- Qin, Cong, Xinfu Chen, Xin Lai, and Wanghui Yu. 2019. Asymptotic behavior of optimal exercise strategy for a small number of executive stock options. *Journal of Mathematical Analysis and Applications* 472: 1253–76. [CrossRef]
- Smith, Donald R. 2009. *Singular-Perturbation Theory: An Introduction with Applications*. Cambridge: Cambridge University Press.
- Tompkins, Robert G. 2001. Stock Index Futures: Stochastic Volatility Models and Smiles. *Journal of Future Markets* 21: 43–78. [CrossRef]
- Yang, Zhaoqiang. 2019. Asymptotic behavior of American lookback options exercise boundary based on random walks of order 2 model. *Journal of Ningxia University (Natural Science Edition)* 21: 33–58.
- Zhu, Song-Ping. 2006. An exact and explicit solution for the valuation of American put options. *Quantitative Finance* 6: 229–42. [CrossRef]
- Zhu, Song-Ping, and Wen-Ting Chen. 2011. A predictor–corrector scheme based on ADI method for pricing American puts with stochastic volatility. *Computers and Mathematics with Applications* 62: 1–26. [CrossRef]