SIAM J. APPL. MATH.
Vol. 64, No. 6, pp. 2034–2049

CLOSED-FORM SOLUTIONS FOR PERPETUAL AMERICAN PUT OPTIONS WITH REGIME SWITCHING*

X. GUO† AND Q. ZHANG‡

Abstract. This paper studies an optimal stopping time problem for pricing perpetual American put options in a regime switching model. An explicit optimal stopping rule and the corresponding value function in a closed form are obtained using the "modified smooth fit" technique. The solution is then compared with the numerical results obtained via a dynamic programming approach and also with a two-point boundary-value differential equation (TPBVDE) method.

Key words. Markov chain, optimal stopping time, American options, regime switching, modified smooth fit principle

AMS subject classifications. 90A09, 60J27

DOI. 10.1137/S0036139903426083

1. Introduction.

Given a probability space $(\Omega, \mathcal{F}, P)$, consider a process $X(t)$ which satisfies (in a strong sense) a stochastic differential equation of the following form:

$$dX(t) = X(t)\mu_{\epsilon(t)}dt + X(t)\sigma_{\epsilon(t)}dW(t), \qquad X(0) = x, \tag{1}$$

where $\epsilon(t) \in \{1, \dots, S\}$ is a finite-state continuous-time Markov chain and $W(t)$ is a standard Wiener process. Here $\epsilon(t)$ and $W(t)$ are defined on $(\Omega, \mathcal{F}, P)$ and are independent. Moreover, for a given $\epsilon(t) = i$, μ_i and σ_i $(i = 1, \dots, S)$ are constants and known.

The $X(t)$ governed by (1) is generally referred to as a process with "regime switching (or shifts)" or "a Markov modulated (geometric) Brownian motion." There is a substantial body of literature on this type of model studied from different perspectives. See, for instance, Di Masi, Kabanov, and Runggaldier [3] for mean variance hedging issues; Guo [5, 7] for closed-form solutions for pricing European and perpetual lookback options; Yao, Zhang, and Zhou [23] for numerical algorithms for computing European stock options; Zhang [24] for suboptimal selling rules for investors; and Zhang and Yin [25] for portfolio optimization problems.

In light of the celebrated Black–Scholes geometric Brownian motion model (see Black and Scholes [1] and Samuelson [20]), which corresponds to a special case of (1) with $\mu_1 = \cdots = \mu_S$ and $\sigma_1 = \cdots = \sigma_S$, the primary motivation for the incorporation of the Markov chain $\epsilon(t)$ is the conviction that various economic factors (e.g., interest rates, quarterly GDP) and general information (e.g., corporate news releases, quarterly earnings reports) could be major catalysts for stock fluctuations. In addition, a finite-state Markov chain has been proved to be simple yet rich enough to characterize the uncertainty in many discrete events. These convictions have been further substantiated by numerical studies: Yao, Zhang, and Zhou [23] showed that the infamous "volatility smile" can be created with a Markov chain of a single jump, instead of the more complicated stochastic volatility model by Renault and Touzi [17].

*Received by the editors April 13, 2003; accepted for publication (in revised form) February 10, 2004; published electronically August 19, 2004.
http://www.siam.org/journals/siap/64-6/42608.html

†School of ORIE, Cornell University, Ithaca, NY 14853 (xinguo@orie.cornell.edu).
‡Department of Mathematics, University of Georgia, Athens, GA 30602 (qingz@math.uga.edu).

Our results. In this paper we consider an optimal stopping problem that arises in pricing American put options, in the framework of this regime switching model. An American option is a derivative that gives its holder the option but not the obligation of exercising a share of stock at his/her choice of time τ $(T \geq \tau \geq 0)$, with a payoff of $(K - X_\tau)^+ = \max(0, K - X_\tau)$. Here, T is the *expiration date* and K is the *strike price*. It is well known that under a risk-neutral measure, the value (or the price) of this option is the expected discounted value of its future cash flow. (For more details, readers are referred to Duffie [4] and the references therein for risk-neutral option pricing for general models, to Guo [6] for the regime switching models, and to Karatzas [10] for the mathematical formulation of the American option pricing problem in the context of optimal stopping problems.) In particular, when $T = \infty$, the option becomes perpetual, and our optimal stopping problem becomes the evaluation of

$$V^*(x,i) = \sup_{0 \leq \tau \leq \infty} E[e^{-r\tau}(K - X(\tau))^+ \mid X(0) = x, \epsilon(0) = i]. \tag{2}$$

Here, $r > 0$ is the discounted factor, and τ is an $\mathcal{F}_t = \sigma\{(W(s), \epsilon(s)) \mid s \leq t\}$-stopping time.

We derive an optimal stopping rule for (2) and its corresponding value functions for $S = 2$ (see Remark 3.5). We show that the optimal stopping times are of threshold type, with the technique of modified smooth fit. The main ingredient of the optimality proof is Dynkin's formula.

It is worth mentioning that a special case of this problem with no switching (i.e., $\mu_1 = \mu_2, \sigma_1 = \sigma_2$) was solved by McKean [14], and it is referred to in what follows as "the McKean problem." His result is the earliest instance in which optimal stopping problems were related to option pricings. See also Jacka [9] and Robbins, Sigmund, and Chow [19] for related literature on optimal stopping.

Organization. In section 2, we provide a detailed derivation of the closed-form solution to (2). The optimality proof is given in section 3. In section 4, we numerically compare the closed-form solution with numerical results derived from other previous approaches, namely the dynamic programming approach (see Guo [7]) and the TPB-VDE (two-point boundary-value differential equation) method (see Zhang [24]). The paper concludes with additional discussion and open problems in section 5.

2. The derivation of solutions.

Given (1), we will study problem (2) with a two-state Markov chain (see Remark 3.5) for the general case K. Without loss of generality, we assume that $\sigma_1 \neq \sigma_2$ (see Remark 3.1) and that the Markov chain has a generator of the form

$$\begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}, \tag{3}$$

with $\lambda_1, \lambda_2 > 0$.

Recall that when there is no regime switching, this problem corresponds to a McKean problem [14] for which there exists a threshold x^* such that the optimal stopping rule is $\tau^* = \inf\{t > 0 : X(t) \notin (x^*, \infty)\}$, and the corresponding value function

$$\begin{aligned} V^*(x) &= \sup_{0 \leq \tau \leq \infty} E[e^{-r\tau}(K - X(\tau))^+ | X(0) = x] \\ &= E[e^{-r\tau^*}(K - X(\tau^*))^+ | X(0) = x] \end{aligned}$$

is given by

$$V^*(x) = \begin{cases} (K-x^*)(x/x^*)^\gamma & \text{if } x > x^*, \\ K-x & \text{if } x \le x^*. \end{cases}$$

Now, with a two-state Markov chain and with $\sigma_1 \neq \sigma_2$, it is easy to see that $(X(t), \epsilon(t))$ is a joint Markov process (see Guo [7]). Therefore, it is natural to conjecture that the optimal stopping rule is also of threshold type, except that the threshold should vary depending on the state $\epsilon(t)$. In other words, we expect the existence of two thresholds $x_1, x_2 \le K$, so that the optimal stopping rule is given as

$$\tau^* = \inf\{t \ge 0 \ \mid \ (X(t), \epsilon(t)) \notin D\},$$

where

$$D = \{(x,i) \ \mid \ V^*(x,i) > (K-x)^+\}.$$

The set D is referred to as the *continuation region.* Using τ^*, the corresponding value functions are

$$V^*(x,i) = E[e^{-r\tau^*}(K - X(\tau^*))^+ \ \mid \ X(0) = x, \epsilon(0) = i]. \tag{4}$$

We consider the case when D can be represented by two threshold levels x_1 and x_2, i.e.,

$$D = \{(x,1) \ \mid \ x \in (x_1, \infty)\} \cup \{(x,2) \ \mid \ x \in (x_2, \infty)\}.$$

Notice that x_1 and x_2 should depend on r, K, $\mu_i, \sigma_i, \lambda_i$. For any x_1 and x_2, there are only three possibilities, $x_1 < x_2$, $x_1 > x_2$, and $x_1 = x_2$. In the next sections we discuss each of these cases and derive the values of these thresholds x_i as well as the corresponding value functions (denoted as $V_i(x)$) obtained from exercising this type of stopping rule. We will then prove the optimality of these value functions, i.e., $V^*(x,i) = V_i(x)$, in Theorem 3.1.

2.1. Case 1: $x_1 < x_2 \le K$. At any given time t, if $\epsilon(t) = 1$ and $X(t) \le x_1$, then one should stop immediately and obtain a payoff of $(K - X(t))^+$; this follows from the definition of x_1 and x_2. However, if $X(t) \le x_1$ with $\epsilon(t) = 2$, it is not optimal to stop until $X(t) \le x_2$. In view of Ito's differential rule, this is translated into a set of differential equations. For $x \in [x_1, x_2]$, we have

$$\begin{cases} (r+\lambda_1)V_1(x) &=& x\mu_1 V_1'(x) + \dfrac{1}{2}x^2\sigma_1^2 V_1''(x) + \lambda_1(K-x), \\ V_2(x) &=& K-x; \end{cases} \tag{5}$$

for $x \in [x_2, \infty)$,

$$\begin{cases} (r+\lambda_1)V_1(x) &=& x\mu_1 V_1'(x) + \dfrac{1}{2}x^2\sigma_1^2 V_1''(x) + \lambda_1 V_2(x), \\ (r+\lambda_2)V_2(x) &=& x\mu_2 V_2'(x) + \dfrac{1}{2}x^2\sigma_2^2 V_2''(x) + \lambda_2 V_1(x); \end{cases} \tag{6}$$

and for $x \in [0, x_1]$,

$$V_1(x) = V_2(x) = K - x. \tag{7}$$

Now, (6) has an associated characteristic function

$$g_1(\beta) g_2(\beta) = \lambda_1 \lambda_2, \tag{8}$$

where

$$g_1(\beta) = \lambda_1 + r - \left(\mu_1 - \frac{1}{2}\sigma_1^2\right)\beta - \frac{1}{2}\sigma_1^2 \beta^2,$$
$$g_2(\beta) = \lambda_2 + r - \left(\mu_2 - \frac{1}{2}\sigma_2^2\right)\beta - \frac{1}{2}\sigma_2^2 \beta^2.$$

Moreover, this characteristic function has four distinct roots $\beta_1 < \beta_2 < 0 < \beta_3 < \beta_4$ (see Guo [7]), such that the general form of the solution to (6) is given by

$$\begin{aligned} V_1(x) &= \sum_{i=1}^{4} A_i x^{\beta_i}, \\ V_2(x) &= \sum_{i=1}^{4} B_i x^{\beta_i}, \end{aligned}$$

with $B_i = l_i A_i$ and $l_i = l(\beta_i) = \frac{g_1(\beta_i)}{\lambda_1} = \frac{\lambda_2}{g_2(\beta_i)}$.

Note that when $x \to \infty$, $V_1(x)$ and $V_2(x)$ are bounded. Thus, the positive powers of x should be eliminated so that

$$\begin{aligned} V_1(x) &= A_1 x^{\beta_1} + A_2 x^{\beta_2}, \\ V_2(x) &= B_1 x^{\beta_1} + B_2 x^{\beta_2}. \end{aligned} \tag{9}$$

Next, we turn our attention to (5). The first equation is an inhomogeneous equation whose solution can be written as

$$V_1(x) = C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \phi(x), \tag{10}$$

where $\phi(x)$ is a special solution and γ_1, γ_2 are the two real roots of

$$\mu_1 \gamma + \frac{1}{2}\sigma_1^2 \gamma(\gamma - 1) = r + \lambda_1.$$

In particular, when $r + \lambda_1 - \mu_1 \neq 0$, one can choose

$$\phi(x) = \frac{\lambda_1 K}{r + \lambda_1} - \frac{\lambda_1 x}{r + \lambda_1 - \mu_1}. \tag{11}$$

Now, we want to solve for A_1, A_2, C_1, C_2, x_1, and x_2. To this end, appropriate boundary conditions are needed. Applying the smooth fit at x_2, conditions $V_2(x+) = V_2(x-)$ and $V_2'(x+) = V_2'(x-)$ suggest

$$\left\{ \begin{aligned} l_1 A_1 x_2^{\beta_1} + l_2 A_2 x_2^{\beta_2} &= K - x_2, \\ \beta_1 l_1 A_1 x_2^{\beta_1} + \beta_2 l_2 A_2 x_2^{\beta_2} &= -x_2. \end{aligned} \right. \tag{12}$$

Similarly, the smoothness of $V_1(x)$ at x_1 and x_2 yields

$$\left\{\begin{array}{lcl} A_1 x_2^{\beta_1} + A_2 x_2^{\beta_2} & = & C_1 x_2^{\gamma_1} + C_2 x_2^{\gamma_2} + \phi(x_2), \\ \beta_1 A_1 x_2^{\beta_1} + \beta_2 A_2 x_2^{\beta_2} & = & \gamma_1 C_1 x_2^{\gamma_1} + \gamma_2 C_2 x_2^{\gamma_2} + x_2\phi'(x_2), \end{array}\right. \tag{13}$$

and

$$\left\{\begin{array}{lcl} C_1 x_1^{\gamma_1} + C_2 x_1^{\gamma_2} + \phi(x_1) & = & K - x_1, \\ \gamma_1 C_1 x_1^{\gamma_1} + \gamma_2 C_2 x_1^{\gamma_2} + x_1\phi'(x_1) & = & -x_1. \end{array}\right. \tag{14}$$

Combining the above three equations and following some algebraic manipulation, we obtain an algebraic equation for x_1 and x_2:

$$\begin{pmatrix} x_1^{-\gamma_1} & 0 \\ 0 & x_1^{-\gamma_2} \end{pmatrix} F_1(x_1) = \begin{pmatrix} x_2^{-\gamma_1} & 0 \\ 0 & x_2^{-\gamma_2} \end{pmatrix} F_2(x_2), \tag{15}$$

where

$$F_1(x_1) = \begin{pmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{pmatrix}^{-1} \begin{pmatrix} K - x_1 - \phi(x_1) \\ -x_1 - x_1\phi'(x_1) \end{pmatrix}$$

and

$$F_2(x_2) = \begin{pmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{pmatrix}^{-1} \left[\begin{pmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{pmatrix}^{-1} \begin{pmatrix} K - x_2 \\ -x_2 \end{pmatrix} - \begin{pmatrix} \phi(x_2) \\ x_2\phi'(x_2) \end{pmatrix} \right].$$

In particular, if $r + \lambda_1 - \mu_1 \neq 0$, where $\phi(x_1)$ is in the form of (11), then

$$F_1(x_1) = a_1 + a_2 x_1$$

and

$$F_2(x_2) = b_1 + b_2 x_2.$$

Here

$$a_1 = \begin{pmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{rK}{r+\lambda_1} \\ 0 \end{pmatrix}, \qquad a_2 = \begin{pmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{\mu_1 - r}{r+\lambda_1-\mu_1} \\ \frac{\mu_1 - r}{r+\lambda_1-\mu_1} \end{pmatrix},$$

$$b_1 = \begin{pmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{pmatrix}^{-1} \left[\begin{pmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{pmatrix}^{-1} \begin{pmatrix} K \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{\lambda_1 K}{r+\lambda_1} \\ 0 \end{pmatrix} \right],$$

$$b_2 = \begin{pmatrix} 1 & 1 \\ \gamma_1 & \gamma_2 \end{pmatrix}^{-1} \left[- \begin{pmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} l_1 & l_2 \\ \beta_1 l_1 & \beta_2 l_2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{\lambda_1}{r+\lambda_1-\mu_1} \\ \frac{\lambda_1}{r+\lambda_1-\mu_1} \end{pmatrix} \right].$$

The coefficients are given by

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} l_1 x_2^{\beta_1} & l_2 x_2^{\beta_2} \\ \beta_1 l_1 x_2^{\beta_1} & \beta_2 l_2 x_2^{\beta_2} \end{pmatrix}^{-1} \begin{pmatrix} K - x_2 \\ -x_2 \end{pmatrix}, \qquad \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} l_1 A_1 \\ l_2 A_2 \end{pmatrix},$$

and

$$\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} x_1^{\gamma_1} & x_1^{\gamma_2} \\ \gamma_1 x_1^{\gamma_1} & \gamma_2 x_1^{\gamma_2} \end{pmatrix}^{-1} \begin{pmatrix} K - x_1 - \phi(x_1) \\ -x_1 - x_1 \phi'(x_1) \end{pmatrix}.$$

With these coefficients, the value functions become

$$V_1(x) = \begin{cases} A_1 x^{\beta_1} + A_2 x^{\beta_2} & \text{if } x > x_2, \\ C_1 x^{\gamma_1} + C_2 x^{\gamma_2} + \phi(x) & \text{if } x_1 < x \le x_2, \\ K - x & \text{if } x \le x_1, \end{cases}$$
$$V_2(x) = \begin{cases} B_1 x^{\beta_1} + B_2 x^{\beta_2} & \text{if } x > x_2, \\ K - x & \text{if } x \le x_2. \end{cases} \tag{16}$$

2.2. Case 2: $x_2 < x_1 \le K$. The derivation of this case is analogous to that of $x_1 < x_2$, and we only summarize the results below.

Let $\widetilde{\gamma}_1$ and $\widetilde{\gamma}_2$ be the roots of

$$\mu_2 \gamma + \frac{1}{2}\sigma_2^2 \gamma(\gamma - 1) = r + \lambda_2,$$

and $\widetilde{\phi}(x)$ be a solution to

$$(r + \lambda_2) V_2(x) = x\mu_2 V_2'(x) + \frac{1}{2} x^2 \sigma_2^2 V_2''(x) + \lambda_2 (K - x).$$

Then, x_1, x_2 satisfy

$$\begin{pmatrix} x_1^{-\widetilde{\gamma}_1} & 0 \\ 0 & x_1^{-\widetilde{\gamma}_2} \end{pmatrix} \widetilde{F}_1(x_1) = \begin{pmatrix} x_2^{-\widetilde{\gamma}_1} & 0 \\ 0 & x_2^{-\widetilde{\gamma}_2} \end{pmatrix} \widetilde{F}_2(x_2), \tag{17}$$

with

$$\widetilde{F}_1(x_1) = \begin{pmatrix} 1 & 1 \\ \widetilde{\gamma}_1 & \widetilde{\gamma}_2 \end{pmatrix}^{-1} \left[\begin{pmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} \widetilde{l}_1 & \widetilde{l}_2 \\ \beta_1 \widetilde{l}_1 & \beta_2 \widetilde{l}_2 \end{pmatrix}^{-1} \begin{pmatrix} K - x_1 \\ -x_1 \end{pmatrix} - \begin{pmatrix} \phi(x_1) \\ x_1 \phi'(x_1) \end{pmatrix} \right]$$

and

$$\widetilde{F}_2(x_2) = \begin{pmatrix} 1 & 1 \\ \widetilde{\gamma}_1 & \widetilde{\gamma}_2 \end{pmatrix}^{-1} \begin{pmatrix} K - x_2 - \phi(x_2) \\ -x_2 - x_2\phi'(x_2) \end{pmatrix},$$

where $\widetilde{l}_i = 1/l_i$.

In particular, if $r + \lambda_2 - \mu_2 \neq 0$, then $\widetilde{\phi}(x)$ is given by

$$\widetilde{\phi}(x) = \frac{\lambda_2 K}{r + \lambda_2} - \frac{\lambda_2 x}{r + \lambda_2 - \mu_2},$$

and

$$\widetilde{F}_1(x_1) = \widetilde{a}_1 + \widetilde{a}_2 x_1,$$

$$\widetilde{F}_2(x_2) = \widetilde{b}_1 + \widetilde{b}_2 x_2,$$

where

$$\widetilde{a}_1 = \begin{pmatrix} 1 & 1 \\ \widetilde{\gamma}_1 & \widetilde{\gamma}_2 \end{pmatrix}^{-1} \left[\begin{pmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} \widetilde{l}_1 & \widetilde{l}_2 \\ \beta_1\widetilde{l}_1 & \beta_2\widetilde{l}_2 \end{pmatrix}^{-1} \begin{pmatrix} K \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{\lambda_2 K}{r+\lambda_2} \\ 0 \end{pmatrix} \right],$$

$$\widetilde{a}_2 = \begin{pmatrix} 1 & 1 \\ \widetilde{\gamma}_1 & \widetilde{\gamma}_2 \end{pmatrix}^{-1} \left[- \begin{pmatrix} 1 & 1 \\ \beta_1 & \beta_2 \end{pmatrix} \begin{pmatrix} \widetilde{l}_1 & \widetilde{l}_2 \\ \beta_1\widetilde{l}_1 & \beta_2\widetilde{l}_2 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{\lambda_2}{r+\lambda_2-\mu_2} \\ \frac{\lambda_2}{r+\lambda_2-\mu_2} \end{pmatrix} \right],$$

$$\widetilde{b}_1 = \begin{pmatrix} 1 & 1 \\ \widetilde{\gamma}_1 & \widetilde{\gamma}_2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{rK}{r+\lambda_2} \\ 0 \end{pmatrix}, \qquad \widetilde{b}_2 = \begin{pmatrix} 1 & 1 \\ \widetilde{\gamma}_1 & \widetilde{\gamma}_2 \end{pmatrix}^{-1} \begin{pmatrix} \frac{\mu_2 - r}{r+\lambda_2-\mu_2} \\ \frac{\mu_1 - r}{r+\lambda_2-\mu_2} \end{pmatrix}.$$

In short, if $x_1 > x_2$, then the corresponding value functions are

$$\begin{aligned} V_1(x) &= \begin{cases} \widetilde{A}_1 x^{\beta_1} + \widetilde{A}_2 x^{\beta_2} & \text{if } x > x_1, \\ K - x & \text{if } x \leq x_1, \end{cases} \\ V_2(x) &= \begin{cases} \widetilde{B}_1 x^{\beta_1} + \widetilde{B}_2 x^{\beta_2} & \text{if } x > x_1, \\ \widetilde{C}_1 x^{\widetilde{\gamma}_1} + \widetilde{C}_2 x^{\widetilde{\gamma}_2} + \widetilde{\phi}(x) & \text{if } x_2 < x \leq x_1, \\ K - x & \text{if } x \leq x_2, \end{cases} \end{aligned} \tag{18}$$

with

$$\begin{pmatrix} \widetilde{A}_1 \\ \widetilde{A}_2 \end{pmatrix} = \begin{pmatrix} x_1^{\beta_1} & x_1^{\beta_2} \\ \beta_1 x_1^{\beta_1} & \beta_2 x_1^{\beta_2} \end{pmatrix}^{-1} \begin{pmatrix} K - x_1 \\ -x_1 \end{pmatrix}, \qquad \begin{pmatrix} \widetilde{B}_1 \\ \widetilde{B}_2 \end{pmatrix} = \begin{pmatrix} l_1 \widetilde{A}_1 \\ l_2 \widetilde{A}_2 \end{pmatrix},$$

and

$$\begin{pmatrix} \widetilde{C}_1 \\ \widetilde{C}_2 \end{pmatrix} = \begin{pmatrix} x_2^{\widetilde{\gamma}_1} & x_2^{\widetilde{\gamma}_2} \\ \widetilde{\gamma}_1 x_2^{\widetilde{\gamma}_1} & \widetilde{\gamma}_2 x_2^{\widetilde{\gamma}_2} \end{pmatrix}^{-1} \begin{pmatrix} K - x_2 - \widetilde{\phi}(x_2) \\ -x_2 - x_2\widetilde{\phi}'(x_2) \end{pmatrix}.$$

2.3. Case 3: $x_1 = x_2 = x^* \leq K$. In this case, we have, for $x \geq x^*$,

$$V_1(x) = A_1 x^{\beta_1} + A_2 x^{\beta_2},$$
$$V_2(x) = B_1 x^{\beta_1} + B_2 x^{\beta_2},$$

and $V_1(x) = V_2(x) = K - x$ for $x \in [0, x^*]$. The smooth fit scheme leads to

$$\begin{cases} A_1(x^*)^{\beta_1} + A_2(x^*)^{\beta_2} & = & K - x^*, \\ \beta_1 A_1(x^*)^{\beta_1} + \beta_2 A_2(x^*)^{\beta_2} & = & -x^*, \end{cases} \tag{19}$$

and

$$\begin{cases} B_1(x^*)^{\beta_1} + B_2(x^*)^{\beta_2} & = & K - x^*, \\ \beta_1 B_1(x^*)^{\beta_1} + \beta_2 B_2(x^*)^{\beta_2} & = & -x^*. \end{cases} \tag{20}$$

Necessarily, we have $A_1 = B_1$ and $A_2 = B_2$, and therefore, $V_1 = V_2$.

Defining $V(x) = V_1(x) = V_2(x)$, then for $x > x^*$, the first equation in (6) reduces to

$$rV(x) = x\mu_i V'(x) + \frac{1}{2}x^2\sigma_i^2 V''(x),$$

for both $i = 1, 2$. This implies

$$V_1(x) = V_2(x) = \begin{cases} \dfrac{(K - x^*)x^\beta}{(x^*)^\beta} & \text{if } x > x^*, \\ K - x & \text{if } x \leq x^*, \end{cases}$$

where $x^* = K\beta/(\beta - 1)$ and β is the negative solution of

$$r - \left(\mu_i - \frac{1}{2}\sigma_i^2\right)\beta - \frac{1}{2}\sigma_i^2\beta^2 = 0$$

for $i = 1$ or $i = 2$.

3. Optimality of the solution. Now, we prove the optimality of $V_i(x)$ and x_i for $i = 1, 2$ derived in the previous section. For general results on stochastic calculus, we refer to the books by Karatzas and Shreve [11], McKean [15], and Revuz and Yor [18].

Recall

$$V^*(x, i) = \sup_\tau E[e^{-r\tau}(K - X(\tau))^+ \mid X(0) = x, \epsilon(0) = i].$$

Then we must prove the following claim.

THEOREM 3.1. *Suppose that* (15) (resp., (17)) *has a solution* (x_1^*, x_2^*) *such that* $0 < x_1^* \le K$ *and* $0 < x_2^* \le K$. *Assume* $V_i(x) > (K-x)^+$ *on* (x_i^*, ∞) *and* $\mu_i \ge 0$ *for* $i = 1, 2$. *Define*

$$D = \{(x, i) \ \mid \ V_i(x) > (K-x)^+\},$$

and let

$$\tau^* = \inf\{t \ge 0 \ \mid \ (X(t), \epsilon(t)) \notin D\}.$$

Then τ^* *is an optimal stopping time, and* $V_i(x)$ *are value functions (i.e.,* $V_i(x) = V^*(x, i)$*) and are given by* (16) (resp., (18)).

Proof. It is easy to see that $V_i(\infty) = 0$, $i = 1, 2$, and

$$D = \{(x, 1) \ \mid \ x \in (x_1^*, \infty)\} \cup \{(x, 2) \ \mid \ x \in (x_2^*, \infty)\}.$$

For any $v(x, i) \in C^2$, define

$$\mathcal{L}v(x, i) = x\mu_i \frac{\partial v(x, i)}{\partial x} + \frac{1}{2} x^2 \sigma_i^2 \frac{\partial^2 v(x, i)}{\partial x^2} + \lambda_i (v(x, 3-i) - v(x, i)) - rv(x, i).$$

Let $v(x, i) = V_i(x)$. Then $\mathcal{L}v \le 0$ on $(x, i) \in D$. Using Dynkin's formula, we have

$$d(e^{-rt} v(X(t), \epsilon(t))) = e^{-rt} \mathcal{L}v(X(t), \epsilon(t)) dt + d(\text{martingale}).$$

For any stopping time τ, it follows, from a smooth approximation approach for variational inequalities in Øksendal [16, p. 204], that

$$v(x, i) \ge E[e^{-r\tau} v(X(\tau), \epsilon(\tau))] \ge E[e^{-r\tau} (K - X(\tau))^+]. \tag{21}$$

To show the optimality of τ^*, note that if $\tau^* < \infty$, then $v(X(\tau^*), \epsilon(\tau^*)) = (K - X(\tau^*))^+$. In this case, Dynkin's formula yields $v(x, i) = E[e^{-r\tau^*} (K - X(\tau^*))^+]$. Otherwise, let

$$D_k = D \cap \{x < k\}, \quad \text{ for } k = 1, 2, \ldots.$$

Let $\tau_k = \inf\{t \ge 0 \ \mid \ (X(t), \epsilon(t)) \notin D_k\}$. Then we can show that $\tau_k \to \tau^*$ a.s. Moreover, as in Zhang [24, Theorems 4.5 and 4.6], we can show that, for each k, $\tau_k < \infty$ a.s. Using the definition of τ_k, we have, for $k > K$,

$$v(X(\tau_k), \epsilon(\tau_k)) = v(X(\tau_k), \epsilon(\tau_k)) I_{\{X(\tau_k) = k\}} + v(X(\tau_k), \epsilon(\tau_k)) I_{\{X(\tau_k) < k\}}.$$

Note that

$$v(X(\tau_k), \epsilon(\tau_k)) I_{\{X(\tau_k) < k\}} = (K - X(\tau_k))^+ I_{\{X(\tau_k) < k\}} \le (K - X(\tau_k))^+.$$

Moreover, note that $0 \le v(x, i) \le K$ and $e^{-r\tau_k} I_{\{X(\tau_k) = k\}} \to 0$, as $k \to \infty$, a.s. It follows that

$$E[e^{-r\tau_k} v(X(\tau_k), \epsilon(\tau_k)) I_{\{X(\tau_k) = k\}}] \to 0.$$

Therefore, we have, as $k \to \infty$,

$$v(x, i) \le E[e^{-r\tau_k} v(X(\tau_k), \epsilon(\tau_k))] \le E[e^{-r\tau^*} (K - X(\tau^*))^+].$$

Combining this with (21), we have

$$v(x,i) = E[e^{-r\tau^*}(K - X(\tau^*))^+].$$

This completes the proof.

Remark 3.1. As mentioned earlier, when $\sigma_1 \neq \sigma_2$, $\epsilon(t)$ becomes observable from the quadratic variation of $X(t)$ by Ito's calculus (see McKean [14]) and yields the joint Markov structure of $(X(t), \epsilon(t))$. This is one of the key points for our analysis. Although the case $\sigma_1 = \sigma_2$ is of independent interest from the filtering perspective since $\epsilon(t)$ is no longer observable (see Wonham [22] for estimating the probability distribution of $\epsilon(t)$, Liptser and Shiryayev [13] for general filtering, and Zhang [26, 27] for state detection and hybrid filtering), the option pricing problem is exactly the McKean problem, since a Girsanov transformation will reduce the regime switching model to the Black–Scholes model.

Remark 3.2. When $\lambda_1\lambda_2 = 0$, the corresponding $\epsilon(t)$ reduces to a single jump process, and the value functions can be solved sequentially using our method.

Remark 3.3. The optimality proof in Theorem 3.1 indicates the uniqueness of the value functions and that of the corresponding x_i's. Moreover, the assumption $V_i(x) > (K-x)^+$ or the existence of x_1, x_2 would be redundant if we assume the C^1 smoothness at the boundary x_1, x_2.

Remark 3.4. The assumption $\mu_i \geq 0$ guarantees that $e^{-rt}v(X(t), \epsilon(t))$ is a supermartingale. This is not restrictive in general. Indeed, it is standard in risk-neutral option pricing to have $\mu_1 = \mu_2 = r \geq 0$, following a change of measure via the Girsanov transformation.

Remark 3.5. It is clear from our analysis that a closed-form solution is possible if and only if K, the number of states of $\epsilon(t)$, equals two, since in general an algebraic equation of order $2K$ needs to be solved.

4. Numerical simulation. In this section we perform numerical experiments to compare the analytical solutions with the TPBVDE solutions studied in Zhang [24], together with the numerical results derived from a dynamic programming (DP) approach.

To this end, we first briefly review both DP and TPBVDE methods.

4.1. Dynamic programming. The DP approach we adopt here is built on the discretization method of the regime switching model proposed by Guo [6].

For a fixed T, let us divide the interval $[0,T]$ into N subintervals such that $T = Nh$. Moreover, if we define

$$u_i = e^{\sigma_i\sqrt{h}}, \quad l_i = e^{-\lambda_i h}, \quad d = e^{-rh}, \tag{22}$$

$$p_i = \frac{\mu_i h + \sigma_i\sqrt{h} - 0.5\sigma_i^2 h}{2\sigma_i\sqrt{h}}, \quad p_i + q_i = 1, \tag{23}$$

then the discrete counterpart of the process $(X(t), \epsilon(t))$ becomes the two-dimensional Markov chain (X_n, ϵ_n) that satisfies the recurrence

$$(X_n, \epsilon_n) = \eta_n^{(\epsilon_n, \epsilon_{n-1})}(X_{n-1}, \epsilon_{n-1}), \tag{24}$$

where $\eta_n^{i,j}$ are independently and identically distributed (i.i.d.) random variables taking values u_j with probability $p_j(\chi_{i,1-j} + (-1)^{\chi_{i,1-j}} e^{-\lambda_j h})$ and $1/u_j$ with probability

$(1-p_j)(\chi_{i,1-j}+(-1)^{\chi_{i,1-j}}e^{-\lambda_j h})$, respectively, where $(i,j=1,2)$ and

$$\chi(i,j)=\begin{cases}1, & i=j=1,2,\\ 0, & i\neq j.\end{cases}$$

In other words, (X_n,ϵ_n) is a random walk taking values on the set $(u_1^m u_2^n, i)$ with $i=1,2$ and $m,n=0,\pm1,\pm2,\ldots$ such that X_n represents the stock price at time n and ϵ_n the state of the market at time n.

Furthermore, the optimal stopping problem in question becomes

$$\widetilde{V}_i(x)=\sup_{\tau\in\{1,2,\ldots,\}} E[d^n(K-X_n)^+|\epsilon_0=i,X_0=x]. \tag{25}$$

Given the Markov chain $X=((X_n,\epsilon_n),\mathcal{F}_n,P)$, the optimal stopping problem (25) can be derived via the following dynamic programming principle:

$$\begin{aligned}
W_0(x) &= (K-x)^+,\\
Z_0(x) &= (K-x)^+,\\
W_m(x) &= \max\left\{W_{m-1}(x), dp_1l_1W_{m-1}(u_1x)+dl_1q_1W_{m-1}\left(\tfrac{x}{u_1}\right)\right.\\
&\qquad \left.+\,d(1-l_1)p_2Z_{m-1}(u_2x)+d(1-l_1)q_2Z_{m-1}\left(\tfrac{x}{u_2}\right)\right\},\\
Z_m(x) &= \max\left\{Z_{m-1}(x), dp_2l_2Z_{m-1}(u_2x)+dl_2q_2Z_{m-1}\left(\tfrac{x}{u_2}\right)\right.\\
&\qquad \left.+\,(1-l_2)dp_1W_{m-1}(u_1x)+(1-l_2)dq_1W_{m-1}\left(\tfrac{x}{u_1}\right)\right\}.
\end{aligned}$$

It is clear that $W_m(x), Z_m(x)$ are nondecreasing sequences, and

$$\widetilde{V}_1(x)=\lim_{n\to\infty}W_n(x),$$

$$\widetilde{V}_2(x)=\lim_{n\to\infty}Z_n(x).$$

Evidently, $\widetilde{V}_1(x)$ and $\widetilde{V}_2(x)$ are bounded nonnegative decreasing functions, and $\widetilde{V}_1(x)\geq(K-x)^+$, $\widetilde{V}_2(x)\geq(K-x)^+$. They are also called the least excessive dominating functions.

If we define

$$x_1=\min\left\{x\geq 0, \min_{i\in\{1,2\}}\widetilde{V}_i(x)=(K-x)^+\right\}$$

and

$$x_2=\min\left\{x\geq 0, \max_{i\in\{1,2\}}\widetilde{V}_i(x)=(K-x)^+\right\},$$

then x_1, x_2 are the so-called free boundary for the stopping rule.

With proper smooth conditions, $\widetilde{V}_i(x)$ coincides with $V(x,i)$ and hence with $V^*(x,i)$. For more detailed discussions on the least excessive dominating function and its application in option pricing, interested readers are referred to Guo [7] and Shiryayev et al. [21].

4.2. The TPBVDE approach. The TPBVDE approach was proposed by Zhang [24] to derive certain selling rules of threshold type. The stopping rule is to stop whenever the underlying stock price reaches two predefined bounds, an upper bound B or a lower bound A:

$$\tau_0 = \inf \{t > 0 \ \mid \ X(t) \notin (A, B)\} .$$

This rule is suboptimal since it limits the holder's choice to a smaller class of stopping times. If one takes $A = x^*$ and $B = \infty$ in Case 3, then it leads to a preferable stopping rule of $\tau_0 = \tau^*$.

The basic idea is to first choose a region of (A, B) so that for any given $0 \le a < b$,

$$X(0)e^{-b} \le A \le X(0)e^{-a},$$

$$X(0)e^{a} \le B \le X(0)e^{b}.$$

Next, we choose A and B within this interval to maximize

$$E[e^{-r\tau}(K - X(\tau)^+].$$

With this given A and B, the value function can thus be derived via analysis of a TPBVDE. (See [24] for details.)

4.3. Numerics. This section will report the numerical comparison results. First, we take

$$r = 3, \quad \mu_1 = \mu_2 = 3, \quad K = 5,$$

$$\lambda_1 = \lambda_2 = 100, \quad \sigma_1 = 9, \quad \sigma_2 = 5,$$

and compare the closed-form solution with the numerical solutions from the DP and TPBVDE methods; for the latter, we use the lower bound $a = 0$ and upper bounds $b = 3, b = 10$. The numerical results are plotted in Figure 1 and labeled with $V^{\mathrm{e}}(x, i)$, $V^{\mathrm{DP}}(x, i)$, $V^{\mathrm{b=3}}(x, i)$, and $V^{\mathrm{b=10}}(x, i)$, accordingly.

After 4000 iterations, with $N = 100,000$ and $h = 0.0001$, we obtain the threshold levels $(x_1^*, x_2^*) = (0.454, 0.617)$ for the DP approach, in comparison to the $(x_1^*, x_2^*) = (0.441, 0.614)$ derived from the closed-form solution.

Figure 2 confirms $V_i(x) \ge (K - x)^+$ and illustrates the differences of these value functions. As is shown, the accuracy of the two-point value method improves with increases in the upper bound b. The DP approach approximates the exact solutions better than the TPBVDE method for $b = 3$, while the converse is true with $b = 10$. In addition, these differences equal zero on the intervals (x_1^*, ∞) and (x_2^*, ∞) for $\epsilon(0) = 1$ or 2, respectively.

Next, we check the monotonicity of these threshold levels with respect to σ_i and λ_i. First, we vary σ_1 and keep all other parameters fixed. The resulting (x_1^*, x_2^*) are listed in Table 1. Both threshold levels x_1^* and x_2^* decrease with decreasing σ_1. This shows that a larger σ_1 leads to a higher option premium and therefore a smaller threshold level.

We then vary λ_1. The result in Table 2 implies that both x_1^* and x_2^* increase with λ_1 increasing: this is because a larger λ_1 implies a shorter period for $\epsilon(t)$ staying at

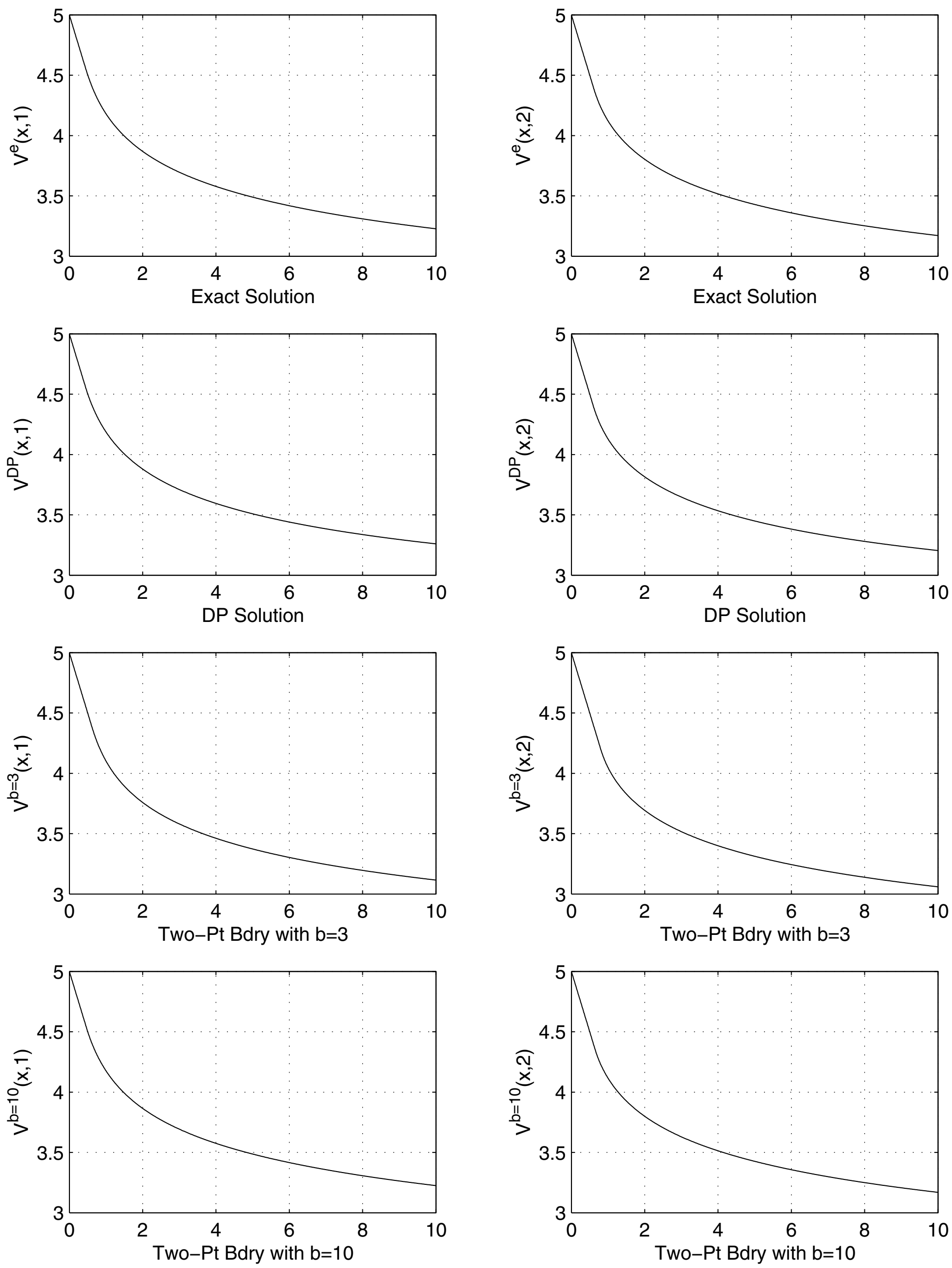

Fig. 1. *Value functions.*

$\epsilon(t) = 1$ and a smaller weight on $\sigma_1 = 9$ $(> \sigma_2 = 5)$, which leads to smaller average volatility.

These monotonicity properties may be better explained using the average volatility $\overline{\sigma} = \sqrt{\nu_1\sigma_1^2 + \nu_2\sigma_2^2}$, where (ν_1, ν_2) is the stationary distribution corresponding to the generator of $\epsilon(t)$. The results in Tables 1 and 2 suggest that both x_1^* and x_2^*

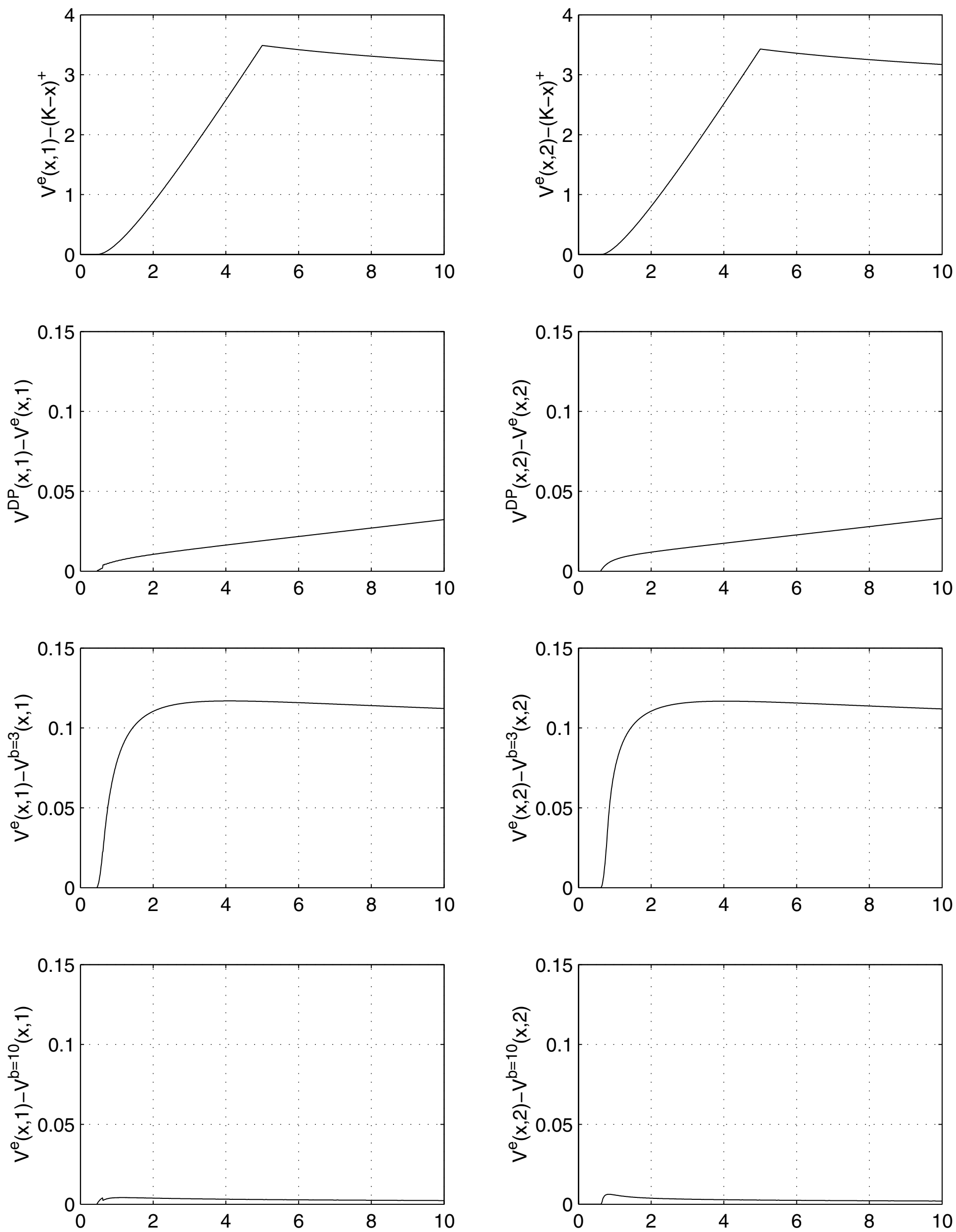

Fig. 2. *Differences between value functions.*

decrease with decreasing $\overline{\sigma}$.

Not surprisingly, the convergence rate of the DP approach depends on the choice of parameters. This in essence has to do with the specific discretization method of the underlying diffusion process. For example, with the same parameters specified above and with perturbations on the magnitude of r, we found that the smaller the r, the longer the computational time.

Table 1
Dependency on σ_1, given $\sigma_2 = 5$.

σ_1	7	8	9	10	11	12
Exact	(.646,.764)	(.531,.683)	(.441,.614)	(.369,.554)	(.312,.505)	(.266,.462)
DP	(.660,.773)	(.545,.687)	(.454,.617)	(.381,.557)	(.324,.506)	(.277,.465)

Table 2
Dependency on λ_1, given $\lambda_2 = 100$.

λ_1	80	90	100	110	120	130
Exact	(.425,.596)	(.433,.605)	(.441,.614)	(.448,.621)	(.456,.629)	(.463,.637)
DP	(.437,.599)	(.446,.607)	(.454,.617)	(.461,.624)	(.469,.632)	(.476,.640)

As far as total CPU usage is concerned, the DP approach took substantially longer time than the closed-form and the TPBVDE methods. For example, with a basic Linux 7.2 i386 system, it took a little more than 30 minutes for our DP solution to complete 4000 iterations, while it took just seconds for both the exact and TPBVDE methods.

5. Concluding remarks. In this paper we have derived a closed-form solution to the optimal stopping problem for pricing perpetual American put options in a regime switching model.

It remains to be seen whether there are alternative methods for deriving the solution. One obvious candidate is the first passage time technique, which was exploited in solving the McKean problem (McKean [14] and Karlin and Taylor [12]). However, despite the two promising features that (i) $(X(t), \epsilon(t))$ is jointly Markovian and (ii) the free boundaries are of threshold type, it seems hard to explicitly solve the integral equation system using results of the first passage time for regime switching models (derived in Guo [8]). The main obstacle seems to be the instantaneous jump due to the regime switching.

It is also of interest to extend our analysis to the case when T is finite. Needless to say, this case would be mathematically interesting and practically appealing. However, a closed-form solution for a finite time horizon problem with regime switching is difficult to obtain. Even the special case with no regime switching remains an open problem to date. Moreover, with all the structural insights gained from the infinite case, it is not even clear whether the boundary is monotonic; i.e., will $x_1 < x_2$ imply $x_1(T) < x_2(T)$? Assuming this monotonicity condition a priori, Buffington and Elliott [2] extended our analysis and obtained certain properties for the value functions of American put options with $T < \infty$.

Nevertheless, our hope is that the closed-form solutions in this paper will provide better understanding of and some insight into the nature of optimal stopping rules, and our approach can be useful for numerical approximations of long-term American options.

Acknowledgments. We thank the referees for a very careful reading of the manuscript and many constructive suggestions.

REFERENCES

[1] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, J. Political Economy, 81 (1973), pp. 637–654.
[2] J. Buffington and R. J. Elliott, *American options with regime switching*, Int. J. Theor. Appl. Finance, 5 (2002), pp. 497–514.
[3] G. B. Di Masi, Yu. M. Kabanov, and W. J. Runggaldier, *Mean-variance hedging of options on stocks with Markov volatilities*, Theory Probab. Appl., 39 (1994), pp. 172–181.
[4] D. Duffie, *Dynamic Asset Pricing Theory*, 2nd ed., Princeton University Press, Princeton, NJ, 1996.
[5] X. Guo, *Inside Information and Stock Fluctuations*, Ph.D. Dissertation, Department of Mathematics, Rutgers University, Newark, NJ, 1999.
[6] X. Guo, *Inside information and option pricings*, Quant. Finance, 1 (2000), pp. 38–44.
[7] X. Guo, *An explicit solution to an optimal stopping problem with regime switching*, J. Appl. Probab., 38 (2001), pp. 464–481.
[8] X. Guo, *When the "bear" meets the "bull": A first passage time problem in a hidden Markov process*, Methodol. Comput. Appl. Probab., 3, (2001), pp. 134–143.
[9] S. D. Jacka, *Optimal stopping and the American put*, Math. Finance, 1 (1991), pp. 1–14.
[10] I. Karatzas, *On the pricing of American options*, Appl. Math. Optim., 17 (1988), pp. 37–60.
[11] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, Springer-Verlag, New York, 1998.
[12] S. Karlin and H. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 1975.
[13] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes*, Springer-Verlag, New York, 1977.
[14] H. P. McKean, *A free boundary problem for the heat equation arising from a problem in mathematical economics*, Ind. Management Rev., 6 (1965), pp. 32–39.
[15] H. P. McKean, *Stochastic Integrals*, Academic Press, New York, 1969.
[16] B. Øksendal, *Stochastic Differential Equations*, 4th ed., Springer-Verlag, New York, 1995.
[17] E. Renault and N. Touzi, *Option hedging and implied volatilities in a stochastic volatility model*, Math. Finance, 6 (1996), pp. 279–302.
[18] D. Revuz and M. Yor, *Continuous Martingale and Brownian Motion*, Springer-Verlag, New York, 1991.
[19] H. Robbins, D. Sigmund, and Y. Chow, *Great Expectations: The Theory of Optimal Stopping*, Houghton Mifflin, Boston, 1971.
[20] P. A. Samuelson, *Mathematics of speculative price (with an appendix on continuous-time speculative processes by R. C. Merton)*, SIAM Rev., 15 (1973), pp. 1–42.
[21] A. N. Shiryaev, Yu. M. Kabanov, O. D. Kramkov, and A. V. Mel'Nikov, *Toward the theory of pricing of options of both European and American types* I*, Discrete time*, Theory Probab. Appl., 39 (1994), pp. 14–60.
[22] W. M. Wonham, *Some applications of stochastic differential equations to optimal nonlinear filtering*, SIAM J. Control, 2 (1964), pp. 347–369.
[23] D. D. Yao, Q. Zhang, and X. Y. Zhou, *Option Pricing with Markov-Modulated Volatility*, preprint.
[24] Q. Zhang, *Stock trading: An optimal selling rule*, SIAM J. Control Optim., 40 (2001), pp. 64–87.
[25] Q. Zhang and G. Yin, *Nearly optimal asset allocation in hybrid stock-investment models*, J. Optim. Theory Appl., to appear.
[26] Q. Zhang, *Nonlinear filtering and control of a switching diffusion with small observation noise*, SIAM J. Control Optim., 36 (1998), pp. 1638–1668.
[27] Q. Zhang, *Hybrid filtering for linear systems with non-Gaussian disturbances*, IEEE Trans. Automat. Control, 45 (2000), pp. 50–61.