THE EFFECT OF NON-SMOOTH PAYOFFS ON THE PENALTY APPROXIMATION OF AMERICAN OPTIONS

S. D. HOWISON, C. REISINGER, AND J. H. WITTE

Abstract. This article combines various methods of analysis to draw a comprehensive picture of penalty approximations to the value, hedge ratio, and optimal exercise strategy of American options. While convergence of the penalised solution for sufficiently smooth obstacles is well established in the literature, see e.g. [4, 34], sharp rates of convergence and particularly the effect of gradient discontinuities (i.e. the omni-present 'kinks' in option payoffs) on this rate have not been fully analysed so far. This effect becomes important not least when using penalisation as a numerical technique. We use matched asymptotic expansions to characterise the boundary layers between exercise and hold regions, and to compute first order corrections for representative payoffs on a single asset following a diffusion or jump-diffusion model. Furthermore, we demonstrate how the viscosity theory framework in [15] can be applied to this setting to derive upper and lower bounds on the value. In a small extension to [4], we derive weak convergence rates also for option sensitivities for convex payoffs under jump-diffusion models. Finally, we outline applications of the results, including accuracy improvements by extrapolation.

Key Words: American Option, Jump-Diffusion Model, Penalty Method, Penalization Error, Non-Smooth Payoff

2010 Mathematics Subject Classification: 60G40, 47G20

1. Introduction

An American option is a financial instrument that gives its holder the right to claim a specified payoff on an asset at any time up to a certain date. Pricing an American option involves determining an optimal exercise strategy in addition to the price itself.

There are two main equivalent approaches to formulating this problem: a probabilistic approach based on stopping times, and a deterministic one based on a linear complementarity problem (free boundary problem). The optimal stopping formulation in diffusion models was first introduced in [3] and [16]; a concise outline can be found in [25]. For simplicity, we discuss first the Black-Scholes setting (cf. [5]), i.e., the stock price follows

$$dS_t/S_t = \mu \, dt + \sigma \, dW_t, \tag{1.1}$$

where σ is the volatility, μ the drift rate, and W a standard Brownian motion.

In [8], it is described how an American option in this setting can be priced using a linear complementarity problem (LCP)

$$\min(-\mathcal{L}_{\mathrm{BS}} V, V - \Psi) = 0, \tag{1.2}$$

Date: May 2012.

J. H. Witte acknowledges support from Balliol College, University of Oxford, the UK Engineering and Physical Sciences Research Council (EPSRC).

where Ψ is the payoff and $\mathcal{L}$ is the Black-Scholes operator

$$\mathcal{L}_{\mathrm{BS}}V := \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV, \tag{1.3}$$

where r is the risk-free interest rate and q a continuously paid dividend yield. The relation between stopping times and PDEs in this context is further analysed in [17].

In this paper, we are concerned in particular with the effects of the payoff function on a so-called *penalty approximation* to the value of such an option. Penalty approximations are useful both for the analysis [4, 34] and numerical analysis [15] of the limiting problem, but also lend themselves to arguably the most efficient numerical approximation methods presently available for American option valuation [10].

Penalisation of (parabolic) variational inequalities is classical (cf. [4]). The canonical penalty approximation of (1.2) is

$$-\mathcal{L}_{\mathrm{BS}}V^\epsilon = \frac{1}{\epsilon}\max(\Psi - V^\epsilon, 0) \tag{1.4}$$

for $\epsilon > 0$ (cf. [3]). The penalty term on the right-hand side is only active when $V^\epsilon < \Psi$, and then it serves to push V upwards towards the payoff.

In the context of American options, in chronological order, [32, 31, 30] study the penalisation error for the Black-Scholes model and different penalty terms, and [1] uses penalisation implicitly to solve a calibration problem. [11, 10, 26, 9, 27] introduce penalty approximations as a means of solving the discretised variational inequality.

Effect of the Penalisation Error on Pricing and Hedging. This paper analyses the accuracy of a penalty approximation to the value function for relevant payoff functions. To do this, we first address the question of what a relevant measure of accuracy should be, which clearly depends on what the solution will normally be used for.

Hedging American options requires knowledge of the hedge ratio, i.e. the amount of stocks held short in the hedging portfolio per unit long position in the option before it is exercised. In the complete market case of the Black-Scholes model, the hedge ratio is the so-called *Delta*, $\Delta_t = (\partial V/\partial S)(S_t, t)$.

Because we do not know the exact option value, but only its penalty approximation, we are exposed to three sources of error if we, say, buy and hedge an American option:

a) We bid the – lower, as we shall see – price $V^\epsilon(S_0, 0)$ instead of $V(S_0, 0)$ for the option at the outset.
b) We hedge with the wrong hedge ratio
$$\Delta_t^\epsilon = \frac{\partial V^\epsilon}{\partial S}(S_t, t)$$
instead of the exact Delta Δ_t.
c) We exercise at the wrong time
$$\tau^\epsilon := \inf\{t : V^\epsilon(S_t, t) \le \Psi(S_t)\},$$
which is before the optimal exercise time τ, since $V^\epsilon(S_t, t) \le V(S_t, t)$.

The value of our hedge portfolio at $t < \tau^\epsilon$ is, by a classical replication argument (e.g. [25]),

$$X_t^\epsilon = X_0^\epsilon + \sigma \int_0^t \Delta_u^\epsilon S_u \mathrm{e}^{(t-u)}\, \mathrm{d}W_u^Q,$$

where $dW_t^Q = dW_t + (\mu - r)/\sigma\, dt$ is the increment of a Brownian motion under the risk-neutral measure Q.

Consider the stopping time $\tau^\epsilon \wedge t$. Then, by the Optional Stopping Theorem,

$$\mathbb{E}^Q[X_{\tau^\epsilon\wedge t}^\epsilon - X_{\tau^\epsilon\wedge t}] = X_0^\epsilon - X_0$$

and, by Itô isometry,

$$\begin{aligned}
\mathbb{V}^Q[X_{\tau^\epsilon\wedge t}^\epsilon - X_{\tau^\epsilon\wedge t}] &= \mathbb{E}\int_0^{\tau^\epsilon\wedge t} \sigma^2 S_u^2 \mathrm{e}^{2(t-u)} (\Delta_u^\epsilon - \Delta_u)^2\, \mathrm{d}u \\
&\le \mathbb{E}\int_0^{T} \sigma^2 S_u^2 \mathrm{e}^{2(t-u)} (\Delta_u^\epsilon - \Delta_u)^2\, \mathrm{d}u \\
&= \int_0^{T} \sigma^2 \mathrm{e}^{2(t-u)} \mathbb{E}[S_u^2 (\Delta_u^\epsilon - \Delta_u)^2]\, \mathrm{d}u \\
&\le \sigma^2 \mathrm{e}^{2T} \int_0^T \int_0^\infty p(S_0, 0; S, u)\, S^2 \left(\frac{\partial V^\epsilon}{\partial S}(S,u) - \frac{\partial V}{\partial S}(S,u)\right)^2 \mathrm{d}S\, \mathrm{d}u \\
&\le C \int_0^T \int_0^\infty \left(\frac{\partial V^\epsilon}{\partial S}(S,u) - \frac{\partial V}{\partial S}(S,u)\right)^2 \mathrm{d}S\, \mathrm{d}u,
\end{aligned}$$

where $p(S_0, 0; S, u)$ is the transition density from S_0 at time 0 to S at time u, and the last inequality follows because $p\, S^2$ is bounded. The variance of the replication error at any time prior to exercise is therefore controlled by the squared H^1 distance between the true and penalty value function,

$$\|V^\epsilon - V\|_1^2 = \int_0^T \int_0^\infty \left(\frac{\partial V^\epsilon}{\partial S}(S,u) - \frac{\partial V}{\partial S}(S,u)\right)^2 + (V^\epsilon(S,u) - V(S,u))^2\, \mathrm{d}S\, \mathrm{d}u,$$

which is one of the error measures we will consider. Next, the loss incurred by exercising too early is

$$V(S_{\tau^\epsilon}, \tau^\epsilon) - V^\epsilon(S_{\tau^\epsilon}, \tau^\epsilon) = V(S_{\tau^\epsilon}, \tau^\epsilon) - \Psi(S_{\tau^\epsilon}), \tag{1.5}$$

which is the (positive) difference between true and penalised solution at the (sub-optimal) penalty exercise boundary. Viewed differently, by hedging with Δ^ϵ we are replicating an option which is exercised not at the optimal exercise time, but at the crossing time of an approximate exercise boundary. An alternative measure of error is therefore the maximum distance

$$\|V^\epsilon - V\|_\infty = \sup_{0\le t\le T, 0\le S} |V(S,t) - V^\epsilon(S,t)|,$$

which is also an upper bound for $|X_0 - X_0^\epsilon| = V(S_0, 0) - V^\epsilon(S_0, 0)$. We will study the convergence in this norm also.

Extension to Jump Models. The suggestion to allow the asset to jump occasionally to reflect the possibility of sudden changes in the market was first made in [20]; an extensive overview and detailed discussion of jump models and their use in modern mathematical finance can be found in [6]. The pricing of American options in the presence of jumps has been developed and studied in [34, 22], which remain the main references on the topic.

We consider models where the underlying asset follows a jump-diffusion process,

$$\mathrm{d}S_t/S_t = \mu\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + (J-1)\,\mathrm{d}N_t, \tag{1.6}$$

where J is a random jump amplitude lying in $[0,\infty)$, and N a compound Poisson process with jump rate $\lambda \geq 0$. The special case $\lambda = 0$ recovers the Black-Scholes model.

Under the assumption that jump risk is unpriced, the value of an American option under jump diffusion can still be described by an equation of the type (1.2), but with

$$\mathcal{L}_{\mathrm{BSJ}} V := \frac{\partial V}{\partial t} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + (r - q - \omega\lambda) S \frac{\partial V}{\partial S} - rV + \lambda \mathbb{E}[V(JS,t) - V(S,t)], \tag{1.7}$$

where the expectation is taken with respect to the jump size J, for fixed S, and $\omega = \mathbb{E}[J - 1]$. This can be re-written as a partial integro-differential equation (PIDE) in terms of the probability density function g of J via

$$\mathbb{E}[V(JS,t) - V(S,t)] = \int_0^\infty V(SJ,t) g(J)\,\mathrm{d}J \,-\, V(S,t).$$

It will also be useful to consider the PIDE in log-coordinates, $x = \log(S/S_0)$, $u(x,t) = V(S_0\exp(x),t)$, where

$$\mathcal{L}u = \frac{\partial u}{\partial t} + \tfrac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2} + (r - q - \omega\lambda - \sigma^2/2)\frac{\partial u}{\partial x} - ru + \lambda\left[\int_{-\infty}^{\infty} u(x+z,t)\nu(z)\,\mathrm{d}z \,-\, u(x,t)\right] \tag{1.8}$$

in (1.2) and $\phi(z) = \Phi(S_0\exp(z))$ is the new payoff and ν the density of $Z = \log(J)$. The pricing equation is then still (1.2), the penalised equation (1.4), now with the operator (1.8).

We will see that the inclusion of finite activity jumps does not alter the properties of penalty approximations qualitatively, and all the general results later on in the paper are derived for this class of models. Some of the specific examples used for numerical examples and for which we derive asymptotic expansions use the Black-Scholes model for ease of exposition. It will be stated clearly at the start of all sections where this is the case.

Structure of this Paper. In Section 2, we discuss a few representative examples of typical payoffs and present numerical results as motivation for the following analysis. In Section 3, we derive the leading order corrections to the penalty solution and the exercise boundary by matched asymptotic expansions, for the American put and butterfly. This analysis is useful because the localised structure of the error, especially for the derivative, is not visible from the more global functional analysis to follow and, to our knowledge, this is the first study of this behaviour. Section 4 generalises the convergence order of the penalisation error of the value to more general classes of convex (order ϵ) and non-convex (order $\epsilon^{1/2}$) payoffs, and gives sharp upper and lower bounds on the solution, following the framework of [15] and extending it to jump processes. Section 5 derives H^1 errors, showing that the rate $\epsilon^{1/2}$ derived in [4] also holds under jump-diffusions and for non-smooth but convex obstacles. Finally, in Section

6, we discuss the results and their applications; in particular, we show how extrapolation can be used for accuracy improvement.

2. Different Payoffs and Their Implications

Part of the appeal of penalty methods as a computational tool is that the resulting algorithms do not depend on the shape of the payoff. In contrast to the formulation as free boundary problem (e.g. 'front-fixing' methods), the topology of exercise and continuation regions is irrelevant for the definition of the penalty approximation and (iterative) solution algorithms based on it.

We will now illustrate how the shape and regularity of the payoff does, however, influence the approximation error.

Example Payoffs and their Exercise Strategies. Two typical payoffs are the standard put payoff

$$\Psi(S) = \max(K - S, 0), \tag{2.1}$$

with strike $K > 0$, and a butterfly spread

$$\Psi(S) = \max(V_0 - \alpha|S - K|, 0), \tag{2.2}$$

for some $\alpha, V_0 > 0$. We also consider an academic example of a 'modified' put, which is a sum of a put and a butterfly spread, and has a piecewise linear payoff.

For American options, typically a exercise boundary determines the asset price(s) at which, for a given point in time, the optimal policy switches from holding the option to exercising it.

Figures 1, 2, and 3 show value functions with their free boundaries for different payoffs. For illustrative purposes, we use a Black-Scholes framework with no dividends, interest rate $r = 0.05$, volatility $\sigma = 0.4$, and maturity $T = 1$. We will see later that jumps do not change the results qualitatively.

We discuss the three examples in turn.

(1) For the standard put, a smooth free boundary $S^*(t)$ separates an exercise region $S < S^*(t)$ from a hold region $S > S^*(t)$, and decreases strictly as we move away from expiry. Denote this vanilla American put value by $P(S, t)$. Note that, at $S = S^*(t)$,

$$P(S^*(t), t) = K - S^*(t), \qquad \frac{\partial P}{\partial S} = -1, \qquad \lim_{S \downarrow S^*} \frac{\partial^2 P}{\partial S^2} = \frac{2rK}{\sigma^2 {S^*}^2}, \tag{2.3}$$

see e.g. [8]. Before expiry, the solution is continuously differentiable with a jump in the second derivative (the *Gamma*) at the free boundary.

(2) For the butterfly spread, short before expiry, the option value is greater than the payoff on the call-like side $S < K$, and is similar to before on the put-like side $S > K$ with an exercise boundary between 100 and 150. The solution smooths the convex kinks which the payoff has at 50 and 100; however, at all times, the solution has a kink at 100. As we move away from expiry, the free boundary decreases strictly from 150 at first and then, having reached $S = 100$, remains there.

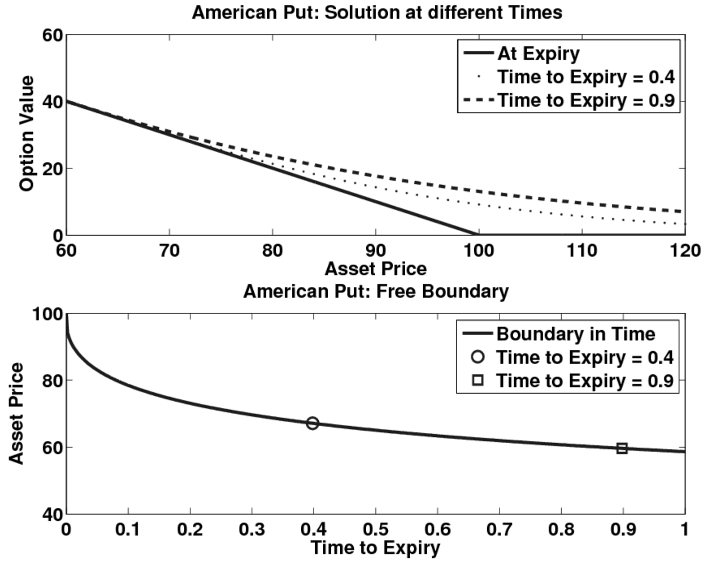

Figure 1. *The value of an American put at different points in time (above) and the evolution of the corresponding free boundary (below). Before expiry, where it equals the payoff, the value function looks smooth. The boundary is strictly decreasing as we move away from expiry.*

(3) For the modified put, whose payoff is defined in the caption of Figure 3, there is again a single free boundary which separates an exercise region for small S from a hold region. Before expiry, the solution smooths the convex kink which the payoff has at about 137, but, near expiry, it still has a kink at 105; however, far away from expiry, the solution looks smooth. As we move away from expiry, the free boundary decreases strictly at first, stagnates (until smooth pasting is reached) and then decreases strictly again. A discussion on this 'waiting time' phenomenon in the context of diffusion problems can be found in [23], to which the present case adds a further example.

Numerical Penalisation Error. We now analyse numerically the penalty convergence for these three examples. We use a Crank-Nicolson finite difference scheme to discretise (1.4) and solve the resulting non-linear discrete system by a Newton iteration (cf. [10]). We compute a reference solution to the original LCP using PSOR (cf. [7]). We measure spatial errors pointwise in the maximum norm, and similarly for the derivative, which is approximated by finite differences of the numerical solution. We measure the convergence rate in the penalty parameter ϵ from (1.4) as $\epsilon \to 0$ by regression of the errors for small ϵ. Generally, we use very fine time and space grids to suppress possible discretisation errors.

Figure 4 shows the numerically computed penalisation error for the standard put as a function of S and t. It appears that the error is constant in the exercise region, jumps to about half this value at the exercise boundary, and decays for large S. The irregular behaviour of the error surface close to the exercise boundary is due to the movement of the exercise boundary between time steps relative to the grid points. Figure 5 is the corresponding picture for the

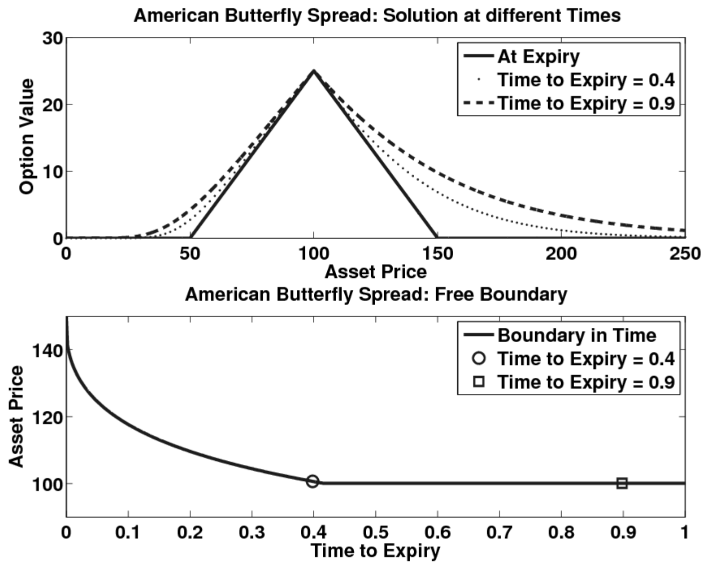

Figure 2. *The value of an American butterfly spread at different points in time (above) and the evolution of the corresponding free boundary (below). (Possibly hard to see in the plot, there is only one free boundary which lies between 100 and 150.) At all times, the value function has a kink at S=100. The boundary decreases strictly at first and then stagnates.*

penalisation error in the first S-derivative. The error appears localised in a very narrow region around the exercise boundary. The jagged shape of the surface results again from an interplay of the penalisation error and discretisation. For the chosen time step and mesh width, the width of the region of large error is small compared to the grid size, and, from one time step to the next, has a different location relative to its nearest grid points. For time steps, where the location of the maximum is close to a grid point, the plotted spike is large, whereas if the maximum lies between grid points, it is small.

For the butterfly, as seen from Figure 6, there is an asymmetry in the penalisation error between the call-like side, where the error grows more steeply in time-to-expiry, and the put-like side, where the error is flat up to the the point in backward-time where the exercise boundary hits the top of the payoff, and from then on increases for larger time-to-maturity. The error is largest at the strike, constant in time, and decays rapidly on either side.

The penalisation error for the modified put is shown in Figure 7.

Table 1 shows the convergence rates in the penalty parameter ϵ for the three payoffs. For the standard put, we find results for the spatial errors in V^ϵ and $\partial V^\epsilon/\partial S$ which appear consistent with $O(\epsilon)$ and $O(\epsilon^{1/2})$, respectively. For the butterfly spread, we find $O(\epsilon^{1/2})$ for the spatial error in V^ϵ , but convergence in $\partial V^\epsilon/\partial S$ seems to be very slow. Looking at the solutions (cf. Figures 1, 2), one readily suspects the concave kink of the butterfly spread, which is prevalent in the solution at all times, to be the reason for the slower convergence. This observation is further supported by the fact that, for the modified put, we find convergence rates

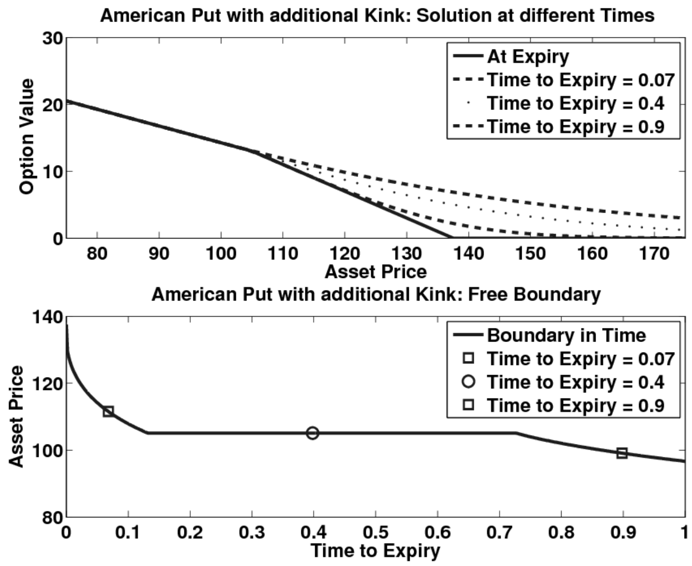

Figure 3. *An American option on a sum of a put and a butterfly payoff, i.e., a put with an additional concave kink at S=105. We see the value at different points in time (above) and the evolution of the corresponding free boundary (below). Near expiry, the value function has a kink at S=105, but, further away from expiry, it looks smooth. The boundary decreases strictly at first, stagnates, and then decreases strictly again.*

Penalty Approximation	Put		Butterfly Spread		Modified Put		
Time to Expiry	0.4	0.9	0.4	0.9	0.07	0.4	0.9
Order in $\lvert V - V^\epsilon \rvert$	1.00	1.00	0.50	0.50	0.51	0.51	0.53
Order in $\lvert \partial V/\partial S - \partial V^\epsilon/\partial S \rvert$	0.55	0.57	0.07	0.08	0.07	0.06	0.61

Table 1. *At different points t in time, we measure the convergence rates in $|V(\cdot,t) - V^\epsilon(\cdot,t)|$ and $|\partial V/\partial S(\cdot,t) - \partial V^\epsilon/\partial S(\cdot,t)|$, where V denotes the true solution, as $\epsilon \to 0$. For the put and the butterfly spread, the rates are the same at all times. For the modified put, the convergence rate in $\partial V^\epsilon/\partial S$ improves hugely when expiry is far into the future (at which point the free boundary has overcome the concave kink).*

comparable to the rates of the butterfly spread near expiry, but, further away from expiry, where the concave kink has been smoothed out (cf. Figure 3), convergence rates improve to $O(\epsilon^{1/2})$ for the spatial errors in V^ϵ and $\partial V^\epsilon/\partial S$.

In the next section, we will use matched asymptotic expansions for small ϵ to explain this behaviour and to derive the leading order corrections to the penalty solution.

3. Approximation by Matched Asymptotic Expansions

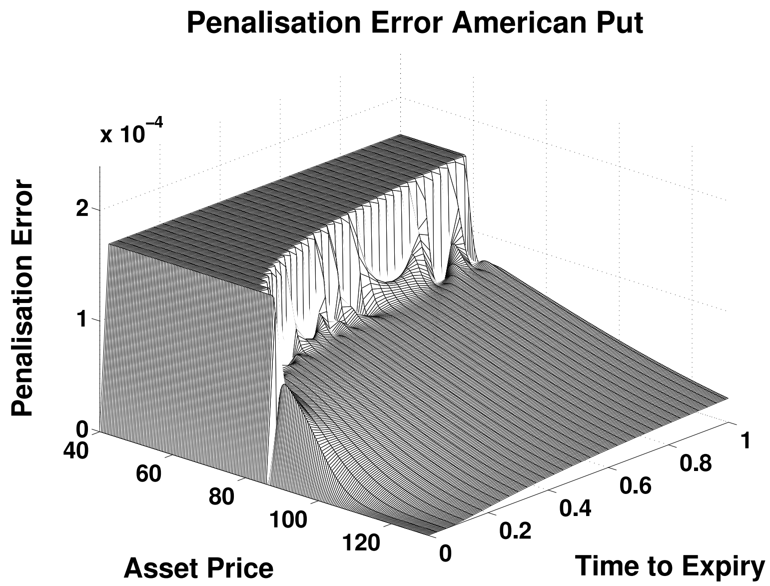

Figure 4. *Local structure of the penalisation error for the put. The error is largest, and roughly constant, in the exercise region, and decays rapidly over a small layer around the exercise boundary.*

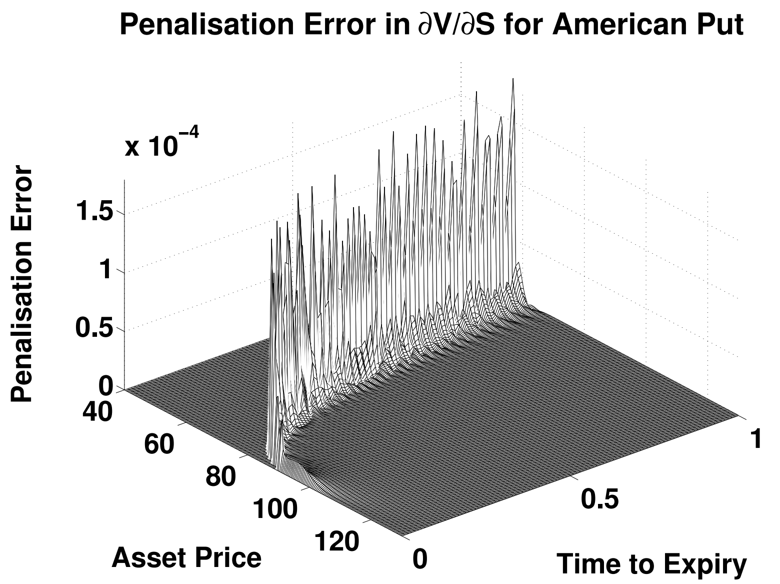

Figure 5. *Local structure of the penalisation error for the put delta. The error is largest in a small layer around the exercise boundary.*

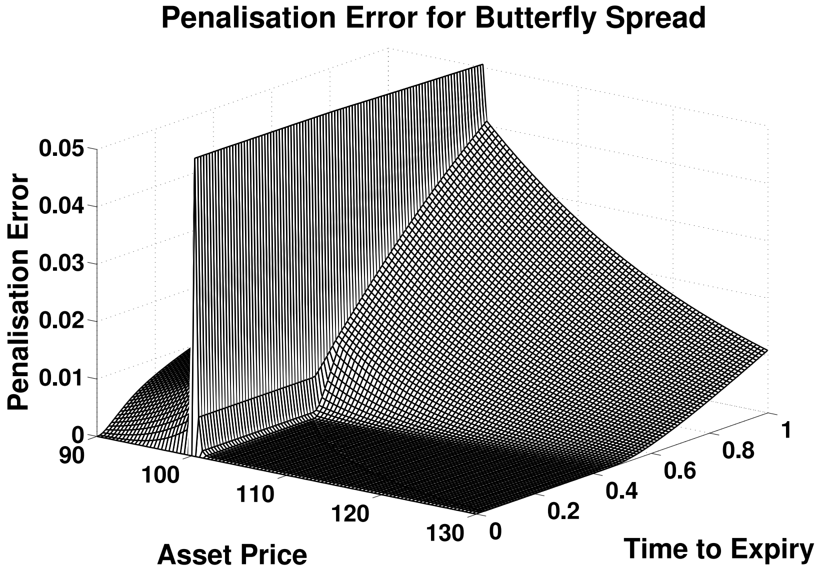

Figure 6. *Local structure of the penalisation error for the butterfly spread. The error is largest in a narrow region around the kink of the payoff. It is negligible on the put-like side up to the point where it is optimal not to exercise the option.*

3.1. **Asymptotics for Put.** We first consider an American put option on an asset in the Black-Scholes model with no dividends. The penalty equation is written

$$\mathcal{L}_{\rm BS} V^\epsilon = -\frac{1}{\epsilon}\max(K-S-V^\epsilon,0) = -\frac{\sigma^2}{\delta^2}\max(K-S-V^\epsilon,0), \tag{3.1}$$

where we have introduced a dimensionless parameter $\delta = \sigma\,\epsilon^{\frac{1}{2}} \ll 1$.[1]

There is a three-region structure, with 'outer' regions above and below an 'inner' region which spans the point at which the penalty solution crosses the payoff. We will see later that δ is the characterisitic width of the inner region between the exercise and hold regions of the option. Initial (i.e. near expiry) transients are ignored (this means both the penalty term transient and the American-put transient).

3.1.1. *Outer region $S > S^*(t)$ ('hold').* First, write

$$V^\epsilon(S,t) = K - S + W^\epsilon(S,t) \tag{3.2}$$

and expand

$$W^\epsilon(S,t) \sim W_0(S,t) + \delta W_1(S,t) + \delta^2 W_2(S,t) + \cdots.$$

We expect $W_0(S,t) = P(S,t) - (K-S)$ and $W_1(S,t) = 0$ (because smooth pasting always leads to a smaller error than a barrier-type 'pinned' condition, cf. the asymptotic results on Bermudan options and discrete barrier options in [28, 13]). W_0 satisfies $\mathcal{L}_{BS} W_0 = -rK$ and

[1] We could have used r or T instead to scale δ, with the same eventuel answer. The choice we have made makes the intermediate calculations simpler. We also keep writing V^ϵ to keep the notation simple.

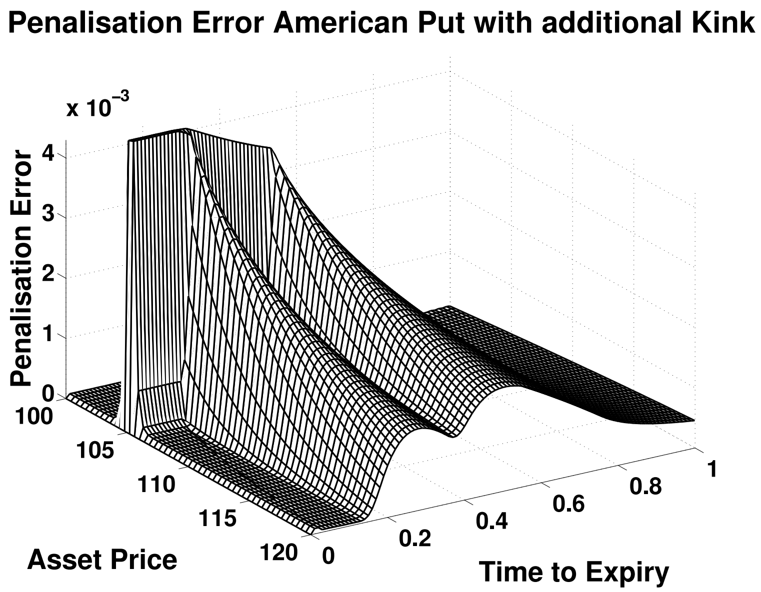

Figure 7. *Local structure of the penalisation error for the modified put. It shows a combination of features of the put and butterfly payoff, and a decay in time resulting from the waiting time phenomenon.*

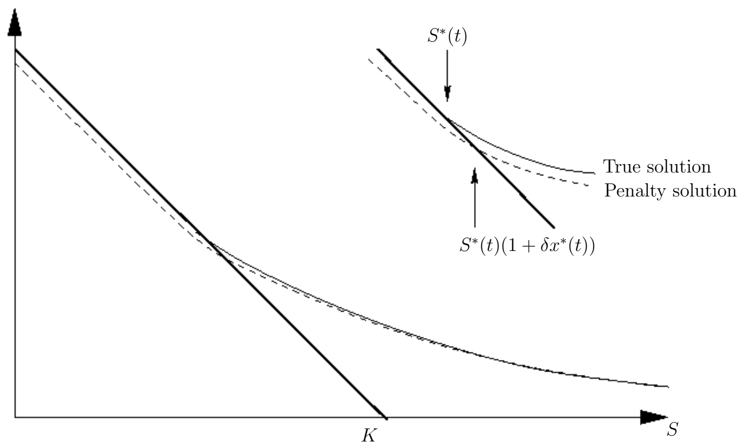

Figure 8. *Schematic of a put option solution and blow-up of inner region. True solution is the solid curve, penalty solution is dashed.*

all W_i for $i > 0$ satisfy the homogeneous Black-Scholes PDE, because the penalty is not active.

Introduce the inner variable

$$S = S^*(t)(1 + \delta x).$$

Then Taylor-expanding $W(S,t)$ about $S = S^*(t)$ and writing the result in terms of x gives the outer expansion expanded in inner variables as

$$\begin{aligned} W \sim \quad & W_0^*(t) \\ & + \delta\left(xS^*(t)W_{0S}^*(t) + W_1^*(t)\right) \\ & + \delta^2\left(\frac{1}{2}x^2 S^*(t)^2 W_{0SS}^*(t) + xS^*(t)W_{1S}^*(t) + W_2^*(t)\right) \\ & + \cdots , \end{aligned} \tag{3.3}$$

where

$$W_0^*(t) = W_0(S^*(t),t), \qquad W_{0S}^*(t) = \frac{\partial W_0}{\partial S}(S^*(t),t) \quad \text{etc.}$$

are functions of t alone and as-yet unknown. Hence, in the absence of spatial boundary conditions, we can do no more in this region for now.

3.1.2. *Outer region: $S < S^*(t)$ ('exercise').* This is the region below the exercise point. The penalty equation becomes

$$\mathcal{L}_{\mathrm{BS}} W^\epsilon = rK + \frac{\sigma^2}{\delta^2} W^\epsilon .$$

Simple iteration gives that

$$W^\epsilon \sim -\delta^2 rK/\sigma^2 + O(\delta^4). \tag{3.4}$$

This dictates the scaling in the inner region.

3.1.3. *Inner region.* Make the change of variables to x and t, and expand

$$W^\epsilon(S,t) = w^\epsilon(x,t) \tag{3.5}$$

$$\sim w_0(x,t) + \delta w_1(x,t) + \delta^2 w_2(x,t) + \cdots . \tag{3.6}$$

The penalty equation becomes

$$\begin{aligned} & \frac{\partial w^\epsilon}{\partial t} - \frac{\dot{S}^*}{\delta S^*}(1+\delta x)\frac{\partial w^\epsilon}{\partial x} + \frac{1}{2}\sigma^2 \frac{(1+\delta x)^2}{\delta^2}\frac{\partial^2 w^\epsilon}{\partial x^2} + r\frac{1+\delta x}{\delta}\frac{\partial w^\epsilon}{\partial x} - rw^\epsilon \\ & \qquad = rK + \begin{cases} 0 & x > x^*, \\ \frac{\sigma^2}{\delta^2} w^\epsilon & x < x^*, \end{cases} \end{aligned} \tag{3.7}$$

where $\dot{S}^* = \mathrm{d}S^*/\mathrm{d}t$, and with $w^\epsilon = 0$ and $\partial w^\epsilon/\partial x$ continuous at $x = x^*$ (i.e., x^* is the crossing point of the penalty solution).

3.1.4. *Matching.* As $x \to -\infty$, we have (cf. (3.4))

$$w_0 \to 0, \qquad w_1 \to 0, \qquad w_2 \to -rK/\sigma^2,$$

and as $x \to +\infty$ we have (compare (3.3))

$$\begin{aligned} w_0(x,t) &\sim W_0^*(t) + o(1), \\ w_1(x,t) &\sim xS^*(t)W_{0S}^*(t) + W_1^*(t) + o(1), \\ w_2(x,t) &\sim \frac{1}{2}x^2 S^*(t)^2 W_{0SS}^*(t) + xS^*(t)W_{1S}^*(t) + W_2^*(t) + o(1). \end{aligned}$$

The largest terms in (3.7) are $O(1/\delta^2)$. When we substitute the expansion (3.6) in and collect terms, we get, at $O(1/\delta^2)$,

$$\tfrac{1}{2}\sigma^2 \frac{\partial^2 w_0}{\partial x^2} = \begin{cases} 0 & x > x^*, \\ \sigma^2 w_0 & x < x^*, \end{cases}$$

and the only solution that vanishes at $x = -\infty$, has continuous first derivative at $x = x^*$, and tends to a constant at $x = +\infty$, is $w_0(x, t) \equiv 0$. This tells us that

$$W_0^*(t) = 0$$

as expected. Because $W_0(S, t)$ is the difference between the vanilla value of the put and the payoff, its S-derivative vanishes at $S = S^*(t)$ — this is smooth pasting. (In more detail, because $W_0(S, t)$ has the right value at $S = S^*(t)$ and the right payoff, uniqueness for solutions of the BSPDE in a parabolic domain tells us that it *is* the vanilla put value.) Hence, $W_{0S}^*(t) = 0$. Now, at $O(1/\delta)$ in (3.7), we get

$$\tfrac{1}{2}\sigma^2 \frac{\partial^2 w_1}{\partial x^2} = \begin{cases} 0 & x > x^*, \\ \sigma^2 w_1 & x < x^*. \end{cases}$$

As $W_{0S}^*(t) = 0$, $w_1(x, t)$ has no linear term at $x = +\infty$, and so, by the same argument as above, it vanishes too, confirming that the inner scaling for W is indeed $O(\delta^2)$. Hence, $W_1^*(t) = 0$, and we can return to the outer region $S > S^*(t)$ to show that $W_1(S, t) \equiv 0$ (zero payoff, zero value on $S = S^*(t)$). Now we come to the first non-trivial term. At $O(1)$ in (3.7), we have

$$\tfrac{1}{2}\sigma^2 \frac{\partial^2 w_2}{\partial x^2} = rK + \begin{cases} 0 & x > x^*, \\ \sigma^2 w_0 & x < x^*. \end{cases}$$

For $x < x^*$, the solution that tends to $-rK/\sigma^2$ at $-\infty$ and vanishes at $x = x^*$ is

$$w_2^-(x, t) = rK\left(\mathrm{e}^{\sqrt{2}(x - x^*(t))} - 1\right)/\sigma^2. \tag{3.8}$$

For $x > x^*$, the solution that vanishes at $x = x^*$ and whose derivative matches (3.8) is

$$w_2^+(x, t) = \frac{rK}{\sigma^2}(x - x^*(t))^2 + rK\sqrt{2}/\sigma^2(x - x^*(t)). \tag{3.9}$$

Now comes the key point. From the matching, we now know that

$$w_2^+(x, t) \sim \tfrac{1}{2}x^2 S^*(t)^2 W_{0SS}^*(t) + W_2^*(t), \qquad x \to \infty.$$

There is no linear term because $W_1(S, t) = 0$. Comparing with (3.9), we find that

$$rK/\sigma^2 = \frac{1}{2}S^*(t)^2 W_{0SS}^*(t) \qquad \text{(coefficient of } x^2\text{)},$$
$$-2rKx^*(t)/\sigma^2 + rK\sqrt{2}/\sigma^2 = 0 \qquad \text{(coefficient of } x\text{)},$$
$$rKx^*(t)^2/\sigma^2 - \sqrt{2}rKx^*(t)/\sigma^2 = W_2^*(t) \qquad \text{(constant coefficient)}.$$

The first of these confirms the boundary Gamma of the vanilla put. The second gives

$$x^*(t) = 1/\sqrt{2}.$$

The third gives

$$W_2^*(t) = -\tfrac{1}{2}rK/\sigma^2.$$

Penalisation Error	Computed	Predicted	Relative Difference
Value in Exercise Region	4.9975e-04	5.0000e-04	-5.0075e-04
Value in Hold Region	2.5070e-04	2.5000e-04	0.0028
Exercise Boundary	0.0174	0.0165	0.0516

Table 2. *For the American put, the maximum penalisation error in exercise and hold regions separately, and the error in the exercise boundary, for* $\epsilon^{-1} = 100$, $\sigma = 0.4$, $r = 0.05$, $K = 100$, $T = 1$, *hence* $\delta = 0.04$. *The numerical result is compared with the first order correction from the asymptotic analysis.*

In original variables, the crossing point is at

$$\begin{aligned} S &= S^*(t)(1 + \delta x^*(t)) \\ &= S^*(t)\left(1 + \sqrt{\epsilon}\sigma/\sqrt{2}\right) \end{aligned}$$

as $\delta^2 = \sigma^2\epsilon$, and the boundary value of the correction is

$$\delta^2 W_2^*(t) = -\tfrac{1}{2} rK\epsilon$$

(independently of σ).

The correction is interpretable as the value of an option with zero payoff at maturity which pays a fixed amount when (if) the stock crosses $S^*(t)$ from above. Interestingly, the continuity correction for a Bermudan option has the same boundary value if we set $\delta^2 = \sigma^2\Delta T/2$, where ΔT is the interval between exercise dates [13].

3.1.5. *Numerical verification.* Table 2 compares the corrections based on W_2^* and x^* against the numerically computed penalisation error and finds excellent agreement.

3.2. **Butterfly Spread.** Still in the Black-Scholes framework without dividends, consider now a butterfly spread with payoff

$$\begin{cases} \max(V_0 + \alpha_1(S - K), 0), & S < K, \\ \max(V_0 + \alpha_2(K - S), 0), & S > K, \end{cases}$$

where $V_0, \alpha_i > 0$. Denote again the penalty value by $V^\epsilon(S,t)$, the true value $B(S,t)$. Consider the situation where $B(S,t)$ has a free boundary on the put-like bit of the payoff ($S > K$) but on the call-like bit, the value of $B(K,t)$ is anchored to V_0. This certainly happens for short times before expiry (as the Black-Scholes operator on the payoff is positive). The former is analysed as before so we focus on the region around the convex kink at $S = K$.

The inner variable is now $S = K(1 + \delta x)$, the outer solution for $S < K$ is of the form

$$V^\epsilon(S,t) = B(S,t) + \delta V_1(S,t) + \cdots$$

and its inner expansion near $S = K$ is

$$V_0 + \delta(xB_S^*(t) + V_1^*(t)) + \cdots,$$

where $B_S^*(t) = \partial B/\partial S\,(K,t)$ and $V_1^*(t) = V_1(K,t)$. The payoff in inner variables is

$$\begin{cases} V_0 + \delta\alpha_1 Kx, & x < 0, \\ V_0 - \delta\alpha_2 Kx, & x > 0. \end{cases}$$

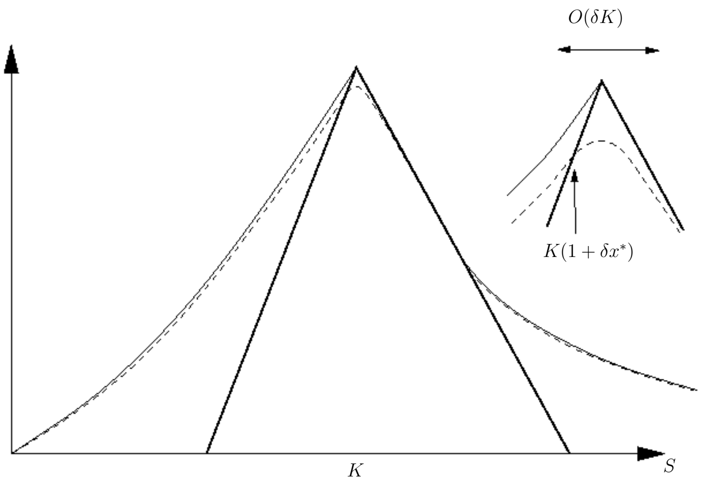

Figure 9. *Schematic of a butterfly option solution and blow-up of the inner region near the peak. The true solution is the solid curve, the penalty solution is dashed.*

This suggests that the inner solution is of size $O(\delta)$, not $O(\delta^2)$, and this is consistent with the left-hand outer solution meeting the payoff at an angle (not smooth pasting). Write the inner expansion in the form

$$V^\epsilon(S,t) \sim V_0 + \delta v_1(x,t) + \cdots .$$

Also let x^* (which is negative) be the point at which the penalty solution crosses the payoff. The leading order inner equation is

$$\tfrac{1}{2}\sigma^2 \frac{\partial^2 v_1}{\partial x^2} = \sigma^2 \begin{cases} v_1 - \alpha_1 K x, & x < 0, \\ v_1 + \alpha_2 K x, & x > 0. \end{cases}$$

The solution is C^1 at both $x = 0$ and $x = x^*$, where C^1 is the space of continuously differentiable functions. As $x \to \infty$, $v_1 \sim -\alpha_2 K x$ because the solution is $O(\delta^2)$ in the 'exercise' region to the right of $S = K$. So,

$$v_1 = \begin{cases} \alpha_1 K x + a(t) K \cosh(x\sqrt{2}) + b(t) K \sinh(x\sqrt{2}), & x < 0, \\ -\alpha_2 K x + a(t) K \mathrm{e}^{-x\sqrt{2}}, & x > 0, \end{cases}$$

for some $a(t)$, $b(t)$. This is continuous at $x = 0$, and continuity of $\partial v_1/\partial x$ at $x = 0$ gives

$$\alpha_1 + b(t)\sqrt{2} = -\alpha_2 - a(t)\sqrt{2}. \tag{3.10}$$

Now, at $x = x^*$, v_1 meets the payoff and joins onto the outer solution:

$$v_1(x^*(t), t) = \alpha_1 K x^*(t), \qquad \frac{\partial v_1}{\partial x} = B^*_S(t),$$

from which

$$\alpha_1 x^*(t) + a(t)\cosh(x^*(t)\sqrt{2}) + b(t)\sinh(x^*(t)\sqrt{2}) = \alpha_1 x^*(t), \tag{3.11}$$

$$\alpha_1 + \sqrt{2}\left(a(t)\sinh(x^*(t)\sqrt{2}) + b(t)\cosh(x^*(t)\sqrt{2})\right) = B^*_S(t). \tag{3.12}$$

This, with (3.10), is three equations for $a(t)$, $b(t)$ and $x^*(t)$. From these, we readily find that

$$x^*(t) = -\frac{1}{\sqrt{2}} \log \frac{\alpha_1 + \alpha_2}{\alpha_1 - B_S^*(t)}.$$

This is clearly negative since we have $0 < B_S^* < \alpha_1$. Also, it tends to $-\infty$ as $B_S^* \to \alpha_1$ from below – that is, as (if) the free boundary moves away from $S = K$.

We also derive the value $V_1(K,t) = \alpha_1 K x^*(t)$, which lets us compute the correction term, although not explicitly because of the complicated time dependence.

3.3. **American Put under Jump-Diffusion.** We now extend the analysis under Black–Scholes to include jumps of relative size J, at the jump time of a compound Poisson process with rate λ. We only do this for the standard put payoff to illustrate the extensions over Black-Scholes. The results are qualitatively very similar to the Black-Scholes case and the analysis suggest this will also be the case for other payoffs. In addition, we also account for continuously paid proportional dividends of rate q, where we assume $q \le r$. (The solution for $q > r$ is qualitatively different.)

The penalised equation is defined as in (3.1), specifically, for the put under a jump model

$$\mathcal{L}_{\text{BSJ}} V^\epsilon = -\frac{1}{\epsilon} \max(K - S - V^\epsilon, 0) = -\frac{\sigma^2}{\delta^2} \max(K - S - V^\epsilon, 0), \tag{3.13}$$

where $\mathcal{L}_{\text{BSJ}}$ is defined in (1.7). Smooth pasting still holds for the vanilla American put value $P(S,t)$, and applying these conditions just to the right of the exercise boundary now gives the boundary Gamma as

$$\left.\frac{\partial^2 P}{\partial S^2}\right|_{S=S^*(t)} = \frac{2}{\sigma^2 S^*(t)^2} \left(rK - (q + \omega\lambda) S^*(t) - \lambda \mathbb{E}[P(JS^*(t), t) - P(S^*(t), t)] \right) =: \frac{2}{S^*(t)^2} \Gamma^*(t). \tag{3.14}$$

Note that $\Gamma^*(t)$ depends on $P(S,t)$ for all $S > 0$ via the term $\mathbb{E}[P(JS^*(t),t) - P(S^*(t),t)]$.

3.3.1. *Outer region $S > S^*(t)$ ('hold').* It will be useful again to introduce $W^\epsilon(S,t)$ as in (3.2). As $S\mathrm{e}^{-q(T-t)}$ and $K\mathrm{e}^{-r(T-t)}$ satisfy $\mathcal{L}_{BSJ} V = 0$ individually (in fact their difference is the value of a forward), one gets

$$\mathcal{L}_{\text{BSJ}} V^\epsilon(S,t) = \mathcal{L}_{\text{BSJ}} W^\epsilon(S,t) - \mathcal{L}_{\text{BSJ}}(S - K) = \mathcal{L}_{\text{BSJ}} W^\epsilon(S,t) - qS + rK.$$

3.3.2. *Outer region: $S < S^*(t)$ ('exercise').* With the above substitution, as $W(S,t) \le 0$ in this region by assumption,

$$\mathcal{L}_{\text{BSJ}} W^\epsilon(S,t) = rK - qS + \frac{\sigma^2}{\delta^2} W^\epsilon(S,t).$$

Inserting the expansion for $W^\epsilon(S,t)$, $W_0(S,t)$ and $W_1(S,t)$ vanish in the outer region $S < S^*(t)$ as before; however, when determining $W_2(S,t)$ in the 'exercise' region, we have to account for jumps into the other regions, particularly into the outer 'hold' region $S > S^*(t)$, where $W^\epsilon(S,t)$ will not be small. Thus, at $O(1)$,

$$\lambda \mathbb{E}[W_0(JS,t) - W_0(S,t)] = rK - qS + \sigma^2 W_2(S,t),$$

and, solving for $W_2(S,t)$,

$$\begin{aligned} W_2(S,t) &= -\left((rK - qS) - \lambda \mathbb{E}[P(SJ,t) - (K - JS)]\right)/\sigma^2 \\ &= -\left((rK - qS) - \lambda \mathbb{E}[P(SJ,t) - P(S,t) + (J-1)S]\right)/\sigma^2. \end{aligned}$$

It is seen from the above equations that $W_2(S,t)$ is negative and inreases with increasing S, from $W_2(0,t) = -rK/\sigma^2$ at the origin, i.e. the penalisation error is largest for $S = 0$, where it has a simple explicit value. This also follows from a maximum principle argument, as we will see in Section 4.

3.3.3. *Inner region.* Similar to (3.7), we now have

$$\begin{aligned}\frac{\partial w^\epsilon}{\partial t} - \frac{\dot{S}^*}{\delta S^*}(1+\delta x)\frac{\partial w^\epsilon}{\partial x} + \frac{1}{2}\sigma^2\frac{(1+\delta x)^2}{\delta^2}\frac{\partial^2 w^\epsilon}{\partial x^2} + r\frac{1+\delta x}{\delta}\frac{\partial w^\epsilon}{\partial x} - rw^\epsilon \\ +\ \mathbb{E}[W^\epsilon(JS^*(1+\delta x),t) - W^\epsilon(S^*(1+\delta x),t)] \\ = rK - qS^*(1+\delta x) + \begin{cases} 0 & x > x^*, \\ \frac{\sigma^2}{\delta^2}w^\epsilon & x < x^*. \end{cases}\end{aligned} \tag{3.15}$$

The non-local term is written in terms of the outer solutions because it acts on the scale of the outer variables. Writing the expectation term as integral and expanding in δ gives, at leading order,

$$\delta S^* x \int_0^\infty J\, W_S^\epsilon(JS^*,t)\, g(J)\, dJ.$$

The simple expansion has a natural interpretation: given a jump size, all jumps starting from the inner region and ending in the outer region end up close to each other. If we were going to a higher order of accuracy (which we are not), we would have to treat the small jumps – those which both start and end in the inner region – separately. So the integral for the expectation would have its range split into inner and outer parts, and so on.

Comparing terms $O(1/\delta^2)$ and $O(1/\delta)$ gives again that $w_0(x,t)$ and $w_1(x,t)$ vanish, and now, at $O(1)$,

$$\tfrac{1}{2}\sigma^2\frac{\partial^2 w_2}{\partial x^2} = \sigma^2\Gamma^*(t) + \begin{cases} 0 & x > x^*, \\ \sigma^2 w_2 & x < x^*. \end{cases}$$

3.3.4. *Matching.* First, we match the inner solution with the outer solution in the exercise region. The matching of w_2^- for $x \to -\infty$ is now to a non-constant value, but it is clear that $W_2(S,t)$ from (3.15) approaches $\Gamma^*(t)$ for $S \to S^*$ in the 'outer' variables, and matching in an overlap region demands that, as $x \to -\infty$, $w_2(x,t) \to -\Gamma^*(t)$. Then calculations identical to before give, for $x < x^*$,

$$w_2^-(x,t) = \Gamma^*(t)\left(\mathrm{e}^{\sqrt{2}(x-x^*(t))} - 1\right),$$

and, for $x > x^*$,

$$w_2^+(x,t) = \Gamma^*(t)\left((x-x^*(t))^2 + \sqrt{2}(x-x^*(t))\right). \tag{3.16}$$

Matching with the outer region $S > S^*$ as before gives

$$\begin{aligned}\Gamma^*(t) &= \frac{1}{2}S^*(t)^2 W^*_{0SS}(t), \\ \Gamma^*(t)(-2x^*(t) + \sqrt{2}) &= 0, \\ \Gamma^*(t)(x^*(t)^2 - \sqrt{2}x^*(t)) &= W_2^*(t).\end{aligned}$$

The first equation recovers the jump diffusion gamma from earlier. Interestingly, the relative position x^* of the penalty crossing point in relation to the exercise boundary, which is given by the second equation, is unaffected by the jumps. The last equation, upon inserting x^*, shows again that the penalisation error at the free boundary is half the value one would get by extrapolation from the outer exercise region.

4. General Upper and Lower Value Bounds

In the previous section, we computed the penalisation error to leading order in the penalty parameter, and noted a distinct difference in the convergence order for the put, $O(\epsilon)$, and butterfly payoffs, $O(\epsilon^{1/2})$. We now show that a distinction into categories of piecewise smooth payoffs with convex and non-convex kinks allows us to derive general upper and lower bounds on the value function. We work with jump-diffusion models.

4.1. **A Maximum Principle Argument.** Considering the penalised equation

$$-\mathcal{L}_{\mathrm{BSJ}} V^\epsilon = \frac{1}{\epsilon}\max(\Psi - V^\epsilon, 0), \tag{4.1}$$

it is automatically true that $-\mathcal{L}_{\mathrm{BSJ}} V^\epsilon \geq 0$, and if (where) $V^\epsilon > \Psi$, then $\mathcal{L}_{\mathrm{BSJ}} V^\epsilon = 0$, such that a complementarity condition is satisfied and $\min(-\mathcal{L}_{\mathrm{BSJ}} V^\epsilon, V^\epsilon - \Psi) \leq 0$. Hence, V^ϵ only fails to be an exact solution to

$$\min(-\mathcal{L}_{\mathrm{BSJ}} V^\epsilon, V^\epsilon - \Psi) \leq 0$$

where $V^\epsilon \geq \Psi$ is violated.

We begin with an elementary analysis of this inequality constraint. To this end, consider $W^\epsilon = V^\epsilon - \Psi$. The biggest violation of $V^\epsilon \geq \Psi$ is given at a global negative minimum of W^ϵ (if one is attained). Note that, for $t < T$, the solution V^ϵ to (4.1) is twice continuously differentiable everywhere in S, by standard regularity arguments. We first consider points at which Ψ is also smooth, i.e. excluding the kinks. Then, at any such negative minimum $S = S^*$ of W^ϵ, by inspection of the individual terms,

$$\begin{aligned}\mathcal{L}_{\mathrm{BSJ}} W^\epsilon = \frac{\partial W^\epsilon}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 W^\epsilon}{\partial S^2} + (r - q - \omega\lambda) S \frac{\partial W^\epsilon}{\partial S} - rW^\epsilon \\ + \lambda \mathbb{E}[W^\epsilon(JS,t) - W^\epsilon(S,t)] > 0.\end{aligned} \tag{4.2}$$

From

$$-\frac{1}{\epsilon}(\Psi - V^\epsilon) = \mathcal{L}_{\mathrm{BSJ}} V^\epsilon = \mathcal{L}_{\mathrm{BSJ}} W^\epsilon + \mathcal{L}_{\mathrm{BSJ}} \Psi\,,$$

it follows that

$$V^\epsilon > \Psi + \epsilon\, \mathcal{L}_{\mathrm{BSJ}} \Psi. \tag{4.3}$$

For piecewise linear payoffs Ψ, it is straighforward to show that, again excluding kinks, $\mathcal{L}_{\mathrm{BSJ}}\Psi$ is bounded from below, uniformly for all S. Also, $W^\epsilon = V^\epsilon - \Psi$ does not have any negative minima at convex kinks of Ψ, i.e. points with $\Psi'(S-) < \Psi'(S+)$.

Summarising, the biggest violation of the inequality $V^\epsilon \geq \Psi$ is either bounded by (4.3), taken over the smooth intervals of Ψ, or at concave kinks, or at one of the boundary $S = 0$ or $S \to \infty$. We will come back to this observation later to obtain easily computable bounds on the solution.

4.2. **Constructing Bounds on the Value Function.** To treat the solution uniformly in the hold and exercise regions, inclusive of kinks, it is convenient to work in the framework of viscosity solutions. We use the equation

$$\min(-\mathcal{L}u, u - \psi) = 0 \tag{4.4}$$

in log coordinates on $\mathbb{R}$, with $\mathcal{L}$ as in (1.8), $\psi(x) = \Psi(S) = \Psi(S_0 \exp(x))$, and its penalised version

$$-\mathcal{L}u^\epsilon = \frac{1}{\epsilon}\max(\psi - u^\epsilon, 0). \tag{4.5}$$

This avoids technicalities of boundary conditions and discontinuous viscosity solutions, and we can use the definition from [22], which we tailor to our setting for convenience:

Definition 4.1 (Viscosity Solution). *$u \in C([0,T]\times\mathbb{R})$ is a* viscosity supersolution (subsolution) *of (4.4), if*

$$\min(-(\mathcal{L}\phi)(x,t), \phi(x,t) - \psi(x)) \geq 0 \quad (\leq 0)$$

whenever $\phi \in C^2([0,T]\times\mathbb{R}) \cap C_2([0,T]\times\mathbb{R})$ and $u - \phi$ has a global minimum (maximum) at $(x,t) \in [0,T) \times \mathbb{R}$ with $v(x,t) = \phi(x,t)$. u is a viscosity solution *iff it is a super- and subsolution.*

Here, C^2 is the space of twice continuously differentiable functions, and C_2 the space of continuous functions with at most quadratic growth at ∞. This includes the put and butterfly payoffs, but *not* the call payoff in log-coordinates. It is clear though that the results can be extended (e.g. by a coordinate transformation identical to the logarithm for small values, the identity for large values, and a smoothly increasing transition in between). We further assume that the density ν gives bounded third moments.

Pham [22] shows that under these conditions (1.2) satisfies a comparison principle.

Theorem 4.2 (Theorem 4.1 in [22]). *If u and v are uniformly continuous sub- and supersolutions of (1.2) respectively, and $u(x,T) \leq v(x,T)$ for all x, then $u \leq v$ everywhere.*

It is clear that u^ϵ is a classical subsolution of (4.4),

$$\min(-\mathcal{L}u^\epsilon, u^\epsilon - \psi) = \min(\epsilon \max(\psi - u^\epsilon, 0), u^\epsilon - \psi) \leq 0,$$

and therefore also a viscosity subsolution, thus $u^\epsilon \leq u$ is a lower bound for the true solution.

We now seek to construct an upper bound by setting

$$\overline{u}^\epsilon := u^\epsilon + \lambda^\epsilon, \tag{4.6}$$

$$\lambda^\epsilon := \min\{\lambda \in \mathbb{R}, \lambda \geq 0 : u^\epsilon + \lambda \geq \psi\} = \max\{(\psi - u^\epsilon)^+\}. \tag{4.7}$$

Indeed, $\overline{u}^\epsilon$ is a (classical and viscosity) supersolution of (4.4),

$$\min(-\mathcal{L}\overline{u}^\epsilon, \overline{u}^\epsilon - \psi) = \min(r\lambda^\epsilon + \epsilon\max(\psi - u^\epsilon, 0), \overline{u}^\epsilon - \psi) \geq 0,$$

and therefore $\overline{u}^\epsilon \geq u$.

This is closely related to the regularised Lagrange multiplier approximation in [14], who propose to solve

$$-\mathcal{L}u = \max\left((\psi - u)/\epsilon + \bar{\lambda}, 0\right)$$

for some fixed function $\bar{\lambda} > 0$ large enough to make the solution feasible. This is possible for sufficiently regular obstacles, i.e. for convex kinks in our framework, as it essentially

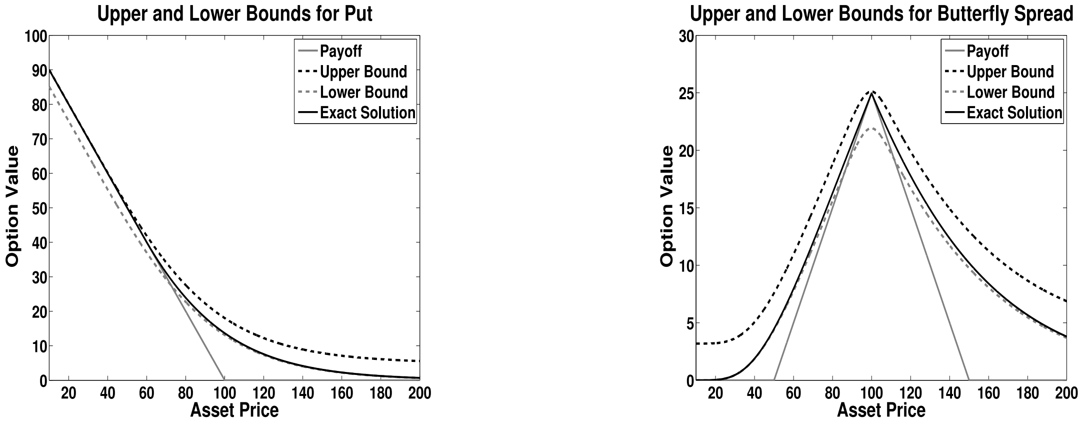

Figure 10. *Illustration of the lower bound V^ϵ, upper bound $\overline{V}^\epsilon = V^\epsilon + \max\{(\Psi - V^\epsilon)^+\}$, payoff Ψ, and true value function V for the American put (left) and the American butterfly spread (right), for $\epsilon = 100$ (put) and $\epsilon = 0.00005$ (butterfly). We are thinking of ϵ as a* small *number, however for very small values, the bounds for the put become optically indistinguishable from the solution.*

corresponds to $\lambda^\epsilon = \epsilon\bar{\lambda} = O(\epsilon)$. In fact, in the following section, we show that the order of λ^ϵ is either $O(\epsilon)$ (no 'active' concave kinks) or $O(\epsilon^{1/2})$ ('active' concave kinks), as expected from the asymptotic expansions.

From (4.3) and the discussion thereafter, we know that λ^ϵ can be estimated from (4.7) by using the right-hand side from (4.3) and values at concave kinks and boundaries. We can use this fact to compute simple lower and upper bounds, which converge to the true solution. Figure 10 illustrates this for the put and butterfly. Note the different magnitude of the penalty parameter required for the put and butterfly to achieve similar accuracy.

4.3. **Convergence Rates and Further Properties.** The following results follow directly from the comparison principle.

Lemma 4.3. *Let ψ and $\widehat{\psi}$ be as before. Denote by u and $\widehat{u}$ the solutions to Problem 4.4 that correspond to ψ and $\widehat{\psi}$, and, similarly, denote by u^ϵ and $\widehat{u}^\epsilon$ the corresponding solutions to Problem 4.5. If $\psi \leq \widehat{\psi}$ everywhere, then we have*

$$u^\epsilon \leq \widehat{u}^\epsilon \quad \textit{and} \quad u \leq \widehat{u}.$$

Trivially, u^ϵ and u are nonnegative if $\psi \geq 0$. Moreover, denote by u^{ϵ_1} and u^{ϵ_2} the solutions to (4.5) corresponding to penalty parameters $\epsilon_1 > \epsilon_2 > 0$, respectively. Then $u^{\epsilon_1} \geq u^{\epsilon_2}$.

Proof. For the first part, consider

$$\min(-\mathcal{L}u, \phi - u) \leq \min(-\mathcal{L}u, \widehat{\phi} - u)$$

and

$$-\mathcal{L}u^\epsilon = \frac{1}{\epsilon}\max(\psi - u^\epsilon, 0) \leq \frac{1}{\epsilon}\max(\widehat{\psi} - u^\epsilon, 0),$$

such that u and u^ϵ are subsolutions to their governing equations with ψ replaced by $\widehat{\psi}$. Similarly, for the second part, apply the same argument to

$$-\mathcal{L}u^{\epsilon_1} = \frac{1}{\epsilon_1}\max(\psi - u^{\epsilon_1}, 0) \le \frac{1}{\epsilon_2}\max(\psi - u^{\epsilon_1}, 0).$$

□

We can apply the framework of [15], pp. 4–8, to estimate λ^ϵ.

Theorem 4.4. *If ψ is Lipschitz continuous and piecewise C^1 with linear growth and*

(1) *convex kinks, then*

$$0 \le u - u^\epsilon \le C\epsilon;$$

(2) *concave kinks, then*

$$0 \le u - u^\epsilon \le C\epsilon^{1/2}.$$

Proof. This follows precisely the steps in the proof of Theorem 2.1 in [15]. Although the context there is that of non-linear PDEs, the results are sufficiently abstract to accommodate PIDEs given a comparison principle. The main steps are based on smoothing the payoff with mollifiers, and bounding the approximation error in the two cases. □

5. Solution of a Variational Formulation

From the previous section, we know $\max_x |u(x,t) - u^\epsilon(x,t)| = O(\epsilon)$ for payoffs with convex kinks. Combined with the differentiability of u^ϵ w.r.t. x, where the size of the derivative is independent of ϵ ,

$$\frac{\partial u}{\partial x}(x,t) = \frac{u(x+\epsilon^{1/2}) - u(x,t)}{\epsilon^{1/2}} + O(\epsilon^{1/2}) = \frac{u^\epsilon(x+\epsilon^{1/2}) - u^\epsilon(x,t)}{\epsilon^{1/2}} + O(\epsilon^{1/2})$$

allows us to estimate the derivative up to $\epsilon^{1/2}$ by a finite difference, which naturally 'regularises' the differentiation. We know e.g. for the put from the asymptotic expansion that convergence will be better behaved everywhere except in a small neighbourhood (of width $\epsilon^{1/2}$) of the exercise boundary. For non-convex kinks, convergence will be slower.

This section develops estimates of the penalisation error for the derivative directly, via analysis in the H^1 norm. We follow here the set-up of [34], who show convergence of penalisation in jump-diffusion models, but do not derive convergence orders.

5.1. **Set-up.** We study problems (4.4) and (4.5), but on a localised domain $\Omega := \{x \in \mathbb{R} : |x| < l\}$ with boundary $\partial\Omega := \{x \in \mathbb{R} : |x| = l\}$. It would be possible to work on $\mathbb{R}$, but this would require us to introduces weighted norms to be able to deal with functions that do not decay (sufficiently fast) for large x (such that their Sobolev norms are well defined), making the variational formulation more cumbersome to write out. Instead, on the finite domain, we can use the standard (separable Hilbert) spaces $H := L^2(\Omega)$ and $V := \{u \in H : \partial u/\partial x \in H\} = H^1(\Omega)$ [2]. For $\Sigma \in \{H\,, V\}$, define $L^2(0,T;\Sigma)$ as the (separable Hilbert) spaces of measurable functions $u : [0,T] \to \Sigma$ satisfying $u(\cdot,t) \in \Sigma$ for almost every $t \in [0,T]$ and for which $\int_0^T |u(\cdot,t)|_\Sigma^2 \, dt < \infty$, equipped with their canonical inner products (cf. [19]). The (Banach) space $L^\infty(0,T;\Sigma)$ is defined to contain all measurable functions $u : [0,T] \to \Sigma$ satisfying $u(\cdot,t) \in \Sigma$ for almost every $t \in [0,T]$ and for which $|u|_{L^\infty(0,T;\Sigma)} := \operatorname{ess\,sup}_{t\in[0,T]} |u(\cdot,t)|_\Sigma < \infty$ (cf. [19]).

Now, for $u \in H$ and $x \in \Omega$, define the jump operator

$$(Bu)(x) := \lambda\Big[\int_{z,z+x\in\Omega} u(x+z)\ \nu(z)\ dz - u(x)\Big].$$

For $u,\, v \in V$, define the bilinear forms

$$a(u,v) := \frac{\sigma^2}{2}\int_\Omega \frac{\partial u}{\partial x}\frac{\partial v}{\partial x}\ dx + \int_\Omega ruv\ dx - \int_\Omega \Big(\mu - \frac{\sigma^2}{2}\Big)\frac{\partial u}{\partial x} v\ dx,$$

where $\mu = r - q - \omega\lambda$ and

$$b(u,v) := -\int_\Omega (Bu)v\ dx.$$

We assume in the following that $\psi : \mathbb{R} \to \mathbb{R}$ is continuous and there exists a constant $M > 0$ such that

$$|\psi(x)| \le Me^{M|x|}, \quad x \in \mathbb{R}, \tag{5.1}$$

then set

$$f(x) := \lambda \int_{z,z+x\notin\Omega} \psi(x+z)\ \nu(z)\ dz, \quad x \in \Omega$$

and assume that $f \in H$.

The following two lemmas are taken from [33], to where we also refer for the proof of the subsequent theorem.

Lemma 5.1. *We have $B \in \mathcal{L}(H, H)$, i.e. $B : H \to H$ is a bounded linear operator.*

Proof. See [33]. □

Lemma 5.2. *There exist constants $\vartheta,\, \xi > 0$ such that*

$$a(u,u) + b(u,u) \ge \vartheta|u|_V^2 - \xi|u|_H^2, \quad u \in V.$$

Proof. See [33]. □

We are now in a position to formulate the variational inequality the solution of which is the value function of the American option.

Problem 5.3. *Find a function $u \in L^2\big(0,T;V\big)$, $\partial u/\partial t \in L^2\big(0,T;H\big)$, such that*

$$u(\cdot,T) = \psi,\ u(x,\cdot) = \psi(x) \text{ for } x \in \partial\Omega,\ u \ge \psi \text{ a.e. on } \Omega\times[0,T]$$

and, a.e. on $[0,T]$, it is

$$-\Big(\frac{\partial u}{\partial t}, v - u\Big) + a(u, v-u) + b(u, v-u) - (f, v-u) \ge 0 \tag{5.2}$$

for all $v \in V$ with $v \ge \psi$.

We emphasise that the payoff function ψ is not confined to the interval $[-l, l]$ and that the function f is given by a nonlocal integral.

Theorem 5.4. *There exists a unique solution u to Problem 5.3 with $u \in L^\infty(0,T;V)$.*

Proof. See [33] or [34]. □

5.2. **Penalisation and Basic Properties.**

Problem 5.5. *Let $\epsilon > 0$ and define $\beta(\cdot) := -(\psi - \cdot)^+$. Find a function $u^\epsilon \in L^2\big(0,T;V\big)$, $\partial u^\epsilon/\partial t \in L^2\big(0,T;H\big)$, such that*

$$u^\epsilon(\cdot,T) = \psi,\ \ u^\epsilon(x,\cdot) = \psi(x) \text{ for } x \in \partial\Omega$$

and, a.e. on $[0,T]$, we have

$$-\Big(\frac{\partial u^\epsilon}{\partial t}, v\Big) + a(u^\epsilon, v) + b(u^\epsilon, v) - (f,v) + \frac{1}{\epsilon}\big(\beta(u^\epsilon), v\big) = 0 \tag{5.3}$$

for all $v \in V$.

Theorem 5.6. *There exists a unique solution u^ϵ to Problem 5.5. Furthermore, there exists a constant $C > 0$, C independent of ϵ, such that*

$$|u^\epsilon|_{L^\infty(0,T;V)} + \frac{1}{\epsilon^{1/2}}|\beta(u^\epsilon)|_{L^2(0,T;H)} + \Big|\frac{\partial u^\epsilon}{\partial t}\Big|_{L^2(0,T;H)} \le C. \tag{5.4}$$

As $\epsilon \to 0$, we have that $u^\epsilon \to u$ strongly and $\partial u^\epsilon/\partial t \to \partial u/\partial t$ weakly in $L^2(0,T;H)$, where u is the solution to Problem 5.3.

Proof. A result of this form comes up naturally when using penalisation to prove the existence of a solution to a variational inequality. In this particular case, it can be directly obtained by adapting the proof given in [33] for the non-localised problem. Alternatively, one can slightly extend a similar result given in [4]. □

Remark 5.7. *The results of Lemma 4.3 still hold for the localised variational problem, but we omit the proof of this.*

5.3. **The American Put and Other Payoffs with Convex Kinks.** The following result is an extension of the one in [4] to jump diffusion, and to accommodate kinks. In the proof, we work with weak coercivity instead of coercivity, and account for the loss of regularity at kinks explicitly.

Theorem 5.8. *Consider an American put option, i.e. let ψ be given by*

$$\psi(x) = (K - e^x)^+, \quad x \in \mathbb{R},$$

and suppose $f \in H$. There exists a constant $C > 0$, C independent of ϵ, such that

$$|u - u^\epsilon|_{L^2(0,T;V)} + |u - u^\epsilon|_{L^\infty(0,T;H)} \le \epsilon^{1/2} C.$$

Proof. Again, we extend a proof that was given, for standard parabolic variational inequalities in H_0^1, in [4]. All constants C_i, with i an integer, are taken to be independent of ϵ and t. Plugging $-\beta(u^\epsilon) \in V$ into (5.3) gives

$$\Big(\frac{\partial u^\epsilon}{\partial t}, \beta(u^\epsilon)\Big) - a\big(u^\epsilon, \beta(u^\epsilon)\big) - b\big(u^\epsilon, \beta(u^\epsilon)\big) + \big(f, \beta(u^\epsilon)\big) - \frac{1}{\epsilon}\big(\beta(u^\epsilon), \beta(u^\epsilon)\big) = 0,$$

which is equivalent to

$$\begin{aligned}\Big(\frac{\partial \beta}{\partial t}(u^\epsilon), \beta(u^\epsilon)\Big) &- a\big(\beta(u^\epsilon), \beta(u^\epsilon)\big) - \frac{1}{\epsilon}\big(\beta(u^\epsilon), \beta(u^\epsilon)\big) \\ &= a\big(\psi, \beta(u^\epsilon)\big) + b\big(u^\epsilon, \beta(u^\epsilon)\big) - \big(f, \beta(u^\epsilon)\big) - \Big(\frac{\partial \psi}{\partial t}, \beta(u^\epsilon)\Big).\end{aligned}$$

Note that ψ is independent of t. Integrating from t to T, $t \in [0,T]$, we obtain

$$\frac{1}{2}\big|\beta\big(u^\epsilon(t)\big)\big|_H^2 + \int_t^T a\big(\beta(u^\epsilon),\beta(u^\epsilon)\big)\ ds + \frac{1}{\epsilon}|\beta(u^\epsilon)|^2_{L^2(t,T;H)}$$
$$= -\int_t^T a\big(\psi,\beta(u^\epsilon)\big) + b\big(u^\epsilon,\beta(u^\epsilon)\big) - \big(f,\beta(u^\epsilon)\big)\ ds. \tag{5.5}$$

Recall Lemma 5.1 and note that

$$-\int_t^T a\big(\psi,\beta(u^\epsilon)\big)\ ds = \int_t^T \Big(\frac{\sigma^2}{2}\int_\Omega \frac{\partial\psi}{\partial x}\frac{\partial}{\partial x}(\psi-u^\epsilon)^+\ dx$$
$$+ r\int_\Omega \psi\,(\psi-u^\epsilon)^+\ dx - (\mu-\frac{\sigma^2}{2})\int_\Omega \frac{\partial\psi}{\partial x}\,(\psi-u^\epsilon)^+\ dx\Big)ds, \tag{5.6}$$

in which

$$\int_t^T \Big(\int_\Omega \frac{\partial\psi}{\partial x}\frac{\partial}{\partial x}(\psi-u^\epsilon)^+\ dx\Big)ds = -\int_t^T \Big(\int_\Omega \frac{\partial^2\psi}{\partial x^2}\,(\psi-u^\epsilon)^+\ dx\Big)ds.$$

As ψ is known explicitly, we can write

$$\int_t^T \int_\Omega \frac{\partial\psi}{\partial x}\frac{\partial}{\partial x}(\psi-u^\epsilon)^+\ dx\,ds \tag{5.7}$$
$$= \int_t^T \Big(-\Big[e^x(\psi-u^\epsilon)^+\Big]_{-l}^{\log K} + \int_{-l}^{\log K} e^x\,(\psi-u^\epsilon)^+\ dx\Big)ds \tag{5.8}$$
$$\le \int_t^T \Big(\int_{-l}^{\log K} e^x\,(\psi-u^\epsilon)^+\ dx\Big)ds \ \le\ K\int_t^T \Big(\int_\Omega (\psi-u^\epsilon)^+\ dx\Big)ds. \tag{5.9}$$

To get from (5.8) to (5.9), we used the following fact: since $u^\epsilon \in L^2(0,T;V)$, a monotonicity result in Remark 5.7 gives that $u^\epsilon(s) \ge 0$ for almost every $s \in [0,T]$; hence, $\big(\psi(\log K) - u^\epsilon(\log K)\big)^+ = \big(-u^\epsilon(\log K)\big)^+ = 0$ almost everywhere on $[0,T]$. Having observed this, applying (5.9) to (5.6), we then obtain

$$-\int_t^T a\big(\psi,\beta(u^\epsilon)\big)\ ds \le C_0|\psi|_{H^1(\Omega)}\,|\beta(u^\epsilon)|_{L^2(t,T;H)}, \tag{5.10}$$

which, applied to (5.5), gives

$$\int_t^T a\big(\beta(u^\epsilon),\beta(u^\epsilon)\big)\ ds + \frac{1}{\epsilon}|\beta(u^\epsilon)|^2_{L^2(t,T;H)} \le C_1\Big(|\psi|_{H^2(\Omega)}\,|\beta(u^\epsilon)|_{L^2(t,T;H)}$$
$$+ |u^\epsilon|_{L^2(t,T;H)}\,|\beta(u^\epsilon)|_{L^2(t,T;H)} + |f|_H\,|\beta(u^\epsilon)|_{L^2(t,T;H)}\Big).$$

The splitting of the integral was necessary because of the kink of ψ, whereas for $\psi\big|_\Omega \in H^2(\Omega)$ the last inequality follows directly by integration by parts. Applying Lemma 5.2 and (5.4) to the last expression, we then get

$$\frac{1}{\epsilon}|\beta(u^\epsilon)|_{L^2(t,T;H)} \le C_2 \tag{5.11}$$

for $0 < \epsilon < 1$. Next, applying Lemma 5.2, (5.4) and (5.11) to equation (5.5) yields

$$|\beta(u^\epsilon)|_{L^2(t,T;V)} + |\beta(u^\epsilon)|_{L^\infty(t,T;H)} \le \epsilon^{1/2}C_3\,. \tag{5.12}$$

We define $r^\epsilon := \psi - u + (\psi - u^\epsilon)^-$, where $(\psi - u^\epsilon)^- := -\min\{\psi - u^\epsilon, 0\}$; in particular, this means $u^\epsilon - u = r^\epsilon + \beta(u^\epsilon)$. Owing to (5.12), to prove the theorem, it is now sufficient to show that

$$|r^\epsilon|_{L^2(t,T;V)} + |r^\epsilon|_{L^\infty(t,T;H)} \le \epsilon^{1/2} C_4.$$

We set $v = r^\epsilon + u = \psi + (\psi - u^\epsilon)^- \ge \psi$ in (5.2) and $v = -r^\epsilon \in V$ in (5.3) and sum the two expressions to obtain

$$-\Big(\frac{\partial}{\partial t}(u - u^\epsilon), r^\epsilon\Big) + a(u - u^\epsilon, r^\epsilon) + b(u - u^\epsilon, r^\epsilon) + \frac{1}{\epsilon}\big(\beta(u^\epsilon), u - \psi\big) \ge 0.$$

As $-\frac{1}{\epsilon}\big(\beta(u^\epsilon), u - \psi\big) \ge 0$, we further get

$$-\Big(\frac{\partial}{\partial t}(u^\epsilon - u), r^\epsilon\Big) + a(u^\epsilon - u, r^\epsilon) + b(u^\epsilon - u, r^\epsilon) \le 0,$$

which in return gives

$$-\Big(\frac{\partial r^\epsilon}{\partial t}, r^\epsilon\Big) + a(r^\epsilon, r^\epsilon) + b(r^\epsilon, r^\epsilon) \le \Big(\frac{\partial \beta}{\partial t}(u^\epsilon), r^\epsilon\Big) - a\big(\beta(u^\epsilon), r^\epsilon\big) - b\big(\beta(u^\epsilon), r^\epsilon\big).$$

We define $\widehat{r^\epsilon} := e^{\xi t} r^\epsilon$, where we use ξ from Lemma 5.2. Multiplying both sides of the last inequality by $e^{2\xi t}$, we get

$$\begin{aligned}
&-\Big(\frac{\partial \widehat{r^\epsilon}}{\partial t}, \widehat{r^\epsilon}\Big) + \xi(\widehat{r^\epsilon}, \widehat{r^\epsilon}) + a(\widehat{r^\epsilon}, \widehat{r^\epsilon}) + b(\widehat{r^\epsilon}, \widehat{r^\epsilon}) \\
&\qquad \le \Big(\frac{\partial}{\partial t}\big[e^{\xi t}\beta(u^\epsilon)\big], \widehat{r^\epsilon}\Big) - \xi\big(e^{\xi t}\beta(u^\epsilon), \widehat{r^\epsilon}\big) - a\big(e^{\xi t}\beta(u^\epsilon), \widehat{r^\epsilon}\big) - b\big(e^{\xi t}\beta(u^\epsilon), \widehat{r^\epsilon}\big).
\end{aligned}$$

Noting that $\hat{r^\epsilon}(T) = e^{\xi T}\big[\psi(T) - u(T) + \big(\psi(T) - u^\epsilon(T)\big)^-\big] = 0$, integrating from t to T, $t \in [0, T]$, gives

$$\begin{aligned}
(5.13)\quad & \frac{1}{2}|\widehat{r^\epsilon}(t)|_H^2 + \int_t^T \xi(\widehat{r^\epsilon}, \widehat{r^\epsilon}) + a(\widehat{r^\epsilon}, \widehat{r^\epsilon}) + b(\widehat{r^\epsilon}, \widehat{r^\epsilon})\, ds \\
&\qquad \le -\big(e^{\xi t}\beta(u^\epsilon)(t), \widehat{r^\epsilon}(t)\big) - \int_t^T e^{\xi t}\Big(\beta(u^\epsilon), \frac{\partial \widehat{r^\epsilon}}{\partial t}\Big)\, ds \\
&\qquad\qquad - \int_t^T \xi\big(e^{\xi t}\beta(u^\epsilon), \widehat{r^\epsilon}\big) + a\big(e^{\xi t}\beta(u^\epsilon), \widehat{r^\epsilon}\big) + b\big(e^{\xi t}\beta(u^\epsilon), \widehat{r^\epsilon}\big) ds.
\end{aligned}$$

We now make three observations, which, taken together, will yield the desired result. First, according to Lemma 5.2, we have

$$\int_t^T \xi(\widehat{r^\epsilon}, \widehat{r^\epsilon}) + a(\widehat{r^\epsilon}, \widehat{r^\epsilon}) + b(\widehat{r^\epsilon}, \widehat{r^\epsilon})\, ds \ge \vartheta \int_t^T |\widehat{r^\epsilon}|_V^2\, ds.$$

Second, according to (5.4),

$$\begin{aligned}
&-\int_t^T \Big(e^{\xi t}\beta(u^\epsilon), \frac{\partial \widehat{r^\epsilon}}{\partial t}\Big)\, ds \\
&\qquad \le \xi e^{2\xi T}|\beta(u^\epsilon)|_{L^2(t,T;H)}\, |r^\epsilon|_{L^2(t,T;H)} + e^{2\xi T}|\beta(u^\epsilon)|_{L^2(t,T;H)} \Big|\frac{\partial u}{\partial t}\Big|_{L^2(t,T;H)} \\
&\qquad \le C_5\, |\beta(u^\epsilon)|_{L^2(t,T;H)}\, |r^\epsilon|_{L^2(t,T;H)} + \epsilon C_6\,.
\end{aligned}$$

Third, we have

$$
\begin{aligned}
&-\big(e^{\xi t}\beta(u^\epsilon)(t),\widehat{r^\epsilon}(t)\big)-\int_t^T \xi\big(e^{\xi t}\beta(u^\epsilon),\widehat{r^\epsilon}\big)+a\big(e^{\xi t}\beta(u^\epsilon),\widehat{r^\epsilon}\big)+b\big(e^{\xi t}\beta(u^\epsilon),\widehat{r^\epsilon}\big)\ ds\\
&\qquad\le C_7|\beta(u^\epsilon)(t)|_H\,|r^\epsilon(t)|_H+C_8|\beta(u^\epsilon)|_{L^2(t,T;V)}\,|r^\epsilon|_{L^2(t,T;V)}.
\end{aligned}
$$

Applying the last three statements as well as (5.12) to (5.13) completes the proof. □

Finally, we formulate a corollary which states that the result just given for the American put also holds for a wider class of functions including a number of traditional option payoffs.

Corollary 5.9. *If there is a finite number of disjoint open intervals $I_i := (x_i\,,x_{i+1})$, $0\le i\le N$, such that $\bigcup_{i=0}^N [x_i\,,x_{i+1}]=[-l,l]$, $\psi\big|_{I_i}\in H^2(I_i)$ for $0\le i\le N$, and, additionally,*

$$
\lim_{x\uparrow x_i}\frac{\partial\psi}{\partial x}(x_i)\le\lim_{x\downarrow x_i}\frac{\partial\psi}{\partial x}(x_i)
$$

for $1\le i\le N$, then the result of Theorem 5.8 also holds. In particular, this includes piecewise smooth functions that are convex, a straddle and an American call.

Proof. Integrating by parts, we can write (5.7) as

$$
\int_\Omega\frac{\partial\psi}{\partial x}\,\frac{\partial}{\partial x}(\psi-u^\epsilon)^+\ dx=\sum_{i=0}^N\Big[\frac{\partial\psi}{\partial x}(\psi-u^\epsilon)^+\Big]_{x_i}^{x_{i+1}}-\sum_{i=0}^N\int_{I_i}\frac{\partial^2\psi}{\partial x^2}(\psi-u^\epsilon)^+\ dx,
$$

in which the first term on the right hand side equals

$$
\sum_{i=1}^N\Big(\lim_{x\uparrow x_i}\frac{\partial\psi}{\partial x}(x)\big(\psi(x_i)-u^\epsilon(x_i)\big)^+-\lim_{x\downarrow x_i}\frac{\partial\psi}{\partial x}(x)\big(\psi(x_i)-u^\epsilon(x_i)\big)^+\Big)\le 0,
$$

and we can replace (5.10) by

$$
-\int_t^T a\big(\psi,\beta(u^\epsilon)\big)\ ds\le C_7\sum_{i=0}^N|\psi|_{H^2(I_i)}|\beta(u^\epsilon)|_{L^2(t,T;H)}.
$$

Having done this, we then proceed as in the proof of Theorem 5.8. □

At this point, essentially, what we have found is that convex kinks in otherwise smooth payoff functions will not affect the known results on the convergence rate of the penalty approximation in the H^1 norm, and that these results carry over to jump-diffusion models as expected.

6. Discussion and Applications

6.1. **Summary of Results.** In this section, we summarise the estimates on the penalisation error obtained in this paper. We focus on two settings: that of convex kinks between otherwise smooth (usually linear) payoffs, and that of concave kinks. As concave kinks result in lower convergence order, it is clear that in situations with mixed convexity this behaviour is dominant.

Table 3 confirms that the convergence order predicted by asymptotic analysis is in line with the numerically estimated one in all situations.

	Convex kinks				Concave kinks			
	L_∞	W^1_∞	L_2	H^1	L_∞	W^1_∞	L_2	H^1
Numerical estimate	1	0.55	1	0.76	0.5	0.07	0.61	0.29
Asymptotic expansions	1	0.5	1	0.75	0.5	$\star$	0.5	0.25
Functional analysis	1	—	0.5	0.5	0.5	—	0	0

Table 3. *Order of convergence in the penalty parameter, for different measures and payoff types, and as predicted by different methods of analysis. 0 indicates convergence, but of no positive order; '—' indicates no known result; '$\star$' indicates no convergence. The numerical estimate was obtained by regression of the errors for different numerically computed penalised solutions. These and the results under asymptotic expansions were computed for representative payoffs (put and butterfly).*

The lack of uniform convergence of the penalty butterfly Delta, denoted by the '$\star$', results from the jump of the exact Delta at the strike, which cannot be matched simultaneously on both sides by the continuous penalty Delta. However, the asymptotic analysis also reveals that the error in the Delta is $O(\epsilon^{1/2})$ except in a region which is also of width $O(\epsilon^{1/2})$.

The rates in the H^1 norm can be explained by the asymptotic analysis as follows: for the put, e.g., it gives an error in the derivative, of $O(\epsilon^{1/2})$ in the inner region of width $O(\epsilon^{1/2})$, resulting in an L_2 error of the derivative of

$$\sqrt{O\big((\epsilon^{1/2})^2\epsilon^{1/2}\big)} = O(\epsilon^{3/4}).$$

The error in the outer region is integrable and $O(\epsilon)$ (we can just differentiate the outer expansion) and therefore negligible. The contribution of the solution to the zero order term in the H^1 norm is also of order 1. A similar argument explains the order 1/4 for the butterfly.

The higher level functional analytic estimates are not sharp in these cases.

6.2. **Interplay Between Penalisation and Discretisation.** A comment is due on the effect of discretisation on the penalty convergence order, and, conversely, of penalisation on the grid convergence order.

Penalisation of discrete systems. Consider a discretisation with mesh width h and time step k, say a central difference fully implicit scheme, then (1.2) becomes

$$\min(Ax - b, x - c) = 0, \tag{6.1}$$

where, for Black-Scholes on equally spaced grid points S_i, $1 \le i \le N$,

$$(Ax)_i = x_i - k\frac{1}{2}\sigma^2 S_i^2 h^2 \frac{x_{i+1} - 2x_i + x_{i-1}}{h^2} - krS_i\frac{x_{i+1} - x_{i-1}}{2h} + krx_i, \tag{6.2}$$

such that $A \in \mathbb{R}^{N\times N}$ is an M-matrix, $b, c, x \in \mathbb{R}^N$ (we assume x_0 and x_{N+1} are fixed by boundary conditions). The solution of (6.1) for fixed N is invariant to scaling of the two arguments of the min by a positive constant. However, for asymptotic considerations for large N, the scaling is crucial for the direct iterative solution of (1.2), as discussed in [12].

Scaling is also relevant for the size of the penalty term. In particular, [10] consider a penalised equation

$$Ax^\epsilon - b = \mathit{Large}\max(c - x^\epsilon, 0),$$

for a large positive parameter *Large*, and show that

$$\|x - x^\epsilon\| \leq C/\mathit{Large}. \tag{6.3}$$

We now elucidate the relation between *Large* and ϵ. A key estimate on p. 2117 in [10] is

$$\|Ac\| \leq \mathit{const},$$

for some positive constant and where c is the payoff, and $\|\cdot\|$ the maximum norm. For a Lipschitz payoff, finite differences for the first derivative, say central differences as per (6.2), applied to the payoff, are bounded as $h \to 0$, and second order differences are $O(1/h)$. So as A contains the spatial finite differences *multiplied* by k, $\|Ac\|_\infty = \max_i (A_c)_i \sim k/h$ for a Lipschitz payoff c, and therefore, as long as k/h is kept fixed, C above is independent of the grid sizes.

This does not contradict the differentiation between convex and concave payoffs found in this paper, as for $k, h \to 0$, the above penalised equation is *not* consistent with a penalised PDE with penalty parameter *Large*, instead, as the "effective" penalty parameter Large/k in

$$\frac{1}{k} A x^\epsilon - \frac{1}{k} b = \frac{\mathit{Large}}{k} \max(c - x^\epsilon, 0)$$

increases with $k \to 0$, it is consistent with the obstacle problem itself.
If we replaced *Large* by k/ϵ in the above to get a scheme consistent with the penalised PDE, (6.3) would still hold, however, the constant C will generally depend on k and h. For the butterfly, in particular, C will go to ∞ for $h \to 0$, k/h fixed. This reflects the fact that, for the butterfly, the limiting continuous problem shows reduced convergence order in ϵ. For the put, in contrast, the required penalty parameter is asymptotically independent of the grid parameters. Therefore, the analysis of the limiting continuous problem informs the choice of penalty parameter for the discretised system.

Smoothing and discretisation of penalised equations. The penalised PDE does not have a (known) closed-form solution and has to be solved numerically. Error estimates for a finite element approximation to the penalised heat equation have been given, e.g., by [24] and [29],

$$\|u^\epsilon - \widehat{u}^\epsilon\| \leq \left(c + \frac{C}{\epsilon^{1/2}} \right) \cdot \left(k + h^2 \right), \tag{6.4}$$

where h is the mesh size and k the time-step of an implicit Euler or θ-method respectively, and $\|\cdot\|$ the L_2 norm.

In contrast, error bounds for the unpenalised problem in [18] have reduced order in k and h,

$$\|u - \widehat{u}\| \leq c \cdot \left(k^{1/2} + h \right). \tag{6.5}$$

These results reflect the fact that penalisation smooths the solution. Consequently, the finite element error bounds (6.4) deteriorate for decreasing ϵ, and the order of convergence in the mesh parameters is lower for the limiting variational inequality. The above results are based on the assumption of sufficiently smooth obstacles, such that the penalisation error is determined by smooth pasting at the free boundary and not any kinks of the payoff.

	Value	Delta
Penalty	1.0000	0.4815
Extrapolation	2.0061	1.5015

Table 4. *Order of convergence with respect to the penalty parameter, in the maximum norm, for original and extrapolated value and its derivative. The setting is the Black-Scholes model with parameters as earlier.*

This technique of smoothing the solutions to non-linear PDEs by penalisation in order to derive grid convergence rates for the limiting problem is used in the more general context of HJB and Isaacs equations in [15].

We should remark that, in *practice*, one can observe $O(k^{3/2} + h^2)$ convergence also for the limiting problem, and even $O(k^2 + h^2)$ for a suitably adaptive time stepping scheme, see [10], which accounts for the singular behaviour of the solution close to expiry.

6.3. **Richardson Extrapolation.** We now show how extrapolation using the asymptotic results can be used to generate more accurate numerical solutions.

Consider here the American put. We know from Section 3.1 (see also Table 3) that the leading order correction to the penalty solution is proportional to ϵ. So doing the calculation with ϵ and 2ϵ, then taking one twice minus the other,

$$\widehat{V}^\epsilon = 2V^\epsilon - V^{2\epsilon},$$

will be a second order approximation to V (assuming the next term in the expansion is quadratic).

For the finite difference computation of V^ϵ, we choose a mesh size $h \sim \epsilon^{1/2}$, for two reasons. The (empirically observed) finite difference error is $O(h^2)$, whereas the penalisation error is $O(\epsilon)$, so the above choice makes both terms the same order of magnitude. Also, although the convergence of the penalised PDE solution is $O(\epsilon)$, overlaid is a displacement of the exercise boundary by $\epsilon^{1/2}$ (see Section 3.1), at which the penalisation error changes rapidly, so extrapolation of the continuous equation (or, in practice, one with very small fixed grid size) does *not* result in an order improvement of the maximum error. However, extrapolation with mesh width $O(\epsilon^{1/2})$ coupled to the penalisation, gives the desired numerical results. This is because the 'inner region' is not resolved within grid cells of width $O(\epsilon^{1/2})$ and therefore does not destroy the convergence order of Richardson extrapolation. The results are summarised in Table 4. Note that by this procedure we gain a full convergence order in the derivative as well.

The inner (asymptotic) analysis in Section 3.1 is independent of the volatility to the order of accuracy we have given. One could use a local volatility model and simply freeze the volatility at its local value. So even in non-Black–Scholes models, Richardson extrapolation may be a good way of using this to get a more accurate outer put value with little extra effort. The strategy should also work for multi-factor models.

6.4. **Extensions.** While the analysis in this article focuses on Black-Scholes and jump-diffusion models, the main results, especially of Section 4 and the applicability of extrapolation, should extend to other settings, including local volatility models and derivatives on more than one underlying or on an asset modelled by additional stochastic factors, e.g. stochastic volatility or interest rates. Another interesting extension would be to free-boundary problems arising from portfolio selection under transaction costs, but we anticipate especially the matched asymptotic expansions to differ more substantially here due to the presence of first order derivatives in the penalty term (c.f. [21]).

References

[1] Y. Achdou. An inverse problem for a parabolic variational inequality arising in volatility calibration with American options. *SIAM Journal on Control and Optimization*, 43(5):1583–1615, 2005.
[2] R. A. Adams. *Sobolev spaces.* Amsterdam, Oxford: Academic Press, 2nd edition, 2003.
[3] A. Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematicae*, 2(2):139–158, 1984.
[4] A. Bensoussan and J. L. Lions. *Applications of variational inequalities in stochastic control*, volume 12 of *Studies in mathematics and its applications.* Amsterdam, New York, Oxford: North-Holland Pub. Co. , 1982.
[5] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
[6] R. Cont and P. Tankov. *Financial modelling with jump processes.* Chapman and Hall/CRC, 2004.
[7] C. W. Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM Journal on Control*, 9(3):385–392, 1971.
[8] P. Wilmott, J. Dewynne and S. Howison. *Option pricing: mathematical models and computation.* Oxford: Oxford Financial Press, 1993.
[9] Y. d'Halluin, P. A. Forsyth and G. Labahn. A penalty method for American options with jump diffusion processes. *Journal of Computational and Applied Mathematics*, 91:321–352, 2003.
[10] P. A. Forsyth and K. R. Vetzal. Quadratic convergence for valuing American options using a penalty method. *SIAM Journal on Scientific Computing*, 23(6):2095–2122, 2002.
[11] R. Zvan, P. A. Forsyth and K. R. Vetzal. Penalty methods for American options with stochastic volatility. *Journal of Computational and Applied Mathematics*, 91(2):199–218, 1998.
[12] Y. Huang, P.A. Forsyth and G. Labahn. Inexact arithmetic considerations for direct control and penalty methods: American options under jump diffusion. Working paper, David R. Cheriton School of Computer Science, University of Waterloo, 2011.
[13] S.D. Howison. A matched asymptotic expansions approach to continuity corrections for discretely sampled options. Part 2: Bermudan options. *Applied Mathematical Finance*, 14:91–104, 2007.
[14] K. Ito and K. Kunisch. Parabolic variational inequalities: The Lagrange multiplier approach. *Journal de Mathématiques Pures et Appliqués*, 85(3):415–449, 2006.
[15] E. R. Jakobsen. On error bounds for monotone approximation schemes for multi-dimensional Isaacs equations. *Asymptotic Analysis*, 49(3,4):249–273, 2006.
[16] I. Karatzas. On the pricing of American options. *Applied Mathematics and Optimization*, 17(1):37–60, 1988.
[17] P. Jaillet, D. Lamberton, B. Lapeyre and C. La Courtine. Variational inequalities and the pricing of American options. *Acta Applicandae Mathematicae*, 21:263–289, 1990.
[18] W. Allegretto, Y. Lin and H. Yang. Finite element error estimates for a nonlocal problem in American option valuation. *SIAM Journal on Numerical Analysis*, 39(3):834–857, 2001.
[19] J. L. Lions and E. Magenes. *Problèmes aux limites non homogènes et applications*, volume 1. Paris: Dunod, 1968.
[20] R. C. Merton. Option pricing when underlying stock returns are discontinuous. Working papers 787-75, Massachusetts Institute of Technology (MIT), Sloan School of Management, 1975.
[21] M. H. A. Davis, A. R. Norman. Portfolio selection with transaction costs. *Mathematics of Operations Research*, 15(4):676–713, 1990.
[22] H. Pham. Optimal stopping, free boundary, and American option in a jump-diffusion model. *Applied Mathematics and Optimization*, 35(2):145–164, 1997.

[23] A. Fasano, A. Mancini, M. Primicerio and B. Zaltzman. Waiting time phenomena forced by critical boundary conditions in classical diffusion problems. *Quarterly of Applied Mathematics*, 69:105–122, 2011.
[24] R. Scholz. Numerical solution of the obstacle problem by the penalty method. *Numerische Mathematik*, 49(2-3):255–268, 1986.
[25] S. Shreve. *Stochastic calculus for finance II: continous-time models*. New York: Springer, 2008.
[26] B. F. Nielsen, O. Skavhaug and A. Tveito. Penalty and front-fixing methods for the numerical solution of American option problems. *Journal of Computational Finance*, 5(2):69–97, 2002.
[27] B. F. Nielsen, O. Skavhaug and A. Tveito. Penalty methods for the numerical solution of American multi-asset option problems. *Journal of Computational and Applied Mathematics*, 222(1):3–16, 2008.
[28] S.D. Howison, M. Steinberg. A matched asymptotic expansions approach to continuity corrections for discretely sampled options. Part 1: Barrier options. *Applied Mathematical Finance*, 14:63–89, 2007.
[29] J.H. Witte. On penalty based finite element methods for the pricing of American options. Transfer report, Mathematical Institute, University of Oxford, 2009.
[30] K. Zhang, S. Wang, X. Q. Yang and K. L. Teo. A power penalty approach to numerical solutions of two-asset American options. *Numerical Mathematics: Theory, Methods and Applications*, 2(2):202–223, 2009.
[31] K. Zhang, X. Yang and K. L. Teo. Convergence analysis of a monotonic penalty method for American option pricing. *Journal of Mathematical Analysis and Applications*, 348(2):915–926, 2008.
[32] S. Wang, X. Q. Yang and K. L. Teo. Power penalty method for a linear complementarity problem arising from American option valuation. *Journal of Optimization Theory and Applications*, 129(2):227–254, 2006.
[33] X. L. Zhang. *Méthodes numériques pour le calcul des options américaines dans un modèle de diffusion avec des sauts.* Thèse de doctorat, École Nationale des Ponts et Chaussées, 1994.
[34] X. L. Zhang. Numerical analysis of American option pricing in a jump-diffusion model. *Mathematics of Operations Research*, 22(3):668–690, 1997.

Mathematical Institute and Oxford-Man Institute of Quantitative Finance, University of Oxford
E-mail address: `[howison,reisinge,witte]@maths.ox.ac.uk`