TI 2022-007/IV
Tinbergen Institute Discussion Paper

Solving penalised American options for jump diffusions using the POST algorithm

Jean-Claude Hessing[1]
Rutger-Jan Lange[1]
Daniel Ralph[2]

1 Erasmus School of Economics Rotterdam

2 Judge Business School, University of Cambridge

Solving penalised American options for jump diffusions using the POST algorithm ⋆

Jean-Claude Hessing [a], Rutger-Jan Lange [a], Daniel Ralph [b]

[a] *Econometric Institute, Erasmus School of Economics, Rotterdam, Netherlands*

[b] *Judge Business School, University of Cambridge, Cambridge, UK*

Abstract

This article establishes the Poisson optional stopping times (POST) method by [22] as a near-universal method for solving liquidity-constrained American options, or, equivalently, penalised optimal-stopping problems. In this setup, the decision maker is permitted to "stop", i.e. exercise the option, only at a set of Poisson arrival times; this can be viewed as a liquidity constraint or "penalty" that limits access to optionality. We use monotonicity arguments in function space to establish that the POST algorithm either (i) finds the solution or (ii) demonstrates that no solution exists. The monotonicity of POST carries over to the discretised setting, where we additionally show geometric convergence and provide convergence bounds. For jump-diffusion processes, dense matrix factorisation may be avoided by using a suitable operator-splitting method for which we prove convergence. We also highlight a connection with linear complementarity problems (LCPs). We use the POST algorithm to value American options and compute early-exercise boundaries for Kou's jump-diffusion model [20] and Heston's stochastic volatility model [14], illustrating the breadth of application and numerical reliability of the method.

Key words: Optimal stopping; Penalty method; HJB equation; Contraction; Fixed Point; Operator Splitting; Implicit Explicit; Linear Complementarity Problem.

1 Introduction

Black and Scholes [5] laid down in their seminal work the fundamentals of option pricing, leading to a Hamilton-Jacobi-Bellman (HJB) equation in the form of a partial differential equation (PDE), which can be analytically solved. Limiting ourselves to problems with analytic solutions — the "analytic straitjacket" [1] — is both unfortunate and, given the many available numerical methods, unnecessary. Yet, such methods tend to be bespoke and "there is a limit to the amount of time one wants to spend on one example" [33, p. 115-6]. This article demonstrates the universality of a recently proposed method for solving multidimensional American options in function spaces, dubbed the Poisson optional stopping times (POST) method [22], which leads to an efficient numerical implementation for the pricing of American options on jump-diffusion processes.

The POST method restricts opportunities to exercise the option to a sequence of times generated by an independent Poisson process with intensity $\lambda > 0$. In financial terms, this can be seen as imposing a "liquidity constraint" on the available exercise times. Other papers using Poisson-generated stopping times, though not the POST algorithm, include, in chronological order, [32,11,8,24,25,27,16,17]. In fact, the POST algorithm is used in [15, eq. 5], in the context of liquidity-constrained stopping, for one-dimensional diffusions.

As noted in [22], the parameter λ can also be interpreted as a "penalty parameter" in the penalty formulation of optimal-stopping problems, which dates back to at least [31] and [4]. This and related approaches have been used in numerical studies by many authors, e.g. chronologically, [44,12,40,41,39,42]. The

⋆ Corresponding author: Rutger-Jan Lange (Postbus 1738, 3000 DR, Rotterdam, Netherlands)

Email addresses: lange@ese.eur.nl (Rutger-Jan Lange), d.ralph@jbs.cam.ac.uk (Daniel Ralph).

POST algorithm offers a monotonic approach in function spaces to solving penalised (or finite-liquidity) options or stopping problems.

We present notation and formulation of penalised (or finite-liquidity) options or stopping problems in Section 2, where, given $\lambda > 0$, the value function is denoted by V_λ. Section 3 shows that the POST algorithm in function space is perfectly adapted (Theorem 1) to finite-liquidity optimal stopping: Broadly speaking, if there is a solution V_λ, the algorithm will find it. This robust function-space result underlies the numerical implementation of the algorithm, for which we discretise. In section 4 we show that the discretised version of the POST algorithm is geometrically convergent (Proposition 2) to the unique discretised solution. Section 5 tackles jump-diffusion processes, which are ubiquitous in finance. We propose an operator-splitting method (Proposition 3) that allows POST to be implemented using sparse linear solvers. In section 6 we note that the discretised finite-liquidity optimal-stopping problem can be equivalently written as a linear complementarity problem (LCP), and we observe (Proposition 4) that the POST algorithm is a new, provably convergent method for the corresponding class of LCPs.

Section 7, on numerical application, opens with a five-step “POST recipe” to implement the discretised POST algorithm. This makes numerical implementation straightforward, and easy to modify to account for different model features. To demonstrate the validity of POST, we use this recipe to solve the jump-diffusion model of [20,21] and compare it against the well-known analytic solution. We then change the distribution of the jumps, thereby ruling out an analytic solution. Our numerical treatment is accomplished by changing a single line of code in the discretised POST implementation. Finally, we consider the valuation of a finite-maturity American option under Heston’s stochastic volatility model [14], as considered in e.g. [6], demonstrating that a plain-vanilla application of POST can solve options with three state variables.

2 Liquidity-constrained American options or penalised optimal-stopping problems

Liquidity-constrained American options are optimal-stopping problems in which the decision maker is prohibited from stopping except at Poisson arrival times. Such problems are also known as "penalty formulations" of stopping problems and were initially described in [31] and [4]. As is standard, the stochastic process $\{X_t\}_{t\geq 0}$ takes values in the state space $\mathcal{X}$, which we think of as a nonempty and closed set in a finite-dimensional Euclidean space $\mathbb{R}^d$, but which in principle could be any locally compact, separable and topological space. We also assume that $\{X_t\}_{t\geq 0}$ satisfies the strong Markov property, and has paths that are right continuous and allow limits from the left (càdlàg). This makes $\{X_t\}$ a Hunt process, prominent examples of which include jump diffusions and Lévy-type processes. The decision maker is permitted to "stop", i.e. exercise the option, only at a set of independently generated Poisson arrival times with intensity $\lambda > 0$. She is forced to continue prior to the first Poisson arrival time and in between any two Poisson arrival times. Classic stopping problems, where stopping is permissible at any time, form a limiting case where $\lambda \to \infty$.

As is standard in this setting, the following Hamilton-Jacobi-Bellman (HJB) equation should be satisfied for all $x \in \mathcal{X}$:

$$r\, V_\lambda(x) \;=\; \mathrm{L}V_\lambda(x) \,+\, f(x) \,+\, \lambda\, [g(x) - V_\lambda(x)]^+, \tag{1}$$

where $[\cdot]^+ = \max\{\cdot, 0\}$. The infinitesimal generator L is defined via its operation on a test function $h : \mathcal{X} \to \mathbb{R}$ as $(\mathrm{L}h)(x) := \lim_{t\downarrow 0}(\mathbb{E}^x[h(X_t)] - h(x))/t$, assuming this limit exists for each x. HJB Eq. (1) further features the discount rate $r > 0$; the Poisson arrival rate $\lambda > 0$; the flow gain, stopping gain and value function denoted by $f, g, V_\lambda : \mathcal{X} \to \mathbb{R}$, where f, g represent the problem data, while the dependence of the solution V_λ on the Poisson intensity λ is indicated by the subscript. In financial terms, HJB Eq. (1) indicates that the return, $r\,V_\lambda$, is the sum of three terms: the expected change in the value function as measured by $\mathrm{L}V_\lambda$, the flow gain f, and the additional value of exchanging V_λ for g whenever the latter exceeds the former and a Poisson arrival time is generated. The term $\lambda[g - V_\lambda]^+$ is also known as a "penalty term", as it takes strictly positive values only in the interior of the "stopping set" $\{g \geq V_\lambda\}$, where the decision maker wants to stop but is prohibited from doing so except at Poisson arrival times. As this penalty term illustrates, it is optimal to stop at the first Poisson arrival time t for which $g(X_t) \geq V_\lambda(X_t)$.

In our applications, we consider the case where $\{X_t\}$ is a jump-diffusion process on $\mathcal{X} = \mathbb{R}^d$. The generator L, operating on the test function h, then has the following explicit representation (see [19, p. 360] or [3, p. 158]):

$$\begin{aligned} \mathrm{L}h &= \mathrm{D}h + \mathrm{J}h, \qquad (2)\\ (\mathrm{D}h)(x) &= \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma_{ij}(x)\frac{\mathrm{d}^2 h(x)}{\mathrm{d}x_i\,\mathrm{d}x_j} + \sum_{i=1}^{d}\mu_i(x)\frac{\mathrm{d}h(x)}{\mathrm{d}x_i},\\ (\mathrm{J}h)(x) &= \int\limits_{\mathbb{R}^d\setminus\{x\}} h(y)\,\nu(x,\mathrm{d}y). \end{aligned}$$

This illustrates that for jump-diffusion processes on $\mathcal{X} = \mathbb{R}^d$, the generator L is a linear but unbounded (in the supremum norm) integro-differential operator. The generator L consists of a "diffusive" part D and a "jump" part J. The diffusive part D contains the $d \times 1$ drift vector $\mu_i(x)$ for $i = 1, \ldots, d$ and the symmetric nonnegative definite $d \times d$ diffusion matrix $\sigma_{ij}(x)$ for $i, j = 1, \ldots, d$. The jump part J contains the nonnegative Borel jump measure $\mathcal{B} \mapsto \nu(x, \mathcal{B})$, where $\mathcal{B}$ can be any Borel subset of the space $\mathbb{R}^d \setminus \{x\}$.

As representation (2) illustrates, HJB Eq. (1) is typically a partial integro-differential equation (PIDE), which cannot be solved analytically except in special cases. The unavailability of closed-form solutions calls for the development of approximation methods, ideally in function space, i.e. prior to discretising; the next section describes such a method.

3 POST dichotomy: Using the POST algorithm to characterise the existence of V_λ

HJB Eq. (1) can be solved in function space using the POST algorithm introduced in [22], which can be motivated by adding $(\lambda - \mathrm{L})V_\lambda$ to both sides to obtain

$$(r + \lambda - \mathrm{L})V_\lambda = f + \lambda \max\{g, V_\lambda\}, \tag{3}$$

where for notational simplicity we omit the argument $x \in \mathcal{X}$ for functions f, g, V_λ. By attaching the superscripts n and $n-1$ to the value functions appearing on the left and right-hand sides, we obtain the POST algorithm:

$$(r + \lambda - \mathrm{L})V_\lambda^{(n)} = f + \lambda \max\{g, V_\lambda^{(n-1)}\}, \tag{4}$$

for $n = 1, 2, \ldots$ and some initialisation $V_\lambda^{(0)}$.

It is easy to see [22] that if $V_\lambda^{(0)} \leq V_\lambda^{(1)}$ (pointwise) then the entire POST sequence is pointwise non-decreasing. Moreover $V_\lambda^{(0)} \leq V_\lambda^{(1)}$ follows immediately [22,15] if we take $V_\lambda^{(0)}$ to be a solution of

$$(r + \lambda - \mathrm{L})V_\lambda^{(0)} = f + \lambda g, \tag{5}$$

which corresponds to the (suboptimal) policy of stopping at the first Poisson arrival time. We use this starting point in numerical experiments in section 7.

For this initialisation, monotonicity of the POST iterates can be understood by noting that $V_\lambda^{(n)}$ represents the value when n (Poisson-generated) optional stopping times remain, after which the decision maker receives $V_\lambda^{(0)}$ if she has not stopped already. That is, each iteration permits one more opportunity to exercise.

A convergence analysis of the POST algorithm (4) is given in [22, Eq. B1] under a somewhat restrictive technical condition on the data functions f, g. Here, we give a more fundamental result that is simpler to derive and involves almost no assumptions, indicating that convergence of the POST algorithm (4) is equivalent to well posedness of the penalised HJB Eq. (1).

Theorem 1 (POST dichotomy in function space) *Fix $r, \lambda > 0$, and let $f, g : \mathcal{X} \to \mathbb{R}$ be such that $V_\lambda^{(0)}$ can be computed, as a finite-valued function, from* (5)*. Then $\{V_\lambda^{(n)}\}$ defined by algorithm* (4) *is pointwise non-decreasing and, as $n \to \infty$, one of two mutually exclusive statements holds:*

(1) $\{V_\lambda^{(n)}\}$ converges, pointwise, to a (finite-valued) solution V_λ of HJB equations (1)*.*

(2) HJB (1) *does not have finite-valued solution.*

The POST dichotomy implies that the POST algorithm (4) is perfectly adapted to solving HJB Eq. (1). The proof, which follows by a direct application of Lebesgue's monotone convergence theorem, is presented in Appendix A; it shows that for any initialisation with $V_\lambda^{(0)} \leq V_\lambda^{(1)}$, the POST sequence either converges to a (finite-valued) solution V_λ or diverges to ∞ at some x. In fact the theorem can be slightly extended to the situation when the suboptimal policy of stopping at the first Poisson arrival time takes the value $+\infty$ at some x. Clearly then, the optimal value V_λ also takes an infinite value there, and we are in case 2 of the POST dichotomy.

Interestingly, Theorem 1 requires neither bounded nor smooth problem data f, g. This contrasts with the majority of the vast literature on solutions to penalised HJB equations, which typically assumes bounded if not smooth problem data, c.f., in chronological order, [31, p. 9], [36, p. 180], [37, p. 274], [43, p. 921], [28, p. 3785], [18, p. 106], [29, p. 2365], [38, p. 1084], [13, p. 272], [35, p. 319], [27, p. 2661], [26, p. 434], [10, p. 38], and [9, p. 23]. Our extension to unbounded data is important in fields as diverse as finance, economics and operations, where problems with unbounded problem data are the norm rather than the exception. The conditions of Theorem 1 are mild; we are unaware of other results proving existence in this general setting.

4 Discretised POST algorithm

While the POST algorithm (4) operates in function space, for computation we work in a bounded domain of some finite-dimensional (vector) space. This approach, also used in [22] and [23], involves fixing a natural number N to discretise the computational domain within the state space $\mathcal{X}$, using grid points $\{x_i : i = 1, \ldots, N\}$. Discretising the problem data f, g yields two vectors $\mathbf{f}, \mathbf{g} \in \mathbb{R}^N$. For example, $\mathbf{f}$ could be defined by setting its i-th component to be the function value $f(x_i)$. Discretising the operator L yields an $N \times N$ matrix $\mathbf{L}$, described below (vectors and matrices are given in bold font). The discretised version of the POST fixed-point Eq. (3) reads

$$\Big[(r+\lambda)\,\mathbf{I} - \mathbf{L}\Big]\,\mathbf{V}_\lambda \;=\; \mathbf{f} + \lambda \max\{\mathbf{g}, \mathbf{V}_\lambda\}\,, \tag{6}$$

where $\mathbf{I}$ is an identity matrix of appropriate size, the max operator is applied elementwise to a pair of vectors, and $\mathbf{V}_\lambda$ denotes the discretised solution. The discretised POST algorithm reads

$$\Big[(r+\lambda)\,\mathbf{I} - \mathbf{L}\Big]\,\mathbf{V}_\lambda^{(0)} = \mathbf{f} + \lambda\,\mathbf{g}, \tag{7}$$

$$\Big[(r+\lambda)\,\mathbf{I} - \mathbf{L}\Big]\,\mathbf{V}_\lambda^{(n)} = \mathbf{f} + \lambda \max\Big\{\mathbf{g}, \mathbf{V}_\lambda^{(n-1)}\Big\}\,, \tag{8}$$

for $n = 1, 2, \ldots$. where we take the same initialisation as in [22] and $\mathbf{V}_\lambda^{(n)} \in \mathbb{R}^N$ is the discretised version of the approximate solution $V_\lambda^{(n)}$ in Eq. (4). The monotonicity of Theorem 1 carries over to the discretised setting if L is appropriately discretised to give $\mathbf{L}$.

For a diffusion process, both $\mathbf{L}$ and the matrix $(r+\lambda)\,\mathbf{I} - \mathbf{L}$ are sparse. It is thus computationally efficient to solve (8) using sparse factorisation of that matrix, rather than forming its (dense) inverse. The next result characterises the theoretical properties of discretised POST.

Proposition 2 (Convergence of Discretised POST) *Let $r, \lambda > 0$, $\mathbf{f}, \mathbf{g} \in \mathbb{R}^N$ and $\mathbf{L} \in \mathbb{R}^{N\times N}$ be weakly diagonally dominant with nonpositive diagonal* *elements and nonnegative off-diagonal elements. Then:*

(1) *Discretised POST,* (8) *initialised via* (7)*, generates a sequence* $\{\mathbf{V}_\lambda^{(n)}\}_{n\in\mathbb{N}}$ *that is monotonically non-decreasing, i.e.* $\mathbf{V}_\lambda^{(n)} \geq \mathbf{V}_\lambda^{(n-1)}$ *for all* n*, and converges geometrically, at rate* $\lambda/(\lambda+r)$ *in the max-norm, to the unique fixed point of Eq.* (6)*.*

(2) *We arrive at the following online computable bound at iteration* $n \geq 1$*:*

$$\left\|\mathbf{V}_\lambda - \mathbf{V}_\lambda^{(n)}\right\|_\infty \leq \frac{\lambda}{r}\left\|\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}\right\|_\infty . \tag{9}$$

The convergence proof appears in Appendix B. Geometric convergence in part 1 and the bound in part 2 rely on the classic bound on the inverse of a diagonally dominant matrix [2], which we apply to a matrix of the form $\mu\mathbf{I}-\mathbf{L}$ where $\mu > 0$ and $\mathbf{L}$ is itself diagonally dominant with nonpositive diagonal entries, giving $\|[\mu\mathbf{I}-\mathbf{L}]^{-1}\|_\infty \leq 1/\mu$. Monotonicity, in part 1, uses an additional sign property which is that off diagonal elements of $\mathbf{L}$ are nonpositive. This ensures for any $\mu > 0$ that the Z matrix $\mu\mathbf{I}-\mathbf{L}$ is an M matrix, such that $[\mu\mathbf{I}-\mathbf{L}]^{-1}$ consists of nonnegative entries; see [30].

5 Discretised POST with operator splitting for jump diffusions

For jump diffusions, in which case $\mathrm{L} = \mathrm{D} + \mathrm{J}$ as in the representation (2), we must construct the discretisation $\mathbf{L} = \mathbf{D} + \mathbf{J} \in \mathbb{R}^{N\times N}$. Although $\mathbf{D}$ is sparse, $\mathbf{J}$ and hence $\mathbf{L}$ will be dense due to the integral nature of the jump operator J. The integral part of J can be discretised using a straightforward Riemann approximation. We take a simple view of jump diffusions, given by the second part of (2), in supposing that $\nu(x, \mathrm{d}y) = \nu(x)\, p(x, \mathrm{d}y)$. That is, we suppose that for each $x \in \mathcal{X}$ there exists a bounded jump intensity $0 \leq \nu(x) < \infty$, such that the probability of a jump in a short time interval equals $\nu(x)\,\mathrm{d}t$. The (marginal) density conditional on the presence of a jump originating from $X_0 = x$ and ending up at some $y \in \mathcal{X}$ is given by $p(x, y)\,\mathrm{d}y$, where $p(x, \cdot)$ is a probability density over $\mathcal{X}$. For a pure jump process, J can be derived by considering

$$
\begin{aligned}
(\mathrm{J}h)(x) &= \frac{\mathbb{E}^x h(X_{\mathrm{d}t}) - h(x)}{\mathrm{d}t}, \\
&= \frac{\nu(x)\,\mathrm{d}t \int_{\mathcal{X}} p(x,y)\,h(y)\,\mathrm{d}y + (1-\nu(x)\,\mathrm{d}t)h(x) - h(x)}{\mathrm{d}t}, \\
&= \nu(x) \int_{\mathcal{X}} p(x,y)\,h(y)\,\mathrm{d}y - \nu(x)\,h(x),
\end{aligned} \tag{10}
$$

which is well-defined for any test function $h : \mathcal{X} \to \mathbb{R}$ such that the integral exists for all x. The function-space action of the integral operator J in Eq. (10) allows a straightforward Riemann discretisation as $\mathrm{diag}(\boldsymbol{\nu})\,\mathbf{P}$ where $\boldsymbol{\nu} \in \mathbb{R}^N$ is a (nonnegative) discretisation of the jump intensity $\nu : \mathcal{X} \to \mathbb{R}$, the diag operator takes a vector and puts it on the diagonal of a matrix that contains zeros elsewhere, while $\mathbf{P}$ is the $N \times N$ matrix whose (i,j)th entry is $\beta_i\, p(x_i, x_j)\alpha_j$, given a measure $\alpha_j \geq 0$ of the area of $\mathcal{X}$ associated with the grid point x_j and a scaling $\beta_i := (\sum_j p(x_i, x_j)\alpha_j)^{-1}$. It follows that $\mathbf{P}$ is a stochastic matrix: It is nonnegative and every row sums to 1, which is consistent with $p(x, \cdot)$ being a density. In summary, the matrix discretisation $\mathbf{J} \in \mathbb{R}^{N\times N}$ is

$$
\mathbf{J} \;:=\; \mathrm{diag}(\boldsymbol{\nu})\,\mathbf{P} \;-\; \mathrm{diag}(\boldsymbol{\nu}). \tag{11}
$$

We observe that $\mathbf{J}$ in Eq. (11) satisfies the sign requirements and weak diagonal dominance requirements as required by Proposition 2. The discretised POST method, (7)–(8), will therefore solve the optimal-stopping problem for a pure jump process. However, $\mathbf{J}$, $\mathbf{L}$ and $(r+\lambda)\mathbf{I} - \mathbf{L}$ are dense, and iterating the purely implicit scheme (8) would require a costly dense matrix factorisation. To alleviate this computational burden, it is attractive to decompose $\mathbf{L}$ into a sparse part $\mathbf{L}_{\mathrm{im}} := \mathbf{D} - \mathrm{diag}(\boldsymbol{\nu})$ and a dense part $\mathbf{L}_{\mathrm{ex}} := \mathrm{diag}(\boldsymbol{\nu})\,\mathbf{P}$,

$$
\mathbf{L} = \mathbf{L}_{\mathrm{im}} + \mathbf{L}_{\mathrm{ex}}. \tag{12}
$$

Then we convert equations (7) and (8) to an implicit-explicit format as follows:

$$\Big[(r+\lambda)\,\mathbf{I}\,-\,\mathbf{L}_{\text{im}}\Big]\,\mathbf{V}_\lambda^{(0)} = \mathbf{L}_{\text{ex}}\,\mathbf{g}+\mathbf{f}+\lambda\,\mathbf{g}, \tag{13}$$

$$\Big[(r+\lambda)\,\mathbf{I}\,-\,\mathbf{L}_{\text{im}}\Big]\,\mathbf{V}_\lambda^{(n)} = \mathbf{L}_{\text{ex}}\,\mathbf{V}_\lambda^{(n-1)}+\mathbf{f}+\lambda\max\Big\{\mathbf{g},\mathbf{V}_\lambda^{(n-1)}\Big\}, \tag{14}$$

for $n = 1, 2, \ldots$. That is, we place the *known* quantity $\mathbf{L}_{\text{ex}}\,\mathbf{V}_\lambda^{(n-1)}$ on the right to compensate for omitting the *unknown* quantity $\mathbf{L}_{\text{im}}\,\mathbf{V}_\lambda^{(n)}$ that would otherwise have appeared on the left. In terms of numerical linear algebra, each iteration requires a dense matrix-vector multiplication and a sparse linear-system solve, the latter similar to POST under a pure diffusion operator [22].

Next we check that the linear system (14) can be solved stably. Since $\mathbf{L}_{\text{im}}$ has a more negative diagonal than $\mathbf{D}$, the intuition is that $(r+\lambda-\mathbf{L}_{\text{im}})^{-1}$ will be more contractive than $(r+\lambda-\mathbf{D})^{-1}$. To explore this, let

$$\delta \;:=\; \min_i \nu_i \;\geq\; 0, \tag{15}$$

where ν_i denotes the i-th element of the vector $\boldsymbol{\nu} \in \mathbb{R}^N$. Next, we observe that $\delta\mathbf{I}+\mathbf{D}$ has nonnegative off-diagonal elements, nonpositive diagonals and is weakly diagonally dominant. Proposition 2 applied to $(r+\lambda)\mathbf{I}-\mathbf{L}_{\text{im}} = (\delta+r+\lambda)\mathbf{I}-(\delta+\mathbf{L}_{\text{im}})$ therefore says that $\|((r+\lambda)\mathbf{I}-\mathbf{L}_{\text{im}})^{-1}\|_\infty \leq 1/(\delta+r+\lambda)$. This is the basis for Proposition 3 below, which shows that if $\|\mathbf{L}_{\text{ex}}\|_\infty < \delta+r$, then (14) defines a contraction that converges to a solution of the discretised PIDE (6). Of course, if $\|\mathbf{L}_{\text{ex}}\|_\infty$ is too large, we can decrease it without affecting δ by shifting any of the upper or lower off-diagonal elements of $\mathbf{L}_{\text{ex}}$ to $\mathbf{L}_{\text{im}}$ at the expense of increasing the density of the latter.

Proposition 3 (POST with operator splitting) *Take $\lambda, r > 0$, $\mathbf{f}, \mathbf{g} \in \mathbb{R}^N$, and $\mathbf{L}_{\text{im}}, \mathbf{L}_{\text{ex}} \in \mathbb{R}^{N\times N}$. Assume that $\mathbf{L}_{\text{im}}$ has positive off-diagonal elements and there exists $\delta > \|\mathbf{L}_{\text{ex}}\|_\infty - r$ such that $\delta\mathbf{I}+\mathbf{L}_{\text{im}}$ has negative diagonal elements and is weakly diagonally dominant. For any initial point $\mathbf{V}_\lambda^{(0)} \in \mathbb{R}^N$, the iterative scheme* (14) *then generates a sequence $\{\mathbf{V}_\lambda^{(n)}\}$ that converges geometrically at rate $(\|\mathbf{L}_{\text{ex}}\|_\infty+\lambda)/(\delta+r+\lambda)$ to a solution of the fixed point*

(6).

The convergence proof appears in Appendix C. When the jump intensity $\nu(x)$ is constant on the state space, i.e. when $\nu(x) = \nu_0 > 0$ for all $x \in \mathcal{X}$, as is often the case in practice, then convergence is guaranteed. To see why, note that $\delta := \nu_0$ and $\|\mathbf{J}\|_\infty = \nu_0 \max_i \sum_j P_{ij} = \nu_0$. Thus $\|\mathbf{J}\|_\infty = \delta < \delta + r$, which guarantees convergence.

6 Discretised POST for solving a class of LCPs

The discretised problem (6) can be reformulated as an equivalent linear complementarity problem (LCP). While classic LCP methods feature decreasing monotonicity [7, Theorem 5.3.17], our algorithm features increasing monotonicity. Moreover, finding an initial feasible point requires no work. The POST algorithm can thus be used to solve LCPs; to the best of our knowledge, this method for solving LCPs is new.

To transform (6) to LCP format, start by subtracting $\mathbf{g}$ from both sides; rewrite the left-hand side $\mathbf{V}_\lambda - \mathbf{g}$ as the difference between its positive part $\mathbf{x}$ and negative part $\mathbf{w}$, i.e.

$$\mathbf{V}_\lambda - \mathbf{g} = \mathbf{x} - \mathbf{w},\ \mathbf{x} := (\mathbf{V}_\lambda - \mathbf{g})^+,\ \mathbf{w} := (\mathbf{V}_\lambda - \mathbf{g})^-. \tag{16}$$

After some rearrangements with $\mathbf{K} := \lambda\Big((r+\lambda)\mathbf{I} - \mathbf{L}\Big)^{-1}$ and $\mathbf{q} := \Big((r+\lambda)\mathbf{I} - \mathbf{L}\Big)^{-1}\mathbf{c} + (\mathbf{I} - \mathbf{K})\mathbf{g}$, we see that (6) is equivalent to the following LCP with variables $\mathbf{x}, \mathbf{w} \in \mathbb{R}^N$,

$$\mathbf{0} \ \le\ \mathbf{w} = (\mathbf{I} - \mathbf{K})\mathbf{x} + \mathbf{q} \quad \perp \quad \mathbf{x} \ \ge\ \mathbf{0}, \tag{17}$$

where $\perp$ denotes orthogonality or complementarity. That is, a solution $\mathbf{V}_\lambda$ to (6) generates a solution $\mathbf{x} = (\mathbf{V}_\lambda - \mathbf{g})^+$ and $\mathbf{w} = (\mathbf{I} - \mathbf{K})\mathbf{x} + \mathbf{q}$ of (17); and, conversely, if $(\mathbf{x}, \mathbf{w})$ solves the LCP then $\mathbf{V}_\lambda - \mathbf{g} = \mathbf{x} - \mathbf{w}$ solves (6).

We note that the iterative method (8) is viewed in the literature on LCPs [7, Section 5.2] as a projective splitting or projective Jacobi method. Consider the projective splitting scheme initialised with $\mathbf{x}^{(0)} = \mathbf{0} \in \mathbb{R}^N$, such that $\mathbf{x}^{(1)} = (-\mathbf{q})^+$ and, in general, for $n = 1, 2, \ldots$

$$\mathbf{x}^{(n)} := (\mathbf{K}\mathbf{x}^{(n-1)} - \mathbf{q})^+, \tag{18}$$

$$\mathbf{w}^{(n)} := \mathbf{x}^{(n)} - (\mathbf{K}\mathbf{x}^{(n-1)} - \mathbf{q}). \tag{19}$$

This corresponds to (8) starting from $n = 1$, because

$$\mathbf{V}_\lambda^{(1)} = -\mathbf{q} + \mathbf{g} = \mathbf{x}^{(1)} - \mathbf{w}^{(1)} + \mathbf{g}$$

and, for $n \geq 2$, we can see that $\mathbf{V}_\lambda^{(n)} = \mathbf{x}^{(n)} - \mathbf{w}^{(n)} + \mathbf{g}$. This leads to a new LCP splitting convergence result, which is a corollary of Proposition 2.

Proposition 4 (Discretised POST solves a class of LCPs) *Let $\mathbf{q} \in \mathbb{R}^N$ and $\mathbf{K} \in \mathbb{R}^{N\times N}$ be such that $\mathbf{K}$ has nonnegative entries and $\|\mathbf{K}\|_\infty < 1$. If $\mathbf{x}^{(0)} = \mathbf{0}$ then* (18) *generates a non-decreasing sequence $\{\mathbf{x}^{(n)}\}$, i.e. $\mathbf{x}^{(n)} \geq \mathbf{x}^{(n-1)}$, such that $\{(\mathbf{x}^{(n)}, \mathbf{w}^{(n)})\}$ converges to a solution of* (17)*.*

Monotonicity can be seen directly given that $\mathbf{x}^{(1)} \geq \mathbf{0} = \mathbf{x}^{(0)}$. Thus, for $n \geq 2$, induction on $\mathbf{x}^{(n-1)} \geq \mathbf{x}^{(n-2)}$ gives $\mathbf{K}\mathbf{x}^{(n-1)} - \mathbf{q} \geq \mathbf{K}\mathbf{x}^{(n-2)} - \mathbf{q}$, hence $\mathbf{x}^{(n)} \geq \mathbf{x}^{(n-1)}$.

In an interesting contrast, applying LCP theory [7, Theorem 5.3.17] in the context of the previous result shows that the projective splitting method is guaranteed to generate a *non-increasing* sequence $\{\mathbf{x}^{(n)}\}$ that converges to a solution of the LCP if the initial point is feasible, i.e. $\mathbf{x}^{(1)} \geq \mathbf{0}$ and $(\mathbf{I} - \mathbf{K})\mathbf{x}^{(1)} + \mathbf{q} \geq \mathbf{0}$. We do not use this result, however, as obtaining an initial feasible point requires considerable effort.

7 Numerical applications

To solve optimal-stopping problems in practice, we provide the following five-step recipe:

(1) **Localise**: Decide on a bounded domain of interest.
(2) **Discretise**: Fix N grid points and construct discretised problem data $\mathbf{f}, \mathbf{g} \in \mathbb{R}^N$ and jump-diffusion matrix $\mathbf{L} \in \mathbb{R}^{N \times N}$ to satisfy the conditions of Proposition 2 or 3.
(3) **Impose boundary conditions**: Adjust $\mathbf{f}, \mathbf{g}$ and $\mathbf{L}$ to incorporate relevant boundary conditions on the edge of the grid.
(4) **Run discretised POST algorithm**: Initialise via (7). Iterate via (8), i.e. without operator splitting, or (14), i.e. with operator splitting, until some convergence criterion is satisfied.
(5) **Perform robustness checks**: Repeat steps (1)–(4) after changing the bounded domain, refining the grid, and increasing the Poisson intensity $\lambda > 0$. Stop if the previous and current discretised solutions are "close".

Below, we apply this approach to valuing American options under Kou's jump-diffusion model and Heston's stochastic volatility model and compare with results available in the literature.

7.1 Application 1: Kou's jump-diffusion model

To demonstrate the accuracy of POST, we consider an American put option for a jump-diffusion process for which an analytic solution is available in [20]. The stock price $\{S_t\}$ is a geometric Brownian motion started at $S_0 = 100$, with volatility $\sigma = 0.2$, risk-neutral drift, and a sequence of independent identically distributed shocks that arrive at the constant jump intensity $\nu > 0$. Jumps are doubly exponentially distributed, where $p = 0.6$ and $q = 1 - p = 0.4$ are the probabilities of upward and downward jumps. Upward (downward) jumps are exponentially distributed with parameter $\eta_1 = 25$ ($\eta_2 = 25$). We take the state variables to be $X_t = \log(S_t/S_0) \in \mathbb{R}$ and $t \leq T$, where $T = 0.25$ is the option's maturity date in years. The stopping gain equals $g(X_t, t) = [K - S_t]^+ \mathbb{1}_{t \leq T}$, where $K = 100$ is the strike price. Time to maturity is denoted by $\tau = T - t$. There are no dividends, hence the flow gain f equals zero. All gains are discounted using $r = 0.05$.

The five-step POST recipe is implemented constructing 250 equally spaced gridpoints relating to time, 1,100 gridpoints relating to the stock price, which range from $S = 10^{-6}$ to $S = 10K$ and have a higher density around the strike price $S = K$. In total, $N = 250 \times 1{,}100 = 275{,}000$. We use standard finite-difference stencils to discretise the diffusive part of $\mathbf{L} \in \mathbb{R}^{N \times N}$. The jump part is discretised by utilising the cumulative distribution function of the double exponential distribution, which is known in closed form. Jumps that move the stochastic process outside the discretised domain are disallowed and reassigned to the boundary of the domain. We truncate the state space at $t = T$ and impose the terminal boundary condition $g(X, T) = [K - S_0 \exp(X)]^+$. On the edge of the grid where $S = 10K$, the option value is set to zero. Operator-splitting algorithm (14) is run until a convergence criterion is satisfied, ensuring that the change in the value function and exercise boundary are sufficiently small. We refine and extend the grid, noting that these changes do not noticeably change the value function or the corresponding policy. We take $\lambda = 2^{10} = 1024$, which corresponds to $\sim$4 exercise opportunities per day, and note that this value is large enough to closely approximate the case $\lambda = \infty$. This was verified numerically: Beyond $\lambda = 2^8 = 128$, the policy and value function are insensitive to changes in λ.

The results shown in Table 1 demonstrate that POST closely approximates the true option value in most scenarios. The difference with the analytic solution by Kou [20] is under 0.5% in most cases, while the exercise boundary typically

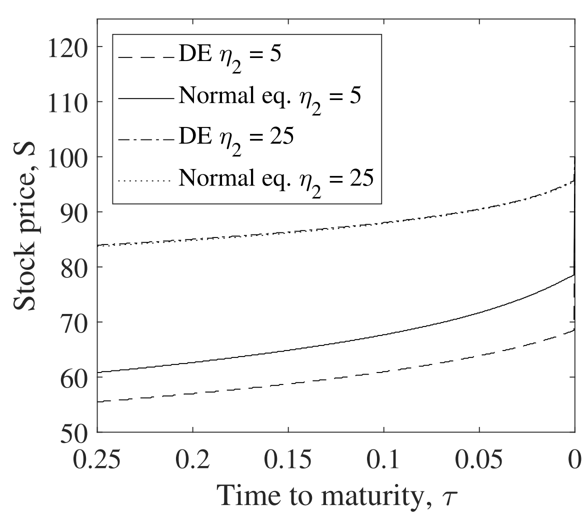
Table 1

American put value for jump-diffusion model with strike price $K = 100$

Parameters					Value at $S_0 = 100$				Boundary at $S_0 = 100$			
σ	ν	p	η_1	η_2	POST	Kou	Δ	$\Delta\%$	POST	Kou	Δ	$\Delta\%$
0.01	3	0.6	25	25	0.906	0.949	-0.043	-4.48	97.43	97.19	0.25	0.25
0.2	3	0.6	25	25	3.878	3.871	0.006	0.17	85.30	85.87	-0.57	-0.66
0.5	3	0.6	25	25	9.567	9.545	0.021	0.22	62.50	65.06	-2.56	-3.93
0.7	3	0.6	25	25	13.439	13.410	0.028	0.21	50.16	53.11	-2.95	-5.55
0.2	3	0.6	25	25	3.878	3.871	0.006	0.17	85.30	85.87	-0.57	-0.66
0.2	7	0.6	25	25	4.386	4.368	0.018	0.41	82.94	83.77	-0.83	-0.99
0.2	3	0.1	25	25	3.870	3.884	-0.014	-0.37	85.98	86.30	-0.32	-0.37
0.2	3	0.3	25	25	3.862	3.878	-0.016	-0.40	85.81	86.13	-0.32	-0.37
0.2	3	0.5	25	25	3.882	3.873	0.009	0.23	85.38	85.96	-0.57	-0.67
0.2	3	0.6	25	25	3.878	3.871	0.006	0.17	85.30	85.87	-0.57	-0.66
0.2	3	0.9	25	25	3.874	3.868	0.006	0.15	84.96	85.59	-0.63	-0.74
0.2	3	0.6	5	25	7.627	7.612	0.016	0.21	54.88	57.15	-2.27	-3.97
0.2	3	0.6	15	25	4.206	4.209	-0.003	-0.07	83.03	83.60	-0.57	-0.68
0.2	3	0.6	25	25	3.878	3.871	0.006	0.17	85.30	85.87	-0.57	-0.66
0.2	3	0.6	50	25	3.720	3.705	0.015	0.39	86.16	86.78	-0.63	-0.72

stays within a 1% margin. This demonstrates the validity of POST for a model that permits an analytic solution; next, we consider the more general case when a closed-form solution is unavailable.

Fig. 1. Exercise boundaries comparing normal jumps with double exponential jumps; the cases for $\eta_2 = 25$ are indistinguishable



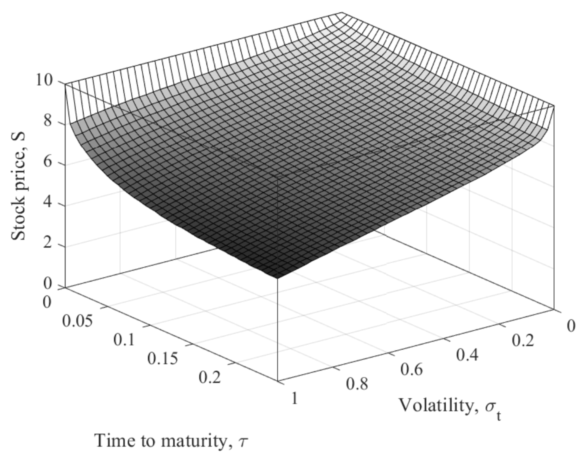
7.2 Application 2: Variation on Kou's model with no analytic solution

Kou's [20] double exponential jump density was selected to facilitate the analytic solution. The POST method allows us to explore alternative jump distributions with minimal user effort by changing one line of code. To investigate the sensitivity of Kou's results to the jump density, we consider a log-normal jump density, with mean and volatility parameters μ_n and σ_n, such that the expected effect on the stock price is a multiplicative factor $\exp(\mu_n + 0.5\sigma_n)$. We vary the jump frequency ν along with the parameters of the jump density (e.g. η_1 and η_2) to ensure that the risk-neutral drift remains constant; otherwise, any differences in the solution could be due to variation in the drift rather than the jump process. When comparing normally and exponentially distributed jumps, we match the mean and variance of jumps in $X_t = \log(S_t/S_0)$. Figure 1 shows the exercise boundaries obtained for the double exponential (DE) distribution and its "normal equivalent", abbreviated as "Normal eq.". When $\eta_2 = 25$, such that the DE distribution is symmetric, the exercise boundaries corresponding to the DE distribution and its normal equivalent are practically indistinguishable. When $\eta_2 = 5$, however, the boundary corresponding to the DE distribution lies substantially below that corresponding to the normal distribution. When downward jumps are more likely, and have heavier tails than the equivalent normal distribution, it is optimal to wait longer before exercising the option. In sum, negative skew and tail thickness of the jump density have a sizeable effect on the optimal policy.

7.3 Application 3: Heston's stochastic volatility model

We price a finite-maturity American put option using Heston's stochastic volatility model [14]. The stock price follows a geometric Brownian motion, in which the volatility parameter σ_t is mean reverting. There are three state variables: time t, stock price S_t, and volatility σ_t. Our three-dimensional grid consists of 50 gridpoints for t, 300 gridpoints for S_t, and 30 gridpoints for σ_t, such that $N = 50 \times 300 \times 30 = 450{,}000$. We take the parameter values as in [6, sec. 3.1]. The resulting exercise boundary is shown in Figure 2. The key finding here is that a plain-vanilla application of the POST method can replicate, with minimal user effort, the output of specialised methods such as [6] for solving American options with stochastic volatility.

Fig. 2. Exercise boundary (i.e. surface) for an American put option with three state variables as in Heston's stochastic volatility model



8 Conclusion

We have extended the applicability of the POST method by [22] to virtually all problems of interest and established it as a rigorous method for solving liquidity-constrained American options with multiple state variables. We have shown by monotonicity arguments in function space that the POST algorithm either finds the solution or demonstrates that no solution exists. This extends the scope of the method relative to [22], who rely on a technical condition that may be hard to verify in practice. Theorem 1 is unlike most results in the literature in that it requires neither bounded nor smooth problem data. The monotonicity of POST carries over to the discretised setting, in which case we additionally show geometric convergence and provide associated convergence bounds. For jump-diffusion processes, dense matrix factorisation may be avoided by an operator-splitting method for which we prove convergence.

Finally, we note that the POST method is a workhorse, not a racehorse, in the sense that for individual problems it is almost certainly possible to develop faster, specialised algorithms. While such algorithms may require less CPU time, they typically require more researcher time, and it is the latter we wish to minimise. In the class of stopping problems that are amenable to discretisation, POST is a near-universal solution method built on rigorous function-space results.

References

[1] Y. Achdou, J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll. Income and wealth distribution in macroeconomics: A continuous-time approach. *The Review of Economic Studies*, Advance online publication, 2021.

[2] J. Ahlberg and E. Nilson. Convergence properties of the spline fit. *Journal of the Society for Industrial and Applied Mathematics*, 11(1):95–104, 1963.

[3] D. Applebaum. *Lévy processes and stochastic calculus.* CUP, 2009.

[4] A. Bensoussan and J.-L. Lions. *Applications of variational inequalities in stochastic control.* Elsevier, 1982.

[5] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–659, 1973.

[6] A. Chockalingam and K. Muthuraman. American options under stochastic volatility. *Operations Research*, 59(4):793–809, 2011.

[7] R. Cottle, J. Pang, and R. Stone. *The linear complementarity problem.* AP, New York, 1992.

[8] M. Dai, Y. K. Kwok, and H. You. Intensity-based framework and penalty formulation of optimal stopping problems. *Journal of Economic Dynamics and Control*, 31(12):3860–3880, 2007.

[9] S. Dai. *Viscosity solutions of optimal stopping problems for Feller processes and their applications*. PhD thesis, University of Liverpool, 2018.

[10] S. Dai and O. Menoukeu-Pamen. Viscosity solution for optimal stopping problems of Feller processes. *arXiv preprint arXiv:1803.03832*, 2018.

[11] P. Dupuis and H. Wang. Optimal stopping with random intervention times. *Advances in Applied probability*, 34(1):141–157, 2002.

[12] 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.

[13] D. Hernández-Hernández and J. A. Minjárez-Sosa. *Optimization, Control, and Applications of Stochastic Systems: In Honor of Onésimo Hernández-Lerma*. Springer, 2012.

[14] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[15] D. Hobson. The shape of the value function under Poisson optimal stopping. *Stochastic Processes and their Applications*, 133:229–246, 2021.

[16] D. Hobson and M. Zeng. Constrained optimal stopping, liquidity and effort. *Stochastic Processes and their Applications*, Advance online publication, 2019.

[17] D. Hobson and M. Zeng. Randomised rules for stopping problems. *Journal of Applied Probability*, 57(3):981–1004, 2020.

[18] Y. Ishikawa. Optimal stopping problem associated with jump-diffusion processes. In *Stochastic Analysis with Financial Applications*, pages 99–120. Springer, 2011.

[19] N. Jacob. *Pseudo differential operators & Markov processes: Markov processes and applications*, volume 3. IPC, 2001.

[20] S. G. Kou. A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.

[21] S. G. Kou and H. Wang. Option pricing under a double exponential jump diffusion model. *Management science*, 50(9):1178–1192, 2004.

[22] R.-J. Lange, D. Ralph, and K. Støre. Real-option valuation in multiple dimensions using Poisson optional stopping times. *Journal of Financial and Quantitative Analysis*, 55(2):653–677, 2020.

[23] R.-J. Lange and C. N. Teulings. The option value of vacant land: Don't build when demand for housing is booming. *Tinbergen Institute Discussion Paper*, 2021-022/IV, 2021.

[24] J. Lempa. Optimal stopping with information constraint. *Applied Mathematics & Optimization*, 66(2):147–173, 2012.

[25] G. Liang and W. Wei. Optimal switching at Poisson random intervention times. *Discrete & Continuous Dynamical Systems-B*, 21(5):1483, 2016.

[26] J.-L. Menaldi and M. Robin. On optimal stopping and impulse control with constraint. In *Modeling, Stochastic Control, Optimization, and Applications*, pages 427–450. Springer, 2019.

[27] J.-L. Menaldi and M. Robin. On some optimal stopping problems with constraint. *SIAM Journal on Control and Optimization*, 54(5):2650–2671, 2016.

[28] H. Morimoto. A singular control problem with discretionary stopping for geometric Brownian motions. *SIAM Jornal on Control and Optimization*, 48(6):3781–3804, 2010.

[29] J. Palczewski and Ł. Stettner. Stopping of functionals with discontinuity at the boundary of an open set. *Stochastic Processes and their Applications*, 121(10):2361–2392, 2011.

[30] R. J. Plemmons. M-matrix characterizations. i — nonsingular m-matrices. *Linear Algebra and its Applications*, 18(2):175–188, 1977.

[31] M. Robin. *Contrôle impulsionnel des processus de Markov.* PhD thesis, Université Paris Dauphine-Paris IX, 1978.

[32] L. C. G. Rogers. Arbitrage with fractional Brownian motion. *Mathematical Finance*, 7(1):95–105, 1997.

[33] L. C. G. Rogers. *Optimal investment.* Springer, 2013.

[34] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes and martingales: Volume 2, Itô calculus.* CUP, 2000.

[35] Ł. Stettner. On general optimal stopping problems using penalty method. *Demonstratio Mathematica*, 45(2):309–323, 2012.

[36] Ł. Stettner and J. Zabczyk. Stochastic version of a penalty method. In *Optimization Techniques*, pages 179–183. Springer, 1980.

[37] Ł. Stettner and J. Zabczyk. Strong envelopes of stochastic processes and a penalty method. *Stochastics*, 4(4):267–280, 1981.

[38] Ł. Stettner. Penalty method for finite horizon stopping problems. *SIAM Journal on Control and Optimization*, 49(3):1078–1099, 2011.

[39] J. H. Witte and C. Reisinger. A penalty method for the numerical solution of Hamilton–Jacobi–Bellman (HJB) equations in finance. *SIAM Journal on Numerical Analysis*, 49(1):213–231, 2011.

[40] K. Zhang, X. Yang, S. Wang, and K. L. Teo. Augmented Lagrangian method applied to American option pricing. *Automatica*, 42:1407–1416, 2006.

[41] K. Zhang, X. Yang, S. Wang, and K. L. Teo. Numerical performance of penalty method for American option pricing. *Optimization Methods & Software*, 25(5):737–752, 2010.

[42] K. Zhang, X. Yang, and Y. Hu. Power penalty method for solving HJB equations arising from finance. *Automatica*, 111:108668, 2020.

[43] 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.

[44] 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.

A Proof of Theorem 1

Take the initialisation $V_\lambda^{(0)}$ from (5), which satisfies $V_\lambda^{(0)} \leq V_\lambda^{(1)}$, [22,15].

We rewrite the POST algorithm (4) as

$$V_\lambda^{(n)}(x) = (r + \lambda - \mathrm{L})^{-1}\,(f + \lambda \max\{g, V_\lambda^{(n-1)}\})(x).$$

for $n = 1, 2, \ldots$. The inverse $(r + \lambda - \mathrm{L})^{-1}$ can be defined via the resolvent formalism, see e.g. [34, pp. 234–238]. This allows the POST algorithm to be written in terms of an expectation operator:

$$V_\lambda^{(n)}(x) = \int_0^\infty \mathrm{e}^{-(\lambda+r)t}\mathbb{E}^x[f(X_t) + \lambda \max\{g(X_t), V_\lambda^{(n-1)}(X_t)\}]\mathrm{d}t, \tag{A.1}$$

where $\mathbb{E}^x$ is an expectation conditional on $X_0 = x$. Here, the probability density $\lambda \exp(-\lambda\, t)\mathrm{d}t$ is recognised as the probability density of the first Poisson arrival time. To simplify the notation, we define the POST operator $\mathcal{T}_\lambda$ operating on a test function $h : \mathcal{X} \to \mathbb{R}$ as

$$(\mathcal{T}_\lambda h)(x) \int_0^\infty \mathrm{e}^{-(\lambda+r)t}\mathbb{E}^x[f(X_t) + \lambda \max\{g(X_t), h(X_t)\}]\mathrm{d}t. \tag{A.2}$$

The POST operator $\mathcal{T}_\lambda$ is pointwise monotonic ($h_1 \leq h_2$ implies $\mathcal{T}_\lambda h_1 \leq \mathcal{T}_\lambda h_2$, both pointwise) so that if $V_\lambda^{(0)} \leq V_\lambda^{(1)}$, then the entire POST sequence is pointwise non-decreasing [22].

The POST algorithm (A.1) can now be written as

$$V_\lambda^{(n)}(x) = (\mathcal{T}_\lambda V_\lambda^{(n-1)})(x), \quad n = 1, 2, \ldots. \tag{A.3}$$

If statement 2 in Theorem 1 holds then statement 1 cannot. The contrapositive is that if statement 2 fails then, by monotonicity of $\{V_\lambda^{(n)}\}$, there exists the pointwise (real-valued) limit $\hat{V} = \lim_{n\to\infty}(\mathcal{T}_\lambda)^n V_\lambda^{(0)}$. In the latter case, Lebesgue monotone convergence theory applies to the monotonic sequence $\{(\mathcal{T}_\lambda)^n V_\lambda^{(0)}\}$ giving, for each x, a limit "under the integral":

$$\begin{aligned}\lim_{n\to\infty}(\mathcal{T}_\lambda)^{n+1}\big(V_\lambda^{(0)}\big)(x) &= \lim_{n\to\infty}\mathcal{T}_\lambda\Big(\mathcal{T}_\lambda)^n V_\lambda^{(0)}\Big)(x)\\ &= \mathcal{T}_\lambda\left(\lim_{n\to\infty}(\mathcal{T}_\lambda)^n\Big(V_\lambda^{(0)}\Big)\right)(x) \;=\; \mathcal{T}_\lambda(\hat{V})(x).\end{aligned}$$

Thus $\hat{V} = \mathcal{T}_\lambda(\hat{V})$, giving statement 1 with $V_\lambda = \hat{V}$.

B Proof of Proposition 2

Linear convergence boils down to the classic bound on the inverse of a diagonally dominant matrix [2, p. 96]. Suppose $\mathbf{L}$ is weakly diagonally dominant: for each row, the sum of the absolute values of the off-diagonal elements does not exceed the absolute value of the diagonal. Suppose further that $\mathbf{L}$ has with nonpositive diagonal elements. Then for any $\mu > 0$, $\mu\mathbf{I} - \mathbf{L}$ is invertible and the norm of $[\mu\mathbf{I} - \mathbf{L}]^{-1}$, induced by the max-norm on $\mathbb{R}^N$, is bounded above by μ^{-1}. Discretised POST corresponds to $\mu = \lambda + r$, which combines with the factor λ on the right hand side of (8) to give a contraction of $\lambda/(\lambda + r)$.

Nonnegative off-diagonal elements of $\mathbf{L}$ ensure for any $\mu > 0$ that the Z-matrix $\mu\mathbf{I} - \mathbf{L}$ is an M-matrix, see [30], thus $[\mu\mathbf{I} - \mathbf{L}]^{-1}$ consists of nonnegative entries. As in the continuous case, it is now straightforward to show that the initialisation in Eq. (7) yields, first, $\mathbf{V}_\lambda^{(1)} \geq \mathbf{V}_\lambda^{(0)}$, hence, by induction, $\mathbf{V}_\lambda^{(n)} \geq \mathbf{V}_\lambda^{(n-1)}$ for all n. Part 2 relies entirely on the nonnegativity of the operator $((r + \lambda)\mathbf{I} - \mathbf{L})^{-1}$, which is contractive having norm bounded above by

$(r+\lambda)^{-1}$, to construct a non-decreasing, geometrically convergent sequence as follows:

$$\left\|\mathbf{V}_\lambda - \mathbf{V}_\lambda^{(n)}\right\|_\infty \leq \sum_{k\geq n}\left\|\mathbf{V}_\lambda^{(k+1)} - \mathbf{V}_\lambda^{(k)}\right\|_\infty \leq \left\|\mathbf{V}_\lambda^{(n+1)} - \mathbf{V}_\lambda^{(n)}\right\|_\infty \sum_{k\geq 0}\left(\frac{\lambda}{\lambda+r}\right)^k$$
$$\leq \frac{\lambda}{r+\lambda}\left\|\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}\right\|_\infty \frac{\lambda+r}{r} \;=\; \frac{\lambda}{r}\left\|\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}\right\|_\infty .$$

C Proof of Proposition 3

We prove Proposition 3 by letting $r+\lambda \to r+\lambda+\delta$ and $\mathbf{L} \to \mathbf{L}_{\text{im}}+\delta$ in Proposition 2 to show that the inverse $[(r+\lambda)\mathbf{I}-\mathbf{L}_{\text{im}}]^{-1}$ exists and has a norm bounded above by $(r+\lambda+\delta)^{-1}$. Successive iterates of (14) give

$$\mathbf{V}_\lambda^{(n+1)} - \mathbf{V}_\lambda^{(n)} = \Big[(r+\lambda)\mathbf{I}-\mathbf{L}_{\text{im}}\Big]^{-1}\times$$
$$\Big(\mathbf{L}_{\mathbf{ex}}(\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}) + \lambda\max\Big\{\mathbf{g},\mathbf{V}_\lambda^{(n)}\Big\} - \lambda\max\Big\{\mathbf{g},\mathbf{V}_\lambda^{(n-1)}\Big\}\Big),$$

such that

$$\left\|\mathbf{V}_\lambda^{(n+1)} - \mathbf{V}_\lambda^{(n)}\right\|_\infty \leq \frac{\|\mathbf{L}_{\text{ex}}\|_\infty}{r+\lambda+\delta}\left\|\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}\right\|_\infty$$
$$+\frac{\lambda}{\delta+r+\lambda}\left\|\max\Big\{\mathbf{g},\mathbf{V}_\lambda^{(n)}\Big\} - \max\Big\{\mathbf{g},\mathbf{V}_\lambda^{(n-1)}\Big\}\right\|_\infty,$$
$$\leq \frac{\|\mathbf{L}_{\text{ex}}\|_\infty}{r+\lambda+\delta}\left\|\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}\right\|_\infty$$
$$+\frac{\lambda}{\delta+r+\lambda}\left\|\mathbf{V}_\lambda^{(n)} - \mathbf{V}_\lambda^{(n-1)}\right\|_\infty .$$

The first inequality uses the fact that the norm is sub-additive as well as sub-multiplicative. where the last term uses monotonicity of $\|\cdot\|_\infty$. Hence contractivity, and geometric convergence at rate $\frac{\|\mathbf{L}_{\text{ex}}\|_\infty+\lambda}{\delta+r+\lambda}$, follows if $\|\mathbf{L}_{\text{ex}}\|_\infty < \delta + r$.