SIAM J. NUMER. ANAL.
Vol. xx, No. x, pp. x–x

WEAKLY CHAINED MATRICES AND IMPULSE CONTROL

P. AZIMZADEH* AND P.A. FORSYTH†

Abstract. This work is motivated by numerical solutions to *Hamilton-Jacobi-Bellman quasi-variational inequalities* (HJBQVIs) associated with *combined stochastic and impulse control* problems. In particular, we consider (i) *direct control*, (ii) *penalized*, and (iii) *explicit control* schemes applied to the HJBQVI problem. Scheme (i) takes the form of a *Bellman problem* involving an operator which is not necessarily contractive. We consider the well-posedness of the Bellman problem and give sufficient conditions for convergence of the corresponding policy iteration. To do so, we use *weakly chained diagonally dominant matrices*, which give a graph-theoretic characterization of weakly diagonally dominant M-matrices. We compare schemes (i)–(iii) under the following examples: (a) optimal control of the exchange rate, (b) optimal consumption with fixed and proportional transaction costs, and (c) pricing guaranteed minimum withdrawal benefits in variable annuities. Perhaps controversially, we find that one should abstain from using scheme (i).

Key words. Hamilton-Jacobi-Bellman (HJB) equation, combined stochastic and impulse control, Bellman problem, policy iteration, optimal exchange rate, optimal consumption, GMWB

AMS subject classifications. 65N06, 93E20

1. Introduction.

This work is motivated by the computation of numerical solutions to *Hamilton-Jacobi-Bellman quasi-variational inequalities* (HJBQVI) associated with *combined stochastic and impulse control*. These problems are of the form:

PROBLEM 1.1. *Find a viscosity solution (see [16, Definition 2.2]) of the HJBQVI*

$$\max\left(\sup_{w\in W}\left\{\frac{\partial u}{\partial t}+\text{Ł}^{w}u-\rho u+f^{w}\right\},\mathcal{M}u-u\right)=0 \text{ on } [0,T)\times\left(\overline{\Omega}\setminus\Lambda\right) \tag{1.1}$$

$$\max\left(g-u,\mathcal{M}u-u\right)=0 \text{ on } \left([0,T)\times\Lambda\right)\cup\left(\{T\}\times\overline{\Omega}\right) \tag{1.2}$$

where $\Omega\subset\mathbb{R}^{d}$ is open, $\Lambda\subset\partial\Omega$, $\text{Ł}^{w}:=\text{Ł}(t,x,w)$ is the generator of an SDE, $f^{w}:=f(t,x,w)$ is a forcing term, and $\mathcal{M}$ is the impulse (a.k.a. intervention) operator

$$\mathcal{M}u\left(t,x\right):=\sup_{z\in Z(t,x)}\left\{u\left(t,\Gamma\left(t,x,z\right)\right)+K\left(t,x,z\right)\right\}. \tag{1.3}$$

If $Z(t,x)$ is empty at a particular point (t,x), $\mathcal{M}u(t,x)$ is understood to take the value $-\infty$, corresponding to no impulses being allowed at that point.

After discretizing the above HJBQVI, we obtain the fixed point problem

$$\text{find } v\in\mathbb{R}^{M} \text{ such that } v=\max\left(\sup_{w\in\mathcal{W}}L\left(w\right)v+c\left(w\right),\sup_{z\in\mathcal{Z}}B\left(z\right)v+k\left(z\right)\right) \tag{1.4}$$

where $L(w)$ and $B(z)$ are contractive and nonexpansive operators, respectively. It is understood that the supremum and maximum are element-wise and controls are "row-decoupled" (see §2). [9] gives sufficient conditions for convergence of a policy iteration to the unique solution of (1.4).

However, convergence in [9] is conditional on the choice of initial guess [9, Theorem 2 (iii)]. We remove this constraint. More importantly, [9] restricts the admissible set

*David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 pazimzad@uwaterloo.ca
†David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 paforsyt@uwaterloo.ca

of controls and imposes a strong assumption on $B(z)$ (of which assumption (H2) in this work is an analogue) to ensure convergence. Unfortunately, reasonable instances of problem (1.4) (including examples in this work) do not necessarily satisfy this condition. We show that, under a much weaker assumption, a solution to (1.4) is unique. When (H2) is not satisfied directly, we provide a way to construct this solution by considering a "modified problem" that satisfies (H2). Roughly speaking, one arrives at the modified problem by removing some suboptimal controls from the control set. However, this procedure is ad hoc (i.e. problem dependent).

To establish the above relaxations, we use *weakly chained diagonally dominant* (WCDD) matrices. WCDD matrices give a graph-theoretic characterization of weakly diagonally dominant M-matrices (Theorem 3.5). The WCDD matrix approach to the convergence of policy iteration applied to (1.4) is intuitive and appeals to well-known results on policy iteration (Proposition 2.2).

We focus only on *unconditionally stable* schemes for the HJBQVI problem:

- A "*direct control*" scheme takes the form of problem (1.4). As discussed above, some suboptimal controls may need to be removed from the original problem to ensure convergence of a policy iteration.
- On the other hand, convergence of a policy iteration associated with a "*penalized*" scheme [27] (a.k.a. a penalty method) is a trivial consequence of the strict diagonal dominance of the input matrices to policy iteration.
- An "*explicit control*" scheme (first used for HJBQVIs in [11]) handles controlled terms using information from the previous timestep, and hence does not need an iterative method altogether. However, this scheme requires that the control w in $Ł^w$ appears only in the coefficient of the first-order term.

The ad hoc removal of suboptimal controls makes the direct control scheme less robust than its counterparts, for which control sets need not be altered to ensure convergence. It is thus natural to ask if there is an advantage to using a direct control scheme. To answer this, we apply each scheme to the following examples:

- optimal control of the exchange rate;
- optimal consumption with fixed and proportional transaction costs;
- pricing guaranteed minimum withdrawal benefits in variable annuities.

The explicit control scheme only requires a single linear solve per timestep since no iterative method is needed. However, an explicit control scheme cannot be used if the control w appears in the diffusion coefficient of $Ł^w$ (or if the underlying process is a Lévy process with controlled arrival rate). We find that the penalized scheme performs at least as well the direct control scheme. Both produce near identical results and often require roughly the same amount of computation. In the specific case of the optimal consumption problem, the penalized scheme even outperforms the direct control scheme, taking only a few policy iterations to converge per timestep. Since control sets need not be altered for a penalized scheme, we recommend–perhaps controversially–against a direct control scheme.

We mention that in the infinite-horizon setting ($T = \infty$), optimal consumption with fixed and proportional transaction costs was considered numerically in [10] using iterated optimal stopping, a theoretical tool [23, Chapter 7, Lemma 7.1] for the construction of solutions that has found its way into numerical implementations [19, 4]. Computationally, for finite-horizon problems ($T < \infty$), iterated optimal stopping has high space complexity [2], and is thus not considered here. Also not considered here is the simulation of penalized backward stochastic differential equations [17], a recent alternative well-suited to high-dimensional problems.

In this work, we restrict our attention to problems of dimension three or lower. To keep focus on the interesting aspects of impulse control, we assume that between impulses, the underlying stochastic process associated with the HJBQVI is a Brownian motion with drift $\mu \coloneqq \mu(t,x,w)$ and scaling $\sigma \coloneqq \sigma(t,x,w)$ (for Lévy process with nontrivial arrival rate see, e.g., [12]). This allows us to write

$$\text{Ł}^{w} u\left(t,x\right) \coloneqq \frac{1}{2} \operatorname{trace}\left(\sigma\left(t,x,w\right) \sigma^{\dagger}\left(t,x,w\right) D_x^2 u\left(t,x\right)\right) + \left\langle \mu\left(t,x,w\right), D_x u\left(t,x\right)\right\rangle .$$

We mention here that problem (1.4) can also be interpreted as a Bellman problem associated with an infinite-horizon Markov decision process (MDP) with vanishing discount (Example 4.1). In fact, (1.4) appears to be a generalization of a reflecting boundary problem (see, e.g., the monograph of Kushner and Dupuis [18, pg. 39–40]). In the context of MDPs, $L(w)$ and $B(z)$ capture the transition probabilities at states with nonvanishing and vanishing discount factors, respectively. A WCDD matrix condition (Corollary 4.3) guarantees the convergence of policy iteration to the unique solution of the optimal control problem. Intuitively, this condition ensures that the underlying Markov chain arrives (with positive probability) at a state with nonvanishing discount independent of the initial state.

We summarize some of our main findings below:

- Policy iteration applied to monotone direct control schemes for the HJBQVI frequently fails due to the possible singularity of the matrix iterates.
- We establish provably convergent techniques to eliminate singularity. However, applying these techniques is problem-dependent.
- We show that a monotone penalized schemes never fails. Numerical tests on three classic HJBQVI problems confirm that penalized schemes perform at least as well as (and sometimes much better than) direct control schemes.

2. Policy iteration. The discretized equations corresponding to (1.1)–(1.3) can be posed as a Bellman problem:

$$\text{find } v \in \mathbb{R}^M \text{ such that } \sup_{P \in \mathcal{P}} \left\{-A\left(P\right) v + b\left(P\right)\right\} = 0 \tag{2.1}$$

where $A\colon \mathcal{P} \to \mathbb{R}^{M\times M}$ and $b\colon \mathcal{P} \to \mathbb{R}^M$. It is understood that(i) $\mathcal{P} \coloneqq \prod_{i=1}^M \mathcal{P}_i$ is a finite product of nonempty sets, (ii) controls are row-decoupled:

$$\left[A\left(P\right)v\right]_i \text{ and } \left[b\left(P\right)\right]_i \text{ depend only on } P_i \in \mathcal{P}_i,$$

(iii) the order on $\mathbb{R}^M$ (resp. $\mathbb{R}^{M\times M}$) is element-wise:

$$\text{for } x, y \in \mathbb{R}^M,\ x \geq y \text{ if and only if } x_i \geq y_i \text{ for all } i,$$

and (iv) the supremum is induced by this order:

$$\text{for } \left\{x\left(P\right)\right\}_{P\in\mathcal{P}} \subset \mathbb{R}^M,\ \sup_{P\in\mathcal{P}} x\left(P\right) \text{ is a vector with components } \sup_{P\in\mathcal{P}} \left[x\left(P\right)\right]_i .$$

Let $\textsc{Solve}(A, b, x^0)$ denote a call to a linear solver for $Ax = b$ with initial guess x^0 (algebraically, $\textsc{Solve}$ computes x exactly; in practice, an iterative solver is used and the choice of x^0 affects performance). A *policy iteration* algorithm is given by:

$\textsc{Policy-Iteration}(\mathcal{P}, A(\cdot), b(\cdot), v^0)$

for $\ell = 1, 2, \ldots$
  Pick P^ℓ such that $-A(P^\ell)v^{\ell-1} + b(P^\ell) = \sup_{P\in\mathcal{P}}\{-A(P)v^{\ell-1} + b(P)\}$
  $v^\ell \coloneqq \textsc{Solve}(A(P^\ell), b(P^\ell), v^{\ell-1})$

Definition 2.1 (Monotone matrix). *A real square matrix A is monotone (in the sense of Collatz) if for all real vectors v, $Av \geq 0$ implies $v \geq 0$.*

We use the following assumptions:

(H0) $P \mapsto A(P)^{-1}$ is bounded (on the subset of $\mathcal{P}$ on which $A(P)$ is nonsingular).

(H1) (i) A and b are bounded and (ii) for all x in $\mathbb{R}^M$, there exists P_x in $\mathcal{P}$ such that $-A(P_x)x + b(P_x) = \sup_{P \in \mathcal{P}}\{-A(P)x + b(P)\}$.

Proposition 2.2 (Convergence of policy iteration). *Suppose* (H0), (H1), *and that $A(P)$ is a monotone matrix for all P in $\mathcal{P}$. $(v^\ell)_{\ell=1}^\infty$ defined by* Policy-Iteration *is nondecreasing and converges to the unique solution v of* (2.1). *Moreover, if $\mathcal{P}$ is finite, convergence occurs in at most $|\mathcal{P}|$ iterations (i.e. $v^{|\mathcal{P}|} = v^{|\mathcal{P}|+1} = \cdots$).*

The monotone convergence of $(v^\ell)_{\ell=1}^\infty$ to the unique solution of (2.1) can be proven similarly to Theorem A.3 of the appendix. See [6, Theorem 2.1] for a proof of finite termination when $\mathcal{P}$ is finite.

In practice, $\mathcal{P}$ is often finite, in which case (H0) and (H1) are trivial.

Remark 2.3. *Theorem* A.3 *gives a constructive proof of the existence and uniqueness of solutions to* (2.1) *independent of* (H1.ii). *Owing to this, results that rely on Proposition* 2.2 *can be relaxed to exclude* (H1.ii), *with the caveat that when $\mathcal{P}$ is infinite,* Policy-Iteration *be replaced by* Sequential-Policy-Iteration *(see appendix) and that the resulting sequences $(v_\ell)_{\ell=1}^\infty$ are not necessarily nondecreasing.*

3. Weakly chained diagonally dominant matrices.

We say row i of a complex matrix $A \coloneqq (a_{ij})$ is strictly diagonally dominant (SDD) if $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$. We say A is SDD if all of its rows are SDD. Weakly diagonally dominant (WDD) is defined with weak inequality instead.

Definition 3.1. *A complex square matrix A is said to be a weakly chained diagonally dominant (WCDD) if it satisfies:*

(i) A is WDD;

(ii) for each row r, there exists a path in the graph of A from r to an SDD row p.

Recall that the directed graph of an $M \times M$ complex matrix $A \coloneqq (a_{ij})$ is given by the vertices $\{1, \ldots, M\}$ and edges defined as follows: there exists an edge from i to j if and only if $a_{ij} \neq 0$. Note that if r is itself an SDD row, the trivial path $r \rightarrow r$ satisfies the requirement of (ii) in the above.

The nonsingularity of WCDD matrices is proven in [26] as a corollary to other results. We provide an elementary proof for the convenience of the reader:

Lemma 3.2. *A WCDD matrix is nonsingular.*

Proof. Suppose $\lambda = 0$ is an eigenvalue of $A \coloneqq (a_{ij})$. Let $v \neq 0$ be an associated eigenvector with components of modulus at most unity. Let r be such that $|v_r| = 1 \geq |v_j|$ for all j. By the Gershgorin circle theorem,

$$|\lambda - a_{rr}| = |a_{rr}| \leq \sum_{j \neq r} |a_{rj}| \, |v_j| \leq \sum_{j \neq r} |a_{rj}| \, .$$

Since A is WDD, it follows that $|a_{rr}| = \sum_{j \neq r} |a_{rj}|$, and hence r is not an SDD row. Therefore, there exists a path $r \rightarrow p_1 \rightarrow \cdots \rightarrow p_k$ where p_k is an SDD row. Since

$$|a_{rr}| = \sum_{j \neq r} |a_{rj}| \, |v_j| = \sum_{j \neq r} |a_{rj}| \, ,$$

it follows that $|v_j| = 1$ whenever $|a_{rj}| \neq 0$. Because $|a_{rp_1}| \neq 0$, $|v_{p_1}| = 1$. Repeating the same argument as above with $r = p_1$ yields $|a_{p_1 p_1}| = \sum_{j \neq p_1} |a_{p_1 j}|$, and hence p_1 is not an SDD row. Continuing the procedure, p_k is not SDD, a contradiction. ∎

We recall some well-known classes of matrices:

Definition 3.3. *A Z-matrix is a real matrix with nonpositive off-diagonals.*

Definition 3.4. *An M-matrix is a monotone Z-matrix.*

We are now ready to state a fundamental characterization of WDD M-matrices:

Theorem 3.5 (Characterization theorem). *The following are equivalent:*
(i) A is a WCDD Z-matrix with positive diagonals;
(ii) A is a WDD M-matrix.

Proof. Since a nonsingular WDD Z-matrix with positive diagonals is an M-matrix (a consequence of, e.g., [24, Theorem 1.A_3]), (i) implies (ii) follows by Lemma 3.2.

As for the converse, since an M-matrix has positive diagonal elements (a consequence of, e.g., [24, Theorem 1.K_{35}]), it is sufficient to show that any WDD Z-matrix with positive diagonals not satisfying (ii) of Definition 3.1 is singular. Let A be such a matrix. It follows that there exists a WDD row r of A from which there is no path to an SDD row of A. Let $V \coloneqq \{i_1, \ldots, i_k\} \subset \{1, \ldots, \dim A\}$ be the set of rows reachable from r (note that $r \in V$ since $a_{rr} > 0$). Let v_i be unity if i is in V and zero otherwise. Then, $Av = 0$ and $v \neq 0$, and hence (ii) implies (i). ∎

This characterization is tight: an M-matrix need not be WCDD (e.g. $\begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}$).

We mention that (i) $\implies$ (ii) of Theorem 3.5 appears in [7]. Therein, WCDD Z-matrices with positive diagonals are referred to as matrices with *positive type.* To the authors' best knowledge, the converse does not appear in the literature.

4. The fixed point problem (1.4).

4.1. Convergence of policy iteration.

We assume $\mathcal{W} \coloneqq \prod_{i=1}^{M} \mathcal{W}_i$ and $\mathcal{Z} \coloneqq \prod_{i=1}^{M} \mathcal{Z}_i$ appearing in problem (1.4) are finite products of nonempty sets. Let

$$\mathcal{P} \coloneqq \textstyle\prod_{i=1}^{M} \mathcal{P}_i \text{ where } \mathcal{P}_i \coloneqq \mathcal{W}_i \times \mathcal{Z}_i \times \mathcal{D}_i \text{ and } \emptyset \neq \mathcal{D}_i \subset \{0, 1\}. \tag{4.1}$$

We associate with each $\psi \coloneqq (\psi_1, \ldots, \psi_M)$ in $\mathcal{D} \coloneqq \prod_{i=1}^{M} \mathcal{D}_i$ a diagonal matrix Ψ with $[\Psi]_{ii} = \psi_i$. We use ψ and Ψ interchangeably. We write $P \coloneqq (w, z, \psi) \in \mathcal{P}$ where $w \in \mathcal{W}$ and $z \in \mathcal{Z}$. We can transform problem (1.4) into the form (2.1) by taking

$$\begin{aligned} A(P) &\coloneqq (I - \Psi)(I - L(w)) + \Psi(I - B(z)); \\ b(P) &\coloneqq (I - \Psi)c(w) + \Psi k(z). \end{aligned} \tag{4.2}$$

Before considering the well-posedness of problem (2.1) subject to (4.1) and (4.2), we visit, as a motivating example, an infinite-horizon MDP with vanishing discount:

Example 4.1. *Let $(X^n)_{n \geq 0}$ be a controlled homogeneous Markov chain on a finite state space $\{1, \ldots, M\}$. A control at state i is a member of $\mathcal{P}_i$ in (4.1) and written $P_i \coloneqq (w_i, z_i, \psi_i)$. The controlled transition probability of the Markov chain is*

$$\mathbb{P}\left(X^{n+1} = j \mid X^n = i\right) = \begin{cases} w_{ij} & \text{if } \psi_i = 0 \\ z_{ij} & \text{if } \psi_i = 1 \end{cases}$$

where $w_i \coloneqq (w_{i1}, \ldots, w_{iM}) \geq 0$ and $w_i e = 1$ (similarly for z_i). That is, members of $\mathcal{W}_i$ and $\mathcal{Z}_i$ are M-dimensional probability vectors. For discount factor $\rho > 0$, let

$$v_i \coloneqq \sup_{P \in \mathcal{P}} \mathbb{E}\left[\sum_{n=0}^{\infty} U(X^n, P) \prod_{m=0}^{n} D(X^m, P) \,\middle|\, X^0 = i\right] \text{ for all } 1 \leq i \leq M \tag{4.3}$$

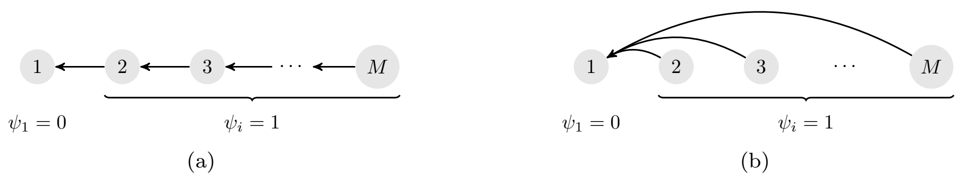


FIGURE 4.1.1: *Graphs of possible matrices $B(z)$ from Example* 4.4

where

$$U(i,P) \coloneqq \begin{cases} \tilde{c}_i(w_i) & \text{if } \psi_i = 0 \\ k_i(z_i) & \text{if } \psi_i = 1 \end{cases} \text{ and } D(i,P) \coloneqq \begin{cases} 1/(1+\rho) & \text{if } \psi_i = 0; \\ 1 & \text{if } \psi_i = 1. \end{cases}$$

The problem of finding v satisfying (4.3) *is put into the form of problem* (2.1) *subject to* (4.1) *and* (4.2) *by taking $[L(w)]_{ij} \coloneqq w_{ij}/(1+\rho)$, $[c(w)]_i \coloneqq \tilde{c}_i(w_i)/(1+\rho)$, $[B(z)]_{ij} \coloneqq z_{ij}$, and $[k(z)]_i \coloneqq k_i(z_i)$ [9, Lemma 5].*

In the above, states i on which $\psi_i = 1$ are the "trouble" states with vanishing discount factor. In fact, requiring $\psi_i = 0$ for all i returns us to a nonvanishing discount factor problem whose well-posedness is easy to establish.

The following assumptions will prove paramount:

(H2) For each $P \coloneqq (w,z,\psi)$ in $\mathcal{P}$ and state i with $\psi_i = 1$, there exists a path in the graph of $B(z)$ from i to a state j with $\psi_j = 0$.

(H3) For all $P \coloneqq (w,z,\psi)$ in $\mathcal{P}$, $I - L(w)$ and $I - B(z)$ are Z-matrices with nonnegative diagonals that are SDD and WDD, respectively.

THEOREM 4.2. *Suppose* (H0)–(H3). $(v^\ell)_{\ell=1}^\infty$ *defined by* POLICY-ITERATION *is nondecreasing and converges to the unique solution v of problem* (2.1) *subject to* (4.1) *and* (4.2). *Moreover, if $\mathcal{P}$ is finite, convergence occurs in at most $|\mathcal{P}|$ iterations.*

Proof. (H2) and (H3) ensure that $A(P)$ is a WCDD Z-matrix with positive diagonals. The desired result follows from Theorem 3.5 and Proposition 2.2. ∎

COROLLARY 4.3. *Consider Example* 4.1. *Suppose* (H0)–(H2). $(v^\ell)_{\ell=1}^\infty$ *defined by* POLICY-ITERATION *converges to v in $\mathbb{R}^M$ satisfying* (4.3).

An example satisfying (H2) is given:

EXAMPLE 4.4. *Consider Example* 4.1. *Suppose all $P \coloneqq (w,z,\psi)$ in $\mathcal{P}$ satisfy*

$$\psi_1 = 0 \text{ and } z_{ij} = 0 \text{ if } 1 < i \leq j. \tag{4.4}$$

This corresponds to(i) a nonvanishing discount at state 1 *and that (ii) transitions from a state with vanishing discount are unidirectional (if $\psi_{X^n} = 1$, $X^{n+1} < X^n$ a.s.). See Figure* 4.1.1 *for example graphs of $B(z)$ subject to* (4.4).

4.2. Uniqueness.

Let

$$\mathbb{L}v \coloneqq \sup_{w \in \mathcal{W}} \{L(w)v + c(w)\} \text{ and } \mathbb{B}v \coloneqq \sup_{z \in \mathcal{Z}} \{B(z)v + k(z)\}. \tag{4.5}$$

The condition (H2) turns out to be too restrictive for some problems of interest. However, the following weaker property of $\mathbb{B}$ is not unusual:

(H4) For each state i and solution v of (2.1) subject to (4.1) and (4.2), there exist integers $m(i)$ and n such that $0 \leq n \leq m(i)$ and $[\mathbb{B}^{m(i)+1}v]_i < [\mathbb{B}^n v]_i$.

LEMMA 4.5. *Suppose* (H4) *and that* $B(z)$ *is a matrix for all* z *in* $\mathcal{Z}$. *Let* $(P^\ell)_{\ell \geq 0}$ *be a sequence in* $\mathcal{P}$ *and* v *a solution of problem* (2.1) *subject to* (4.1) *and* (4.2) *with*

$$-A\left(P^{\ell}\right) v+b\left(P^{\ell}\right) \rightarrow \sup _{P \in \mathcal{P}}\{-A(P) v+b(P)\}=0.$$

There exists $\ell_0 \geq 0$ *such that for each* $\ell \geq \ell_0$ *and state* i *with* $\psi_i^\ell = 1$, *there exists a path in the graph of* $B(z^\ell)$ *from* i *to a state* j *with* $\psi_j^\ell = 0$.

Proof. Suppose the contrary. A pigeonhole principle argument yields the existence of a subsequence $(P^{\ell_q})_{q \geq 0}$ of $(P^\ell)_{\ell \geq 0}$ such that for some i and all q, $\psi_i^{\ell_q} = 1$ and there does not exist a path in the graph of $B(z^{\ell_q})$ from i to a state j with $\psi_j = 0$. Since $v_r = [\mathbb{B}v]_r$ for all r with $\psi_r = 1$ and the limit of a convergent sequence equals to the limit of any of its subsequences, $v_i = [\mathbb{B}v]_i = [\mathbb{B}^2 v]_i = \cdots = [\mathbb{B}^n v]_i = \cdots = [\mathbb{B}^{m(i)+1} v]_i$. Therefore, $[\mathbb{B}^n v]_i = [\mathbb{B}^{m(i)+1} v]_i$. But $[\mathbb{B}^{m(i)+1} v]_i < [\mathbb{B}^n v]_i$ by (H4), a contradiction. ☐

If we take the trivial path $i \to i$ as having length zero, the proof above also implies that for ℓ sufficiently large, there exists a path of length at most $m(i)$ (given by (H4)) in the graph of $B(z^\ell)$ from any state i to a state j with $\psi_j^\ell = 0$. An example is given:

EXAMPLE 4.6. *Consider Example* 4.1 *with* $\mathcal{Z}_i = \mathcal{Z}_j$ *for all states* i *and* j. *For all states* i, *let* $k_i(z_i) \coloneqq -C < 0$. *It follows that for all* x *in* $\mathbb{R}^M$,

$$\mathbb{B}^2 x=\sup _{z, z^{\prime} \in \mathcal{Z}}\left\{B(z) B\left(z^{\prime}\right) x\right\}-2 C<\sup _{z \in \mathcal{Z}}\{B(z) x\}-C=\mathbb{B} x,$$

so that (H4) *is satisfied with* $m(i) = n = 1$ *for all* i. *In this case, the control shown in Figure* 4.1.1a *cannot be optimal (*$P^\star$ *is optimal if for a solution* v, $-A(P^\star)v + b(P^\star) = \sup_{P \in \mathcal{P}}\{-A(P)v + b(P)\}$*) since any path from* $i > 2$ *to* $j = 1$ *is of length greater than* $m(i) = 1$. *Intuitively, the controller pays twice the cost to apply* $\mathbb{B}$ *twice.*

We can now prove uniqueness independent of (H2):

THEOREM 4.7 (Uniqueness). *Suppose* (H0), (H3), *and* (H4). *A solution of problem* (2.1) *subject to* (4.1) *and* (4.2) *is unique.*

Proof. Let x and y be two solutions and $(P^\ell)_{\ell \geq 0}$ be a sequence in $\mathcal{P}$ such that

$$-A\left(P^{\ell}\right) y+b\left(P^{\ell}\right) \rightarrow \sup _{P \in \mathcal{P}}\{-A(P) y+b(P)\}=0.$$

It follows from (H3), (H4), and Lemma 4.5 that for ℓ sufficiently large, $A(P^\ell)$ is a WCDD Z-matrix with positive diagonals, and hence an M-matrix by Theorem 3.5. For some sequence $(\epsilon^\ell)_{\ell \geq 0}$ in $\mathbb{R}^M$ with $\epsilon^\ell \to 0$, we can write

$$-A\left(P^{\ell}\right) y+b\left(P^{\ell}\right)+\epsilon^{\ell}=0=\sup _{P \in \mathcal{P}}\{-A(P) x+b(P)\} \geq-A\left(P^{\ell}\right) x+b\left(P^{\ell}\right),$$

so that $A(P^\ell)(x-y) \geq -\epsilon^\ell$. Since the inverse of a monotone matrix has nonnegative elements and $P \mapsto A(P)^{-1}$ is bounded by (H0), $x - y \geq 0$. Similarly, $y - x \geq 0$. ☐

REMARK 4.8. *Unfortunately, the conditions of Theorem* 4.7 *cannot guarantee that the iterates* $(v^\ell)_{\ell=1}^\infty$ *given by policy iteration are well-defined, as* $A(P^\ell)$ *may be singular for some* $\ell \geq 1$. *This is demonstrated in the following example:*

EXAMPLE 4.9 (Failure of policy iteration). *Consider Example* 4.1. *For all states* i, *let* $\mathcal{Z}_i \coloneqq \{e^j\}_{j=1}^M$ *be the set of standard basis vectors and*

$$k_i\left(z_i\right) \coloneqq -C-\sum_j z_{i j}|i-j|<0 \text { where } C>0.$$

As in Example 4.6, (H4) *is satisfied due to the fixed cost* $C > 0$.

Suppose there exists a state r with $1 \in \mathcal{D}_r$ and $\tilde{c}_r(w_r) < -C(1+\rho)$ for all controls w_r in $\mathcal{W}_r$. It is readily verified that Policy-Iteration *initialized with the zero vector $v^0 := 0$ picks $P^1 := (w^1, z^1, \psi^1)$ with $z_r^1 = e^r$ and $\psi_r^1 = 1$. It follows that*

$$\left[A\left(P^1\right)\right]_{rj} = \left[I - B\left(z^1\right)\right]_{rj} = [I]_{rj} - [I]_{rj} = 0 \text{ for all } j$$

so that $A(P^1)$ is singular, and hence v^1 is undefined.

For any $\ell \geq 1$, it is possible to construct more complicated examples in which the matrices $A(P^1), \ldots, A(P^{\ell-1})$ are nonsingular while $A(P^\ell)$ is singular. That is, policy iteration can fail at any iterate.

4.3. Policy iteration on a modified problem. As demonstrated in the previous section, if (H2) is not satisfied, policy iteration may fail. We may, however, hope to construct a solution by performing policy iteration on a "modified problem" with control set $\mathcal{P}'$ obtained by removing controls P in $\mathcal{P}$ that render $A(P)$ singular.

We define (H1)$'$ by replacing all occurrences of $\mathcal{P}$ with $\mathcal{P}'$ in the definition of (H1). (H2)$'$ and (H3)$'$ are defined similarly. We can now state the above idea precisely:

Theorem 4.10. *Let $\mathcal{P}' := \prod_{i=1}^M \mathcal{P}_i'$ where $\mathcal{P}_i' \subset \mathcal{P}_i$ is nonempty. Suppose* (H0), (H1)$'$, (H2)$'$, (H3), (H4), *and*

$$\text{for all } v \text{ in } \mathbb{R}^M, \quad \sup_{P \in \mathcal{P}'} \{-A(P)v + b(P)\} = 0 \implies \sup_{P \in \mathcal{P}} \{-A(P)v + b(P)\} = 0. \tag{4.6}$$

$(v^\ell)_{\ell=1}^\infty$ defined by Policy-Iteration$(\mathcal{P}', \ldots)$ *is nondecreasing and converges to the unique solution of problem* (2.1) *subject to* (4.1) *and* (4.2)*. Moreover, if $\mathcal{P}'$ is finite, convergence occurs in at most $|\mathcal{P}'|$ iterations.*

Proof. Since $\mathcal{P}' \subset \mathcal{P}$, it follows immediately that (H0)$'$ and (H3)$'$ are satisfied, so that by Theorem 4.2, $(v^\ell)_{\ell=1}^\infty$ is well-defined and converges to the unique solution v of the modified problem (under $\mathcal{P}'$). By (4.6), v solves (2.1). Since solutions to (2.1) are unique by Theorem 4.7, the desired result follows. ∎

Example 4.11. *Consider Example* 4.1*. Let(i) $\mathcal{Z}$ and k be given as in Example* 4.9*, (ii) $\mathcal{W}_i$ be the set of all M-dimensional probability vectors w_i with $w_{ij} = 0$ whenever $|i-j| > 1$, (iii) $\tilde{c}$ be continuous and bounded, and (iv) $\bar{c} := \max_{w \in \mathcal{W}} \tilde{c}(w)$ with $\bar{c}_{i-1} \geq \bar{c}_i$ for all $1 < i \leq M$. As in Example* 4.9, (H2) *is not satisfied.*

If $\mathcal{P}'$ is taken to be all $P := (w, z, \psi)$ in $\mathcal{P}$ satisfying (4.4)*, the conditions of Theorem* 4.10 *are satisfied.*

Proof. It is straightforward to verify (H0), (H1)$'$, (H2)$'$, (H3), and (H4). We show that the solution of the modified problem solves the original problem (i.e. (4.6) is satisfied). We write $\mathcal{P}_i' := \mathcal{W}_i \times \mathcal{Z}_i' \times \mathcal{D}_i'$ and define $[\mathbb{B}'x]_i$ for $i > 1$ by replacing $\mathcal{Z}$ with $\mathcal{Z}' := \prod_{i=1}^M \mathcal{Z}_i'$ in (4.5).

We first show that the solution v to the modified problem is nonincreasing:

$$v_{i-1} \geq v_i \text{ for all } 1 < i \leq M.$$

Suppose the contrary. Let $r > 1$ be the minimal element such that $v_{r-1} < v_r$. If $v_r = [\mathbb{B}'v]_r$, then $v_r = v_j - C - |r-j|$ for some $j < r$. Either $j = r-1$ or

$$v_{r-1} \geq [\mathbb{B}'v]_{r-1} \geq v_j - C - |(r-1) - j| \geq v_r$$

(both are contradictions). It follows that $v_r = [\mathbb{L}v]_r$. Letting $w_0, v_0 := 0$ for notational convenience, assumption (ii) implies

$$v_{r-1} \geq [\mathbb{L}v]_{r-1} = \max_{w \in \mathcal{W}} \left\{ \sum_{j=r-2}^{r} w_{r-1,j} v_j + [\tilde{c}(w)]_{r-1} \right\} / (1+\rho) \geq \frac{v_{r-1} + \bar{c}_{r-1}}{1+\rho}$$

so that $\rho v_{r-1} \geq \overline{c}_{r-1}$. If $v_r \geq v_{r+1}$, it follows similarly from $v_r = [\mathbb{L}v]_r$ that $\rho v_r \leq \overline{c}_r$ so that $v_{r-1} \geq v_r$ (a contradiction) and hence it must be the case that $v_r < v_{r+1}$. We can repeat this argument inductively to arrive at the contradiction

$$v_{r-1} < v_r < \cdots < v_M \text{ and } \rho v_{r-1} \geq \overline{c}_{r-1} \geq \overline{c}_M \geq \rho v_M.$$

Since v is nonincreasing, $v \geq \mathbb{B}v$ (it is suboptimal to take $\psi_i = 1$ and $z_{ij} = 1$ for states i and j with $j \geq i$) and hence v solves the original problem. □

5. Numerical schemes for the HJBQVI problem (1.1)–(1.3). Barles and Souganidis [3] show that a numerical scheme converges to the unique viscosity solution of a fully nonlinear second order equation satisfying a *comparison result* if it is ℓ_∞ *stable*, *consistent* (with respect to a discretization parameter h approaching zero), and *monotone* in the viscosity sense (not to be confused with Definition 2.1). Comparison results for the HJBQVI (1.1)–(1.3) are provided in [25, Theorem 5.11].

All numerical schemes herein are on a rectilinear grid

$$\left\{t^1, \ldots, t^N\right\} \times \left\{x_1^1, x_2^1, \ldots\right\} \times \cdots \times \left\{x_1^d, x_2^d, \ldots\right\}$$

where $0 =: t^1 < \cdots < t^N := T$ and $x_1^j < x_2^j < \cdots$ for all j. Multi-indices are used (i.e. $x_i := (x_{i_1}, \ldots, x_{i_d})$). M denotes the number of spatial points x_i. For functions $q := q(t, x)$ defined on $[0, T] \times \mathbb{R}^d$, the shorthands $q_i^n := q(t^n, x_i)$ and $q^n(x) := q(t^n, x)$ are employed. In the absence of ambiguity, we use q^n to denote the vector with components q_i^n and take $\Delta t := t^{n+1} - t^n$.

Control sets W and $Z(t, x)$ are approximated by finite sets $\emptyset \neq W_h \subset W$ and $Z_h(t, x) \subset Z(t, x)$. The reader concerned with consistency should impose some regularity to justify this approximation, such as:(i) W is compact, (ii) Z is everywhere compact and continuous with respect to the Hausdorff metric, and (iii) $\max_{w \in W} \min_{w_h \in W_h} |w - w_h| \to 0$ as $h \to 0$ along with an identical pointwise condition for Z and Z_h. We do not discuss consistency further.

The discretized impulse operator (1.3) is

$$[\mathcal{M}_h^n u^n]_i := \max_{z \in (Z_h)_i^n} \left\{ u^n \left[\!\left[\Gamma_i^n(z) \right]\!\right] + K_i^n(z) \right\}$$

where $\varphi[\![x]\!]$ denotes linear interpolation using the value of φ on grid nodes. $Ł_h^n(w)$ is used to denote a consistent discretization of $Ł^w$ with coefficients frozen at $t = t^n$.

Recall that in (1.2), $\Lambda \subset \partial\Omega$ is a special subset of the boundary at which a Dirichlet-like condition is applied. To distinguish points, we denote by Φ a diagonal matrix satisfying $[\Phi]_{ii} = 0$ whenever x_i is in Λ and $[\Phi]_{ii} = 1$ otherwise.

Since the Dirichlet-like condition (1.2) is imposed at the final time $t = T$, the numerical method proceeds backwards in time (i.e. from t^{n+1} to t^n). More precisely, letting $u_i^N := g_i^N$, the numerical solution u^n at timestep $1 \leq n < N$ produced by each scheme (given the solution at the previous timestep, u^{n+1}) is written as a solution of (2.1) with A and b picked appropriately. Control sets are given by (4.1) and

$$\mathcal{W}_i := W_h,\ \mathcal{Z}_i := \begin{cases} (Z_h)_i^{n_0} & \text{if } (Z_h)_i^{n_0} \neq \emptyset \\ \{\emptyset\} & \text{otherwise} \end{cases},\text{ and } \mathcal{D}_i := \begin{cases} \{0,1\} & \text{if } (Z_h)_i^{n_0} \neq \emptyset \\ \{0\} & \text{otherwise} \end{cases} \tag{5.1}$$

where n_0 is $n + 1$ for the explicit control scheme (see §5.3) and n otherwise. As a technical detail, we take $\mathcal{Z}_i$ to be a nonempty set (we choose $\{\emptyset\}$ arbitrarily) whenever $(Z_h)_i^{n_0}$ is empty to ensure that the product $\mathcal{W}_i \times \mathcal{Z}_i \times \mathcal{D}_i$ of (4.1) is nonempty.

We make the following assumptions:

(A0) $\mathcal{W} \coloneqq \prod_{i=1}^{M} \mathcal{W}_i$ and $\mathcal{Z} \coloneqq \prod_{i=1}^{M} \mathcal{Z}_i$ are finite.
(A1) For all w in $\mathcal{W}$, $-Ł_h^n(w)$ is a WDD Z-matrix with nonnegative diagonals.
(A2) For all z in $\mathcal{Z}$, $B^n(z)$ is a right Markov matrix with $[B^n(z)v]_i = v[\![\Gamma_i^n(z_i)]\!]$.
(A3) $\rho \geq 0$ and $\Upsilon, \epsilon > 0$.

Since (A0) ensures that $\mathcal{P}$ is finite, all schemes below trivially satisfy (H0) and (H1). (A1) and (A2) ensure monotonicity in the viscosity sense (see [22, Section 1.3] for an example of a stable nonmonotone scheme that fails to converge).

5.1. Direct control. In a direct control formulation, either the PDE ($\sup_{w \in W}\{\partial u/\partial t + Ł^w u - \rho u + f^w\}$) or impulse ($\mathcal{M}u - u$) component is active at any grid point. Since these have different units, comparing them in floating point arithmetic requires a scaling factor $\Upsilon > 0$ to ensure good convergence properties [15]. Scaling by Υ and discretizing (1.1) directly yields

$$\max\left(\max_{w \in W_h}\left\{\frac{u_i^{n+1} - u_i^n}{\Delta t} + \left[Ł_h^n(w)\, u^n\right]_i - \rho u_i^n + f_i^n(w)\right\}, \Upsilon\left(\left[\mathcal{M}_h^n u^n\right]_i - u_i^n\right)\right) = 0.$$

Including boundary conditions, this is put in the form of problem (2.1) subject to (4.1) and (4.2) by taking

$$\begin{aligned} L(w) &\coloneqq \Phi\left(Ł_h^n(w)\,\Delta t - \rho I \Delta t\right); & c(w) &\coloneqq \Phi\left(u^{n+1} + f^n(w)\,\Delta t\right) + (1 - \Phi)\, g^n; \\ B(z) &\coloneqq B^n(z); & k(z) &\coloneqq K^n(z). \end{aligned} \tag{5.2}$$

With B and k given above, the operator $\mathcal{M}_h^n$ is equivalent to $\mathbb{B}$ defined in (4.5).

L and B given above satisfy (H3) due to (A1)–(A3). Therefore, (H4) is a sufficient condition for uniqueness of solutions (Theorem 4.7). Similarly, (H2) is a sufficient condition for convergence of the corresponding policy iteration (Theorem 4.2).

5.2. Penalized. A penalized formulation (treated in detail in [27]) imposes a penalty scaled by $1/\epsilon' \gg 0$ whenever $\mathcal{M}u > u$. To arrive at this scheme, we perform the following string of manipulations on (1.1):

$$\boxed{\text{Penalize}} \sup_{w \in W}\left\{\frac{\partial u}{\partial t} + Ł^w u - \rho u + f^w\right\} + \max\left(\mathcal{M}u - u, 0\right)/\epsilon' = 0;$$

$$\boxed{\text{Discretize}} \max_{w \in W_h}\left\{\frac{u_i^{n+1} - u_i^n}{\Delta t} + \left[Ł_h^n(w)\, u^n\right]_i - \rho u_i^n + f_i^n(w)\right\} + \max\left(\left[\mathcal{M}_h^n u^n\right]_i - u_i^n, 0\right)/\epsilon' = 0.$$

Let $\epsilon' \coloneqq \epsilon \Delta t$. This is put in the form (2.1) by taking

$$\begin{aligned} A(P) &\coloneqq I + \Phi\left(\rho I \Delta t - Ł_h^n(w)\,\Delta t\right) + \Psi\left(I - B^n(z)\right)/\epsilon; \\ b(P) &\coloneqq \Phi\left(u^{n+1} + f^n(w)\,\Delta t\right) + (1 - \Phi)\, g^n + \Psi K^n(z)/\epsilon. \end{aligned}$$

Convergence of the corresponding policy iteration is trivial since $A(P)$ is an SDD Z-matrix with positive diagonals (by virtue of (A1)–(A3)), and hence an M-matrix.

5.3. Explicit control. The crux of an explicit control scheme is the use of a *Lagrangian derivative* to remove the D_x coefficient's dependency on the control w. It is assumed that(i) σ is independent of the control and (ii) the drift μ and forcing term f can be split into (sufficiently regular) controlled and uncontrolled components:

$$\mu(y, w) = \hat{\mu}(y) + \hat{\hat{\mu}}(y, w) \text{ and } f(y, w) = \hat{f}(y) + \hat{\hat{f}}(y, w).$$

We now give some intuition behind an explicit control scheme. Consider a generator $\hat{\text{Ł}}$ corresponding to an uncontrolled SDE:

$$\hat{\text{Ł}} u(y) \coloneqq \text{Ł}(w) u(y) - \left\langle \hat{\mu}(y, w), D_x u(y) \right\rangle .$$

Letting $X \coloneqq X(t)$ denote a d-dimensional trajectory satisfying

$$X(t^n) = x_i \text{ and } dX(t) = \hat{\mu}(t, X(t), w)\, dt \text{ on } \left(t^n, t^{n+1}\right]$$

so that $X(t^{n+1}) \approx X(t^n) + \hat{\mu}(t^n, X(t^n), w)\Delta t = x_i + \hat{\mu}_i^n(w)\Delta t$. We define the Lagrangian derivative with respect to X as

$$\frac{Du}{Dt}(t, X(t), w) \coloneqq \frac{\partial}{\partial t}\left[u(t, X(t))\right] = \frac{\partial u}{\partial t}(t, X(t)) + \left\langle \hat{\mu}(t, X(t), w), D_x u(t, X(t)) \right\rangle$$

and perform the following string of manipulations on (1.1):

$$\boxed{\text{Substitute } \frac{Du}{Dt}} \max\left(\sup_{w \in W}\left\{\frac{Du}{Dt}^w + \hat{\text{Ł}} u - \rho u + f^w\right\}, \mathcal{M}u - u\right) = 0;$$

$$\boxed{\text{Discretize}} \max\Bigg(\max_{w \in W_h}\left\{u^{n+1}\left[\!\left[x_i + \hat{\mu}_i^n(w)\Delta t\right]\!\right] + \hat{f}_i^{n+1}(w)\Delta t\right\},$$
$$\left[\mathcal{M}_h^{n+1} u^{n+1}\right]_i\Bigg) - u_i^n + \left(\left[\hat{\text{Ł}}_h^n u^n\right]_i - \rho u_i^n + \hat{f}_i^n\right)\Delta t = 0.$$

Consistency of this scheme (subject to some mild assumptions) can be shown similarly to [11, Lemma 6.6]. In lieu of (A1), we assume:

(A1′) $-\hat{\text{Ł}}_h^n$ is a WDD Z-matrix with nonnegative diagonals.

Let $\mathbf{x}$ denote a vector with components x_i. In the form (2.1):

$$A \coloneqq I + \Phi\left(\rho I \Delta t - \hat{\text{Ł}}_h^n \Delta t\right);$$
$$b(P) \coloneqq \Phi\left(\hat{f}^n \Delta t + (I - \Psi)\left(u^{n+1}\left[\!\left[\mathbf{x} + \hat{\mu}^n(w)\Delta t\right]\!\right] + \hat{f}^{n+1}(w)\Delta t\right)\right)$$
$$+ (1 - \Phi) g^n + \Psi\left(B^{n+1}(z) u^{n+1} + K^{n+1}(z)\right).$$

Since A is independent of P, (2.1) becomes $Av = \max_{P \in \mathcal{P}}\{b(P)\}$; no iterative method is required. A is nonsingular since it is SDD (by virtue of (A1′) and (A3)).

6. Examples.

The remainder of this work focuses on numerical examples.

6.1. Optimal combined control of the exchange rate.

The following is studied in [21, 8]. Consider a government able to influence the foreign exchange (FEX) rate of its currency by:

- choosing the domestic interest rate **(stochastic control)**;
- buying or selling foreign currency **(impulse control)**.

Let $(r_t)_{t \geq 0}$ denote the domestic interest rate process and $\overline{r}$ the foreign interest rate. At any point in time, the government can buy ($z > 0$) or sell ($z < 0$) foreign currency to influence the FEX market. $(X_t)_{t \geq 0}$, the log of the FEX rate, follows

$$\begin{aligned} dX_t &= -a(r_t - \overline{r})\, dt + \sigma d\mathfrak{W}_t \qquad \text{if } \tau_j < t < \tau_{j+1} && \text{(stochastic control)}; \\ X_{\tau_{j+1}} &= X_{\tau_{j+1}-} + z_{\tau_{j+1}} && \text{(impulse control)}. \end{aligned}$$

Parameter		Value	
Discount factor	ρ	2%	per annum
Volatility	σ	30%	per annum
Expiry	T	10	years
Optimal parity	$x^{\star}$	0	
Interest rate differential	W	0-7%	per annum
Interest rate differential effect	a	1/4	
Interest rate differential cost	b	3	
Scaled transaction cost	λ	1	
Fixed transaction cost	C	1/10	

TABLE 6.1.1
Optimal combined control of the exchange rate: parameters

$(\mathfrak{W}_t)_{t\geq 0}$ is a standard Brownian motion under the $\mathbb{P}$ measure. $a > 0$ parameterizes the effect of the interest rate differential on the FEX rate. Let $w_t \coloneqq r_t - \overline{r}$ denote the interest rate differential at time t with $w_{\min} \leq w_t \leq w_{\max}$.

Let $\theta \coloneqq (w, \tau_1, \tau_2, \ldots, z_{\tau_1}, z_{\tau_2}, \ldots)$ where(i) $(w_t)_{t\geq 0}$ is an adapted process, (ii) $\tau_1 \leq \tau_2 \leq \ldots \leq T$ are stopping times, and (iii) z_{τ_k} is a τ_k-measurable random variable taking values from some set $Z(\tau_k, X_{\tau_k})$. Any such θ satisfying these properties is referred to as a *combined control.*

The optimal cost at time t when $X_t = x$ is given by

$$u\left(t,x\right) \coloneqq e^{\rho t} \sup_{\theta} \mathbb{E}^{(t,x)} \left[-\int_t^T e^{-\rho s} \left(p\left(X_s\right) + b w_s^2 \right) ds - \sum_{\tau_j \leq T} e^{-\rho \tau_j} \left(\lambda \left| z_{\tau_j} \right| + C \right) \right]. \tag{6.1}$$

The cost of the distance of the FEX rate to the optimal parity $x^{\star}$ is parameterized by the function p. We take $p(x) \coloneqq (\max(x - x^{\star}, 0))^2$. The constant $b \geq 0$ parameterizes the cost associated with a nonzero interest rate differential. $\lambda \geq 0$ and $C > 0$ parameterize the cost of an impulse. $\rho \geq 0$ is a discount factor.

It is well-known [5] that the dynamic programming equation associated to (6.1) is the HJBQVI on $\Omega \coloneqq \mathbb{R}$ and $\Lambda \coloneqq \emptyset$ given by (1.1)–(1.3) with $g(T,x) \coloneqq 0$ and

$$\begin{aligned} W &\coloneqq [w_{\min}, w_{\max}]; & Z(t,x) &\coloneqq \mathbb{R}; \\ Ł(t,x,w) &\coloneqq -aw\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2}; & \Gamma(t,x,z) &\coloneqq x + z; \\ f(t,x,w) &\coloneqq -p(x) - bw^2; & k(t,x,z) &\coloneqq -\lambda|z| - C. \end{aligned}$$

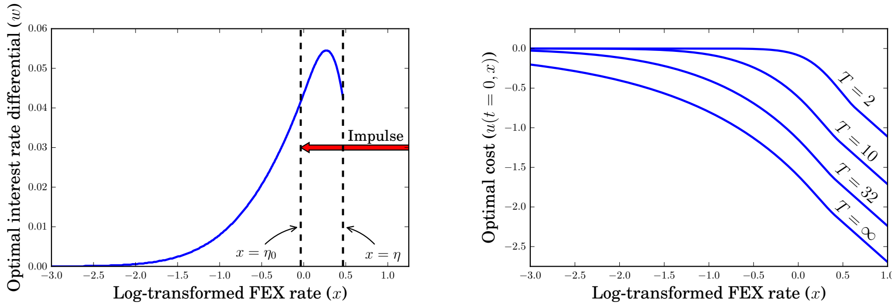
We use the parameters in Table 6.1.1 in our computations.

6.1.1. Convergence of the direct control scheme. Discretization requires that we truncate $[0,T]\times\mathbb{R}$ to $[0,T]\times[x_1,x_M]$ and $Z(t,x) = \mathbb{R}$ to $[x_1,x_M]-x$ so that the exchange rate rate after an impulse, $\Gamma(t,x,z) = x + z$, remains in the computational domain. Let $\Delta z > 0$ divide $x_M - x_1$. A discretization of the truncated $Z(t,x)$ is

$$\left(Z_h^n\right)_i \coloneqq \{0, \Delta z, 2\Delta z, \ldots, x_M - x_1\} + (x_1 - x_i).$$

An artificial Neumann boundary condition $\partial u/\partial x = 0$ is used at x_1 and x_M so that the first and last rows of $Ł_h^n(w)$ are zero. In particular, we assume an upwind three-point stencil [14] so that

$$\left[Ł_h^n\left(w\right)v\right]_i \coloneqq \begin{cases} 0 & \text{if } i = 1 \text{ or } i = M \\ \left(v_{i-1} - v_i\right)\alpha_i^n\left(w\right) + \left(v_{i+1} - v_i\right)\beta_i^n\left(w\right) & \text{otherwise} \end{cases}$$



(a) *Optimal control ($T = 10$)*　(b) *Values for several times-to-expiry*

FIGURE 6.1.1: *Optimal combined control of the exchange rate*

where $\alpha_i^n(w)$ and $\beta_i^n(w)$ are nonnegative.

The direct control problem is given by (2.1) subject to (4.1), (4.2), (5.1), and (5.2). It is easy to verify that $\mathbb{B}^2 x < \mathbb{B}x$ for all x so that (H4) is satisfied (recall $\mathbb{B} = \mathcal{M}_h^n$). By Theorem 4.7, solutions to the problem are unique. However, policy iteration may fail since (H2) is not satisfied. A trivial example violating (H2) is that of a cycle between two nodes $x_i \neq x_j$ (e.g. $\Gamma(t, x_i, x_j - x_i) = x_j$ and $\Gamma(t, x_j, x_i - x_j) = x_i$).

We perform policy iteration on a modified problem with control set $\mathcal{P}' \subsetneq \mathcal{P}$ consisting of all controls $P \coloneqq (w, z, \psi)$ in $\mathcal{P}$ satisfying

$$\psi_1 = 0 \text{ and } z_i < 0 \text{ for all } i > 1 \tag{6.2}$$

so that (H2)$'$ holds. If u^{n+1} is nonincreasing (i.e. $u_{i-1}^{n+1} \geq u_i^{n+1}$), we can use the same arguments as in Example 4.11 to establish that the solution $v = u^n$ of the modified problem solves the original problem (i.e. (4.6) is satisfied) and is nonincreasing. Since the initial solution $u^N = 0$ is nonincreasing, induction establishes the desired result.

REMARK 6.1. *The condition $z_i < 0$ appeals to intuition: the domestic government should never perform an impulse that weakens the domestic currency (i.e. $z_i \geq 0$).*

6.1.2. Optimal control. The optimal cost u for varying expiry times T is shown in Figure 6.1.1b. If the currency is sufficiently weak, the government intervenes in the FEX market. That is, at time t, the impulse occurs only on $[\eta(t), \infty)$ for some $\eta(t)$ (the region $(-\infty, \eta(t))$ on which the impulse is not applied is referred to as the *continuation region*, corresponding to nodes i with $\psi_i = 0$ in the numerical solution). When the FEX rate at time t enters $[\eta(t), \infty)$, the government intervenes to bring it back to $\eta_0(t) < \eta(t)$. This phenomenon is shown in Figure 6.1.1a.

6.1.3. Convergence tests. Convergence tests are shown in Table 6.1.3. Times are normalized to the fastest explicit control solve. The ratio of successive changes in the solution (at a point) is reported.

BiCGSTAB with an ILUT preconditioner is used for the SOLVE routine (line 3 of POLICY-ITERATION) in this and all subsequent sections. In the specific case of the explicit control scheme for the exchange rate problem, a simple tridiagonal solve can be used since the problem is a one-dimensional diffusion.

h	x nodes	w nodes	z nodes	Timesteps
1	32	8	16	16
1/2	64	16	32	32
⋮	⋮	⋮	⋮	⋮

TABLE 6.1.2
Optimal combined control of the exchange rate: numerical grid

h	$u(t=0,x=0)$	Avg. policy its.	Avg. BiCGSTAB its.	Ratio	Norm. time
1	-0.606852556383	3.13	0.74		1.3208e+01
1/2	-0.611872282846	2.88	0.90		6.9925e+01
1/4	-0.613009251583	2.58	0.93	4.42	3.9823e+02
1/8	-0.613175772538	2.49	0.94	6.83	2.7661e+03
1/16	-0.613212923394	2.48	0.95	4.48	2.0892e+04
1/32	-0.613219033234	2.46	0.95	6.08	1.6100e+05

(a) *Direct control*

h	$u(t=0,x=0)$	Avg. policy its.	Avg. BiCGSTAB its.	Ratio	Norm. time
1	-0.607176517477	3.19	0.71		1.3828e+01
1/2	-0.611949599597	2.88	0.76		6.9611e+01
1/4	-0.613029734206	2.55	0.91	4.42	3.9498e+02
1/8	-0.613179661308	2.48	1.28	7.20	2.7647e+03
1/16	-0.613213897805	2.48	1.33	4.38	2.0884e+04
1/32	-0.613219278808	2.46	0.99	6.36	1.6131e+05

(b) *Penalized*

h	$u(t=0,x=0)$	Ratio	Norm. time
1	-0.692778046681		1.0000e+00
1/2	-0.648067155768		6.4890e+00
1/4	-0.628659648118	2.30	4.8988e+01
1/8	-0.620278222694	2.32	3.8620e+02
1/16	-0.616535114461	2.24	3.1665e+03
1/32	-0.614801227389	2.16	2.6442e+04
1/64	-0.613983119004	2.12	2.1737e+05

(c) *Explicit control*

TABLE 6.1.3
Optimal combined control of the exchange rate: convergence tests

POLICY-ITERATION is terminated upon achieving a desired error tolerance:

$$\max_i \left\{ \frac{\left|v_i^k - v_i^{k-1}\right|}{\max\left(\left|v_i^k\right|, \text{scale}\right)} \right\} < \text{tol}.$$

The scale parameter ensures that unrealistic levels of accuracy are not imposed on the solution. We take tol $= 10^{-6}$ and scale $= 1$ for this and all future tests. The initial guess v^0 is taken to be the solution at the previous timestep, u^{n+1}.

Following [15], we take $\epsilon \coloneqq D\Delta t$ and $\Upsilon \coloneqq 1/\epsilon$ with $D = 10^{-2}$.

The direct control and penalized schemes converge superlinearly. We speculate that this occurs since $x \mapsto u(t,x)$ is linear to the right of $x = \eta_0(t)$, and hence no error is made in approximating the term $D_x u$ and $D_x^2 u$ there. At t^n, the solution

u^{n+1} of the explicit control scheme is linear to the right of $\eta_0(t^{n+1})$. Hence, error is introduced due to the approximation of $\eta_0(t^n)$ by $\eta_0(t^{n+1})$. This suggests that the direct control and penalized schemes may outperform the explicit control scheme for problems with simple continuation regions and linear transactional costs.

Unsurprisingly, the direct control and penalized schemes are near-identical in performance and accuracy since the scaling and penalty factors are chosen identically (i.e. $\Upsilon = 1/\epsilon$). We mention that the choice of $\Upsilon \coloneqq 1$ (i.e. no scaling) yields poor performance in the direct control setting (see [15] for an explanation).

Note that the average number of BiCGSTAB iterations per call to Solve can be less than one, suggesting that sometimes, no BiCGSTAB iterations are required on line 3 of Policy-Iteration. This occurs when the initial residual, $b(P^\ell) - A(P^\ell)v^{\ell-1}$, is small enough in magnitude (i.e. at the last policy iteration before convergence).

6.2. Optimal consumption and portfolio with both fixed and proportional transaction costs. The following is studied in [10]. Consider an investor that, at any point in time, has two investment opportunities: a stock and a bank account. Let $(S_t)_{t\geq 0}$ and $(B_t)_{t\geq 0}$ denote the amount of money invested in these two, respectively. The investor is able to

- consume continuously **(stochastic control)**;
- transfer money from the bank to the stock (or vice versa) subject to a transaction cost **(impulse control)**.

Denote by $(w_t)_{t\geq 0}$ the consumption rate with $0 \leq w_t \leq w_{\max}$. At any point in time, the investor can move money to $(z > 0)$ or from $(z < 0)$ the stock incurring a transaction cost of $\lambda |z| + c$ where $C > 0$ and $0 < \lambda < 1$. This is captured by

$$
\begin{aligned}
dS_t &= \mu S_t dt + \xi S_t d\mathfrak{W}_t && \text{if } \tau_j < t < \tau_{j+1} \quad \text{(stochastic control);}\\
dB_t &= (rB_t - w_t)\, dt && \text{if } \tau_j < t < \tau_{j+1} \quad \text{(stochastic control);}\\
S_{\tau_{j+1}} &= S_{\tau_{j+1}-} + z_{\tau_{j+1}} && \text{(impulse control);}\\
B_{\tau_{j+1}} &= B_{\tau_{j+1}-} - z_{\tau_{j+1}} - \lambda \left| z_{\tau_{j+1}} \right| - C && \text{(impulse control).}
\end{aligned}
$$

A combined control $\theta \coloneqq (w, \tau_1, \tau_2, \ldots, z_{\tau_1}, z_{\tau_2}, \ldots)$ is *admissible* if at all times, the stock holdings and bank account are a.s. nonnegative.

The investor's maximal expected utility at time t with amount $S_t = s$ in the stock and $B_t = b$ in the bank account is given by

$$
u(t, s, b) \coloneqq e^{\rho t} \sup_{\theta} \mathbb{E}^{(t,s,b)} \left[\int_t^T e^{-\rho t'} \frac{w_{t'}^{\gamma}}{\gamma} dt' + e^{-\rho T} \frac{\max\left(B_T + (1-\lambda) S_T - C, 0\right)^{\gamma}}{\gamma} \right]
$$

where $0 \leq 1-\gamma < 1$ is the investor's relative risk-aversion and $\rho \geq 0$ is the rate of time preference. The utility received at the expiry corresponds to liquidating the asset and consuming everything instantaneously. The supremum is over admissible controls.

The associated HJBQVI on $\Omega \coloneqq (0,\infty)^2$ and $\Lambda \coloneqq \emptyset$ is given by (1.1)–(1.3) with $g(T, x) \coloneqq \max(b + (1-\lambda)s - C, 0)^{\gamma}/\gamma$ and

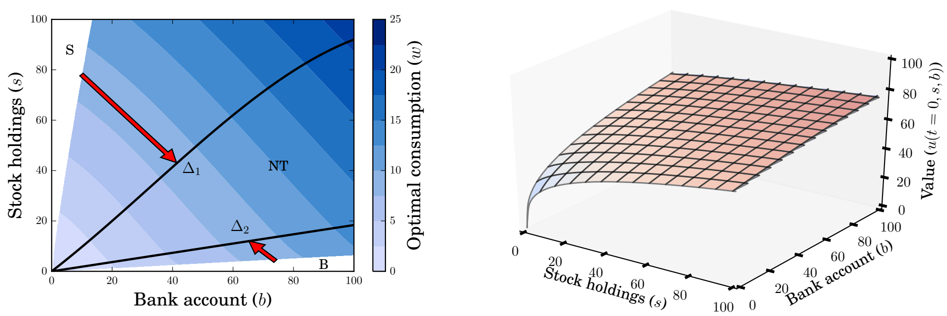
$$
W \coloneqq [0, w_{\max}]; \qquad Z(t,x) \coloneqq \left[-s, \frac{b - C}{1+\lambda}\right];
$$

$$
Ł^w \coloneqq \frac{1}{2}\xi^2 s^2 \frac{\partial^2}{\partial s^2} + \mu s \frac{\partial}{\partial s} + \begin{cases} (rb - w)\, \frac{\partial}{\partial b} & \text{if } b > 0;\\ 0 & \text{otherwise;} \end{cases} \qquad \Gamma(t, x, z) \coloneqq x + \big(z, -z - \lambda |z| - C\big);
$$

$$
f^w \coloneqq \begin{cases} w^{\gamma}/\gamma & \text{if } b > 0;\\ 0 & \text{otherwise;} \end{cases} \qquad k(t, x, z) \coloneqq 0.
$$

Parameter		Value	
Discount factor	ρ	10%	per annum
Interest rate	r	7%	per annum
Drift	μ	11%	per annum
Volatility	ξ	30%	per annum
Expiry	T	40	years
Relative risk aversion	$1-\gamma$	0.7	
Scaled transaction cost	λ	0.1	
Fixed transaction cost	C	0.05	
Maximum withdrawal rate	$w_{\max}$	100	
Initial stock value	s_0	\$45.20	
Initial bank account value	b_0	\$45.20	

Table 6.2.1
Optimal consumption: parameters from [10]



(a) *Optimal control* (b) *Value*

Figure 6.2.1: *Optimal consumption (compare with [10, Figures 1 and 2])*

$Z(t,x)$ should be understood subject to the convention $[q_1,q_2]=\emptyset$ if $q_1>q_2$.

We use the parameters in Table 6.2.1 in our computations.

6.2.1. Convergence of the direct control scheme. As in §6.1.1, the domain $[0,T]\times[0,\infty)^2$ and $Z(t,x)$ are truncated so that state after an impulse $\Gamma(t,x_i,z_i)$ remains in the truncated domain. We use the notation $x_i=(s_i,b_i)$. The direct control problem is given by (2.1) subject to (4.1), (4.2), (5.1), and (5.2).

Suppose there exists a grid node x_{i^1} and $P\coloneqq(w,z,\psi)$ such that $\psi_{i^1}=1$ and that there exists no path in $B(z)$ from i^1 to some j with $\psi_j=0$. Since $C>0$, there exists a path $i^1\to i^2\to\cdots$ of infinite length such that $s_{i^1}+b_{i^1}>s_{i^2}+b_{i^2}>\cdots$ and $\psi_{i^q}=1$ for all q. Due to the finitude of the grid, $x_{i^q}=x_{i^\ell}$ (and hence $s_{i^q}+b_{i^q}=s_{i^\ell}+b_{i^\ell}$) for some $q<\ell$, a contradiction. It follows that no such x_{i^1} exists: (H2) is satisfied.

6.2.2. Optimal control. As in [10], three regions are observed in an optimal control: the *buy* (B), *sell* (S), and *continuation/no transaction* (NT) regions. In the B and S regions, the controller intervenes by jumping back to the closest of the two lines marked Δ_1 and Δ_2. In NT, the controller consumes continuously.

6.2.3. Convergence tests. Convergence tests are shown in Table 6.2.3. We mention that artificial Neumann boundary conditions $\partial^q u/\partial s^q=0$ and $\partial u/\partial b=0$ are used at the truncated boundaries $s=s_{\max}$ and $b=b_{\max}$. The results for the

h	s nodes	b nodes	w nodes	z nodes	Timesteps
1	20	20	15	15	32
1/2	40	40	30	30	64
⋮	⋮	⋮	⋮	⋮	⋮

TABLE 6.2.2
Optimal consumption: numerical grid

h	$u(t = 0, s_0, b_0)$	Avg. policy its.	Avg. BiCGSTAB its.	Ratio	Norm. time
1	56.0621229141	7.63	1.28		1.6345e+01
1/2	58.7392240395	8.80	1.90		2.9329e+02
1/4	59.4201246475	10.4	2.28	3.93	5.6638e+03
1/8	59.6584129364	11.8	3.28	2.86	1.0282e+05
1/16	59.7547798553	13.3	4.45	2.47	1.8548e+06
1/32	59.7972061330	14.2	6.54	2.27	3.0458e+07

(a) *Direct control*

h	$u(t = 0, s_0, b_0)$	Avg. policy its.	Avg. BiCGSTAB its.	Ratio	Norm. time
1	56.0584963190	4.09	1.43		1.0303e+01
1/2	58.7390408653	3.95	1.77		1.4742e+02
1/4	59.4200754123	3.40	2.35	3.94	1.9949e+03
1/8	59.6583990235	3.04	3.69	2.86	2.6855e+04
1/16	59.7547779953	2.80	5.50	2.47	4.0166e+05
1/32	59.7972150004	2.58	5.98	2.27	5.8572e+06

(b) *Penalized*

h	$u(t = 0, s_0, b_0)$	Avg. BiCGSTAB its.	Ratio	Norm. time
1	55.6216321734	1.00		1.0000e+00
1/2	58.7820641022	2.00		1.5465e+01
1/4	59.4045764001	3.00	5.08	2.5950e+02
1/8	59.5693702945	4.00	3.78	4.0526e+03
1/16	59.6511861506	6.00	2.01	6.6804e+04
1/32	59.7053148416	8.00	1.51	1.1079e+06
1/64	59.7483254658	10.8	1.26	1.8531e+07

(c) *Explicit control*
TABLE 6.2.3
Optimal consumption: convergence tests

direct control and penalized schemes are near-identical, though the former requires significantly more policy iterations per timestep. The rate of convergence for the explicit control scheme becomes sublinear for higher levels of refinement.

6.3. Guaranteed minimum withdrawal benefit (GMWB) in variable annuities. Guaranteed minimum withdrawal benefits (GMWB) in variable annuities provide investors with the tax-deferred nature of variable annuities along with a guaranteed minimum payment. GMWB pricing has been previously considered as a singular control problem in [20, 13] and as an impulse control problem in [11]. Optimal controls for GMWBs with annual withdrawals is considered in [1].

A GMWB is composed of investment and guarantee accounts, $(S_t)_{t \geq 0}$ and $(A_t)_{t \geq 0}$, respectively. It is bootstrapped via a lump sum payment s_0 to an insurer, placed in the (risky) investment account (i.e. $S_0 = s_0$). A GMWB promises to pay back at least the lump sum s_0, assuming that the holder of the contract does not withdraw above a certain rate. This is captured by setting $A_0 = s_0$ and reducing both investment and guarantee accounts on a dollar-for-dollar basis upon withdrawals. The holder can continue to withdraw as long as the guarantee account remains positive. In particular, at any point in time until the expiry of the contract T, the holder may:

- withdraw continuously at a rate of $G \geq 0$ per annum regardless of the performance of the investment **(stochastic control)**;
- withdraw a finite amount instantaneously reduced by the lapsation rate $0 \leq \kappa \leq 1$ **(impulse control)**.

The holder gets the larger of the investment account and a full withdrawal at expiry.

The guarantee account can be withdrawn from continuously or instantaneously:

$$\begin{aligned} dA_t &= -w_t dt && \text{if } \tau_j < t < \tau_{j+1} && \text{(stochastic control);} \\ A_{\tau_{j+1}} &= A_{\tau_{j+1}-} - z_{\tau_{j+1}} && && \text{(impulse control).} \end{aligned}$$

Let $\rho \geq 0$ denote the risk-free rate. Consider an index $(Y_t)_{t\geq 0}$ following

$$dY_t = \rho Y_t dt + \xi Y_t d\mathfrak{W}_t$$

under the risk-neutral measure (that which renders the discounted index into a martingale). The investment account tracks the index and is adjusted by withdrawals from the guarantee account:

$$dS_t = (\rho - \eta) S_t dt + \xi S_t d\mathfrak{W}_t + dA_t \text{if } \tau_j < t < \tau_{j+1}.$$

$0 \leq \eta \leq \rho$ is the proportional rate deducted from the investment account and serves as a premium for the guarantee. A combined control $\theta \coloneqq (w, \tau_1, \tau_2, \ldots, z_1, z_2, \ldots)$ is admissible if at all times, the guarantee account is a.s. nonnegative.

The insurer's worst-case cost of hedging a GMWB at time t with amount $S_t = s$ in the risky account and amount $A_t = a$ is

$$\begin{aligned} u(t,s,a) \coloneqq e^{\rho t} \sup_{\theta} \mathbb{E}^{(t,s,a)} \Bigg[\int_t^T e^{-\rho t'} w_{t'} dt' + e^{-\rho T} \max\left(S_T, (1-\kappa) A_T - C\right) \\ + \sum_{\tau_j \leq T} e^{-\rho \tau_j} \left((1-\kappa) z_{\tau_j} - C \right) \Bigg] \end{aligned}$$

where $C > 0$ is a fixed transaction cost. The terminal payoff corresponds to the maximum of the investment account or withdrawing the entirety of the guarantee account at the lapsation rate. The supremum is over admissible controls.

Let $x \coloneqq (s, a)$ and $\zeta \coloneqq \rho - \eta$. The associated HJBQVI on $\Omega \coloneqq (0,\infty)^2$ and $\Lambda \coloneqq \emptyset$ is given by (1.1)–(1.3) with $g(T,x) \coloneqq \max(s, (1-\kappa)a - C)$ and

$$W \coloneqq [0, G]; \qquad Z(t,x) \coloneqq [0, a];$$

$$Ł^w \coloneqq \frac{\xi^2 s^2}{2} \frac{\partial^2}{\partial s^2} + \zeta s \frac{\partial}{\partial s} + \begin{cases} -w\left(\frac{\partial}{\partial a} + \frac{\partial}{\partial s}\right) & \text{if } s, a > 0; \\ -w\frac{\partial}{\partial a} & \text{if } a > 0; \\ 0 & \text{otherwise;} \end{cases} \qquad \Gamma(t,x,z) \coloneqq \max(x - z, 0);$$

$$f^w \coloneqq \begin{cases} w & \text{if } a > 0; \\ 0 & \text{otherwise;} \end{cases} \qquad k(t,x,z) \coloneqq (1-\kappa) z - C.$$

Throughout, we consider the parameters in Table 6.3.1.

Parameter		Value	
Risk-free rate	ρ	5%	per annum
Premium	η	0%	per annum
Volatility	ξ	30%	per annum
Expiry	T	10	years
Withdrawal rate	G	\$10	per annum
Lapsation rate	κ	10%	
Fixed transaction cost	c	0	
Initial lump sum payment	s_0	\$100	

TABLE 6.3.1
GMWB: parameters from [11]

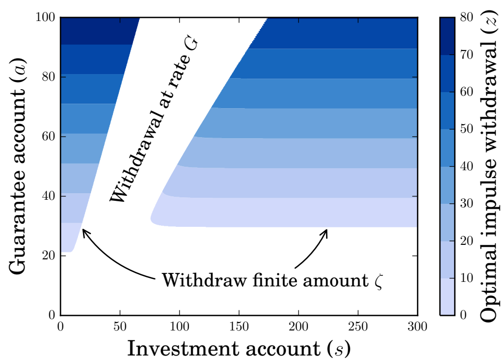


FIGURE 6.3.1: *GMWB: optimal control at initial time with* $\eta = 0.03126$ *from [11]*

6.3.1. Convergence of the direct control scheme. We use the notation $x_i = (s_i, a_i)$ and assume the origin $(0, 0)$ is part of the numerical grid. The direct control problem is given by (2.1) subject to (4.1), (4.2), (5.1), and (5.2).

Suppose (H4) is not satisfied so that for some solution v, there exists i such that $v_i = [\mathbb{B}v]_i = [\mathbb{B}^2 v]_i = \cdots$. Since $C > 0$, it follows that $v_i = -\infty$, a contradiction. Hence, (H4) holds.

We perform policy iteration on a modified problem with control set $\mathcal{P}'$ consisting of all controls $P \coloneqq (w, z, \psi)$ in $\mathcal{P}$ satisfying

$$\psi_i = 0 \text{ whenever } a_i = 0 \text{ and } z_i \neq 0 \text{ whenever } a_i \neq 0.$$

As in Example 4.4, (H2)$'$ follows from the undirectionality of z_i. (4.6) is established by noting that $z_i = 0$ incurs an infinite cost (and is therefore suboptimal). Convergence then follows from an application of Theorem 4.10

REMARK 6.2. *The condition* $z_i \neq 0$ *appeals to intuition: the holder should never pay* $C > 0$ *for a withdrawal of zero dollars.*

6.3.2. Optimal control. Figure 6.3.1 shows an optimal control for a GMWB, corresponding to a worst-case cost of hedging from the perspective of the insurer. We refer to [11] for an explanation of the three distinct withdrawal regions.

6.3.3. Convergence tests. Convergence tests are shown in Table 6.3.3. Since $w \mapsto Ł(t, x, w)$ is linear, we take $W_h = \{0, G\}$ independent of h. A linear boundary condition $(s \mapsto u(t, s, a)) \in \Theta(s)$ is used at the truncated boundary $s = s_{\max}$ (no

h	w nodes	a nodes	z nodes	Timesteps
1	64	50	2	32
1/2	128	100	4	64
⋮	⋮	⋮	⋮	⋮

TABLE 6.3.2
GMWB: numerical grid

h	$u(t=0,s_0,s_0)$	Avg. policy its.	Avg. BiCGSTAB its.	Ratio	Norm. time
1	107.683417498	3.47	1.47		1.7066e+01
1/2	107.706787394	4.25	1.64		2.0338e+02
1/4	107.718780318	4.34	1.85	1.95	2.6029e+03
1/8	107.725782831	4.43	2.22	1.71	3.4646e+04
1/16	107.729641357	4.31	2.71	1.81	4.7536e+05
1/32	107.731755456	4.15	3.40	1.83	7.5549e+06

(a) *Direct control*

h	$u(t=0,s_0,s_0)$	Avg. policy its.	Avg. BiCGSTAB its.	Ratio	Norm. time
1	107.682425551	3.47	1.58		1.8034e+01
1/2	107.706388904	4.08	1.65		2.0576e+02
1/4	107.718700668	3.95	1.76	1.95	2.4530e+03
1/8	107.725763667	3.98	1.97	1.74	3.2215e+04
1/16	107.729637421	3.71	2.39	1.82	4.3381e+05
1/32	107.731754559	3.32	3.01	1.83	6.6197e+06

(b) *Penalized*

h	$u(t=0,s_0,s_0)$	Avg. BiCGSTAB its.	Ratio	Norm. time
1	107.423506170	1.00		1.0000e+00
1/2	107.684431768	1.00		1.0195e+01
1/4	107.708405901	1.00	10.9	1.3915e+02
1/8	107.722569027	1.00	1.70	2.0349e+03
1/16	107.730146084	1.00	1.87	3.3067e+04
1/32	107.732241020	1.98	3.62	6.1040e+05
1/64	107.733372937	2.90	1.85	1.1311e+07

(c) *Explicit control*
TABLE 6.3.3
GMWB: convergence tests

boundary condition is needed at $a = a_{\max}$ since the characteristics are outgoing in the a direction). For details, see [11]. The direct control and penalized scheme produce near-identical results and exhibit similar execution times.

7. Concluding remarks. This work establishes the well-posedness of (1.4) and gives sufficient conditions for convergence of the corresponding policy iteration. (1.4) has applications to the numerical solutions of HJBQVIs (§5–6) and infinite-horizon MDPs with vanishing discount factor (Corollary 4.3).

An explicit control scheme for the HJBQVI (1.1)–(1.3) is both easy to implement and requires only one linear solve per timestep. However, it cannot be used if the diffusion or jump arrival rate of the underlying stochastic process are control-dependent.

The direct control and penalized schemes do not suffer these limitations. Numerical evidence suggests that both schemes perform similarly. However, policy iteration applied to the direct control scheme can fail (Example 4.9) unless additional care is taken to remove certain suboptimal controls. The removal of these controls is ad hoc (i.e. problem dependent). Due to this, we recommend (perhaps controversially) abstaining from a direct control scheme.

Appendix A. General well-posedness of the Bellman problem (2.1)**.**

By modifying policy iteration, it is possible to arrive at a version of Proposition 2.2 independent of (H1.ii). We can interpret this algorithm as taking into account the error from approximating the supremum in Policy-Iteration. The algorithm, closely related to [6, Algorithm Ho-4], is given below:

Sequential-Policy-Iteration($\mathcal{P}, A(\cdot), b(\cdot), v^0$)

Pick a positive sequence $(\epsilon^\ell)_{\ell\geq 0}$ in $\mathbb{R}^M$ such that $\sum_{\ell\geq 0}\epsilon^\ell$ converges
for $\ell = 1, 2, \ldots$
  Pick P^ℓ such that $-A(P^\ell)v^{\ell-1} + b(P^\ell) + \epsilon^\ell \geq \sup_{P\in\mathcal{P}}\{-A(P)v^{\ell-1} + b(P)\}$
  Solve $A(P^\ell)v^\ell = b(P^\ell)$ for v^ℓ

The following appears in [6]:

Lemma A.1. *A bounded sequence $(v^\ell)_{\ell\geq 0}$ in $\mathbb{R}^M$ converges whenever there exists a sequence $(\epsilon^\ell)_{\ell\geq 0}$ in $\mathbb{R}^M$ such that $\sum_{\ell\geq 0}\epsilon^\ell$ converges and $v^\ell - v^{\ell-1} \geq -\epsilon^\ell$ for $\ell \geq 1$.*

We require the following lemma, whose proof is trivial and thus omitted:

Lemma A.2. *Let X be a set, Y a normed linear space, $T\colon X\times Y\to\mathbb{R}$, and $Q\colon X\to\mathbb{R}$. Suppose that for each x in X, $T_x\colon Y\to\mathbb{R}$ defined by $T_x(y) \coloneqq T(x,y)$ is linear and that T_x has operator norm bounded uniformly with respect to x. The map $y\mapsto\sup_{x\in X}\{T(x,y)+Q(x)\}$ is uniformly continuous.*

Theorem A.3 (Convergence of sequential policy iteration). *Suppose* (H0), (H1.i)*, and that $A(P)$ is a monotone matrix for all P in $\mathcal{P}$. $(v^\ell)_{\ell=1}^\infty$ defined by* Sequential-Policy-Iteration *converges to the unique solution v of* (2.1)*.*

Proof. First, note that

$$A(P^\ell)\left(v^\ell - v^{\ell-1}\right) = -A(P^\ell)v^{\ell-1} + b(P^\ell) \geq \sup\left\{-A\left(P\right)v^{\ell-1} + b\left(P\right)\right\} - \epsilon^\ell. \quad \text{(A.1)}$$

For $\ell > 1$,

$$\sup\left\{-A\left(P\right)v^{\ell-1} + b\left(P\right)\right\} \geq -A(P^{\ell-1})v^{\ell-1} + b(P^{\ell-1}) = 0.$$

Combining this with (A.1), it follows from the boundedness of $P\mapsto A(P)^{-1}$ in (H0) that there exists some α in $\mathbb{R}$ such that for all $\ell > 1$

$$v^\ell - v^{\ell-1} \geq [A(P^\ell)]^{-1}\epsilon^\ell \geq -\alpha\epsilon^\ell.$$

By Lemma A.1, $v^\ell \to v$ for some v in $\mathbb{R}^M$. Taking limits on both sides of (A.1) and applying Lemma A.2 yields

$$0 = \lim_{\ell\to\infty}\left(\sup\left\{-A\left(P\right)v^{\ell-1} + b\left(P\right)\right\}\right) = \sup\left\{-A\left(P\right)v + b\left(P\right)\right\}.$$

Hence, v is a solution to (2.1). Uniqueness is proven similarly to Theorem 4.7. ∎

REFERENCES

[1] P. Azimzadeh and P. A. Forsyth. The existence of optimal bang-bang controls for GMxB contracts. *SIAM J. Financial Math.*, 6(1):117–139, 2015.

[2] J. Babbin, P. A. Forsyth, and G. Labahn. A comparison of iterated optimal stopping and local policy iteration for American options under regime switching. *J. Sci. Comput.*, 58(2):409–430, 2014.

[3] G. Barles and P. E. Souganidis. Convergence of approximation schemes for fully nonlinear second order equations. *Asymptot. Anal.*, 4(3):271–283, 1991.

[4] E. Bayraktar and H. Xing. Pricing Asian options for jump diffusion. *Math. Finance*, 21(1):117–143, 2011.

[5] A. Bensoussan and J. L. Lions. *Impulse control and quasi-variational inequalities.* Gauthier-Villars, Paris, 1984.

[6] O. Bokanowski, S. Maroso, and H. Zidani. Some convergence results for Howard's algorithm. *SIAM J. Numer. Anal.*, 47(4):3001–3026, 2009.

[7] J. H. Bramble and B. E. Hubbard. On a finite difference analogue of an elliptic boundary problem which is neither diagonally dominant nor of non-negative type. *J. Math. Phys.*, 43(2):117, 1964.

[8] A. Cadenillas and F. Zapatero. Optimal central bank intervention in the foreign exchange market. *J. Econom. Theory*, 87(1):218–242, 1999.

[9] J. P. Chancelier, M. Messaoud, and A. Sulem. A policy iteration algorithm for fixed point problems with nonexpansive operators. *Math. Methods Oper. Res.*, 65(2):239–259, 2007.

[10] J. P. Chancelier, B. Øksendal, and A. Sulem. Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control. *Proc. Steklov Inst. Math.*, 237(0):149–172, 2002.

[11] Z. Chen and P. A. Forsyth. A numerical scheme for the impulse control formulation for pricing variable annuities with a guaranteed minimum withdrawal benefit (GMWB). *Numer. Math.*, 109(4):535–569, 2008.

[12] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential lévy models. *SIAM J. Numer. Anal.*, 43(4):1596–1626, 2005.

[13] M. Dai, Y. K. Kwok, and J. Zong. Guaranteed minimum withdrawal benefit in variable annuities. *Math. Finance*, 18(4):595–611, 2008.

[14] P. A. Forsyth and G. Labahn. Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *J. Comput. Finance*, 11(2):1, 2007.

[15] Y. Huang, P. A. Forsyth, and G. Labahn. Inexact arithmetic considerations for direct control and penalty methods: American options under jump diffusion. *Appl. Numer. Math.*, 72:33–51, 2013.

[16] K. Ishii. Viscosity solutions of nonlinear second order elliptic PDEs associated with impulse control problems. *Funkcial. Ekvac.*, 36(1):123–141, 1993.

[17] I. Kharroubi, J. Ma, H. Pham, J. Zhang, et al. Backward SDEs with constrained jumps and quasi-variational inequalities. *Ann. Appl. Probab.*, 38(2):794–840, 2010.

[18] H. J. Kushner and P. G. Dupuis. *Numerical methods for stochastic control problems in continuous time.* Springer-Verlag, New York, 1992.

[19] H. Le and C. Wang. A finite time horizon optimal stopping problem with regime switching. *SIAM J. Control Optim.*, 48(8):5193–5213, 2010.

[20] M. A. Milevsky and T. S. Salisbury. Financial valuation of guaranteed minimum withdrawal benefits. *Insurance Math. Econom.*, 38(1):21–38, 2006.

[21] G. Mundaca and B. Øksendal. Optimal stochastic intervention control with application to the exchange rate. *J. Math. Econom.*, 29(2):225–243, 1998.

[22] A. M. Oberman. Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems. *SIAM J. Numer. Anal.*, 44(2):879–895, 2006.

[23] B. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions*, volume 498. Springer, 2005.

[24] R. J. Plemmons. M-matrix characterizations.I–Nonsingular M-matrices. *Linear Algebra Appl.*, 18(2):175–188, 1977.

[25] R. C. Seydel. Existence and uniqueness of viscosity solutions for QVI associated with impulse control of jump-diffusions. *Stochastic Process. Appl.*, 119(10):3719–3748, 2009.

[26] P. N. Shivakumar and K. H. Chew. A sufficient condition for nonvanishing of determinants. *Proc. Amer. Math. Soc.*, pages 63–66, 1974.

[27] J. H. Witte and C. Reisinger. Penalty methods for the solution of discrete HJB equations-continuous control and obstacle problems. *SIAM J. Numer. Anal.*, 50(2):595–625, 2012.