Preprint Cermics

A policy iteration algorithm for fixed point problems with nonexpansive operators

Jean-Philippe Chancelier[1], **Marouen Messaoud**[2], **Agnès Sulem**[2]

[1] Cermics, École Nationale des Ponts et Chaussées, 6 et 8 avenue Blaise Pascal, 77455, Marne la Vallée, Cedex, France, e-mail: `jpc@cermics.enpc.fr`

[2] Inria Domaine de Voluceau,BP 105 Rocquencourt, 78153 Le Chesnay Cedex France, e-mail: `Marouen.Messaoud@inria.fr`, e-mail: `Agnes.Sulem@inria.fr`

Received: date / Revised version: date

Abstract The aim of this paper is to solve the fixed point problems :

$$v = \mathcal{O}v, \quad \text{with} \quad \mathcal{O}v(x) \stackrel{\text{def}}{=} \max(Lv(x), Bv(x)), \quad x \in \mathcal{E} \ , \tag{1}$$

where $\mathcal{E}$ is a finite set, L is contractive and B is a nonexpansive operator and

$$v = \mathcal{O}v, \quad \text{with} \quad \mathcal{O}v(x) \stackrel{\text{def}}{=} \max(\sup_w L^w v(x), \sup_z B^z v(x)), \quad x \in \mathcal{E} \ , \tag{2}$$

where $\mathcal{W}$ and $\mathcal{Z}$ are general control sets, the operators L^w are contractive and operators B^z are nonexpansive. For these two problems, we give conditions which imply existence and uniqueness of a solution and provide a policy iteration algorithm which converges to the solution. The proofs are slightly different for the two problems since the set of controls is finite for (1) while it is not necessary the case for problem (2). Equation (2) typically arises in numerical analysis of quasi variational inequalities and variational inequalities associated to impulse or singular stochastic control.

Key words Howard Algorithm – Policy iteration – Impulse control – Quasi-variational inequalities – Fixed point problems – Optimal control of Markov Chains – Nonexpansive operators

1 Introduction

The case of Bellman equations associated to optimal control of Markov chains on an infinite horizon with discount factor $\lambda > 0$ has been studied for a long time by many authors (see e.g the monographs by [1] and [8] and the references herein). Typically these equations are of the form $v = \sup_{w \in \mathcal{W}} \{ \frac{1}{1+\lambda} M^w v + c^w \}$ where M^w is the transition matrix of the Markov chain, c^w is the running utility and w is the control variable with values in some control set $\mathcal{W}$. We know that the iteration policy algorithm converges to the solution of the Bellman equation since the operator $\frac{1}{1+\lambda} M^w + c^w$ is contractive and satisfies a discrete maximum principle.

The problem addressed in this paper concerns more general fixed point problems on a finite state space. Typically the operator we will consider is the maximum of a contractive operator and a nonexpansive one which satisfy some appropriate properties. We refer to [1] for the study of shortest path problems which also lead to some fixed point problems with nonexpansive operators but in a rather different context or to [4, p.39] where reflecting boundaries lead to nonexpansive operators on the boundary. This last problem appears to be a special case of ours.

The paper is organized as follows: In Section 2, we study the problem:

$$v(x) = \max(Lv(x), Bv(x)), \quad x \in \mathcal{E} \ , \tag{3}$$

where $\mathcal{E}$ is a finite set, L is contractive and B is a nonexpansive operator. We prove the convergence of an iteration policy algorithm to the solution of (3) provided that the operators L and M fulfill some conditions. In section 3, we turn to

$$v(x) = \max(\sup_w L^w v(x), \sup_z B^z v(x)), \quad x \in \mathcal{E} \ , \tag{4}$$

where $\mathcal{W}$ and $\mathcal{Z}$ are general control sets, the operators L^w are contractive and operators B^z are nonexpansive. Now the set of controls is infinite and the proof of convergence of the policy iteration has to be adapted. These problems are illustrated by examples in optimal control of Markov chains.

Finally Section 4 concerns an application of the results of Section 3 to the numerical analysis of quasi variational inequalities (QVIs) associated to combined impulse/stochastic optimal controls. Indeed, stable and consistent finite difference approximations of these QVIs lead to fixed point problems of type (4) where L^w comes from the approximation of the underlying controlled diffusion and B^z comes from the approximation of the intervention operator. We refer to [5] for an general exposition on impulse control problems and to [7]-[3] for the study of impulse control problems associated to portfolio optimization with fixed transaction costs. Results of section 2 can be applied to the numerical analysis of variational inequalities with gradient constraints associated to singular optimal controls [2].

2 A fixed point problem with a nonexpansive operator and a finite set of controls

2.1 Formulation of the problem and hypotheses

We consider the fixed point problem :

$$v(x) = \max(Lv(x), Bv(x)), \quad x \in \mathcal{E} \tag{5}$$

where $\mathcal{E}$ is a finite set of cardinal n, v is a function defined on $\mathcal{E}$ taking values in $\mathbb{R}$, L is contractive and B is nonexpansive.

If A is a subset of $\mathcal{E}$, we denote $V_A \stackrel{\text{def}}{=} \{v : A \mapsto \mathbb{R}\}$. When $A = \mathcal{E}$ we simply write V for $V_{\mathcal{E}}$. We identify functions belonging to V_A with vectors of dimension equal to $\mathrm{card}A$.

If T is a subset of $\mathcal{E}$, we denote by $\mathcal{O}_T$ the following operator :

$$\mathcal{O}_T v(x) \stackrel{\text{def}}{=} \begin{cases} Lv(x) & \text{if } x \in C \quad \text{where} \quad C \stackrel{\text{def}}{=} \mathcal{E}\backslash T, \\ Bv(x) & \text{if } x \in T. \end{cases} \tag{6}$$

Problem (5) is equivalent to

$$v(x) = \max_{T \in \mathcal{P}(\mathcal{E})} \mathcal{O}_T v(x). \tag{7}$$

We restrict now the set of admissible controls. Define

$$T_{\text{ad}} \stackrel{\text{def}}{=} \mathcal{P}(\mathcal{E})\backslash\mathcal{E}. \tag{8}$$

We thus assume that the choice $T = \mathcal{E}$ is not admissible. Note that T_{ad} is a finite set since $\mathcal{E}$ is a finite set. Set:

$$\mathcal{O}v(x) \stackrel{\text{def}}{=} \max_{T \in T_{\text{ad}}} \mathcal{O}_T v(x). \tag{9}$$

We look for $(v^\star, T^\star) \in V \times T_{\text{ad}}$, a solution of :

$$v^\star = \max_{T \in T_{\text{ad}}} \mathcal{O}_T v^\star = \mathcal{O}_{T^\star} v^\star. \tag{10}$$

For $T \in T_{\text{ad}}$ and $v \in \mathbb{R}^n$ we denote by (v_C, v_T) the decomposition of v on the partition C, T of $\mathcal{E}$. We make the following assumptions:

H$_1$ For each $T \in T_{\text{ad}}$, there exist two operators $\overline{L} : V_C \mapsto V_C$, $\overline{B} : V_C \mapsto V_T$ and a function $\overline{k} : T \mapsto T$ such that :

$$\mathcal{O}_T v = v \Leftrightarrow \overline{L} v_C = v_C \text{ and } v_T = \overline{B} v_C + \overline{k} \tag{11}$$

$$\mathcal{O}_T v^1 - \mathcal{O}_T v^2 \leq v^1 - v^2 \Rightarrow \overline{L} v_C^1 - \overline{L} v_C^2 \leq v_C^1 - v_C^2 \tag{12}$$

$$\text{and } \overline{B}(v_C^1 - v_C^2) \leq (v_T^1 - v_T^2) \tag{13}$$

$$v_C \geq 0 \Rightarrow \overline{B} v_C \geq 0. \tag{14}$$

H$_2$ The operator $\overline{L}$ defined in **H**$_1$ is contractive, i.e. satisfies

$$|\overline{L} v^1 - \overline{L} v^2|_\infty < |v^1 - v^2|_\infty \tag{15}$$

and satisfies a discrete maximum principle (DMP in short):

$$\overline{L} v^1 - \overline{L} v^2 \leq v^1 - v^2 \Rightarrow v^1 - v^2 \geq 0. \tag{16}$$

Under these hypothesis we will prove the convergence of a policy iteration algorithm to solve problem (10). We start with two lemmas.

Lemma 1 *For each $T \in T_{ad}$, $\mathcal{O}_T$ satisfies a discrete maximum principle (16) and $\mathcal{O}$ satisfies also a discrete maximum principle (16) .*

Proof : Let $T \in T_{\text{ad}}$, given v_1 and v_2, suppose that

$$\mathcal{O}_T v^1 - \mathcal{O}_T v^2 \leq v^1 - v^2.$$

From (12), we get

$$\overline{L} v_C^1 - \overline{L} v_C^2 \leq v_C^1 - v_C^2.$$

Then the discrete maximum principle (16) applied to $\overline{L}$ implies

$$v_C^1(x) - v_C^2(x) \geq 0 \quad \text{for all} \quad x \in C. \tag{17}$$

Now from (13) we have

$$\overline{B}(v_C^1 - v_C^2) \leq (v_T^1 - v_T^2)$$

and since $v_C^1(x) - v_C^2(x) \geq 0$, (14) implies

$$\overline{B}(v_C^1 - v_C^2) \geq 0.$$

We conclude that

$$v_T^1(x) - v_T^2(x) \geq 0 \quad \text{for all} \quad x \in T. \tag{18}$$

Together with (17), this implies

$$v^1(x) - v^2(x) \geq 0 \quad \text{for all } x \in \mathcal{E}.$$

Suppose now that

$$\mathcal{O}v^1 - \mathcal{O}v^2 \leq v^1 - v^2$$

and let $T_2 \in T_{\text{ad}}$ such that $\mathcal{O}_{T_2}v^2 = \mathcal{O}v^2$. We have :

$$\mathcal{O}_{T_2}v^1 - \mathcal{O}_{T_2}v^2 \leq \mathcal{O}v^1 - \mathcal{O}v^2 \leq v^1 - v^2.$$

Since $\mathcal{O}_{T_2}$ satisfies the DMP, this implies $v_1 - v_2 \geq 0$. □

Lemma 2 *For each $T \in T_{ad}$, $\mathcal{O}_T$ has a unique fixed point. If a fixed point exists for $\mathcal{O}$ it is unique.*

Proof : Lemma 1 implies that if a fixed point exists for $\mathcal{O}_T$ or $\mathcal{O}$, it is unique. Now, from (11) the existence of a fixed point of $\mathcal{O}_T$ is equivalent to the existence of a fixed point of the associated $\overline{L}$ operator. But $\overline{L}$ has indeed a fixed point since it is contractive. □

2.2 A policy iteration algorithm

The policy iteration algorithm for solving (7) consists in constructing two sequences of admissible policies $(T_k, k \in \mathbb{N}^\star)$ and functions $(v_k, k \in \mathbb{N})$ as follows: Note that here an admissible policy is nothing but a given admissible partition of $\mathcal{E}$. Let $v_0 \in V$ a given function. For $k \geq 0$ we do the following iterations :

- (step $2k$) Given v_k compute a policy T_{k+1} such that

$$T_{k+1} \in \operatorname*{Argmax}_{T \in T_{\text{ad}}} \{\mathcal{O}_T v_k\}. \tag{19}$$

  We may for example set :

$$T_{k+1} = \{x \in \mathcal{E}\,,\, Bv_k(x) > Lv_k(x)\}.$$

- (step $2k+1$) Let (T_{k+1}) be a given admissible policy, compute v_{k+1} as the solution of

$$v_{k+1} = \mathcal{O}_{T_{k+1}} v_{k+1}. \tag{20}$$

  Set $k \leftarrow k+1$ and return to step $2k$.

Theorem 1 – *(i) The sequence $(v_k(x), k \in \mathbb{N})$ is well defined and nondecreasing.*

- *(ii) Suppose that there exists a function v^+ which satisfies $\mathcal{O}_T v^+ \leq v^+$ for all $T \in T_{ad}$. Then the sequence $(v_k(x), k \in \mathbb{N})$ is bounded by $v^+(x)$ and converges for all $x \in \mathcal{E}$.*
- *(iii) The limit $v^\sharp$ of the sequence $(v_k, k \in \mathbb{N})$ is a fixed point of $\mathcal{O}$.*

Proof : (i) Given T_k, v_k is defined as the fixed point of $\mathcal{O}_{T_k}$, which exists and is unique by Lemma 2. Using the definition of T_{k+1} we have :

$$\mathcal{O}_{T_k} v_k \leq \mathcal{O}_{T_{k+1}} v_k. \tag{21}$$

So

$$\mathcal{O}_{T_{k+1}} v_{k+1} - \mathcal{O}_{T_{k+1}} v_k \leq \mathcal{O}_{T_{k+1}} v_{k+1} - \mathcal{O}_{T_k} v_k. \tag{22}$$

Now, using $v_k = \mathcal{O}_{T_k} v_k$ and $v_{k+1} = \mathcal{O}_{T_{k+1}} v_{k+1}$, we get

$$\mathcal{O}_{T_{k+1}} v_{k+1} - \mathcal{O}_{T_{k+1}} v_k \leq v_{k+1} - v_k. \tag{23}$$

Using Lemma 1 we conclude that $v_{k+1} \geq v_k$.

(ii) One sees easily that

$$\mathcal{O}_{T_k} v^+ - \mathcal{O}_{T_k} v_k \leq v^+ - v_k \tag{24}$$

which implies $v^+ - v_k \geq 0$.

(iii) In what follows, the existence of v^+ is not used. Let $(T_k, k \in \mathbb{N})$ be the sequence of partitions computed by the policy iteration algorithm. Since T_{ad} is finite we can find (k, k') such that $k < k'$ and $T_k = T_{k'}$. Lemma 2 which gives the uniqueness of the fixed point of $\mathcal{O}_{T_k}$ implies that $v_k = v_{k'}$. Combined with the monotonicity of the sequence (v_k), this implies that $v_k = v_{k+1}$. Thus :

$$\mathcal{O} v_k = \mathcal{O}_{T_{k+1}} v_k = \mathcal{O}_{T_{k+1}} v_{k+1} = v_{k+1} = v_k$$

and we conclude that v_k is a fixed point of $\mathcal{O}$. Moreover since a fixed point of $\mathcal{O}$ is also a super-solution of $\mathcal{O}$ it can play the role of v^+ and it gives an upper bound of the sequence v_k. Then $v_{k'}(x)$ is constant for $k' \geq k$ and the sequence $(v_k(x))_{k \in \mathbb{N}}$ converges in a finite number of iterations. □

2.3 An example related to Bellman equations

Let $\mathcal{E} \stackrel{\text{def}}{=} \{1, \ldots, n\}$ be the state space. A partition T of $\mathcal{E}$ will be admissible if $1 \notin T$. Let M be a $n \times n$ stochastic matrix, c and k two vectors of dimension n, $\lambda > 0$, $\sigma : [1, n] \mapsto [1, n]$ an integer function such that $\sigma(i) < i$ and $\mathbf{B}$ a $n \times n$ stochastic matrix (except for the first line) defined by :

$$\mathbf{B}_{i,j} = \begin{cases} 1 & \text{if } j = \sigma(i) \text{ and } i \neq 1 \\ 0 & \text{elsewhere.} \end{cases} \tag{25}$$

The matrix $\mathbf{B}$ is thus lower tridiagonal with zeros on the diagonal. We define

$$Lv \stackrel{\text{def}}{=} \frac{1}{1+\lambda}(Mv + c) \quad \text{and} \quad Bv \stackrel{\text{def}}{=} (\mathbf{B}v + k). \tag{26}$$

Lemma 3 *Let $T \in T_{ad}$ and let $v = (v_C, v_T)$ be the decomposition of the vector v on the set T and its complementary $C \stackrel{def}{=} \mathcal{E}\backslash T$. Let σ^p denote the p-composition of the function σ ($\sigma^0(x) \stackrel{def}{=} x$) and for each $x \in T$, let :*

$$p(x) \stackrel{def}{=} inf\{p \geq 1 \text{ such that } \sigma^p(x) \in \mathcal{E}\backslash T\}.$$

Define the operator $\overline{B}$ as

$$(\overline{B}v)(x) = v(\sigma^{p(x)}(x))$$

and the vector $\overline{k}$ as

$$\overline{k}(x) = \sum_{i=0}^{p(x)-1} k(\sigma^i(x)).$$

The value of $\overline{B}v$ only depends on the value of v on C and we thus write $\overline{B}v_C$. We have

- *(i) $\overline{B}$ fulfills property (14).*
- *(ii) $v = Bv$ on $T \Leftrightarrow v_T = \overline{B}v_C + \overline{k}_T$.*
- *(iii) $Bv^1 - Bv^2 \leq v^1 - v^2$ on $T \Rightarrow \overline{B}v^1_C - \overline{B}v^2_C \leq v^1_T - v^2_T$.*

Proof : (i) For i in T, let $\sigma^p(i)$ denote the p-composition of the function σ. Since σ is strictly decreasing and $\sigma(2) = 1 \in \mathcal{E}\backslash T$, the pth-composition of σ starting from a point in T will end up in $\mathcal{E}\backslash T$ after a finite number of steps. Thus for i in T we can define $p(i)$ the smallest value of $p \geq 1$ (such that $\sigma^p(i) \in \mathcal{E}\backslash T$ and $\overline{B}$ is well defined. It is clear that $\overline{B}$ satisfies property (14) and only depends on the value of v on C.

(ii) Suppose that v is a function such that $v = Bv$ on the set T. We then have for $x \in T$

$$v(x) = Bv(x) = v(\sigma(x)) + k(x).$$

By iteration we have

$$v(x) = v(\sigma^p(x)) + \sum_{i=0}^{p(x)-1} k(\sigma^i(x)) \quad \text{for} \quad p \leq p(x)\,.$$

Taking $p = p(x)$, we get $v(x) = (\overline{B}v)(x) + \overline{k}(x)$.

Suppose now that

$$v(x) = (\overline{B}v)(x) + \overline{k}(x) \quad \text{for all} \quad x \in T\,. \tag{27}$$

Let $x \in T$. If $p(x) = 1$ we get immediately $v(x) = (Bv)(x)$.

If $p(x) > 1$ then $\sigma(x) \in T$ and $p(\sigma(x)) = p(x) - 1$. We have, using the definition of $\overline{B}$

$$(\overline{B}v)(x) = v(\sigma^{p(x)}(x))$$

and

$$(\overline{B}v)(\sigma(x)) = v(\sigma^{p(\sigma(x))}\sigma(x)) = v(\sigma^{p(x)}(x)).$$

Consequently

$$(\overline{B}v)(x) = (\overline{B}v)(\sigma(x)).$$

We also have :

$$\begin{aligned}\overline{k}(x) = \sum_{i=0}^{p(x)-1} k(\sigma^i(x)) &= k(x) + \sum_{i=1}^{p(\sigma(x))} k(\sigma^i(x)) \\ &= k(x) + \overline{k}(\sigma(x)).\end{aligned}$$

Using now (27), we get $v(x) = v(\sigma(x)) + k(x) = Bv(x)$. We conclude that $v = Bv$ on T.

(iii) We prove *(iii)* similarly. □

We can suppose, modulo a permutation on the elements of $\mathcal{E}$, that the first elements of $\mathcal{E}$ belong to C and the others belong to T. Then the matrices M and $\mathbf{B}$ can be written as :

$$M = \begin{pmatrix} M_{CC} & M_{CT} \\ M_{TC} & M_{TT} \end{pmatrix} \quad \text{and} \quad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{CC} & \mathbf{B}_{CT} \\ \mathbf{B}_{TC} & \mathbf{B}_{TT} \end{pmatrix}$$

and

$$\mathcal{O}_T v = \begin{pmatrix} \frac{1}{1+\lambda} M_{CC} & \frac{1}{1+\lambda} M_{CT} \\ \mathbf{B}_{TC} & \mathbf{B}_{TT} \end{pmatrix} \begin{pmatrix} v_C \\ v_T \end{pmatrix} + \begin{pmatrix} c_C \\ k_T \end{pmatrix}$$

Define $\overline{L}$ and $\overline{M}$ as :

$$\overline{M} \stackrel{\text{def}}{=} M_{CC} + M_{CT} * \overline{B} \text{ and } \overline{L} v_C \stackrel{\text{def}}{=} \frac{1}{1+\lambda}(\overline{M} v_C + \overline{c}_C) \tag{28}$$

where $\overline{B}$ is given in Lemma 3 and $\overline{c}_C \stackrel{\text{def}}{=} M_2 \overline{k} + c_C$

Lemma 4 *Using $\overline{B}$ and $\overline{L}$ the hypothesis $\mathbf{H}_1$ and $\mathbf{H}_2$ are fulfilled.*

Proof : The entries of the matrix $\overline{M}$ are nonnegative since $\overline{B}$ has nonnegative entries and M is stochastic. Since $1 \notin T$ we have $(\mathbf{B}e)_T = e_T$ where $e = (1, \ldots, 1)^T$. Proceeding as in Lemma 3 *(ii)*, we get $e_T = \overline{B} e_C$. Then :

$$\begin{aligned} \overline{M} e_C &= M_{CC} e_C + M_{CT} \overline{B} e_C \\ &= M_{CC} e_C + M_{CT} e_T = \begin{pmatrix} M_{CC} & M_{CT} \end{pmatrix} e = e_C \end{aligned} \tag{29}$$

The matrix $\overline{M}$ is thus stochastic and $\overline{L}$ is linear, contractive and satisfies the discrete maximum principle. So $\mathbf{H}_2$ is satisfied.

Suppose now that v is a solution of $\mathcal{O}_T v = v$. This is equivalent to

$$v_T = Bv \text{ on } T \tag{30}$$

$$v_C = Lv \text{ on } C. \tag{31}$$

By Lemma 3, (30) is equivalent to

$$v_T = \overline{B} v_C + \overline{k}. \tag{32}$$

Using (32), we can write

$$\begin{aligned} \overline{L} v_C &= \frac{1}{1+\lambda}((M_{CC} + M_{CT} * \overline{B}) v_C + \overline{c}_C) = \frac{1}{1+\lambda}(M_{CC} v_C + M_{CT} v_T + c_C) \\ &= \frac{1}{1+\lambda}((Mv)_C + c_C) = Lv(x) \text{ on } C. \end{aligned}$$

Combined with (31), this gives $v_C = \overline{L} v_C$. Using Lemma 3 again the converse implication follows easily.

Suppose now that

$$\mathcal{O}_T v^1 - \mathcal{O}_T v^2 \leq v^1 - v^2.$$

This implies on T

$$Bv^1 - Bv^2 \leq v^1 - v^2.$$

Using lemma 3 this implies that

$$(v^1 - v^2)_T \geq \overline{B}(v^1 - v^2)_C. \tag{33}$$

On C, we have

$$Lv^1 - Lv^2 \leq v^1 - v^2.$$

Moreover

$$\begin{aligned}
Lv^1 - Lv^2 &= \frac{1}{1+\lambda}(Mv^1 - Mv^2) \\
&= \frac{1}{1+\lambda}\left(\begin{pmatrix} M_{CC} & M_{CT}\end{pmatrix}\begin{pmatrix} v_C^1 - v_C^2 \\ v_T^1 - v_T^2 \end{pmatrix}\right) \\
&= \frac{1}{1+\lambda}\left(M_{CC}(v_C^1 - v_C^2) + M_{CT}(v_T^1 - v_T^2)\right) \\
&\geq \frac{1}{1+\lambda}\left(M_{CC}(v_C^1 - v_C^2) + M_{CT}\overline{B}(v_C^1 - v_C^2)\right).
\end{aligned}$$

by using (33). We thus get

$$\frac{1}{1+\lambda}\left(M_{CC} + M_{CT}\overline{B}\right)(v_C^1 - v_C^2) \leq v_C^1 - v_C^2.$$

Consequently

$$\frac{1}{1+\lambda}\left(\overline{M}v_C^1 - \overline{M}v_C^2\right) \stackrel{\text{def}}{=} \overline{L}v_C^1 - \overline{L}v_C^2 \leq v_C^1 - v_C^2$$

and (12) is obtained. □

Remark 1 When $k \equiv 0$, one can easily checks that $v^+ \stackrel{\text{def}}{=} |c|_\infty/\lambda$ is, for all $T \in T_{\text{ad}}$ a super-solution of $\mathcal{O}_T$ ($\mathcal{O}_T v^+ \leq v^+$). When $k \neq 0$, Lemma 6 below gives an upper bound for $v^\star$.

A probabilistic interpretation in terms of an optimal control problem. Let $(X_n, n \in \mathbb{N})$ be an homogeneous controlled Markov chain with transition matrix M^u defined on a finite state space $\mathcal{E}$. We assume that the control u has only two possible values: $u \in \{\mathbf{c}, \mathbf{t}\}$. The transition matrix takes thus two values $M^{\mathbf{c}} \stackrel{\text{def}}{=} M$ and $M^{\mathbf{t}} \stackrel{\text{def}}{=} B$. We consider the following optimal control problem:

$$v^\star(x) \stackrel{\text{def}}{=} \max_{\mathbb{U}\in\mathcal{U}} v^{\mathbb{U}}(x) \quad \text{with} \quad v^{\mathbb{U}}(x) \stackrel{\text{def}}{=} \mathbb{E}_x\left[\sum_{k=0}^{+\infty}\prod_{i=0}^{k}\xi(X_i, U_i)C(X_k, U_k)\right] \tag{34}$$

where $\mathcal{U}$ stands for the stationary Markovian strategies. The profit and discount rate functions are defined by :

$$\xi(x,u) \stackrel{\text{def}}{=} \begin{cases} (1+\lambda)^{-1} & \text{if } u = \mathbf{c} \\ 1 & \text{if } u = \mathbf{t}. \end{cases} \quad C(x,u) \stackrel{\text{def}}{=} \begin{cases} c(x) & \text{if } u = \mathbf{c} \\ k(x) & \text{if } u = \mathbf{t}. \end{cases} \tag{35}$$

To each subset T of $\mathcal{E}$ we associate a stationary policy $\mathbb{U}_T = (U_i, i \geq 0)$ where $U_i = U(X_i)$ for all $i \geq 0$, and U is a function: $\mathcal{E} \mapsto \{\mathbf{c}, \mathbf{t}\}$ such that $U(x) = \mathbf{t}$ for $x \in T$ and $U(x) = \mathbf{c}$ for $x \in \mathcal{E}\backslash T$.

Lemma 5 *$v^{\mathbb{U}_T}$ is the fixed point of the operator $\mathcal{O}_T$ and $v^\star$ is the fixed point of $\mathcal{O}$.*

Proof : For an initial state $x \in \mathcal{E}$, we define the stopping time :

$$\tau_T^x = \inf\{k \geq 0 \text{ such that } X_k^x \in T\}.$$

Let C_∞ be the subset of $\mathcal{E}\backslash T$ consisting of the states x such that $\tau_T^x = \infty$. When $x \in C_\infty$, $v^{\mathbb{U}_T}(x)$ can be rewritten as :

$$v^{\mathbb{U}_T}(x) = \mathbb{E}_x \left[\sum_{k=0}^{+\infty} \frac{1}{(1+\lambda)^{k+1}} c(X_k) \right]$$

and the restriction of $v^{\mathbb{U}_T}$ to C_∞ is the solution of the equation :

$$v^{\mathbb{U}_T}(x) = \frac{1}{1+\lambda}(M_{C_\infty} v^{\mathbb{U}_T}(x) + c(x)) \text{ for } x \in C_\infty \tag{36}$$

where M_{C_∞} is the restriction of M to C_∞.

For $x \in \mathcal{E}\backslash C_\infty$, the strong Markov property implies :

$$v^{\mathbb{U}_T}(x) \stackrel{\text{def}}{=} \mathbb{E}_x \Bigg[\sum_{k=0}^{\tau_T^x - 1} \prod_{i=0}^{k} \xi(X_i, U_i) C(X_k, U_k) + \prod_{i=0}^{\tau_T^x} \xi(X_i, U_i) C(X_{\tau_T^x}, U_{\tau_T^x}) + \prod_{i=0}^{\tau_T^x} \xi(X_i, U_i) v^{\mathbb{U}_T}(X_{\tau_T^x + 1}) \Bigg].$$

Using the definition of τ_T^x, we obtain

$$v^{\mathbb{U}_T}(x) = \mathbb{E}_x \left[\sum_{k=0}^{\tau_T^x - 1} \frac{1}{(1+\lambda)^{k+1}} c(X_k) + \frac{1}{(1+\lambda)^{\tau_T^x}} \Psi_T(X_{\tau_T^x}) \right]$$

$$\Psi_T(x) \stackrel{\text{def}}{=} B v^{\mathbb{U}_T}(x) + k(x)$$

It is well known that $v^{\mathbb{U}_T}(x)$ then satisfies the Kolmogorov equation

$$v^{\mathbb{U}_T}(x) \stackrel{\text{def}}{=} \begin{cases} \frac{1}{1+\lambda}(M v^{\mathbb{U}_T}(x) + c(x)) & \text{if } x \in \mathcal{E}\backslash T,\ x \notin C_\infty. \\ B v^{\mathbb{U}_T}(x) + k(x) & \text{if } x \in T. \end{cases} \tag{37}$$

Combining (36) and (37), we obtain that $v^{\mathbb{U}_T}$ is a fixed point of $\mathcal{O}_T$. Since the set of stationary Markovian strategies is finite, $v^\star$ is well defined and there exists $T^\star$ such that $v^\star = v^{\mathbb{U}_{T^\star}}$. This implies that $v^\star$ coincides with the unique fixed point of $\mathcal{O}$. □

Lemma 6 *We have $v^\star \leq (|c|_\infty + n|k|_\infty)/\lambda$ where n is the cardinal of $\mathcal{E}$.*

Proof : For a given strategy $\mathbb{U}_T$, let $(X_{\rho(n)})$ be the subsequence of X_n which belongs to subset $C = \mathcal{E}\backslash T$. Using the probabilistic interpretation of $v^{\mathbb{U}_T}(x)$ we can write :

$$v^{\mathbb{U}_T}(x) = \mathbb{E}_x \left[\sum_{k=0}^{+\infty} \frac{1}{(1+\lambda)^{k+1}} \left(c(X_{\rho(k)}) + \sum_{j=\rho(k)+1}^{\rho(k+1)-1} k(X_j) \right) \right] \tag{38}$$

But we also have $|\rho(k+1) - \rho(k)| \leq n$ since after at most n successive transitions in T the chain is in C. The result follows. □

3 A fixed point problem with a nonexpansive operator and non finite sets of controls

We consider now the fixed point problem :

$$v(x) = \max(\sup_{w\in\mathcal{W}} L^w v(x), \sup_{z\in\mathcal{Z}} B^z v(x)), \quad x \in \mathcal{E} \tag{39}$$

where $\mathcal{E}$ is a finite set, $\mathcal{W}$ and $\mathcal{Z}$ are general control sets. The operators L^w and B^z depend now on control variables w and z.

Let T be a subset of $\mathcal{E}$, $w \in \mathcal{W}$ and $z \in \mathcal{Z}$ and denote by $\mathcal{O}_{T,w,z}$ the operator :

$$\mathcal{O}_{T,w,z} v(x) \stackrel{\text{def}}{=} \begin{cases} L^w v(x) & \text{if } x \in C \quad \text{where} \quad C \stackrel{\text{def}}{=} \mathcal{E}\backslash T, \\ B^z v(x) & \text{if } x \in T. \end{cases} \tag{40}$$

Problem (39) is equivalent to

$$v(x) = \sup_{T\in\mathcal{P}(\mathcal{E}), w\in\mathcal{W}, z\in\mathcal{Z}} \mathcal{O}_{T,w,z}.$$

Set

$$\mathcal{O}v(x) \stackrel{\text{def}}{=} \sup_{T\in T_{\text{ad}}, w\in\mathcal{W}, z\in\mathcal{Z}} \mathcal{O}_{T,w,z}.$$

where T_{ad} is defined in (8). We restrict ourselves to the following problem

$$v(x) = \sup_{T\in T_{\text{ad}}, w\in\mathcal{W}, z\in\mathcal{Z}} \mathcal{O}_{T,w,z}. \tag{41}$$

We make the following assumptions :

H$_1$ For each given $T \in T_{\text{ad}}$, $W : \mathcal{E} \to \mathcal{W}$ and $Z : \mathcal{E} \to \mathcal{Z}$, we can build two operators $\overline{L}^{T,W,Z}$, $\overline{B}^{T,W,Z}$ and a function $\overline{k}^{T,Z}$ such that :

$$\mathcal{O}_{T,W,Z} v = v \Leftrightarrow \overline{L}^{T,W,Z} v_C = v_C \text{ and } v_T = \overline{B}^{T,W,Z} v_C + \overline{k}^{T,Z} \tag{42}$$

$$\mathcal{O}_{T,W,Z} v^1 - \mathcal{O}_{T,w,z} v^2 \le v^1 - v^2 \Rightarrow \overline{L}^{T,W,Z} v_C^1 - \overline{L}^{T,W,C} v_C^2 \le v_C^1 - v_C^2 \tag{43}$$

$$\text{and } \overline{B}^{T,W,Z} (v_C^1 - v_C^2) \le (v_T^1 - v_T^2) \tag{44}$$

$$v_C \ge 0 \Rightarrow \overline{B}^{T,W,Z} v_C \ge 0. \tag{45}$$

H$_2$ For each (T, W, Z), the operator $\overline{L}^{T,W,Z}$ defined by **H**$_1$ is contractive and satisfies a discrete maximum principle :

$$\overline{L}^{T,W,Z} v^1 - \overline{L}^{T,W,Z} v^2 \le v^1 - v^2 \Rightarrow v^1 - v^2 \ge 0. \tag{46}$$

H$_3$
- For each $v \in V$, $\text{Argmax}\{L^w v(x), w \in \mathcal{W}\}$ and $\text{Argmax}\{B^z v(x), z \in \mathcal{Z}\}$ exist.
- For each $T \in T_{ad}$, $W : \mathcal{E} \to \mathcal{W}$ and $Z : \mathcal{E} \to \mathcal{Z}$, the operator $\mathcal{O}_{T,W,Z}$ is nondecreasing.
- The operator $\mathcal{O}$ is continuous and nondecreasing.

Lemma 7
- *For each $T \in T_{ad}$, $W : \mathcal{E} \to \mathcal{W}$ and $Z : \mathcal{E} \to \mathcal{Z}$, the operator $\mathcal{O}_{T,W,Z}$ satisfies a discrete maximum principle and $\mathcal{O}$ satisfies also a discrete maximum principle.*
- *For each $T \in T_{ad}$, $W : \mathcal{E} \to \mathcal{W}$ and $Z : \mathcal{E} \to \mathcal{Z}$, the operator $\mathcal{O}_{T,W,Z}$ has a unique fixed point. If a fixed point exists for $\mathcal{O}$ it is unique.*

Proof : We use similar arguments as in the proof of lemma 1 and lemma 2. □

3.1 A policy iteration algorithm

The policy iteration algorithm for solving (41) consists in constructing two sequences of admissible policies $((T_k, W_k, Z_k), k \in \mathbb{N})$ and functions $(v_k, k \in \mathbb{N})$ as follows: Let $v_0 \in V$ a given function. For $k \geq 0$ we do the following iterations :

– (step $2k$) Given v_k, compute an admissible policy $(T_{k+1}, W_{k+1}, Z_{k+1})$ such that

$$(T_{k+1}, W_{k+1}, Z_{k+1}) \in \operatorname*{Argmax}_{T,W,K}\{\mathcal{O}_{T,W,Z} v_k\}. \tag{47}$$

In other words $\mathcal{O}v_k = \mathcal{O}_{T_{k+1},W_{k+1},Z_{k+1}} v_k$.

– (step $2k+1$) Let $(T_{k+1}, W_{k+1}, Z_{k+1})$ be a given admissible policy, compute v_{k+1} as the solution of

$$v_{k+1} = \mathcal{O}_{T_{k+1},W_{k+1},Z_{k+1}} v_{k+1}. \tag{48}$$

Set $k \leftarrow k+1$ and go to step $2k$.

Theorem 2 – *(i) The sequence $(v_k(x), k \in \mathbb{N})$ is well defined and nondecreasing.*
– *(ii) Suppose that there exists a function v^+ which satisfies $\mathcal{O}_{T,W,Z} v^+ \leq v^+$ for all (T, W, Z).Then the sequence $(v_k(x), k \in \mathbb{N})$ is bounded by $v^+(x)$ and converges for all $x \in \mathcal{E}$.*
– *(iii) There exists v_0 such that $v_0 \leq \mathcal{O}v_0$. The limit $v^\sharp$ of the sequence $(v_k, k \in \mathbb{N})$ starting from this v_0 is a fixed point of $\mathcal{O}$.*

Proof :
(i) and *(ii)* are proved similarly as in the proof of Theorem 1.
The proof of *(iii)* is different as explained now: Set for writing simplicity

$$\mathcal{O}_{(k)} \stackrel{\text{def}}{=} \mathcal{O}_{T_k,W_k,Z_k}.$$

For all $x \in \mathcal{E}$, the sequence $(v_k(x), k \in \mathbb{N})$ is nondecreasing and bounded by $v^+(x)$, it converges to $v^\sharp(x) \leq v^+(x)$. Let us prove by induction that

$$\mathcal{O}^k v_0 \leq v_k. \tag{49}$$

We have by (47) at step $k = 0$,

$$\mathcal{O}v_0 = \mathcal{O}_{(1)} v_0.$$

and by (48)

$$v_1 = \mathcal{O}_{(1)} v_1.$$

So $\mathcal{O}v_0 \leq v_1$ if and only if $\mathcal{O}_{(1)} v_0 \leq \mathcal{O}_{(1)} v_1$ which is true since $\mathcal{O}_{(1)}$ is nondecreasing (by $\mathbf{H}_3$) and $v_0 \leq v_1$.

Suppose (49) holds. By $\mathbf{H}_3$, $\mathcal{O}$ is nondecreasing, so

$$\mathcal{O}^{k+1} v_0 \leq \mathcal{O}v_k.$$

by (47),

$$\mathcal{O}v_k = \mathcal{O}_{(k+1)} v_k.$$

By $\mathbf{H}_3$, $\mathcal{O}_{(k+1)}$ is nondecreasing. Since $v_k \leq v_{k+1}$, $\mathcal{O}_{(k+1)} v_k \leq \mathcal{O}_{(k+1)} v_{k+1}$. By (48)

$$v_{k+1} = \mathcal{O}_{(k+1)} v_{k+1}.$$

Consequently

$$\mathcal{O}^{k+1} v_0 \leq v_{k+1}.$$

Suppose that there exists v_0 such that $v_0 \leq \mathcal{O}v_0$. The sequence $\gamma_k = \mathcal{O}^k v_0$ is nondecreasing and bounded and thus converges to $\overline{v}$. Taking the limit when $k \to \infty$ in (49), we get

$$\overline{v} \leq v^\sharp \leq v^+.$$

Moreover $\overline{v}$ is a fixed point of $\mathcal{O}$ since γ_k converges, $\gamma_{k+1} = \mathcal{O}\gamma_k$ and $\mathcal{O}$ is continuous. Now

$$v_k = \mathcal{O}_{(k)} v_k \leq \mathcal{O}_{(k+1)} v_k = \mathcal{O} v_k$$

Taking the limit when $k \to \infty$, we get

$$v^\sharp \leq \mathcal{O} v^\sharp. \tag{50}$$

Suppose that there exists $\overline{v}$ such that

$$\overline{v} = \mathcal{O}\overline{v}. \tag{51}$$

Subtracting (51) from (50) and using a discrete maximum principle for $\mathcal{O}$ (lemma 7), we get $v^\sharp \leq \overline{v}$.

We conclude that $v^\sharp = \overline{v}$ and $v^\sharp$ is a fixed point of $\mathcal{O}$.

It remains to prove that v_0 such that $v_0 \leq \mathcal{O}v_0$ exists. For $v \in V$, set

$$Lv = \sup_{w \in \mathcal{W}} L^w v.$$

Consider the special strategy $T = \emptyset$. Hypothesis $\mathbf{H}_2$implies that L is contractive. It thus has a fixed point v_0 and $\mathcal{O}v_0 \geq Lv_0 = v_0$. □

3.2 An example

Let $(M^w, w \in \mathcal{W})$ be a family of $n \times n$ stochastic matrices, $(c^w, w \in \mathcal{W})$ and $(k^z, z \in \mathcal{Z})$ two families of vectors of dimension n, λ a strictly positive real number, $\sigma : [1, n] \times \mathcal{Z} \mapsto [1, n]$ an integer function such that $\sigma(i, z) < i$ and $(\mathbf{B}^z, z \in \mathcal{Z})$ a family of $n \times n$ stochastic matrices (except for the first line) defined by :

$$\mathbf{B}^z_{i,j} = \begin{cases} 1 & \text{if } j = \sigma(i,z) \text{ and } i \neq 1 \\ 0 & \text{elsewhere.} \end{cases} \tag{52}$$

We define

$$L^w v \stackrel{\text{def}}{=} \frac{1}{1+\lambda}(M^w v + c^w) \quad \text{and} \quad B^z v \stackrel{\text{def}}{=} (\mathbf{B}^z v + k^z). \tag{53}$$

We define the state space $\mathcal{E} \stackrel{\text{def}}{=} \{1, \ldots, n\}$ and the admissible partitions as the ones such that $1 \notin T$.

Let $T \in T_{\text{ad}}$, $Z : \mathcal{E} \to \mathcal{Z}$ and $W : \mathcal{E} \to \mathcal{W}$. Define the function $\bar{\sigma}$ on $\mathcal{E}$ by $\bar{\sigma}(x) = \sigma(x, Z(x))$. For each $x \in T$, let :

$$p^{T,Z}(x) \stackrel{\text{def}}{=} \inf\{p \geq 1 \text{ such that } \bar{\sigma}^p(x) \in \mathcal{E} \backslash T\}.$$

Define the operator $\overline{B}^{T,Z}$ as

$$(\overline{B}^{T,Z} v)(x) = v(\bar{\sigma}^{p^{T,Z}(x)}).$$

and the vector $\overline{k}^{T,Z}$ as

$$\overline{k}^{T,Z}(x) = \sum_{i=0}^{p^{T,Z}(x)-1} k(\bar{\sigma}^i(x)).$$

We can suppose, modulo a permutation on the elements of $\mathcal{E}$ that the first elements of $\mathcal{E}$ belong to C and the others belong to T. Then the matrices M^w and $\mathbf{B}^z$ can be written as :

$$M^w = \begin{pmatrix} M^w_{CC} & M^w_{CT} \\ M^w_{TC} & M^w_{TT} \end{pmatrix} \quad \text{and} \quad \mathbf{B}^z = \begin{pmatrix} \mathbf{B}^z{}_{CC} & \mathbf{B}^z{}_{CT} \\ \mathbf{B}^z{}_{TC} & \mathbf{B}^z{}_{TT} \end{pmatrix}$$

and

$$\mathcal{O}_{T,Z,W} v = \begin{pmatrix} \frac{1}{1+\lambda} M^w_{CC} & \frac{1}{1+\lambda} M^w_{CT} \\ \mathbf{B}^z{}_{TC} & \mathbf{B}^z{}_{TT} \end{pmatrix} \begin{pmatrix} v_C \\ v_T \end{pmatrix} + \begin{pmatrix} c^w_C \\ k^z_T \end{pmatrix}$$

Define $\overline{L}^{T,W,Z}$ and $\overline{M}^{T,W,Z}$ as :

$$\overline{M}^{T,W,Z} \stackrel{\text{def}}{=} M^W_{CC} + M^W_{CT} * \overline{B}^{T,Z}$$
$$\overline{L}^{T,W,Z} v_C \stackrel{\text{def}}{=} \frac{1}{1+\lambda} (\overline{M}^{T,W,Z} v_C + \overline{c}_C^{T,W,Z}).$$

where $\overline{B}$ is given in lemma 3 and $\overline{c}_C^{T,W,Z} \stackrel{\text{def}}{=} M_{CT} \overline{k}^{T,Z} + c^W_C$

Lemma 8 *$\overline{B}^{T,Z}$ and $\overline{L}^{T,W}$ satisfy the hypothesis $\mathbf{H}_1$, $\mathbf{H}_2$. Moreover, suppose: (i) $\mathcal{W}$ and $\mathcal{Z}$ are compact sets, (ii) $w \mapsto (M^w, c^w)$ is continuous, (iii) $z \mapsto k^z$ is continuous. Then $\mathbf{H}_3$ is also satisfied.*

Proof : The proof that $\mathbf{H}_1$ and $\mathbf{H}_2$ are fulfilled is similar to the proof of lemma 4. For $v \in V$ and for all $x \in \mathcal{E}$, Argmax$\{L^w v(x), w \in \mathcal{W}\}$ exists since $w \mapsto L^w v(x)$ is continuous in w on a compact set $\mathcal{W}$. Note that $\mathbf{B}^z v(x)$ can take a finite number of values and consequently Argmax$\{B^z v(x), z \in \mathcal{Z}\}$ exists.

For all (T, W, Z) the operator $\mathcal{O}_{T,W,Z}$ is nondecreasing since M^W and $\mathbf{B}^Z$ are matrices with nonnegative entries. The operator $\mathcal{O}$ is also nondecreasing since it is the maximum of nondecreasing operators. The operator $\mathcal{O}$ is nonexpansive and thus continuous. □

A probabilistic interpretation. Let $(X_n, n \in \mathbb{N})$ be an homogeneous controlled Markov chain with transition matrix M^U defined on a finite state space $\mathcal{E}$. We assume here that the control $U = (u, w, z)$ where u has only two possible values $u \in \{\mathbf{c}, \mathbf{t}\}$, $w \in W$ and $z \in Z$.

$$\mathbb{M}^{(u,z,w)} \stackrel{\text{def}}{=} \begin{cases} M^w & \text{if } u = \mathbf{c}, \\ B^z & \text{if } u = \mathbf{t}. \end{cases}$$

We consider the optimal control problem :

$$v^\star(x) \stackrel{\text{def}}{=} \sup_{\mathbb{U} \in \mathcal{U}} v^{\mathbb{U}}(x) \tag{54}$$

with :

$$v^{\mathbb{U}}(x) \stackrel{\text{def}}{=} \mathbb{E}_x \left[\sum_{n=0}^{+\infty} \prod_{i=0}^{n} \xi(X_i, U_i) C(X_n, U_n, W_n) \right]$$

where $\mathcal{U}$ stands for the stationary Markovian strategies. The profit and discount rate functions are defined by :

$$\xi(x,u) \stackrel{\text{def}}{=} \begin{cases} (1+\lambda)^{-1} & \text{if } u = \mathbf{c} \\ 1 & \text{if } u = \mathbf{t}. \end{cases} \quad C(x,u,w) \stackrel{\text{def}}{=} \begin{cases} c(x,w) & \text{if } u = \mathbf{c} \\ k(x,z) & \text{if } u = \mathbf{t}. \end{cases} \tag{55}$$

To each subset T of $\mathcal{E}$ and given functions $W : \mathcal{E} \mapsto \mathcal{W}$ and $Z : \mathcal{E} \mapsto \mathcal{Z}$ we associate a stationary policy denoted by $\mathbb{U}_{T,Z,W} = (U_i, i \geq 0)$ where $U_i = U(X_i)$ and $U(x) = (\mathbf{t}, Z(x))$ for $x \in T$ and $U(x) = (\mathbf{c}, W(x))$ for $x \in \mathcal{E} \backslash T$ (To be rigorous we should write $U(x) = (\mathbf{t}, Z(x), W(x))$ but w (resp. z) is not involved when $u = \mathbf{t}$ (resp. $u = \mathbf{c}$)).

Lemma 9 *Let (T, W, Z) be an admissible strategy and $\mathcal{O}_{T,W,Z}$ the operator defined with B^Z and L^W as in equations* (52) *and* (53)*. Let v be a fixed function. The following property holds :*

$$\mathcal{O}_{T,W,Z}(v + \alpha) \geq \mathcal{O}_{T,W,Z}(v) + \frac{\alpha}{1+\lambda} \quad \forall \alpha \in \mathbb{R}.$$

Proof : This is straightforward since

$$B^Z(v + \alpha) = B^Z(v) + \alpha \geq B^Z(v) + \frac{\alpha}{1+\lambda}$$

and

$$L^W(v + \alpha) \geq L^W v + \frac{\alpha}{1+\lambda}.$$

□

Lemma 10 *$v^{\mathbb{U}_{T,Z,W}}$ is the fixed point of the operator $\mathcal{O}_{T,W,Z}$ and $v^\star$ is the fixed point of $\mathcal{O}$.*

Proof : Using Remark 1 and Theorem 2 we know that a fixed point of $\mathcal{O}$ exists. The fact that, for a fixed strategy (T, W, Z), $v^{\mathbb{U}_{T,Z,W}}$ is the fixed point of the operator $\mathcal{O}_{T,W,Z}$ can be proved similarly as in the proof of Lemma 5. Note also that $v^{\mathbb{U}_{T,Z,W}} \leq v^+$, thus $v^\star$ defined in (54) is well defined. It remains to show that $v^\star$ coincides with the fixed point $\hat{v}$ of $\mathcal{O}$. Let $\epsilon > 0$ be fixed and let $T_\epsilon, W_\epsilon, Z_\epsilon$ be an admissible policy such that

$$v^\star - \epsilon \leq v_\epsilon$$

where we have denoted

$$v_\epsilon \stackrel{\text{def}}{=} v^{\mathbb{U}_{T_\epsilon, W_\epsilon, Z_\epsilon}}.$$

We have

$$v_\epsilon = \mathcal{O}_{T_\epsilon, W_\epsilon, Z_\epsilon} v_\epsilon \leq \mathcal{O} v_\epsilon.$$

Subtracting this inequality to the equality

$$\hat{v} = \mathcal{O}\hat{v}$$

we obtain, by using the DMP for $\mathcal{O}$

$$\hat{v} \geq v_\epsilon. \tag{56}$$

We then conclude that $\hat{v} \geq v^\star$. Let now $(T_\epsilon, W_\epsilon, Z_\epsilon)$ be a strategy such that

$$\mathcal{O}\hat{v} - \epsilon \leq \mathcal{O}_{T_\epsilon, W_\epsilon, Z_\epsilon} \hat{v}. \tag{57}$$

Denote v_ϵ the fixed point of $\mathcal{O}_{T_\epsilon, W_\epsilon, Z_\epsilon}$:

$$\mathcal{O}_{T_\epsilon, W_\epsilon, Z_\epsilon} v_\epsilon = v_\epsilon \tag{58}$$

Since $\hat{v} = \mathcal{O}\hat{v}$, (57) can be rewritten as

$$-\mathcal{O}_{T_\epsilon, W_\epsilon, Z_\epsilon} \hat{v} \leq -\hat{v} + \epsilon.$$

Together with (58), this gives

$$\mathcal{O}_{T_\epsilon,W_\epsilon,Z_\epsilon} v_\epsilon - \mathcal{O}_{T_\epsilon,W_\epsilon,Z_\epsilon} \hat{v} \le \epsilon + v_\epsilon - \hat{v}$$

which combined with Lemma 9 leads to :

$$\mathcal{O}_{T_\epsilon,W_\epsilon,Z_\epsilon} v_\epsilon - \mathcal{O}_{T_\epsilon,W_\epsilon,Z_\epsilon} (\hat{v} - \beta) \le v_\epsilon - (\hat{v} - \beta) \text{ with } \beta = \frac{1+\lambda}{\lambda}\epsilon$$

Using the DMP again we conclude that $v_\epsilon \ge \hat{v} - \beta$. This leads to $v^\star \ge \hat{v} - \beta$ and to $v^\star \ge \hat{v}$ as $\epsilon \mapsto 0$. □

4 Application to the numerical analysis of quasi-variational inequalities associated to combined stochastic and impulse control problems

Let $B(t) = B(t,\omega)\ ;\ t \ge 0,\ \omega \in \Omega$ be a d-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, P)$, $B(0) = 0$ a.s. Let U be a given compact set of $\mathbb{R}^\ell$. Let $b : \mathbb{R}^k \times U \to \mathbb{R}^k$ and $\sigma : \mathbb{R}^k \times U \to \mathbb{R}^{k\times d}$ be given continuous functions. We consider a system whose state $Y(t) \in \mathbb{R}^k$ follows the dynamics

$$dY(t) = b(Y(t), u(Y(t))dt + \sigma(Y(t), u(t))dB(t)\ ; \qquad \tau_j \le t < \tau_{j+1} \tag{59}$$

$$Y(\tau_{j+1}) = \Gamma(Y(\tau_{j+1}^-), \zeta_{j+1})\ ; \qquad j = 0, 1, 2, \ldots \tag{60}$$

where $u(t)$ is a control process with values in U and $v = (\tau_1, \tau_2, \ldots; \zeta_1, \zeta_2, \ldots) \in \mathcal{V}$ is an impulse control and $\tau_0 = 0$. Here $\tau_1 < \tau_2 < \cdots$ are $\mathcal{F}_t$-stopping times (intervention times), $\zeta_j, j \ge 1$ are $\mathcal{F}_{\tau_j}$-measurable random variables representing the corresponding impulses, $\zeta_j \in Z \subset \mathbb{R}^\ell$, $Z = Z(y)$ is a given set which may depend on y. The result of giving an impulse ζ when the state of the system is y is that the state jumps immediately from $Y(t^-) = y$ to $Y(t) = \Gamma(y, \zeta)$, where $\Gamma : \mathbb{R}^k \times Z \to \mathbb{R}^k$ is a given function.

Let $\mathcal{W}$ be the set of *admissible* combined controls $w = (u, v)$ such that a unique strong solution $Y^{(w)}(t)$ of (59), (60) exists and $\lim_{j\to\infty} \tau_j = \infty$ a.s.

Let S be an open set of $\mathbb{R}^k$ and define $T = \inf\{t > 0; Y^{(w)}(t) \notin S\}$. Let f be a *profit/utility rate* function and g a *bequest function*. Moreover, suppose the *profit/utility* of performing an intervention with impulse $\zeta \in Z$ when the system is in state y is $K(y, \zeta)$, where $K : S \times Z \to \mathbb{R}$ is a given function. We assume that for all $y \in \mathbb{R}^k,\ w \in \mathcal{W}$,

$$E^y\left[\int_0^T |f(Y^{(w)}(t), u(t))|dt\right] < \infty, \quad E^y\left[|g(Y^{(w)}(T))\right] < \infty,$$

$$E^y\left[\sum_{\tau_j < T} |K(Y^{(w)}(\tau_j^-), \zeta_j)|\right] < \infty.$$

The performance is given by

$$\begin{aligned} J^{(w)}(y) = E^y\Big[\int_0^T e^{-\lambda t} f(Y(t), u(t))dt + e^{-\lambda T} g(Y(T)) \\ + \sum_j e^{-\lambda \tau_j} K(Y(\tau_j^-), \zeta_j)\Big] \end{aligned} \tag{61}$$

and we want to find the value function Φ defined by

$$\Phi(y) = \sup_{w\in\mathcal{W}} J^{(w)}(y)\ . \tag{62}$$

The function Φ is associated to the HJBQVI

$$\max\left(\sup_{\alpha\in U}\{L^\alpha\Phi(y)+f(y,\alpha)\},\mathbb{M}\Phi(y)-\Phi(y)\right)=0\ ,\qquad y\in S \tag{63}$$

with boundary values

$$\Phi(y)=g(y)\ ;\qquad y\in\partial S \tag{64}$$

where

$$L^\alpha\Phi(y)=\sum_{i=1}^k b_i(y,\alpha)\frac{\partial\Phi}{\partial y_i}+\tfrac{1}{2}\sum_{i,j=1}^k\left(\sigma\sigma^T\right)_{ij}(y,\alpha)\frac{\partial^2\Phi}{\partial y_i\partial y_j}-\lambda\Phi \tag{65}$$

and

$$\mathbb{M}\Phi(y)=\sup\left\{\Phi(\Gamma(y,\zeta))+K(y,\zeta);\zeta\in Z,\Gamma(y,\zeta)\in S\right\}\ . \tag{66}$$

We discretize (63) by using a consistent and stable finite difference approximation with spatial step δ. Following the method explained in [3], we obtain the following discrete-time problem in the grid S_δ:

$$\mathcal{O}\Phi_\delta(y)\stackrel{\text{def}}{=}\max\left(\sup_{\alpha\in U}\{L_\delta^\alpha\Phi_\delta(y)+f_\delta^\alpha(y)\},\mathbb{M}_\delta\Phi_\delta(y)-\Phi_\delta(y)\right)=0\ ,\qquad y\in S_\delta \tag{67}$$

where L_δ^α is diagonally dominant and

$$\mathbb{M}_\delta\Phi(y)=\sup\left\{\Phi(\Gamma(y,\zeta))+K(y,\zeta);\zeta\in Z,\Gamma(y,\zeta)\in S_\delta\right\}.$$

Equation (67) can be rewritten as

$$\Phi_\delta(y)=\max\left(\sup_{\alpha\in U}\{\mathcal{L}_\delta^\alpha\Phi_\delta(y)\},\sup_{\zeta\in\mathcal{Z}_\delta}B^\zeta\Phi_\delta(y)\right)\ ,\qquad y\in S_\delta \tag{68}$$

where

$$\begin{aligned}
\mathcal{L}_\delta^\alpha v &= \frac{1}{1+\lambda k}(M_\delta^\alpha v+kf_\delta^\alpha)\\
k &\le \frac{1}{|\lambda+(L_\delta^\alpha)_{ii}|}\\
M_\delta^\alpha &= kL_\delta^\alpha+(1+\lambda k)I\quad\text{(Stochastic matrix)}\\
B^\zeta v(y) &= v(\Gamma(y,\zeta))+K(y,\zeta)\\
\mathcal{Z}_\delta &= \{\zeta\in Z,\Gamma(y,\zeta)\in S_\delta\}.
\end{aligned}$$

The grid S_δ consists is a finite set of N points of $S:S_\delta=\{x_i,i=1,\dots N\}$.
We make the following assumptions:

- There exists an integer function $\sigma:\{1,2,\dots N\}\times\mathcal{Z}\to\{1,2,\dots N\}$ such that for all $\zeta\in\mathcal{Z}$

$$\Gamma(x_i,\zeta)=x_{\sigma(i,\zeta)}\text{ for all }x_i\in S_\delta\quad\text{with}\quad\sigma(i,\zeta)<i$$

- It is not admissible to perform an intervention in $x_1\in S_\delta$.

Problem (68) is of the form (41). Assuptions $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ are satisfied and Problem (68) can thus be solved by using the policy iteration described in section (3.1).

References

1. D.P. Bertsekas. *Dynamic Programming and Optimal control Vol I and II.* Athena Scientific, Belmont Massachusetts, 2001.
2. J-Ph. Chancelier: A policy iteration algorithm for a singular control problem, manuscript.
3. J-Ph. Chancelier, B. Øksendal and A. Sulem, "Combined stochastic control and optimal stopping, and application to numerical approximation of combined stochastic and impulse control", *Stochastic Financial Mathematics, Proc. Steklov Math. Inst., Moscou, Vol. 237, pp 149 –173, editeur A.N. Shiryaev, 2002*
4. H.J. Kushner and P. Dupuis. *Numerical Methods for stochastic Control Problems in Continuous Time*. Springer Verlag, 1992.
5. B. Øksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions"*. Book project, Universitext, Springer Verlag.
6. B. Lapeyre, A. Sulem and D. Talay : *Understanding Numerical Analysis for Financial Models*, Cambridge University Press, to appear
7. B. Øksendal and A. Sulem, "Optimal Consumption and Portfolio with both fixed and proportional transaction costs: ", *SIAM J. Control and Optim.*, 2002, Vol.40, No. 6, pp 1765-1790.
8. M.L Puterman. *Markov Decision Processes: Discrete Stochastics Dynamic Programming*. Probability and Mathematical Statistics: applied probability and statistics section. Wiley, 1994.