Applied Probability Trust (4 March 2019)

A PROBABILISTIC VERIFICATION THEOREM FOR THE FINITE HORIZON TWO-PLAYER ZERO-SUM OPTIMAL SWITCHING GAME IN CONTINUOUS TIME

S. HAMADÈNE,* *Le Mans Université*

R. MARTYR,** *Queen Mary University of London*

J. MORIARTY,** *Queen Mary University of London*

Abstract

In this paper we study continuous-time two-player zero-sum optimal switching games on a finite horizon. Using the theory of doubly reflected backward stochastic differential equations (DRBSDE) with interconnected barriers, we show that this game has a value and an equilibrium in the players' switching controls.

Keywords: optimal switching; optimal switching games; stopping times; optimal stopping problems; optimal stopping games; backward stochastic differential equations

2010 Mathematics Subject Classification: Primary 91A55
Secondary 60G40;91A05;49N25

1. Zero-sum optimal switching game

Optimal switching is a generalisation of optimal stopping which has various applications in economics and mathematical finance. It consists of one or more agents and a system which they control by successively switching the system's operational mode according to a discrete set of choices. A typical example is the economic valuation of a power plant which manages its fuel mix for electricity production. One can also think of a trader who changes the composition of her portfolio according to the returns of the assets. There are several works on optimal switching problems in continuous time, and a survey of the literature identifies two main approaches used to analyse these problems: an analytical approach using partial differential equations (PDEs) and a probabilistic one.

Methods based on PDEs and associated variational inequalities appeared as early as the 1970s, under the topic of impulsive control for diffusion processes (see [28] and the references therein). A viscosity solutions approach to this type of PDE appeared in the late 1980s to early 1990s (for instance, [31]) and is still the topic of active research [23]. For example, the Hamilton-Jacobi-Bellman equation corresponding to

* Postal address: LMM, Avenue Olivier Messiaen, 72085 Le Mans, Cedex 9, France
** Postal address: School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK
* Email address: said.hamadene@univ-lemans.fr
** Email address: r.martyr@qmul.ac.uk; j.moriarty@qmul.ac.uk

the optimal switching problem for a diffusion process $(X_t)_{t\geq 0}$ on a finite time horizon $[0,T]$ is the following system of PDEs with obstacles depending on the solution $(v^i)_{i\in\Gamma^1}$, with $\Gamma^1 := \{1,\dots,m_1\}$: $\forall i \in \Gamma^1$

$$\begin{cases} \min\Big\{ v^i(s,x) - \max_{k\in\Gamma^1-\{i\}}[v^k(s,x) - g^{i,k}(s,x)]\,, \\ \qquad\qquad\qquad\qquad \big(-\partial_s - \mathrm{L}^X\big)(v^i)(s,x) - f^i(s,x)\Big\} = 0\,; \\ v^i(T,x) = h^i(x), \end{cases} \tag{1}$$

where: i) Γ^1 is the set of available switching modes; ii) $g^{i,k}$ is the function which gives the cost of switching from mode i to mode k; iii) f^i and h^i are the functions which give respectively the instantaneous and terminal yield of the system when it is working in mode i; and iv) L^X is the infinitesimal generator associated with the diffusion process X. If a sufficiently regular solution of the system (1) exists, then $v^i(s,x)$ is nothing else but the optimal profit that can be generated by switching, with initial conditions i for the system's mode and x for the process X at time $s\in[0,T]$.

Probabilistic solution methods for optimal switching problems have been investigated since the 1970s and 1980s in various degrees of generality (see [4, 26, 28, 34, 35] for instance), and most of the recent research in this area has been a combination of the martingale approach via Snell envelopes ([12, 24]) and the theory of backward stochastic differential equations (BSDE) ([7, 14, 20]). The latter methods lead to the study of the following system of reflected BSDEs with lower interconnected obstacles: find adapted stochastic processes $(Y^i, Z^i, K^i)_{i\in\Gamma^1}$, with K^i an increasing process, satisfying: $\forall i\in\Gamma^1$ and $s\in[0,T]$,

$$\begin{cases} Y^i_s = h^i + \int_s^T f^i_t dt - \int_s^T Z^i_t dB_t + K^i_T - K^i_s\ ; \\ Y^i_s \geq L^i_s(\boldsymbol{Y}) := \max_{k\in\Gamma^1-\{i\}}\{Y^k_s - g^{i,k}_s\}; \\ \int_s^T (Y^i_t - L^i_t(\boldsymbol{Y}))dK^i_t = 0, \end{cases} \tag{2}$$

where h^i, f^i, and $g^{i,k}$ play the same role as above, but here they are not formulated as functions: instead, h^i is a random variable and f^i, $g^{i,k}$ are stochastic processes. Similar to the PDE system (1), the BSDE system (2) can concisely represent solutions to the optimal switching problem and is investigated in several papers including [12, 20, 7, 17, 19]. For any $i\in\Gamma^1$, Y^i_s is then the system's value (that is, its optimal yield in the sense of expectation) when it starts from mode i at time s. On the other hand, this system (2) allows for the construction of the optimal strategy (see e.g. [12, 20, 7, 17, 19] for more details). All of the aforementioned references are concerned with single-person optimisation problems. Multiple-person optimal switching problems in a continuous-time stochastic setting, the topic under which the present work falls, have been studied less frequently in the literature (there is related work for deterministic systems such as [33, 32]). In the two-player zero-sum game of optimal switching, which is the setting of the present work, there are works such as [30, 21, 11, 18] and related studies for impulse control games [29, 10].

Let the finite, discrete sets $\Gamma^k = \{1,\dots,m_k\}$, $k\in\{1,2\}$ represent the operating modes that player k can choose. Letting $\Gamma = \Gamma^1\times\Gamma^2$ denote the product space of operating modes $\gamma = (\gamma^{(1)},\gamma^{(2)})$, having cardinality $|\Gamma| = m = m_1\times m_2$, the following costs and rewards are taken into account:

- For $(i,j)\in\Gamma$, $f^{i,j}$ defines a running reward paid by player 2 to player 1 and $h^{i,j}$ a terminal reward paid by player 2 to player 1, when player 1's (resp. player 2's) active mode is i (resp. j).
- For $i_1,i_2\in\Gamma^1$, $\hat{g}^{i_1,i_2}$ defines a non-negative payment from player 1 to player 2 when the former switches from mode i_1 to i_2.
- For $j_1,j_2\in\Gamma^2$, $\check{g}^{j_1,j_2}$ defines a non-negative payment from player 2 to player 1 when the former switches from mode j_1 to j_2.

For all $(i,j)\in\Gamma$ and $t\in[0,T]$ we set $\hat{g}^{i,i}_t=\check{g}^{j,j}_t=0$.

From the probabilistic point of view, since there are two players and their advantages are antagonistic, the zero-sum switching game leads to the study of a system of reflected BSDEs with inter-connected bilateral obstacles which is an extension of (2): For each $(i,j)\in\Gamma$ and $s\in[0,T]$ we should have:

$$\begin{cases} Y^{i,j}_s=h^{i,j}+\int_s^T f^{i,j}_t dt+K^{i,j}_T-K^{i,j}_s-\int_s^T Z^{i,j}_t dB_t\,; \\ Y^{i,j}_s\le U^{i,j}_s(\boldsymbol{Y}) \text{ and } Y^{i,j}_s\ge L^{i,j}_s(\boldsymbol{Y})\,; \\ \int_s^T (Y^{i,j}_t-U^{i,j}_t(\boldsymbol{Y}))dK^{i,j,-}_t=\int_s^T (L^{i,j}_t(\boldsymbol{Y})-Y^{i,j}_t)dK^{i,j,+}_t=0, \end{cases} \tag{3}$$

where: i) $L^{i,j}_s(\boldsymbol{Y}):=\max_{k\in\Gamma^1-\{i\}}\{Y^{k,j}_s-\hat{g}^{i,k}_s\}$; ii) $U^{i,j}_s(\boldsymbol{Y}):=\min_{\ell\in\Gamma^2-\{j\}}\{Y^{i,\ell}_s+\check{g}^{j,\ell}_s\}$; and iii) $K^{i,j}=K^{i,j,+}-K^{i,j,-}$ ($K^{i,j,\pm}$ are increasing processes). While this system is formulated precisely in Section 2, an intuitive interpretation is provided here.

Since the game is zero-sum we suppose that player 1 is the maximiser while player 2 is the minimiser. A system of processes $(Y^{i,j},Z^{i,j},K^{i,j})_{(i,j)\in\Gamma}$ is sought which may be understood as follows. Firstly, $Y^{i,j}_t$ is the value of the game played on the time interval $[t,T]$ when the players begin at time t in mode (i,j). Secondly, $Z^{i,j}_t$ is the volatility of $Y^{i,j}_t$. Thirdly the term $K^{i,j}_t-K^{i,j}_T$ captures, in the sense of expectation, the loss of yield which would be caused by remaining in the state (i,j) over the time interval $[t,T]$. Set $\boldsymbol{Y}=(Y^{i,j})_{(i,j)\in\Gamma}$. Since player 1 may switch mode at any time, the process $L^{i,j}(\boldsymbol{Y})$ provides a lower bound for $Y^{i,j}$. Further, consider any time interval $[t_1,t_2]\subset[0,T]$ in which $Y^{i,j}$ strictly exceeds this bound: then since there is no incentive for player 1 to switch from mode i to any other mode, $K^{i,j}$ should not increase on $[t_1,t_2]$. Symmetric arguments apply to player 2. Therefore the processes $(Y^{i,j},Z^{i,j},K^{i,j})_{(i,j)\in\Gamma}$ should verify (3).

System (3) is studied in rather few papers including [21, 11, 18]. In [21], the authors have shown that a solution exists when $\check{g}^{j,\ell}$ and $\hat{g}^{i,k}$ are constant. It is also studied in [11] in a Markovian framework where, by combining techniques of PDEs and ones of backward equations, the authors proved that system (3) has a unique solution (see Theorem 2). Finally, more recently there is a paper by Hamadène-Mu [18], where it is studied in a more general framework (see Theorem 1) and the authors have shown the existence of a solution when the switching costs $\check{g}^{j,\ell}$ and $\hat{g}^{i,k}$ are only Itô processes.

Concerning system (3), it is expected that $Y^{i,j}_s$ is the value of the zero-sum switching game when the system starts from $(i,j)\in\Gamma$ in the same way as Y^i_s of system (2) is the value of the control problem when the system starts from $i\in\Gamma^1$. However this fact is not obvious at all and proved only in some specific cases. Actually in [11], the authors have shown when $f^{i,j}$ and $h^{i,j}$ are separated in the sense that for any $(i,j)\in\Gamma$, $f^{i,j}=\hat{f}^i+\check{f}^j$ and $h^{i,j}=\hat{h}^i+\check{h}^j$, then $Y^{i,j}_s$ is the value of the zero-sum switching game. This latter assumption allows the game to be decomposed into two optimal switching problems which are simpler and decoupled, but limits the potential for practical applications. The case of arbitrary $f^{i,j}$, $\check{g}^{j,\ell}$, $\hat{g}^{i,k}$ and $h^{i,j}$ is still open and this is the main objective of the present work. We prove *without the separability assumption*, and this is the novelty of this paper, that a solution of system (3) provides the game's value. Furthermore, we derive equilibrium strategies for the players which adapt to the opponent's switching decisions and provide robust performance guarantees.

This paper is organised as follows. In Section 2, we introduce the zero-sum switching game and provide some results related to the existence of a solution of system (3). In Section 3, we show the main result, that $Y^{i,j}$ coincides with the value of the zero-sum game. Moreover, we provide results on the existence of optimal strategies in the game. For completeness, we also interpret our findings in the diffusion framework.

2. Probabilistic setup and notation

We follow closely the setup in [11], working on a finite horizon $[0,T]$ and filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ where $\mathbb{F} = (\mathcal{F}_t)_{0 \le t \le T}$ is the usual completion of the natural filtration of $B = (B_t)_{0 \le t \le T}$, a d-dimensional standard Brownian motion, where $d \ge 1$. Since our method of proof does not depend on the dimension d, without loss of generality we assume henceforth that $d = 1$.

- Let $\mathcal{T}$ be the set of $\mathbb{F}$-stopping times bounded above by T, and for a given $\nu \in \mathcal{T}$, $\mathcal{T}_\nu$ the set of all $\tau \in \mathcal{T}$ satisfying $\tau \ge \nu$ a.s.
- For any sub-σ-algebra $\mathcal{F}^o$ of $\mathcal{F}$, let $L^p(\mathcal{F}^o)$, $1 \le p < \infty$, denote the set of p-integrable $\mathcal{F}^o$-measurable random variables, and set $L^p := L^p(\mathcal{F})$.
- Let $\mathcal{H}^2$ be the set of $\mathbb{F}$-progressively measurable processes $w = (w_t)_{0 \le t \le T}$ which are $dt \otimes d\mathbb{P}$-square integrable.
- Let $\mathcal{S}^2$ be the set of $\mathbb{F}$-adapted processes $w = (w_t)_{0 \le t \le T}$ with paths that are right-continuous with left limits satisfying,
$$\sup_{0 \le t \le T} |w_t| \in L^2.$$
Let $\mathcal{S}^2_c \subset \mathcal{S}^2$ denote the subset of processes $w \in \mathcal{S}^2$ with continuous paths.
- Let $\mathcal{K}^2$ denote the set of $\mathbb{F}$-adapted right-continuous with left limits processes K of finite variation satisfying $K_0 = 0$ and,
$$\int_0^T |dK_t| \in L^2,$$
where $|dK_t(\omega)|$ is the total variation measure on $[0,T]$. Let $\mathcal{K}^2_c$ denote the subset of processes $K \in \mathcal{K}^2$ with continuous paths.

Definition 1. Let Y be a right-continuous with left limits semi-martingale having decomposition $Y_t = Y_0 + M_t + K_t$ where M is a local martingale, K has finite variation, and $M_0 = K_0 = 0$. Note that M is continuous due to the choice of filtration $\mathbb{F}$ (see Lemma 14.5.2 of [9]). We say that Y is square-integrable and write $Y \in \mathcal{W}^2$ if
$$Y_0 \in L^2, M \in \mathcal{S}^2 \text{ and } K \in \mathcal{K}^2.$$
If Y is continuous then we write $Y \in \mathcal{W}^2_c$.

Let $\mathcal{S}_c^{2,m}$ denote the m-product of $\mathcal{S}_c^2$. Similarly we define $\mathcal{H}^{2,m}$, $L^{2,m}$, $\mathcal{S}^{2,m}$, $\mathcal{K}^{2,m}$, $\ldots$, for the m-products of the spaces $\mathcal{H}^2$, L^2, $\mathcal{S}^2$, $\mathcal{K}^2$, and so on.

2.1. Costs, rewards, switching controls and solutions of system (3)

Let $\Gamma^k = \{1, \ldots, m_k\}$, $k \in \{1, 2\}$, be a finite, discrete set representing the operating modes that player k can choose. Let $\Gamma = \Gamma^1 \times \Gamma^2$ denote the product space of operating modes $\gamma = (\gamma^{(1)}, \gamma^{(2)})$, having cardinality $|\Gamma| = m = m_1 \times m_2$. Next for $(i,j) \in \Gamma$, $i_1, i_2 \in \Gamma^1$ and $j_1, j_2 \in \Gamma^2$, let $f^{i,j}$ be a process of $\mathcal{H}^2$, $\hat{g}^{i_1,i_2}$, $\check{g}^{j_1,j_2}$ are non-negative processes of $\mathcal{S}_c^2$ and finally $h^{i,j}$ a random variable belonging to $L^2(\mathcal{F}_T)$. We also assume that for all $(i,j) \in \Gamma$ and $t \in [0,T]$, $\hat{g}_t^{i,i} = \check{g}_t^{j,j} = 0$. As a matter of convention, for players 1 and 2 we use the respective "hat" and "check" notation to distinguish objects that are similar otherwise.

Definition 2. For $N \geq 2$ a *loop* in Γ of length $N-1$ is a sequence $\{(i_1, j_1), \ldots, (i_N, j_N)\}$ of elements in Γ with $N-1$ distinct members such that $(i_N, j_N) = (i_1, j_1)$ and either $i_{q+1} = i_q$ or $j_{q+1} = j_q$ for any $q = 1, \ldots, N-1$.

We now present conditions under which a solution for system (3) exists. To begin with, throughout this paper we make the following assumption.

Assumption 1. *We impose the following conditions on the switching costs:*

1. *Consistency:*

   (a) For all sequences $\{i_1, i_2, i_3\} \in \Gamma^1$ and $\{j_1, j_2, j_3\} \in \Gamma^2$ with $i_1 \neq i_2$, $i_2 \neq i_3$ and $j_1 \neq j_2$, $j_2 \neq j_3$, we have for all $t \in [0,T]$, $\mathbb{P}$-a.s.,

   $$\hat{g}_t^{i_1,i_3} < \hat{g}_t^{i_1,i_2} + \hat{g}_t^{i_2,i_3} \quad \text{and} \quad \check{g}_t^{j_1,j_3} < \check{g}_t^{j_1,j_2} + \check{g}_t^{j_2,j_3} \quad . \tag{4}$$

   (b) For all $(i,j) \in \Gamma$ we have, $\mathbb{P}$-a.s.,

   $$\max_{i_1 \in \Gamma^1 - \{i\}} \{h^{i_1,j} - \hat{g}_T^{i,i_1}\} \leq h^{i,j} \leq \min_{j_1 \in \Gamma^2 - \{j\}} \{h^{i,j_1} + \check{g}_T^{j,j_1}\}. \tag{5}$$

2. *Non-free loop property: For any loop $\{(i_1, j_1), \ldots, (i_N, j_N)\}$ in Γ we have for all $t \in [0,T]$,*

   $$\sum_{q=1}^{N-1} \varphi_t^{q,q+1} \neq 0 \;\; \mathbb{P}\text{-}a.s., \tag{6}$$

   where $\varphi_t^{q,q+1} = -\hat{g}_t^{i_q,i_{q+1}} \mathbf{1}_{\{i_q \neq i_{q+1}\}} + \check{g}_t^{j_q,j_{q+1}} \mathbf{1}_{\{j_q \neq j_{q+1}\}}$.

In Hamadène-Mu [18], the following result is obtained:

Theorem 1. *Suppose that:*

i) Assumption 1 holds;
ii) For any $i_1, i_2 \in \Gamma^1$ and $j_1, j_2 \in \Gamma^2$, $\hat{g}^{i_1,i_2}$ and $\check{g}^{j_1,j_2}$ are Itô processes: for any $0 \leq t \leq T$,

$$\hat{g}^{i_1,i_2}(t) = \hat{g}^{i_1,i_2}(0) + \int_0^t \hat{b}_s^{i_1,i_2} ds + \int_0^t \hat{\phi}_s^{i_1,i_2} dB_s$$

and

$$\check{g}^{j_1,j_2}(t) = \check{g}^{j_1,j_2}(0) + \int_0^t \check{b}_s^{j_1,j_2} ds + \int_0^t \check{\phi}_s^{j_1,j_2} dB_s$$

where $\hat{b}^{i_1,i_2}$, $\check{b}^{j_1,j_2}$, $\hat{\phi}^{i_1,i_2}$ and $\check{\phi}^{j_1,j_2}$ belong to $\mathcal{H}^2$ with $\mathbb{E}[\sup_{s\leq T}(|\hat{b}_s^{i_1,i_2}|+|\check{b}_s^{j_1,j_2}|)^2]<\infty$ moreover.

Then there exists an m-tuple of processes $(Y^{i,j},Z^{i,j},K^{i,j,+},K^{i,j,-})_{(i,j)\in\Gamma}$, with $K^{i,j,\pm}$ increasing, such that $Y^{i,j}\in\mathcal{S}_c^2$, $Z^{i,j}\in\mathcal{H}^2$, $K^{i,j,\pm}\in\mathcal{K}^2$ and $(Y^{i,j},Z^{i,j},K^{i,j,+}-K^{i,j,-})_{(i,j)\in\Gamma}$ satisfies (3).

Remark 1. Under the consistency condition (4), it is more economical for the players to switch directly from one state to another, rather than doing so via an intermediary third state. If these conditions are not satisfied then, mathematically, there can be issues related to the well-posedness of the system (3). Regarding condition (5), note that it is a contradiction to require simultaneously that Y^{ij} is continuous on $[0,T]$, $Y_T^{ij}=h^{i,j}$ and $\max_{i_1\in\Gamma^1-\{i\}}(Y_s^{i_1,j}-\hat{g}_s^{i,i_1})\leq Y_s^{i,j}\leq\min_{j_1\in\Gamma^2-\{j\}}(Y_s^{i,j}+\check{g}_s^{j,j_1})$ for any $s<T$, without assuming (5). If this latter condition is not satisfied, $Y^{i,j}$ may not be continuous and then the pair (α^*,β^*) of (20) may not exist.

We now highlight a result in [11] on the existence of a solution to system (3) when the randomness is modelled by a diffusion process. Let $X^{s,x}$ be the solution to the following stochastic differential equation with initial condition $(s,x)\in[0,T]\times\mathbb{R}^k$:

$$\begin{cases}\forall t\in[s,T],\;\; X_t^{s,x}=x+\displaystyle\int_s^t b(r,X_r^{s,x})dr+\int_s^t\phi(r,X_r^{s,x})dB_r\,;\\ X_r^{s,x}=x,\;\; r\in[0,s].\end{cases}\tag{7}$$

where the functions b and ϕ are both continuous with values in $\mathbb{R}^k$ and $\mathbb{R}^{k\times 1}$ respectively, Lipschitz with respect to x uniformly in t. Next assume that

$$f_t^{i,j}=\bar{f}^{i,j}(t,X_t^{s,x}),\; h^{i,j}=\bar{h}^{i,j}(X_T^{s,x}),\; \hat{g}_t^{i,k}=\hat{\bar{g}}^{i,k}(t,X_t^{s,x}),\; \check{g}_t^{j,\ell}=\check{\bar{g}}^{j,\ell}(t,X_t^{s,x}).\tag{8}$$

In this context, Assumption 1 is derived from the following structural conditions on the functions $\bar{f}^{i,j}$, $\bar{h}^{i,j}$, $\hat{\bar{g}}^{i,k}$ and $\check{\bar{g}}^{j,\ell}$:

Assumption 2.

1. *Non-negativity:* $\min\limits_{i_1\in\Gamma^1}\hat{\bar{g}}^{i,i_1}\geq 0$ *and* $\min\limits_{j_1\in\Gamma^2}\check{\bar{g}}^{j,j_1}\geq 0$ *for all* $i\in\Gamma^1$, $j\in\Gamma^2$.

2. *Consistency:*

   (a) *For all sequences $\{i_1,i_2,i_3\}\in\Gamma^1$ and $\{j_1,j_2,j_3\}\in\Gamma^2$ with $i_1\neq i_2$, $i_2\neq i_3$ and $j_1\neq j_2$, $j_2\neq j_3$, we have for all t,x,*

$$\begin{cases}\hat{\bar{g}}^{i_1,i_3}(t,x)<\hat{\bar{g}}^{i_1,i_2}(t,x)+\hat{\bar{g}}^{i_2,i_3}(t,x),\\ \check{\bar{g}}^{j_1,j_3}(t,x)<\check{\bar{g}}^{j_1,j_2}(t,x)+\check{\bar{g}}^{j_2,j_3}(t,x).\end{cases}\tag{9}$$

   (b) *For all $(i,j)\in\Gamma$ we have, for all x,*

$$\begin{aligned}\max_{i_1\in\Gamma^1-\{i\}}\{\bar{h}^{i_1,j}(x)-\hat{\bar{g}}^{i,i_1}(T,x)\}&\leq\bar{h}^{i,j}(x)\\ &\leq\min_{j_1\in\Gamma^2-\{j\}}\{\bar{h}^{i,j_1}(x)+\check{\bar{g}}^{j,j_1}(T,x)\}\end{aligned}\tag{10}$$

3. *Non-free loop property: For any loop* $\{(i_1, j_1), \ldots, (i_N, j_N)\}$ *in* Γ *we have for all* t, x,
$$\sum_{q=1}^{N-1} \bar{\varphi}^{q,q+1}(t,x) \neq 0, \tag{11}$$
where $\bar{\varphi}^{q,q+1}(t,x) = -\hat{\bar{g}}^{i_q,i_{q+1}}(t,x)\mathbf{1}_{\{i_q \neq i_{q+1}\}} + \breve{\bar{g}}^{j_q,j_{q+1}}(t,x)\mathbf{1}_{\{j_q \neq j_{q+1}\}}$.

Theorem 2. (see [11].) *In the Markovian setting* (7)–(8)*, suppose that:*

i) Assumption 2 holds;
ii) The functions $\breve{\bar{g}}^{j,\ell}$*, where* $j, \ell \in \Gamma^2$*, or* $\hat{\bar{g}}^{i,k}$*, where* $i, k \in \Gamma^1$*, are* $\mathcal{C}^{1,2}$ *and their derivatives are of polynomial growth.*

Then there exists an m*-tuple of processes* $(Y^{i,j}, Z^{i,j}, K^{i,j,+}, K^{i,j,-})_{(i,j)\in\Gamma}$ *depending on the initial condition* (s,x) *for* $X^{s,x}$*, such that* $Y^{i,j} \in \mathcal{S}^2_c$*,* $Z^{i,j} \in \mathcal{H}^2$*,* $K^{i,j,\pm} \in \mathcal{K}^2$ *(where* $K^{i,j,\pm}$ *are increasing processes) and* $(Y^{i,j}, Z^{i,j}, K^{i,j,+} - K^{i,j,-})_{(i,j)\in\Gamma}$ *satisfies* (3)*. Moreover, there also exist deterministic continuous functions with polynomial growth* $(v^{i,j})_{(i,j)\in\Gamma}$ *such that for any* (i,j) *and* $t \in [0, T-s]$,
$$Y^{i,j}_{s+t} = v^{i,j}(s+t, X^{s,x}_{s+t})$$
and $(v^{i,j})_{(i,j)\in\Gamma}$ *is the unique solution in viscosity sense of the following Hamilton-Jacobi-Bellman system of PDEs with obstacles: For any* $(i,j) \in \Gamma$ *and* $(s,x) \in [0,T] \times \mathbb{R}^k$,
$$\begin{cases} \min\Big\{ v^{i,j}(s,x) - L^{i,j}(\boldsymbol{v})(s,x)\,,\, \max\big\{ v^{i,j}(s,x) - U^{i,j}(\boldsymbol{v})(s,x), \\ \qquad\qquad\qquad\qquad \big(-\partial_s - L^X\big)(v^{i,j})(s,x) - f^{i,j}(s,x)\big\}\Big\} = 0\,; \\ v^{i,j}(T,x) = h^{i,j}(x), \end{cases} \tag{12}$$
where L^X *is the generator associated with* $X^{s,x}$*; and* $\boldsymbol{v} = (v^{i,j})_{(i,j)\in\Gamma}$*,* $L^{i,j}(\boldsymbol{v}) \coloneqq \max_{k\in\Gamma^1-\{i\}}\{v^{k,j} - \hat{g}^{i,k}\}$ *and* $U^{i,j}(\boldsymbol{v}) \coloneqq \min_{\ell\in\Gamma^2-\{j\}}\{v^{i,\ell} + \breve{g}^{j,\ell}\}$.

2.1.1. *Individual switching controls and strategies.*

Definition 3. (*Switching controls.*) A *control for player* 1 *is a sequence* $\alpha = (\sigma_n, \xi_n)_{n\geq 0}$ *such that,*

1. *for all* $n \geq 0$*,* $\sigma_n \in \mathcal{T}$ *and is such that* $\sigma_n \leq \sigma_{n+1}$*,* $\mathbb{P}$*-a.s., and* $\mathbb{P}(\{\sigma_n < T\ \forall n \geq 0\}) = 0$*;*
2. *for all* $n \geq 0$*,* ξ_n *is an* $\mathcal{F}_{\sigma_n}$*-measurable* Γ^1*-valued random variable;*
3. *for* $n \geq 1$*, on* $\{\sigma_n < T\}$ *we have* $\sigma_n < \sigma_{n+1}$ *and* $\xi_n \neq \xi_{n-1}$*, while on* $\{\sigma_n = T\}$ *we have* $\xi_n = \xi_{n-1}$*.*

Let A *denote the set of controls for player* 1*. The set* B *of controls* $\beta = (\tau_n, \zeta_n)_{n\geq 0}$ *for player* 2*, where the* ζ_n *are* Γ^2*-valued, is defined analogously. Denoting by* C^α_N *the cost of the first* $N \geq 1$ *switches,*
$$C^\alpha_N \coloneqq \sum_{n=1}^{N} \hat{g}^{\xi_{n-1},\xi_n}_{\sigma_n},$$
note that the limit $\lim\limits_{N\to\infty} C^\alpha_N$ *is well defined.*

Remark 2. Under the consistency condition (5) of Assumption 1, which has also been used previously in papers such as [11, 21], switching at time T provides no additional benefit for either player. Consequently, it is without loss of generality that the definition of controls excludes switching at time T. Furthermore, as Remark 1 explains, this consistency condition ensures well-posedness of the system (3).

Definition 4. A control $\alpha \in \mathsf{A}$ for player 1 is said to be *square-integrable if,*

$$\lim_{N \to \infty} C_N^{\alpha} \in L^2.$$

Let $\mathcal{A}$ denote the set of such controls. Similarly, the set $\mathcal{B}$ of square-integrable controls for player 2 consists of those $\beta \in \mathsf{B}$ satisfying,

$$\lim_{N \to \infty} C_N^{\beta} \in L^2,$$

where

$$C_N^{\beta} \coloneqq \sum_{n=1}^{N} \breve{g}_{\tau_n}^{\zeta_{n-1},\zeta_n}.$$

Definition 5. (*Non-anticipative switching strategies.*) Let $s \in [0,T]$ and $\nu \in \mathcal{T}_s$. Two controls $\alpha^1, \alpha^2 \in \mathsf{A}$ with $\alpha^1 = \left(\sigma_n^1, \xi_n^1\right)_{n \geq 0}$ and $\alpha^2 = \left(\sigma_n^2, \xi_n^2\right)_{n \geq 0}$ are said to be *equivalent*, denoting this by $\alpha^1 \equiv \alpha^2$, *on* $[s, \nu]$ if we have a.s.,

$$\xi_0^1 \mathbf{1}_{[\sigma_0^1, \sigma_1^1]}(t) + \sum_{n \geq 1} \xi_n^1 \mathbf{1}_{(\sigma_n^1, \sigma_{n+1}^1]}(t) = \xi_0^2 \mathbf{1}_{[\sigma_0^2, \sigma_1^2]}(t) + \sum_{n \geq 1} \xi_n^2 \mathbf{1}_{(\sigma_n^2, \sigma_{n+1}^2]}(t), \quad s \leq t \leq \nu.$$

A *non-anticipative strategy* for player 1 is a mapping $\overline{\alpha} \colon \mathsf{B} \to \mathsf{A}$ such that:

- *Non-anticipativity:* for any $s \in [0,T]$, $\nu \in \mathcal{T}_s$, and $\beta^1, \beta^2 \in \mathsf{B}$ such that $\beta^1 \equiv \beta^2$ on $[s, \nu]$, we have $\overline{\alpha}(\beta^1) \equiv \overline{\alpha}(\beta^2)$ on $[s, \nu]$.
- *Square-integrability:* for any $\beta \in \mathcal{B}$ we have $\overline{\alpha}(\beta) \in \mathcal{A}$.

In a similar manner we define non-anticipative strategies for player 2. Let $\mathscr{A}$ and $\mathscr{B}$ denote the set of non-anticipative strategies for players 1 and 2 respectively.

Definition 6. For $s \in [0,T]$ and $i \in \Gamma^1$, let A_s^i denote the set of controls $\alpha \in \mathsf{A}$ satisfying $\xi_0 = i$ and $\sigma_0 = s$. Similarly, define B_s^j for $s \in [0,T]$ and $j \in \Gamma^2$. Analogous notation will be used below for other classes of controls, for example square-integrable controls $\mathcal{A}_s^i$, $\mathcal{B}_s^j$, and strategies $\mathscr{A}_s^i$, $\mathscr{B}_s^j$.

2.1.2. *Coupling of controls.* We now define the coupling of two controls $\alpha \in \mathsf{A}$ and $\beta \in \mathsf{B}$ under the following assumption: player 1's switch is implemented first if both players decide to switch at the same instant.

Definition 7. Given controls $\alpha \in \mathsf{A}$ and $\beta \in \mathsf{B}$, define the *coupling* $\gamma(\alpha, \beta) = (\rho_n, \gamma_n)_{n \geq 0}$ where $\rho_n \in \mathcal{T}$ is defined by,

$$\rho_n = \sigma_{r_n} \wedge \tau_{s_n}, \tag{13}$$

with $r_0 = s_0 = 0$, $r_1 = s_1 = 1$ and for $n \geq 2$,

$$r_n = r_{n-1} + \mathbf{1}_{\{\sigma_{r_{n-1}} \leq \tau_{s_{n-1}}\}}, \quad s_n = s_{n-1} + \mathbf{1}_{\{\tau_{s_{n-1}} < \sigma_{r_{n-1}}\}},$$

and γ_n is a Γ-valued random variable such that $\gamma_0 = (\xi_0, \zeta_0)$ and for $n \geq 1$,

$$\gamma_n = \begin{cases} \big(\xi_{r_n}, \gamma^{(2)}_{n-1}\big), & \text{on } \{\sigma_{r_n} \leq \tau_{s_n},\ \sigma_{r_n} < T\} \\ \big(\gamma^{(1)}_{n-1}, \zeta_{s_n}\big), & \text{on } \{\tau_{s_n} < \sigma_{r_n}\} \\ \gamma_{n-1}, & \text{on } \{\tau_{s_n} = \sigma_{r_n} = T\}. \end{cases} \tag{14}$$

Define for all $0 \leq t \leq T$,

$$u_t = \gamma_0 \mathbf{1}_{[\rho_0, \rho_1]}(t) + \sum_{n \geq 1} \gamma_n \mathbf{1}_{(\rho_n, \rho_{n+1}]}(t), \tag{15}$$

where $(\rho_n, \rho_{n+1}] = \emptyset$ on $\{\rho_n = \rho_{n+1}\}$.

Note that the coupling $\gamma(\alpha, \beta) = (\rho_n, \gamma_n)_{n \geq 0}$ of the controls $\alpha \in \mathsf{A}^i_s$ and $\beta \in \mathsf{B}^j_s$ has the following properties:

1. $\rho_0 = s$ and for all $n \geq 0$ we have $\rho_n \in \mathcal{T}$ and $\rho_n \leq \rho_{n+1}$ $\mathbb{P}$-a.s., and $\mathbb{P}(\{\rho_n < T\ \forall n \geq 0\}) = 0$;
2. $\gamma_0 = (i, j)$ and for all $n \geq 0$ the random variable γ_n is $\mathcal{F}_{\rho_n}$-measurable, Γ-valued and $\gamma_{n+1} \neq \gamma_n$ on $\{\rho_{n+1} < T\}$.

Hereafter, we use notation such as f^γ, where γ is an Γ-valued random variable, with the interpretation $f^\gamma = f^{i,j}$ on $\{\gamma = (i, j)\}$. Write $C^{\gamma(\alpha,\beta)}_N$ for the joint cumulative cost of the first N switches,

$$C^{\gamma(\alpha,\beta)}_N = \sum_{n=1}^{N} \Big[\hat{g}^{\gamma^{(1)}_{n-1}, \gamma^{(1)}_n}_{\rho_n} - \check{g}^{\gamma^{(2)}_{n-1}, \gamma^{(2)}_n}_{\rho_n} \Big], \qquad N \geq 1.$$

Definition 8. The coupling $\gamma(\alpha, \beta) = (\rho_n, \gamma_n)_{n \geq 0}$ of the controls $\alpha \in \mathsf{A}^i_s$ and $\beta \in \mathsf{B}^j_s$ is said to be *admissible*, writing $\gamma(\alpha, \beta) \in \mathcal{G}^{i,j}_s$ to indicate this, if $\sup_{N \geq 1} \big| C^{\gamma(\alpha,\beta)}_N \big| \in L^2$.

Note that for every $\alpha \in \mathsf{A}$ and $\beta \in \mathsf{B}$ we have $\lim_{N\to\infty} C^{\gamma(\alpha,\beta)}_N = \lim_{N\to\infty} C^\alpha_N - \lim_{N\to\infty} C^\beta_N$. Using the triangle inequality, we see that every pair of square-integrable controls (α, β), $\alpha \in \mathcal{A}^i_s$ and $\beta \in \mathcal{B}^j_s$, satisfies $\gamma(\alpha, \beta) \in \mathcal{G}^{i,j}_s$.

2.2. The zero-sum switching game

For the zero-sum game we assume that player 1 is the maximiser and define the total reward from its perspective. Letting $(s, i, j) \in [0, T) \times \Gamma$ be the initial state and recalling (15), we have

$$J^{i,j}_s(\gamma(\alpha, \beta)) = \mathbb{E}\left[\int_s^T f^{u_t}_t \, dt - \sum_{n=1}^{\infty} \Big[\hat{g}^{\gamma^{(1)}_{n-1}, \gamma^{(1)}_n}_{\rho_n} - \check{g}^{\gamma^{(2)}_{n-1}, \gamma^{(2)}_n}_{\rho_n} \Big] + h^{u_T} \;\Big|\; \mathcal{F}_s \right], \qquad \alpha \in \mathsf{A}^i_s,\ \ \beta \in \mathsf{B}^j_s. \tag{16}$$

The *lower and upper values* for this game, denoted respectively by $\check{V}^{i,j}_s$ and $\hat{V}^{i,j}_s$, are defined as follows:

$$\begin{cases} \check{V}^{i,j}_s := \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} \operatorname*{ess\,inf}_{\beta \in \mathcal{B}^j_s} J^{i,j}_s(\gamma(\alpha, \beta)) \\ \hat{V}^{i,j}_s := \operatorname*{ess\,inf}_{\beta \in \mathcal{B}^j_s} \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} J^{i,j}_s(\gamma(\alpha, \beta)). \end{cases} \tag{17}$$

Note that $\check{V}^{i,j}_s \leq \hat{V}^{i,j}_s$ a.s.

Definition 9. The game is said to have a *value at* (s,i,j) if

$$\check{V}^{i,j}_s = \hat{V}^{i,j}_s \quad \text{a.s.} \tag{18}$$

The common value $V^{i,j}_s$, when it exists, is referred to as the game's *solution at* (s,i,j). When $s = T$ we formally set $\check{V}^{i,j}_T = \hat{V}^{i,j}_T = h^{i,j}$.

In this paper we construct a pair of controls $(\alpha^*, \beta^*) \in \mathsf{A}^i_s \times \mathsf{B}^j_s$ such that $\gamma(\alpha^*, \beta^*) \in \mathcal{G}^{i,j}_s$ and the game has a value $V^{i,j}_s = J^{i,j}_s(\gamma(\alpha^*, \beta^*))$ (see Theorem 3 below). Our result is obtained by dynamic programming and the connection between doubly reflected backward stochastic differential equations (DRBSDEs) with implicitly defined barriers and zero-sum optimal stopping games. We also prove the existence of optimal non-anticipative strategies $\overline{\alpha^*} \in \mathscr{A}^i_s$ and $\overline{\beta^*} \in \mathscr{B}^j_s$ which are *robust* in the sense that each is a best response to the worst-case opponent.

3. A probabilistic verification theorem for the zero-sum game

Theorem 3 uses the system (3) to prove the existence of a value for the zero-sum game. Recall that $m = |\Gamma|$ is the number of joint operating modes $(i,j) \in \Gamma$. For $(i,j) \in \Gamma$ define the *lower and upper switching operators*, $L^{i,j} \colon \mathcal{S}^{2,m}_c \to \mathcal{S}^2_c$ and $U^{i,j} \colon \mathcal{S}^{2,m}_c \to \mathcal{S}^2_c$ respectively, as follows: for $\boldsymbol{Y} \in \mathcal{S}^{2,m}_c$,

$$\begin{cases} L^{i,j}(\boldsymbol{Y}) = \max\limits_{i_1 \in \Gamma^1 - \{i\}} \{Y^{i_1,j} - \hat{g}^{i,i_1}\}, \\ U^{i,j}(\boldsymbol{Y}) = \min\limits_{j_1 \in \Gamma^2 - \{j\}} \{Y^{i,j_1} + \check{g}^{j,j_1}\}. \end{cases} \tag{19}$$

Let $\boldsymbol{L} \colon \mathcal{S}^{2,m}_c \to \mathcal{S}^{2,m}_c$ and $\boldsymbol{U} \colon \mathcal{S}^{2,m}_c \to \mathcal{S}^{2,m}_c$ be the operators defined, using matrix notation, by $\boldsymbol{L} = (L^{i,j})_{(i,j)\in\Gamma}$ and $\boldsymbol{U} = (U^{i,j})_{(i,j)\in\Gamma}$. The following definition formalises the concept of a solution to (3).

Definition 10. A *solution* to the system of DRBSDEs with terminal value $\boldsymbol{h} \in L^{2,m}(\mathcal{F}_T)$, driver $\boldsymbol{f} \in \mathcal{H}^{2,m}$, and implicit barriers $\boldsymbol{L}$ and $\boldsymbol{U}$, is a triple $(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{K}) \in \mathcal{S}^{2,m}_c \times \mathcal{H}^{2,m} \times \mathcal{K}^{2,m}_c$ such that a.s. for all $(i,j) \in \Gamma$ and all $0 \leq s \leq T$,

$$\begin{aligned} &(i) \quad Y^{i,j}_s = h^{i,j} + \int_s^T f^{i,j}_t dt + K^{i,j}_T - K^{i,j}_s - \int_s^T Z^{i,j}_t dB_t\,; \\ &(ii) \quad Y^{i,j}_s \leq U^{i,j}_s(\boldsymbol{Y}) \text{ and } Y^{i,j}_s \geq L^{i,j}_s(\boldsymbol{Y})\,; \\ &(iii) \quad \int_s^T (Y^{i,j}_t - U^{i,j}_t(\boldsymbol{Y}))dK^{i,j,-}_t = \int_s^T (L^{i,j}_t(\boldsymbol{Y}) - Y^{i,j}_t)dK^{i,j,+}_t = 0, \end{aligned} \tag{3 revisited}$$

where $K^{i,j,+}$ and $K^{i,j,-}$ are the increasing processes in the orthogonal decomposition $K^{i,j} := K^{i,j,+} - K^{i,j,-}$.

Note that for any solution to (3), the stochastic integral $\int_0^t Z^{i,j}_s dB_s$ is well-defined, and is a martingale belonging to $\mathcal{S}^2_c$ (see Chapter 3 of [8]). Since we have provided conditions, besides Assumptions 1 or 2, under which such a solution exists (see Theorems 1 or 2), hereafter we work under Assumption 1 and assume that a solution to (3) exists.

Theorem 3. *Suppose there exists a solution $(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{K})$ to the DRBSDE* (3). *For every initial state $(s, i, j) \in [0, T] \times \Gamma$,*

(i) Existence of value and optimal controls: the switching game has a value with,

$$Y_s^{i,j} = V_s^{i,j} = J_s^{i,j}(\gamma(\alpha^*, \beta^*)) \ \ a.s., \tag{20}$$

where $(\alpha^*, \beta^*) \in \mathsf{A}_s^i \times \mathsf{B}_s^j$ are a pair of controls satisfying $\gamma(\alpha^*, \beta^*) \in \mathcal{G}_s^{i,j}$.

(ii) Existence of optimal strategies: there exist non-anticipative strategies $\overline{\alpha^*} \in \mathscr{A}_s^i$ and $\overline{\beta^*} \in \mathscr{B}_s^j$ that are optimal in the robust sense:

$$\begin{cases} \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} J_s^{i,j}\big(\gamma(\overline{\alpha^*}(\beta), \beta)\big) = \operatorname*{ess\,sup}_{\overline{\alpha} \in \mathscr{A}_s^i} \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} J_s^{i,j}\big(\gamma(\overline{\alpha}(\beta), \beta)\big) \\ \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}_s^i} J_s^{i,j}\big(\gamma(\alpha, \overline{\beta^*}(\alpha))\big) = \operatorname*{ess\,inf}_{\overline{\beta} \in \mathscr{B}_s^j} \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}_s^i} J_s^{i,j}\big(\gamma(\alpha, \overline{\beta}(\alpha))\big) \end{cases}$$

Furthermore, these robust values are equal to the game's value,

$$\operatorname*{ess\,sup}_{\overline{\alpha} \in \mathscr{A}_s^i} \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} J_s^{i,j}\big(\gamma(\overline{\alpha}(\beta), \beta)\big) = V_s^{i,j} = \operatorname*{ess\,inf}_{\overline{\beta} \in \mathscr{B}_s^j} \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}_s^i} J_s^{i,j}\big(\gamma(\alpha, \overline{\beta}(\alpha))\big).$$

This concept of robustness, which is well known in the optimal control and differential games literature [22, 1, 2], is natural in the context of zero-sum games [10].

Remark 3. Since the switching costs are non-negative we get the following type of *Mokobodski's condition*: there exists a system of processes $\boldsymbol{w} = \{w^{i,j}\}_{(i,j)\in\Gamma}$ belonging to $\mathcal{W}_c^{2,m}$ such that for all $(i, j) \in \Gamma$: for all $0 \le t \le T$ a.s.,

$$\max_{i_1 \in \Gamma^1 - \{i\}} \{w_t^{i_1,j} - \hat{g}_t^{i,i_1}\} \le w_t^{i,j} \le \min_{j_1 \in \Gamma^2 - \{j\}} \{w_t^{i,j_1} + \check{g}_t^{j,j_1}\}. \tag{21}$$

Indeed, by taking $\boldsymbol{w}$ to be the m-dimensional null process, $\boldsymbol{w} \equiv \boldsymbol{0}$, it is easily verified that $\boldsymbol{w} \in \mathcal{W}_c^{2,m}$ and (21) holds. Mokobodski's condition (21) is an extension of that typically assumed for single-agent switching problems in a variety of settings [5, 6, 14, 24], or for two-player Dynkin games or DRBSDEs [3, 16, 25, 27, 13], both of which are special, somewhat degenerate, cases of the optimal switching game studied here.

Let us point out that for any solution $(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{K})$ to the DRBSDE (3), $\boldsymbol{Y}$ satisfies Mokobodski's condition (21) and, a posteriori, also belongs to $\mathcal{W}_c^{2,m}$. Condition (21) can therefore be seen as a feasibility check for the inequality constraint (3)–(ii): there exists at least one system of processes $\boldsymbol{Y}$ which satisfies (3)–(ii) within a suitable class of candidates. Actually, we know from the results in [27] that well-posedness of (3) is intricately linked to Mokobodski's condition (21).

3.1. Proof of Theorem 3

The existence of a solution to the DRBSDE (3) is closely related to the existence of both a value and a Nash equilibrium in the following Dynkin game (see for example [16, 13, 15], and also [29] for the relation to impulse control games with delay).

Proposition 1. *Suppose there exists a solution $(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{K})$ to the DRBSDE* (3). *Then for all $(s, i, j) \in [0, T] \times \Gamma$ a.s.:*

(a)

$$Y_s^{i,j} = \operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s} \operatorname*{ess\,sup}_{\sigma\in\mathcal{T}_s} \mathcal{J}_s^{i,j}(\sigma,\tau) = \operatorname*{ess\,sup}_{\sigma\in\mathcal{T}_s} \operatorname*{ess\,inf}_{\tau\in\mathcal{T}_s} \mathcal{J}_s^{i,j}(\sigma,\tau), \tag{22}$$

where,

$$\begin{aligned} \mathcal{J}_s^{i,j}(\sigma,\tau) \coloneqq \mathbb{E}\Big[\int_s^{\sigma\wedge\tau} f_t^{i,j}dt + \mathbf{1}_{\{\tau<\sigma\}}U_\tau^{i,j}(\boldsymbol{Y}) + \mathbf{1}_{\{\sigma\le\tau,\ \sigma<T\}}L_\sigma^{i,j}(\boldsymbol{Y}) \;\Big|\; \mathcal{F}_s\Big] \\ + \mathbb{E}\big[h^{i,j}\mathbf{1}_{\{\sigma=\tau=T\}} \;\big|\; \mathcal{F}_s\big], \end{aligned} \tag{23}$$

and $\boldsymbol{h}$*,* $\boldsymbol{f}$*,* $\boldsymbol{L}$ *and* $\boldsymbol{U}$ *are the data for* (3) *(see Definition 10).*

(b) we have $Y_s^{i,j} = \mathcal{J}_s^{i,j}(\sigma_s^{i,j},\tau_s^{i,j})$ *where* $\sigma_s^{i,j}\in\mathcal{T}_s$ *and* $\tau_s^{i,j}\in\mathcal{T}_s$ *are stopping times defined by,*

$$\begin{cases} \sigma_s^{i,j} = \inf\{s\le t\le T\colon Y_t^{i,j} = L_t^{i,j}(\boldsymbol{Y})\}\wedge T, \\ \tau_s^{i,j} = \inf\{s\le t\le T\colon Y_t^{i,j} = U_t^{i,j}(\boldsymbol{Y})\}\wedge T, \end{cases} \tag{24}$$

and we use the convention that $\inf\emptyset = +\infty$*. Moreover,* $\big(\sigma_s^{i,j},\tau_s^{i,j}\big)$ *is a Nash equilibrium for the Dynkin game,*

$$\mathcal{J}_s^{i,j}(\sigma,\tau_s^{i,j}) \le \mathcal{J}_s^{i,j}(\sigma_s^{i,j},\tau_s^{i,j}) \le \mathcal{J}_s^{i,j}(\sigma_s^{i,j},\tau) \quad \forall\sigma\in\mathcal{T}_s \text{ and } \tau\in\mathcal{T}_s. \tag{25}$$

Proof. Recalling the ordering (3)-(ii), the result follows from Proposition 2.2.1 of [15], for example. □

We will use Proposition 1 and a dynamic programming argument to first establish claim *(i)* of Theorem 3, then obtain *(ii)* as a corollary. Since (20) trivially holds when $s=T$, let $s\in[0,T)$ and $(i,j)\in\Gamma$ be arbitrary. Define a sequence $(\rho_n,\gamma_n)_{n\ge0}$ as follows,

$$\rho_0 = s,\quad \gamma_0 = (i,j) \text{ and for } n\ge 1, \tag{26}$$

$$\rho_n = \sigma_{\rho_{n-1}}^{\gamma_{n-1}} \wedge \tau_{\rho_{n-1}}^{\gamma_{n-1}},\quad \gamma_n = \begin{cases} \big(\mathcal{L}_{\rho_n}^{\gamma_{n-1}}(\boldsymbol{Y}),\gamma_{n-1}^{(2)}\big), & \text{on} \quad \mathcal{M}_n^+ \\ \big(\gamma_{n-1}^{(1)},\mathcal{U}_{\rho_n}^{\gamma_{n-1}}(\boldsymbol{Y})\big), & \text{on} \quad \mathcal{M}_n^- \\ \gamma_{n-1}, & \text{otherwise} \end{cases} \tag{27}$$

where $\sigma_{\rho_{n-1}}^{\gamma_{n-1}}$ and $\tau_{\rho_{n-1}}^{\gamma_{n-1}}$ are defined using (24) above, $\mathcal{L}_{\rho_n}^{\gamma_{n-1}}$ and $\mathcal{U}_{\rho_n}^{\gamma_{n-1}}$ are obtained from the switching selectors,

$$\begin{cases} \mathcal{L}_t^{i,j}(\boldsymbol{Y}) \in \operatorname*{arg\,max}_{i_1\in\Gamma^1-\{i\}}\{Y_t^{i_1,j} - \hat{g}_t^{i,i_1}\}, \\ \mathcal{U}_t^{i,j}(\boldsymbol{Y}) \in \operatorname*{arg\,min}_{j_1\in\Gamma^2-\{j\}}\{Y_t^{i,j_1} + \check{g}_t^{j,j_1}\}, \end{cases} \tag{28}$$

and for $n\ge1$, $\mathcal{M}_n^+$ and $\mathcal{M}_n^-$ are the events,

$$\begin{cases} \mathcal{M}_n^+ = \big\{\sigma_{\rho_{n-1}}^{\gamma_{n-1}} \le \tau_{\rho_{n-1}}^{\gamma_{n-1}},\ \sigma_{\rho_{n-1}}^{\gamma_{n-1}} < T\big\}, \\ \mathcal{M}_n^- = \big\{\tau_{\rho_{n-1}}^{\gamma_{n-1}} < \sigma_{\rho_{n-1}}^{\gamma_{n-1}}\big\}. \end{cases}$$

Lemma 1. *Under the conditions of Theorem 3 we have $\gamma(\alpha^*,\beta^*) \in \mathcal{G}_s^{i,j}$ and $Y_s^{i,j} = J_s^{i,j}(\gamma(\alpha^*,\beta^*))$ a.s., where $\alpha^* = (\sigma_n^*,\xi_n^*)_{n\geq 0}$ and $\beta^* = (\tau_n^*,\zeta_n^*)_{n\geq 0}$ are sequences defined from $(\rho_n,\gamma_n)_{n\geq 0}$ as follows,*

$$\sigma_0^* = \tau_0^* = s, \quad (\xi_0^*,\zeta_0^*) = (i,j) \text{ and for } n \geq 1, \tag{29}$$

$$\begin{cases} \sigma_n^* = \inf\{t \geq \sigma_{n-1}^* \colon u_t^{(1)} \neq \xi_{n-1}^*\} \wedge T, \quad \xi_n^* = u_{\sigma_n^*+}^{(1)}, \\ \tau_n^* = \inf\{t \geq \tau_{n-1}^* \colon u_t^{(2)} \neq \zeta_{n-1}^*\} \wedge T, \quad \zeta_n^* = u_{\tau_n^*+}^{(2)}, \end{cases} \tag{30}$$

where u is defined using (15).

Proof. We begin by establishing that $\alpha^* \in \mathsf{A}_s^i$. The non-free loop property (6) prevents accumulation of the switching times ρ_n^*, in the sense that $\mathbb{P}(\{\rho_n^* < T \ \forall n \geq 0\}) = 0$ (see, for example, [17, pp. 192–193]). Since $\sigma_n^* \geq \rho_n$ for $n \geq 0$, it follows that $\mathbb{P}(\{\sigma_n^* < T \ \forall n \geq 0\}) = 0$. Also, the consistency property (4) ensures that it is not optimal for a single player to switch twice at the same instant, so we have $\sigma_n^* < \sigma_{n+1}^*$ on $\{\sigma_n^* < T\}$ for $n \geq 1$ (see [24] or [17]). By the construction of α^*, noting that $u_{\sigma_n^*+}^{(1)}$ is $\mathcal{F}_{\sigma_n^*}$-measurable since $\mathbb{F}$ is right-continuous, the remaining parts of Definition 3 are satisfied, and $\alpha^* \in \mathsf{A}_s^i$. Similarly $\beta^* \in \mathsf{B}_s^j$.

We now prove that $\gamma(\alpha^*,\beta^*) \in \mathcal{G}_s^{i,j}$ by proceeding in a similar manner to [17]. Using (3)-(i) and (3)-(iii) together with the construction of ρ_1 gives $\mathbb{P}$-a.s.,

$$\begin{aligned} Y_s^{i,j} &= \int_s^{\rho_1} f_t^{i,j} dt + h^{i,j}\mathbf{1}_{\{\rho_1 = T\}} + Y_{\rho_1}^{i,j}\mathbf{1}_{\{\rho_1 < T\}} + \int_s^{\rho_1} dK_t^{i,j,+} - \int_s^{\rho_1} dK_t^{i,j,-} \\ &\quad - \int_s^{\rho_1} Z_t^{i,j} dB_t \\ &= \int_s^{\rho_1} f_t^{i,j} dt + h^{i,j}\mathbf{1}_{\{\rho_1 = T\}} + Y_{\rho_1}^{i,j}\mathbf{1}_{\{\rho_1 < T\}} - \int_s^{\rho_1} Z_t^{i,j} dB_t. \end{aligned}$$

By considering the first switch for either player we have

$$\begin{aligned} Y_s^{i,j} &= \int_s^{\rho_1} f_t^{i,j} dt + h^{i,j}\mathbf{1}_{\{\rho_1 = T\}} + \left(Y_{\sigma_s^{i,j}}^{\gamma_1^{(1)},j} - \hat{g}_{\sigma_s^{i,j}}^{i,\gamma_1^{(1)}}\right)\mathbf{1}_{\{\sigma_s^{i,j} < T\}}\mathbf{1}_{\{\sigma_s^{i,j} \leq \tau_s^{i,j}\}} \\ &\quad + \left(Y_{\tau_s^{i,j}}^{i,\gamma_1^{(2)}} + \check{g}_{\tau_s^{i,j}}^{j,\gamma_1^{(2)}}\right)\mathbf{1}_{\{\tau_s^{i,j} < \sigma_s^{i,j}\}} - \int_s^{\rho_1} Z_t^{i,j} dB_t \\ &= \int_s^{\rho_1} f_t^{u_t} dt + Y_{\rho_1}^{\gamma_1}\mathbf{1}_{\{\rho_1 < T\}} + h^{\gamma_0}\mathbf{1}_{\{\rho_1 = T\}} - \left[\hat{g}_{\rho_1}^{\gamma_0^{(1)},\gamma_1^{(1)}} - \check{g}_{\rho_1}^{\gamma_0^{(2)},\gamma_1^{(2)}}\right] - \int_s^{\rho_1} Z_t^{u_t} dB_t, \end{aligned}$$

(to account for the event $\{\rho_1 = T\}$, recall that $\hat{g}_t^{i,i} = \check{g}_t^{j,j} = 0$). Proceeding iteratively for $n = 1,\ldots,N$ we obtain by substitution

$$\begin{aligned} Y_s^{i,j} &= \int_s^{\rho_N} f_t^{u_t} dt + \sum_{n=1}^{N} h^{\gamma_{n-1}}\mathbf{1}_{\{\rho_n = T,\ \rho_{n-1} < T\}} - \sum_{n=1}^{N}\left[\hat{g}_{\rho_n}^{\gamma_{n-1}^{(1)},\gamma_n^{(1)}} - \check{g}_{\rho_n}^{\gamma_{n-1}^{(2)},\gamma_n^{(2)}}\right] \\ &\quad + Y_{\rho_N}^{\gamma_N}\mathbf{1}_{\{\rho_N < T\}} - \int_s^{\rho_N} Z_t^{u_t} dB_t, \end{aligned} \tag{31}$$

from which we obtain

$$C_N^{\gamma(\alpha^*,\beta^*)} = Y_{\rho_N}^{\gamma_N}\mathbf{1}_{\{\rho_N<T\}} - Y_s^{i,j} + \int_s^{\rho_N} f_t^{u_t}dt + \sum_{n=1}^{N} h^{\gamma_{n-1}}\mathbf{1}_{\{\rho_n=T,\ \rho_{n-1}<T\}} - \int_s^{\rho_N} Z_t^{u_t}dB_t. \tag{32}$$

Let $M^u = (M_t^u)_{s\le t\le T}$ denote the stochastic integral $M_t^u = \int_s^t Z_r^{u_r}dB_r$, which is a well-defined square-integrable martingale on $[s,T]$ [8]. Continuing from (32) we have a.s.,

$$\sup_{N\ge1}\big|C_N^{\gamma(\alpha^*,\beta^*)}\big| \le \int_s^T |f_t^{u_t}|dt + \max_{(i,j)\in\Gamma}|h^{i,j}| + |Y_s^{i,j}| + \max_{(i,j)\in\Gamma}\sup_{s\le t\le T}|Y_s^{i,j}| + \sup_{s\le t\le T}|M_t^u|. \tag{33}$$

The right-hand side of (33) is a square-integrable random variable, thereby proving $\gamma(\alpha^*,\beta^*)\in\mathcal{G}_s^{i,j}$.

It is now straightforward to prove $Y_s^{i,j} = J_s^{i,j}(\gamma(\alpha^*,\beta^*))$ a.s. by taking conditional expectations in (31) then passing to the limit $N\to\infty$, which is justified since $\gamma(\alpha^*,\beta^*)\in\mathcal{G}_s^{i,j}$,

$$\begin{aligned} Y_s^{i,j} &= \mathbb{E}\left[\int_s^T f_t^{u_t}dt + h^{u_T} - \sum_{n=1}^{\infty}\Big[\hat{g}_{\rho_n}^{\gamma_{n-1}^{(1)},\gamma_n^{(1)}} - \check{g}_{\rho_n}^{\gamma_{n-1}^{(2)},\gamma_n^{(2)}}\Big] \,\Big|\, \mathcal{F}_s\right] \\ &= J_s^{i,j}(\gamma(\alpha^*,\beta^*)). \end{aligned} \tag{34}$$

□

For a given $\alpha = (\sigma_n,\xi_n)_{n\ge0}\in\mathsf{A}_s^i$, let $\overline{\beta^*}(\alpha) = (\tau_n,\zeta_n)_{n\ge0}$ be the control for player 2 defined similarly to (29) with the sequence $(\rho_n,\gamma_n)_{n\ge0}$ constructed by,

$$\rho_0 = s,\quad \gamma_0 = (i,j) \text{ and for } n\ge1, \tag{35}$$

$$\rho_n = \sigma_{r_n}\wedge\tau_{\rho_{n-1}}^{\gamma_{n-1}},\quad \gamma_n = \begin{cases}\big(\xi_{r_n},\gamma_{n-1}^{(2)}\big), & \text{on } \mathcal{M}_n^+\\ \big(\gamma_{n-1}^{(1)},\mathcal{U}_{\rho_n}^{\gamma_{n-1}}(\boldsymbol{Y})\big), & \text{on } \mathcal{M}_n^-\\ \gamma_{n-1}, & \text{otherwise}\end{cases} \tag{36}$$

where, reusing earlier notation, $\mathcal{U}_{\rho_n}^{\gamma_{n-1}}$ is obtained from (28), $\{r_n\}_{n\ge0}$ is defined iteratively by $r_0=0$, $r_1=1$ and for $n\ge2$,

$$r_n = r_{n-1} + \mathbf{1}_{\{\sigma_{r_{n-1}}\le\tau_{\rho_{n-2}}^{\gamma_{n-2}}\}},$$

and for $n\ge1$, $\mathcal{M}_n^+$ and $\mathcal{M}_n^-$ are the events,

$$\begin{cases}\mathcal{M}_n^+ = \{\sigma_{r_n}\le\tau_{\rho_{n-1}}^{\gamma_{n-1}},\ \sigma_{r_n}<T\},\\ \mathcal{M}_n^- = \{\tau_{\rho_{n-1}}^{\gamma_{n-1}}<\sigma_{r_n}\}.\end{cases}$$

In an analogous manner using the lower switching selector $\boldsymbol{\mathcal{L}}(\boldsymbol{Y})$ in (28), for each $\beta\in\mathsf{B}_s^j$ we define $\overline{\alpha^*}(\beta)\in\mathsf{A}_s^j$ for player 1. The following lemma points out key properties of $\overline{\alpha^*}$ and $\overline{\beta^*}$ utilised below to finish the proof of Theorem 3.

Lemma 2.

(i) We have $\overline{\alpha^*} \in \mathscr{A}_s^i$ *and* $\overline{\beta^*} \in \mathscr{B}_s^j$.

(ii) We have

$$\operatorname*{ess\,sup}_{\alpha \in \mathcal{A}_s^i} J_s^{i,j}\big(\gamma(\alpha, \overline{\beta^*}(\alpha))\big) = Y_s^{i,j} = \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} J_s^{i,j}\big(\gamma(\overline{\alpha^*}(\beta), \beta)\big). \tag{37}$$

Proof.

Proof of (i): We only show $\overline{\beta^*} \in \mathscr{B}_s^j$ since the proof that $\overline{\alpha^*} \in \mathscr{A}_s^i$ follows by similar arguments. Just as in the proof of Lemma 1, the construction of $\overline{\beta^*}(\alpha)$ together with the non-free loop and consistency properties are sufficient to establish that $\overline{\beta^*}(\alpha) \in \mathsf{B}_s^j$ for each $\alpha \in \mathsf{A}_s^i$. Moreover, $\overline{\beta^*}$ satisfies the non-anticipative property in Definition 5 by construction. Let $\alpha \in \mathcal{A}_s^i$ be given and let $\overline{\beta^*}(\alpha) = \beta = (\tau_n, \zeta_n)_{n \geq 0} \in \mathsf{B}_s^j$. To show that this control is square-integrable we will proceed as in the proof of Lemma 1, to obtain that a.s.,

$$\begin{aligned}
Y_s^{i,j} = & \int_s^{\rho_1} f_t^{i,j} dt + h^{i,j} \mathbf{1}_{\{\rho_1 = T\}} + Y_{\rho_1}^{i,j} \mathbf{1}_{\{\rho_1 < T\}} + \int_s^{\rho_1} dK_t^{i,j,+} - \int_s^{\rho_1} dK_t^{i,j,-} \\
& - \int_s^{\rho_1} Z_t^{i,j} dB_t \\
\geq & \int_s^{\rho_1} f_t^{u_t} dt + h^{i,j} \mathbf{1}_{\{\rho_1 = T\}} - \big[\hat{g}_{\rho_1}^{i,\gamma_1^{(1)}} - \check{g}_{\rho_1}^{j,\gamma_1^{(2)}}\big] + Y_{\rho_1}^{\gamma_1} \mathbf{1}_{\{\rho_1 < T\}} - \int_s^{\rho_1} Z_t^{u_t} dB_t,
\end{aligned}$$

where, in contrast to the proof of Lemma 1, here α is arbitrary and so $\gamma_1^{(1)}$ is not necessarily optimal at time ρ_1. This means the inequality $L_{\rho_1}^{i,j}(\boldsymbol{Y}) \leq Y_{\rho_1}^{i,j}$ must be enforced and the non-negative term $\int_s^{\rho_1} dK_t^{i,j,+}$ cannot be neglected. Proceeding iteratively for $n = 1, \ldots, N$ it follows that

$$\begin{aligned}
Y_s^{i,j} \geq & \int_s^{\rho_N} f_t^{u_t} dt + \sum_{n=1}^{N} h^{\gamma_{n-1}} \mathbf{1}_{\{\rho_n = T,\, \rho_{n-1} < T\}} - \sum_{n=1}^{N} \Big[\hat{g}_{\rho_n}^{\gamma_{n-1}^{(1)}, \gamma_n^{(1)}} - \check{g}_{\rho_n}^{\gamma_{n-1}^{(2)}, \gamma_n^{(2)}}\Big] \\
& + Y_{\rho_N}^{\gamma_N} \mathbf{1}_{\{\rho_N < T\}} - \int_s^{\rho_N} Z_t^{u_t} dB_t,
\end{aligned} \tag{38}$$

from which we obtain

$$\begin{aligned}
\sum_{n=1}^{N} \check{g}_{\rho_n}^{\gamma_{n-1}^{(2)}, \gamma_n^{(2)}} \leq & - \int_s^{\rho_N} f_t^{u_t} dt - \sum_{n=1}^{N} h^{\gamma_{n-1}} \mathbf{1}_{\{\rho_n = T,\, \rho_{n-1} < T\}} + \sum_{n=1}^{N} \hat{g}_{\rho_n}^{\gamma_{n-1}^{(1)}, \gamma_n^{(1)}} \\
& + Y_s^{i,j} - Y_{\rho_N}^{\gamma_N} \mathbf{1}_{\{\rho_N < T\}} + \int_s^{\rho_N} Z_t^{u_t} dB_t.
\end{aligned} \tag{39}$$

We note that $\mathbb{P}(\{\rho_N < T \ \forall N \geq 1\}) = 0$ and the limits as $N \to \infty$ on both sides of (39) are well defined. As the switching costs are non-negative we have

$$0 \leq \sum_{n \geq 1} \check{g}_{\tau_n}^{\zeta_{n-1}, \zeta_n} \leq - \int_s^T f_t^{u_t} dt - h^{u_T} + \sum_{n \geq 1} \hat{g}_{\sigma_n}^{\xi_{n-1}, \xi_n} + Y_s^{i,j} + \int_s^T Z_t^{u_t} dB_t. \tag{40}$$

Since $\alpha \in \mathcal{A}^i_s$, $Y^{i,j} \in \mathcal{S}^2_c$, $h^{i,j} \in L^2(\mathcal{F}_T)$, and $f^{i,j}$, $Z^{i,j}$ belong to $\mathcal{H}^2$ for all $(i,j) \in \Gamma$, the random variable on the right-hand side of (40) belongs to L^2 and we conclude that the control β is square-integrable.

Proof of (ii): We only show the first equality in (37) as the second follows via similar arguments. We proceed by showing that for every $\alpha \in \mathcal{A}^i_s$ we have,

$$Y^{i,j}_s \geq J^{i,j}_s\big(\gamma(\alpha, \overline{\beta^*}(\alpha))\big). \tag{41}$$

Taking conditional expectations in (38) above we get,

$$\begin{aligned} Y^{i,j}_s \geq \mathbb{E}\Bigg[\int_s^{\rho_N} f^{u_t}_t dt + \sum_{n=1}^{N} h^{\gamma_{n-1}} \mathbf{1}_{\{\rho_n = T,\ \rho_{n-1} < T\}} - \sum_{n=1}^{N} \Big[\hat{g}^{\gamma^{(1)}_{n-1}, \gamma^{(1)}_n}_{\rho_n} - \check{g}^{\gamma^{(2)}_{n-1}, \gamma^{(2)}_n}_{\rho_n}\Big] \;\Big|\; \mathcal{F}_s\Bigg] \\ + \mathbb{E}\left[Y^{\gamma_N}_{\rho_N} \mathbf{1}_{\{\rho_N < T\}} | \mathcal{F}_s\right]. \end{aligned} \tag{42}$$

Using *(i)* above we have $\gamma\big(\alpha, \overline{\beta^*}(\alpha)\big) \in \mathcal{G}^{i,j}_s$, so taking the limit $N \to \infty$ in (42) proves the inequality (41).

Next, for each integer $k \geq 0$ let α^*_k denote the truncation of the control α^* from Lemma 1 to the first k switches: $\alpha^*_k = \big(\sigma^*_n, \xi^*_n\big)_{0 \leq n \leq k}$ with (T, ξ^*_k) appended. Then $\alpha^*_k \in \mathcal{A}^i_s$ for each k and $J^{i,j}_s\big(\gamma\big(\alpha^*_k, \overline{\beta^*}(\alpha^*_k)\big)\big) \to J^{i,j}_s\big(\gamma\big(\alpha^*, \overline{\beta^*}(\alpha^*)\big)\big)$ by the non-anticipative properties of $\overline{\beta^*}$ and as $\gamma\big(\alpha^*, \overline{\beta^*}(\alpha^*)\big) \in \mathcal{G}^{i,j}_s$. The claim

$$\operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} J^{i,j}_s\big(\gamma(\alpha, \overline{\beta^*}(\alpha))\big) = Y^{i,j}_s,$$

is then proved by passing to the limit $k \to \infty$ in,

$$J^{i,j}_s\big(\gamma\big(\alpha^*_k, \overline{\beta^*}(\alpha^*_k)\big)\big) \leq \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} J^{i,j}_s\big(\gamma(\alpha, \overline{\beta^*}(\alpha))\big) \leq Y^{i,j}_s,$$

and using Lemma 1. □

Proof of Theorem 3.

Proof of (i): By construction we have $\alpha^* = \overline{\alpha^*}(\beta^*)$ and $\beta^* = \overline{\beta^*}(\alpha^*)$ so that, by Lemma 1,

$$Y^{i,j}_s = J^{i,j}_s\big(\gamma(\alpha^*, \beta^*)\big) = J^{i,j}_s\big(\gamma(\overline{\alpha^*}(\beta^*), \beta^*)\big) = J^{i,j}_s\big(\gamma(\alpha^*, \overline{\beta^*}(\alpha^*))\big), \tag{43}$$

and by Lemma 2,

$$\operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} J^{i,j}_s\big(\gamma(\alpha, \overline{\beta^*}(\alpha))\big) = Y^{i,j}_s = \operatorname*{ess\,inf}_{\beta \in \mathcal{B}^j_s} J^{i,j}_s\big(\gamma(\overline{\alpha^*}(\beta), \beta)\big).$$

Since $\overline{\beta^*}(\alpha) \in \mathcal{B}^j_s$ for every $\alpha \in \mathcal{A}^i_s$ and $\overline{\alpha^*}(\beta) \in \mathcal{A}^i_s$ for every $\beta \in \mathcal{B}^j_s$, almost surely we have,

$$\hat{V}^{i,j}_s := \operatorname*{ess\,inf}_{\beta \in \mathcal{B}^j_s} \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} J^{i,j}_s\big(\gamma(\alpha, \beta)\big) \leq Y^{i,j}_s \leq \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}^i_s} \operatorname*{ess\,inf}_{\beta \in \mathcal{B}^j_s} J^{i,j}_s\big(\gamma(\alpha, \beta)\big) =: \check{V}^{i,j}_s,$$

which completes the proof since $\hat{V}^{i,j}_s \geq \check{V}^{i,j}_s$ a.s.

Proof of (ii): For all $\overline{\alpha} \in \mathscr{A}_s^i$ we have a.s.,

$$\begin{aligned} \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} J_s^{i,j}\big(\gamma(\overline{\alpha}(\beta),\beta)\big) &\leq \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}_s^i} J_s^{i,j}\big(\gamma(\alpha,\beta)\big) \\ &= Y_s^{i,j} = \operatorname*{ess\,inf}_{\beta \in \mathcal{B}_s^j} J_s^{i,j}\big(\gamma(\overline{\alpha^*}(\beta),\beta)\big), \end{aligned}$$

and the corresponding statement for $\overline{\beta^*}$ is proved analogously. Since $\overline{\alpha^*} \in \mathscr{A}_s^i$ and $\overline{\beta^*} \in \mathscr{B}_s^j$ the proof is complete. □

Remark 4. In proving Theorem 3 we established the following. For players 1 and 2 respectively there exist *non-anticipative strategies* $\overline{\alpha^*}$ and $\overline{\beta^*}$ as well as *controls* α^* and β^* which satisfy the following,

- the controls α^*, β^* and non-anticipative strategies $\overline{\alpha^*}$, $\overline{\beta^*}$ are related by $\alpha^* = \overline{\alpha^*}(\beta^*)$ and $\beta^* = \overline{\beta^*}(\alpha^*)$;
- α^* and β^* are jointly admissible;
- when player 2 (the minimiser) uses the non-anticipative strategy $\overline{\beta^*}$, then the use of the control α^* by player 1 (the maximiser) gives the maximum possible value for the switching game over all controls α such that $(\alpha, \overline{\beta^*}(\alpha))$ is jointly admissible, including all square-integrable controls α;
- when player 1 uses the non-anticipative strategy $\overline{\alpha^*}$, then the use of the control β^* by player 2 gives the minimum possible value for the switching game over all controls β such that $(\overline{\alpha^*}(\beta), \beta)$ is jointly admissible, including all square-integrable controls β;
- the strategies $\overline{\alpha^*}$ and $\overline{\beta^*}$ are best responses in the robust sense [22, 1, 2].

Let us emphasise that $\overline{\alpha^*}$ is not necessarily a best response strategy in the sense,

$$J_s^{i,j}\big(\gamma(\overline{\alpha^*}(\beta),\beta)\big) = \operatorname*{ess\,sup}_{\alpha \in \mathcal{A}_s^i} J_s^{i,j}\big(\gamma(\alpha,\beta)\big) \quad \forall \beta \in \mathcal{B}_s^j,$$

and correspondingly for $\overline{\beta^*}$. In the game with initial data (s,i,j), for player 1 we can define a mapping $\overline{\alpha} \colon \mathcal{B}_s^j \to \mathcal{A}_s^i$ such that for each $\beta \in \mathcal{B}_s^j$ a.s.,

$$J_s^{i,j}\big(\gamma(\overline{\alpha}(\beta),\beta)\big) \geq J_s^{i,j}\big(\gamma(\alpha,\beta)\big) \text{ a.s.} \quad \forall \alpha \in \mathcal{A}_s^i,$$

but this mapping is generally not non-anticipative since its output $\overline{\alpha}(\beta)$ can depend on the entire trajectory corresponding to the input β. For example, let $\beta \in \mathcal{B}_s^j$ be a given switching control for player 2. We denote by u^β the stochastic process that indicates player 2's current mode according to β,

$$u_t^\beta = \zeta_0 \mathbf{1}_{[\tau_0,\tau_1]}(t) + \sum_{n \geq 1} \zeta_n \mathbf{1}_{(\tau_n,\tau_{n+1}]}(t), \quad t \in [s,T].$$

For $i \in \Gamma^1$ and $t \in [s,T]$, let $\tilde{f}_t^i(\beta) = f_t^{i,u_t^\beta}$ denote the controllable part of player 1's instantaneous reward when player 2's switching control β is given and fixed. Similarly, let $\tilde{h}_{s,T}^i(\beta) := h^{i,u_T^\beta} + \sum_{n=1}^\infty \check{g}_{\tau_n}^{\zeta_{n-1},\zeta_n} \mathbf{1}_{\{\tau_n \geq s\}}$ denote the controllable part of player 1's terminal reward given player 2's switching control β, and inclusive of the payments received from player 2's switching costs $\breve{g}_{\tau_n}^{\zeta_{n-1},\zeta_n}$ from time s onwards. With these definitions we can evaluate a control $\alpha \in \mathcal{A}_s^i$ for player 1 according to,

$$\tilde{J}_s^i(\alpha;\beta) = \mathbb{E}\left[\int_s^T \tilde{f}_t^{u_t^\alpha}(\beta)\,dt - \sum_{n=1}^{\infty} \hat{g}_{\sigma_n}^{\xi_{n-1},\xi_n} + \tilde{h}_{s,T}^{u_T^\alpha}(\beta)\;\Big|\;\mathcal{F}_s\right], \quad \alpha \in \mathcal{A}_s^i, \tag{44}$$

and define the corresponding optimal value $\tilde{V}_s^i(\beta) = \operatorname{ess\,sup}_{\alpha\in\mathcal{A}_s^i} \tilde{J}_s^i(\alpha;\beta)$. Using the results in [11, 24], we can prove the existence of an optimal control $\alpha^* \in \mathcal{A}_s^i$ for this problem. The non-anticipativity issue arises from the dependence of (44) on the expected future rewards due to player 2's control β.

Acknowledgement

Randall Martyr and John Moriarty would like to thank the UK Engineering and Physical Sciences Research Council (EPSRC) for its financial support via Grant EP/N013492/1 and Grant EP/P002625/1 respectively.

References

[1] Ball, J. A., Chudoung, J. and Day, M. V. (2002). Robust Optimal Switching Control for Nonlinear Systems. *SIAM Journal on Control and Optimization* **41,** 900–931.

[2] Bayraktar, E., Cosso, A. and Pham, H. (2016). Robust Feedback Switching Control: Dynamic Programming and Viscosity Solutions. *SIAM Journal on Control and Optimization* **54,** 2594–2628.

[3] Bismut, J. M. (1977). Sur un problème de dynkin. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **39,** 31–53.

[4] Bismut, J.-M. (1979). Contrôle de processus alternants et applications. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **47,** 241–288.

[5] Bismut, J.-M. (1981). Convex inequalities in stochastic control. *Journal of Functional Analysis* **42,** 226–270.

[6] Bouchard, B. (2009). A stochastic target formulation for optimal switching problems in finite horizon. *Stochastics An International Journal of Probability and Stochastic Processes* **81,** 171–197.

[7] Chassagneux, J. F., Elie, R. and Kharroubi, I. (2011). A note on existence and uniqueness for solutions of multidimensional reflected BSDES. *Electronic Communications in Probability* **16,** 120–128.

[8] Chung, K. L. and Williams, R. J. (2014). *Introduction to stochastic integration* second ed. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, NY.

[9] Cohen, S. N. and Elliott, R. J. (2015). *Stochastic Calculus and Applications.* Probability and Its Applications. Springer New York, New York, NY.

[10] Cosso, A. (2013). Stochastic Differential Games Involving Impulse Controls and Double-Obstacle Quasi-variational Inequalities. *SIAM Journal on Control and Optimization* **51,** 2102–2131.

[11] Djehiche, B., Hamadène, S., Morlais, M.-A. and Zhao, X. (2017). On the equality of solutions of maxmin and minmax systems of variational inequalities with interconnected bilateral obstacles. *Journal of Mathematical Analysis and Applications* **452,** 148–175.

[12] Djehiche, B., Hamadène, S. and Popier, A. (2009). A Finite Horizon Optimal Multiple Switching Problem. *SIAM Journal on Control and Optimization* **48,** 2751–2770.

[13] Dumitrescu, R., Quenez, M.-c. and Sulem, A. (2016). Generalized Dynkin games and doubly reflected BSDEs with jumps. *Electronic Journal of Probability* **21,** 1–32.

[14] Elie, R. and Kharroubi, I. (2014). Adding constraints to BSDEs with jumps: an alternative to multidimensional reflections. *ESAIM: Probability and Statistics* **18,** 233–250.

[15] Hamadène, S. and Hassani, M. (2004). BSDEs with two reflecting barriers : the general result. *Probability Theory and Related Fields* **132,** 237–264.

[16] Hamadène, S. and Hassani, M. (2006). BSDEs with two reflecting barriers driven by a Brownian motion and Poisson noise and related Dynkin game. *Electronic Journal of Probability* **11,** 121–145.

[17] Hamadène, S. and Morlais, M. A. (2013). Viscosity Solutions of Systems of PDEs with Interconnected Obstacles and Switching Problem. *Applied Mathematics & Optimization* **67,** 163–196.

[18] Hamadene, S. and Mu, T. (2019). Systems of reflected BSDEs with interconnected bilateral obstalces: Existence, Uniqueness and Applications. *HAL preprint: hal-01973450, version 1, https://hal.archives-ouvertes.fr/hal-01973450*.

[19] Hamadène, S. and Zhang, J. (2010). Switching problem and related system of reflected backward SDEs. *Stochastic Processes and their Applications* **120,** 403–426.

[20] Hu, Y. and Tang, S. (2010). Multi-dimensional BSDE with oblique reflection and optimal switching. *Probability Theory and Related Fields* **147,** 89–121.

[21] Hu, Y. and Tang, S. (2015). Switching game of backward stochastic differential equations and associated system of obliquely reflected backward stochastic differential equations. *Discrete and Continuous Dynamical Systems* **35,** 5447–5465.

[22] Krasovskiĭ, N. N. and Subbotin, A. I. (1988). *Game-Theoretical Control Problems*. Springer Series in Soviet Mathematics. Springer New York, New York, NY.

[23] Lundström, N. L. P., Nyström, K. and Olofsson, M. (2014). Systems of variational inequalities in the context of optimal switching problems and operators of Kolmogorov type. *Annali di Matematica Pura ed Applicata* **193,** 1213–1247.

[24] Martyr, R. (2016). Finite-Horizon Optimal Multiple Switching with Signed Switching Costs. *Mathematics of Operations Research* **41,** 1432–1447.

[25] Martyr, R. (2016). Solving finite time horizon Dynkin games by optimal switching. *Journal of Applied Probability* **53,** 957–973.

[26] Morimoto, H. (1987). Optimal switching for alternating processes. *Applied Mathematics & Optimization* **16,** 1–17.

[27] Pham, T. and Zhang, J. (2013). Some norm estimates for semimartingales. *Electronic Journal of Probability* **18,** 1–26.

[28] Robin, M. (1976). Some optimal control problems for queueing systems. In *Stochastic Systems: Modeling, Identification and Optimization, II (Mathematical Programming Studies Vol. 6)*. ed. R. J.-B. Wets. vol. 6. Springer Berlin Heidelberg ch. 10, pp. 154–169.

[29] Stettner, L. (1982). Zero-sum Markov games with stopping and impulsive strategies. *Applied Mathematics & Optimization* **9,** 1–24.

[30] Tang, S. and Hou, S.-h. (2007). Switching Games of Stochastic Differential Systems. *SIAM Journal on Control and Optimization* **46,** 900–929.

[31] Tang, S. and Yong, J. (1993). Finite horizon stochastic optimal switching and impulse controls with a viscosity solution approach. *Stochastics An International Journal of Probability and Stochastic Processes* **45,** 145–176.

[32] Yong, J. (1990). A Zero-Sum Differential Game in a Finite Duration with Switching Strategies. *SIAM Journal on Control and Optimization* **28,** 1234–1250.

[33] Yong, J. (1990). Differential games with switching strategies. *Journal of Mathematical Analysis and Applications* **145,** 455–469.

[34] Zabczyk, J. (1973). Optimal control by means of switching. *Studia Mathematica* **45,** 161–171.

[35] Zvonkin, A. K. (1971). On Sequentially Controlled Markov Processes. *Mathematics of the USSR-Sbornik* **15,** 607–617.