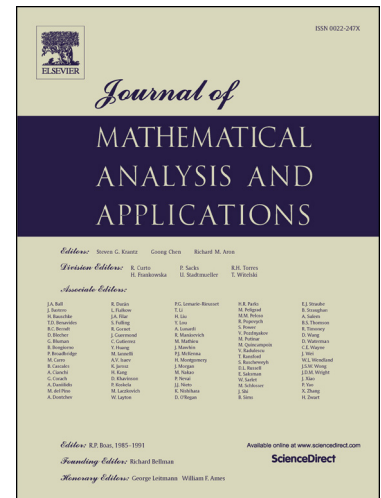


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Existence, uniqueness and regularity of solutions to systems of nonlocal obstacle problems related to optimal switching

Niklas L.P. Lundström* Marcus Olofsson† Thomas Önskog‡

November 5, 2018

Abstract

We study viscosity solutions to a system of nonlinear degenerate parabolic partial integro-differential equations with interconnected obstacles. This type of problem occurs in the context of optimal switching problems when the dynamics of the underlying state variable is described by an n -dimensional Lévy process. We first establish a continuous dependence estimate for viscosity sub- and supersolutions to the system under mild regularity, growth and structural assumptions on the partial integro-differential operator and on the obstacles and terminal conditions. Using the continuous dependence estimate, we obtain the comparison principle and uniqueness of viscosity solutions as well as Lipschitz regularity in the spatial variables. Our main contribution is construction of suitable families of viscosity sub- and supersolutions which we use as “barrier functions” to prove Hölder continuity in the time variable, and, through Perron’s method, existence of a unique viscosity solution. This paper generalizes parts of the results of Biswas, Jakobsen and Karlsen (2010) [BJK10] and of Lundström, Nyström and Olofsson (2014) [LNO14, LNO14b] to hold for more general systems of equations.

2000 *Mathematics Subject Classification*: 35R09, 49L25, 45G15.

Keywords and phrases: variational inequality, existence, viscosity solution, nonlocal operator, partial integro-differential operator, Lévy process, jump diffusion, optimal switching problem, regularity, continuous dependence, well posed

1 Introduction

This paper deals with systems of partial integro-differential equations with interconnected obstacles. The problem considered can be stated as

$$\begin{aligned} \min \Big\{ & -\partial_t u_i(x, t) + \mathcal{F}_i(x, t, u_i(x, t), Du_i(x, t), D^2 u_i(x, t), u_i(\cdot, t)), \\ & u_i(x, t) - \max_{j \neq i} \{u_j(x, t) - c_{ij}(x, t)\} \Big\} = 0, \\ & u_i(x, T) = g_i(x), \end{aligned} \quad (1.1)$$

*Department of Mathematics and Mathematical Statistics, Umeå University, SE-90187 Umeå, Sweden. E-mail: niklas.lundstrom@umu.se

†Department of Mathematics, Uppsala University, SE-75106 Uppsala, Sweden. E-mail: marcus.olofsson@math.uu.se

‡Department of Mathematics, KTH Royal Institute of Technology, SE-10044 Stockholm, Sweden. E-mail: onskog@kth.se

for $x \in \mathbb{R}^n$, $t \in [0, T]$ and $i \in \{1, 2, \dots, m\}$. Here the \mathcal{F}_i are partial integro-differential operators of the form

$$\mathcal{F}_i(x, t, r, p, X, \varphi(\cdot)) = -\mathcal{L}_i(x, t, r, p, X) - \mathcal{J}_i(x, t, p, \varphi(\cdot)) - f^i(x, t),$$

for local second order, possibly degenerate, linear operators \mathcal{L}_i , nonlocal operators \mathcal{J}_i and functions f^i .

A particular case of problem (1.1) arises in multi-modes switching problems, in which the operators \mathcal{F}_i and \mathcal{I}_i can be interpreted as the infinitesimal generators of the stochastic processes underlying the optimization problem and where c_{ij} is the cost of switching from state i to state j . The viscosity solution of (1.1) is then the value function of the multi-modes switching problem from which one can find the sought optimal strategy. For the sake of completeness, a brief outline of the optimal switching problem and its relation to systems like (1.1) is given in Section 5.

Since the pioneering work of Brennan and Schwarz [BS85] dealing with a two-modes switching problem describing the life cycle of an investment in the natural resource industry, systems of variational inequalities with interconnected obstacles of type (1.1) has been extensively studied. We here give a non-exhaustive list of contributions. Starting in the purely local setting, i.e., when the operators $\mathcal{I}_i \equiv 0$, problem (1.1) and its connection to multi-modes switching have been studied by, e.g., El Asri and Hamadène [AH09], Djehiche, Hamadène, and Popier [DHP10], Hamadène and Morlais [HM12], Hu and Tang [HT07], Lundström, Nyström, and Olofsson [LNO14b] and Djehiche, Hamadène, Morlais and Zhao [DHMZ14]. In the majority of the above cited papers, the main focus lies on existence and uniqueness of solutions to (1.1).

The nonlocal setting has not been studied to the same extent, but the amount of literature is steadily growing. Biswas, Jakobsen, and Karlsen [BJK10] show that the dynamic programming principle for multi-modes switching problems holds also for a nonlocal underlying stochastic process, i.e., when the process is allowed to jump, and this is used to prove that the value function of such a multi-modes switching problem indeed satisfies a system like (1.1). The authors also proceed to prove existence, uniqueness and some regularity of the viscosity solution. The nonlocal setting has also been studied by Lundström, Nyström, and Olofsson [LNO14], using PDE techniques, and recently by Hamadène and Zhao [HZ15], and Klimsiak [K16, K16b] using mainly stochastic techniques. Hamadène and Morlais [HM16] have recently studied the non-local setting and proven existence and uniqueness results similar to those presented here when the function f^i has a non-local dependence. However, their work deal with a single (decoupled) obstacle ($m = 1$) and thus dodge much of the technical difficulties arising when constructing barriers in this paper. Also, [HM16] deal with a finite jump-measure λ , a restriction which is later relaxed in [H16]. Another note is that the latter papers rely mainly on stochastic techniques, i.e., backward SDEs, while the present paper is based on PDE-theory. In connection to the above, important contributions to the mathematical theory of viscosity solutions to second order partial integro-differential equations that deserve to be mentioned are given by Barles and Imbert [BI08] and Jakobsen and Karlsen [JK06].

A special technical difficulty when studying systems like (1.1) is the treatment of the obstacles. Therefore, many of the above references choose to impose rather strong assumptions on the switching costs c_{ij} , e.g., positive and constant switching costs. In fact, to the authors knowledge, possibly negative switching costs are only treated by El Asri and Fakhouri [AF17] and Martyr [M14] using stochastic techniques and recently in [LNO14] and [LNO14b] using PDE techniques. We note that it is natural to allow for negative switching costs as these allow one to

model the situation when, e.g., a government through environmental policies provides subsidies, grants or other financial support to energy production facilities in case they switch to more ‘green’ production. In this case it is not a cost for the facility to switch, it is a gain.

In this paper, we first establish a continuous dependence estimate (Theorem 3.1) by using standard methods for viscosity solutions of integro PDEs. Such an estimate bounds the difference between a solution to (1.1) and a solution to a slightly different version of (1.1) in terms of the difference between the equations and the terminal conditions. Continuous dependence estimates of this type are important in themselves as they quantify the stability properties of viscosity solutions, but are also useful in numerical analysis, see, e.g., Krylov [Kr05] and Barles and Jakobsen [BJ07]. In this paper, the continuous dependence estimate is used to obtain the comparison principle and uniqueness of solutions (Corollary 3.2), as well as regularity of the solution, both in space and time (Corollary 3.3 and Theorem 3.4). The main contributions of this paper are Hölder continuity in the time variable, and the existence of a unique viscosity solution to system (1.1) (Theorem 3.5). The proofs of these theorems rely on nontrivial constructions of families of viscosity sub- and supersolutions (Lemma 4.8 and Lemma 4.9) which we use as barrier functions in order to trap the viscosity solution via the comparison principle.

Our results generalize regularity and existence results of Biswas, Jakobsen, and Karlsen [BJK10] and of Lundström, Nyström, and Olofsson [LNO14] in the sense that we allow for more general systems of equations. We impose weaker assumptions on the operator and less regularity and structural assumptions on the spatially dependent switching costs c_{ij} . Our barrier functions also imply a more general existence result in the setting of Kolmogorov operators studied in Lundström, Nyström, and Olofsson [LNO14b].

2 Preliminaries

In this section we introduce the notation used in the paper, discuss some preliminaries and state the assumptions imposed on the system (1.1).

For smooth functions $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ we let $D\varphi = (\partial_{x_1}\varphi, \dots, \partial_{x_n}\varphi)$ denote the spatial gradient of φ and $D^2\varphi$ the Hessian matrix of φ . We denote the set of natural numbers by $\mathcal{N} = \{1, 2, 3, \dots\}$ and let \mathcal{I}_m be the integer set $\{1, 2, \dots, m\}$. For any positive integer p we let $LSC_p(\mathbb{R}^n \times [0, T])$ and $USC_p(\mathbb{R}^n \times [0, T])$ denote the spaces of lower- and upper semicontinuous functions, respectively, on $\mathbb{R}^n \times [0, T]$, whose elements h satisfy the growth condition

$$|h(x, t)| \leq K(1 + |x|^p), \quad \text{whenever } (x, t) \in \mathbb{R}^n \times [0, T]. \quad (2.1)$$

Here and in the following, K denotes a generic constant, $1 < K < \infty$, which may change value from line to line. The space $C_p^{a,b}(\mathbb{R}^n \times [0, T])$ contains all functions $\varphi : \mathbb{R}^n \times [0, T] \rightarrow \mathbb{R}$ which are a times continuously differentiable in the spatial variables and b times continuously differentiable in the time variable and which satisfy the polynomial growth condition (2.1). From here on in, we fix the growth parameter p . We denote the indicator function for the closed unit ball in \mathbb{R}^l by $\chi_{\{|z| \leq 1\}}$ and let $B(x, r)$ be the closed ball in \mathbb{R}^n which has radius r and is centered at x . We let I_n denote the $n \times n$ identity matrix, and let \mathbb{S}^n denote the space of $n \times n$ real symmetric matrices equipped with the positive semi-definite ordering, i.e., for $X, Y \in \mathbb{S}^n$, we write $X \leq Y$ if $\langle (X - Y)\xi, \xi \rangle \leq 0$ for all $\xi \in \mathbb{R}^n$. We will also make use of the matrix norm notation

$$\|A\| := \sup\{|\lambda| : \lambda \text{ is an eigenvalue of } A\} = \sup\{|\langle A\xi, \xi \rangle| : |\xi| \leq 1, \xi \in \mathbb{R}^n\}.$$

The supremum norm is denoted $\|h\|_\infty := \sup_{x \in \mathbb{R}^n} |h(x)|$ for any function h defined on \mathbb{R}^n .

We assume that the operator \mathcal{F}_i can be decomposed into a local second order operator \mathcal{L}_i , a nonlocal integral operator \mathcal{J}_i and a function f^i such that

$$\mathcal{F}_i(x, t, r, p, X, \varphi(\cdot)) = -\mathcal{L}_i(x, t, r, p, X) - \mathcal{J}_i(x, t, p, \varphi(\cdot)) - f^i(x, t), \quad (F_1)$$

for $(x, t, r, p, X) \in \mathbb{R}^n \times [0, T] \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n$ and any smooth function $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$. We assume that the local operators \mathcal{L}_i can be written as

$$\mathcal{L}_i(x, t, r, p, X) = \sum_{k,l=1}^n a_{kl}^i(x, t) X_{kl} + \sum_{k=1}^n b_k^i(x, t) p_k - c^i(x, t) r,$$

for continuous functions a_{kl}^i , b_k^i and c^i . We denote by a^i the $n \times n$ matrix with elements a_{kl}^i , and by b^i the vector of length n with elements b_k^i . Moreover, we assume that

$$a_{kl}^i(x, t) = (\sigma^i(x, t)(\sigma^i)^*(x, t))_{kl}, \quad \text{for } i \in \mathcal{I}_m, \quad k, l \in \mathcal{I}_n,$$

for an $n \times n$ matrix σ^i and where $(\sigma^i)^*$ is the transpose of σ^i . The functions σ_{kl}^i , b_k^i , c^i and f^i are assumed to satisfy

$$\begin{aligned} |\sigma_{kl}^i(x, t) - \sigma_{kl}^i(y, t)| + |b_k^i(x, t) - b_k^i(y, t)| + |c^i(x, t) - c^i(y, t)| &\leq K|x - y|, \\ |f^i(x, t) - f^i(y, t)| &\leq K(1 + |x|^{p-1} + |y|^{p-1})|x - y| \\ |b_k^i(0, t)| + |\sigma_{kl}^i(0, t)| + |f^i(0, t)| - c^i(x, t) &\leq K \end{aligned} \quad (F_2)$$

whenever $k, l \in \mathcal{I}_n$, $i \in \mathcal{I}_m$, $x, y \in \mathbb{R}^n$ and $t \in [0, T]$. Concerning the nonlocal operators \mathcal{J}_i we assume they can be written as

$$\mathcal{J}_i(x, t, p, \varphi(\cdot)) = \int_{\mathbb{R}^l \setminus \{0\}} \varphi(x + \eta^i(x, t, z)) - \varphi(x) - \chi_{\{|z| \leq 1\}} \langle \eta^i(x, t, z), p \rangle \nu^i(dz),$$

where ν^i is a positive Radon measure defined on $\mathbb{R}^l \setminus \{0\}$ and η^i is an \mathbb{R}^n -valued function, continuous in x and t and Borel measurable in z . We assume that ν^i and η^i satisfy

$$\begin{aligned} \int_{0 < |z| \leq 1} |z|^2 \nu^i(dz) + \int_{|z| > 1} |z|^p \nu^i(dz) &< K, \\ |\eta_k^i(x, t, z) - \eta_k^i(y, t, z)| &\leq K(1 + |z|)|x - y| \quad \text{and} \quad |\eta_k^i(x, t, z)| \leq K(1 + |x|)|z| \end{aligned} \quad (F_3)$$

whenever $k \in \mathcal{I}_n$, $i \in \mathcal{I}_m$, $x, y \in \mathbb{R}^n$, $t \in [0, T]$ and $z \in \mathbb{R}^l$.

The functions c_{ij} appearing in the obstacle are called “switching costs” due to the connection between (1.1) and optimal switching problems. In light of this connection, the following definition makes sense.

Definition 2.1 *A switching chain from state i to state j is a sequence of indices $(i_1, \dots, i_l) \in \mathcal{I}_m^l$ such that $i_1 = i$ and $i_l = j$. The set of switching chains from i to j is denoted \mathcal{A}_{ij} .*

We assume c_{ij} to be continuous functions satisfying the classical no-loop condition, i.e.,

$$\min_{(i_1, \dots, i_l) \in \mathcal{A}_{ii}} \sum_{k=1}^{l-1} c_{i_k i_{k+1}}(x, t) > 0, \quad (O_1)$$

for all $i \in \mathcal{I}_m$ and $(x, t) \in \mathbb{R}^n \times [0, T]$. Moreover, we will need the stronger structural assumption

$$c_{ik}(x, t) \leq c_{ij}(x, t) + c_{jk}(x, t), \quad (O_2)$$

whenever $i, j, k \in \mathcal{I}_m$, $x \in \mathbb{R}^n$ and $t \in [0, T]$. The assumption (O_2) is needed for our existence and time-regularity results, (in particular, to prove Lemma 4.8), but we stress that this assumption can be made without loss of generality in the context of optimal switching, see Section 5.

For Lemma 4.8 we also need to assume that c_{ij} is locally semi-concave in space, locally Lipschitz continuous in both space and time, and satisfy a polynomial growth condition in space. In particular, we assume that

$$\begin{aligned} |c_{ij}(x, t) - c_{ij}(y, t')| &\leq K (1 + |x|^{p-1} + |y|^{p-1}) |(x, t) - (y, t')|, \\ D^2 c_{ij}(z, s) &\leq K (1 + |z|^{p-2}) I_n, \end{aligned} \quad (O_3)$$

whenever $i, j \in \mathcal{I}_m$, $x, y \in \mathbb{R}^n$, $t, t' \in [0, T]$ and for almost every $(z, s) \in \mathbb{R}^n \times [0, T]$.

Moreover, we assume that the terminal data g_i is locally Lipschitz continuous, and, to be able to achieve continuity up to the terminal time T , that g_i are consistent with the obstacle, i.e.,

$$|g_i(x) - g_i(y)| \leq K (1 + |x|^{p-1} + |y|^{p-1}) |x - y|, \quad g_i(x) \geq \max_{j \neq i} \{g_j(x) - c_{ij}(x, T)\}, \quad (G)$$

whenever $i \in \mathcal{I}_m$ and $x, y \in \mathbb{R}^n$.

Since the matrices in the local operators \mathcal{L}_i and the jump vectors η^j are allowed to vanish, we cannot expect any smoothing from the equation itself. Therefore, a notion of weak solutions is needed and we will consider solutions in the viscosity sense.

Definition 2.2 A vector $u = (u_1, \dots, u_m)$, where $u_i \in USC_p(\mathbb{R}^n \times [0, T])$ (or $u_i \in LSC_p(\mathbb{R}^n \times [0, T])$) for all $i \in \mathcal{I}_m$, is a viscosity subsolution (supersolution) to system (1.1) if $u_i(x, T) \leq g_i(x)$ ($u_i(x, T) \geq g_i(x)$) whenever $x \in \mathbb{R}^n$, $i \in \mathcal{I}_m$, and if the following holds. For every $(x_0, t_0) \in \mathbb{R}^n \times [0, T)$ and $\varphi \in C_p^{2,1}(\mathbb{R}^n \times [0, T])$ such that (x_0, t_0) is a global maximum (minimum) of $u_i - \varphi$, for some $i \in \mathcal{I}_m$, we have

$$\begin{aligned} \min \left\{ -\partial_t \varphi_i(x_0, t_0) + \mathcal{F}_i(x_0, t_0, u_i(x_0, t_0), D\varphi(x_0, t_0), D^2\varphi(x_0, t_0), \varphi(\cdot, t_0)), \right. \\ \left. u_i(x_0, t_0) - \max_{j \neq i} \{u_j(x_0, t_0) - c_{ij}(x_0, t_0)\} \right\} \leq (\geq) 0, \end{aligned}$$

A vector $u = (u_1, \dots, u_m)$, where $u_i \in C_p(\mathbb{R}^n \times [0, T])$ for all $i \in \mathcal{I}_m$, is a viscosity solution to system (1.1) if it is both a viscosity subsolution and a viscosity supersolution.

Note that the test function appears in the nonlocal slot of the operator \mathcal{I}_i in Definition 2.2. This is necessary due to the infinite activity of the jump measure ν^i close to the origin. However, away from the origin, the important property for \mathcal{I}_i to be well-defined is not regularity but rather restrictions on its growth at infinity, see (F_3) . Therefore, outside of the origin one may replace the test function φ with the solution itself, u_i , and get an equivalent definition of a viscosity solution, see Lemma 2.1 of [BJK10]. When constructing barrier super- and subsolutions in Lemma 4.8 and Lemma 4.9, we will use the above definition. However, when proving the continuous dependence estimate (Theorem 3.1) and the existence of a solution (Theorem 3.5), some necessary calculations follow those of [BJK10] and [LNO14b]. As [BJK10] and [LNO14b] use the latter definition, we will do the same in the proofs of Theorem 3.1 and Theorem 3.5 to avoid repetition of lengthy calculations.

3 Main results

In this section, we list the main results of the paper. All proofs are postponed to Section 4. In the following, we write ‘depending on the data’ to indicate dependence on (at most) the constants K and p introduced in assumptions (F_1) – (F_3) , (O_1) – (O_3) and (G) , as well as dependence on the dimension n and the terminal time T . To state our first result, which is a continuous dependence estimate, we let, for all $i \in \mathcal{I}_m$, $\widehat{\mathcal{F}}_i$ denote the operator \mathcal{F}_i , but with σ_{kl}^i , b_k^i , c^i , f^i , η_k^i and ν^i replaced by $\widehat{\sigma}_{kl}^i$, \widehat{b}_k^i , \widehat{c}^i , \widehat{f}^i , $\widehat{\eta}_k^i$ and $\widehat{\nu}^i$, respectively.

Theorem 3.1 (Continuous dependence estimate) *Let $u = (u_1, \dots, u_m)$ be a viscosity subsolution of system (1.1) and let $\widehat{u} = (\widehat{u}_1, \dots, \widehat{u}_n)$ be a viscosity supersolution of another system of the form (1.1) defined with $\widehat{\mathcal{F}}_i$, \widehat{g}_i and \widehat{c}_{ij} in place of \mathcal{F}_i , g_i and c_{ij} . Assume that both systems satisfy (F_1) – (F_3) , (O_1) and (G) . Then there exists a positive constant C , depending only on the data, such that*

$$\begin{aligned} u_i(x, t) - \widehat{u}_i(x, t) &\leq \max_{j \in \mathcal{I}_m} \|g_j - \widehat{g}_j\|_\infty + T \max_{j \in \mathcal{I}_m} \|f^j - \widehat{f}^j\|_\infty \\ &\quad + C \max_{j \in \mathcal{I}_m} \left\{ \|c^j - \widehat{c}^j\|_\infty + \|b^j - \widehat{b}^j\|_\infty + \|\sigma^j - \widehat{\sigma}^j\|_\infty \right. \\ &\quad \left. + \left\| \int |\bar{\eta}^j|^2 |\nu^j - \widehat{\nu}^j|(dz) \right\|_\infty^{1/2} + \left\| \int |\eta^j - \widehat{\eta}^j|^2 \bar{\nu}^j(dz) \right\|_\infty^{1/2} \right\}, \end{aligned}$$

whenever $(x, t) \in \mathbb{R}^n \times [0, T]$ and $i \in \mathcal{I}_m$, where $\bar{\eta}^i = \max\{\eta^i, \widehat{\eta}^i\}$ and $\bar{\nu}^i = \max\{\nu^i, \widehat{\nu}^i\}$.

The classical comparison principle and Lipschitz regularity in the spatial variables easily follows from Theorem 3.1. In particular, setting $\widehat{\mathcal{F}}_i = \mathcal{F}_i$, $\widehat{g}_i = g_i$ and $\widehat{c}_{ij} = c_{ij}$ in Theorem 3.1 gives the following corollary.

Corollary 3.2 (Comparison principle) *Let $u^- = (u_1^-, \dots, u_n^-)$ and $u^+ = (u_1^+, \dots, u_n^+)$ be a viscosity subsolution and a viscosity supersolution of system (1.1), respectively. Assume (F_1) – (F_3) , (O_1) and (G) . Then $u_i^-(x, t) \leq u_i^+(x, t)$ for all $(x, t) \in \mathbb{R}^n \times [0, T]$ and $i \in \mathcal{I}_m$. As a consequence, viscosity solutions to system (1.1) are unique (in the class of polynomial growth).*

With the above results in place, consider a viscosity solution u of system (1.1) and define, for all $i, j \in \mathcal{I}_m$, $k, l \in \mathcal{I}_n$ and $h \in \mathbb{R}^n$,

$$\left(\widehat{\sigma}_{kl}^i, \widehat{b}_k^i, \widehat{c}^i, \widehat{f}^i, \widehat{c}_{ij}, \widehat{\eta}_k^i \right) (x, t) := \left(\sigma_{kl}^i, b_k^i, c^i, f^i, c_{ij}, \eta_k^i \right) (x + h, t).$$

Setting $\widehat{\nu}^i = \nu^i$ and $\widehat{g}_i(x) = g_i(x + h)$ it follows that $\widehat{u}(x, t) = u(x + h, t)$ is a viscosity solution to (1.1) with $\widehat{\mathcal{F}}_i$, \widehat{g}_i and \widehat{c}_{ij} in place of \mathcal{F}_i , g_i and c_{ij} . By the assumptions of Section 2 it follows that we can bound the right-hand side of the estimate in Theorem 3.1 by $C(1 + |x|^{p-1} + |x + h|^{p-1})|h|$, where C is a positive constant depending only on the data. Hence, by an application of Theorem 3.1 we have the following corollary.

Corollary 3.3 (Lipschitz continuity in space) *Assume (F_1) – (F_3) , (O_1) and (G) . Then there exists a constant C , depending only on the data, such that for any viscosity solution $u = (u_1, \dots, u_m)$ to system (1.1) satisfying $u_i \in C_p(\mathbb{R}^n \times [0, T])$ for all $i \in \mathcal{I}_m$, it holds that*

$$|u_i(x, t) - u_i(y, t)| \leq C(1 + |x|^{p-1} + |y|^{p-1})|x - y|,$$

for any $i \in \mathcal{I}_m$, $x, y \in \mathbb{R}^n$ and $t \in [0, T]$.

We proceed by stating our results on Hölder continuity in time and on existence of solutions. To prove these theorems we construct families of viscosity super- and subsolutions (Lemma 4.8 and Lemma 4.9) which we use as barrier functions in the comparison principle. To this end, we need to impose additional assumptions on the switching costs.

Theorem 3.4 (Hölder continuity in time) *Assume (F_1) – (F_3) , (O_1) and (G) . Suppose also that either $c_{ij} \geq 0$ in $\mathbb{R}^n \times [0, T]$ for all $i, j \in \mathcal{I}_m$, or that (O_2) and (O_3) holds. Then there exists a constant C , depending only on the data, such that for any viscosity solution $u = (u_1, \dots, u_m)$ to system (1.1) satisfying $u_i \in C_p(\mathbb{R}^n \times [0, T])$ for all $i \in \mathcal{I}_m$, it holds that*

$$|u_i(x, s) - u_i(x, t)| \leq C (1 + |x|^p) |s - t|^{1/2}$$

for all $i \in \mathcal{I}_m$, $x \in \mathbb{R}^n$ and $s, t \in [0, T]$.

Finally, the following existence theorem is proved via Perron's method. Here, the barrier functions from Lemma 4.8 and Lemma 4.9 are used to ensure that the Perron solution is bounded and attains the terminal data.

Theorem 3.5 (Existence) *Assume (F_1) – (F_3) , (O_1) and (G) . Suppose also that either $c_{ij} \geq 0$ in $\mathbb{R}^n \times [0, T]$ for all $i, j \in \mathcal{I}_m$, or that (O_2) and (O_3) holds. Then there exists a unique viscosity solution $u = (u_1, \dots, u_m)$ of system (1.1) satisfying $u_i \in C_p(\mathbb{R}^n \times [0, T])$ for all $i \in \mathcal{I}_m$.*

Remark 3.6 *There is an $|x|^p$ -dependence in the right hand side of Theorem 3.4, whereas the $|x|$ -dependence in the corresponding Hölder estimate in [BJK10] (Lemma 5.3) is linear. This is due to relaxed growth assumptions on f_i , c_{ij} and g_i . In particular, setting $p = 1$ in Theorem 3.4 we retrieve the result of [BJK10] in the more general setting studied here.*

Remark 3.7 *Concerning generality we note that it should be possible to further relax the assumptions (F_1) – (F_3) , by applying the full generality of the results of [BI08] and [JK06]. In particular, the continuous dependence estimate may be generalized using [BI08] and [JK06]. Given the validity of a continuous dependence estimate, if the assumptions on the operator then implies Lipschitz continuity in space and the validity of (4.21), then our barrier constructions hold and all our main results follows. We have chosen to stay within the “standard” assumptions (F_1) – (F_3) in this paper to avoid additional technicalities and lengthy assumptions that are hard to interpret.*

4 Proofs of the main results

In this section we prove Theorems 3.1, 3.4 and 3.5.

Proof of Theorem 3.1 (Continuous dependence estimate) We proceed along the lines of [BJK10, Theorem 5.1] to which we refer for additional details.

For constants $\lambda, \theta, \gamma, \epsilon > 0$ we define the test function

$$\phi(t, x, y) = e^{\lambda(T-t)} \frac{\theta}{2} |x - y|^2 + e^{\lambda(T-t)} \frac{\epsilon}{2 + \gamma} (|x|^{2+\gamma} + |y|^{2+\gamma}),$$

on $[0, T] \times \mathbb{R}^n \times \mathbb{R}^n$. We double the variables by defining for $i \in \mathcal{I}_m$,

$$\Psi_i(t, x, y) = u_i(x, t) - \widehat{u}_i(y, t) - \phi(t, x, y) - \frac{\delta(T-t)}{T}\sigma - \frac{\bar{\epsilon}}{t}$$

where $0 < \delta < 1$, $\bar{\epsilon} > 0$, and

$$\sigma = \sup_{i, t, x, y} \left\{ u_i(x, t) - \widehat{u}_i(y, t) - \phi(t, x, y) - \frac{\bar{\epsilon}}{t} \right\} - \sigma_T,$$

$$\sigma_T = \sup_{i, x, y} \left\{ u_i(x, T) - \widehat{u}_i(y, T) - \phi(T, x, y) - \frac{\bar{\epsilon}}{T} \right\}.$$

From this we see that

$$u_i(x, t) - \widehat{u}_i(x, t) - e^{\lambda T} \epsilon |x|^{2+\gamma} - \frac{\bar{\epsilon}}{t} \leq \sigma + \sigma_T, \quad \text{whenever } (x, t) \in \mathbb{R}^n \times [0, T], i \in \mathcal{I}_m, \quad (4.1)$$

and thus the main steps of the proof is to derive an upper bound on σ and σ_T . We start by establishing a bound for σ . If $\sigma \leq 0$ we can take 0 as the upper bound and we are done. Therefore we will assume in the following that $\sigma > 0$. By the upper semicontinuity of $u_i - \widehat{u}_i$, the growth assumptions (provided $2 + \gamma > p$), and the penalization term $-\bar{\epsilon}/t$, there exists $(i_0, t_0, x_0, y_0) \in \mathcal{I}_m \times (0, T] \times \mathbb{R}^n \times \mathbb{R}^n$ such that

$$\Psi_{i_0}(t_0, x_0, y_0) = \sup_{i, t, x, y} \Psi_i(t, x, y).$$

The assumption $\sigma > 0$ forces $t_0 < T$. To see this we observe that

$$\Psi_{i_0}(t_0, x_0, y_0) \geq \sup_{i, t, x, y} \left\{ u_i(x, t) - \widehat{u}_i(y, t) - \phi(t, x, y) - \frac{\bar{\epsilon}}{t} \right\} - \delta\sigma = \sigma_T + (1 - \delta)\sigma > \sigma_T,$$

as $\delta < 1$, while on the other hand $t_0 = T$ would imply $\Psi_{i_0}(t_0, x_0, y_0) = \sigma_T$.

Now we are in a position to apply the maximum principle for semicontinuous functions adapted to nonlocal systems. As we allow switching costs c_{ij} to depend on (x, t) , as well as polynomial growth of viscosity solutions, we may not apply [BJK10, Lemma 4.1] immediately. However, we may apply the generalized version of this result found in [LNO14b, Proof of Theorem 1.1], to retrieve the analogue of estimate (5.3) in [BJK10]. For each $0 < \kappa \leq 1$ there are symmetric matrices X and Y , and we can chose the index i_0 such that

$$\begin{aligned} & -\phi_t(t_0, x_0, y_0) + \frac{\delta\sigma}{T} + \frac{\bar{\epsilon}}{t_0^2} \\ & \leq -\widehat{\mathcal{L}}_{i_0}(y_0, t_0, \widehat{u}^{i_0}(y_0, t_0), -D_y\phi(t_0, x_0, y_0), Y) - \widehat{f}^i(y_0, t_0) \\ & \quad - \widehat{\mathcal{J}}_{i_0, \kappa}(y_0, t_0, -D_y\phi(t_0, x_0, y_0), -\phi(t_0, x_0, \cdot)) - \widehat{\mathcal{J}}_{i_0}^\kappa(y_0, t_0, -D_y\phi(t_0, x_0, y_0), \widehat{u}^{i_0}(\cdot, t_0)) \\ & \quad + \mathcal{L}_{i_0}(x_0, t_0, u^{i_0}(x_0, t_0), D_x\phi(t_0, x_0, y_0), X) + f^i(x_0, t_0) \\ & \quad + \mathcal{J}_{i_0, \kappa}(x_0, t_0, D_x\phi(t_0, x_0, y_0), -\phi(t_0, \cdot, y_0)) + \mathcal{J}_{i_0}^\kappa(x_0, t_0, D_x\phi(t_0, x_0, y_0), u^{i_0}(\cdot, t_0)), \end{aligned} \quad (4.2)$$

where the matrices X and Y satisfy standard upper bounds depending on the second derivatives of $\phi(t, x, y)$. We remind the reader that we here consider an alternative but equivalent definition

of viscosity solutions and refer to [BJK10] and [LNO14b] for details. In (4.2), the splitting of the nonlocal term is defined as in [BJK10, Definition 2.1].

Following [BJK10] we obtain the estimates

$$\begin{aligned} & \operatorname{tr} (a^i(x_0, t_0)X) - \operatorname{tr} (a^i(y_0, t_0)Y) \\ & \leq C e^{\lambda(T-t_0)} \left\{ \theta |x_0 - y_0|^2 + \theta \|\sigma^i - \widehat{\sigma}^i\|_\infty^2 + \epsilon (1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma}) \right\}, \end{aligned} \quad (4.3)$$

$$\begin{aligned} & \widehat{b}^i(y_0, t_0) D_y \phi(t_0, x_0, y_0) + b^i(x_0, t_0) D_x \phi(t_0, x_0, y_0) \\ & \leq C e^{\lambda(T-t_0)} \left\{ \theta |x_0 - y_0|^2 + \theta \|b^i - \widehat{b}^i\|_\infty^2 + \epsilon (1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma}) \right\}, \end{aligned} \quad (4.4)$$

$$|\widehat{f}^i(y_0, t_0) - f^i(x_0, t_0)| \leq \|f^i - \widehat{f}^i\|_\infty + C (1 + |x_0|^{p-1} + |y_0|^{p-1}) |x_0 - y_0|, \quad (4.5)$$

for any $i \in \mathcal{I}_m$. Applying the polynomial growth assumptions of u, \widehat{u}, c and \widehat{c} yields

$$\begin{aligned} & |\widehat{c}^i(y_0, t_0) \widehat{u}(y_0, t_0) - c^i(x_0, t_0) u(x_0, t_0)| \\ & \leq C (1 + |x_0|^p + |y_0|^p) \|c^i - \widehat{c}^i\|_\infty + C (1 + |x_0|^p + |y_0|^p) |x_0 - y_0|. \end{aligned} \quad (4.6)$$

For the nonlocal terms we obtain

$$\begin{aligned} & \mathcal{J}_{i_0, \kappa}(x_0, t_0, D_x \phi(t_0, x_0, y_0), -\phi(t_0, \cdot, y_0)) + \mathcal{J}_{i_0}^\kappa(x_0, t_0, D_x \phi(t_0, x_0, y_0), u^{i_0}(\cdot, t_0)), \\ & - \widehat{\mathcal{J}}_{i_0, \kappa}(y_0, t_0, -D_y \phi(t_0, x_0, y_0), -\phi(t_0, x_0, \cdot)) - \widehat{\mathcal{J}}_{i_0}^\kappa(y_0, t_0, -D_y \phi(t_0, x_0, y_0), \widehat{u}^{i_0}(\cdot, t_0)) \\ & \leq C \theta e^{\lambda(T-t_0)} \left\{ |x_0 - y_0|^2 + \left\| \int |\overline{\eta}^{i_0}|^2 |\nu^{i_0} - \widehat{\nu}^{i_0}|(dz) \right\|_\infty + \left\| \int |\eta^{i_0} - \widehat{\eta}^{i_0}|^2 \overline{\nu}^{i_0}(dz) \right\|_\infty \right\} \\ & + \mathcal{O}(\kappa) + C e^{\lambda(T-t_0)} \epsilon (1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma}), \end{aligned} \quad (4.7)$$

where $\overline{\eta}^{i_0} = \max\{\eta^{i_0}, \widehat{\eta}^{i_0}\}$ and $\overline{\nu}^{i_0} = \max\{\nu^{i_0}, \widehat{\nu}^{i_0}\}$. Now by (4.2), estimates (4.3)-(4.7), and the form of ϕ_t , it follows that

$$\begin{aligned} & \lambda \left[e^{\lambda(T-t_0)} \frac{\theta}{2} |x_0 - y_0|^2 + e^{\lambda(T-t_0)} \frac{\epsilon}{2+\gamma} (|x_0|^{2+\gamma} + |y_0|^{2+\gamma}) \right] + \frac{\delta\sigma}{T} + \frac{\bar{\epsilon}}{t_0^2} \\ & \leq C \theta e^{\lambda(T-t_0)} \max_{i \in \mathcal{I}_m} \left\{ \|\sigma^i - \widehat{\sigma}^i\|_\infty^2 + \|b^i - \widehat{b}^i\|_\infty^2 + \left\| \int |\overline{\eta}^i|^2 |\nu^i - \widehat{\nu}^i|(dz) \right\|_\infty + \left\| \int |\eta^i - \widehat{\eta}^i|^2 \overline{\nu}^i(dz) \right\|_\infty \right\} \\ & + \max_{i \in \mathcal{I}_m} \|f^i - \widehat{f}^i\|_\infty + C (1 + |x_0|^p + |y_0|^p) \max_{i \in \mathcal{I}_m} \|c^i - \widehat{c}^i\|_\infty + C \theta e^{\lambda(T-t_0)} |x_0 - y_0|^2 \\ & + C (1 + |x_0|^p + |y_0|^p) |x_0 - y_0| + C e^{\lambda(T-t_0)} \epsilon (1 + |x_0|^{2+\gamma} + |y_0|^{2+\gamma}) + \mathcal{O}(\kappa) \end{aligned}$$

where the constant C is not necessarily the same at each occurrence but may depend only on the data. In the above estimate, t_0, x_0 and y_0 are independent of κ , so we can let $\kappa \rightarrow 0$ and ignore the term $\mathcal{O}(\kappa)$. By taking λ large enough, its magnitude depending only on the data, we can conclude that

$$\begin{aligned} \delta\sigma & \leq C T \theta \max_{i \in \mathcal{I}_m} \left\{ \|\sigma^i - \widehat{\sigma}^i\|_\infty^2 + \|b^i - \widehat{b}^i\|_\infty^2 + \left\| \int |\overline{\eta}^i|^2 |\nu^i - \widehat{\nu}^i|(dz) \right\|_\infty + \left\| \int |\eta^i - \widehat{\eta}^i|^2 \overline{\nu}^i(dz) \right\|_\infty \right\} \\ & + T \max_{i \in \mathcal{I}_m} \|f^i - \widehat{f}^i\|_\infty + T \sup_{x, y} \Gamma(x, y), \end{aligned} \quad (4.8)$$

where

$$\begin{aligned}\Gamma(x, y) &= C(1 + |x|^p + |y|^p) \max_{i \in \mathcal{I}_m} \|c^i - \tilde{c}^i\|_\infty - \theta |x - y|^2 + C(1 + |x|^p + |y|^p) |x - y| \\ &\quad - \epsilon(1 + |x|^{2+\gamma} + |y|^{2+\gamma}) + C\epsilon.\end{aligned}$$

In fact, by increasing λ even further we see that we can take

$$\begin{aligned}\Gamma(x, y) &= C(1 + |x| + |y|)^p \max_{i \in \mathcal{I}_m} \|c^i - \tilde{c}^i\|_\infty - \theta |x - y|^2 + C(1 + |x| + |y|)^p |x - y| \\ &\quad - \epsilon(1 + |x| + |y|)^{2+\gamma} + C\epsilon,\end{aligned}$$

and after a maximization with respect to $|x - y|$ we have

$$\Gamma(x, y) \leq C(1 + |x| + |y|)^p \max_{i \in \mathcal{I}_m} \|c^i - \tilde{c}^i\|_\infty + C \frac{(1 + |x| + |y|)^{2p}}{\theta} - \epsilon(1 + |x| + |y|)^{2+\gamma} + C\epsilon,$$

and so

$$\Gamma(x, y) \leq C(1 + |x| + |y|)^{2p} \left(\max_{i \in \mathcal{I}_m} \|c^i - \tilde{c}^i\|_\infty + \frac{1}{\theta} \right) - \epsilon(1 + |x| + |y|)^{2+\gamma} + C\epsilon.$$

Now, pick $\gamma = 4p - 2$ and maximize anew, this time with respect to $(1 + |x| + |y|)$. The result is

$$\Gamma(x, y) \leq \frac{C}{\epsilon} \left(\max_{i \in \mathcal{I}_m} \|c^i - \tilde{c}^i\|_\infty^2 + \frac{1}{\theta^2} \right) + C\epsilon.$$

and by now choosing $\epsilon = 1/\theta$ we can conclude that

$$\Gamma(x, y) \leq C\theta \max_{i \in \mathcal{I}_m} \|c^i - \tilde{c}^i\|_\infty^2 + \frac{C}{\theta}. \quad (4.9)$$

We next estimate σ_T . We have, using (G), that

$$\begin{aligned}\sigma_T &= \sup_{i, x, y} \left\{ g_i(x) - \hat{g}_i(y) - \phi(T, x, y) - \frac{\bar{\epsilon}}{T} \right\} \\ &\leq \max_{i \in \mathcal{I}_m} \|g^i - \hat{g}^i\|_\infty + \sup_{x, y} \left\{ C(1 + |x|^{p-1} + |y|^{p-1}) |x - y| - \phi(T, x, y) \right\} \\ &= \max_{i \in \mathcal{I}_m} \|g^i - \hat{g}^i\|_\infty + \sup_{x, y} \Gamma_T(x, y),\end{aligned}$$

where $\Gamma_T(x, y)$ can be bounded in a similar way as $\Gamma(x, y)$, i.e.,

$$\begin{aligned}\Gamma_T(x, y) &\leq C(1 + |x| + |y|)^{p-1} |x - y| - \frac{\theta}{2} |x - y|^2 - \frac{\epsilon}{C} (1 + |x| + |y|)^{2+\gamma} \\ &\leq C \frac{(1 + |x| + |y|)^{2p}}{\theta} - \frac{\epsilon}{C} (1 + |x| + |y|)^{4p} \leq \frac{C}{\theta^2 \epsilon} = \frac{C}{\theta}\end{aligned}$$

and hence

$$\sigma_T \leq \max_{i \in \mathcal{I}_m} \|g^i - \hat{g}^i\|_\infty + \frac{C}{\theta}. \quad (4.10)$$

Collecting the estimates (4.8), (4.9) and (4.10), sending $\delta \rightarrow 1$ and inserting them in (4.1) yields, after noting that the term $-e^{\lambda T} \epsilon |x|^{2+\gamma}$ in (4.1) can be absorbed by $\Gamma(x, y)$ (by an increase in λ),

$$\begin{aligned} u_i(x, t) - \widehat{u}_i(x, t) - \frac{\bar{\epsilon}}{t} &\leq \sigma + \sigma_T + e^{\lambda T} \epsilon |x|^{2+\gamma} \\ &\leq CT\theta \max_{j \in \mathcal{I}_m} \left\{ \|\sigma^j - \widehat{\sigma}^j\|_\infty^2 + \|b^j - \widehat{b}^j\|_\infty^2 + \|c^j - \widehat{c}^j\|_\infty^2 + \left\| \int |\overline{\eta}^j|^2 |\nu^j - \widehat{\nu}^j|(dz) \right\|_\infty \right. \\ &\quad \left. + \left\| \int |\eta^j - \widehat{\eta}^j|^2 \widehat{\nu}^j(dz) \right\|_\infty \right\} + \max_{j \in \mathcal{I}_m} \|g^j - \widehat{g}^j\|_\infty + T \max_{j \in \mathcal{I}_m} \|f^j - \widehat{f}^j\|_\infty + \frac{CT}{\theta}, \end{aligned}$$

whenever $(x, t) \in \mathbb{R}^n \times [0, T]$ and $i \in \mathcal{I}_m$. After minimizing the right hand side with respect to θ and sending $\bar{\epsilon} \rightarrow 0$, we have

$$\begin{aligned} u_i(x, t) - \widehat{u}_i(x, t) &\leq CT \max_{j \in \mathcal{I}_m} \left\{ \|\sigma^j - \widehat{\sigma}^j\|_\infty + \|b^j - \widehat{b}^j\|_\infty + \|c^j - \widehat{c}^j\|_\infty + \left\| \int |\overline{\eta}^j|^2 |\nu^j - \widehat{\nu}^j|(dz) \right\|_\infty^{1/2} \right. \\ &\quad \left. + \left\| \int |\eta^j - \widehat{\eta}^j|^2 \widehat{\nu}^j(dz) \right\|_\infty^{1/2} \right\} + \max_{j \in \mathcal{I}_m} \|g^j - \widehat{g}^j\|_\infty + T \max_{j \in \mathcal{I}_m} \|f^j - \widehat{f}^j\|_\infty, \end{aligned}$$

whenever $(x, t) \in \mathbb{R}^n \times [0, T]$ and $i \in \mathcal{I}_m$. This completes the proof of the theorem. \square

We will now prove Theorem 3.4 and Theorem 3.5 by building appropriate barrier functions. The barrier functions will be constructed as viscosity super- and subsolutions to (1.1) which, by the comparison principle (Corollary 3.2), will give bounds for the unique viscosity solution from above and below, respectively. Before going into the proof we note that the main difficulty lies in constructing an appropriate family of supersolutions which exceed the obstacle in the case when the switching costs c_{ij} are allowed to be negative. Our main idea for this construction is to include the switching costs explicitly in the barrier. Since the switching costs are allowed to be non-smooth, the operator cannot be applied directly and we need to consider approximation arguments and viscosity solution theory. Lemma 4.8 gives an appropriate family of viscosity supersolutions in this case. A suitable subsolution, and a supersolutions when $c_{ij} \geq 0$, can be constructed independent of the switching costs and is therefore much simpler. Lemma 4.9 gives the appropriate barriers in this case, which do not need assumptions (O_2) and (O_3) . The barrier constructions was inspired by related arguments in Lundström, Nyström and Olofsson [LNO14, LNO14b], Biswas, Jacobsen and Karlsen [BJK10], Ishii and Sato [IS04] and Lundström and Önskog [LÖ15].

Lemma 4.8 *Assume (F_1) – (F_3) , (O_1) – (O_3) and let $h = (h_1, \dots, h_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a function satisfying (G) for some positive integer p . Then there exists a positive constant c , depending only on K, n and p , such that for all $(y, s) \in \mathbb{R}^n \times [0, T]$ and all $i \in \mathcal{I}_m$, the function $\psi^{i,y,s} = (\psi_1^{i,y,s}, \dots, \psi_m^{i,y,s})$, defined as*

$$\psi_j^{i,y,s}(x, t) = c \lambda e^{c(s-t)} \left\{ A c(s-t) + \frac{A}{\lambda^2} + B|x-y|^2 + |x-y|^p \right\} + h_i(y) + c_{ij}(x, t)$$

for all $j \in \mathcal{I}_m$, where $A = (1 + |y|^p)$ and $B = (1 + |y|^{p-2})$, is a viscosity supersolution to (1.1) with terminal condition given by h , in $\mathbb{R}^n \times [0, s)$ and whenever $\lambda \geq 1$.

Lemma 4.9 Assume (F_1) – (F_3) and (O_1) and let h, c, A and B be as in Lemma 4.8. Then for all $(y, s) \in \mathbb{R}^n \times [0, T]$ the function $\check{\psi}^{y,s} = (\check{\psi}_1^{y,s}, \dots, \check{\psi}_m^{y,s})$ defined as

$$\check{\psi}_j^{y,s}(x, t) = -c \lambda e^{c(s-t)} \left\{ A c(s-t) + \frac{A}{\lambda^2} + B|x-y|^2 + |x-y|^p \right\} + h_j(y)$$

for all $j \in \mathcal{I}_m$ is a viscosity subsolution to (1.1), with terminal condition given by h , in $\mathbb{R}^n \times [0, s)$ and whenever $\lambda \geq 1$.

Moreover, if $c_{ij} \geq 0$ in $\mathbb{R}^n \times [0, s)$ for all $i, j \in \mathcal{I}_m$, then $\hat{\psi}^{y,s} = (\hat{\psi}_1^{y,s}, \dots, \hat{\psi}_m^{y,s})$ defined as

$$\hat{\psi}_j^{y,s}(x, t) = c \lambda e^{c(s-t)} \left\{ A c(s-t) + \frac{A}{\lambda^2} + B|x-y|^2 + |x-y|^p \right\} + h_j(y)$$

is a viscosity supersolution to (1.1).

Proof. That $\check{\psi}_j^{y,s}$ is a subsolution follows by repeating steps 1 and 3 in the proof of Lemma 4.8 given below. Both steps are simpler in this case since $\check{\psi}_j^{y,s}$ does not involve the switching costs c_{ij} . That $\hat{\psi}_j^{y,s}$ is a supersolution follows similarly by noting that step 2 holds trivially since $c_{ij} \geq 0$ in $\mathbb{R}^n \times [0, s)$ for all $i, j \in \mathcal{I}_m$. \square

Before proving Lemma 4.8 we recall two well-known results in Lemma 4.10 and Lemma 4.11, needed when we prove that our family of functions in Lemma 4.8 consists of viscosity supersolutions.

Lemma 4.10 Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ be semiconvex and \hat{x} be a strict local maximum point of φ . For $p \in \mathbb{R}^n$, set $\varphi_p(x) = \varphi(x) + \langle p, x \rangle$. Then for $r, \delta > 0$,

$$K = \{x \in B(\hat{x}, r) : \text{there exists } p \in B(0, \delta) \text{ for which } \varphi_p \text{ has a local maximum at } x\}$$

has positive measure.

Proof. This result is given as Lemma A.3 in [CIL92] to which we refer for a proof. \square

Lemma 4.11 Let $\Omega \subset \mathbb{R}^n$. A function $\varphi : \Omega \rightarrow \mathbb{R}$ is differentiable with derivative satisfying $|D\varphi(x) - D\varphi(y)| \leq K|x-y|$ for all $x, y \in \Omega$ if and only if φ is both K -semiconvex and K -semiconcave on Ω , i.e., both $\varphi(x) + \frac{K}{2}|x|^2$ and $-\varphi(x) + \frac{K}{2}|x|^2$ are convex functions on Ω .

Proof. A proof can be found in, e.g., Harvey and Lawson [HL13, Theorem A.1]. \square

Armed with Lemmas 4.10 and 4.11 we are ready to prove Lemma 4.8.

Proof of Lemma 4.8. The proof naturally split into three steps.

Step 1: $\psi_j^{i,y,s}$ satisfies the terminal condition. We have to show that

$$\psi_j^{i,y,s}(x, s) \geq h_j(x) \quad \text{whenever } x \in \mathbb{R}^n. \quad (4.11)$$

Using the fact that $ab \leq a^2 + b^2$ we see that, for all $\lambda \geq 1$,

$$\begin{aligned} \psi_j^{i,y,s}(x, s) &= c \left\{ \lambda B|x-y|^2 + \frac{A}{\lambda} + \lambda|x-y|^p \right\} + h_i(y) + c_{ij}(x, s) \\ &\geq c \left\{ \sqrt{AB}|x-y| + \lambda|x-y|^p \right\} + h_i(y) + c_{ij}(x, s) \\ &\geq \frac{c}{2} \{1 + |y|^{p-1} + |x-y|^{p-1}\} |x-y| + h_i(y) + c_{ij}(x, s). \end{aligned}$$

Now for any $k > 0$ it holds that

$$|y|^k + |x - y|^k \geq \frac{1}{2^{k+1}} \left(|x|^k + |y|^k \right) \quad (4.12)$$

and, using this inequality and that h_i satisfies (G) we can conclude that for $c \geq K2^{p+1}$

$$\begin{aligned} \psi_j^{i,y,s}(x, s) &\geq \frac{c}{2^{p+1}} (1 + |x|^{p-1} + |y|^{p-1}) |x - y| + h_i(y) + c_{ij}(x, s) \\ &\geq h_i(x) + c_{ij}(x, s) \geq h_j(x), \end{aligned}$$

for all $\lambda \geq 1$, all $j \in \mathcal{I}_m$, and all $x \in \mathbb{R}^n$. This proves (4.11) and therefore the terminal condition is fulfilled.

Step 2: $\psi_j^{i,y,s}$ exceeds the obstacle. To show that $\psi_j^{i,y,s}$ exceeds the obstacle we have to show that

$$\psi_j^{i,y,s}(x, t) - \max_{k \neq j} \left\{ \psi_k^{i,y,s}(x, t) - c_{jk}(x, t) \right\} \geq 0 \quad (4.13)$$

at all points $(x, t) \in \mathbb{R}^n \times [0, s]$ and whenever $i, j, k \in \mathcal{I}_m$. To do so we note that assumption (O_2) reads

$$c_{ij}(x, t) + c_{jk}(x, t) - c_{ik}(x, t) \geq 0,$$

for all $i, j, k \in \mathcal{I}_m$, and hence

$$\psi_j^{i,y,s}(x, t) - (\psi_k^{i,y,s}(x, t) - c_{jk}(x, t)) = c_{ij}(x, t) - (c_{ik}(x, t) - c_{jk}(x, t)) \geq 0.$$

Therefore, inequality (4.13) is satisfied and our supersolution candidate exceeds the obstacle.

Step 3: $\psi_j^{i,y,s}$ satisfies the equation of being a supersolution. We cannot apply the operator directly to $\psi_j^{i,y,s}$ since the switching costs c_{ij} are in general not differentiable. Instead, we consider a viscosity solution approach. According to Definition 2.2 we need to show that, for c large enough,

$$-\partial_t \varphi(\hat{x}, \hat{t}) + \mathcal{F}_j(\hat{x}, \hat{t}, \psi_j^{i,y,s}(\hat{x}, \hat{t}), D\varphi(\hat{x}, \hat{t}), D^2\varphi(\hat{x}, \hat{t}), \varphi(\cdot, \hat{t})) \geq 0, \quad (4.14)$$

whenever φ is a $C_p^{2,1}(\mathbb{R}^n \times [0, s])$ function such that, for some $j \in \mathcal{I}_m$, $\psi_j^{i,y,s} - \varphi$ has a global minimum at $(\hat{x}, \hat{t}) \in \mathbb{R}^n \times [0, s]$. We may w.l.o.g. assume that the minimum is strict.

Using the notation

$$\psi_j^{i,y,s}(x, t) = \gamma^{i,y,s}(x, t) + c_{ij}(x, t),$$

where

$$\gamma^{i,y,s}(x, t) = c \lambda e^{c(s-t)} \left\{ A c (s-t) + \frac{A}{\lambda^2} + B|x-y|^2 + |x-y|^p \right\} + h_i(y),$$

we find from assumption (O_3) that

$$\partial_t \varphi(\hat{x}, \hat{t}) \leq \partial_t \gamma^{i,y,s}(\hat{x}, \hat{t}) + K(1 + |\hat{x}|^p) \quad \text{and} \quad |D_{x_k} \varphi(\hat{x}, \hat{t})| \leq D_{x_k} \gamma^{i,y,s}(\hat{x}, \hat{t}) + K(1 + |\hat{x}|^{p-1}),$$

for all $k \in \mathcal{I}_n$. Thus

$$-\partial_t \varphi(\hat{x}, \hat{t}) \geq c^2 \lambda e^{c(s-\hat{t})} \{ A + B|\hat{x} - y|^2 + |\hat{x} - y|^p \} - K(1 + |\hat{x}|^p),$$

and, as (4.12) implies

$$\{A + |\hat{x} - y|^p\} \geq \frac{1}{2} \{1 + |y|^p + |y|^p + |\hat{x} - y|^p\} \geq \frac{1}{2^{p+2}} \{1 + |y|^p + |\hat{x}|^p\},$$

we conclude that

$$-\partial_t \varphi(\hat{x}, \hat{t}) \geq c^2 \lambda e^{c(s-\hat{t})} \frac{1}{2^{p+2}} \{1 + |\hat{x}|^p + |y|^p\} - K(1 + |\hat{x}|^p), \quad (4.15)$$

whenever $(\hat{x}, \hat{t}) \in \mathbb{R}^n \times [0, T]$. Similarly, for the first derivative in space we have that

$$\begin{aligned} |D\varphi(\hat{x}, \hat{t})| &\leq c \lambda e^{c(s-\hat{t})} \{2B|\hat{x} - y| + p|\hat{x} - y|^{p-1}\} + K(1 + |\hat{x}|^{p-1}) \\ &\leq c \lambda e^{c(s-\hat{t})} C(1 + |\hat{x}|^{p-1} + |y|^{p-1}) \end{aligned} \quad (4.16)$$

since

$$\frac{\partial}{\partial x_k} \gamma^{i,y,s}(x, t) = c \lambda e^{c(s-t)} \{2B + p|x - y|^{p-2}\} (x_k - y_k).$$

Here and in the following, by C we will denote a constant, $1 \leq C < \infty$, not necessarily the same at each occurrence, which may depend only on K, n and p .

To establish an upper bound for $D^2\varphi$, we first note that since φ is twice differentiable in space, it is, by Lemma 4.11, locally semi-concave in space. Hence, by assumption (O_3) the function $\psi_j^{i,y,s} - \varphi$ is also semi-concave. We can thus apply Lemma 4.10 to obtain a sequence $(x_k, q_k) \in \mathbb{R}^n \times \mathbb{R}^n$ such that $x_k \rightarrow \hat{x}$, $q_k \rightarrow 0$ and such that the function

$$\psi_j^{i,y,s}(x, \hat{t}) - \varphi(x, \hat{t}) + \langle q_k, x \rangle$$

has a local minimum at (x_k, \hat{t}) . Furthermore, from Lemma 4.10 it follows that we may assume that $D^2\psi_j^{i,y,s}(x_k, \hat{t})$ exists for all k . Therefore,

$$D^2\varphi(x_k, \hat{t}) \leq D^2\psi_j^{i,y,s}(x_k, \hat{t}) = D^2\gamma^{i,y,s}(x_k, \hat{t}) + D^2c_{ij}(x_k, \hat{t}),$$

for all $i, j \in \mathcal{I}_m$. Taking the limit as $k \rightarrow \infty$ and using (O_3) then yields

$$D^2\varphi(\hat{x}, \hat{t}) \leq D^2\gamma^{i,y,s}(\hat{x}, \hat{t}) + K(1 + |\hat{x}|^{p-2}) I_n. \quad (4.17)$$

Now, let δ_{kl} denote the Kronecker-delta and observe that

$$\begin{aligned} \frac{\partial^2}{\partial x_k \partial x_l} \gamma^{i,y,s}(x, t) &= c \lambda e^{c(s-t)} \{2B + p|x - y|^{p-2}\} \delta_{kl} \\ &\quad + c \lambda e^{c(s-t)} p(p-2) |x - y|^{p-4} (x_l - y_l) (x_k - y_k), \end{aligned}$$

from which we get, using (4.17), that

$$\begin{aligned} \|D^2\varphi(\hat{x}, \hat{t})\| &\leq c \lambda e^{c(s-\hat{t})} \{2B + p(p-1)|\hat{x} - y|^{p-2}\} + K(1 + |\hat{x}|^{p-2}) \\ &\leq c \lambda e^{c(s-\hat{t})} C(1 + |\hat{x}|^{p-2} + |y|^{p-2}). \end{aligned} \quad (4.18)$$

Next, by noting that (F_2) implies $|a_{kl}^j(\hat{x}, \hat{t})| \leq C(1 + |\hat{x}|^2)$, $c^j(\hat{x}, \hat{t}) \geq -K$ and $|b_k^j(\hat{x}, \hat{t})| \leq K(1 + |\hat{x}|)$, we conclude that

$$\begin{aligned}
& -\mathcal{L}_j\left(\hat{x}, \hat{t}, \psi_j^{i,y,s}(\hat{x}, \hat{t}), D\varphi(\hat{x}, \hat{t}), D^2\varphi(\hat{x}, \hat{t})\right) \\
&= -\operatorname{tr}\left(a^j(\hat{x}, \hat{t})D^2\varphi(\hat{x}, \hat{t})\right) - \sum_{k=1}^n b_k^j(\hat{x}, \hat{t})\left(D\varphi(\hat{x}, \hat{t})\right)_k + c^j(\hat{x}, \hat{t})\psi_j^{i,y,s}(\hat{x}, \hat{t}) \\
&\geq -n\|a^j(\hat{x}, \hat{t})\|\|D^2\varphi(\hat{x}, \hat{t})\| - |b_k^j(\hat{x}, \hat{t})|\|D\varphi(\hat{x}, \hat{t})\| - K|\psi_j^{i,y,s}(\hat{x}, \hat{t})| \\
&\geq -C(1 + |\hat{x}|^2)c\lambda e^{c(s-\hat{t})}C(1 + |\hat{x}|^{p-2} + |y|^{p-2}) - K(1 + |\hat{x}|)c\lambda e^{c(s-\hat{t})}C(1 + |\hat{x}|^{p-1} + |y|^{p-1}) \\
&\quad - Kc\lambda e^{c(s-\hat{t})}C(1 + |\hat{x}|^p + |y|^p) \\
&\geq -c\lambda e^{c(s-\hat{t})}C(1 + |\hat{x}|^p + |y|^p). \tag{4.19}
\end{aligned}$$

Regarding the nonlocal term \mathcal{J}_j , we first note that since φ is $C_p^{2,1}(\mathbb{R}^n \times [0, s])$, it follows from Taylor's theorem that

$$\varphi(\hat{x} + \eta^j, \hat{t}) - \varphi(\hat{x}, \hat{t}) - \langle D\varphi(\hat{x}, \hat{t}), \eta^j \rangle = \langle D^2\varphi(\hat{x}, \hat{t})\eta^j, \eta^j \rangle + \mathcal{O}(|\eta^j|^2).$$

Here and in the following, we have used the notation $\eta^j = \eta^j(\hat{x}, \hat{t}, z)$. Hence

$$\begin{aligned}
\mathcal{J}_j(\hat{x}, \hat{t}, D\varphi(\hat{x}, \hat{t}), \varphi(\cdot, \hat{t})) &= \int_{\mathbb{R}^d \setminus \{0\}} \left(\varphi(\hat{x} + \eta^j, \hat{t}) - \varphi(\hat{x}, \hat{t}) - \chi_{\{|z| \leq 1\}} \langle \eta^j, D\varphi(\hat{x}, \hat{t}) \rangle \right) \nu^j(dz) \\
&= \int_{B_1(0) \setminus \{0\}} \mathcal{O}(|\eta^j|^2) \nu^j(dz) + \int_{B_1(0) \setminus \{0\}} \|D^2\varphi(\hat{x}, \hat{t})\| |\eta^j|^2 \nu^j(dz) \\
&\quad + \int_{\mathbb{R}^d \setminus B_1(0)} \left(\varphi(\hat{x} + \eta^j, \hat{t}) - \varphi(\hat{x}, \hat{t}) \right) \nu^j(dz) =: I_1 + I_2 + I_3
\end{aligned}$$

and to bound \mathcal{J}_j it suffices to establish an appropriate upper bound on the integrals I_1, I_2 and I_3 . Assumption (F_3) yields

$$I_1 = \int_{B_1(0) \setminus \{0\}} \mathcal{O}(|\eta^j|^2) \nu^j(dz) \leq C(1 + |\hat{x}|^2) \int_{B_1(0) \setminus \{0\}} |z|^2 \nu^j(dz) \leq C(1 + |\hat{x}|^2),$$

while assumption (F_3) and (4.18) gives

$$\begin{aligned}
I_2 &= \int_{B_1(0) \setminus \{0\}} \|D^2\varphi(\hat{x}, \hat{t})\| |\eta^j|^2 \nu^j(dz) \\
&\leq c\lambda e^{c(s-\hat{t})}C(1 + |\hat{x}|^p + |y|^{p-2}) \int_{B_1(0) \setminus \{0\}} |z|^2 \nu^j(dz) \\
&\leq c\lambda e^{c(s-\hat{t})}C(1 + |\hat{x}|^p + |y|^{p-2}).
\end{aligned}$$

To estimate I_3 we observe that

$$\varphi(\hat{x} + \eta^j, \hat{t}) - \varphi(\hat{x}, \hat{t}) \leq \psi_j^{i,y,s,\delta}(\hat{x} + \eta^j, \hat{t}) - \psi_j^{i,y,s,\delta}(\hat{x}, \hat{t})$$

since $\psi_j^{i,y,s,\delta} - \varphi$ has a minimum at (\hat{x}, \hat{t}) . Therefore

$$\begin{aligned} I_3 &= \int_{\mathbb{R}^\ell \setminus B_1(0)} \left(\varphi(\hat{x} + \eta^j, \hat{t}) - \varphi(\hat{x}, \hat{t}) \right) \nu^j(dz) \leq \int_{\mathbb{R}^\ell \setminus B_1(0)} \left(\psi_j^{i,y,s,\delta}(\hat{x} + \eta^j, \hat{t}) - \psi_j^{i,y,s,\delta}(\hat{x}, \hat{t}) \right) \nu^j(dz) \\ &\leq c \lambda e^{c(s-\hat{t})} \int_{\mathbb{R}^\ell \setminus B_1(0)} \left((1 + |y|^{p-2}) (|\hat{x} + \eta^j - y|^2 - |\hat{x} - y|^2) + |\hat{x} + \eta^j - y|^p - |\hat{x} - y|^p \right) \nu^j(dz) \\ &\quad + \int_{\mathbb{R}^\ell \setminus B_1(0)} \left(c_{ij}(\hat{x} + \eta^j, \hat{t}) - c_{ij}(\hat{x}, \hat{t}) \right) \nu^j(dz). \end{aligned}$$

Using (F_3) and (O_3) we obtain

$$\begin{aligned} I_3 &\leq c \lambda e^{c(s-\hat{t})} C (1 + |\hat{x}|^p + |y|^p) \int_{\mathbb{R}^\ell \setminus B_1(0)} \nu^j(dz) + c \lambda e^{c(s-\hat{t})} (1 + |y|^{p-2}) \int_{\mathbb{R}^\ell \setminus B_1(0)} |\eta^j|^2 \nu^j(dz) \\ &\quad + c \lambda e^{c(s-\hat{t})} \int_{\mathbb{R}^\ell \setminus B_1(0)} |\eta^j|^p \nu^j(dz) + C \int_{\mathbb{R}^\ell \setminus B_1(0)} \left(1 + |\hat{x}|^p + |\eta^j|^p \right) \nu^j(dz) \\ &\leq c \lambda e^{c(s-\hat{t})} C \{ (1 + |\hat{x}|^p + |y|^p) + |y|^{p-2} (1 + |\hat{x}|^2) + (1 + |\hat{x}|^p) \} + C (1 + |\hat{x}|^p) \\ &\leq c \lambda e^{c(s-\hat{t})} C (1 + |\hat{x}|^p + |y|^p). \end{aligned}$$

Summing up bounds for I_1 , I_2 and I_3 implies

$$-\mathcal{J}_j(\hat{x}, \hat{t}, D\varphi(\hat{x}, \hat{t}), \varphi(\cdot, \hat{t})) \geq -c \lambda e^{c(s-\hat{t})} C (1 + |\hat{x}|^p + |y|^p), \quad (4.20)$$

where C may depend only on K, n and p .

Finally, assumption (F_2) yields $|f^i(\hat{x}, \hat{t})| \leq C (1 + |\hat{x}|^p)$ and plugging this inequality, (4.15), (4.19) and (4.20) into (4.14) gives us

$$\begin{aligned} &-\partial_t \varphi(\hat{x}, \hat{t}) + \mathcal{F}_j \left(\hat{x}, \hat{t}, \psi_j^{i,y,s}(\hat{x}, \hat{t}), D\varphi(\hat{x}, \hat{t}), D^2\varphi(\hat{x}, \hat{t}), \varphi(\cdot, \hat{t}) \right) \\ &\geq c^2 \lambda e^{c(s-\hat{t})} \frac{1}{2^{p+2}} \{ 1 + |\hat{x}|^p + |y|^p \} \\ &\quad - c \lambda e^{c(s-\hat{t})} C (1 + |\hat{x}|^p + |y|^p) - C (1 + |\hat{x}|^p) \geq 0, \end{aligned} \quad (4.21)$$

where the last inequality is based on choosing c large enough, our choice depending only on K, n and p . Hence, we conclude that $\psi_j^{i,y,s}$ satisfies (4.14) and therefore the equation of being a viscosity supersolution is fulfilled.

This completes step 3 in the proof of Lemma 4.8 and in light of step 1 and step 2 the proof of the Lemma is complete. \square

Proof of Theorem 3.4 (Hölder continuity in time) To prove Hölder continuity in the time variable, let u be a viscosity solution to (1.1) and fix arbitrary $(y, s) \in \mathbb{R}^n \times [0, T]$ and $i \in \mathcal{I}_m$. In the case allowing for signed switching costs we apply Lemma 4.8 with $h_i(x) = u_i(x, s)$. Note that since u is a viscosity solution to (1.1), Corollary 3.3 asserts that h_i satisfies (G) . By the comparison principle, we have

$$u_j(x, t) \leq \psi_j^{i,y,s}(x, t)$$

for all $j \in \mathcal{I}_m$ and $(x, t) \in \mathbb{R}^n \times [0, s]$. In particular, setting $j = i$ and $x = y$ and using $c_{ii} = 0$, this reduces to

$$u_i(y, t) - u_i(y, s) \leq c^2 e^{c(s-t)} A \left(\lambda(s-t) + \frac{1}{\lambda} \right)$$

where $A = (1 + |y|^p)$. For s and t fixed, we set $\lambda = (s - t)^{-1/2}$ and get

$$u_i(y, t) - u_i(y, s) \leq 2c^2 e^{c(s-t)} A(s - t)^{1/2}.$$

Without loss of generality we may assume that $|s - t| \leq 1$, and, therefore,

$$u_i(y, t) - u_i(y, s) \leq C(1 + |y|^p)(s - t)^{1/2},$$

where C may depend only on K, n, p and T . This proves the Hölder continuity in the time variable ‘from above’.

The upper bound on $u_i(y, t) - u_i(y, s)$ in the case with non-negative switching costs, as well as the lower bound on $u_i(y, t) - u_i(y, s)$, follows by the similar argument but uses $\hat{\psi}^{y,s}$ and $\check{\psi}^{y,s}$ from Lemma 4.9 as barriers. The proof of Theorem 3.4 is complete. \square

Proof of Theorem 3.5 (Existence) We construct a viscosity solution to problem (1.1) using Perron’s method. To this end, we define $u = (u_1, \dots, u_m)$ as

$$u_i(x, t) := \inf\{u_i^+(x, t) : u_i^+ = (u_1^+, \dots, u_m^+) \text{ is a supersolution to (1.1)}\}$$

Note that this construction is well defined given the explicit viscosity sub- and supersolutions $\check{\psi}_j^{y,T}(x, t)$, $\hat{\psi}_j^{y,T}(x, t)$ and $\psi_j^{i,y,T}(x, t)$ constructed in Lemma 4.8 and Lemma 4.9. We now intend to prove that u^* and u_* , the upper- and lower semicontinuous envelopes of u , are, respectively, a subsolution and a supersolution to (1.1). It then follows by the comparison principle that $u^* \leq u_*$ and hence $u = u_* = u^*$ is a viscosity solution to (1.1).

We first prove that u^* satisfies the terminal condition of being a subsolution. To this end, we make use of the explicit viscosity supersolutions from Lemmas 4.8 and 4.9 again. (We show the case allowing for signed switching costs, the other case is similar) Fix a component $i \in \mathcal{I}_m$ and a point $y \in \mathbb{R}^n$. By construction of u_i we have $u_i(x, t) \leq \psi_i^{j,y,T}(x, t)$ whenever $(x, t) \in \mathbb{R}^n \times [0, T]$. Moreover, since $\psi_i^{j,y,T}(x, t)$ is continuous, it follows that $u_i^*(x, t) \leq (\psi_i^{j,y,T}(x, t))^* = \psi_i^{j,y,T}(x, t)$ for every $\varepsilon > 0$ and $(x, t) \in \mathbb{R}^n \times [0, T]$. In particular, setting $j = i$ and $(x, t) = (y, T)$, we deduce

$$u_i^*(y, T) \leq \lim_{\lambda \rightarrow \infty} \psi_i^{i,y,T}(y, T) = u_i(y, T) + c_{ii}(y, T) = g_i(y).$$

Since i and y are arbitrary in this argument, we conclude that u^* satisfies the terminal condition. We prove that u_* satisfies the terminal condition in a similar way using the comparison principle and $\check{\psi}_j^{y,T}(x, t)$ from Lemma 4.9 as a barrier from below.

After noticing that our switching costs are continuous, the sub- and supersolution properties are shown in the same way as outlined by Biswas, Jakobsen, and Karlsen [BJK10, pages 70–72]. We refer the interested reader there and conclude that the proof of Theorem 3.5 is complete. \square

5 A brief outline of the optimal switching problem

Consider a production facility which can be run in m different modes of production (henceforth called ‘states’). Given an underlying stochastic process $X = \{X_t\}_{t \geq 0}$, the task of the facility manager is to switch between different states to maximize the total payoff generated by the plant up to some terminal time T . Since switching between states in general comes with a cost, the manager needs to make a trade-off between being in the momentarily best state i^* and minimizing the cost of switching. In particular, if the plant generates a running reward of

$f_i(t, X_t)$ when in state i and the cost of switching from state i to state j at time t is $c_{ij}(X_t, t)$, the manager wishes to maximize

$$\mathbb{E} \left[\int_t^T f_{\mu_s}(X_s, s) ds - \sum_{\substack{k \geq 1 \\ \tau_k \leq T}} c_{\xi_{k-1}, \xi_k}(X_{\tau_k}, \tau_k) \middle| X_t = x \right]. \quad (5.1)$$

In the above, $(\xi_k, \tau_k)_{k \geq 1}$ and μ_s represent the *strategy* of the manager, where $(\xi_k, \tau_k)_{k \geq 1}$ is a sequence of random variables ξ_k and stopping times τ_k , both adapted to the filtration of the stochastic process X_t , indicating that the production should be switched to state ξ_k at time τ_k , and μ_s a function indicating the current state of the production. The task in the optimal switching problem is now to find the strategy μ^* from a given set of *admissible* strategies \mathcal{A} which maximizes the profit in (5.1) and the corresponding *value* of the problem

$$v_i(x, t) = \sup_{\mathcal{A}} \mathbb{E} \left[\int_t^T f_{\mu_s}(X_s, s) ds - \sum_{\substack{k \geq 1 \\ \tau_k \leq T}} c_{\xi_{k-1}, \xi_k}(X_{\tau_k}, \tau_k) \middle| X_t = x, \mu_0 = i \right].$$

When the underlying process X_t is Markovian, it can be shown that, under some assumptions, the value functions $\{v_i(x, t)\}_{i \in \{1, 2, \dots, m\}}$ coincide with a Hamilton–Jacobi–Bellman-type equation with interconnected obstacles. In particular, the value functions satisfy a system of equations of the form

$$\begin{aligned} \min \left\{ -\partial_t v_i(x, t) - \mathcal{F}_i v_i(x, t) - f_i(x, t), \right. \\ \left. v_i(x, t) - \max_{j \neq i} \{-c_{i,j}(x, t) + v_j(x, t)\} \right\} = 0, \\ v_i(x, T) = 0, \end{aligned} \quad (5.2)$$

where \mathcal{F}_i is to be interpreted as the infinitesimal generator of the underlying stochastic process X_t in state i .

The above connection is proved rigorously in the local setting, i.e., when the underlying process has no jumps, in, e.g., [DHP10], where the authors rely on the use of Snell envelopes and backwards SDEs. The paper [TY93] deals with the optimal switching from a viscoisty solutions approach as in the present paper and also provide a proof of the so called dynamic programming principle for optimal switching problems. In the non-local setting, i.e., when the infinitesimal generator F_i is of non-local type, meaning that the underlying process X_t may have jumps, similar results on the connection between (5.2) and optimal switching can be found in, e.g., [BJK10].

We proceed by giving some remarks on our assumptions in the setting of optimal switching problems. Indeed, in this setting, assumption (O_2) is no restriction as we may assume, without loss of generality, that

$$c_{i i_2}(x, t) + c_{i_2 i_3}(x, t) + \dots c_{i_{l-1} j}(x, t) \geq c_{ij}(x, t), \quad (5.3)$$

for all switching chains $(i, i_2, \dots, i_{l-1}, j) \in \mathcal{A}_{ij}$ and any $i, j \in \mathcal{I}_m$. In particular, if (5.3) does not hold, we can construct new switching costs \tilde{c}_{ij} by

$$\tilde{c}_{ij}(x, t) = \min_{(i_1, \dots, i_l) \in \mathcal{A}_{ij}} \sum_{k=1}^{l-1} c_{i_k i_{k+1}}(x, t)$$

which we then consider in place of c_{ij} . Since the regularity assumptions (O_3) on the original switching costs c_{ij} only assume semi-concavity and Lipschitz continuity, the new switching costs \tilde{c}_{ij} will satisfy (O_3) by construction. The same is true for the classical no-loop condition in (O_1). Moreover, using $\tilde{c}_{ij}(x, t)$ in place of c_{ij} will not alter the cost structure of the problem and hence the value function will remain unchanged. In the setting of optimal switching, the switching cost $\tilde{c}_{ij}(x, t)$ represent switching using the “cheapest” switching chain from state i to j and assumption (5.3) means that it is always cheaper to switch directly to a state than to go through some intermediate state. More explicitly, this implies that at any time t at most one switch is made. Note that the switching costs \tilde{c}_{ij} can be no more than Lipschitz continuous, regardless of the regularity of c_{ij} . Hence, it is essential that the regularity assumptions of Lundström, Nyström, and Olofsson [LNO14b] are relaxed in order to assume (5.3) without loss of generality in the context of optimal switching.

References

- [AF17] Brahim El Asri and Imade Fakhouri, *Viscosity solutions for a system of PDEs and optimal switching*, IMA Journal of Mathematical Control and Information, **34**, 937-960.
- [AH09] Brahim El Asri and Saïd Hamadène, *The finite horizon optimal multi-modes switching problem: The viscosity solution approach* Applied Mathematics & Optimization, **60** (2009), 213-235.
- [BI08] Guy Barles and Cyril Imbert, *Second-order elliptic integro-differential equations: viscosity solutions theory revisited*, Annales de l’Institut Henri Poincaré, **25** (2008), 567-585.
- [BJ07] Guy Barles and Espen R. Jakobsen, *Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations*, Math. Comp. **76** (2007), 1861-1893.
- [BJK10] Imran H. Biswas, Espen R. Jakobsen, and Kenneth H. Karlsen, *Viscosity solutions for a system of integro-PDEs and connections to optimal switching and control of jump-diffusion processes*, Applied Mathematics & Optimization, **62** (2010), 47-80.
- [BS85] Michael J. Brennan and Eduardo S. Schwartz, *Evaluating natural resource investments*, Journal of Business, **58** (1985), 135-157.
- [CIL92] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions, *User’s guide to viscosity solutions of second order partial differential equations*, Bulletin of the American Mathematical Society, **27** (1992), 1-67.
- [DHP10] Boualem Djehiche, Saïd Hamadène, and Alexandre Popier, *A Finite Horizon Optimal Multiple Switching Problem*, SIAM Journal on Control and Optimization, **48** (2010), 2751-2770.
- [DHMZ14] Boualem Djehiche, Saïd Hamadène, Marie-Amelie Morlais, and Xuzhe Zhao, *On the equality of solutions of max-min and min-max systems of variational inequalities with interconnected bilateral obstacles*, arXiv preprint arXiv:1408.4282 (2014).

- [HM12] Saïd Hamadène and Marie-Amélie Morlais, *Viscosity Solutions of Systems of PDEs with Interconnected Obstacles and Switching Problem* Applied Mathematics & Optimization, **67**(2013), 163-196.
- [H16] Saïd Hamadène, *Viscosity solutions of second order integro-partial differential equations without monotonicity condition: A new result*, Non-linear analysis **147** (2016): 213-235.
- [HM16] Saïd Hamadène and Marie-Amélie Morlais, *Viscosity solutions for second order integro-differential equations without monotonicity condition: the probabilistic approach*, Stochastics **88** (2016): 632-649.
- [HZ15] Saïd Hamadène and Xuzhe Zhao *Systems of integro-PDE with interconnected obstacles and multi-modes switching problem driven by Lévy process*, Nonlinear differential equations and Applications, **22** (2015), 1607-1660.
- [HL13] F. Reese Harvey, and H. Blaine Lawson, *Notes on the differentiation of quasi-convex functions*, arXiv preprint arXiv:1309.1772 (2013).
- [HT07] Ying Hu and Shanjian Tang, *Multi-dimensional BSDE with Oblique Reflection and Optimal Switching*, Probability Theory and Related Fields, **147** (2010), 89-121.
- [IK91] Hitoshi Ishii and Shigeaki Koike, *Viscosity solutions of a system of nonlinear second-order elliptic PDEs arising in switching games*, Funkcialaj Ekvacioj, **34** (1991), 143-155.
- [IS04] Hitoshi Ishii, Moto-Hiko Sato, *Nonlinear bolique derivative problems for singular degenerate parabolic equations on a general domain*, Nonlinear analysis, **57** (2004), 1077-1098.
- [JK06] Espen R. Jakobsen and Kenneth H. Karlsen, *A Maximum Principle for Semicontinuous Functions Applicable to Integro-Partial Differential Equations*, Nonlinear differential equations and applications, **13** (2006), 137-165.
- [K16] Tomasz Klimsiak, *Obstacle problem for evolution equations involving measure data and operator corresponding to semi-Dirichlet form*, arXiv preprint arXiv:1612.07274 (2016).
- [K16b] Tomasz Klimsiak, *Systems of quasi-variational inequalities related to the switching problem*, arXiv preprint arXiv:1609.02221 (2016).
- [Kr05] Nikolai V. Krylov, *The rate of convergence of finite-difference approximations for Bellman equations with Lipschitz coefficients*, Appl. Math. Optim., **52** (2005) 365-399.
- [LNO14] Niklas L.P. Lundström, Kaj Nyström and Marcus Olofsson, *Systems of variational inequalities for nonlocal operators related to optimal switching problems: existence and uniqueness*, Manuscripta Mathematica, **145** (2014), 407-432.
- [LNO14b] Niklas L.P. Lundström, Kaj Nyström, and Marcus Olofsson, *Systems of variational inequalities in the context of Optimal Switching Problems and Operators of Kolmogorov Type*, Annali di Matematica Pura ed Applicata, **193** (2014), 1213-1247.

- [LÖ15] Niklas L.P. Lundström and Thomas Önskog, *Stochastic and partial differential equations on non-smooth time-dependent domains*, arXiv preprint arXiv:1503.05433 (2015).
- [M14] Randall Martyr, *Finite-horizon optimal multiple switching with signed switching costs*, arXiv preprint arXiv:1411.3971 (2014).
- [TY93] Shanjian Tang and Jiongmin Yong, *Finite horizon stochastic optimal switching and impulse controls with a viscosity solution approach*, Stochastics and Stochastics Reports, **45** (1993), 145-176.