

HÉCTOR JASSO-FUENTES (Ciudad de México)
JOSE-LUIS MENALDI (Detroit, MI)

RELAXATION AND LINEAR PROGRAMS IN A HYBRID CONTROL MODEL

Abstract. Some optimality results for hybrid control problems are presented. The hybrid model under study consists of two subdynamics, one of a standard type governed by an ordinary differential equation, and the other of a special type having a discrete evolution. We focus on the case when the interaction between the subdynamics takes place only when the state of the system reaches a given fixed region of the state space. The controller is able to apply *two* controls, each applied to one of the two subdynamics, whereas the state follows a composite evolution, of continuous type and discrete type. By the relaxation technique, we prove the existence of a pair of controls that minimizes an incurred (discounted) cost. We conclude the analysis by introducing an auxiliary infinite-dimensional linear program to show the equivalence between the initial control problem and its associated relaxed counterpart.

1. Introduction. Hybrid control systems can be considered as a subclass of controlled dynamical systems with the key property that the associated dynamic may undergo structural modifications from time to time, exerted by the own controller or by means of the location of the state of the system. The hybrid control system we are interested in is composed of two subdynamics: one of a standard type that runs in almost all situations, and another of a special type that is activated under extreme circumstances. Any change of subdynamic may produce a structural modification in the system and, at the same time, an opportunity for an instantaneous (and sizeable)

2010 *Mathematics Subject Classification:* 49J15, 49N25, 93C15.

Key words and phrases: optimal control, hybrid systems, relaxed controls, infinite linear programming.

Received 5 February 2019.

Published online 19 September 2019.

change in the state of the system. Naturally, at any given time, only one of the two subdynamics (standard/special) may be active.

The dynamic is a key feature in hybrid control models, but the form of the state and control is also important: The state of the system does not only provide a “usual” description of the phenomenon, but also has a record keeping mechanism. Specifically, the state is represented as a pair where the first entry describes the standard evolution of the system (a continuous-type variable) and the second one records the structural changes (a discrete-type variable). As for the control variable, the controller is able to apply two controls: one acting only on the standard subdynamic and the other acting only on the special subdynamic.

The interaction between the subdynamics is only possible when the state variable reaches a specific region of the state space. In this situation, the hybrid control model is said to satisfy the *automaton property*, as the only possible switch from one subdynamics to the other is carried out automatically in that region. The details are provided in the next section.

In the above set-up, the aim of the controller is to find a control policy, regarded as a pair (with each component controlling each of the two subdynamics), so as to minimize an infinite-horizon discounted cost introduced later. To accomplish this goal, we rely on the relaxation technique, which is a useful and well-known tool for providing optimal control policies. We shall produce optimality results in this framework by applying two different types of hypotheses. In addition, an auxiliary infinite-dimensional linear program is analyzed in order to show the equivalence between the control problem under study and the associated relaxed counterpart.

Related literature on hybrid control systems is vast and it covers both theoretical and practical results. To mention a handful of related works with a more theoretical inclination, we cite Azhmyakov et al. [3], Barles et al. [4], Bensoussan and Menaldi [6], Branicky et al. [7], Dharmatti and Ramaswamy [8], Lygeros [12], Riedinger et al. [14], Shaikh and Caines [15], Zhang and James [18], among others. All these references are based on dynamic programming (through the analysis of some quasi-variational inequalities) or on the use of the well-known maximum principle. For the point of view of applications, specifically in robotics, aircraft planning, and automata, we mention, for instance, Posa et al. [13], Soler et al. [16], Tavernini [17], although there are many others in the literature.

This paper is somehow a continuation of Bensoussan and Menaldi [6]. Indeed, this reference provided conditions ensuring the existence of the optimal value $u(x, n)$, regarded as a continuous viscosity solution of certain quasi-variational inequalities (QVI). However, the existence of optimal controls was not studied in that paper; in fact, the question of existence becomes delicate because there is not enough regularity to provide optimal control

policies as a straightforward consequence of the QVI. However, from the use of relaxation methods either on the control variable or on the state-control variables, it is possible to overcome the difficulties and find optimal controls in this hybrid environment.

To the best of our knowledge there is no literature that uses the same methodology. The preprint of Zhao et al. [19] is somewhat related to ours. It analyzes the optimality of a *finite-horizon* cost through the use of linear programs and occupation measures. Optimal policies are obtained under convexity and affine assumptions on the cost function. In our paper, however, we tackle an optimal control problem for an *infinite-horizon discounted* cost criterion and the techniques used go in two directions: (1) we use a relaxed control approach when certain regularity in some parts of our model is satisfied, and (2) we use occupation measures and linear programs when such regularity is not known in advance. In both cases, the techniques used to find optimal control policies differ considerably from the arguments provided by Zhao et al. [19].

Our paper is divided into six sections. Section 2 presents the details of our dynamical system and introduces some elements of it, such as state, action, and interface spaces, the dynamic of the system, types of control policies, the payoff to be optimized, as well as our main assumptions. Section 3 provides optimality results for the control model under the so-called transversality assumption. To this end, the control is regarded as a “distribution” of controls, i.e., the concept of relaxed control and the corresponding optimality criterion are used. A kind of continuity of the trajectories of the system with respect to the relaxed controls is necessary here, which is the key to finding optimal results. In Section 4, the same problem as in Section 3 is studied, but without assuming the transversality condition; this forces us to regard the state as a “distribution” of states. In this scenario, the control problem is rewritten as an infinite-dimensional linear problem, in which the control policies are replaced by measures with some characteristics. An important feature of the space where these measures live is its relative compactness. Then standard results on continuity-compactness are applied to show the existence of an optimal measure that optimizes our performance criterion. Section 5 is devoted to showing the equivalence between the control problem under study and the control problem associated to the relaxed policies. For this purpose, an auxiliary infinite-dimensional linear program and the corresponding dual are discussed. By studying the restrictions of the dual problem, it is possible to deduce the above equivalence under an additional hypothesis on the costs. Finally, Section 6 provides a discussion of the transversality condition, which allows us to give a detailed proof of Proposition 3.2.

Warning: For simplicity of notation, throughout this article, we shall use (x, n) to represent the initial condition of the state of the system $(x(\cdot), n(\cdot))$

in (2.1) below, but sometimes it will be denoted as (x_0, n_0) or $(x(0), n(0))$ when the context requires it.

2. Model definition. The controlled dynamic system we are interested in is formed by the state space $S = \mathbb{R}^d \times \mathcal{N}$ with $\mathcal{N} \subset \mathbb{R}^l$, the control spaces $V \subset \mathbb{R}^p$ and $K \subset \mathbb{R}^q$, and the interface set $D \subset S$. We have two subdynamics, one of a standard type governed by an ordinary differential equation (ODE) and the other with an (instantaneous) impulsive or transitional character. The state of the system is denoted by the pair $(x, n) \in S$, where x and n represent the continuous-type and discrete-type states, respectively.

Changes of (x, n) over time are effected through the two subdynamics (standard and special) as well as through interventions of the controller carried out by the selection over time of two control parameters $v \in V$ and $k \in K$ acting on the standard and special subdynamics respectively.

The activation of each subdynamic is decided automatically depending on the location of the state variable. To be more specific, when (x, n) belongs to $S \setminus D$, the standard dynamic is turned on and it is affected by the control v . When (x, n) touches D , the special subdynamic acts whose control variable is now k . Certainly, one and only one of the two subdynamics is active at any given time.

Formally, the aforementioned dynamic is represented as follows:

$$(2.1) \quad \underbrace{(\dot{x}(t), \dot{n}(t)) = (g(x(t), n(t_i), v(t)), 0)}_{\text{continuous subdynamic}} \quad \text{for } t \in [t_i, t_{i+1}[,$$

$$\underbrace{(x(t_i), n(t_i)) = (X(x(t_i-), n(t_i-), k_i), N(x(t_i-), n(t_i-), k_i))}_{\text{discrete subdynamic}}, \quad i = 0, 1, \dots,$$

$$t_{i+1} := \inf\{t \geq t_i : (x(t-), n(t_i)) \in D\}, \quad \text{when } t_i < \infty.$$

with initial condition $t_0 = 0$, $(x(0), n(0)) = (x_0, n_0) \in S$, and

$$(2.2) \quad (X(x(t_0-), n(t_0-), k_0), N(x(t_0-), n(t_0-), k_0))$$

$$:= \begin{cases} (X(x_0, n_0, k_1), N(x_0, n_0, k_1)) & \text{when } (x_0, n_0) \in D, \\ (x_0, n_0) & \text{when } (x_0, n_0) \in S \setminus D. \end{cases}$$

In the above dynamic, $t-$ means the left limit of t , and the value $v(t)$ represents the control action used by the controller at time t , which is exerted in the standard subdynamic; in contrast, the values $\{k_i : i \geq 1\}$ are the control actions applied by the controller at each time t_i in the special subdynamic. For convenience, we also use an extra variable k_0 . This variable becomes a fictitious element in our model and just makes sense when $(x_0, n_0) \in D$, whose value is actually k_1 ; in the other situation (when $(x_0, n_0) \in S \setminus D$)

this variable is not used. Moreover, the sequence $\{t_i\}$ is referred to as a *time-interface* set and it is obtained by means of the last line in (2.1).

In other words, the standard subdynamic evolves as an ODE in the continuous-type variable x , with drift (or vector field) expressed by the function $g : S \setminus D \times V \rightarrow \mathbb{R}^d$, while the discrete-type variable n remains constant. In contrast, the special subdynamic is expressed by the *transition function* $(X, N) : D \times K \rightarrow S$.

REMARK 2.1. (a) In previous works (see, for instance, Branicky et al. [7]), the ODE in the hybrid dynamic (2.1) has been considered to change dimension each time the special subdynamic is turned on. In this case, the sequence $\{d_k\}$, whose elements give the dimension of $x \in \mathbb{R}^{d_k}$ at the k th activation of n , could be either bounded or unbounded. In the former situation, we can define $d := \max_k d_k$ and work with a single dimension; whereas in the latter case, we can consider as the range of x the space of sequences $\mathbb{R}_F^\infty := \bigcup_{k \geq 1} \mathbb{R}^k$, i.e., sequences of real numbers with only a finite number of nonzero terms. The convergence of elements in \mathbb{R}_F^∞ is with respect to the inductive topology, i.e., $x^{(n)} \rightarrow x$ if and only if (i) all the $x^{(n)}$ and x belong to the same \mathbb{R}^k for some k (sufficiently large) and (ii) $x^{(n)} \rightarrow x$ in \mathbb{R}^k . Our analysis here is based on the former scenario, that is, when the ODE is of a single finite (and likely large) dimension without any changes.

(b) As part of our hypotheses, we have assumed that $S_n := \{x \in \mathbb{R}^d : (x, n) \in S\} = \mathbb{R}^d$ for all $n \in \mathcal{N}$. Otherwise, if $S_n \subsetneq \mathbb{R}^d$, a more detailed analysis would apply. Namely, by defining

$$(2.3) \quad D_n = \{x \in \mathbb{R}^d : (x, n) \in D\},$$

the case $\partial S_n \subset D_n$ (imposed for instance as the assumption (A2) in Barles et al. [4]) does not cause any inconvenience, because when the state (x, n) reaches the boundary ∂S_n , it also reaches D_n , so a discrete transition is triggered according to the rule in (2.1) (see also (2.6) below). However, the case $\partial S_n \setminus D_n \neq \emptyset$ is more delicate, since there might be situations when the standard subdynamic x gets away from the set S_n in *finite* time. In this case, some additional conditions must be imposed on the vector field g in order to ensure the state x stays inside or even in the border of S_n . We may then have the following conditions: if S_n is a closed set, then the continuous dynamics cannot leave the set $S_n \setminus D_n$. This condition can be achieved if the boundary ∂S_n is piecewise smooth and the vector field $g(\cdot, n, \cdot)$ never points to the exterior of S_n at points of the active boundary $\partial S_n \setminus D_n$. Alternatively, if S_n is an open set, then on $\partial S_n \setminus D_n$ (if non-empty), the vector field $g(\cdot, n, \cdot)$ must point strictly to the interior of S_n . Other possibilities may be used; for instance, we can stop the system evolution when leaving $S_n \setminus D_n$. Note also that the case $\partial S_n \setminus D_n \neq \emptyset$ does not necessarily force the boundary ∂D_n to satisfy a

transversality condition (but such a condition must probably be imposed on the region $\partial S_n \setminus D_n$).

(c) The dynamic (2.1) allows, in principle, the case of multiple (instantaneous) transitions triggered by the discrete subdynamic at some time t_i . Later, we will impose assumptions on the model to avoid this possibility.

Control policies. An *admissible control policy* will be a pair $(v(\cdot), \{k_i\})$ consisting of:

- A continuous-type control that is a Borel measurable V -valued function $v(\cdot)$ on $[0, \infty[$. We denote by \mathcal{V} the set of all continuous-type control policies.
- An impulse-type (or discrete-type) control that consists of a sequence $\{k_i\}$ such that $k_i \in K \subset \mathbb{R}^q$. We denote by \mathcal{K} the set of all impulse-type controls.

Throughout this paper we will assume that the following assumptions hold true.

(a) The interface and control sets satisfy

$$(2.4) \quad \begin{cases} \text{the set interface } D \text{ is closed,} \\ \text{the control spaces } V \text{ and } K \text{ are compact.} \end{cases}$$

(b) There exists a positive constant M , such that

$$(2.5) \quad \begin{cases} g : S \setminus D \times V \rightarrow \mathbb{R}^d \text{ is continuous,} \\ |g(x, n, v)| \leq M, \quad \forall x, n, v, \\ |g(x, n, v) - g(x', n, v)| \leq M|x - x'|, \quad \forall x, x', n, v. \end{cases}$$

(c) The transition function

$$(2.6) \quad (X, N) : D \times K \rightarrow S \text{ is uniformly continuous.}$$

REMARK 2.2. Since K is compact, by a simple use of Tikhonov’s theorem the space \mathcal{K} of impulse controls is compact too in the product topology.

REMARK 2.3. It is easy to verify that conditions in (2.4) and (2.5) ensure that, for any admissible policy $v(\cdot)$, the solution

$$x(t) = x(s) + \int_s^t g(x(r), n, v(r)) dr$$

of the ODE $\dot{x}(t) = g(x(t), n, v(t))$ exists and is unique for $t \geq s$ and each $n \in \mathcal{N}$ (see, for instance, Fleming and Rishel [9]).

Construction of the controlled paths. We now present the construction of the state of the system along time (controlled paths) by means of an *algorithm*. The algorithm provides, in particular, the existence and unique-

ness of the evolution of the state $t \mapsto (x(t), n(t))$. Furthermore, it also generates the sequence $\{t_i : i \geq 1\}$ of time-interfaces, the path $t \mapsto (x(t), n(t))$ defined on each interval of times $[t_i, t_{i+1}[$, and the sequence of transitions $\{(x(t_i), n(t_i)) : i \geq 1\}$ whenever $t_i < \infty$. For simplicity, and when the context requires it, we write $(x(t_i-), n(t_i-))$ as (x_i, n_i) or (x_{i-}, n_{i-}) .

Suppose that a pair $(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}$ is given. Then the algorithm runs as follows:

(0) *Initialization.* Assume (x_0, n_0) is a given initial state at time $t_0 = 0$. According to whether (x_0, n_0) belongs or not to D , the counter i is set: If $(x_0, n_0) \in S \setminus D$, set $i = 0$, $(x_i, n_i) = (x_0, n_0)$, and go to step (1); else if $(x_0, n_0) \in D$, set $i = 1$, $(x_{i-}, n_{i-}) = (x_0, n_0)$, and go to step (2).

(1) *Continuous type.* If $(x_i, n_i) \in S \setminus D$, then the standard subdynamic is activated at time $t_i < \infty$ and the continuous-type state evolves as $x(t) = x_i + \int_{t_i}^t g(x(s), n_i, v(s)) ds$ for $t_i \leq t < t_{i+1}$ (with t_i as in (2.1)), whereas the discrete-type state remains constant, with value $n(t) = n_i$. Thus, $(x_{i+1-}, n_{i+1-}) = (x(t_{i+1-}), n(t_{i+1-}))$ belongs to D , and $x(t) \in S \setminus D$ for every $t \in [t_i, t_{i+1}[$. If $t_{i+1} = \infty$, then stop successfully.

(2) *Discrete type.* If $(x_{i-}, n_{i-}) \in D$, then the special subdynamic is activated at time $t_i < \infty$ and a new state $(X, N)(x_{i-}, n_{i-}, k_i) = (x, n)$ is produced. Now, if $(x, n) \in S \setminus D$ then set $(x_i, n_i) = (x, n)$. Otherwise, i.e., if $(x, n) \in D$, then set $(x_{i+1-}, n_{i+1-}) = (x, n)$ and a discrete-type transition, either $(x_{i+1}, n_{i+1}) = (X, N)(x_{i+1-}, n_{i+1-}, k_{i+1})$ with (x_{i+1}, n_{i+1}) in $S \setminus D$ or $(x_{i+1-}, n_{i+1-}) = (X, N)(x_{i+1-}, n_{i+1-}, k_{i+1})$ with (x_{i+1-}, n_{i+1-}) in D , is triggered again. This is repeated until the state (x_{i+j}, n_{i+j}) belongs to $S \setminus D$. In this case, $t_i = t_{i+1} = \dots = t_{i+j}$, all $(x_{i-}, n_{i-}), (x_{i+1-}, n_{i+1-}), \dots, (x_{i+j-}, n_{i+j-})$ belong to D , and (x_{i+j}, n_{i+j}) belongs to $S \setminus D$. Note that $j + 1$ discrete-type transitions occurred at the same instant of time t_i , and if a finite j as above is not found then this step never ends and the construction fails.

(3) *Iteration.* Now repeat (1) and (2) alternately, i.e., after (1) go to (2) and after (2) go to (1).

(4) *Ending.* If this construction gets trapped in (2), then the hybrid evolution exists only up to time $t_i < \infty$. Otherwise this iteration may end only after step (1) is completed successfully with $t_{i+1} = \infty$, or it may be repeated to generate an infinite sequence $t_0 \leq \dots \leq t_i \leq t_{i+1} \leq \dots$ of impulse/switching times. In any case, the hybrid trajectory $t \mapsto (x(t), n(t))$ is defined as a cadlag function on $[t_0, t_i[$ for any $i \geq 0$, with t_i being either $< +\infty$ or $+\infty$.

In the above algorithm, (i) each k_i is used only when $t_i < \infty, i \geq 1$; thus, the variable i counts the impulse/switching times, (ii) $t_0 = 0$ is the initial time, (iii) the first impulse/switching time t_1 may be equal to t_0 .

As pointed out earlier, some conditions on the data are necessary to ensure that (a) the procedure (0), . . . , (3) does not end with step (2), i.e., after a finite number of instantaneous transitions, a state in $S \setminus D$ is reached; and (b) the evolution runs over time, i.e., the sequence $\{t_i : i \geq 1\}$ diverges to $+\infty$. A sufficient condition that overcomes these drawbacks is the following: There exist constants c and C satisfying

$$(2.7) \quad 0 < c \leq \{|\xi - X(x, n, k)| + |\eta - N(x, n, k)|\} \leq C, \\ \forall (x, n), (\xi, \eta) \in D, k \in K.$$

The following result ensures that the trajectories in (2.1) are well defined in the following sense.

PROPOSITION 2.4. *Under assumptions (2.4)–(2.7), for any pair of controls $(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}$, the trajectory $t \mapsto (x(t), n(t))$ obtained from the hybrid algorithm exists, is unique, and it does not allow simultaneous jumps, i.e., there exists a constant $\mathbf{h} := (M + 1)^{-1} \log[1 + c(M + 1)/M] > 0$ such that the sequence $\{t_i : i \geq 1\}$ of impulse times satisfies $t_{i+1} \geq t_i + \mathbf{h}$ for all $i = 0, 1, \dots$. As a consequence, $t_i \rightarrow \infty$ as $i \rightarrow \infty$, and the trajectories are right continuous with left limits on $[0, \infty[$.*

Proof. The existence and uniqueness follows from the definition of each transition $(X, N)(x(t_i-), n(t_i-), k_i)$ and from the existence and uniqueness of each trajectory $x(t)$ on $[t_i, t_{i+1}[$, both constructed via the hybrid algorithm. Namely, when the hybrid algorithm is at step (1), the continuous-type variable x evolves as an ODE with usual assumptions guaranteeing existence and uniqueness (see Remark 2.3), whereas the discrete-type variable n is a constant. On the other hand, when the hybrid algorithm is at stage (2), the transition function (X, N) acts, which of course produces one and only one value (x, n) from some (x_{i-}, n_{i-}) . Therefore, the existence and uniqueness of the whole evolution $t \mapsto (x(t), n(t))$ follows by linking together the trajectory in accordance with the hybrid algorithm. The last part of the theorem follows from Bensoussan and Menaldi [6, Theorem 2.1]. ■

DEFINITION 2.5. Let $\mathcal{X} \subset \mathbb{R}^j$ with $j \geq d + l + p + q$ (recall the dimension of the state-action spaces).

- (a) We denote by $B_b(\mathcal{X})$ the space of all Borel measurable and bounded real-valued functions on \mathcal{X} , endowed with the supremum norm $\|\cdot\|$.
- (b) $C_b(\mathcal{X})$ and $C_b^u(\mathcal{X})$ are subspaces of $B_b(\mathcal{X})$ consisting of all continuous and all uniformly continuous functions, respectively.
- (c) Consider the special case $\mathcal{X} \equiv S$. We denote by $C_b^{1,0}(S)$ the following set of real-valued functions defined on S :

$$(2.8) \quad C_b^{1,0}(S) := \{\varphi \in C_b(S) : \partial_{x_i} \varphi \in C_b(S), i = 1, \dots, d\},$$

with norm defined by

$$\|\varphi\|_1 := \|\varphi\| + \sum_{i=1}^d \|\partial_{x_i}\varphi\|.$$

Note that in (2.8), the derivative of φ is only applied to the variable x but not to n .

For every pair $(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}$, the dynamic (2.1) can be characterized by the following integration by parts formula (similar to one in Bensoussan and Lions [5, p. 87]): for each $\varphi \in C_b^{1,0}(S)$ and $t \geq 0$,

$$\begin{aligned} (2.9) \quad & e^{-\alpha t} \varphi(x(t), n(t)) - \varphi(x, n) \\ &= \int_{t_0}^t e^{-\alpha s} [g(x(s), n(s), v(s)) \cdot \nabla_x \varphi(x(s), n(s)) - \alpha \varphi(x(s), n(s))] ds \\ &+ \sum_{i=0}^{\infty} e^{-\alpha t_i} [\varphi(X(x(t_i-), n(t_i-), k_i), N((x(t_i-), n(t_i-), k_i)) \\ &\quad - \varphi(x(t_i-), n(t_i-))] \mathbf{1}_{\{t_i \leq t\}}, \end{aligned}$$

where the time-interface set $\{t_i : i \geq 0\}$ is generated by the hybrid algorithm.

Performance index. We introduce the instantaneous and switching cost rates $f : S \setminus D \times V \rightarrow \mathbb{R}$ and $\ell : D \times K \rightarrow \mathbb{R}$, respectively, satisfying the following conditions:

$$(2.10) \quad \begin{cases} f \geq 0 & \text{and} & f \in C_b^u(S \setminus D \times V), \\ \ell \geq 0 & \text{and} & \ell \in C_b^u(D \times K). \end{cases}$$

With the above ingredients, if (x, n) denotes the initial state at time $t_0 = 0$, then, for each pair $(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}$ of controls, the total cost incurred by the controller is defined as

$$(2.11) \quad \begin{aligned} J(x, n; v(\cdot), \{k_i\}) &= \int_0^{\infty} e^{-\alpha t} f(x(t), n(t), v(t)) dt \\ &+ \sum_{i=0}^{\infty} e^{-\alpha t_i} \ell(x(t_i-), n(t_i-), k_i), \end{aligned}$$

where the set $\{t_i : i \geq 1\}$ of impulse times is generated by the set D through the hybrid algorithm as explained earlier. Note that Proposition 2.4 and assumption (2.10) imply that the total cost (2.11) is *finite* for every $(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}$.

The *value function* or *optimal cost* is defined to be

$$(2.12) \quad u(x, n) = \inf_{(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}} J(x, n; v(\cdot), \{k_i\}).$$

Moreover, if there exists $(\hat{v}(\cdot), \{\hat{k}_i\}) \in \mathcal{V} \times \mathcal{K}$ satisfying $J(x, n; \hat{v}(\cdot), \{\hat{k}_i\}) = u(x, n)$, then we will refer to it as an *optimal pair*.

A direct consequence of the above is that, in principle, the optimal cost u is an element of $B_b(S)$.

3. Regular case. The existence and characterization of the optimal cost (2.12) has been studied by Bensoussan and Menaldi [6], who provided conditions ensuring the existence of the optimal value $u(x, n)$, regarded as a continuous viscosity solution of certain quasi-variational inequalities (QVI). However, the existence of optimal controls was not studied; in fact, the existence issue becomes delicate because there is not enough regularity to provide optimal control policies as a straightforward consequence of the QVI.

An effective method of finding optimal controls is the relaxation technique. In this method, the set \mathcal{V} is embedded into a bigger set \mathbf{V} that is convex and compact (in an appropriate topology). Working in this new set, it is possible to show the existence of an element, say $\mathbf{v} \in \mathbf{V}$, together with a suitable impulse control $\{k_i\}$ such that both controls become optimal for the minimization problem (2.12).

In the rest of this section,

- we introduce the concept of relaxed controls and show that this set is compact in a suitable topology;
- we define a new optimal control problem related to the set of relaxed controls;
- we impose a transversality condition on the set-interface D that ensures the continuity (in a certain sense) of the trajectories $(x(\cdot), n(\cdot))$, with respect to the control variables $\mathbf{v} \in \mathbf{V}$ and $\{k_i\} \in \mathcal{K}$;
- we prove the existence of a pair $(\hat{\mathbf{v}}(\cdot), \{\hat{k}_i\}) \in \mathbf{V} \times \mathcal{K}$ such that, for this pair, the total cost defined in (3.3) below equals the value function (2.12).

To begin, we denote by $\mathcal{P}(V)$ the set of all probability measures on V . Let \mathbf{V} be the set of functions $\mathbf{v} : [0, \infty[\rightarrow \mathcal{P}(V)$. We can identify every $v \in \mathcal{V}$ as an element in \mathbf{V} by the relation $v(t)$ “is isomorphic to” $\delta_{v(t)}(\cdot)$, where δ_a denotes the Dirac measure at a . From this last relation, we can interpret \mathcal{V} as a subset of \mathbf{V} . The latter set is known as the set of *relaxed controls*.

The following proposition ensures important properties of \mathbf{V} (for a proof, see for instance Gamkrelidze [10, Theorem 8.1]).

PROPOSITION 3.1. *Under the compactness assumption on V given in (2.4), the set \mathbf{V} of relaxed controls is weakly sequentially compact: for any sequence $\{\mathbf{v}^m(\cdot)\}$ in \mathbf{V} , there exists $\mathbf{v}(\cdot) \in \mathbf{V}$ and a subsequence of $\{\mathbf{v}^m(\cdot)\}$*

(still denoted by $\{\mathbf{v}^m(\cdot)\}$) such that, for all $g \in C_b([0, T] \times V)$ and $T > 0$,

$$(3.1) \quad \int_0^T \int_V g(t, v) \mathbf{v}_t^m(dv) dt \rightarrow \int_0^T \int_V g(t, v) \mathbf{v}_t(dv) dt \quad \text{as } m \rightarrow \infty.$$

In this case, we write $\mathbf{v}^m \xrightarrow{w} \mathbf{v}$ as $m \rightarrow \infty$.

By applying a relaxed control $\mathbf{v}(\cdot) \in \mathbf{V}$ (in lieu of $v(\cdot) \in \mathcal{V}$) together with an impulse-type control $\{k_i\}$ to the hybrid dynamic (2.1), the following *relaxed dynamic* is generated:

$$(3.2) \quad \begin{aligned} (\dot{x}(t), \dot{n}(t)) &= \left(\int_V g(x(t), n(t_i), v) \mathbf{v}_t(dv), 0 \right) \quad \text{for } t \in [t_i, t_{i+1}[, \\ (x(t_i), n(t_i)) &= (X(x(t_i-), n(t_i-), k_i), N(x(t_i-), n(t_i-), k_i)), \\ & \qquad \qquad \qquad i = 0, 1, \dots, \\ t_{i+1} &:= \inf\{t \geq t_i : (x(t-), n(t_i)) \in D\} \quad \text{when } t_i < \infty. \end{aligned}$$

with initial condition as in (2.2).

Following steps (0) to (3) of the hybrid algorithm, we can formally deduce the existence and uniqueness of a solution $t \mapsto (x(t), n(t))$ in (3.2) by mimicking the arguments of Section 2 in the nonrelaxed case. It is also clear that Proposition 2.4 is still valid in the framework of relaxed controls.

For each initial condition (x, n) and any pair $(\mathbf{v}, \{k_i\}) \in \mathbf{V} \times \mathcal{K}$, we define the *relaxed cost function*

$$(3.3) \quad \begin{aligned} \mathbf{J}(x, n; \mathbf{v}(\cdot), \{k_i\}) &= \int_0^\infty e^{-\alpha t} \int_V f(x(t), n(t), v) \mathbf{v}_t(dv) dt \\ &+ \sum_{i=0}^\infty e^{-\alpha t_i} \ell(x(t_i-), n(t_i-), k_i). \end{aligned}$$

Furthermore, we define the *optimal relaxed cost* by

$$(3.4) \quad \mathbf{u}(x, n) = \inf_{(\mathbf{v}, \{k_i\}) \in \mathbf{V} \times \mathcal{K}} \mathbf{J}(x, n; \mathbf{v}(\cdot), \{k_i\}).$$

Recall the set D_n defined in (2.3). We denote by ∂D and ∂D_n the boundaries of D and D_n , respectively; we also write $\overset{\circ}{D}$ for the interior of D .

In order to prove the existence of a pair $(\hat{\mathbf{v}}(\cdot), \{\hat{k}_i\}) \in \mathbf{V} \times \mathcal{K}$ such that $u(x, n) = \mathbf{J}(x, n; \hat{\mathbf{v}}(\cdot), \{\hat{k}_i\})$, we will impose the following *transversality condition* on the boundary of D (and as a consequence on D_n). These conditions have been considered in previous works (see, e.g., Bensoussan and Menaldi [6] or Branicky et al. [7]).

For all $(x, n) \in \partial D$, there exists $\eta(x, n)$, a unit inner normal to ∂D_n , and a positive constant ρ_0 such that, for any n , the function $x \mapsto \eta(x, n)$ belongs to $C_b(\mathbb{R}^d)$, and

$$(3.5) \quad \begin{cases} \partial D_n \text{ is smooth,} \\ |\eta(x, n)| = 1, \quad \forall x \in \partial D_n, & \text{(unit normal)} \\ |\eta(x, n) \cdot g(x, n, v)| \geq \rho_0, \quad \forall (x, n, v) \in \partial D \times V & \text{(transversality).} \end{cases}$$

Our next result concerns the continuity of the trajectories $(x(\cdot), n(\cdot))$ with respect to controls $(\mathbf{v}, \{k_i\}) \in \mathbf{V} \times \mathcal{K}$. The proof, provided in Section 6, strongly uses the transversality condition (3.5).

PROPOSITION 3.2. *Suppose that assumptions (2.4)–(2.7), (2.10), and (3.5) are satisfied. Consider a sequence $\{(\mathbf{v}^m, \{k_i^m\})\}$ of controls in $\mathbf{V} \times \mathcal{K}$ and a pair $(\mathbf{v}^\infty, \{k_i^\infty\}) \in \mathbf{V} \times \mathcal{K}$ such that $\mathbf{v}^m \xrightarrow{w} \mathbf{v}^\infty$ and $\{k_i^m\} \rightarrow \{k_i^\infty\}$ (in the product topology of \mathcal{K}) as $m \rightarrow \infty$. Denote by $(x^m(\cdot), n^m(\cdot))$ the trajectory (3.2) corresponding to $(\mathbf{v}^m, \{k_i^m\})$ for $m \geq 1$. Then $(x^m(\cdot), n^m(\cdot)) \rightarrow (x^\infty(\cdot), n^\infty(\cdot))$ locally uniformly at almost every point. Moreover, the limit trajectory satisfies (3.2) with controls $(\mathbf{v}^\infty, \{k_i^\infty\})$.*

Now we can establish one of our main theorems regarding the existence of relaxed controls that minimize the optimal cost u defined in (2.12).

THEOREM 3.3. *Under the assumptions of Proposition 3.2, there exists a pair $(\hat{\mathbf{v}}, \{\hat{k}_i\})$ consisting of a relaxed control $\hat{\mathbf{v}} \in \mathbf{V}$ and an impulse-type control $\{\hat{k}_i\} \in \mathcal{K}$ such that $\mathbf{J}(x, n, \hat{\mathbf{v}}(\cdot), \{\hat{k}_i\}) = u(x, n)$, where \mathbf{J} is the total cost defined in (3.3).*

Proof. By the definition of infimum, we can select a minimizing sequence $\{(\delta_{v^m(\cdot)}(\cdot), \{k_i^m\})\} \subset \mathbf{V} \times \mathcal{K}$ such that

$$(3.6) \quad \mathbf{J}(x, n, v^m(\cdot), \{k_i^m\}) = \mathbf{J}(x, n, \delta_{v^m(\cdot)}(\cdot), \{k_i^m\}) \downarrow u(x, n) \quad \text{as } m \rightarrow \infty,$$

where

$$(3.7) \quad \begin{aligned} \mathbf{J}(x, n, \delta_{v^m(\cdot)}(\cdot), \{k_i^m\}) &= \int_0^\infty e^{-\alpha t} \int_V f(x^m(t), n^m(t), v) \delta_{v^m(t)}(dv) dt \\ &+ \sum_{i=0}^\infty e^{-\alpha t_i^m} \ell(x(t_i^m-), n(t_i^m-), k_i^m). \end{aligned}$$

Since both \mathbf{V} and \mathcal{K} are compact (the former in a weak sense), we have $\delta_{v^m(\cdot)} \xrightarrow{w} \hat{\mathbf{v}} \in \mathbf{V}$ and $\{k_i^m\} \rightarrow \{\hat{k}_i\} \in \mathcal{K}$ for a subsequence (not relabeled).

The integral in (3.7) can be expressed as follows:

$$\begin{aligned}
 (3.8) \quad & \int_0^\infty e^{-\alpha t} \int_V f(x^m(t), n^m(t), v) \delta_{v^m(t)}(dv) dt \\
 &= \sum_{i=0}^\infty \int_{t_i^\infty}^{t_{i+1}^\infty} e^{-\alpha t} \int_V f(x_i^m(t), n_i^m(t), v) \delta_{v^m(t)}(dv) dt \\
 &= \sum_{i=0}^\infty \int_{t_i^\infty}^{t_{i+1}^\infty} e^{-\alpha t} \int_V [f(x_i^m(t), n_i^m(t), v) - f(x_i^\infty(t), n_i^\infty(t), v)] \delta_{v^m(t)}(dv) dt \\
 &\quad + \sum_{i=0}^\infty \int_{t_i^\infty}^{t_{i+1}^\infty} e^{-\alpha t} \int_V f(x_i^\infty(t), n_i^\infty(t), v) \delta_{v^m(t)}(dv) dt,
 \end{aligned}$$

where the paths $t \mapsto (x_i^m(t), n_i^m(t))$, $i \geq 0$, are introduced by (6.1)–(6.4) in the Appendix, whereas the sequence $\{t_i^\infty\} \subset [0, T]$ is fixed (actually, it is generated according to the proof of Proposition 3.2 in the aforementioned appendix).

By the proof of Proposition 3.2, $(x_i^m(\cdot), n_i^m(\cdot)) \rightarrow (x_i^\infty(\cdot), n_i^\infty(\cdot))$ uniformly as $m \rightarrow \infty$. Hence, as $f \in C_b^u(S \setminus D \times V)$, we have $f(x_i^m(\cdot), n_i^m(\cdot), v) \rightarrow f(x_i^\infty(\cdot), n_i^\infty(\cdot), v)$ on $[t_i^\infty, t_{i+1}^\infty[$ as $m \rightarrow \infty$ for all $v \in V$. Using the latter property, we can easily verify that, for all $t \in [t_i^\infty, t_{i+1}^\infty[$,

$$\int_V [f(x_i^m(t), n_i^m(t), v) - f(x_i^\infty(t), n_i^\infty(t), v)] \delta_{v^m(t)}(dv) \rightarrow 0 \quad \text{as } m \rightarrow \infty.$$

Then, by the dominated convergence theorem,

$$\sum_{i=0}^\infty \int_{t_i^\infty}^{t_{i+1}^\infty} e^{-\alpha t} \int_V [f(x_i^m(t), n_i^m(t), v) - f(x_i^\infty(t), n_i^\infty(t), v)] \delta_{v^m(t)}(dv) dt \rightarrow 0$$

as $m \rightarrow \infty$. Moreover, as the trajectory $(x^\infty(\cdot), n^\infty(\cdot))$ is continuous on $[t_i^\infty, t_{i+1}^\infty[$, the mapping $(t, v) \mapsto f(x^\infty(t), n^\infty(t), v)$ is continuous on $[t_i^\infty, t_{i+1}^\infty[$ too. Now, using the fact that $\delta_{v^m(\cdot)} \xrightarrow{w} \hat{v}$, we deduce that for all $t \in [t_i^\infty, t_{i+1}^\infty[$,

$$(3.9) \quad \int_V f(x_i^\infty(t), n_i^\infty(t), v) \delta_{v^m(t)}(dv) \rightarrow \int_V f(x_i^\infty(t), n_i^\infty(t), v) \hat{v}_t(dv).$$

Hence, by the dominated convergence theorem,

$$\begin{aligned}
 & \sum_{i=0}^\infty \int_{t_i^\infty}^{t_{i+1}^\infty} e^{-\alpha t} \int_V f(x_i^\infty(t), n_i^\infty(t), v) \delta_{v^m(t)}(dv) dt \\
 & \rightarrow \sum_{i=0}^\infty \int_{t_i^\infty}^{t_{i+1}^\infty} e^{-\alpha t} \int_V f(x_i^\infty(t), n_i^\infty(t), v) \hat{v}_t(dv) dt \quad \text{as } m \rightarrow \infty.
 \end{aligned}$$

Thus, we have proved that

$$(3.10) \quad \int_0^\infty e^{-\alpha t} \int_V f(x^m(t), n^m(t), v) \delta_{v^m(t)}(dv) dt \\ \rightarrow \int_0^\infty e^{-\alpha t} \int_V f(x^\infty(t), n^\infty(t), v) \hat{v}_t(dv) dt \quad \text{as } m \rightarrow \infty.$$

On the other hand, the continuity of the mapping $(x, n, v) \mapsto \ell(x, n, k)$, the uniform convergence $(x_i^m(\cdot), n_i^m(\cdot)) \rightarrow (x_i^\infty(\cdot), n_i^\infty(\cdot))$ on $[t_i^\infty, t_{i+1}^\infty[$, and the convergences $t_i^m \rightarrow t_i^\infty$ and $\{k_i^m\} \rightarrow \{\hat{k}_i\}$ established in Proposition 3.2 and in its proof ensure that

$$e^{-\alpha t_i^m} \ell(x_i^m(t_i^m -), n_i^m(t_i^m -), k_i^m) \rightarrow e^{-\alpha t_i^\infty} \ell(x_i^\infty(t_i^\infty -), n_i^\infty(t_i^\infty -), \hat{k}_i).$$

Using the dominated convergence theorem again, we deduce

$$(3.11) \quad \sum_{i=0}^\infty e^{-\alpha t_i^m} \ell(x_i^m(t_i^m -), n_i^m(t_i^m -), k_i^m) \\ \rightarrow \sum_{i=0}^\infty e^{-\alpha t_i^\infty} \ell(x_i^\infty(t_i^\infty -), n_i^\infty(t_i^\infty -), \hat{k}_i).$$

Thus, relying on (3.10) and (3.11), we conclude that

$$(3.12) \quad J(x, n, \delta_{v^m(\cdot)}(\cdot), \{k_i^m\}) = \int_0^\infty \int_V e^{-\alpha t} f(x^m(t), n^m(t), v) \delta_{v^m(t)}(dv) dt \\ + \sum_{i=0}^\infty e^{-\alpha t_i^m} \ell(x_i^m(t_i^m -), n_i^m(t_i^m -), k_i^m) \\ \rightarrow \int_0^\infty \int_V e^{-\alpha t} f(x^\infty(t), n^\infty(t), v) \hat{v}_t(dv) dt \\ + \sum_{i=0}^\infty e^{-\alpha t_i^\infty} \ell(x_i^\infty(t_i^\infty -), n_i^\infty(t_i^\infty -), \hat{k}_i) \\ = J(x, n, \hat{v}, \{\hat{k}_i\}) = u(x, n),$$

which proves the result. ■

4. General case. In this section we drop the transversality condition (3.5). This implies that the continuity of the trajectories established in Proposition 3.2 may no longer be valid. So different arguments must be applied.

In this section we do the following:

- We prove that, for each $t \geq 0$, the family $\{(x^m(t), n^m(t))\}_m$ of trajectories associated to a sequence $\{(\mathbf{v}^m, \{k_i^m\})\}_m \subset \mathbf{V} \times \mathcal{K}$ of controls is bounded for all $m \geq 1$.
- We define the occupation measures associated to the trajectories $(x(\cdot), n(\cdot))$.
- We prove the pre-compactness of the occupation measures with respect to a large set of finite measures satisfying suitable properties.
- We establish the existence of an optimal measure under which $u(x, n)$ attains the minimum.

For this purpose, similar to (3.2), the expression (2.9) can be written as follows: for each $\varphi \in C_b^{1,0}(S)$, $t \geq 0$, and $\mathbf{v} \in \mathbf{V}$,

$$\begin{aligned}
 (4.1) \quad & e^{-\alpha t} \varphi(x(t), n(t)) - \varphi(x, n) \\
 &= \int_0^t e^{-\alpha s} \left[\int_V g(x(s), n(s), v) \mathbf{v}_s(dv) \cdot \nabla_x \varphi(x(s), n(s)) - \alpha \varphi(x(s), n(s)) \right] ds \\
 &+ \sum_{i=0}^{\infty} e^{-\alpha t_i} [\varphi(X(x(t_i-), n(t_i-), k_i), N((x(t_i-), n(t_i-), k_i)) \\
 &\quad - \varphi(x(t_i-), n(t_i-))] \mathbf{1}_{\{t_i \leq t\}}.
 \end{aligned}$$

For $\lambda \geq 1$, we define

$$(4.2) \quad c_0 \equiv c_0(\lambda) := \sup \left\{ \frac{x \cdot g(x, n, v)}{\lambda + |x|^2 + |n|^2} : (x, n) \in S \setminus D, v \in V \right\},$$

where g is the vector field in (3.2). Since g is bounded by the constant M (see (2.5)),

$$\frac{x \cdot g(x, n, v)}{\lambda + |x|^2 + |n|^2} \leq \frac{|x|M}{\lambda + |x|^2 + |n|^2} \leq \frac{|x|}{\sqrt{\lambda + |x|^2 + |n|^2}} \cdot \frac{M}{\sqrt{\lambda + |x|^2 + |n|^2}},$$

i.e., $c_0 \leq M\lambda^{-1/2}$. This implies that given the discount factor $\alpha > 0$ in (2.11) or in (3.3), there exists $\lambda > 0$ large enough that $c_0 < \alpha$, and this is our choice of $\lambda > 0$ for all what follows.

We also define

$$\begin{aligned}
 (4.3) \quad & c_1 \equiv c_1(\lambda) := \\
 & \sup \{ [(\lambda + |X(x, n, k)|^2 + |N(x, n, k)|^2)^{1/2} - (\lambda + |x|^2 + |n|^2)^{1/2}] \\
 & \quad \times (|X(x, n, k) - x|^2 + |N(x, n, k) - n|^2)^{-1/2} : (x, n) \in D, k \in K \}.
 \end{aligned}$$

From the estimate

$$(4.4) \quad |(\lambda + |x|^2 + |n|^2)^{1/2} - (\lambda + |x'|^2 + |n'|^2)^{1/2}| \leq (|x - x'|^2 + |n - n'|^2)^{1/2},$$

we can deduce that $c_1 \in [-1, 1]$.

where for all $(x, n) \in S$, $v \in V$, and $k \in K$,

$$(4.7) \quad \begin{aligned} A^v \varphi(x, n) &:= -g(x, n, v) \cdot \nabla_x \varphi(x, n) + \alpha \varphi(x, n), \\ L^k \varphi(x, n) &:= \varphi(x, n) - \varphi(X(x, n, k), N(x, n, k)). \end{aligned}$$

It is easy to see that under assumptions (2.5) and (2.6), the mapping $(x, n, v) \mapsto A^v \varphi(x, n)$ is in $C_b(S \setminus D \times V)$ and $(x, n, k) \mapsto L^k \varphi(x, n)$ is in $C_b(D \times K)$, for all $\varphi \in C_b^{1,0}(S)$.

Now we introduce the so-called occupation measures. For ease of notation, any sequence $\{k_i\} \in \mathcal{K}$ will be denoted by \mathbf{k} .

Occupation measures. For each control pair $(\mathbf{v}, \mathbf{k}) \in \mathbf{V} \times \mathcal{K}$ and each initial condition (x, n) , we define the *occupation measures*

$$(4.8) \quad \begin{aligned} \mu_{(x,n)}^{\mathbf{v},\mathbf{k}}(\mathbf{A} \times \mathbf{B}) &:= \int_0^\infty \int_V e^{-\alpha t} \mathbf{1}_A(x(t), n(t)) \mathbf{1}_B(v) \nu_t(dv) dt, \\ \nu_{(x,n)}^{\mathbf{v},\mathbf{k}}(\mathbf{A} \times \mathbf{B}) &:= \sum_{i=0}^\infty e^{-\alpha t_i} \mathbf{1}_A(x(t_i-), n(t_i-)) \mathbf{1}_B(k_i), \end{aligned} \quad \forall \mathbf{A} \times \mathbf{B} \subseteq S \setminus D \times V, \quad \forall \mathbf{A} \times \mathbf{B} \subseteq D \times K.$$

We denote the sets of all occupation measures by

$$\begin{aligned} M_{(x,n)}(S \setminus D \times V) &:= \{\mu_{(x,n)}^{\mathbf{v},\mathbf{k}} : (\mathbf{v}, \mathbf{k}) \in \mathbf{V} \times \mathcal{K}\}, \\ M_{(x,n)}(D \times K) &:= \{\nu_{(x,n)}^{\mathbf{v},\mathbf{k}} : (\mathbf{v}, \mathbf{k}) \in \mathbf{V} \times \mathcal{K}\}. \end{aligned}$$

Observe that

$$(4.9) \quad 0 \leq \mu_{(x,n)}^{\mathbf{v},\mathbf{k}} \leq \frac{1}{\alpha} \quad \text{and} \quad 0 \leq \nu_{(x,n)}^{\mathbf{v},\mathbf{k}}(x, n) \leq \frac{1}{1 - e^{-\alpha \mathbf{h}}}, \quad \forall (\mathbf{v}, \mathbf{k}) \in \mathbf{V} \times \mathcal{K},$$

where \mathbf{h} is the constant in Proposition 2.4. In fact, the normalized measures

$$(4.10) \quad \tilde{\mu}_{(x,n)}^{\mathbf{v},\mathbf{k}} := \alpha \mu_{(x,n)}^{\mathbf{v},\mathbf{k}} \quad \text{and} \quad \tilde{\nu}_{(x,n)}^{\mathbf{v},\mathbf{k}} := \hat{S}_{\mathbf{v},\mathbf{k}} \nu_{(x,n)}^{\mathbf{v},\mathbf{k}},$$

with $\hat{S}_{\mathbf{v},\mathbf{k}} := [1 / \sum_{i=0}^\infty e^{-\alpha t_i^{(\mathbf{v},\mathbf{k})}}]$, turn out to be probability measures on $S \setminus D \times V$ and $D \times K$, respectively.

DEFINITION 4.2.

- (a) Let $M_b(S \setminus D \times V)$ (resp. $M_b(D \times K)$) be the space of all signed finite measures on $S \setminus D \times V$ (resp. on $D \times K$).
- (b) Denote by $M_b^+(S \setminus D \times V)$ (resp. $M_b^+(D \times K)$) the subset of all non-negative elements of $M_b(S \setminus D \times V)$ (resp. of $M_b(D \times K)$).

Consider now the following problem:

(W) minimize $\langle (\mu, \nu), (f, \ell) \rangle$ subject to

$$\begin{aligned}
 & \langle \delta_{(x,n)}, \varphi \rangle = \langle (\mu, \nu), (A^v \varphi, L^k \varphi) \rangle, \quad \forall \varphi \in C_b^{1,0}(S), \\
 (4.11) \quad & (\mu, \nu) \leq \left(\frac{1}{\alpha}, \frac{1}{1 - e^{-\alpha h}} \right), \\
 & (\mu, \nu) \in M_b^+(S \setminus D \times V) \times M_b^+(D \times K).
 \end{aligned}$$

Note that the relation (4.6) can be rewritten in terms of the occupation measures as follows:

$$\begin{aligned}
 (4.12) \quad & \langle \delta_{(x,n)}, \varphi \rangle = \langle \mu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, A^v \varphi \rangle + \langle \nu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, L^k \varphi \rangle \\
 & = \langle (\mu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, \nu_{(x,n)}^{\mathbf{v}, \mathbf{k}}), (A^v \varphi, L^k \varphi) \rangle
 \end{aligned}$$

for all $\varphi \in C_b^{1,0}(S)$.

From (4.9) and (4.12), every occupation measure $(\mathbf{v}, \mathbf{k}) \mapsto (\mu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, \nu_{(x,n)}^{\mathbf{v}, \mathbf{k}})$ satisfies the restrictions of the weak problem (W). In terms of occupation measures the total cost (3.3) can be expressed as

$$(4.13) \quad J(x, n, \mathbf{v}(\cdot), \{k_i\}) = \langle \mu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, f \rangle + \langle \nu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, \ell \rangle = \langle (\mu_{(x,n)}^{\mathbf{v}, \mathbf{k}}, \nu_{(x,n)}^{\mathbf{v}, \mathbf{k}}), (f, \ell) \rangle$$

for all $(\mathbf{v}, \mathbf{k}) \in \mathbf{V} \times \mathcal{K}$.

We shall endow the spaces $M_b(S \setminus D \times V)$ and $M_b(D \times K)$ with the standard (Prokhorov) weak convergence: we say that a sequence $\{\mu_m\}$ of measures in $M_b(S \setminus D \times V)$ converges to some $\mu \in M_b(S \setminus D \times V)$ and write $\mu_m \xrightarrow{w} \mu$ if and only if

$$\int_{S \setminus D \times V} h(x, n, v) \mu_m(d(x, n, v)) \xrightarrow{m \rightarrow \infty} \int_{S \setminus D \times V} h(x, n, v) \mu(d(x, n, v))$$

for all $h \in C_b(S \setminus D \times V)$. Similarly, we can define the above convergence in $M_b(D \times K)$.

The following proposition ensures that the sets $M_{(x,n)}(S \setminus D \times V)$ and $M_{(x,n)}(D \times K)$ of occupation measures are pre-compact (in a weak sense) in the spaces $M_b(S \setminus D \times V)$ and $M_b(D \times K)$.

PROPOSITION 4.3. *The sets $M_{(x,n)}(S \setminus D \times V)$ and $M_{(x,n)}(D \times K)$ are weakly pre-compact: for any sequence $\{\mu_{(x,n)}^{\mathbf{v}, \mathbf{k}, m}\}_m \subset M_{(x,n)}(S \setminus D \times V)$, there exists $\mu \in M_b(S \setminus D \times V)$ such that*

$$\mu_{(x,n)}^{\mathbf{v}, \mathbf{k}, m} \xrightarrow{w} \mu \quad \text{as } m \rightarrow \infty$$

along a subsequence. Similarly, for any sequence $\{\nu_{(x,n)}^{\mathbf{v}, \mathbf{k}, m}\}_m \subset M_{(x,n)}(D \times K)$, there exists $\nu \in M_b(D \times K)$ such that, along a subsequence,

$$\nu_{(x,n)}^{\mathbf{v}, \mathbf{k}, m} \xrightarrow{w} \nu \quad \text{as } m \rightarrow \infty.$$

Proof. We proceed by showing that the family of normalized measures (4.10) is tight, which will imply the tightness of the original family (4.8). To this end, fix $T > 0$ and define the compact set

$$\bar{B}_{C_T}(x, n) := \{(z, r) \in S : |z - x| + |r - n| \leq C_T\},$$

with C_T as in (4.5) at $t = T$, and where (x, n) denotes the initial state of the dynamic (3.2). Now take a sequence $\{\mu_{(x,n)}^{v,k,m}\}_m \subset M_{(x,n)}((S \setminus D) \times V)$ and denote by $(x^m(\cdot), n^m(\cdot))$ the corresponding trajectory. Observe that by Proposition 4.1, the path $s \mapsto (x^m(s), n^m(s))$ with $s \in [0, T]$ lies in $\bar{B}_{C_T}(x, n)$ for all $m \geq 1$; in fact, $(x^m(s), n^m(s), v) \in \bar{B}_{C_T}(x, n) \times V$. Then, for every $\varepsilon > 0$, we can choose T sufficiently large such that

$$\begin{aligned} \tilde{\mu}_{(x,n)}^{v,k,m}(\bar{B}_{C_T}(x, n) \times V) &= \alpha \int_0^\infty \int_V e^{-\alpha s} \mathbf{1}_{\bar{B}_{C_T}(x,n)}(x^m(s), n^m(s)) \mathbf{1}_V(v) \mathbf{v}_s(dv) ds \\ &= \alpha \int_0^T e^{-\alpha s} ds = (1 - e^{-\alpha T}) > 1 - \varepsilon, \quad \forall m \geq 1. \end{aligned}$$

Similarly,

$$\begin{aligned} \tilde{\nu}_{(x,n)}^{v,k,m}(\bar{B}_{C_T}(x, n) \times K) &= \hat{S}_{v_m, k_m} \sum_{i=0}^\infty e^{-\alpha t_i^m} \mathbf{1}_{\bar{B}_{C_T}(x,n)}(x^m(t_i-), n^m(t_i-)) \mathbf{1}_K(k_i) \\ &= \hat{S}_{v_m, k_m} \sum_{i=0}^\infty e^{-\alpha t_i^m} \mathbf{1}_{\{t_i \leq T\}} > 1 - \varepsilon, \quad \forall m \geq 1, \end{aligned}$$

where the last inequality follows from $\hat{S}_{v_m, k_m} \sum_{i=0}^\infty e^{-\alpha t_i^m} \mathbf{1}_{\{t_i \leq T\}} \uparrow 1$ as $T \rightarrow \infty$. Then by definition both sequences $\{\tilde{\mu}_{(x,n)}^{v,k,m}\}$ and $\{\tilde{\nu}_{(x,n)}^{v,k,m}\}$ are tight, and thus relatively compact by Prokhorov's theorem. Of course this property of the normalized measures leads to the same property of the occupation measures, which proves the result. ■

The next theorem gives a characterization of the value function $u(x, n)$ in terms of a pair of measures, say (μ, ν) , that satisfies the restrictions of problem (W).

THEOREM 4.4. *Under the assumptions (2.4)–(2.7) and (2.10), there exists $(\hat{\mu}, \hat{\nu}) \in M_b(S \setminus D \times V) \times M_b(D \times K)$ satisfying the restrictions (4.11) of problem (W) and such that the value function $u(x, n)$ in (2.12) becomes $u(x, n) = \langle (\hat{\mu}, \hat{\nu}), (f, \ell) \rangle$.*

Proof. Choose a minimizing sequence of controls $(\delta_{v_m(\cdot)}(\cdot), \{k_i^m\}) \in \mathcal{V} \times \mathcal{K}$ so that

$$(4.14) \quad J(x, n, v_m(\cdot), \{k_i^m\}) = J(x, n, \delta_{v_m(\cdot)}(\cdot), \{k_i^m\}) \downarrow u(x, n) \quad \text{as } m \rightarrow \infty.$$

By (4.13), we can relate to $J((x, n, \delta_{v_m(\cdot)}(\cdot), \{k_i^m\}))$ the occupation measures $(\mu_m, \nu_m) \in M_b(S \setminus D \times V) \times M_b(D \times K)$ defined in (4.8). Moreover, every pair $(\delta_{v_m(\cdot)}(\cdot), \{k_i^m\}) \in \mathcal{V} \times \mathcal{K}$ produces a trajectory given in (2.1), which in turn is equivalent to (2.9). Hence, in virtue of (4.12) together with the fact that (μ_m, ν_m) are in fact measures (and therefore are nonnegative), we deduce that each element of the above sequence satisfies the restrictions of (4.11).

By compactness of the control spaces, we know that $\delta_{v_m(\cdot)} \xrightarrow{w} \hat{v}$ and $\{k_i^m\} \rightarrow \{\hat{k}_i\}$ for some $\hat{v} \in \mathbb{V}$ and $\{\hat{k}_i\} \in \mathcal{K}$, so we can invoke Proposition 4.3 to deduce the existence of $(\hat{\mu}, \hat{\nu}) \in M_b(S \setminus D \times V) \times M_b(D \times K)$ such that $(\mu_m, \nu_m) \xrightarrow{w} (\hat{\mu}, \hat{\nu})$ as $m \rightarrow \infty$. From (4.14), this limit pair is the one associated with the value function $u(x, n)$, i.e., $u(x, n) = \langle (\hat{\mu}, \hat{\nu}), (f, \ell) \rangle$.

It only remains to prove that $(\hat{\mu}, \hat{\nu})$ satisfies the restrictions in (4.11). To do so, note again that for every $m \geq 1$,

$$(4.15) \quad \langle \delta_{(x,n)}, \varphi \rangle = \langle (\mu_m, \nu_m), (A^v \varphi, L^k \varphi) \rangle, \quad \forall \varphi \in C_b^{1,0}(S).$$

Since the mapping $(x, n, v) \mapsto A^v \varphi(x, n)$ is in $C_b(S \setminus D \times V)$ and $(x, n, k) \mapsto L^k \varphi(x, n)$ is in $C_b(D \times K)$, for all $\varphi \in C_b^{1,0}(S)$, and since $(\mu_m, \nu_m) \xrightarrow{w} (\hat{\mu}, \hat{\nu})$ as $m \rightarrow \infty$, we deduce that

$$\begin{aligned} \langle \delta_{(x,n)}, \varphi \rangle &= \lim_{m \rightarrow \infty} \langle (\mu_m, \nu_m), (A^v \varphi, L^k \varphi) \rangle \\ &= \langle (\hat{\mu}, \hat{\nu}), (A^v \varphi, L^k \varphi) \rangle, \quad \forall \varphi \in C_b^{1,0}(S), \end{aligned}$$

and

$$0 \leq \lim_m \mu_m = \hat{\mu} \leq \frac{1}{\alpha} \quad \text{and} \quad 0 \leq \lim_m \nu_m = \hat{\nu} \leq \frac{1}{1 - e^{-\alpha h}}.$$

This proves the result. ■

5. Linear programs. This section is devoted to showing that the control problems (2.12) and (3.4) are equivalent for the regular case of Section 3. Furthermore, we also provide a second equivalence of the former problem with the minimization of an ancillary infinite-dimensional linear program; in particular, the restrictions of the corresponding dual counterpart of this linear program will be of great importance. This material could also be interesting from the point of view of approximations since infinite-dimensional linear programs can be analyzed through approximations of finite-dimensional linear programs—see, for instance, Lasserre [11]. In this way, our results in this section could provide a guide to approximating the value function u in (2.12) by means of finite-dimensional linear programs; such an analysis is, however, outside the scope of this paper.

It is also important to mention that along this section we shall use a regularity-type condition on the value function u that holds under the transversality condition (3.5).

We summarize this section into the following four main facts:

- Problem (W) (see (4.11)) is embedded into a linear program (P) (see (5.3)).
- Problem (P) has an associated dual counterpart (P*) (see (5.4)).
- Under an extra assumption on the value function u (see (5.13)), we prove $\inf(P) \geq u(x, n)$. As a consequence, problems (2.12) and (3.4) are equivalent: $u(x, n) = \mathbf{u}(x, n) = \min(W) = \inf(P)$.
- The transversality condition (3.5) implies that assumption (5.13) is satisfied.

Dual pairs. General results in duality theory show that, for some subset \mathcal{X} of \mathbb{R}^j , with j as in Definition 2.5, the topological (strong) dual of $C_b(\mathcal{X})$ turns out to be $M_b(\mathcal{X})$, with C_b and M_b corresponding to the sets in Definitions 2.5(b) and 4.2, respectively. It follows that these spaces define a dual pair under the duality

$$(5.1) \quad \langle M_b(\mathcal{X}), C_b(\mathcal{X}) \rangle_{\mathcal{X}} = \int_{\mathcal{X}} g \, d\eta, \quad \forall g \in C_b(\mathcal{X}), \eta \in M_b(\mathcal{X}).$$

The above spaces are Banach spaces under the associated norms; nevertheless, hereafter, convergence in $M_b(\mathcal{X})$ is assumed to hold in the weak topology $\sigma(C_b(\mathcal{X}), M_b(\mathcal{X}))$: $\mu_n \xrightarrow{w} \mu$ means $\langle \mu_n, g \rangle \rightarrow \langle \mu, g \rangle$ for all $g \in C_b(\mathcal{X})$.

On the other hand, for $\mathcal{X} \equiv S$, recall the space $C_b^{1,0}(S)$ introduced in Definition 2.5(c). We shall denote its (algebraic) dual by $\mathcal{D}_b(S)$. These spaces endowed with the weak topologies $\sigma(C_b^{1,0}(S), \mathcal{D}_b(S))$ and $\sigma(\mathcal{D}_b(S), C_b^{1,0}(S))$ become a dual pair under the duality

$$\langle \mathcal{D}_b(S), C_b^{1,0}(S) \rangle_S = \int_S h \, d\varrho, \quad \forall h \in C_b^{1,0}(S), \varrho \in \mathcal{D}_b(S).$$

Observe that $\delta_{(x,n)}(\cdot) \in \mathcal{D}_b(S)$ for all $(x, n) \in S$.

Now recall the pair (A^v, L^k) introduced in (4.7). This pair, regarded as a single operator, maps $C_b^{1,0}(S)$ into $C_b(S \setminus D \times V) \times C_b(D \times K)$. A useful property of this operator is

PROPOSITION 5.1. *The operator (A^v, L^k) is continuous with respect to the norm induced by $C_b(S \setminus D \times V) \times C_b(D \times K)$, defined by*

$$\|(g_1, g_2)\|_* := \max\{\|g_1\|, \|g_2\|\}, \quad \forall g_1 \in C_b(S \setminus D \times V), g_2 \in C_b(D \times K).$$

Proof. It is easy to verify that

$$\|A^v\varphi\| \leq (\|g\| + \alpha)\|\varphi\|_1, \quad \|L^k\varphi\| \leq 2\|\varphi\|_1, \quad \forall \varphi \in C_b^{1,0}(S).$$

Then

$$\|(A^v, L^k)\varphi\|_* = \max\{\|A^v\varphi\|, \|L^k\varphi\|\} \leq (\|g\| + \alpha + 2)\|\varphi\|_1.$$

This proves the result. ■

For each (μ, η) in the product space $M_b((S \setminus D) \times V) \times M_b(D \times K)$, define the functional $\Phi_{(\mu, \eta)}$ on $C_b^{1,0}(S)$ as follows:

$$\Phi_{(\mu, \eta)}\varphi := \langle (\mu, \eta), (A^v, L^k)\varphi \rangle = \langle \mu, A^v\varphi \rangle + \langle \eta, L^k\varphi \rangle, \quad \forall \varphi \in C_b^{1,0}(S).$$

Since (A^v, L^k) is continuous, so is $\Phi_{(\mu, \eta)}$. This implies the existence of $\nu_{(\mu, \eta)} \in \mathcal{D}_b(S)$ such that

$$\Phi_{(\mu, \eta)}\varphi = \langle \nu_{(\mu, \eta)}, \varphi \rangle = \langle (\mu, \eta)(A^v, L^k)\varphi \rangle, \quad \forall \varphi \in C_b^{1,0}(S).$$

Since this holds for every $(\mu, \eta) \in M_b((S \setminus D) \times V) \times M_b(D \times K)$, we can define the operator $(A^v, L^k)^* : M_b((S \setminus D) \times V) \times M_b(D \times K) \rightarrow \mathcal{D}_b(S)$ by

$$(A^v, L^k)^*(\mu, \eta) := \nu_{(\mu, \eta)}, \quad \forall (\mu, \eta) \in M_b((S \setminus D) \times V) \times M_b(D \times K).$$

It is clear that $(A^v, L^k)^*$ is the adjoint of (A^v, L^k) , because

$$(5.2) \quad \langle (A^v, L^k)^*(\mu, \eta), \varphi \rangle = \langle (\mu, \eta), (A^v, L^k)\varphi \rangle$$

for all $(\mu, \eta) \in M_b(S \setminus D \times V) \times M_b(D \times K)$ and $\varphi \in C_b^{1,0}(S)$.

Since $\mathcal{D}_b(S)$ is the algebraic dual of $C_b^{1,0}(S)$, by construction of the adjoint operator $(A^v, L^k)^*$ we see that it maps $M_b(S \setminus D \times V) \times M_b(D \times K)$ into $\mathcal{D}_b(S)$. A characterization of this last assertion is given in the following result (a proof can be found e.g. in Aliprantis and Border [1, Theorem 6.43]).

PROPOSITION 5.2. *The operator $(A^v, L^k)^*$ is weakly continuous, i.e., it is continuous with respect to the weak topologies $\sigma(M_b(S \setminus D \times V) \times M_b(D \times K), C_b(S \setminus D \times V) \times C_b(D \times K))$ and $\sigma(\mathcal{D}_b(S), C_b^{1,0}(S))$.*

Cones. Now we define the natural cones of the space $C_b(S \setminus D \times V) \times C_b(S \times K)$:

$$\begin{aligned} & [C_b(S \setminus D \times V) \times C_b(S \times K)]^+ \\ & := \{(\varphi_1, \varphi_2) \in C_b(S \setminus D \times V) \times C_b(D \times K) : \varphi_1, \varphi_2 \geq 0\}. \end{aligned}$$

We also define the corresponding dual cone as

$$\begin{aligned} & [M_b(S \setminus D \times V) \times M_b(D \times K)]^+ \\ & := \{(\mu_1, \mu_2) \in M_b(S \setminus D \times V) \times M_b(D \times K) : \langle (\mu_1, \mu_2), (\varphi_1, \varphi_2) \rangle \geq 0, \\ & \quad \forall (\varphi_1, \varphi_2) \in [C_b(S \setminus D \times V) \times C_b(S \times K)]^+\}. \end{aligned}$$

With all the previous definitions, we define the linear program

(P) minimize $\langle (\mu, \nu), (f, \ell) \rangle$ subject to

$$(5.3) \quad \begin{aligned} \delta_{(x,n)} &= (A^v, L^k)^*(\mu, \nu), \\ (\mu, \nu) &\in [M_b(S \setminus D \times V) \times M_b(D \times K)]^+. \end{aligned}$$

Note that $\langle (\mu, \nu), (f, \ell) \rangle$ means $\langle \mu, f \rangle_{S \setminus D \times V} + \langle \nu, \ell \rangle_{D \times K}$ (see (5.1)).

The dual problem of (P) turns out to be

(P*) maximize $\langle \delta_{(x,n)}, \varphi \rangle$ subject to

$$(5.4) \quad \begin{aligned} (f, \ell) - (A^v, L^k)\varphi &\in [C_b(S \setminus D \times V) \times C_b(S \times K)]^+, \\ \varphi &\in C_b^{1,0}(S). \end{aligned}$$

The latter restriction can be seen as

$$(5.5) \quad A^v \varphi \leq f, \quad L^k \varphi \leq \ell.$$

5.1. Consistency. It is obvious that the dual problem is feasible (i.e., the restrictions are nonempty). Indeed, take the constant function $0 \in C_b^{1,0}$ and use the fact that the costs f and g are nonnegative. On the other hand, we have already verified that the weak problem (W) is feasible, hence so is (P). This implies that problems (P) and (P*) are both consistent. Then we define the value of the primal problem (P) to be

$$(5.6) \quad \inf(\text{P}) := \inf\{\langle (\mu, \nu), (f, \ell) \rangle : (\mu, \nu) \text{ is feasible for (P)}\}.$$

In a similar manner, the value of the dual problem (P*) is defined by

$$(5.7) \quad \sup(\text{P}^*) := \sup\{\langle \delta_{(x,n)}, \varphi \rangle : \varphi \text{ is feasible for (P}^*)\}.$$

Since problem (W) satisfies the restrictions (5.3), we have $\inf(\text{W}) \geq \inf(\text{P})$.

The following result ensures the so-called weak duality between the linear programs (P) and (P*). The proof is provided for general infinite-dimensional linear spaces by Anderson and Nash [2].

PROPOSITION 5.3 (Weak duality). *The values of (P) and (P*) are finite and they satisfy*

$$(5.8) \quad \sup(\text{P}^*) \leq \inf(\text{P}).$$

REMARK 5.4. It is possible to get equality in (5.8) under an extra condition on the cost functions. This condition must guarantee lower bounds $f \geq c_0$ and $\ell \geq c_0$ for some positive constant c_0 .

Relation of (P) and the value function u . First consider those functions $\hat{f}(x, n, v) \in C_b^u(S \setminus D \times V)$ and $\hat{\ell}(x, n, k) \in C_b^u(D \times K)$ which are Lipschitz continuous in x uniformly with respect to the other variables, i.e., there is a positive constant $M_{\hat{f}, \hat{\ell}}$, such that, for all $x, y \in \mathbb{R}^d$,

$$(5.9) \quad \sup_{n,v} |\hat{f}(x, n, v) - \hat{f}(y, n, v)| + \sup_{n,k} |\hat{\ell}(x, n, k) - \hat{\ell}(y, n, k)| \leq M_{\hat{f}, \hat{\ell}} |x - y|.$$

For any initial condition $(x, n) \in S$, and each pair of controls $(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}$, let us associate to \hat{f} and $\hat{\ell}$ the total cost

$$(5.10) \quad \hat{J}(x, n; v(\cdot), \{k_i\}) = \int_0^\infty e^{-\alpha t} \hat{f}(x(t), n(t), v(t)) dt + \sum_{i=0}^\infty e^{-\alpha t_i} \hat{\ell}(x(t_i-), n(t_i-), k_i),$$

as well as the corresponding value function

$$(5.11) \quad \hat{u}(x, n) = \inf_{(v(\cdot), \{k_i\}) \in \mathcal{V} \times \mathcal{K}} \hat{J}(x, n; v(\cdot), \{k_i\}).$$

Now let us go back to the original costs (f, ℓ) defined in (2.10). By the properties of these functions, it is easy to see that the mappings $x \mapsto f(x, n, v)$ and $x \mapsto \ell(x, n, k)$ can be approximated by Lipschitz continuous functions uniformly on compact sets.

So denote by $\{(f^m, \ell^m)\}$ a sequence of Lipschitz continuous functions satisfying

$$(5.12) \quad \begin{aligned} & \sup_{x \in \hat{X}} \sup_{(n,v) \in \mathcal{N} \times V} |f^m(x, n, v) - f(x, n, v)| \rightarrow 0, \\ & \sup_{x \in \hat{X}} \sup_{(n,k) \in \mathcal{N} \times \mathcal{K}} |\ell^m(x, n, k) - \ell(x, n, k)| \rightarrow 0, \\ & \text{as } m \rightarrow \infty, \text{ for every compact set } \hat{X} \subset \mathbb{R}^d. \end{aligned}$$

Let us impose the following condition on the value functions $u^m(x, n)$ associated to the elements of the above convergent sequence:

$$(5.13) \quad \text{There exists a sequence } \{(f^m, \ell^m)\} \text{ satisfying (5.9) and (5.12) under which the corresponding value function } u^m(x, n) \text{ in (5.11) is Lipschitz continuous in } x, \text{ uniformly in } n \in \mathcal{N}, \text{ with Lipschitz constant } M_u^m.$$

REMARK 5.5. Assumption (5.13) may seem to be a little strong, but we will see later that there are situations (such as those when the boundary of the set-interface D is regular) where this condition turns out to be a consequence of our present assumptions.

We now present an ancillary result that is based on the dynamic programming principle. For a proof, see Bensoussan and Menaldi [6, Corollary 3.8].

LEMMA 5.6. *For any pair of functions f^m, ℓ^m satisfying (5.13), the corresponding value function u^m in (5.11) satisfies the following system of quasi-variational inequalities (QVI):*

$$(5.14) \quad \begin{aligned} 0 & \leq f^m(x, n, v) + \langle g(x, n, v), \partial_x u^m(x, n) \rangle - \alpha u^m(x, n), \\ 0 & \leq \ell^m(x, n, k) + u^m(X(x, n, k), N(x, n, k)) - u^m(x, n), \end{aligned}$$

for almost all $x \in \mathbb{R}^d$ and all n, v, k .

Given some $\varepsilon > 0$, let $\varrho_\varepsilon : \mathbb{R}^m \rightarrow \mathbb{R}$ satisfy

- (i) $\varrho_\varepsilon \geq 0$,
- (ii) $\varrho_\varepsilon(x) = 0, \forall |x| \geq \varepsilon$,
- (iii) $\int_{B_\varepsilon(x)} \varrho_\varepsilon(x) dx = 1$,
- (iv) ϱ_ε is infinitely differentiable.

Now consider the convolution of u^m and ϱ_ε :

$$(5.15) \quad \begin{aligned} u_\varepsilon^m(x, n) &= \varrho_\varepsilon * u^m(x, n) \\ &:= \int_{B_\varepsilon(x)} \varrho_\varepsilon(x - y) u^m(y, n) dy = \int_{B_\varepsilon(0)} \varrho_\varepsilon(y) u^m(x - y, n) dy. \end{aligned}$$

As a direct consequence of the definition, $\|u_\varepsilon^m\| \leq \|u^m\|$. The next result regards the regularity of u_ε^m .

LEMMA 5.7. *The function u_ε^m belongs to $C_b^{1,0}(S)$.*

Proof. From the definition of u_ε^m , it is evident that u_ε^m is differentiable with continuous derivatives at $x \in \mathbb{R}^d$, and also the mapping $n \mapsto u_\varepsilon^m(x, n)$ is continuous.

It remains to prove that the derivatives are bounded. We have

$$(5.16) \quad \begin{aligned} &|u_\varepsilon^m(x, n) - u_\varepsilon^m(y, n)| \\ &= \left| \int_{B_\varepsilon(0)} \varrho_\varepsilon(z) u^m(x - z, n) dz - \int_{B_\varepsilon(0)} \varrho_\varepsilon(z) u^m(y - z, n) dz \right| \\ &= \left| \int_{B_\varepsilon(0)} \varrho_\varepsilon(z) [u^m(x - z, n) - u^m(y - z, n)] dz \right| \\ &\leq \int_{B_\varepsilon(0)} \varrho_\varepsilon(z) M_u^m |x - y| dz = M_u^m |x - y|. \end{aligned}$$

Hence

$$\frac{|u_\varepsilon^m(x, n) - u_\varepsilon^m(y, n)|}{|x - y|} \leq M_u^m, \quad \text{implying} \quad |\partial_x u_\varepsilon^m(x, n)| \leq M_u^m. \quad \blacksquare$$

LEMMA 5.8. *The value function u^m can be approximated uniformly by the functions u_ε^m :*

$$\|u_\varepsilon^m - u^m\| \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0.$$

Proof. Observe that

$$\begin{aligned} |u_\varepsilon^m(x, n) - u^m(x, n)| &= \left| \int_{B_\varepsilon(0)} \varrho_\varepsilon(y) u^m(x - y, n) dy - u^m(x, n) \right| \\ &\leq \int_{B_\varepsilon(0)} \varrho_\varepsilon(y) |u^m(x - y, n) - u^m(x, n)| dy \\ &\leq \int_{B_\varepsilon(0)} \varrho_\varepsilon(y) M_u^m |y| dy \leq M_u^m \varepsilon. \end{aligned}$$

Thus,

$$\|u_\varepsilon^m - u^m\| = \sup_{(x,n) \in S} |u_\varepsilon^m(x, n) - u^m(x, n)| \leq M_u^m \varepsilon,$$

which gives $\|u_\varepsilon^m - u^m\| \rightarrow 0$ as $\varepsilon \rightarrow 0$. ■

LEMMA 5.9. *The function u_ε^m satisfies*

$$(5.17) \quad \begin{aligned} \gamma_\varepsilon^m &\leq f^m(x, n, v) + \langle g(x, n, v), \partial_x u_\varepsilon^m(x, n) \rangle - \alpha u_\varepsilon^m(x, n), & \forall (x, n, v) \in S \times V, \\ \beta_\varepsilon^m &\leq \ell^m(x, n, k) + u_\varepsilon^m(X(x, n, k), N(x, n, k)) - u_\varepsilon^m(x, n), & \forall (x, n, k) \in D \times K, \end{aligned}$$

where γ_ε and β_ε are constants with $\gamma_\varepsilon \rightarrow 0$ and $\beta_\varepsilon \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Proof. First recall that u^m satisfies (5.14). Applying to these QVI the convolution with the function ϱ_ε , we obtain

$$\begin{aligned} 0 &\leq f^m(x, n, v) * \varrho_\varepsilon + \langle g(x, n, v), \partial_x u^m(x, n) * \varrho_\varepsilon \rangle - \alpha u^m(x, n) * \varrho_\varepsilon \\ &\qquad\qquad\qquad \text{for almost all } (x, n) \in S \text{ and all } v \in V, \\ 0 &\leq \ell^m(x, n, k) * \varrho_\varepsilon + u^m(X(x, n, k), N(x, n, k)) * \varrho_\varepsilon - u^m(x, n) * \varrho_\varepsilon, \\ &\qquad\qquad\qquad \forall (x, n, k) \in D \times K. \end{aligned}$$

Then, the proof reduces to showing that

- (i) $\|f^m * \varrho_\varepsilon - f^m\| \rightarrow 0$, (ii) $\|\langle g, \partial_x u^m \rangle * \varrho_\varepsilon - \langle g, \partial_x u_\varepsilon^m \rangle\| \rightarrow 0$,
- (iii) $\|\ell^m * \varrho_\varepsilon - \ell^m\| \rightarrow 0$ as $\varepsilon \rightarrow 0$.

Let us prove (i):

$$(5.18) \quad \begin{aligned} |f^m * \varrho_\varepsilon(x, n, v) - f^m(x, n, v)| &= \left| \int_{B_\varepsilon(0)} \varrho_\varepsilon(y) f^m(x - y, n, v) dy - f^m(x, n, v) \right| \\ &\leq M_{f,\ell}^m \int_{B_\varepsilon(0)} \varrho_\varepsilon(y) |y| dy \leq M_{f,\ell}^m \varepsilon. \end{aligned}$$

Then the result follows by taking the supremum over all $(x, n, v) \in S \setminus D \times V$. In the same way, we can prove (iii).

It remains to prove (ii): To this end, observe that

$$\begin{aligned} &|\langle g(x, n, v), \partial_x u^m(x, n) \rangle * \varrho_\varepsilon(x) - \langle g(x, n, v), \partial_x u_\varepsilon^m(x, n) \rangle| \\ &= \left| \sum_{j=1}^d \int_{B_\varepsilon(0)} [g_j(x - y, n, v) - g_j(x, n, v)] \partial_{x_j} u^m(x - y, n) \varrho_\varepsilon(y) dy \right| \\ &\leq \int_{B_\varepsilon(0)} \|\partial_x u^m\| M |y| \varrho_\varepsilon(y) dy \leq M \varepsilon \|\partial_x u^m\|, \end{aligned}$$

where M is the constant introduced in (2.5). This proves (5.17). ■

We now establish the relations between the values of the linear program $\inf(\text{P})$ with the value function $u(x, n)$ in (2.12). To begin, we note that our previous results imply that

$$(5.19) \quad u(x, n) \geq \mathbf{u}(x, n) \geq \inf(\text{W}) \geq \inf(\text{P}).$$

The next theorem states that $u(x, n) \leq \inf(\text{P})$, yielding equality in (5.19). This result uses, in some sense, the restrictions of the dual problem (P^*) .

THEOREM 5.10. *Suppose that assumptions (2.4)–(2.7), (2.10), and (5.13) are satisfied. Then, for each $(x, n) \in S$,*

$$(5.20) \quad \inf(\text{P}) = u(x, n).$$

Proof. Let $u_\varepsilon^m(x, n)$ be the approximation of $u^m(x, n)$ established in (5.15). By (5.17), it satisfies

$$\begin{aligned} \gamma_\varepsilon^m &\leq f^m(x, n, v) + \langle g(x, n, v), \partial_x u_\varepsilon^m(x, n) \rangle - \alpha u_\varepsilon^m(x, n), \quad \forall (x, n, v) \in S \times V, \\ \beta_\varepsilon^m &\leq \ell^m(x, n, k) + u_\varepsilon^m(X(x, n, k), N(x, n, k)) - u_\varepsilon^m(x, n), \\ &\quad \forall (x, n, k) \in D \times K. \end{aligned}$$

Recalling the definitions

$$\begin{aligned} A^v \varphi(x, n) &:= -\langle g(x, n, v), \partial_x \varphi(x, n) \rangle + \alpha \varphi(x, n), \\ L^k \varphi(x, n) &:= \varphi(x, n) - \varphi(X(x, n, k), N(x, n, k)), \end{aligned}$$

we can write

$$(5.21) \quad \begin{aligned} \gamma_\varepsilon^m + A^v u_\varepsilon^m(x, n) &\leq f^m(x, n, v), \quad \forall (x, n, v) \in S \setminus D \times V, \\ \beta_\varepsilon^m + L^k u_\varepsilon^m(x, n) &\leq \ell^m(x, n, k), \quad \forall (x, n, k) \in D \times K. \end{aligned}$$

On the other hand, consider a feasible pair (μ_1, μ_2) for (P) . Then

$$\begin{aligned} \langle (\mu_1, \mu_2), (f^m, \ell^m) \rangle &= \langle \mu_1, f^m \rangle + \langle \mu_2, \ell^m \rangle \\ &\geq \langle \mu_1, A^v u_\varepsilon^m \rangle + \langle \mu_1, \gamma_\varepsilon^m \rangle + \langle \mu_2, L^k u_\varepsilon^m \rangle + \langle \mu_2, \beta_\varepsilon^m \rangle \quad (\text{by (5.21)}) \\ &= \langle (\mu_1, \mu_2), (A^v, L^k) u_\varepsilon^m \rangle + \langle (\mu_1, \mu_2), (\gamma_\varepsilon^m, \beta_\varepsilon^m) \rangle \\ &= \langle (A^v, L^k)^*(\mu_1, \mu_2), u_\varepsilon^m \rangle + \langle (\mu_1, \mu_2), (\gamma_\varepsilon^m, \beta_\varepsilon^m) \rangle \quad (\text{by (5.2)}) \\ &= u_\varepsilon^m(x, n) + \langle (\mu_1, \mu_2), (\gamma_\varepsilon^m, \beta_\varepsilon^m) \rangle \quad (\text{by (5.3)}). \end{aligned}$$

Letting $\varepsilon \rightarrow 0$, we deduce

$$(5.22) \quad \langle (\mu_1, \mu_2), (f^m, \ell^m) \rangle \geq u^m(x, n).$$

On the other hand,

$$\begin{aligned}
 |u^m(x, n) - u(x, n)| &\leq \sup_{v(\cdot), \{k_i\}} |J^m(x, n, v(\cdot), \{k_i\}) - J(x, n, v(\cdot), \{k_i\})| \\
 &\leq \sup_{v(\cdot)} \int_0^\infty e^{-\alpha t} |f^m(x(t), n(t), v(t)) - f(x(t), n(t), v(t))| \\
 &\quad + \sup_{\{k_i\}} \sum_{i=0}^\infty e^{-\alpha t_i} |\ell^m(x(t_i-), n(t_i-), \{k_i\}) - \ell(x(t_i-), n(t_i-), \{k_i\})| \rightarrow 0,
 \end{aligned}$$

where the last convergence follows from (5.12). Finally, by letting $m \rightarrow \infty$ in (5.22), we get

$$\langle (\mu_1, \mu_2), (f, \ell) \rangle \geq u(x, n), \quad \forall (x, n) \in S.$$

Since (μ_1, μ_2) was arbitrary, we get $\inf(P) \geq u(x, n)$. This inequality together with (5.19) yields (5.20). ■

As a consequence, we can deduce the following.

REMARK 5.11. The optimal control problem (2.12) is equivalent to its relaxed counterpart defined in (3.4).

Transversality case. We conclude this section by showing that, under the transversality condition (3.5), it is possible to deduce the assumption (5.13).

Firstly, we assert that the transversality condition provides the regularity of the trajectory $t \mapsto (x(t), n(t))$ with respect to the initial data (x, n) . To be more specific, condition (3.5) allows the construction of a function $\psi : S \setminus D \rightarrow \mathbb{R}$ satisfying, in an appropriate sense (viscosity, distribution or semigroup) the inequality

$$\begin{cases} \langle g(x, n, v), \partial_x \psi(x, n) \rangle - \alpha \psi(x, n) \leq -1, & \forall (x, n) \in S \setminus D, \\ \psi(x, n) = 0, & \forall (x, n) \in D. \end{cases}$$

Then, denoting by t_i and t'_i the i th exit times from $S \setminus D$ for the trajectory with initial conditions (x, n) and (x', n) , respectively, we can obtain continuity of these times in the following sense (see Bensoussan and Menaldi [6, p. 415]):

$$(5.23) \quad |e^{-\alpha t_i} - e^{-\alpha t'_i}| \leq \tilde{C}[|x - x'| + |n - n|] \quad \text{for some positive constant } \tilde{C}.$$

By using (5.23), under a suitable induction procedure, similar to that given in the proof of Proposition 3.2 (see Section 6), it is possible to get Lipschitz continuity of the continuous-type variable $x(\cdot)$ of either (6.2) or (6.4):

$$(5.24) \quad |x(t) - x'(t)| \leq \tilde{M}|x - x'| \quad \text{for some positive constant } \tilde{M}, \forall t \geq 0.$$

As a consequence, we have

PROPOSITION 5.12. *For every pair $(\hat{f}, \hat{\ell})$ of functions satisfying the Lipschitz condition (5.9), the corresponding value function $\hat{u}(x, n)$ is Lipschitz continuous in x with constant $M_{\hat{u}} := \tilde{M}M_{\hat{f}, \hat{\ell}}[1/\alpha + 1/(1 - e^{-\alpha h})]$.*

Proof. To prove this result, let us use the notation (x^z, n^p) to emphasize that the trajectory $x(\cdot), n(\cdot)$ has begun at state $(x(0), n(0)) = (z, p)$. With this in mind, we get

$$\begin{aligned} & |\hat{J}(z, p, v(\cdot), \{k_i\}) - \hat{J}(y, p, v(\cdot), \{k_i\})| \\ &= \left| \int_0^\infty e^{-\alpha t} \hat{f}(x^z(t), n^p(t), v(t)) dt + \sum_{i=0}^\infty e^{-\alpha t_i} \hat{\ell}(x^z(t_i-), n^p(t_i-), k_i) \right. \\ &\quad \left. - \int_0^\infty e^{-\alpha t} \hat{f}(x^y(t), n^p(t), v(t)) dt + \sum_{i=0}^\infty e^{-\alpha t_i} \hat{\ell}(x^y(t_i-), n^p(t_i-), k_i) \right| \\ &\leq \int_0^\infty e^{-\alpha t} |\hat{f}(x^z(t), n^p(t), v(t)) - \hat{f}(x^y(t), n^p(t), v(t))| dt \\ &\quad + \sum_{i=0}^\infty e^{-\alpha t_i} |\hat{\ell}(x^z(t_i-), n^p(t_i-), k_i) - \hat{\ell}(x^y(t_i-), n^p(t_i-), k_i)| \\ &\leq \int_0^\infty e^{-\alpha t} M_{\hat{f}, \hat{\ell}} |x^z(t) - x^y(t)| dt + \sum_{i=0}^\infty e^{-\alpha t_i} M_{\hat{f}, \hat{\ell}} |x^z(t_i-) - x^y(t_i-)|. \end{aligned}$$

By using (5.24), we know that

$$|x^z(t) - x^y(t)| \leq \tilde{M}|z - y| \quad \text{for some constant } \tilde{M},$$

so we conclude that

$$|\hat{J}(z, p, v(\cdot), \{k_i\}) - \hat{J}(y, p, v(\cdot), \{k_i\})| \leq \tilde{M}M_{\hat{f}, \hat{\ell}} \left[\frac{1}{\alpha} + \frac{1}{1 - e^{-\alpha h}} \right] |x - y|.$$

Thus,

$$\begin{aligned} |\hat{u}(z, p) - \hat{u}(y, p)| &\leq \sup_{v(\cdot), \{k_i\}} |\hat{J}(z, p, v(\cdot), \{k_i\}) - \hat{J}(y, p, v(\cdot), \{k_i\})| \\ &\leq M_{\hat{u}} |x - y| \end{aligned}$$

with $M_{\hat{u}} := \tilde{M}M_{\hat{f}, \hat{\ell}}[1/\alpha + 1/(1 - e^{-\alpha h})]$. ■

6. Appendix: Proof of Proposition 3.2. Our first step is to rewrite the dynamic (3.2) as follows. For any fixed $(\mathbf{v}, \{k_i\}) \in \mathbf{V} \times \mathcal{K}$, and a fixed initial condition (x_0, n_0) , we consider two cases.

CASE 1: $(x_0, n_0) \in S \setminus D$. Then

$$(6.1) \quad (x_i(t), n_i(t)) = \left(x_i(t_i) + \int_{t_i}^t \int_V g(x_i(s), n_i(s), v) \mathbf{v}_s(dv) ds, n_i(t_i) \right), \quad t \geq 0,$$

$$(x_i(t_i), n_i(t_i)) = (X(x_{i-1}(t_i-), n_{i-1}(t_i-), k_i), N(x_{i-1}(t_i-), n_{i-1}(t_i-), k_i)),$$

$$t_{i+1} := \inf\{t \geq t_i : (x_i(t-), n_i(t_i)) \in D\} \quad \text{when } t_i < \infty, \quad \forall i = 0, 1, \dots,$$

$$(X(x_{-1}(t_0-), n_{-1}(t_0-), k_0), N(x_{-1}(t_0-), n_{-1}(t_0-), k_0)) := (x_0, n_0), t_0 = 0.$$

In this case, (3.2) is equivalent to

$$(6.2) \quad (x(t), n(t)) = \sum_{i=0}^{\infty} (x_i(t), n_i(t)) \mathbf{1}_{[t_i, t_{i+1}[}(t),$$

$$(x_0(t_0), n_0(t_0)) = (x_0, n_0).$$

CASE 2: $(x_0, n_0) \in D$. Then

$$(6.3) \quad (x_i(t), n_i(t)) = \left(x_i(t_i) + \int_{t_i}^t \int_V g(x_i(s), n_i(s), v) \mathbf{v}_s(dv) ds, n_i(t_i) \right), \quad t \geq 0,$$

$$(x_i(t_i), n_i(t_i)) = (X(x_{i-1}(t_i-), n_{i-1}(t_i-), k_i), N(x_{i-1}(t_i-), n_{i-1}(t_i-), k_i)),$$

$$t_{i+1} := \inf\{t \geq t_i : (x_i(t-), n_i(t_i)) \in D\} \quad \text{when } t_i < \infty, \quad \forall i = 1, 2, \dots,$$

$$(X(x_0(t_1-), n_0(t_1-), k_1), N(x_0(t_1-), n_0(t_1-), k_1))$$

$$=: (X(x_0, n_0, k_1), N(x_0, n_0, k_1)), \quad t_1 = 0,$$

for which (3.2) turns into

$$(6.4) \quad (x(t), n(t)) = \sum_{i=1}^{\infty} (x_i(t), n_i(t)) \mathbf{1}_{[t_i, t_{i+1}[}(t),$$

$$(x_1(t_1), n_1(t_1)) = (X(x_0, n_0, k_1), N(x_0, n_0, k_1)).$$

Now take a sequence $\{(\mathbf{v}^m, \{k_i^m\})\}_m$ of controls. Denote by $(x_i^m(\cdot), n_i^m(\cdot))$ the trajectory in either (6.1) or (6.3) when the control pair $(\mathbf{v}^m, \{k_i^m\})$ is applied. Our first step consists in proving that the convergence $(\mathbf{v}^m, \{k_i^m\}) \rightarrow (\mathbf{v}^\infty, \{k_i^\infty\})$ implies the existence of a trajectory $(x_0^\infty(\cdot), n_0^\infty(\cdot))$ satisfying (6.1) or (6.3) respectively and such that $(x_0^m(\cdot), n_0^m(\cdot)) \rightarrow (x_0^\infty(\cdot), n_0^\infty(\cdot))$ locally uniformly as $m \rightarrow \infty$. This last trajectory is governed by the pair $(\mathbf{v}^\infty, \{k_i^\infty\})$.

Let us proceed to handle Case 1. To begin, we define

$$(X(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m), N(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m))$$

$$:= (x_0, n_0), \quad m \geq 1.$$

Then trivially

$$\lim_{m \rightarrow \infty} (X(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m), N(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m)) = (x_0, n_0),$$

so we define

$$\begin{aligned} & (X(x_{-1}^\infty(t_0-), n_{-1}^\infty(t_0-), k_0), N(x_{-1}^\infty(t_0-), n_{-1}^\infty(t_0-), k_0)) \\ & := \lim_{m \rightarrow \infty} (X(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m), N(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m)). \end{aligned}$$

Now use (6.1) for $i = 0$: for every $t \geq 0$ and $t_0 = 0$,

$$(6.5) \quad \begin{aligned} (x_0^m(t), n_0^m(t)) &= \left(X(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m) \right. \\ & \quad \left. + \int_0^t \int_V g(x_0^m(s), n_0^m(s), v) \mathbf{v}_s^m(dv) ds, N(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m) \right). \end{aligned}$$

It is clear that

$$(6.6) \quad n_0^m(t) = N(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m) = n_0, \quad \forall m \geq 1, \forall t \geq 0,$$

which implies that $\lim_m n_0^m(t) = n_0$ for all $t \geq 0$. Thus, we define $n_0^\infty(t) := \lim_m n_0^m(t) = n_0$.

On the other hand, since by (2.5) the vector field $g(x, n, v)$ is bounded uniformly on $S \times V$, we see that for all $m \geq 1$, $\int_V g(x, n, v) \mathbf{v}_t^m(dv)$ is bounded too. Now fix some $T > 0$. According to (6.5), for all $0 \leq s < t \leq T$,

$$(6.7) \quad \begin{aligned} |x_0^m(t) - x_0^m(s)| &= \left| \int_s^t \int_V g(x_0^m(r), n_0^m(r), v) \mathbf{v}_r^m(dv) dr \right| \\ &\leq M(t - s), \quad \forall m \geq 1, \end{aligned}$$

where M is the constant defined in (2.5). Also observe that for all $m \geq 1$ and all $0 \leq t \leq T$,

$$(6.8) \quad \begin{aligned} |x_0^m(t)| &\leq Mt + |x_0^m(0)| \\ &\leq MT + |X(x_{-1}^m(t_0-), n_{-1}^m(t_0-), k_0^m)| \\ &= MT + x_0 < +\infty. \end{aligned}$$

From (6.7) and (6.8), the family $\{x_0^m(\cdot)\}$ is equicontinuous and bounded, so by the Arzelà–Ascoli theorem, it is relatively uniformly compact on $[0, T]$: there exists a (uniformly) convergent subsequence of $\{x_0^m(\cdot)\}$ (not relabeled) such that $x_0^m(\cdot) \rightarrow x_0^\infty(\cdot) \in C_b([0, T])$ uniformly.

Let us now check that this limit term satisfies the first subdynamic in (6.5) for $m = +\infty$. Write

$$\begin{aligned}
 (6.9) \quad x_0^m(t) &= x_0^m(0) + \int_0^t \int_V g(x_0^m(s), n_0^m(s), v) \mathbf{v}_s^m(dv) ds \\
 &= x_0^m(0) + \int_0^t \int_V [g(x_0^m(s), n_0^m(s), v) - g(x_0^\infty(s), n_0^\infty(s), v)] \mathbf{v}_s^m(dv) ds \\
 &\quad + \int_0^t \int_V g(x_0^\infty(s), n_0^\infty(s), v) \mathbf{v}_s^m(dv) ds, \quad \forall 0 \leq t \leq T.
 \end{aligned}$$

The continuity of both $t \mapsto x_0^\infty(t)$ and $(x, n, v) \mapsto g(x, n, v)$ yields the continuity of $(t, v) \mapsto g(x_0^\infty(t), n_0^\infty(t), v)$. Then from Proposition 3.1 we deduce

$$(6.10) \quad \int_0^t \int_V g(x_0^\infty(s), n_0^\infty(s), v) \mathbf{v}_s^m(dv) ds \rightarrow \int_0^t \int_V g(x_0^\infty(s), n_0^\infty(s), v) \mathbf{v}_s^\infty(dv) ds$$

as $m \rightarrow \infty$, for all $0 \leq t \leq T$. Using the Lipschitz property of the vector field g , we can easily verify that the second term on the right-hand side of (6.9) goes to zero as $m \rightarrow \infty$. Together with (6.10) and (6.6), this yields

$$\begin{aligned}
 (6.11) \quad (x_0^m(t), n_0^m(t)) &\rightarrow (x_0^\infty(t), n_0^\infty(t)) \\
 &= \left(x_0 + \int_0^t \int_V g(x_0^\infty(s), n_0^\infty(s), v) \mathbf{v}_s^\infty(dv) ds, n_0 \right).
 \end{aligned}$$

Let us now analyze the first times when the sequence $\{(x_0^m(\cdot), n_0^m(\cdot))\}$ of processes reaches the set-interface D , for each $m \geq 0$ (also known as exit times from $S \setminus D$), and study the convergence of those times. For this purpose and noting that $n_0^m(t) = n_0$ for all $t \geq 0$ and $m \geq 1$, we define

$$\begin{aligned}
 t_1^m &:= \inf\{t \geq 0 : (x_0^m(t-), n_0) \in D\}, \\
 t_1^\infty &:= \inf\{t \geq 0 : (x_0^\infty(t-), n_0) \in D\}.
 \end{aligned}$$

Recalling the definition of D_{n_0} in (2.3), we can write

$$t_1^m = \inf\{t \geq 0 : x_0^m(t-) \in D_{n_0}\} \quad \text{and} \quad t_1^\infty = \inf\{t \geq 0 : x_0^\infty(t-) \in D_{n_0}\}.$$

Since $(x_0, n_0) \in S \setminus D$ (which implies $x_0 \in \mathbb{R}^d \setminus D_{n_0}$), we have $t_1^\infty > 0$. Therefore, there exists a constant \hat{s} with $0 < \hat{s} < t_1^\infty$. By the properties of $x_0^\infty(\cdot)$, the trajectory $t \mapsto x_0^\infty(t)$ lies in $\mathbb{R}^d \setminus D_{n_0}$ on $[0, \hat{s}]$. Observe that $\{x_0^\infty(s) : s \in [0, \hat{s}]\}$ is a compact set contained in the open set $\mathbb{R}^d \setminus D_{n_0}$. By the uniform convergence $x_0^m(\cdot) \rightarrow x_0^\infty(\cdot)$, there exists a natural number M such that for all $m \geq M$, the set $\{x_0^m(t), x_0^m(t) : 0 \leq t \leq \hat{s}\}$ also lies in $\mathbb{R}^d \setminus D_{n_0}$, yielding $t_1^m \geq \hat{s}$ for $m \geq M$. Hence, $\liminf_m t_1^m \geq \hat{s}$. Since \hat{s} was arbitrary, we can take it close enough to t_1^∞ , implying that $\liminf_m t_1^m \geq t_1^\infty$.

Note that if $t_1^\infty = +\infty$, then $\liminf_m t_1^m = +\infty$ and thus the proof would follow by applying the convergence (6.11) on the interval $[0, T]$ for every $T > 0$. Therefore, we now focus on the case $t_1^\infty < +\infty$; namely,

using the assumptions in (3.5), for any $t > t_1^\infty$, there exists $\hat{t} < t$ such that $x_0^\infty(\hat{t}) \in \mathring{D}_{n_0}$. As \mathring{D}_{n_0} is open, the convergence $x_0^m(\cdot) \rightarrow x_0^\infty(\cdot)$ (in this case we must take the constant T in the Arzelà–Ascoli theorem no smaller than \hat{t}) ensures the existence of some constant M such that for all $m \geq M$, we can guarantee $x_0^m(\hat{t}) \in \mathring{D}_{n_0}$. Then $t_1^m \leq \hat{t}$ and so $\limsup_m t_1^m \leq \hat{t}$. Finally, taking \hat{t} close to t_1^∞ , we can deduce that $\limsup_m t_1^m \leq t_1^\infty$. Combining the previous arguments, we obtain

$$\lim_{m \rightarrow \infty} t_1^m = t_1^\infty.$$

This convergence together with the previous uniform convergence of $(x_0(\cdot), n_0(\cdot))$ implies that

$$(6.12) \quad (x_0^\infty(t_1^\infty -), n_0^\infty(t_1^\infty -)) = \lim_{m \rightarrow \infty} (x_0^m(t_1^m -), n_0^m(t_1^m -)).$$

Next, in virtue of the continuity of the mappings X, N and the convergences in (6.12) and $\{k_i^m\} \rightarrow \{k_i^\infty\}$, we deduce

$$(6.13) \quad (X, N)(x_0^m(t_1^m -), n_0^m(t_1^m -), k_1^m) \rightarrow (X, N)(x_0^\infty(t_1^\infty -), n_0^\infty(t_1^\infty -), k_1^\infty)$$

as $m \rightarrow \infty$.

In general, for $i \geq 1$, we first apply similar arguments to those in (6.5) to obtain, for all $t \geq t_i^m$, and $m \geq 1$,

$$(6.14) \quad (x_i^m(t), n_i^m(t)) = \left(X(x_{i-1}^m(t_i^m -), n_{i-1}^m(t_i^m -), k_i^m) + \int_{t_i^m}^t \int_V g(x_i^m(s), n_i^m(s), v) \mathbf{v}_s^m(dv) ds, N(x_{i-1}^m(t_i^m -), n_{i-1}^m(t_i^m -), k_i^m) \right).$$

Furthermore, the process $n(\cdot)$ behaves as

$$(6.15) \quad n_i^m(t) = N(x_{i-1}^m(t_i^m -), n_{i-1}^m(t_i^m -), k_i^m), \quad \forall t \geq t_i^m, \forall m \geq 1.$$

Since the sequence $\{t_i^m\}$ is convergent, we can define $\bar{t}_i := \inf_m t_i^m$. Using this number, we can define a continuous extension of the trajectory $(x_i^m(t), n_i^m(t))$ on $[\bar{t}_i, +\infty[$ by letting

$$(6.16) \quad (x_i^m(t), n_i^m(t)) := \begin{cases} (x_i^m(t_i^m), n_i^m(t_i^m)) & \text{on } [\bar{t}_i, t_i^m[, \\ (x_i^m(t), n_i^m(t_i^m)) & \text{on } [t_i^m, +\infty[. \end{cases}$$

Similar to the convergence (6.13) and the convergence $k_i^m \rightarrow k_i^\infty$, we can deduce

$$\begin{aligned} n_i^m(t) &= N(x_{i-1}^m(t_i^m -), n_{i-1}^m(t_i^m -), k_i^m) \\ &\rightarrow N(x_{i-1}^\infty(t_i^\infty -), n_{i-1}^\infty(t_i^\infty -), k_i^\infty) = n_i^\infty(t), \quad \forall t \geq \bar{t}_i, \end{aligned}$$

and this convergence is uniform.

On the other hand, using again (2.5), we find that the dynamic $x(\cdot)$ has the following properties. For all $T > \bar{t}_i$ and $\bar{t}_i \leq s < t \leq T$, we have

$$|x_i^m(t) - x_i^m(s)| \leq \int_s^t \int_V |g(x_i^m(r), n_i^m(r), v)| \mathbf{v}_r^m(dv) dr \leq M(t - s),$$

and for all $\bar{t}_i \leq t \leq T$,

$$\begin{aligned} |x_i^m(t)| &\leq M(T - \bar{t}_i) + |X(x_{i-1}^m(\bar{t}_i-), n_{i-1}^m(\bar{t}_i-), k_i^m)| \\ &\leq MT + \sup_{m \geq 1} |X(x_{i-1}^m(\bar{t}_i-), n_{i-1}^m(\bar{t}_i-), k_i^m)| < +\infty, \end{aligned}$$

where the last term is bounded due to the convergences

$$(x_{i-1}^\infty(t_i^\infty-), n_{i-1}^\infty(t_i^\infty-)) = \lim_{m \rightarrow \infty} (x_{i-1}^m(t_i^m-), n_{i-1}^m(t_i^m-))$$

and

$$(6.17) \quad \begin{aligned} (X, N)(x_{i-1}^m(t_i^m-), n_{i-1}^m(t_i^m-), k_i^m) \\ \rightarrow (X, N)(x_{i-1}^\infty(t_i^\infty-), n_{i-1}^\infty(t_i^\infty-), k_i^\infty) \end{aligned}$$

as $m \rightarrow \infty$. Hence, by the simple use of the Arzelà–Ascoli theorem, there exists a uniformly convergent subsequence $\{x_i^m(\cdot)\}$ (not relabeled) such that $x_i^m(\cdot) \rightarrow x_i^\infty(\cdot)$ on $[\bar{t}_i, T]$, in particular on $[t_i^\infty, T]$.

To verify that this limit term satisfies the first subdynamic in (6.14) for the case $m = +\infty$, we proceed as follows: From (6.14) and the continuous extension (6.16) we know that

$$x_i^m(t) = X(x_{i-1}^m(t_i^m-), n_{i-1}^m(t_i^m-), k_i^m) + \int_{t_i^m}^t \int_V g(x_i^m(s), n_i^m(s), v) \mathbf{v}_s^m(dv) ds.$$

By following the same steps as in the 0th case, we can prove that

$$(6.18) \quad \begin{aligned} \int_{t_i^m}^t \int_V g(x_i^m(s), n_i^m(s), v) \mathbf{v}_s^m(dv) ds \\ \rightarrow \int_{t_i^\infty}^t \int_V g(x_i^\infty(s), n_i^\infty(s), v) \mathbf{v}_s^\infty(dv) ds \end{aligned}$$

as $m \rightarrow \infty$. Then, by letting $m \rightarrow \infty$, the convergence in (6.18) together with (6.17) yields

$$(6.19) \quad \begin{aligned} (x_i^m(t), n_i^m(t)) &\rightarrow (x_i^\infty(t), n_i^\infty(t)) \\ &= \left(X(x_{i-1}^\infty(t_i^\infty-), n_{i-1}^\infty(t_i^\infty-), k_i^\infty) \right. \\ &\quad \left. + \int_{t_i^\infty}^t \int_V g(x_i^\infty(s), n_i^\infty(s), v) \mathbf{v}_s(dv) ds, N(x_{i-1}^\infty(t_i^\infty-), n_{i-1}^\infty(t_i^\infty-), k_i^\infty) \right) \end{aligned}$$

for all $t_i^\infty \leq t \leq T$.

Let us now analyze the first times (exit times) when the sequence $\{(x_i^m(\cdot), n_i^m(\cdot))\}$ of processes is outside $S \setminus D$ for each $m \geq 0$ and study the convergence of these times. We define these exit times as

$$(6.20) \quad \begin{aligned} t_{i+1}^m &:= \inf\{t \geq t_i^m : (x_i^m(t-), n_i^m(t_i^m)) \in D\}, \\ t_{i+1}^\infty &:= \inf\{t \geq t_i^\infty : (x_i^\infty(t-), n_i^\infty(t_i^\infty)) \in D\}. \end{aligned}$$

Our aim is to prove the convergence (6.21) below. Indeed, it is clear that $(x_i^\infty(t_i^\infty), n_i^\infty(t_i^\infty)) \in S \setminus D$, so $t_{i+1}^\infty > 0$. As a consequence, there exists a constant \hat{s} satisfying $0 < \hat{s} < t_{i+1}^\infty$.

Using the continuous extension (6.16) it is easy to see that $(x_i^\infty(t), n_i^\infty(t)) \in S \setminus D$ for all $t \in [\hat{t}_i, \hat{s}]$. Furthermore, $\{(x_i^\infty(s), n_i^\infty(s)) : s \in [\hat{t}_i, \hat{s}]\} \in S \setminus D$. Since $S \setminus D$ is open, we can use the uniform convergence of $(x_i^m(\cdot), n_i^m(\cdot)) \rightarrow (x_i^\infty(\cdot), n_i^\infty(\cdot))$ on $[\hat{t}_i, \hat{s}]$ to deduce the existence of a natural number M such that for all $m \geq M$, the set $\{(x_i^m(t), n_i^m(t)), (x_i^\infty(t), n_i^\infty(t)) : \hat{t}_i \leq t \leq \hat{s}\}$ is contained in $S \setminus D$. This implies that $\hat{t}_{i+1}^m \geq \hat{s}$ for $m \geq M$, yielding $\liminf_m \hat{t}_{i+1}^m \geq \hat{s}$. Since \hat{s} was arbitrary, we can take this constant close enough to t_{i+1}^∞ , implying that $\liminf_m \hat{t}_{i+1}^m \geq t_{i+1}^\infty$.

The proof for the converse inequality $\limsup_m \hat{t}_{i+1}^m \leq t_{i+1}^\infty$ is identical to the 0th case, so we omit it.

If $t_{i+1}^\infty = +\infty$, then $\liminf_m t_{i+1}^m = +\infty$ and thus the assertion follows by applying the convergence (6.19) on the interval $[t_i^\infty, T]$ for every $T > 0$.

Combining the previous arguments, we can deduce

$$(6.21) \quad \lim_{m \rightarrow \infty} t_{i+1}^m = t_{i+1}^\infty.$$

Again, this convergence together with the uniform convergence of $(x_i^m(\cdot), n_i^m(\cdot)) \rightarrow (x_i^\infty(\cdot), n_i^\infty(\cdot))$ implies that

$$(6.22) \quad (x_i^\infty(t_{i+1}^\infty-), n_i^\infty(t_{i+1}^\infty-)) = \lim_{m \rightarrow \infty} (x_i^m(t_{i+1}^m-), n_i^m(t_{i+1}^m-)),$$

yielding

$$\begin{aligned} (X, N)(x_i^m(t_{i+1}^m-), n_i^m(t_{i+1}^m-), k_{i+1}^m) \\ \rightarrow (X, N)(x_i^\infty(t_{i+1}^\infty-), n_i^\infty(t_{i+1}^\infty-), k_{i+1}^\infty) \end{aligned}$$

as $m \rightarrow \infty$, and so on.

Now take a sequence of processes as in (6.2):

$$\begin{aligned} (x^m(t), n^m(t)) &= \sum_{i=0}^{\infty} (x_i^m(t), n_i^m(t)) \mathbf{1}_{[t_i^m, t_{i+1}^m[}(t), \\ (x_0^m(t_0), n_0^m(t_0)) &= (x_0, n_0) \end{aligned}$$

and define the limit trajectory

$$(x^\infty(t), n^\infty(t)) = \sum_{i=0}^{\infty} (x_i^\infty(t), n_i^\infty(t)) \mathbf{1}_{[t_i^\infty, t_{i+1}^\infty[}(t),$$

$$(x_0^\infty(t_0), n_0^\infty(t_0)) = (x_0, n_0).$$

By construction, $(x^m(t), n^m(t)) \rightarrow (x^\infty(t), n^\infty(t))$ uniformly on each interval $[t_i^\infty, t_{i+1}^\infty[$; in other words, the above convergence is locally uniform at almost every point of $[0, \infty[$.

Finally, to prove Case 2, when $(x_0, n_0) \in D$, we use the previous steps but starting the analysis from (6.13), and then follow the rest of the proof of Case 1. ■

Acknowledgements. This research was funded in part by Wayne State University and by CONACyT grant no. 238045.

References

- [1] C. Aliprantis and K. Border, *Infinite Dimensional Analysis*, 3rd ed., Springer, 2006.
- [2] E. J. Anderson and P. Nash, *Linear Programming in Infinite-Dimensional Spaces*, Wiley, Chichester, 1987.
- [3] V. Azhmyakov, V. G. Boltyanski and A. Poznyak, *Optimal control of impulsive hybrid systems*, *Nonlinear Anal. Hybrid Systems* 2 (2008), 1089–1097.
- [4] G. Barles, S. Dharmatti and M. Ramaswamy, *Unbounded viscosity solutions of hybrid control systems*, *ESAIM: Control Optim. Calc. Var.* 16 (2010), 176–193.
- [5] A. Bensoussan et J.-L. Lions, *Contrôle impulsionnel et inéquations quasi variationnelles*, Gauthier-Villars, Paris, 1982.
- [6] A. Bensoussan and J.-L. Menaldi, *Hybrid control and dynamic programming*, *Dynam. Contin. Discrete Impulsive Systems* 3 (1997), 395–442.
- [7] M. S. Branicky, V. S. Borkar and S. K. Mitter, *A unified framework of hybrid control: model and optimal control theory*, *IEEE Trans. Autom. Control* 43 (1998), 31–45.
- [8] S. Dharmatti and M. Ramaswamy, *Hybrid control systems and viscosity solutions*, *SIAM J. Control Optim.* 44 (2005), 1259–1288.
- [9] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer, New York, 1975.
- [10] R. V. Gamkrelidze, *Principles of Optimal Control Theory*, Plenum Press, New York, 1978.
- [11] J. B. Lasserre, *Moments, Positive Polynomials and Their Applications*, Imperial College Press, London, 2009.
- [12] J. Lygeros, *An overview of hybrid systems control*, in: *Handbook of Networked and Embedded Control Systems*, Birkhäuser, Boston, MA, 2005, 519–537.
- [13] M. Posa, C. Cantu and R. Tedrake, *A direct method for trajectory optimization of rigid bodies through contact*, *Int. J. Robotics Res.* 33 (2014), 69–81.
- [14] P. F. Riedinger, C. Iung and F. Kratz, *An optimal control approach for hybrid systems*, *Eur. J. Control* 9 (2003), 449–458.
- [15] M. S. Shaikh and P. E. Caines, *On the hybrid optimal control problem: theory and algorithms*, *IEEE Trans. Autom. Control* 52 (2007), 1587–1603.

- [16] M. Soler, A. Olivares and E. Staffetti, *Hybrid optimal control to commercial aircraft trajectory planning*, J. Guidance Control Dynam. 33 (2010), 985–991.
- [17] L. Tavernini, *Differential automata and their discrete simulators*, Nonlinear Anal. 11 (1987), 665–683.
- [18] H. Zhang and M. R. James, *Optimal control of hybrid systems and a system of quasi-variational inequalities*, SIAM J. Control Optim. 45 (2006), 722–761.
- [19] P. Zhao, S. Mohan and R. Vasudevan, *Optimal control for nonlinear hybrid systems via convex relaxations*, arXiv:1702.04310v2 (2018).

Héctor Jasso-Fuentes (corresponding author)
Departamento de Matemáticas
CINVESTAV-IPN
A. Postal 14-740
Ciudad de México, 07000, México
E-mail: hjasso@math.cinvestav.mx

Jose-Luis Menaldi
Department of Mathematics
Wayne State University
Detroit, MI 48202, U.S.A.
E-mail: menaldi@wayne.edu

