

# Linear programming approach to optimal impulse control problems with functional constraints

Piunovskiy, Alexey; Zhang, Yi

DOI:

[10.1016/j.jmaa.2020.124817](https://doi.org/10.1016/j.jmaa.2020.124817)

License:

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Piunovskiy, A & Zhang, Y 2021, 'Linear programming approach to optimal impulse control problems with functional constraints', *Journal of Mathematical Analysis and Applications*, vol. 496, no. 2, 124817.  
<https://doi.org/10.1016/j.jmaa.2020.124817>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# Linear programming approach to optimal impulse control problems with functional constraints\*

Alexey Piunovskiy<sup>†</sup> and Yi Zhang<sup>‡</sup>

**Abstract.** This paper considers an optimal impulse control problem of dynamical systems generated by a flow. The performance criteria are total costs over the infinite time horizon. Apart from the main performance to be minimized, there are multiple constraints on performance functionals of a similar type. Under a natural set of compactness-continuity conditions on the system primitives, we establish a linear programming approach, and prove the existence of a stationary optimal control strategy out of a more general class of randomized strategies. This is done by making use of the tools from Markov decision processes.

**Keywords.** Dynamical System, Impulse Control, Constraints, Randomized Strategy, Markov Decision Process, Linear Programming.

**AMS 2000 subject classification:** Primary 49N25; Secondary 90C40.

## 1 Introduction

The literature of impulse control problems is quite vast in terms of both theoretical developments and applications. One popular method of investigations is dynamic programming, based on solving the quasi-variational inequality, see e.g., [2, 3, 12, 13, 25, 26, 31], where the system under control is deterministic in [2, 31] and stochastic in the others. Various meaningful problems arising from Internet congestion control, finance, epidemiology, inventory management, and insurance were solved using this method in [1, 20, 31, 33, 35]. Below we only focus on the literature on deterministic impulse control problems. Another method is the maximum principle. This was established for deterministic impulse control problems in [6, 8, 23, 27, 34], where in [23, 27], the original impulse control problems were first transformed to equivalent problems without impulses in discrete-time and in continuous-time, respectively, and then the versions of maximum principle were formulated for the transformed problems. The control model in [27] was further generalized in [29], see also the monograph [28]. In relation to the maximum principle, certain impulse control problems with a fixed number of impulses can be viewed as mathematical programming problems with the collection of admissible impulse moments and impulses being the control variables. This method was demonstrated in [19, 24], where the gradient of the performance functionals with respect to the control variables was calculated.

A different method from the aforementioned ones is the linear programming approach. This method is rather powerful in handling problems with functional constraints. The formulation of the linear program is often based on the observation that the performance can be expressed as a linear function of suitably defined occupation measures, which satisfy a certain linear equation. Then one may formulate the linear program on the space of all measures satisfying that linear equation. The

---

\*Declarations of interest: none.

<sup>†</sup>Department of Mathematical Sciences, University of Liverpool, L69 7ZL, UK. E-mail: piunov@liv.ac.uk

<sup>‡</sup>Corresponding author. Department of Mathematical Sciences, University of Liverpool, L69 7ZL, UK. E-mail: yi.zhang@liv.ac.uk

linear programming approach is established if one can show that the linear program and the original impulse control problem are equivalent. In general, this equivalence is a delicate matter, and its validity depends on the nature of the problems. For deterministic optimal control problem without impulses, the linear programming method was investigated and established in e.g., [15, 17, 22]. Similar approach to gradually controlled stochastic systems was justified in e.g., [5, 7, 21].

In the literature on deterministic impulse control problems, to the best of our knowledge, only few works were devoted to the linear programming approach: see [9, 10, 16]. In [9] (resp., [10, 16]), for a gradual-impulse control problem over a finite horizon, the authors introduced suitable occupation measures (resp., made use of tools from the theory of partial differential equations) to formulate linear programs on the space of finite measures. However, the equivalence between the linear program and the original problem was left open in [9], and was discussed in [10, 16]. In the aforementioned works, the flow in the Euclidean space came from an ordinary differential equation, whereas in the present paper the flow is arbitrary enough and lives in a Borel space. Moreover, we allow a simultaneous sequence of impulses which was not the case in [9, 10, 16].

The contributions of the present paper are twofold. First, we establish fully the linear programming approach to a deterministic impulse control problem over an infinite time horizon with total cost criteria, where apart from the main performance to be maximized, there are multiple constraints on other performance functionals of a similar type. We achieve this by transforming the impulse control problem to an equivalent Markov decision process (MDP) with total cost, and exploit the relevant results obtained for MDPs in [11]. For this reason, the occupation measures introduced here are different from those introduced in [9, 10, 16]. Also, since we allow accumulation of impulses, the linear program is on the space of infinite-valued measures. In Remark 3.3 we introduce ‘aggregated occupation measures’, which have more in common with the occupation measures in [9, 10, 16]. The corresponding linear programs in terms of such measures can be found in the follow-up paper [32], where we compare our approach with [9, 10, 16].

The second contribution of this paper is that we prove the solvability of the linear program and the original impulse control problem, under a natural set of compactness-continuity conditions on the system primitives. For problems with functional constraints, it is natural to consider randomized control strategies, see [30]. We show that there exists a stationary control strategy, which is optimal in that class. It turns out that the imposed compactness-continuity conditions on the system primitives of the impulse control problem, while being natural, do not imply the equivalent MDP to be a semicontinuous model. To get over this difficulty, we introduce an auxiliary MDP with an extended state space, and, by introducing a suitable metric on it, we show that the auxiliary MDP model is semicontinuous. From this, we eventually retrieve the desired solvability result for the original impulse control problem.

The rest of this paper is organized as follows. In Section 2 we state the concerned deterministic impulse control problem under consideration. In Section 3 we establish the linear programming approach for the impulse control problem by reformulating it as a constrained MDP. In Section 4, we obtain the main solvability result for the impulse control problem. We finish this paper with a conclusion in Section 5. To improve the readability, some technical results are collected in the appendix.

## 2 Problem Statement

In what follows,  $\mathbb{I}\{\cdot\}$  denotes the indicator function.

We will deal with an impulse control model  $\{\mathbf{X}, \mathbf{A}, \phi, l, \{(C_j^g, C_j^I)\}_{j=0}^J\}$  defined through the following system primitives:

- $\mathbf{X}$  is the state space, which is a Borel space, i.e., a topological space that is homeomorphic to a Borel subset of a complete separable metric (i.e., Polish) space. We endow  $\mathbf{X}$  with its Borel

$\sigma$ -algebra  $\mathcal{B}(\mathbf{X})$ .

- $\phi : \mathbf{X} \times [0, \infty) \rightarrow \mathbf{X}$  is the measurable flow possessing the semigroup property  $\phi(x, t + s) = \phi(\phi(x, s), t)$  for all  $x \in \mathbf{X}$  and  $(t, s) \in [0, \infty) \times [0, \infty)$ ;  $\phi(x, 0) = x$  for all  $x \in \mathbf{X}$ . Between the consecutive impulses, the state evolves according to the flow.
- $\mathbf{A}$  is the action space, which is a Borel space and endowed with its Borel  $\sigma$ -algebra  $\mathcal{B}(\mathbf{A})$ .
- $l : \mathbf{X} \times \mathbf{A} \rightarrow \mathbf{X}$  is the mapping describing the new state after the corresponding impulse is applied at the current state.
- For each  $j = 0, 1, \dots, J$ , where  $J$  is a fixed natural number,  $C_j^g : \mathbf{X} \rightarrow [0, \infty)$  is the gradual (or say running) cost rate.
- For each  $j = 0, 1, \dots, J$ ,  $C_j^I : \mathbf{X} \times \mathbf{A} \rightarrow [0, \infty)$  is the cost function associated with the impulses applied at the corresponding states.

All the mappings  $\phi, l, \{C_j^g\}_{j=0}^J$  and  $\{C_j^I\}_{j=0}^J$  are assumed to be (Borel) measurable in their arguments.

Let us describe verbally the system dynamics under control as follows. If the current state is  $x_0 \in \mathbf{X}$ , the decision maker chooses a pair  $(\theta_1, a_1) \in [0, \infty] \times \mathbf{A}$ , where  $\theta_1$  represents the time to wait until the next impulse is applied, and  $a_1$  is the next impulse. Over  $[0, \theta_1)$ , the state process evolves according to  $\phi(x_0, t)$ , and at  $\theta_1$ , the state is  $l(\phi(x_0, \theta_1), a_1)$ . Then the decision maker chooses  $(\theta_2, a_2)$ , and so on. Note that a simultaneous sequence of impulses is not forbidden:  $\theta_i = \theta_{i+1} = \dots \theta_k = 0$ , and at the corresponding time moment the process has  $2 + k - i$  different values.

If  $\theta_1, \theta_2, \dots, \theta_{i-1} < \infty$ , at the current state  $x_{i-1}$ , the total cost accumulated over the next interval of length  $\theta_i$  equals

$$\int_{[0, \theta_i)} C_j^g(\phi(x_{i-1}, u)) du + \mathbb{I}\{\theta_i < \infty\} C_j^I(\phi(x_{i-1}, \theta_i), a_i). \quad (1)$$

In the previous sum, the last summand is absent when  $\theta_i = \infty$  because if  $\theta_i = \infty$  is selected, it means waiting indefinitely until the next impulse moment, i.e., no further impulse will be applied. When  $\theta_i = \infty$ , we will artificially put the next state  $x_i = \Delta$  for some isolated point  $\Delta \notin \mathbf{X}$ . We put  $\mathbf{X}_\Delta := \mathbf{X} \cup \{\Delta\}$ . On the other hand, the selection of  $\theta_i = 0$  means an instantaneous application of the impulse  $a_i$  at the current state  $x_{i-1}$ , leading to the next state  $x_i = l(x_{i-1}, a_i)$ .

Mathematically, the system dynamics (also called trajectories henceforth) can be represented as sequences in one of the following two forms:

$$\begin{aligned} & x_0 \rightarrow (\theta_1, a_1) \rightarrow x_1 \rightarrow (\theta_2, a_2) \rightarrow \dots; \quad \theta_i < \infty \text{ for all } i = 1, 2, \dots, \\ \text{or} & \\ & x_0 \rightarrow (\theta_1, a_1) \rightarrow \dots \rightarrow x_n \rightarrow (\infty, a_{n+1}) \rightarrow \Delta \rightarrow (\theta_{n+2}, a_{n+2}) \rightarrow \Delta \rightarrow \dots, \end{aligned} \quad (2)$$

where  $x_0 \in \mathbf{X}$  is the initial state of the controlled process and  $\theta_i < \infty$  for all  $i = 1, 2, \dots, n$ .

Let the space of all the trajectories (2) be denoted by

$$\begin{aligned} \Omega = & \bigcup_{n=1}^{\infty} [\mathbf{X} \times (([0, \infty) \times \mathbf{A}) \times \mathbf{X})^n \times (\{\infty\} \times \mathbf{A}) \times \{\Delta\} \times (([0, \infty] \times \mathbf{A}) \times \{\Delta\})^\infty] \\ & \cup [\mathbf{X} \times (([0, \infty) \times \mathbf{A}) \times \mathbf{X})^\infty]. \end{aligned}$$

We fix the natural Borel  $\sigma$ -algebra  $\mathcal{F}$  on  $\Omega$ . Finite sequences in the form

$$h_i = (x_0, (\theta_1, a_1), x_1, (\theta_2, a_2), \dots, x_i)$$

will be called  $i$ -histories;  $i = 0, 1, 2, \dots$ . The space of all such  $i$ -histories will be denoted as  $\mathbf{H}_i$ . We endow it with the  $\sigma$ -algebra  $\mathcal{F}_i := \mathcal{B}(\mathbf{H}_i)$ , which is the restriction of  $\mathcal{F}$  to  $\mathbf{H}_i$ .

In line with the previous verbal description, for  $i = 1, 2, \dots$ , at the current state  $x_{i-1} \in \mathbf{X}$ , if the pair  $(\theta_i, a_i) \in [0, \infty] \times \mathbf{A}$  is the selected control, then after  $\theta_i$  time units, the impulse  $a_i$  will be applied, leading instantaneously to the new state

$$x_i = \begin{cases} l(\phi(x_{i-1}, \theta_i), a_i), & \text{if } \theta_i < \infty; \\ \Delta, & \text{if } \theta_i = \infty. \end{cases} \quad (3)$$

Once  $\Delta$  appeared for the first time, it will appear indefinitely in the system dynamics, and one may regard it as an absorbing state.

Next, let us describe the selection of the pairs  $(\theta_i, a_i)$ . If  $\theta_1, \theta_2, \dots, \theta_{i-1} < \infty$ , at the current state  $x_{i-1}$ , the decision maker has in hand the information about the  $i - 1$ -history, that is, the sequence  $h_{i-1} = (x_0, (\theta_1, a_1), x_1, \dots, (\theta_{i-1}, a_{i-1}), x_{i-1})$ . The selection of the next pair  $(\theta_i, a_i) \in [0, \infty] \times \mathbf{A}$  is based on this information, and we further allow the selection of the pair  $(\theta_i, a_i)$  to be randomized. For this reason, even though the original uncontrolled process evolves deterministically according to the flow  $\phi$ , a suitable probability space will be constructed.

Control strategies (or simply say strategies) specify the selection of the pair  $(\theta_i, a_i)$  given the  $i - 1$ -history  $h_{i-1}$ , and are defined as follows.

- Definition 2.1** (a) A (control) strategy  $\pi = \{\pi_i\}_{i=1}^\infty$  is a sequence of stochastic kernels  $\pi_i$  on  $[0, \infty] \times \mathbf{A}$  given  $\mathbf{H}_{i-1}$ .
- (b) A Markov strategy is defined by the sequence of stochastic kernels in the form  $\{\pi_i(d\theta \times da | x_{i-1})\}_{i=1}^\infty$ .
- (c) A strategy is called stationary and denoted as  $\tilde{\pi}$ , if there is a stochastic kernel  $\tilde{\pi}$  on  $[0, \infty] \times \mathbf{A}$  given  $\mathbf{X}_\Delta$  such that  $\pi_i(d\theta \times da | h_{i-1}) = \tilde{\pi}(d\theta \times da | x_{i-1})$  for all  $i = 1, 2, \dots$ .
- (d) A strategy  $\pi = \{\pi_i\}_{i=1}^\infty$  is called deterministic stationary and denoted as  $f$ , if for all  $i = 1, 2, \dots$ ,  $\pi_i(d\theta \times da | h_{i-1}) = \delta_{f(x_{i-1})}(d\theta \times da)$ , where  $f : \mathbf{X}_\Delta \rightarrow [0, \infty] \times \mathbf{A}$  is a measurable mapping.

In terms of interpretation, under a strategy  $\pi = \{\pi_i\}_{i=1}^\infty$ ,  $\pi_i(d\theta \times da | h_{i-1})$  is the (regular) conditional distribution of  $(\Theta_i, A_i)$ <sup>1</sup> given the  $i - 1$ -history  $H_{i-1} = h_{i-1}$ . This is in line with the following construction. Here and below, the capital letters  $X_i, T_i, \Theta_i, A_i$  and  $H_i$  denote the corresponding functions of  $\omega \in \Omega$ , i.e., random elements.

For a given initial distribution  $\nu$  on  $\mathbf{X}$  and a strategy  $\pi$ , by the Ionescu-Tulcea Theorem, see e.g., [4, Prop.7.28], there is a unique probability measure  $P_\nu^\pi$  on  $(\Omega, \mathcal{F})$  satisfying the following conditions:

$$P_\nu^\pi(X_0 \in \Gamma_X) = \nu(\Gamma_X) \quad \forall \Gamma_X \in \mathcal{B}(\mathbf{X}_\Delta);$$

and for all  $i = 1, 2, \dots$ ,  $\Gamma \in \mathcal{B}([0, \infty] \times \mathbf{A})$ ,  $\Gamma_X \in \mathcal{B}(\mathbf{X}_\Delta)$ ,

$$\begin{aligned} P_\nu^\pi((\Theta_i, A_i) \in \Gamma | H_{i-1}) &= \pi_i(\Gamma | H_{i-1}); \\ P_\nu^\pi(X_i \in \Gamma_X | H_{i-1}, (\Theta_i, A_i)) &= \begin{cases} \delta_{l(\phi(X_{i-1}, \Theta_i), A_i)}(\Gamma_X), & \text{if } X_{i-1} \in \mathbf{X}, \Theta_i < \infty; \\ \delta_\Delta(\Gamma_X) & \text{otherwise.} \end{cases} \end{aligned} \quad (4)$$

When the initial distribution  $\nu$  is a Dirac measure concentrated on a singleton, say  $\{x_0\}$ , we write  $P_\nu^\pi$  as  $P_{x_0}^\pi$ . The mathematical expectation with respect to  $P_\nu^\pi$  and  $P_{x_0}^\pi$  is denoted as  $E_\nu^\pi$  and  $E_{x_0}^\pi$ , respectively.

<sup>1</sup>As was mentioned above, the selection of  $(\theta_i, a_i)$  can be randomized, so that any selected  $(\theta_i, a_i)$  can be regarded as the realization of the pair of some random variables  $(\Theta_i, A_i)$ .

Let us introduce the performance measure of a strategy  $\pi$  at an initial distribution  $\nu$ :

$$\mathcal{V}_j(\nu, \pi) := E_\nu^\pi \left[ \sum_{i=1}^{\infty} \mathbb{I}\{X_{i-1} \neq \Delta\} \left\{ \int_{[0, \Theta_i)} C_j^g(\phi(X_{i-1}, u)) du + \mathbb{I}\{\Theta_i < \infty\} C_j^I(\phi(X_{i-1}, \Theta_i), A_i) \right\} \right]$$

for each  $j = 0, 1, \dots, J$ . Again, when  $\nu = \delta_{x_0}$ , we write  $\mathcal{V}_j(x_0, \pi)$  for  $\mathcal{V}_j(\nu, \pi)$ . Note that we do not exclude the possibility of  $\sum_{i=1}^{\infty} \Theta_i < \infty$ , but we will only consider the total cost accumulated over  $[0, \sum_{i=1}^{\infty} \Theta_i)$ . This is consistent with the definition of  $\mathcal{V}_j(\nu, \pi)$ .

The constrained optimal control problem under study is the following one:

$$\begin{aligned} & \text{Minimize with respect to } \pi && \mathcal{V}_0(x_0, \pi) \\ & \text{subject to} && \mathcal{V}_j(x_0, \pi) \leq d_j, \quad j = 1, 2, \dots, J. \end{aligned} \tag{5}$$

Here and below, we take  $x_0 \in \mathbf{X}$  as a fixed initial point, and  $\{d_j\}_{j=1}^J$  as fixed constraint constants.

**Definition 2.2** *A strategy  $\pi$  is called feasible if it satisfies all the constraint inequalities in problem (5). A feasible strategy  $\pi^*$  is called optimal if  $\mathcal{V}_0(x_0, \pi^*) \leq \mathcal{V}_0(x_0, \pi)$  for all feasible strategies  $\pi$ .*

The primary goal of this paper is to show, under a natural set of conditions, the existence of a stationary optimal strategy for the impulse control problem (5), and to establish the linear programming approach that can be used to obtain it. It is thus outside our concern if either problem (5) has no feasible strategy or all the feasible strategies  $\pi$  are with an infinite value, i.e.,  $\mathcal{V}_0(x_0, \pi) = \infty$ . In the former case, problem (5) is not solvable; and in the latter case, any feasible strategy is optimal. Therefore, in the forthcoming discussions, we shall assume that the following condition is in force, regarding the consistency of problem (5).

**Condition 2.1** *There exists some feasible strategy  $\pi$  such that  $\mathcal{V}_0(x_0, \pi) < \infty$ .*

The solvability of problem (5) cannot be guaranteed without imposing further requirements on the system primitives. We shall impose the following set of compactness-continuity conditions, under which, we will actually show the existence of a stationary optimal strategy, and establish the linear programming approach for obtaining it.

**Condition 2.2** (a) *The space  $\mathbf{A}$  is compact, and  $\infty$  is the one-point compactification of  $[0, \infty)$ .*

(b) *The mapping  $(x, a) \in \mathbf{X} \times \mathbf{A} \rightarrow l(x, a)$  is continuous.*

(c) *The mapping  $(x, \theta) \in \mathbf{X} \times [0, \infty) \rightarrow \phi(x, \theta)$  is continuous.*

(d) *For each  $j = 0, 1, \dots, J$ , the function  $(x, a) \in \mathbf{X} \times \mathbf{A} \rightarrow C_j^I(x, a)$  is lower semicontinuous.*

(e) *For each  $j = 0, 1, \dots, J$ , the function  $x \in \mathbf{X} \rightarrow C_j^g(x)$  is lower semicontinuous.*

### 3 Reformulation as an MDP and as a linear program

In this section, we reformulate the impulse control problem (5) as an MDP with total cost criteria. By exploiting some facts about this class of MDPs, we will establish the relation between the impulse control problem (5) and a linear program. While the MDP formulation itself is natural, the resulting MDP is not always convenient to deal with: under Condition 2.2, it is not a semicontinuous model. We shall elaborate on and deal with it in the next section (see the proof of Theorem 4.1 therein), where the main issue of solvability will be addressed.

### 3.1 The MDP formulation

The formulation of the impulse control model  $\{\mathbf{X}, \mathbf{A}, \phi, l, \{(C_j^g, C_j^I)\}_{j=0}^J\}$  as an MDP is done as follows. The state space of the MDP is  $\mathbf{X}_\Delta = \mathbf{X} \cup \{\Delta\}$ , where  $\Delta \notin \mathbf{X}$  is an isolated point. The action space is  $\mathbf{B} := [0, \infty] \times \mathbf{A}$ , which is endowed with the product topology and the corresponding Borel  $\sigma$ -algebra. The transition kernel is defined by

$$Q(dy|x, (\theta, a)) := \begin{cases} \delta_{l(\phi(x, \theta), a)}(dy), & \text{if } x \neq \Delta, \theta \neq \infty; \\ \delta_\Delta(dy) & \text{otherwise,} \end{cases}.$$

The cost functions are given by

$$\bar{C}_j(x, (\theta, a)) = \mathbb{I}\{x \neq \Delta\} \left\{ \int_{[0, \theta]} C_j^g(\phi(x, u)) du + \mathbb{I}\{\theta < \infty\} C_j^I(\phi(x, \theta), a) \right\}, \quad j = 0, 1, \dots, J,$$

and the constraint constants are  $d_j \in [0, \infty)$ ,  $j = 1, 2, \dots, J$ . Here  $J$  is the number of constraints. We may summarize this MDP model as  $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \bar{C}_0, \{\bar{C}_j, d_j\}_{j=1}^J\}$ . We shall often refer to this MDP as the “original” model, since we will introduce an “auxiliary” MDP model in the proof of Theorem 4.1 below.

Now the impulse control problem (5) can be viewed as a (constrained) MDP with total cost criteria.

**Definition 3.1** *For each strategy  $\pi$ , its occupation measure  $\mu^\pi$  in the original MDP model is defined by*

$$\mu^\pi(\Gamma_1 \times \Gamma_2) := E_{x_0}^\pi \left[ \sum_{n=0}^{\infty} \mathbb{I}\{(X_n, B_{n+1}) \in \Gamma_1 \times \Gamma_2\} \right] \quad \forall \Gamma_1 \in \mathcal{B}(\mathbf{X}_\Delta), \Gamma_2 \in \mathcal{B}(\mathbf{B}). \quad (6)$$

Note that the measure  $\mu^\pi$  is  $[0, \infty]$ -valued.

The following fact concerning the occupation measure of a stationary strategy is frequently used below.

**Remark 3.1** *If  $\tilde{\pi}$  is a stationary strategy, then  $\mu^{\tilde{\pi}}(dx \times db) = \tilde{\pi}(db|x) \mu^{\tilde{\pi}}(dx \times \mathbf{B})$ . This follows from  $\mu^\pi(\Gamma_1 \times \Gamma_2) := \sum_{n=0}^{\infty} E_{x_0}^\pi [\mathbb{I}\{(X_n \in \Gamma_1) E_{x_0}^\pi [\mathbb{I}\{B_{n+1} \in \Gamma_2\} | X_n]\}]$  in the notation of Definition 3.1. (For an arbitrary strategy  $\pi$ , the decomposition of its occupation measure into the product of its marginal and a stochastic kernel on  $\mathbf{B}$  given  $\mathbf{X}_\Delta$  is not trivial because the occupation measure in an MDP with total cost is not finite-valued. See Proposition 3.2 below.)*

Let us consider the following linear program for the MDP  $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \bar{C}_0, \{\bar{C}_j, d_j\}_{j=1}^J\}$ :

$$\begin{aligned} \text{Minimize over all measures } \mu \text{ on } \mathbf{X}_\Delta \times \mathbf{B} : & \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) \\ \text{subject to : } & \mu(dx \times \mathbf{B}) = \delta_{x_0}(dx) + \int_{\mathbf{X}_\Delta \times \mathbf{B}} Q(dx|y, b) \mu(dy \times db); \\ & \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) \leq d_j, \quad j = 1, 2, \dots, J. \end{aligned} \quad (7)$$

**Definition 3.2** (a) *A feasible measure  $\mu$  in linear program (7) is called to be with a finite value if it satisfies  $\int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) < \infty$ .*

(b) *In the linear program (7), a measure  $\mu_1$  is said to outperform another measure  $\mu_2$  if*

$$\int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu_1(dx \times db) \leq \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu_2(dx \times db)$$

*for each  $j = 0, 1, \dots, J$ .*

Since we may write  $\mathcal{V}_j(x_0, \pi) = \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu^\pi(dx \times db)$ , the occupation measure of any feasible strategy  $\pi$  for problem (5) satisfies the constraint inequalities in the linear program (7). Moreover, by [18, Lem. 9.4.3], it also satisfies the constraint equality in (7). Thus, for each feasible strategy  $\pi$ , its occupation measure  $\mu^\pi = \mu$  is feasible for the linear program (7). Moreover, under Condition 2.1, there is some feasible measure with a finite value for the linear program (7). We may draw from these observations a first relation between the impulse control problem (5) and the linear program (7):

$$\text{optimal value of problem (5)} \geq \text{optimal value of the linear program (7)}. \quad (8)$$

In the opposite direction, under Conditions 2.1 and 2.2, any feasible measure with a finite value for the linear program (7) is outperformed by the occupation measure of some feasible strategy for problem (5). The exact and more informative formulation of this fact, which is given in Proposition 3.2 below, uses the following notation. Let

$$V := \left\{ x \in \mathbf{X}_\Delta : \inf_{\pi} E_x^\pi \left[ \sum_{n=0}^{\infty} \sum_{j=0}^J \bar{C}_j(X_n, B_{n+1}) \right] > 0 \right\}, \quad (9)$$

where  $\{B_n\}_{n=1}^{\infty}$  is the action process in the MDP  $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \bar{C}_0, \{\bar{C}_j, d_j\}_{j=1}^J\}$ . Note that  $\Delta \in V^c := \mathbf{X}_\Delta \setminus V$  because  $\Delta$  is a costless cemetery in this MDP.

**Proposition 3.1** *Suppose Condition 2.2 is satisfied.*

- (a) *The function  $x \in \mathbf{X}_\Delta \rightarrow \inf_{\pi} E_x^\pi \left[ \sum_{n=0}^{\infty} \sum_{j=0}^J \bar{C}_j(X_n, B_{n+1}) \right]$  is lower semicontinuous, so that in particular, the set  $V^c$  is a closed subset of  $\mathbf{X}_\Delta$ .*
- (b) *There is a deterministic stationary strategy, identified by a measurable mapping  $f^*$  from  $\mathbf{X}_\Delta$  to  $\mathbf{B}$ , such that  $\inf_{\pi} E_x^\pi \left[ \sum_{n=0}^{\infty} \sum_{j=0}^J \bar{C}_j(X_n, B_{n+1}) \right] = E_x^{f^*} \left[ \sum_{n=0}^{\infty} \sum_{j=0}^J \bar{C}_j(X_n, B_{n+1}) \right]$  for each  $x \in \mathbf{X}_\Delta$ , and*

$$Q(V|x, f^*(x)) = 0 \quad \forall x \in V^c. \quad (10)$$

*Proof.* This statement actually concerns an unconstrained version of the impulse control problem, which is the object studied in [31]. Both parts (a) and (b) follow from [31, Thm.1], whereas the last assertion in part (b) is valid further by [11, Prop.3.2] and its proof.  $\square$

By the definition of the set  $V^c$ , we see that the deterministic stationary strategy  $f^*$  coming from Proposition 3.1(b) satisfies

$$\sum_{j=0}^J \bar{C}_j(x, f^*(x)) = 0 \quad \forall x \in V^c. \quad (11)$$

On the set  $V^c$ , it is intuitively clear that one should apply the strategy  $f^*$  from Proposition 3.1 because it leads to null cost thereafter. On the other hand, on the set  $V \times \mathbf{B}$ , any feasible measure with a finite value in the linear program (7) possesses some desired properties as stated in the next proposition.

**Proposition 3.2** *Suppose Conditions 2.1 and 2.2 are satisfied. Then each feasible measure  $\mu$  with a finite value in the linear program (7) is  $\sigma$ -finite on  $V \times \mathbf{B}$ , and there is a stochastic kernel  $\varphi_\mu$  on  $\mathbf{B}$  given  $V$  satisfying*

$$\varphi_\mu(db|x) \mu(dx \times \mathbf{B}) = \mu(dx \times db) \quad \text{on } \mathcal{B}(V \times \mathbf{B}). \quad (12)$$



Moreover, the stationary strategy  $\widetilde{\pi}_\mu$  defined by

$$\widetilde{\pi}_\mu(db|x) := \mathbb{I}\{x \in V\} \varphi_\mu(db|x) + \mathbb{I}\{x \in V^c\} \delta_{f^*(x)}(db) \quad (13)$$

satisfies that

$$\mu^{\widetilde{\pi}_\mu}(\Gamma \times \mathbf{B}) \leq \mu(\Gamma \times \mathbf{B}) \quad \forall \Gamma \in \mathcal{B}(V). \quad (14)$$

*Proof.* The  $\sigma$ -finiteness of  $\mu$  on  $V \times \mathbf{B}$  can be shown as in the proof of [11, Thm.3.2]. Then the existence of the stochastic kernel  $\varphi_\mu$  is a consequence of this and [14, Appendix 4]. The remaining assertions can be shown as in the proof of [11, Thm.3.3, Cor.3.1].  $\square$

**Remark 3.2** *It is a consequence of (11) and (14) that, in the notation of Proposition 3.2, the occupation measure  $\mu^{\widetilde{\pi}_\mu}$  outperforms  $\mu$  in the linear program (7). The verification is as follows:*

$$\begin{aligned} \mathcal{V}_j(x_0, \pi) &= \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\widetilde{\pi}_\mu}(dx \times db) \\ &= \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \widetilde{\pi}_\mu(db|x) \mu^{\widetilde{\pi}_\mu}(dx \times \mathbf{B}) + \int_{V^c \times \mathbf{B}} \bar{C}_j(x, b) \widetilde{\pi}_\mu(db|x) \mu^{\widetilde{\pi}_\mu}(dx \times \mathbf{B}) \\ &= \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \widetilde{\pi}_\mu(db|x) \mu^{\widetilde{\pi}_\mu}(dx \times \mathbf{B}) \leq \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \widetilde{\pi}_\mu(db|x) \mu(dx \times \mathbf{B}) \\ &= \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) \leq \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) \end{aligned}$$

where the second equality holds by Remark 3.1 because  $\widetilde{\pi}_\mu$  is a stationary strategy, the third equality follows from (11) and (13), the first inequality is by (14), the fourth equality is by (12) and (13), and the last inequality holds because  $\bar{C}_j$  is nonnegative. Consequently, under Conditions 2.1 and 2.2,

$$\text{optimal value of problem (5)} \leq \text{optimal value of the linear program (7)}.$$

We draw immediately from Remark 3.2 and (8) the next corollary.

**Corollary 3.1** *Suppose Conditions 2.1 and 2.2 are satisfied. The following assertions hold.*

- (a) *The optimal value of problem (5) coincides with the optimal value of the linear program (7).*
- (b) *The occupation measure of an optimal strategy for problem (5) is an optimal solution to the linear program (7).*

**Remark 3.3** *Every occupation measure as in Definition 3.1 gives rise to the following ‘aggregated occupation measure’  $\eta$  on  $\mathbf{X} \times \mathbf{A}_\square$ , where  $\mathbf{A}_\square := \mathbf{A} \cup \{\square\}$  and  $\square$  is an isolated artificial point:*

$$\begin{aligned} \eta(dy \times \square) &:= \int_{\mathbb{R}_+^0} \int_{\mathbf{X}} \delta_{\phi(x, u)}(dy) \mu(dx \times [u, \infty] \times \mathbf{A}) du, \\ \eta(dy \times da) &:= \int_{\mathbf{X}} \int_{\mathbb{R}_+^0} \delta_{\phi(x, \theta)}(dy) \mu(dx \times d\theta \times da). \end{aligned}$$

One can formulate and investigate the linear program, associated with the optimal control problem (5), also in terms of aggregated occupation measures. Such linear programs have much in common with works [9, 10, 16]. This linear programming approach is presented in the follow-up paper [32].

### 3.2 A simpler linear program and the linear programming approach

In view of Proposition 3.2, Remark 3.2 and Corollary 3.1, the main use of the optimal solution say  $\mu$  to the linear program (7) is to produce via (12) the (stationary) strategy to be applied when the state is in  $V$ . It is natural to “restrict” the linear program (7) to the space of  $\sigma$ -finite measures on  $V \times \mathbf{B}$ . This results in a simpler linear program:

$$\begin{aligned} \text{Minimize} \quad & \int_{V \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) \text{ over } \sigma\text{-finite measures } \mu \text{ on } V \times \mathbf{B} \\ \text{subject to} \quad & \mu(dx \times \mathbf{B}) = \delta_{x_0}(dx) + \int_{V \times \mathbf{B}} Q(dx|y, b) \mu(dy \times db) \text{ on } \mathcal{B}(V); \\ & \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) \leq d_j, \quad j = 1, 2, \dots, J. \end{aligned} \quad (15)$$

(Recall that on  $V^c$ , the strategy  $f^*$  is the best one to be used, and it satisfies for all  $x \in V^c$  that  $\bar{C}_j(x, f^*(x)) = 0$  by the definition of  $V^c$ , and  $Q(V|x, f^*(x)) = 0$  by Proposition 3.1.)

In the next two lemmas, we will elaborate on the relations between the linear programs (7) and (15).

**Lemma 3.1** *Suppose Conditions 2.1 and 2.2 are satisfied. Let  $\mu'$  be a feasible measure with a finite value in the linear program (7). Consider the occupation measure  $\mu^{\widetilde{\pi}_{\mu'}}$  of the stationary strategy  $\widetilde{\pi}_{\mu'}$  defined by (13) with  $\mu$  being replaced by  $\mu'$ . Then the restriction of  $\mu^{\widetilde{\pi}_{\mu'}}$  on  $V \times \mathbf{B}$  defined by*

$$\mu^{\widetilde{\pi}_{\mu'}}|_{V \times \mathbf{B}}(dx \times db) := \mu^{\widetilde{\pi}_{\mu'}}(dx \times db \cap V \times \mathbf{B})$$

*is a feasible measure with a finite value in the linear program (15) satisfying*

$$\int_{\mathbf{X}_{\Delta} \times \mathbf{B}} \bar{C}_j(x, b) \mu'(dx \times db) \geq \int_{\mathbf{X}_{\Delta} \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\widetilde{\pi}_{\mu'}}(dx \times db) = \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\widetilde{\pi}_{\mu'}}|_{V \times \mathbf{B}}(dx \times db).$$

*In particular, the optimal value of the linear program (15) is majorized by the optimal value of the linear program (7).*

*Proof.* For the feasibility of  $\mu^{\widetilde{\pi}_{\mu'}}|_{V \times \mathbf{B}}$  in the linear program (15), one may refer to (10), (11) and the definition of  $\widetilde{\pi}_{\mu'}(db|x) = \delta_{f^*(x)}$  when  $x \in V^c$ . The rest of this lemma follows from applying Remark 3.2 to  $\mu'$  and applying Remark 3.1 to  $\widetilde{\pi}_{\mu'}$ .  $\square$

**Lemma 3.2** *Suppose Conditions 2.1 and 2.2 are satisfied. Let  $\mu$  be a feasible measure with a finite value in the linear program (15), and  $\varphi_{\mu}$  be the stochastic kernel on  $\mathbf{B}$  given  $V$  satisfying (12) and  $\widetilde{\pi}_{\mu}$  be the stationary strategy satisfying (13). Then  $\mu^{\widetilde{\pi}_{\mu}}(dx \times \mathbf{B}) \leq \mu(dx \times \mathbf{B})$  on  $\mathcal{B}(V)$ , and*

$$\int_{\mathbf{X}_{\Delta} \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\widetilde{\pi}_{\mu}}(dx \times db) = \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\widetilde{\pi}_{\mu}}(dx \times db) \leq \int_{V \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db).$$

*In particular, the optimal value of the linear program (15) is greater or equal to the optimal value of the linear program (7)*

*Proof.* The first assertion can be shown as in the proof of [11, Thm.3.3]. The second assertion follows from the first assertion, just as in Remark 3.1. The last assertion immediately follows from the second assertion.  $\square$

We now combine Lemmas 3.1 and 3.2 for the next corollary, which establishes the linear programming approach for solving the impulse control problem (5) and will be referred to in the proof of Theorem 4.1 in the next section.

**Corollary 3.2** *Suppose Conditions 2.1 and 2.2 are satisfied. Then the following assertions hold.*

- (a) *The optimal values of the linear programs (7) and (15) as well as the impulse control problem (5) are all the same.*
- (b) *Suppose the measure  $\mu_*$  is optimal in the linear program (15). Then  $\mu^{\widetilde{\pi}_{\mu_*}}$  defined by (13) with  $\mu$  being replaced by  $\mu_*$  is an optimal solution to the linear program (7), which is feasible with a finite value. In particular, the stationary strategy  $\widetilde{\pi}_{\mu_*}$  is optimal for the impulse control problem (5).*
- (c) *If  $\mu'$  is an optimal solution to the linear program (7), then there is an optimal solution to the linear program (15).*

*Proof.* By the last assertions of Lemma 3.1 and Lemma 3.2, the optimal value of the linear program (15) is equal to the optimal value of the linear program (7). Corollary 3.1(a) asserts that the optimal value of the linear program (7) is the same as the optimal value of the impulse control problem (5). Part (a) is thus verified. Parts (c) and (b) immediately follow from part (a) and Lemmas 3.1 and 3.2, respectively.  $\square$

## 4 Optimality result

In this section, we show the existence of a stationary optimal strategy for the impulsive control problem (5), as well as the solvability of the linear program (15). To this end, we hope to apply the relevant results for MDPs. It will be explained that this application is not immediate, but plausible after additional arguments.

**Theorem 4.1** *Suppose Conditions 2.1 and 2.2 are satisfied. Then the following assertions hold.*

- (a) *There exists an optimal solution say  $\mu^*$  to the linear program (15), and an optimal stationary strategy  $\widetilde{\pi}$  for the impulse control problem (5) such that*

$$\mu^*(dx \times db) = \mu^*(dx \times \mathbf{B})\widetilde{\pi}(db|x) \text{ on } \mathcal{B}(V \times \mathbf{B})$$

*and  $\widetilde{\pi}(db|x) = \delta_{f^*(x)}(db)$  for  $x \in V^c$ , where  $f^*$  is a deterministic stationary strategy coming from Proposition 3.1*

- (b) *The linear program (15) has the same optimal value as the impulse control problem (5). Moreover, if  $\pi^*$  is a strategy such that the restriction of its occupation measure  $\mu^{\pi^*}|_{V \times \mathbf{B}}$  on  $V \times \mathbf{B}$  solves the linear program (15), and for each  $x_i \in V^c$ ,  $\pi_{i+1}^*(db|h_i)$  is concentrated on  $\{b \in \mathbf{B} : \sum_{j=0}^J \bar{C}_j(x_i, b) = 0\}$ , then  $\pi^*$  is an optimal strategy for the impulse control problem (5).*

*Proof.* (a) If we can prove that the linear program (7) has an optimal solution say  $\mu'$ , then we may refer to Corollary 3.2(c) for an optimal solution  $\mu_*$  to the linear program (15), and Corollary 3.2(b) for the statement to be proved with  $\mu^* = \mu^{\widetilde{\pi}_{\mu_*}}$  and  $\widetilde{\pi} = \widetilde{\pi}_{\mu_*}$ . (Recall from Remark 3.1 that  $\mu^{\widetilde{\pi}}(dx \times db) = \mu^{\widetilde{\pi}}(dx \times \mathbf{B})\widetilde{\pi}(db|x)$  on  $\mathcal{B}(\mathbf{X}_\Delta \times \mathbf{B})$  for any stationary strategy  $\widetilde{\pi}$ .)

The rest of the proof of part (a) verifies that the linear program (7) has an optimal solution. To this end, by Corollary 3.1(b), it suffices to show that the impulse control problem (5) has an optimal strategy. Since problem (5) can be regarded as an MDP, we would like to do this by making use of a relevant result of MDPs, though it is not immediately applicable as to be seen below.

Let us quote the relevant fact of MDPs from [11, Thm.4.1]. Namely, for a constrained total cost MDP<sup>2</sup> with Borel state space  $\mathbf{X}_\Delta$ , Borel action space  $\mathbf{B}$ , transition probability  $Q$ , and positive cost

<sup>2</sup>Here the system primitives of the MDP are generic, and are not necessarily defined using the primitives of the impulse control model.

functions  $\{\bar{C}_j\}_{j=0}^J$ , if the model is semicontinuous, then, provided that there exists a feasible strategy with finite value, there is an optimal stationary strategy. Here the model is called semicontinuous if its action space  $\mathbf{B}$  is compact,  $\{\bar{C}_j\}_{j=0}^J$  are all lower semicontinuous, and  $Q$  is continuous, i.e., for each bounded continuous function  $f$  on  $\mathbf{X}_\Delta$ ,  $\int_{\mathbf{X}_\Delta} f(y)Q(dy|x, a)$  is continuous on  $\mathbf{X}_\Delta \times \mathbf{B}$ .

Therefore, under Condition 2.1, if the MDP model  $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \bar{C}_0, \{\bar{C}_j, d_j\}_{j=1}^J\}$  defined in Section 3 using the primitives of the impulse control model were semicontinuous, we would then be able to refer directly to the quoted result to complete the proof of part (a). It is just that under Condition 2.1 and 2.2, this MDP is not necessarily semicontinuous, however.

In greater detail, under Condition 2.2, while the action space  $\mathbf{B}$  is compact, and the cost functions are lower semicontinuous (as verified in the proof of [31, Thm.1]), the transition probability  $Q$  is in general not continuous, because, when  $x \neq \Delta$ ,  $\theta_n \in [0, \infty)$  and  $\theta_n \rightarrow \infty$ , the probabilities  $Q(dy|x, (\theta_n, a))$  do not converge to  $\delta_\Delta(dy)$  in the standard weak topology on the space of Borel probability measures on  $\mathbf{X}_\Delta$  (generated by bounded continuous functions).

We will get over this difficulty by introducing an auxiliary MDP model  $\{\hat{\mathbf{X}}, \mathbf{B}, \hat{Q}, \hat{C}_0, \{\hat{C}_j, d_j\}_{j=1}^J\}$ . Here the state space in the auxiliary model is

$$\hat{\mathbf{X}} := ([0, \infty) \times \mathbf{X}) \cup \{(\infty, \Delta)\}.$$

We endow it with the metric  $\rho_{\hat{\mathbf{X}}}$  introduced in Definition A.1, see also Lemma A.1. By Lemma A.2, this metric space  $\hat{\mathbf{X}}$  is a Borel space. The action space  $\mathbf{B}$  is the same as in the original MDP model.

According to the last assertion of Lemma A.1, with this metric  $\rho_{\hat{\mathbf{X}}}$ ,  $(s_n, x_n) \rightarrow (s, x) \in [0, \infty) \times \mathbf{X}$  if and only if  $s_n \rightarrow s$  and  $x_n \rightarrow x$ , whereas if  $(s, x) = (\infty, \Delta)$ , then the convergence takes place if and only if  $s_n \rightarrow \infty$  in the extended half line  $[0, \infty]$ . Therefore, the relative topology on  $[0, \infty) \times \mathbf{X} \subseteq \hat{\mathbf{X}}$  is the product topology. In that sense we understand the Borel  $\sigma$ -algebra  $\mathcal{B}([0, \infty) \times \mathbf{X})$ . Moreover, by [4, Lem.7.4],  $\mathcal{B}(\hat{\mathbf{X}})$  is generated by  $\mathcal{B}([0, \infty) \times \mathbf{X})$  and  $\{(\infty, \Delta)\}$ . Consequently, we legitimately define the transition kernel  $\hat{Q}$  on  $\hat{\mathbf{X}}$  given  $\hat{\mathbf{X}} \times \mathbf{B}$  by

$$\hat{Q}(\Gamma_1 \times \Gamma_2 | (s, x), (\theta, a)) := \delta_{s+\theta}(\Gamma_1)Q(\Gamma_2 | x, (\theta, a))$$

for each  $\Gamma_1 \in \mathcal{B}([0, \infty))$  and  $\Gamma_2 \in \mathcal{B}(\mathbf{X})$ , and

$$\hat{Q}(\{\infty, \Delta\} | (s, x), (\theta, a)) := \mathbb{I}\{\theta + s = \infty\}.$$

Finally, the cost functions are

$$\hat{C}_j((s, x), (\theta, a)) := \bar{C}_j(x, (\theta, a))$$

for each  $(s, x) \in \hat{\mathbf{X}}$  and  $(\theta, a) \in \mathbf{B}$ , and  $j = 0, 1, \dots, J$ .

Compared to the original MDP model, in this auxiliary MDP model, whose primitives are often signified with a “hat”, the state space has been extended and a state has an additional coordinate. More precisely, the second coordinate of the current state in the auxiliary model records the current state in the original model, and the first coordinate records the time moment when the impulse was applied, leading to that recorded state in the original model. Therefore, we take the initial distribution in the hat MDP model as the product measure  $\delta_0(dt) \times \delta_{x_0}(dx)$ .

Since the consecutive time moments of impulses can be calculated by summing up the time durations between consecutive impulses, i.e., the first coordinate of the action in the original MDP model, any strategy in the auxiliary model admits an equivalent strategy in the original model. Therefore, under Condition 2.1, the auxiliary MDP has a feasible strategy with finite value. Furthermore, if there is an optimal strategy in the auxiliary model, then the original MDP admits an optimal strategy, too.

Thus, to complete this proof, it suffices to show that the auxiliary MDP has an optimal strategy. This is done by verifying that under Condition 2.2 the auxiliary MDP model is semicontinuous, as follows.

The action space  $\mathbf{B}$  is compact trivially.

For the continuity of  $\hat{Q}$ , consider a bounded continuous function  $f$  on  $\hat{\mathbf{X}}$ . Let  $(s_n, x_n) \rightarrow (s, x)$ , and  $(\theta_n, a_n) \rightarrow (\theta, a)$ . Then consider

$$\begin{aligned} \int_{\hat{\mathbf{X}}} f(t, y) \hat{Q}(dt \times dy | (s_n, x_n), (\theta_n, a_n)) &= f(s_n + \theta_n, l(\phi(x_n, \theta_n), a_n)) \mathbb{I}\{s_n + \theta_n < \infty\} \\ &\quad + \mathbb{I}\{s_n + \theta_n = \infty\} f(\infty, \Delta), \end{aligned} \quad (16)$$

If  $s + \theta < \infty$ , then  $s_n, \theta_n < \infty$ , and  $x_n \in \mathbf{X}$  for all large enough  $n \geq 1$ , and the right hand side of the above equality converges to  $f(s + \theta, l(\phi(x, \theta), a)) = \int_{\hat{\mathbf{X}}} f(t, y) \hat{Q}(dt \times dy | (s, x), (\theta, a))$  according to Condition 2.2(b,c). If  $s = \infty$  or  $\theta = \infty$ , then  $s_n + \theta_n \rightarrow \infty$ , and hence  $f(s_n + \theta_n, l(\phi(x_n, \theta_n), a_n)) \rightarrow f(\infty, \Delta)$  according to the definition of the topology on  $\hat{\mathbf{X}}$ . Thus, the right hand side of (16) still converges to  $f(\infty, \Delta) = \int_{\hat{\mathbf{X}}} f(t, y) \hat{Q}(dt \times dy | (s, x), (\theta, a))$ , as required.

For the lower semicontinuity of  $\hat{C}_j$ , where  $j = 0, 1, \dots, J$  is fixed, consider  $(s_n, x_n) \rightarrow (s, x)$ , and  $(\theta_n, a_n) \rightarrow (\theta, a)$ . If  $(s, x) \in [0, \infty) \times \mathbf{X}$ , then  $(s_n, x_n) \in [0, \infty) \times \mathbf{X}$  for all large enough  $n \geq 1$ , and so

$$\begin{aligned} \lim_{n \rightarrow \infty} \hat{C}_j((s_n, x_n), (\theta_n, a_n)) &= \lim_{n \rightarrow \infty} \left\{ \int_{[0, \theta_n]} C_j^g(\phi(x_n, u)) du + \mathbb{I}\{\theta_n < +\infty\} C_j^I(\phi(x_n, \theta_n), a_n) \right\} \\ &\geq \int_{[0, \theta]} C_j^g(\phi(x, u)) du + \mathbb{I}\{\theta < +\infty\} C_j^I(\phi(x, \theta), a) = \hat{C}_j((s, x), (\theta, a)), \end{aligned}$$

where the inequality holds because the function  $(\theta, x) \in [0, \infty] \times \mathbf{X} \rightarrow \int_{[0, \theta]} C_j^g(\phi(x, u)) du + \mathbb{I}\{\theta < \infty\} C_j^I(\phi(x, \theta), a)$  is lower semicontinuous under Condition 2.2(c,d,e), as was proved in the proof of [31, Thm.1], see equation (3.3) therein. If  $(s, x) = (\infty, \Delta)$ , then  $\lim_{n \rightarrow \infty} \hat{C}_j((s_n, x_n), (\theta_n, a_n)) \geq 0 = \hat{C}_j((\infty, \Delta), (\theta, a))$  still holds. The lower semicontinuity of  $\hat{C}_j$  is now seen. This completes the proof of part (a).

(b) By Corollaries 3.1 and 3.2, the linear programs (7) and (15) have the same optimal value as the impulse control problem (5). Now the statement to be proved follows from the observation that  $\mu^{\pi^*}$  is an optimal solution to the linear program (7), which holds because  $\int_{V^c \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\pi^*}(dx \times db) = 0$ .  $\square$

## 5 Conclusion

To sum up, under a natural set of compactness-continuity conditions on the system primitives, we established a linear programming approach to the concerned deterministic optimal impulse control problem, and proved the existence of a stationary optimal control strategy out of a more general class of randomized strategies. This was done by exploiting the relevant facts of MDPs, which were not directly applicable but became plausible after additional arguments.

The linear programming approach proved to be useful for many important control problems, especially in case of constrained optimization. It looks promising to develop such an approach also for non-impulsive deterministic optimal control problems with constraints and to understand its connection with the maximum principle, in case the flow comes from a controlled ordinary differential equation.

## A A metric on the state space of the auxiliary model

In this appendix, we introduce with justifications a metric  $\rho_{\hat{\mathbf{X}}}$  on the state space  $\hat{\mathbf{X}}$  of the auxiliary MDP model in the proof of Theorem 4.1, and verify that it topologizes  $\hat{\mathbf{X}}$  to a Borel space.

Let  $\rho_{\mathbf{X}}$  be a compatible metric on  $\mathbf{X}$ . Since  $\mathbf{X}$  is a Borel space, according to the lemma of Urysohn, see [4, Prop.7.2], it is without loss of generality to assume that  $\rho_{\mathbf{X}}(x, y) \leq 2$  for each  $x, y \in \mathbf{X}$ .

**Definition A.1** Define a function  $\rho_{\hat{\mathbf{X}}} : \hat{\mathbf{X}} \times \hat{\mathbf{X}} \rightarrow [0, \infty)$  as follows. For each  $(x_1, s_1), (x_2, s_2) \in \hat{\mathbf{X}}$ ,

$$\rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_2, x_2)) := \rho_{\mathbf{X}}(x_1, x_2)(1 - g(s_1) \vee g(s_2)) + |g(s_1) - g(s_2)|,$$

where

$$g(s) := \frac{1}{1 + \frac{1}{s}}$$

defines a one-to-one correspondence between  $[0, \infty]$  and  $[0, 1]$ , accepting  $g(0) := 0$  and  $g(\infty) := 1$ . (Note that in the previous definition, if  $x_i = \Delta$  for some  $i = 1, 2$ , then  $\rho_{\mathbf{X}}(x_1, x_2)$  is undefined, but  $1 - g(s_1) \vee g(s_2) = 0$ ; in this case we formally regard  $\rho_{\mathbf{X}}(x_1, x_2)(1 - g(s_1) \vee g(s_2)) := 0$ . This convention should be kept in mind below.)

**Lemma A.1**  $\rho_{\hat{\mathbf{X}}}$  is a metric on  $\hat{\mathbf{X}}$ . With this metric,  $(s_n, x_n) \rightarrow (s, x) \in [0, \infty) \times \mathbf{X}$  if and only if  $s_n \rightarrow s$  and  $x_n \rightarrow x$ , whereas if  $(s, x) = (\infty, \Delta)$ , then the convergence takes place if and only if  $s_n \rightarrow \infty$  in the Euclidean topology. In particular, the relative topology of  $\hat{\mathbf{X}}$  on  $[0, \infty) \times \mathbf{X}$  is the same as the product topology on  $[0, \infty) \times \mathbf{X}$ .

*Proof.* We verify that  $\rho_{\hat{\mathbf{X}}}$  is indeed a metric on  $\hat{\mathbf{X}}$  as follows. Evidently, it is  $[0, \infty)$ -valued and symmetric. Moreover, if  $(s_1, x_1) = (s_2, x_2)$ , then  $\rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_2, x_2)) = 0$ . Let us now show that the opposite direction holds. For notational convenience, below, we write

$$g(s) = \tilde{s} \tag{17}$$

for all  $s \in [0, \infty]$ . Suppose  $\rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_2, x_2)) = 0$ . Then  $\tilde{s}_1 = \tilde{s}_2$  necessarily. We distinguish two possibilities.

- If  $\tilde{s}_1 = \tilde{s}_2 < 1$ , then  $\rho_{\mathbf{X}}(x_1, x_2) = 0$ , i.e.,  $(s_1, x_1) = (s_2, x_2)$ .
- If  $\tilde{s}_1 = \tilde{s}_2 = 1$ , then necessarily  $(s_1, x_1) = (s_2, x_2) = (\infty, \Delta)$ , as required.

It remains to verify that  $\rho_{\hat{\mathbf{X}}}$  satisfies the triangle inequality. Let  $(s_i, x_i) \in \hat{\mathbf{X}}$ ,  $i = 1, 2, 3$ , be arbitrarily fixed. In case  $\tilde{s}_1 \leq \tilde{s}_2 \leq \tilde{s}_3$ ,

$$\begin{aligned} & \rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_3, x_3)) + \rho_{\hat{\mathbf{X}}}((s_3, x_3), (s_2, x_2)) - \rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_2, x_2)) \\ &= \rho_{\mathbf{X}}(x_1, x_3)(1 - \tilde{s}_3) + \tilde{s}_3 - \tilde{s}_1 + \rho_{\mathbf{X}}(x_3, x_2)(1 - \tilde{s}_3) + \tilde{s}_3 - \tilde{s}_2 - \rho_{\mathbf{X}}(x_1, x_2)(1 - \tilde{s}_2) - \tilde{s}_2 + \tilde{s}_1 \\ &= \rho_{\mathbf{X}}(x_1, x_3)(1 - \tilde{s}_3) + \rho_{\mathbf{X}}(x_3, x_2)(1 - \tilde{s}_3) + 2(\tilde{s}_3 - \tilde{s}_2) - \rho_{\mathbf{X}}(x_1, x_2)(1 - \tilde{s}_2) \\ &\geq (1 - \tilde{s}_3)\rho_{\mathbf{X}}(x_1, x_2) - \rho_{\mathbf{X}}(x_1, x_2)(1 - \tilde{s}_2) + 2(\tilde{s}_3 - \tilde{s}_2) \\ &= \rho_{\mathbf{X}}(x_1, x_2)(\tilde{s}_2 - \tilde{s}_3) + 2(\tilde{s}_3 - \tilde{s}_2) \geq 0. \end{aligned}$$

In case  $\tilde{s}_2 \leq \tilde{s}_1 \leq \tilde{s}_3$ ,

$$\begin{aligned} & \rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_3, x_3)) + \rho_{\hat{\mathbf{X}}}((s_3, x_3), (s_2, x_2)) - \rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_2, x_2)) \\ &= \rho_{\mathbf{X}}(x_1, x_3)(1 - \tilde{s}_3) + \tilde{s}_3 - \tilde{s}_1 + \rho_{\mathbf{X}}(x_3, x_2)(1 - \tilde{s}_3) + \tilde{s}_3 - \tilde{s}_2 - \rho_{\mathbf{X}}(x_1, x_2)(1 - \tilde{s}_1) - \tilde{s}_1 + \tilde{s}_2 \\ &\geq \rho_{\mathbf{X}}(x_1, x_2)(\tilde{s}_1 - \tilde{s}_3) - 2(\tilde{s}_1 - \tilde{s}_3) \geq 0. \end{aligned}$$

Finally, in case  $\tilde{s}_2 \leq \tilde{s}_3 \leq \tilde{s}_1$ ,

$$\begin{aligned}
& \rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_3, x_3)) + \rho_{\hat{\mathbf{X}}}((s_3, x_3), (s_2, x_2)) - \rho_{\hat{\mathbf{X}}}((s_1, x_1), (s_2, x_2)) \\
&= \rho_{\mathbf{X}}(x_1, x_3)(1 - \tilde{s}_1) + \tilde{s}_1 - \tilde{s}_3 + \rho_{\mathbf{X}}(x_3, x_2)(1 - \tilde{s}_3) + \tilde{s}_3 - \tilde{s}_2 - \rho_{\mathbf{X}}(x_1, x_2)(1 - \tilde{s}_1) - \tilde{s}_1 + \tilde{s}_2 \\
&= (\rho_{\mathbf{X}}(x_1, x_3) - \rho_{\mathbf{X}}(x_1, x_2))(1 - \tilde{s}_1) + \rho_{\mathbf{X}}(x_3, x_2)(1 - \tilde{s}_3) \\
&\geq (\rho_{\mathbf{X}}(x_1, x_3) - \rho_{\mathbf{X}}(x_1, x_2))(1 - \tilde{s}_1) + \rho_{\mathbf{X}}(x_3, x_2)(1 - \tilde{s}_1) \geq 0,
\end{aligned}$$

as required. The claimed relation holds for all the other possible orders of  $\tilde{s}_1, \tilde{s}_2, \tilde{s}_3$  by symmetry. Thus,  $\rho_{\hat{\mathbf{X}}}$  is a metric on  $\hat{\mathbf{X}}$ .

The last assertion holds automatically from Definition A.1.  $\square$

**Lemma A.2** *Endowed with the metric  $\rho_{\hat{\mathbf{X}}}$ ,  $\hat{\mathbf{X}}$  is a Borel space.*

*Proof.* We show that  $\hat{\mathbf{X}}$  is a Borel subset of some Polish (i.e., complete metric) space, as follows. Let  $\mathbf{X}$  be a Borel subset of a Polish space  $[\mathbf{X}]$ . Consider

$$[\hat{\mathbf{X}}] := ([0, \infty) \times [\mathbf{X}]) \cup \{(\infty, \Delta)\},$$

endowed with  $\rho_{[\hat{\mathbf{X}}]}$  defined as in Definition A.1 with  $\mathbf{X}$  being replaced by  $[\mathbf{X}]$ , see also Lemma A.1 applied to  $[\mathbf{X}]$ .

Let us show that this metric space  $[\hat{\mathbf{X}}]$  is a Polish space. For its separability, the union of  $\{(\infty, \Delta)\}$  and the countable dense subset of  $[0, \infty) \times [\mathbf{X}]$  (endowed with the product topology) provides a required dense subset of  $[\hat{\mathbf{X}}]$ , by the last assertion of Lemma A.1 applied to  $[\hat{\mathbf{X}}]$ . For its completeness, consider a Cauchy sequence  $\{(s_n, x_n)\}_{n=0}^\infty \subseteq [\hat{\mathbf{X}}]$ . We use the notation (17). Then  $\{\tilde{s}_n\}_{n=0}^\infty$  is Cauchy in  $[0, 1]$ , and so it converges to some  $\tilde{s} \in [0, 1]$ . If  $\tilde{s} = 1$ , then  $s_n \rightarrow \infty$  and so  $(s_n, x_n) \rightarrow (\infty, \Delta)$  in the metric space  $[\hat{\mathbf{X}}]$ . If  $\tilde{s} \neq 1$ , then  $s_n \rightarrow s \neq \infty$ , and  $x_n \rightarrow x$  for some  $x \in [\mathbf{X}]$  as  $[\mathbf{X}]$  is Polish. In either case, Cauchy sequences in  $[\hat{\mathbf{X}}]$  converge, as required.

Finally, we observe that  $\hat{\mathbf{X}}$  is a Borel subset of  $[\hat{\mathbf{X}}]$ . In greater detail,  $[0, \infty) \times [\mathbf{X}]$  is an open subset of the Polish space  $[\hat{\mathbf{X}}]$ , because a singleton in any metric space is closed. By [4, Lem.7.4], the trace of  $\mathcal{B}([\hat{\mathbf{X}}])$  on  $[0, \infty) \times [\mathbf{X}]$  is the same as the Borel  $\sigma$ -algebra  $\mathcal{B}([0, \infty) \times [\mathbf{X}])$  when  $[0, \infty) \times [\mathbf{X}]$  is endowed with the relative topology of  $[\hat{\mathbf{X}}]$ , which is the product topology on  $[0, \infty) \times [\mathbf{X}]$  according to the last assertion of Lemma A.1 (applied to  $[\hat{\mathbf{X}}]$ ). Consequently,  $[0, \infty) \times \mathbf{X}$  is a Borel subset of  $[0, \infty) \times [\mathbf{X}]$ , and also of  $[\hat{\mathbf{X}}]$ , and thus  $\hat{\mathbf{X}}$  is a Borel subset of  $[\hat{\mathbf{X}}]$ .  $\square$

## Acknowledgements

This work was partially supported by the Royal Society International Exchanges award IE160503. The authors are grateful to Professor Alexander Plakhov from University of Aveiro (Portugal) and Institute for Information Transmission Problems (Russia) for fruitful discussions.

## References

- [1] Avrachenkov, K., Habachi, O., Piunovskiy, A. and Zhang, Y. (2015). Infinite horizon optimal impulsive control with applications to Internet congestion control. *Intern. J. of Control* **88**, 703–716.
- [2] Barles, G. (1985). Deterministic impulse control problems. *SIAM J. Control Optim.* **23**, 419–432.

- [3] Bensoussan, A. and Lions, J.T. (1975). Optimal impulse and continuous control: method of nonlinear quasi-variational inequalities. *Trudy Mat. Inst. Steklov* **134**, 5–22.
- [4] Bertsekas, D. and Shreve, S. (1978). *Stochastic Optimal Control*. Academic Press, New York.
- [5] Bhatt, A. and Borkar, V. (1996). Occupation measures for controlled Markov processes: characterization and optimality. *Ann. Probab.* **24**, 1531–1562.
- [6] Blaqui re, A. (1985). Impulsive optimal control with finite or infinite time horizon. *J. Optim. Theory Appl.* **46**, 431–439.
- [7] Buckdahn, R., Goreac, D. and Quincampoix, M. (2011). Stochastic optimal control and linear programming approach. *Appl. Math. Optim.* **63**, 257–276.
- [8] Chahim, M., Hartl, R. and Kort, P. (2012). A tutorial on the deterministic impulse control maximum principle: necessary and sufficiency optimality conditions. *Euro. J. Oper. Res.* **219**, 18–26.
- [9] Cl yes, M., Arzelier, D., Henrion, D. and Lasserre, J. (2014). Measures and LMIs for impulsive nonlinear optimal control. *IEEE Trans. on Automatic Control*, **59**, 1374–1379.
- [10] Cl yes, M., Henrion, D. and Kru   k, M. (2017). Semi-definite relaxations for optimal control problems with oscillation and concentration effects. *ESAIM COCV* **23**, 95–117.
- [11] Dufour, F., Horiguchi, M. and Piunovskiy, A. (2012). The expected total cost criterion for Markov decision processes under constraints: a convex analytic approach. *Adv. Appl. Probab.* **44**, 774–793.
- [12] Dufour, F., Horiguchi, M. and Piunovskiy, A. (2016). Optimal impulsive control of piecewise deterministic Markov processes. *Stochastics* **88**, 1073–1098.
- [13] Dufour, F. and Piunovskiy, A. (2015). Impulsive control for continuous-time Markov decision processes. *Adv. Appl. Probab.* **47**, 106–127.
- [14] Dynkin, E. and Yushkevich, A. (1979). *Controlled Markov Processes*. Springer, Berlin.
- [15] Gaitsgory, V. and Quincampoix, M. (2009). Linear programming approach to deterministic infinite horizon optimal control problems with discounting. *SIAM J. Control Optim.* **48**, 2480–2512.
- [16] Henrion, D., Kru   k, M. and Weisser, T. (2019). Optimal control problems with oscillations, concentrations and discontinuities. *Automatica* **103**, 159–165.
- [17] Hern   ndez-Hern   ndez, D., Hern   ndez-Lerma, O. and Taksar, M. (1996). The linear programming approach to deterministic optimal control problems. *Applicaciones Mathematicae* **24**, 17–33.
- [18] Hern   ndez-Lerma, O. and Lasserre, J. (1999). *Further Topics on Discrete-Time Markov Control Processes*. Springer-Verlag, New York.
- [19] Hou, S. and Wong, K. (2011). Optimal impulsive control problem with application to human immunodeficiency virus treatment. *J. Optim. Theory Appl.* **151**, 385–401.
- [20] Korn, R. (1997). Optimal impulse control when control actions have random consequences. *Math. Oper. Res.* **22**, 639–667.
- [21] Kurtz, T. and Stockbridge, R. (1998). Existence of Markov controls and characterization of optimal Markov controls. *SIAM J. Control Optim.* **36**, 609–653.



- [22] Lasserre, J., Henrion, D., Prieur, C. and Trélat, E. (2008). Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM J. Control Optim.* **47**, 1643–1666.
- [23] Leander, R., Lenhart, S. and Protopopescu, V. (2015). Optimal control of continuous systems with impulse controls, *Optim. Control Appl. Meth.* **36**, 535–549.
- [24] Liu, Y., Teo, K., Jennings, L. and Wang, S. (1998). On a class of optimal control problems with state jumps. *J. Optim. Theory Appl.* **98**, 65–82.
- [25] Menaldi, J. L. and Robin, M. (2018). On some ergodic impulse control problems with constraint. *SIAM J. Control Optim.* **56**, 2690—2711.
- [26] Miller, A., Miller B. and Stepanyan, K. (2020). Simultaneous impulse and continuous control of a Markov chain in continuous time. *Automat. Rem. Control* **81**, 469–482.
- [27] Miller, B. and Rubinovich, E. (1992). Optimal impulse control problem with constrained number of impulses. *Math. Comput. Simulat.* **34**, 23–49.
- [28] Miller, B. and Rubinovich, E. (2003). *Impulsive Control in Continuous and Discrete-Continuous Systems*. Springer, New York.
- [29] Miller, B. and Rubinovich, E. (2009). Impulsive control with impulsive actions of two types. *Automat. Rem. Control* **70**, 1795–1813.
- [30] Piunovskiy, A. (1997). *Optimal Control of Random Sequences in Problems with Constraints*, Kluwer, Dordrecht.
- [31] Piunovskiy, A., Plakhov, A., Torres, D. and Zhang, Y. (2019). Optimal impulse control of dynamical systems. *SIAM J. Control Optim.* **57**, 2720–2752.
- [32] Piunovskiy, A. and Zhang, Y. (2020). Aggregated occupation measures and linear programming approach to constrained impulse control problems. In submission. Available at Arxiv...
- [33] Presman, E. and Sethi, S. (2006). Inventory models with continuous and Poisson demands and discounted and average costs. *Prod. Oper. Manag.* **15**, 279–293.
- [34] Rempala, R. and Zabczyk, J. (1988). On the maximum principle for deterministic impulse control problems. *J. Optim. Theory Appl.* **59**, 281–288.
- [35] Wei, J., Yang, H. and Wang, R. (2010). Classical and impulse control for the optimization of dividend and proportional reinsurance policies with regime switching. *J. Optim. Theory Appl.* **147**, 358–377.