

Primal-Dual Programs for the Constrained Optimal Impulse Control: Discounted Model

Alexey Piunovskiy^{1,*}, Yi Zhang²

1. Department of Mathematical Sciences, University of Liverpool, UK

2. School of Mathematics, University of Birmingham, UK

Abstract: This paper studies constrained optimal impulse control problems of a deterministic system described by a (semi)flow, where the performance measures are the discounted total costs including both the costs incurred with applying impulses as well as running costs. We formulate the relaxed problem and the associated primal convex programs in measures together with the dual programs, and establish the relevant duality results. As an application, we formulate and justify a general procedure for obtaining optimal $(J + 1)$ -mixed strategies for the original impulse control problems. This procedure is illustrated with an explicitly solved example.

Keywords: Impulse control, Markov decision process, discounted costs, optimal strategy, convex program, linear program, duality.

2010 Mathematics Subject Classification: 49N25, 90C05, 90C40.

1. Introduction

This paper studies optimal impulse control problems of a deterministic system described by a (semi)flow, where the performance measures are the discounted total costs including both the lump sum cost associated with applying impulses as well as running costs. The decision maker aims at minimizing one performance measure, subject to constraints on the other similar performance measures given by different impulse cost functions and running cost rates.

Optimal impulse control problems have been investigated in various setups, where the process under control can be stochastic or deterministic, and the performance measure can be given by a long-run average or a total accumulated cost. In general, ‘impulse’ means that the decision maker instantly changes, arbitrarily enough, the state of the process. Examples include instant ordering a product in inventory systems, starting vaccination of the population in epidemiology, modifying the portfolio in financial mathematics etc. Below, with few exceptions, we concentrate on the literature on optimal impulse control problem of deterministic systems with total (discounted or undiscounted) cost criteria, in line with the focus of the present paper.

In the absence of functional constraints, the standard method of investigation of optimal impulse control problems is dynamic programming, leading to the so called quasi-variational inequalities, see e.g., [4, 20]. The Pontryagin maximum principle and similar methods were developed in e.g., [7, 8, 18], and the monographs [2, 19].

Another approach that is especially convenient in handling problems with constraints is the convex analytic approach. This method is based on reformulating the original problem as linear or convex programs in the space of suitably defined occupation measures. Such measures and linear programs were introduced and investigated in [5, 9]. In [9], apart from impulse control, the system is also under other types of controls (gradual and singular), the performance is given by the total cost over a finite horizon plus a terminal cost, and a linear program is formulated as a relaxation of the original optimal control problem; see [13] for further extensions. In [5], the author considered an optimal impulse control problem, where a given number (say N) of time instants at which an impulse can be applied are given *a priori*, and there is a constraint requiring that at least one impulse must be applied within a given time duration after some of

*Corresponding author.

E-mail addresses: A. Piunovskiy*: piunov@liv.ac.uk; Y. Zhang: y.zhang.29@bham.ac.uk
<https://doi.org/10.68381/jca34032>

Received date: April 21, 2025; Accepted date: February 24, 2026

© 2027 Published by Banach Press LLC. (Dedicated to Alexander Plakhov on the occasion of his 65th birthday.)

these time instants. For the latter reason, this problem is formulated as a nonlinear integer program, where the control variable is a vector in $\{0, 1\}^N$, indicating whether an impulse is applied at each of the N time instants, which is relaxed to a linear program. Such a control variable can be viewed as a more detailed occupation measure than the one introduced in this paper. In the existing literature on optimal impulse control problems with functional constraints, where the performance criteria are the total undiscounted costs over an infinite horizon, two types of occupation measures were introduced, leading to two different linear programs as relaxations of the original problem, and the relaxation gap was shown to be absent. If the functional constraints are absent, then the aforementioned two linear programs can be viewed as dual programs corresponding to two forms of optimality equations, both of which were established in [20]. Numerical methods for solving optimal impulse control problems based on solving linear programs can be found in e.g., [5, 9, 13]. Other numerical and simulation methods suitable for optimal control problems including impulse control were developed in [1, 3].

Applications of optimal impulse control problems have been found in economics [7], epidemiology [20], human immunodeficiency virus treatment [15], inventory management [8]; for further examples, see Section 3 in [16], Section 3 in [8] and the monograph [24]. In Section 6 of the present paper, an example in queueing system is explicitly solved.

We remark that in [20], the first step of investigation was to reformulate the impulse control problem as a Markov decision process (MDP), where each step corresponds to a time instant when an impulse is applied. This idea will be further explored in the present paper, too. The connection between optimal impulse control problems and discrete-time optimal control problems was also explored in [16], but for problems without functional constraints, while requiring the total number of impulses to be bounded. The latter requirement, which is not assumed in the present paper, leads to a finite-horizon MDP problem.

The aim of the present paper is to extend and complement the known results to problems with constraints and discounted costs, for which we further investigate the convex analytic approach by focusing on the duality issues and formulate procedures for finding optimal strategies in specific forms, the so called $(J + 1)$ -mixed strategies with J being the number of constraints. To the best of our knowledge, such mixed strategies were not considered in the aforementioned literature on optimal impulse control. Note that, without further conditions, the relaxed problem of the original impulse control problem is in space of infinite measures, even in presence of discounting. This is in contrast with situations encountered in convex analytic approach for (classic) optimal gradual control problems. The primal convex program is in the occupation measures. One may formulate its dual program, and the crude but otherwise useful observation is the absence of the duality gap, which follows from the relevant results in [17]. The dual functional is in the space of non-negative Lagrange multipliers. It is desirable (and is required in our general procedure to solve constrained optimal impulse problems) to compute the value of the dual functional for each Lagrange multiplier by solving an unconstrained impulse control problem with a new cost function modified with the Lagrange multipliers. To guarantee that these two values coincide, we impose suitable conditions on the cost functions, which also imply that the relaxed problem is equivalent to a (so called second) convex program in finite measures.

Having said the above, our main contributions lie in the suitable formulation of the dual program of the second convex program, proving the absence of the duality gap, showing that the dual functional of the relaxed problem can be computed by solving an unconstrained problem using the dynamic programming method, and presenting with justifications the general procedure for computing the optimal $(J + 1)$ -mixed optimal strategies for the impulse control problems. We also demonstrate the procedure by solving in closed form an example that can be understood in the context of fluid queueing systems. We note that while the existence of optimal mixed strategies over $J + 1$ deterministic stationary strategies, which was not addressed in the aforementioned works, follows from the recent paper [22], here we showed that the mixture could be done over deterministic stationary strategies which are uniformly optimal for the unconstrained problem modulated by optimal Lagrange multipliers.

The rest of this article is organized as follows. After the preliminary model description in Section 2, we present the rigorous problem statement. We reformulate the impulse control problem as a constrained MDP problem in Section 3. In Section 4, we present relevant results for the induced MDP, formulate the relaxed problem, the primal convex program and its dual program, and observe the absence of relaxation gap and the duality gap. The main results are presented in Section 5, consisting of the formulation of the second convex program and its dual program, the duality results, and the general procedure for finding optimal $(J + 1)$ -mixed strategies for the original impulse control problems. This procedure is illustrated with a solved example in Section 6. Section 7 contains the proofs of the statements. The paper is ended

with a conclusion in Section 8.

The following notations are used throughout this paper. Measures are $[0, \infty]$ -valued. We use the notations $\mathbb{R}_+ := [0, \infty)$ and $\mathbb{R}_+^I := [0, \infty)^I$. When $\mathbf{Y}$ is considered as a subset of $\mathbf{X}$, the complement of the set $\mathbf{Y}$ in $\mathbf{X}$ is denoted by $\mathbf{Y}^c$. Let $\mathbf{V} \subseteq \mathbf{Y}$ be fixed. The restriction on $\mathbf{V}$ of a $[-\infty, \infty]$ -valued function $f(\cdot)$ defined on $\mathbf{Y}$ is denoted by $f|_{\mathbf{V}}$. This convention applies to measures as functions defined on σ -algebras, too. If $(\mathbf{Y}, \mathcal{F})$ and $(\mathbf{B}, \mathcal{G})$ are measurable spaces, and μ is a measure on $\mathbf{Y} \times \mathbf{B}$ endowed with the product σ -algebra, then the marginal of μ on $\mathbf{Y}$ is denoted by $\mu_{\mathbf{Y}}$ defined by $\mu_{\mathbf{Y}}(dy) = \mu(dy \times \mathbf{B})$. We denote by $\mathcal{M}^{\sigma<\infty}(\mathbf{Y})$ (or $\mathcal{M}^{<\infty}(\mathbf{Y})$) the set of σ -finite (resp., finite) measures on $(\mathbf{Y}, \mathcal{F})$ or say for brevity on $\mathbf{Y}$. The space of all bounded measurable functions on $\mathbf{Y}$ is denoted by $\mathbb{B}(\mathbf{Y})$.

2. Preliminary model description

The state space $\mathbf{X}$ of the system under investigation is a non-empty Borel subset of a complete separable metric space (i.e., Polish space) with metric $\rho_{\mathbf{X}}$ and the Borel σ -algebra $\mathcal{B}(\mathbf{X})$. Typical examples are the finite-dimensional Euclidean spaces $\mathbb{R}^n$. In absence of control, the dynamics is described by the (semi-)flow $\phi(\cdot, \cdot) : \mathbf{X} \times [0, \infty) \rightarrow \mathbf{X}$. This flow possesses the following properties:

- (a) $\phi(x, 0) = x$ for all $x \in \mathbf{X}$;
- (b) the mapping $\phi(\cdot, \cdot)$ is measurable, where and below measurability is understood as Borel-measurability;
- (c) $\lim_{t \downarrow 0} \phi(\phi(x, s), t) = \phi(x, s)$ for all $x \in \mathbf{X}$ and $s \in [0, \infty)$;
- (d) $\phi(x, (s + t)) = \phi(\phi(x, s), t)$ for all $x \in \mathbf{X}$ and $(s, t) \in [0, \infty) \times [0, \infty)$.

Suppose at the initial time moment 0 that the state of the system is $x_0 \in \mathbf{X}$. As time goes on, in absence of impulses, the process moves along the flow ϕ leading to the trajectory $\{\phi(x_0, t), t \geq 0\}$. Property (c) of the flow ϕ means that the trajectories of the (uncontrolled) system are continuous from the right, whereas property (d) means that, in absence of impulses, the movement along the flow from $x \in \mathbf{X}$ to $\phi(x, s)$ on the time interval $(0, s]$ and the further movement from $\phi(x, s)$ to $\phi(\phi(x, s), t)$ on the time interval $(s, s + t]$ can be equivalently represented as the combined movement from x to $\phi(x, s + t)$ on the interval $(0, s + t]$.

Roughly speaking, in the system under (purely) impulsive control, the decision maker decides particular time instants $0 \leq t_1 \leq t_2 \leq \dots \leq \infty$, at which he applies impulsive controls (or say impulses) $a_1, a_2, \dots \in \mathbf{A}$ with $\mathbf{A}$ being a non-empty Borel subset of a complete separable metric space with metric $\rho_{\mathbf{A}}$ and the Borel σ -algebra $\mathcal{B}(\mathbf{A})$. Here, if $t_n = \infty$, the impulse applied at t_n is fictitious, and formally speaking, it means no further impulses are applied after t_{n-1} . Each impulse $a \in \mathbf{A}$ applied at the current state $x \in \mathbf{X}$ results in the immediate jump of the process to the new state $l(x, a) \in \mathbf{X}$, where the mapping $l(\cdot, \cdot) : \mathbf{X} \times \mathbf{A} \rightarrow \mathbf{X}$ is assumed to be measurable. The case when $l(x, a) = x$ for some $a \in \mathbf{A}$ is not excluded. After an impulse say a is applied at $t_n < \infty$ when the state was x , and before the next time instant t_{n+1} , the system dynamics evolves according to the flow $\phi(l(x, a), t)$ with t being the time elapsed from the last impulse time instant t_n . See Fig.1. If two finite consecutive impulse time instants coincide, say $t_n = t_{n+1} < \infty$, then the state resulted in by the impulse a' at t_{n+1} is $l(l(x, a), a')$. This happens when multiple impulses are applied at a single time moment (see $t_3 = t_2$ in Fig.1), and we do not exclude such situations from consideration. It implies that at the time moment t_n , multiple values of the state are assigned. While it is possible to describe such multi-valued trajectories of the system under impulsive control, for our purpose it is equivalent and more convenient to consider this system as one in discrete-time, where each step corresponds to a time instant when an impulse is applied. This becomes clearer after we describe the performance measure in the concerned optimal control problem. Further details are given in the next section when this discrete-time system, or more precisely, this Markov decision process (MDP), is formulated rigorously.

Now we describe the optimal impulse control problem. Assume that in absence of impulses, there is a running cost accumulated at the (gradual) cost rate $C^g(x) \in [0, +\infty]$ associated with any state $x \in \mathbf{X}$. The function $C^g(\cdot)$ is again assumed to be measurable. Besides, whenever an impulse a is applied at state x at some finite time instant, the lump sum cost (or say the impulse cost) $C^I(x, a) \in [0, +\infty]$ is incurred, where the function $C^I(\cdot, \cdot)$ is again measurable. On top of that, we assume that the costs (running cost and impulse cost) are discounted with the discount factor $\alpha \in (0, \infty)$. In the current section, to simplify the presentation, we consider only one performance measure coming from the functions $C^g(\cdot)$ and $C^I(\cdot, \cdot)$, see the next paragraph.

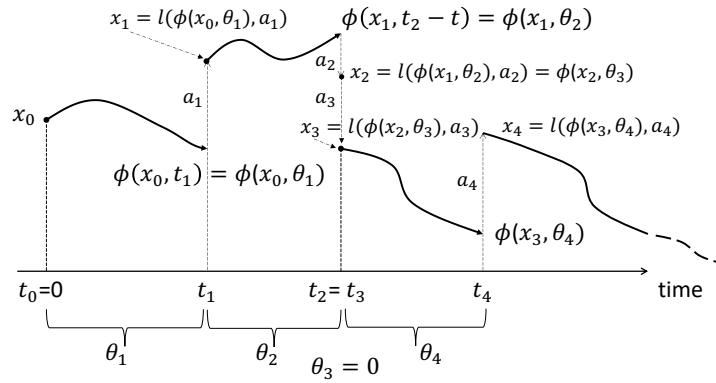


Figure 1. Trajectories of the impulsively controlled dynamical system.

It is convenient to introduce $\theta_i := t_i - t_{i-1}$, $i = 1, 2, \dots$, accepting $\infty - \infty = \infty$, $t_0 := 0$, so that $t_i = \sum_{k=1}^i \theta_k$, $i = 0, 1, 2, \dots$. In terms of interpretation, when it is finite, θ_i is the time duration between two corresponding consecutive time instants when impulses are applied. If, with some $N \in \{1, 2, \dots\}$, there are N impulses $a_1, a_2, \dots, a_N$ being applied at the time moments $t_1 \leq t_2 \leq \dots \leq t_N < \infty$, respectively, and no further impulses beyond t_N , meaning that $\theta_{N+1} = t_{N+1} = t_{N+2} = \dots = \infty$, then the total (discounted) cost equals

$$\sum_{i=1}^N \left\{ \int_0^{\theta_i} e^{-\alpha(t_{i-1}+t)} C^g(\phi(x_{i-1}, t)) dt + e^{-\alpha t_i} C^l(\phi(x_{i-1}, \theta_i), a_i) \right\} + \mathbb{I}\{N < \infty\} \int_0^\infty e^{-\alpha(t_N+t)} C^g(\phi(x_N, t)) dt, \quad (1)$$

where $\alpha > 0$ is the fixed discount factor. In the above formula, $x_i := l(\phi(x_{i-1}, \theta_i), a_i)$, $i = 1, 2, \dots, N$. Expression (1) is the performance measure of a particular (control) strategy applying those specified impulses at the given finitely many time instants. Note that the indicator $\mathbb{I}\{N < \infty\}$ was redundant because $N < \infty$ under the current strategy. Nevertheless, we included it to make the formula valid even when $N = \infty$, provided that we accept that $\mathbb{I}\{N < \infty\} \int_0^\infty e^{-\alpha(t_N+t)} C^g(\phi(x_N, t)) dt = 0$ when $N = \infty$, and $e^{-\alpha t_i} C^l(\phi(x_{i-1}, \theta_i), a_i) = 0$ when $t_i = \infty$.

In general, the number of impulses is not bounded *a priori*, neither are all t_i (or a_i) necessarily deterministic time instants (resp., impulses). They depend on the control strategy. Accordingly, the expression of the concerned performance measures of a general strategy will be given in the next section after the definitions of a strategy and its corresponding so-called strategic measure become clear. For now, one may keep in mind that the optimal control problem investigated in this paper is about minimizing over strategies the expected total (discounted) cost, subject to further constraints to be introduced.

3. Reformulation as an MDP

We now formulate the concerned optimal impulse control problem as an MDP. As mentioned in the last section, this MDP naturally arises from inspecting the system state $\{x_{i-1}\}_{i=1}^\infty$ after each impulse is applied, after the initial state x_0 is observed. Certainly, the sequences $\{\theta_i\}_{i=1}^\infty$ and $\{t_i\}_{i=1}^\infty$ are in one-to-one correspondence with $t_i = \sum_{j=1}^i \theta_j$. The actions in this MDP are the pairs of (θ_i, a_i) with θ_i being the time to wait until the next time instant when the impulse a_i is applied. Then in this MDP, the one-step cost is formed by the (discounted) running cost as well as the impulse cost incurred over two consecutive time instants when impulses are applied in the form of

$$\begin{aligned} & e^{-\alpha t_{i-1}} \int_0^{\theta_i} e^{-\alpha t} C^g(\phi(x_{i-1}, t)) dt + e^{-\alpha t_i} C^l(\phi(x_{i-1}, \theta_i), a_i) \\ &= e^{-\alpha t_{i-1}} \left\{ \int_0^{\theta_i} e^{-\alpha t} C^g(\phi(x_{i-1}, t)) dt + e^{-\alpha \theta_i} C^l(\phi(x_{i-1}, \theta_i), a_i) \right\}. \end{aligned}$$

Here we accept that, if $\theta = \infty$, then $e^{-\alpha\theta} = 0$ and $0 \times \infty = 0$. However, if one pursues this way, then due to the presence of discounting (more precisely, $e^{-\alpha t_{i-1}}$ in the above expression, it would enforce the state of the MDP at step $i - 1$ to be in form of (t_{i-1}, x_{i-1}) . For brevity, we would prefer to avoid the inclusion of the extra coordinate t_{i-1} into the state of the MDP. This is done by observing that $e^{-\alpha t_{i-1}} = \prod_{j=1}^{i-1} e^{-\alpha\theta_j}$ and $e^{-\alpha\theta_j} \in [0, 1]$ can be understood as the conditional probabilities of certain events described as follows. After an impulse a_j is applied at $t_j = t_{j-1} + \theta_j$, the process is artificially terminated (or say “killed”) with the conditional probability $1 - e^{-\alpha\theta_j}$ given that the process has not been killed before. (If $t_{j-1} < \infty$ and $\theta_j = \infty$, it means that no further impulses will be applied beyond t_{j-1} , and in this case, the process is killed with the conditional probability $1 - e^{-\alpha\infty} = 1$.) This amounts to sending the state to Δ , a costless cemetery point, not in $\mathbf{X}$. We join Δ to $\mathbf{X}$, and form $\mathbf{X}_\Delta := \mathbf{X} \cup \{\Delta\}$, where Δ is an isolated point. Δ is absorbing in the sense that once Δ is hit, the system state remains there forever. With this interpretation, one can now take the one-step cost function in the MDP given by

$$\mathbb{I}\{x_{i-1} \in \mathbf{X}\} \left\{ \int_0^{\theta_i} e^{-\alpha t} C^g(\phi(x_{i-1}, t)) dt + e^{-\alpha\theta_i} C^I(\phi(x_{i-1}, \theta_i), a_i) \right\} :$$

the term $e^{-\alpha t_{i-1}}$ has been replaced with $\mathbb{I}\{x_{i-1} \in \mathbf{X}\}$, after we extend the state space to $\mathbf{X}_\Delta$ and modify the transition law in line with what was described. The resulting MDP is what we deal with throughout the rest of this paper.

The rigorous formulation of the MDP model is as follows. It is given by the following elements $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \{\bar{C}_j\}_{j=0}^J, \{d_j\}_{j=1}^J\}$, where

- $\mathbf{X}_\Delta := \mathbf{X} \cup \{\Delta\}$ is the state space;
- $\mathbf{B} := [0, \infty] \times \mathbf{A}$ is the action space equipped with the product σ -algebra, where $[0, \infty]$ is the one-point compactification of $[0, \infty)$;
- $Q(dy|x, b) := \begin{cases} e^{-\alpha\theta} \delta_{I(\phi(x, \theta), a)}(dy) + (1 - e^{-\alpha\theta}) \delta_\Delta(dy) & \text{if } x \neq \Delta, \theta \neq \infty; \\ \delta_\Delta(dy) & \text{otherwise} \end{cases}$ is the transition law, where $b = (\theta, a) \in \mathbf{B}$;
- for each $j \in \{0, \dots, J\}$ with J being a positive integer,

$$\bar{C}_j(x, (\theta, a)) := \begin{cases} \int_0^\theta e^{-\alpha t} C_j^g(\phi(x, t)) dt + e^{-\alpha\theta} C_j^I(\phi(x, \theta), a), & \text{if } x \in \mathbf{X}; \\ 0, & \text{if } x = \Delta, \end{cases}$$

where $C_j^g(\cdot)$ and $C_j^I(\cdot, \cdot)$ are $[0, \infty]$ -valued measurable functions on $\mathbf{X}$ and $\mathbf{X} \times \mathbf{A}$, respectively, and we accept that $e^{-\alpha\theta} C_j^I(\phi(x, \theta), a) = 0$ when $\theta = \infty$;

- for each $1 \leq j \leq J$, d_j is a non-negative real number.

Note that in the above MDP model, instead of a single cost function $\bar{C}(\cdot, \cdot)$, we have $J + 1$ cost functions. This is because we shall consider a constrained optimal impulse control problem in what follows. In this regard, J is the number of constraints, and $\{d_j\}_{j=1}^J$ are the corresponding constraint constants.

Consider the MDP model $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \{\bar{C}_j\}_{j=0}^J, \{d_j\}_{j=1}^J\}$. As usual, we introduce the canonical sample space as the countable product $\Omega := (\mathbf{X}_\Delta \times \mathbf{B})^\infty$ endowed with the product topology, and the corresponding product σ -algebra $\mathcal{F}$. It coincides with the Borel σ -algebra on Ω . An element $\omega = (x_0, b_1 = (\theta_1, a_1), x_1, b_2 = (\theta_2, a_2), x_2, \dots) \in \Omega$ is a trajectory of the MDP. Let, for each $i \in \{0, 1, \dots\}$,

$$X_i(\omega) := x_i, B_{i+1}(\omega) := b_{i+1}, \Theta_{i+1}(\omega) := \theta_{i+1}, A_{i+1}(\omega) := a_{i+1}$$

be the corresponding projection mappings of $\omega \in \Omega$, and let $T_i(\omega) := \sum_{j=1}^i \Theta_j(\omega)$. The process $\{(X_{i-1}, B_i)\}_{i \geq 1}$ is often called an MDP, too, but we use the term for both the bivariate process and the underlying model when there is no danger of confusion. In general and with few exceptional cases, capital letters (not in special fonts) are used to denote random elements defined on Ω , and the argument ω is typically omitted in expressions inside a probability or an expectation.

Finite sequences of the form

$$h_i = (x_0, b_1 = (\theta_1, a_1), x_1, b_2 = (\theta_2, a_2), \dots, x_i)$$

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

A (control) strategy $\pi = \{\pi_i\}_{i=1}^\infty$ is a sequence of stochastic kernels π_i on $\mathbf{B}$ given $\mathbf{H}_{i-1}$. A strategy π is called Markov and denoted by π^M if for each $i \in \{1, 2, \dots\}$, there is some stochastic kernel π_i^M on $\mathbf{B}$ given $\mathbf{X}_\Delta$ such that $\pi_i(db|h_{i-1}) = \pi_i^M(db|x_{i-1})$ for all $h_{i-1} = (x_0, b_1 = (\theta_1, a_1), x_1, b_2 = (\theta_2, a_2), \dots, x_{i-1}) \in \mathbf{H}_{i-1}$. A Markov strategy π^M is called stationary and denoted as π^s , if there is a stochastic kernel π^s on $\mathbf{B}$ given $\mathbf{X}_\Delta$ such that $\pi_i^M(db|x) = \pi^s(db|x)$ for all $x \in \mathbf{X}_\Delta$ and $i \in \{1, 2, \dots\}$. A strategy π (or Markov strategy π^M) is called deterministic (resp., deterministic Markov), if for all $i \in \{1, 2, \dots\}$, there is some measurable mapping $f_i(\cdot) : \mathbf{X}_\Delta \rightarrow \mathbf{B}$ (resp., $f_i^M(\cdot) : \mathbf{X}_\Delta \rightarrow \mathbf{B}$) such that $\pi_i(db|h_{i-1}) = \delta_{f_i(x_{i-1})}(db)$ (resp., $\pi_i^M(db|x_{i-1}) = \delta_{f_i^M(x_{i-1})}(db)$). Finally, a stationary strategy π^s is called deterministic stationary and denoted as f , if $\pi^s(db|x) = \delta_{f(x)}(db)$ for all $x \in \mathbf{X}_\Delta$, where $f(\cdot) : \mathbf{X}_\Delta \rightarrow \mathbf{B}$ is a measurable mapping. Let Π_G be the set of all strategies.

In terms of interpretation, under a strategy π , $\pi_i(db|h_{i-1})$ is the (regular) conditional distribution of $B_i = (\Theta_i, A_i)$ given the $(i-1)$ -history $H_{i-1} = h_{i-1}$. This is in line with the following construction of the underlying probability.

For a given initial state $x_0 \in \mathbf{X}_\Delta$ and a strategy π , by the Ionescu-Tulcea theorem, there is a unique probability measure $\mathbf{P}_{x_0}^\pi$ on $(\Omega, \mathcal{F})$ satisfying the following conditions:

$$\mathbf{P}_{x_0}^\pi(X_0 \in \Gamma_X) = \delta_{x_0}(\Gamma_X) \quad \forall \Gamma_X \in \mathcal{B}(\mathbf{X}_\Delta),$$

and for all $i = 1, 2, \dots$, $\Gamma_B \in \mathcal{B}(\mathbf{B})$, $\Gamma_X \in \mathcal{B}(\mathbf{X}_\Delta)$, $\mathbf{P}_{x_0}^\pi$ -almost surely

$$\begin{aligned} \mathbf{P}_{x_0}^\pi(B_i \in \Gamma_B | H_{i-1}) &= \pi_i(\Gamma_B | H_{i-1}); \\ \mathbf{P}_{x_0}^\pi(X_i \in \Gamma_X | H_{i-1}, B_i = (\Theta_i, A_i)) &= Q(\Gamma_X | X_{i-1}, B_i) \\ &= \begin{cases} e^{-\alpha\Theta_i} \delta_{l(\phi(X_{i-1}, \Theta_i), A_i)}(\Gamma_X) + (1 - e^{-\alpha\Theta_i}) \delta_\Delta(\Gamma_X), & \text{if } X_{i-1} \neq \Delta, \Theta_i \neq \infty; \\ \delta_\Delta(\Gamma_X) & \text{otherwise.} \end{cases} \end{aligned} \quad (2)$$

The probability $\mathbf{P}_{x_0}^\pi$ is called the strategic measure of π (given the initial state x_0). The mathematical expectation with respect to $\mathbf{P}_{x_0}^\pi$ is denoted as $\mathbf{E}_{x_0}^\pi$. Note that $\mathbf{P}_{x_0}^\pi(X_i \in \mathbf{X}) = \mathbf{E}_{x_0}^\pi[e^{-\alpha T_i}]$, $i = 0, 1, \dots$

Now we define the performance measures of each strategy. It will be convenient to introduce the following notation of the performance measure of a generic cost function $\bar{C}(\cdot, \cdot)$ on $\mathbf{X}_\Delta \times \mathbf{B}$ for the MDP in the form of

$$\bar{C}(x, (\theta, a)) := \begin{cases} \int_0^\theta e^{-\alpha t} C^g(\phi(x, t)) dt + e^{-\alpha\theta} C^l(\phi(x, \theta), a), & \text{if } x \in \mathbf{X}; \\ 0, & \text{if } x = \Delta, \end{cases}$$

with $C^l(\cdot, \cdot)$ and $C^g(\cdot)$ being $[0, \infty]$ -valued measurable functions on $\mathbf{X} \times \mathbf{A}$ and on $\mathbf{X}$, respectively. That $\bar{C}(\cdot, \cdot)$ is called generic in the sense that if $C^l(\cdot, \cdot)$ and $C^g(\cdot)$ in its definition are replaced by $C_j^l(\cdot, \cdot)$ and $C_j^g(\cdot)$, then we obtain $\bar{C}_j(\cdot, \cdot)$. Now we define for each $x \in \mathbf{X}_\Delta$ and strategy π the performance measure

$$\mathcal{V}(x, \pi, \bar{C}) := \mathbf{E}_x^\pi \left[\sum_{i=1}^\infty \bar{C}(X_{i-1}, (\Theta_i, A_i)) \right].$$

Below, we take $x_0 \in \mathbf{X}$ as a fixed initial state, since the case of $x_0 = \Delta$ is trivial. Then the constrained optimal control problem under study is the following one:

$$\begin{aligned} &\text{Minimize with respect to } \pi \in \Pi_G && \mathcal{V}(x_0, \pi, \bar{C}_0) \\ &\text{subject to} && \mathcal{V}(x_0, \pi, \bar{C}_j) \leq d_j, \quad j = 1, 2, \dots, J. \end{aligned} \quad (3)$$

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 $\mathcal{V}(x_0, \pi, \bar{C}_j)$.

Definition 3.1. A strategy π is called feasible in problem (3) if it satisfies all the constraint inequalities therein. A strategy π^* is called optimal in problem (3) if it is feasible and $\mathcal{V}(x_0, \pi^*, \bar{C}_0) \leq \mathcal{V}(x_0, \pi, \bar{C}_0)$ for all feasible strategies π in problem (3). A feasible strategy π such that $\mathcal{V}(x_0, \pi, \bar{C}_0) < \infty$ is called feasible with a finite value.

The infimum of $\mathcal{V}(x_0, \pi, \bar{C}_0)$ over all feasible strategies is called the optimal value of problem (3), and is denoted by $val(\text{problem (3)})$. The notation $val(\cdot)$ will also be used for other concerned optimization problems; its meaning lies unambiguously in its argument. Note that, if there are no feasible strategies with a finite value, then either problem (3) is not consistent or any feasible strategy is optimal.

4. Facts and preliminaries

4.1. Semicontinuous induced MDP

The MDP (model) $\{\mathbf{X}_\Delta, \mathbf{B}, Q, \{\bar{C}_j\}_{j=0}^J, \{d_j\}_{j=1}^J\}$ introduced in the previous section is induced by the impulse control model, and thus will be referred to as the induced MDP. If the state space $\mathbf{X}_\Delta$ is replaced with an arbitrary nonempty Borel space say $\mathbf{Y}$, and $\mathbf{B}$, Q and $\{\bar{C}_j\}_{j=0}^J$ are arbitrary nonempty Borel action space, stochastic kernel and $[0, \infty]$ -valued cost functions (not necessarily induced in the above specific form by $C_j^g(\cdot)$ and $C_j^l(\cdot, \cdot)$), then we call the MDP $\{\mathbf{Y}, \mathbf{B}, Q, \{\bar{C}_j\}_{j=0}^J, \{d_j\}_{j=1}^J\}$ generic. The optimal control problem (3), Definition 3.1 and Definition 4.6 below can be understood for a generic MDP, too, with obvious adjustments; in particular, x_0 will be a fixed state in $\mathbf{Y}$.

Definition 4.1. A generic MDP $\{\mathbf{Y}, \mathbf{B}, Q, \{\bar{C}_j\}_{j=0}^J, \{d_j\}_{j=1}^J\}$ is called semicontinuous if the following hold:

- (a) the action space $\mathbf{B}$ is compact;
- (b) for each bounded continuous function $u(\cdot)$ on $\mathbf{Y}$, $\int_{\mathbf{Y}} u(y)Q(dy|x, b)$ is continuous in $(x, b) \in \mathbf{Y} \times \mathbf{B}$;
- (c) for each $j \in \{0, \dots, J\}$, $\bar{C}_j(\cdot, \cdot)$ is lower semicontinuous.

The following condition implies that the induced MDP is semicontinuous.

Assumption 4.2. (a) $\mathbf{A}$ is compact.

- (b) $l(\cdot, \cdot)$ is continuous on $\mathbf{X} \times \mathbf{A}$.
- (c) $\phi(\cdot, \cdot)$ is continuous on $\mathbf{X} \times [0, \infty)$.
- (d) For each $j \in \{0, \dots, J\}$, $C_j^l(\cdot, \cdot)$ is lower semicontinuous on $\mathbf{X} \times \mathbf{A}$, and $C_j^g(\cdot)$ is lower semicontinuous on $\mathbf{X}$.

Proposition 4.3. Under Condition 4.2, the induced MDP is semicontinuous.

Proof. See Section 7. □

Lemma 4.4. If the induced MDP is semicontinuous and admits a feasible strategy with a finite value, then there exists a stationary strategy, which is with a finite value and optimal in problem (3).

Proof. The assertion follows from the fact established in Theorem 4.1 of [10] for a generic MDP. □

Let

$$\mathbf{V} := \left\{ x \in \mathbf{X}_\Delta : \inf_{\pi \in \Pi_G} \mathbb{E}_x^\pi \left[\sum_{i=1}^{\infty} \sum_{j=0}^J \bar{C}_j(X_{i-1}, B_i) \right] = \inf_{\pi \in \Pi_G} \mathcal{V} \left(x, \pi, \sum_{j=0}^J \bar{C}_j \right) > 0 \right\}, \quad (4)$$

Note that $\Delta \in \mathbf{V}^c := \mathbf{X}_\Delta \setminus \mathbf{V}$ because Δ is a costless cemetery in the induced MDP.

Suppose the induced MDP is semicontinuous. Applying Corollary 9.17.2 of [6] to the induced MDP under study with the cost function $\sum_{j=0}^J \bar{C}_j(\cdot, \cdot)$, we obtain that the Bellman function

$$\mathcal{V}^*(x, \sum_{j=0}^J \bar{C}_j) := \inf_{\pi \in \Pi_G} \mathbb{E}_x^\pi \left[\sum_{i=1}^{\infty} \sum_{j=0}^J \bar{C}_j(X_{i-1}, B_i) \right], \quad x \in \mathbf{X}_\Delta \quad (5)$$

is a non-negative lower semicontinuous function, so that, in particular, the set $\mathbf{V}^c$ is a closed subset of $\mathbf{X}_\Delta$ and $\mathbf{V}$ is an open subset of $\mathbf{X}$. Moreover, there is a deterministic stationary strategy f^* providing the infimum in (5) and in the Bellman equation satisfied by $\mathcal{V}^*(\cdot, \sum_{j=0}^J \bar{C}_j)$:

$$\mathcal{V}^*(x, \sum_{j=0}^J \bar{C}_j) = \inf_{b \in \mathbf{B}} \left\{ \sum_{j=0}^J \bar{C}_j(x, b) + \int_{\mathbf{X}_\Delta} \mathcal{V}^*(y, \sum_{j=0}^J \bar{C}_j) Q(dy|x, b) \right\}, \quad x \in \mathbf{X}_\Delta,$$

and

$$\bar{C}_j(x, f^*(x)) = Q(\mathbf{V}|x, f^*(x)) = 0 \quad \forall x \in \mathbf{V}^c, j \in \{0, \dots, J\}. \quad (6)$$

For the last assertion, it is sufficient to notice that for all $x \in \mathbf{V}^c$

$$0 = \mathcal{V}^*(x, \sum_{j=0}^J \bar{C}_j) = \sum_{j=0}^J \bar{C}_j(x, f^*(x)) + \int_{\mathbf{V}} \mathcal{V}^*(y, \sum_{j=0}^J \bar{C}_j) Q(dy|x, f^*(x))$$

and $\mathcal{V}^*(y, \sum_{j=0}^J \bar{C}_j) > 0$ for $y \in \mathbf{V}$. We fix this deterministic stationary strategy f^* .

Now it is clear from (6) that, if $x_0 \in \mathbf{V}^c$, then f^* is an optimal strategy in problem (3) because

$$\mathcal{V}(x, f^*, \bar{C}_j) = 0 \quad \forall x \in \mathbf{V}^c, j = 0, 1, \dots, J.$$

Consequently, in what follows, we suppose that $x_0 \in \mathbf{V}$, and thus $\mathbf{V} \neq \emptyset$. It is also clear that, for each $i \in \{1, 2, \dots\}$, it is optimal to apply the action $f^*(X_{i-1})$ as soon as $X_{i-1} \in \mathbf{V}^c$ because in this way the expected subsequent costs associated with $\bar{C}_j(\cdot, \cdot)$, $j = 0, 1, \dots, J$, would be zero.

Definition 4.5. Suppose that the induced MDP is semicontinuous and admits a feasible strategy with a finite value. We denote by Π the set of feasible strategies with a finite value that satisfy

$$\pi_i(db|h_{i-1}) := \delta_{f^*(x_{i-1})}(db), \quad \forall i = 1, 2, \dots, \quad (7)$$

whenever x_{i-1} , the last component of h_{i-1} , is in $\mathbf{V}^c$.

Now, given that the induced MDP is semicontinuous and admits a feasible strategy with a finite value, it is sufficient to restrict to the strategies from Π and concentrate on the subset $\mathbf{V}$. After some obvious modifications if necessary, one can regard, without loss of generality, that the optimal strategy π in problem (3), existing by Lemma 4.4, is from the set Π .

4.2. Relaxed problem and absence of relaxation gap

Definition 4.6. for each strategy π in the induced MDP, its occupation measure μ^π is defined by

$$\mu^\pi(\Gamma) := \mathbb{E}_{x_0}^\pi \left[\sum_{i=0}^{\infty} \mathbb{I}\{(X_i, B_{i+1}) \in \Gamma\} \right] \quad \forall \Gamma \in \mathcal{B}(\mathbf{X}_\Delta \times \mathbf{B}).$$

Recall that x_0 was assumed to be a fixed initial state in $\mathbf{V}$, though the above definition is meaningful for any initial state $x_0 \in \mathbf{X}_\Delta$.

The space of all occupation measures (under the fixed initial state x_0) is convex in the following sense: if μ_1, μ_2 are occupation measures for the induced MDP and $\beta \in [0, 1]$, then $\beta\mu_1 + (1-\beta)\mu_2$ is also an occupation measure. This fact follows from Chapter 5, Section 5 of [11]; see also Proposition C.2.13 of [21].

The induced MDP is an MDP with a cemetery in the sense of [22]. The next definition is consistent with Definition 7 in [22].

Definition 4.7. Let L be a fixed positive integer. A strategy π is called an L -mixture over $\{\pi_l\}_{l=1}^L \subseteq \Pi_G$ if $\mu^\pi|_{\mathbf{X} \times \mathbf{B}} = \sum_{l=1}^L \gamma_l \mu^{\pi_l}|_{\mathbf{X} \times \mathbf{B}}$ for some $\{\gamma_l\}_{l=1}^L \subseteq [0, 1]$ satisfying $\sum_{l=1}^L \gamma_l = 1$. The constants $\{\gamma_l\}_{l=1}^L$ are referred to as the coefficients or weights in the mixture.

It holds for each $j \in \{0, \dots, J\}$ that

$$\mathcal{V}(x_0, \pi, \bar{C}_j) = \int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu^\pi(dx \times db), \quad j = 0, 1, \dots, J.$$

Hence, problem (3) is equivalent to minimizing $\int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_j(x, b) \mu^\pi(dx \times db)$ over the space of occupation measures of feasible strategies in the induced MDP. As is well known (see Lemma 9.4.3 of [14]), each occupation measure μ satisfies equation

$$\mu(\Gamma \times \mathbf{B}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{X}_\Delta \times \mathbf{B}} Q(\Gamma|y, b) \mu(dy \times db), \quad \Gamma \in \mathcal{B}(\mathbf{X}_\Delta). \quad (8)$$

Lemma 4.8. Suppose the induced MDP is semicontinuous. Then for each measure μ on $\mathbf{X}_\Delta \times \mathbf{B}$, satisfying equation (8) and such that $\int_{\mathbf{X}_\Delta \times \mathbf{B}} \sum_{j=0}^J \bar{C}_j(x, b) \mu(dx \times db) < \infty$, it holds that $\mu_{\mathbf{X}_\Delta}|_{\mathbf{V}} \in \mathcal{M}^{\sigma<\infty}(\mathbf{V})$, where $\mathcal{M}^{\sigma<\infty}(\mathbf{V})$ is the set of σ -finite measures on $(\mathbf{V}, \mathcal{B}(\mathbf{V}))$. In particular, for each feasible strategy π with a finite value, it holds that $\mu_{\mathbf{X}_\Delta}^\pi|_{\mathbf{V}} \in \mathcal{M}^{\sigma<\infty}(\mathbf{V})$.

Proof. For the proof of the first assertion, see Theorem 3.2 of [10]. The last assertion follows from the first one, because for each feasible strategy π with a finite value, its occupation measure μ^π satisfies (8) and $\int_{\mathbf{X}_\Delta \times \mathbf{B}} \sum_{j=0}^J \bar{C}_j(x, b) \mu^\pi(dx \times db) = \sum_{j=0}^J \mathcal{V}(x_0, \pi, \bar{C}_j) < \infty$. $\square$

Suppose the induced MDP is semicontinuous and admits a feasible strategy with a finite value. According to the explanations at the end of the previous subsection (see (6)), for problem (3) it is sufficient to restrict to the set of strategies $\Pi \neq \emptyset$. That, in turn, is equivalent to minimizing $\int_{\mathbf{X}_\Delta \times \mathbf{B}} \bar{C}_0(x, b) \mu^\pi(dx \times db)$ over the space of occupation measures of strategies in Π in the induced MDP. We shall introduce a relaxation of the latter problem. To this end, let us now examine the occupation measures of strategies in Π and make two observations.

(α) If $\pi \in \Pi$, then for each measurable subset $\Gamma_1 \subset \mathbf{V}^c$,

$$\mu^\pi(\Gamma_1 \times \Gamma_2) = \mathbb{E}_{x_0}^\pi \left[\sum_{i=0}^{\infty} \mathbb{I}\{X_i \in \Gamma_1\} \mathbb{I}\{\Gamma_2 \ni f^*(X_i)\} \right] \quad \forall \Gamma_2 \in \mathcal{B}(\mathbf{B}).$$

Therefore,

$$\int_{\mathbf{V}^c \times \mathbf{B}} \bar{C}_j(x, b) \mu^\pi(dx \times db) = \mathbb{E}_{x_0}^\pi \left[\sum_{i=0}^{\infty} \mathbb{I}\{X_i \in \mathbf{V}^c\} \bar{C}_j(X_i, f^*(X_i)) \right] = 0, \quad j = 0, 1, \dots, J \quad (9)$$

by (6), and

$$\mathcal{V}(x_0, \pi, \bar{C}_j) = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^\pi(dx \times db), \quad j = 0, 1, \dots, J.$$

(β) For $\pi \in \Pi$, the measure $\mu = \mu^\pi|_{\mathbf{V} \times \mathbf{B}}$ on $\mathbf{V} \times \mathbf{B}$ satisfies the equation

$$\mu(\Gamma \times \mathbf{B}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{V} \times \mathbf{B}} Q(\Gamma|y, b) \mu(dy \times db), \quad \Gamma \in \mathcal{B}(\mathbf{V}),$$

because for each measurable subset $\Gamma \subset \mathbf{V}$, again by (6),

$$\int_{\mathbf{V}^c \times \mathbf{B}} Q(\Gamma|y, b) \mu^\pi(dy \times db) = \mathbb{E}_{x_0}^\pi \left[\sum_{i=0}^{\infty} \mathbb{I}\{X_i \in \mathbf{V}^c\} Q(\Gamma|X_i, f^*(X_i)) \right] = 0.$$

We are ready to formulate the following relaxation of problem (3) for the induced MDP, assumed to be semicontinuous and to admit a feasible strategy with a finite value:

$$\begin{aligned} &\text{Minimize} && \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) && (10) \\ &&& \text{over measures } \mu \text{ on } \mathbf{V} \times \mathbf{B} \text{ with } \mu_{\mathbf{V}} \in \mathcal{M}^{\sigma<\infty}(\mathbf{V}) \end{aligned}$$

$$\text{subject to} \quad \mu(\Gamma \times \mathbf{B}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{V} \times \mathbf{B}} Q(\Gamma|y, b) \mu(dy \times db), \quad \Gamma \in \mathcal{B}(\mathbf{V}) \quad (11)$$

$$\text{and} \quad \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) \leq d_j, \quad j = 1, 2, \dots, J. \quad (12)$$

The fact that the above problem is a relaxation of the MDP problem (3) is because of Items (α, β) in the above and Lemma 4.8. In particular,

$$\text{val}(\text{relaxed problem (10)-(12)}) \leq \text{val}(\text{problem (3)}) < \infty, \quad (13)$$

given that the induced MDP is semicontinuous and admits a feasible strategy with a finite value. We remark that in general, a measure μ satisfying equation (8) or (11) may not be an occupation measure of any strategy π . Thus, in general, one cannot conclude that the relaxed problem and the original problem are equivalent *a priori*. Theorem 4.11 below shows that there is no relaxation gap if the induced MDP is semicontinuous and has a feasible strategy with a finite value. It will follow from the next few observations.

Definition 4.9. Suppose the induced MDP is semicontinuous. Consider a measure μ on $\mathbf{V} \times \mathbf{B}$ with $\mu|_{\mathbf{V}} \in \mathcal{M}^{\sigma<\infty}(\mathbf{V})$. We say that a stationary strategy π^μ is induced by μ , if it has the following form

$$\pi^\mu(db|x) = \mathbb{I}\{x \in \mathbf{V}\} \varphi_\mu(db|x) + \mathbb{I}\{x \in \mathbf{V}^c\} \delta_{f^*(x)}(db),$$

where φ_μ is a stochastic kernel on $\mathbf{B}$ given $\mathbf{V}$ satisfying

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

The stochastic kernel φ_μ exists by Corollary 7.27.1 of [6] which should be applied to the sets $\{\Gamma_n \times \mathbf{B}\}_{n=1}^\infty$ that decompose $\mathbf{V} \times \mathbf{B}$ with $\mu(\Gamma_n \times \mathbf{B}) < \infty$. There may be multiple induced strategies by μ , different on $\mu(dx \times \mathbf{B})$ -null subsets of $\mathbf{V}$.

Lemma 4.10. Suppose the induced MDP is semicontinuous. Let a measure μ be feasible in problem (10)-(11) such that $\int_{\mathbf{V} \times \mathbf{B}} \sum_{j=0}^J \bar{C}_j(x, b) \mu(dx \times db) < \infty$. Then for each $j = 0, 1, \dots, J$

$$\mathbb{E}_{x_0}^{\pi^\mu} \left[\sum_{i=1}^{\infty} \bar{C}_j(X_{i-1}, B_i) \right] = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\pi^\mu}(dx \times db) \leq \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db).$$

Proof. For the proof, see Corollary 3.1 of [10]. In that corollary, integration was over $\mathbf{X}_\Delta \times \mathbf{B}$, but, similarly to (9), $\int_{\mathbf{V}^c \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\pi^\mu}(dx \times db) = 0$. $\square$

Note that under the conditions imposed in Lemma 4.10, if μ therein is feasible in the relaxed problem (10)-(12), then Π is non-empty, because the strategy π^μ therein belongs to Π .

Theorem 4.11. Suppose the induced MDP is semicontinuous and admits a feasible strategy π with a finite value. Then the following assertions hold.

- (a) $\text{val}(\text{problem (3)}) = \text{val}(\text{relaxed problem (10)-(12)}) < \infty$.
- (b) If a strategy $\pi^* \in \Pi$ is optimal in problem (3), then the measure $\mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$ is optimal in the relaxed problem (10)-(12). Note that an optimal strategy $\pi^* \in \Pi$ exists by Lemma 4.4 and explanations at the end of the previous subsection.
- (c) If a measure μ^* is optimal in the relaxed problem (10)-(12), then the induced strategy $\pi^{\mu^*} \in \Pi$ is optimal in problem (3).

Proof. (a) In view of (13), we only need to show that

$$\text{val}(\text{problem (3)}) \leq \text{val}(\text{relaxed problem (10)-(12)})$$

as follows. Consider an arbitrary measure μ that is feasible in the relaxed problem (10)-(12) and satisfies $\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) < \infty$. For the relaxed problem (10)-(12), it suffices to restrict to the set of such measures, which is non-empty because of (13). Then for the induced (by μ) strategy π^μ ,

$$\mathcal{V}(x_0, \pi^\mu, \bar{C}_j) \leq \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db), \quad j = 0, 1, \dots, J$$

by Lemma 4.10. Therefore, the induced strategy π^μ is feasible in problem (3) and

$$\mathcal{V}(x_0, \pi^\mu, \bar{C}_0) \leq \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db).$$

The arbitrariness of μ and this inequality lead to the required inequality.

(b) Let $\pi^* \in \Pi$ be an optimal strategy in problem (3). Then the restriction of its occupation measure $\mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$ is feasible in the relaxed problem (10)-(12) according to Lemma 4.8 and the observations made below it, see Items (α, β) therein. Moreover,

$$\begin{aligned} \mathcal{V}(x_0, \pi^*, \bar{C}_0) &= \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^{\pi^*}(dx \times db) = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}(dx \times db) \\ &= \text{val}(\text{problem (3)}) = \text{val}(\text{relaxed problem (10)-(12)}), \end{aligned}$$

where the last equality holds by (a). Thus, the measure $\mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$ is optimal in the relaxed problem (10)-(12).

(c) Let the measure μ^* be optimal in the relaxed problem (10)-(12). It holds that

$$\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^*(dx \times db) < \infty,$$

because $val(\text{relaxed problem (10)-(12)}) < \infty$ as was shown in (a). Now, Lemma 4.10 is applicable to μ^* , and we see that the induced (by μ^*) strategy π^{μ^*} is feasible in problem (3), and

$$\begin{aligned} \mathcal{V}(x_0, \pi^{\mu^*}, \bar{C}_0) &\leq \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^*(dx \times db) \\ &= val(\text{relaxed problem (10)-(12)}) = val(\text{problem (3)}) < \infty. \end{aligned}$$

Thus the strategy π^{μ^*} is optimal in problem (3). That π^{μ^*} belongs to Π follows from the last inequalities and the definitions of Π and π^{μ^*} . $\square$

4.3. Primal convex program and duality results

Suppose the induced MDP is semicontinuous and has a feasible strategy with a finite value. Then in view of Theorem 4.11, the relaxed problem (10)-(12) can be rewritten as the following one:

$$\begin{aligned} \text{Minimize} \quad & \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) \text{ over } \mu \in \mathcal{D} \\ \text{subject to} \quad & \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) - d_j \leq 0, \quad j = 1, 2, \dots, J, \end{aligned} \quad (14)$$

where $\mathcal{D}$ is the set of measures μ on $\mathbf{V} \times \mathbf{B}$ satisfying

- (a) $\mu_{\mathbf{V}} \in \mathcal{M}^{\sigma < \infty}(\mathbf{V})$,
- (b) equation (11), i.e., $\mu(dx \times \mathbf{B}) = \delta_{x_0}(dx) + \int_{\mathbf{V} \times \mathbf{B}} Q(dx|y, b) \mu(dy \times db)$, and
- (c)

$$\int_{\mathbf{V} \times \mathbf{B}} \sum_{j=0}^J \bar{C}_j(x, b) \mu(dx \times db) < \infty. \quad (15)$$

In this way, we may view the relaxed problem (10)-(12) as a convex program with J constraint inequalities, because the set $\mathcal{D}$ is convex in the sense that if $c \in [0, 1]$ and $\mu_1, \mu_2 \in \mathcal{D}$, then $c\mu_1 + (1 - c)\mu_2 \in \mathcal{D}$. It is also possible to embed $\mathcal{D}$ in a vector space, although this linear space is never needed in what follows. We call this convex program the primal convex program.

It is straightforward to check that the primal convex program (14) is equivalent to

$$\text{Minimize over } \mu \in \mathcal{D} : \sup_{\bar{g} \in \mathbb{R}_+^J} L_1(\mu, \bar{g}), \quad (16)$$

where $\mathbb{R}_+^J := [0, \infty)^J$, and

$$L_1(\mu, \bar{g}) := \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) + \sum_{j=1}^J g_j \left(\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) - d_j \right)$$

for $\mu \in \mathcal{D}$, $\bar{g} := (g_1, g_2, \dots, g_J) \in \mathbb{R}_+^J$ is the Lagrangian, which is real-valued due to Condition (c) in the definition of $\mathcal{D}$. In greater detail, this follows from the following three observations. (a) For any $\mu \in \mathcal{D}$, if it is infeasible in the primal convex program (14), then $\sup_{\bar{g} \in \mathbb{R}_+^J} L_1(\mu, \bar{g}) = \infty$ for some $\bar{g} \in \mathbb{R}_+^J$, meaning that in problem (16) it is sufficient to restrict to the set of $\mu \in \mathcal{D}$ feasible in the primal convex program (14). (b) This set is non-empty, because the optimal solution $\mu^* \in \mathcal{D}$ in the primal convex program (14) exists by Theorem 4.11(b). (c) If $\mu \in \mathcal{D}$ is feasible in the primal convex program (14), then $\sup_{\bar{g} \in \mathbb{R}_+^J} L_1(\mu, \bar{g}) = L_1(\mu, \mathbf{0}) = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db)$, where $\mathbf{0} = (0, 0, \dots, 0)$. In particular,

$$val(\text{primal convex program (14)}) = val(\text{problem (16)}). \quad (17)$$

and both problems have the same optimal solutions.

The dual problem for the primal convex program (16) reads

$$\text{Maximize over } \bar{g} \in \mathbb{R}_+^J : \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}), \quad (18)$$

It follows that $val(\text{problem (18)}) \leq val(\text{problem (16)})$. There is a duality gap if the strict inequality holds. Next, in Theorem 4.13, we observe the absence of the duality gap, under the following Slater condition, cf. (8.3) in [23].

Assumption 4.12 (Slater condition). There exists $\mu \in \mathcal{D}$ such that

$$\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) < d_j, \quad j = 1, 2, \dots, J.$$

In view of Lemma 4.10, given that the induced MDP is semicontinuous, Condition 4.12 implies that the induced MDP has a feasible strategy with a finite value.

Theorem 4.13. Suppose the induced MDP is semicontinuous and satisfies Condition 4.12 (i.e., the Slater condition). Then the following assertions hold.

(a) The duality gap is absent:

$$\begin{aligned} val(\text{primal convex program (16)}) &= \inf_{\mu \in \mathcal{D}} \sup_{\bar{g} \in \mathbb{R}_+^J} L_1(\mu, \bar{g}) \\ &= \sup_{\bar{g} \in \mathbb{R}_+^J} \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}), \end{aligned}$$

there exists at least one $\bar{g}^* \in \mathbb{R}_+^J$ solving the dual problem (18), and the dual functional $h(\cdot)$ defined for each $\bar{g} \in \mathbb{R}_+^J$ by

$$h(\bar{g}) := \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g})$$

is concave.

(b) Consider a pair $(\mu^*, \bar{g}^*) \in \mathcal{D} \times \mathbb{R}_+^J$. Then μ^* is an optimal solution to the primal convex program (16) and $\bar{g}^*$ is an optimal solution to the dual problem (18) if and only if one of the following two equivalent statements holds.

(i) The pair $(\mu^*, \bar{g}^*)$ is a saddle point of the Lagrangian:

$$L_1(\mu^*, \bar{g}) \leq L_1(\mu^*, \bar{g}^*) \leq L_1(\mu, \bar{g}^*), \quad \forall \mu \in \mathcal{D}, \bar{g} \in \mathbb{R}_+^J.$$

(ii) The following relations hold for the pair $(\mu^*, \bar{g}^*)$:

$$\begin{aligned} \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) &\leq d_j, \quad j = 1, 2, \dots, J; \\ L_1(\mu^*, \bar{g}^*) &= \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*); \\ \sum_{j=1}^J \bar{g}_j^* \left(\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) - d_j \right) &= 0. \end{aligned}$$

The last equality is known as the complementary slackness condition.

Proof. (a) The absence of the duality gap and the solvability of the dual program (18) follow from Theorem 1, Section 8.6 of [17], and the last assertion holds by Proposition 1, Section 8.6 of [17]. Note that in the framework of Section 8.3 and Section 8.6 of [17], $\mathcal{D}$ is embedded in a vector space. One can check that the reasoning therein for the results referred to in this proof holds if $\mathcal{D}$ is convex in the sense that if $c \in [0, 1]$ and $\mu_1, \mu_2 \in \mathcal{D}$, then $c\mu_1 + (1 - c)\mu_2 \in \mathcal{D}$.

(b) The first assertion in (b), i.e., the (primal and dual) optimality of μ^* and $\bar{g}^*$, is equivalent to (b-i) according to (a) and Theorem 2 of [23]. Assume that assertion (b-i) holds. Then, the first relation in (b-ii) holds, because μ^* is optimal and in particular feasible in the primal convex program (16). The second

relation in (b-ii) holds by (b-i). The last relation in (b-ii) holds by Theorem 1, Section 8.6 of [17]. It remains to prove that (b-ii) implies (b-i). Assume that (b-ii) holds. Then from $L_1(\mu^*, \bar{g}^*) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*)$ we see that $L_1(\mu^*, \bar{g}^*) \leq L_1(\mu, \bar{g}^*)$ for all $\mu \in \mathcal{D}$. From the complementary slackness condition, we see that $L_1(\mu^*, \bar{g}^*) = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^*(dx \times db)$. Now from $\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) \leq d_j$, $j = 1, 2, \dots, J$, we see that $L_1(\mu^*, \bar{g}^*) = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^*(dx \times db) \geq \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu^*(dx \times db) + \sum_{j=1}^J g_j \left(\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) - d_j \right) = L_1(\mu^*, \bar{g})$ for all $\bar{g} \in \mathbb{R}_+^J$. Thus,

$$L_1(\mu^*, \bar{g}) \leq L_1(\mu^*, \bar{g}^*) \leq L_1(\mu, \bar{g}^*), \quad \forall \mu \in \mathcal{D}, \quad \bar{g} \in \mathbb{R}_+^J,$$

i.e., (b-ii) implies (b-i). □

5. Main results

In this section we present the main results of this paper. Subsection 5.1 formulates a second pair of primal-dual programs for the original optimal impulse control problem and presents the duality results, whereas Subsection 5.2 presents a procedure for finding an optimal strategy for the optimal impulse control problem, which is in the form of a mixture over $J + 1$ deterministic stationary strategies.

5.1. Second convex program and duality results

Assumption 5.1. (a) $C_0^I(x, a) \geq \delta > 0$ for all $x \in \mathbf{X}$, $a \in \mathbf{A}$.

(b) $C = \max_{j=0,1,\dots,J} \sup_{x \in \mathbf{X}} C_j^g(x) + \max_{j=0,1,\dots,J} \sup_{(x,a) \in \mathbf{X} \times \mathbf{A}} C_j^I(x, a) < \infty$.

Suppose the induced MDP is semicontinuous. Under Condition 5.1(a), for each $x \in \mathbf{V}^c \setminus \{\Delta\}$, by (6), necessarily $f^*(x) = (\infty, \hat{a})$ for some $\hat{a} \in \mathbf{A}$. Hence, under Condition 5.1, for each $x \in \mathbf{V}^c \setminus \{\Delta\}$ $C_j^g(\phi(x, t)) = 0$ for almost all $t \geq 0$ with respect to the Lebesgue measure for all $j = 0, 1, \dots, J$. One can also see that, for each given $x \in \mathbf{X}$, $x \in \mathbf{V}^c$ if and only if $C_j^g(\phi(x, t)) \equiv 0$ a.e. with respect to the Lebesgue measure for all $j = 0, 1, \dots, J$, and $Q(\{\Delta\} | x, f^*(x)) = 1$.

Lemma 5.2. Suppose Condition 5.1(a) is satisfied. For each given strategy π , if $\mu^\pi(\mathbf{X} \times \mathbf{B}) = \infty$, then $\mathcal{V}(x_0, \pi, \bar{C}_0) = \infty$.

Proof. See Section 7. □

Suppose that the induced MDP is semicontinuous. Recall from Theorem 4.11 that the impulse control problem (3) is equivalent to the relaxed problem (10)-(12). This problem has a finite optimal value, and is solvable. In view of Lemma 5.2, if in addition, Condition 5.1(a) is also satisfied, then problem (10)-(12) is equivalent to the following convex program:

$$\text{Minimize} \quad \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) \text{ over } \mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B}) \quad (19)$$

$$\text{subject to} \quad \mu(\Gamma \times \mathbf{B}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{V} \times \mathbf{B}} Q(\Gamma | y, b) \mu(dy \times db), \quad \Gamma \in \mathcal{B}(\mathbf{V}) \quad (20)$$

$$\text{and} \quad \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) \leq d_j, \quad j = 1, 2, \dots, J, \quad (21)$$

where $\mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})$ is the set of finite measures on $(\mathbf{V} \times \mathbf{B}, \mathcal{B}(\mathbf{V} \times \mathbf{B}))$. We shall refer to problem (19)-(21) as the second (primal) convex program with one measure-valued constraint equality and J constraint inequalities to distinguish it from the primal convex program (14) with J constraint inequalities. Note that problem (19)-(21) is a linear program in the linear space of finite signed measures on $\mathbf{V} \times \mathbf{B}$.

Suppose that the induced MDP is semicontinuous and has a feasible strategy with a finite value, and Condition 5.1(a) is satisfied. We shall formulate a dual problem for the second convex program (19)-(21) and show the absence of the duality gap. Corresponding to the constraint equality in problem (19)-(21), let us introduce $\mathbb{B}(\mathbf{V})$, the set of bounded measurable functions on $\mathbf{V}$. Then we introduce the Lagrangian

$L_2(\cdot, \cdot, \cdot)$ defined for each $(\mu, W(\cdot), \bar{g}) \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B}) \times \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J$ by

$$\begin{aligned} L_2(\mu, W, \bar{g}) &:= \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_0(x, b) \mu(dx \times db) \\ &+ \int_{\mathbf{V}} W(x) \left[\delta_{x_0}(dx) + \int_{\mathbf{V}} Q(dx|y, b) \mu(dy \times db) - \mu(dx \times \mathbf{B}) \right] \\ &+ \sum_{j=1}^J g_j \left(\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db) - d_j \right). \end{aligned}$$

Since $\bar{C}_j(\cdot, \cdot)$ and g_j are non-negative, the Lagrangian $L_2(\cdot, \cdot, \cdot)$ is well defined and the terms therein can be arbitrarily rearranged, as $\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu(dx \times db)$ are the only possibly infinite terms, which are non-negative and with non-negative coefficients. Here $0 \times \infty := 0$ is accepted.

It is straightforward to check that the second primal convex (in fact, linear) program (19)-(21) is equivalent to

$$\text{Minimize over } \mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B}) : \sup_{(W(\cdot), \bar{g}) \in \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J} L_2(\mu, W, \bar{g}). \quad (22)$$

In fact, it follows from the observation formulated in the next remark.

Remark 5.3. If equality (20) is violated by some $\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})$, then there exist $\varepsilon > 0$ and $\Gamma \in \mathcal{B}(\mathbf{V})$ such that

$$\delta_{x_0}(\Gamma) + \int_{\mathbf{V}} Q(\Gamma|y, b) \mu(dy \times db) - \mu(\Gamma \times \mathbf{B}) \geq \varepsilon \text{ (or } \leq -\varepsilon).$$

Now for $W_N(x) = N\mathbb{I}\{x \in \Gamma\}$ (resp., for $W_N(x) = -N\mathbb{I}\{x \in \Gamma\}$),

$$\lim_{N \rightarrow \infty} L_2(\mu, W_N, \bar{g}) = +\infty.$$

The similar reasoning applies to the constraints (21). The conclusion is that in problem (22), one can be restricted to the set of measures μ that are feasible in the second primal convex program (19)-(21), and problem (22) restricted on this set becomes the same as the second primal convex program (19)-(21).

Now, we formulate the dual program for the second primal convex program (19)-(21) as

$$\text{Maximize over } (W(\cdot), \bar{g}) \in \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J : \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}). \quad (23)$$

One can solve the dual program (23) in two steps: firstly, for each fixed $\bar{g} \in \mathbb{R}_+^J$, solve problem

$$\text{Maximize over } W(\cdot) \in \mathbb{B}(\mathbf{V}) : \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}); \quad (24)$$

and after that solve problem

$$\text{Maximize over } \bar{g} \in \mathbb{R}_+^J : \sup_{W(\cdot) \in \mathbb{B}(\mathbf{V})} \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}).$$

Up to the additive constant $-\sum_{j=1}^J g_j d_j$, problem (24) can be rewritten as follows:

$$\text{Maximize } W(x_0) \text{ over } W(\cdot) \in \mathbb{B}(\mathbf{V}) \quad (25)$$

$$\begin{aligned} \text{subject to } & \bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) + \int_{\mathbf{V}} W(y) Q(dy|x, b) - W(x) \geq 0, \\ & (x, b) \in \mathbf{V} \times \mathbf{B}. \end{aligned} \quad (26)$$

Indeed, if inequality (26) is violated by some $W(\cdot) \in \mathbb{B}(\mathbf{V})$, then, for some $\varepsilon > 0$, $\bar{C}_0(\hat{x}, \hat{b}) + \sum_{j=1}^J g_j \bar{C}_j(\hat{x}, \hat{b}) + \int_{\mathbf{V}} W(y) Q(dy|\hat{x}, \hat{b}) - W(\hat{x}) < -\varepsilon$ for some $(\hat{x}, \hat{b}) \in \mathbf{V} \times \mathbf{B}$. Now, for the sequence of finite measures defined for each $N \in \{1, 2, \dots\}$ by $\mu_N(dx \times db) = N\delta_{(\hat{x}, \hat{b})}(dx \times db) \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})$, $\lim_{N \rightarrow \infty} L_2(\mu_N, W, \bar{g}) = -\infty$,

and thus $\inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}) = -\infty$. On the other hand, for each $W(\cdot) \in \mathbb{B}(\mathbf{V})$ satisfying (26), $\inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g})$ is provided by the null measure $\mu \equiv 0$ and equals $W(x_0) - \sum_{j=1}^J g_j d_j$.

It is observed in Theorem 5.5 below that for each $\bar{g} \in \mathbb{R}_+^J$, problem (25)-(26) is solvable. In fact, under the conditions imposed in Theorem 5.5, its optimal solution is given by the function $W_{\bar{g}}^*(\cdot)$ defined by

$$\begin{aligned} W_{\bar{g}}^*(x) &:= \inf_{\pi \in \Pi_G} \mathbb{E}_x^\pi \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] \\ &= \inf_{\pi \in \Pi_G} \left\{ \mathcal{V}(x, \pi, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \pi, \bar{C}_j) \right\}, \quad x \in \mathbf{V}. \end{aligned} \quad (27)$$

Thus, $W_{\bar{g}}^*(\cdot)$ is the restriction on $\mathbf{V}$ of the Bellman function $\mathcal{V}^*(\cdot, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$ defined for each $x \in \mathbf{X}_\Delta$ by $\mathcal{V}^*(x, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)(x) := \inf_{\pi \in \Pi_G} \mathcal{V}(x, \pi, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$. We also call $W_{\bar{g}}^*(\cdot)$ the Bellman function because

$$\mathcal{V}^*(x, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) = 0$$

for all $x \in \mathbf{V}^c$. The next lemma shows that $W_{\bar{g}}^*(\cdot)$ is a feasible solution in problem (25)-(26) and provides a useful characterization of the function $W_{\bar{g}}^*(\cdot)$.

Lemma 5.4. *Suppose the induced MDP is semicontinuous and Condition 5.1 is satisfied. Let $\bar{g} \in \mathbb{R}_+^J$ is arbitrarily fixed. Then the following assertions hold.*

- (a) $W_{\bar{g}}^*(\cdot)$ is bounded, non-negative and lower semicontinuous on $\mathbf{V}$, and $W_{\bar{g}}^*(\cdot)$ is a feasible solution in problem (25)-(26).
- (b) $W_{\bar{g}}^*(\cdot)$ is the unique lower semicontinuous bounded function on $\mathbf{V}$ satisfying the following equation:

$$W(x) = \inf_{b \in \mathbf{B}} \left\{ \bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) + \int_{\mathbf{V}} W(y) Q(dy|x, b) \right\}, \quad \forall x \in \mathbf{V}.$$

Moreover, if a function $W(\cdot) \in \mathbb{B}(\mathbf{V})$ satisfies

$$\begin{aligned} W(x) &= \inf_{b \in \mathbf{B}} \left\{ \bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) + \int_{\mathbf{V}} W(y) Q(dy|x, b) \right\} \\ &= \bar{C}_0(x, \hat{f}(x)) + \sum_{j=1}^J g_j \bar{C}_j(x, \hat{f}(x)) + \int_{\mathbf{V}} W(y) Q(dy|x, \hat{f}(x)), \quad x \in \mathbf{V} \end{aligned} \quad (28)$$

for a measurable mapping $\hat{f}(\cdot) : \mathbf{V} \rightarrow \mathbf{B}$, then the deterministic stationary strategy $\hat{f}$ (supplemented by $\hat{f}(x) := f^*(x)$ for $x \in \mathbf{V}^c$, see Definition 4.5) is uniformly optimal in the sense that

$$\mathcal{V}(x, \hat{f}, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \hat{f}, \bar{C}_j) = \mathcal{V}^*(x, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j), \quad x \in \mathbf{X}_\Delta.$$

Proof. See Section 7. □

Theorem 5.5. *Suppose the induced MDP is semicontinuous, and Conditions 4.12 and 5.1 are satisfied. Then the following assertions hold.*

- (a) For a fixed $\bar{g} \in \mathbb{R}_+^J$, the Bellman function $W_{\bar{g}}^*(\cdot)$ defined by (27) is bounded, non-negative and lower semicontinuous, and solves problem (25)-(26) with

$$val(\text{problem (25)-(26)}) = W_{\bar{g}}^*(x_0).$$

(b) $W_{\bar{g}}^*(x_0) - \sum_{j=1}^J g_j d_j = h(\bar{g})$, where $h(\cdot)$ was defined in Theorem 4.13(a). Thus, a maximizer $\bar{g}^* \in \mathbb{R}_+^J$ for

$$\max_{\bar{g} \in \mathbb{R}_+^J} \left\{ \max_{W(\cdot) \in \mathbb{B}(\mathbf{V})} \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}) \right\} = \max_{\bar{g} \in \mathbb{R}_+^J} h(\bar{g})$$

exists.

(c) The pair $(W_{\bar{g}^*}^*, \bar{g}^*)$ is a solution to the dual program (23).

(d) There is no duality gap:

$$\begin{aligned} & \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} \sup_{(W(\cdot), \bar{g}) \in \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J} L_2(\mu, W, \bar{g}) \\ &= \sup_{(W(\cdot), \bar{g}) \in \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J} \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}). \end{aligned}$$

Proof. See Section 7. □

In view of (27) and the definition of $h(\bar{g}) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g})$ given in Theorem 4.13(a), under the conditions imposed in Theorem 5.5, part (b) therein asserts that

$$\inf_{\pi \in \Pi_G} \mathbb{E}_x^\pi \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] - \sum_{j=1}^J g_j d_j = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g})$$

for any $\bar{g} \in \mathbb{R}_+^J$. This equality is attractive from the practical point of view, because it shows that the value of the dual functional $h(\bar{g})$ can be calculated using the dynamic programming approach, and this is further elaborated on in the next subsection. On the other hand, here we remark that Condition 5.1 is important for the last equality.

5.2. Procedures for finding $(J + 1)$ -mixed optimal strategies in the original impulse control problem

Throughout this subsection, we suppose that the induced MDP is semicontinuous, and Conditions 4.12 and 5.1 are satisfied, so that all the above results are applicable. In view of Theorem 5.5, we now present a general procedure for solving the primal convex problem (14) and the relaxed problem (10)-(12), leading to an optimal strategy in the form of a $(J + 1)$ -mixture of deterministic stationary strategies. The main result in this subsection is the justification of this procedure, see Theorem 5.6 below. An application of this procedure is illustrated in the next section.

Formulation of the procedure.

Step 1. For each fixed $\bar{g} \in \mathbb{R}_+^J$, compute the dual functional $h(\bar{g}) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g})$. According to Theorem 5.5(b), $W_{\bar{g}}^*(x_0) - \sum_{j=1}^J g_j d_j = h(\bar{g})$. Thus, it suffices to compute the Bellman function $W_{\bar{g}}^*(\cdot)$, which is bounded. To this end, since the induced MDP is positive and semicontinuous, one may apply the successive approximations, see Corollary 9.17.2 of [6].

Step 2. Find any $\bar{g}^*$ providing $\max_{\bar{g} \in \mathbb{R}_+^J} h(\bar{g})$. By Theorem 4.13(a) and Theorem 5.5(b), such $\bar{g}^*$ exists and is optimal in problem (18), and the function $h(\cdot)$ is concave. In this step, one can use any standard method of maximizing concave functions on $\mathbb{R}_+^J$.

Step 3. For $\bar{g}^*$ obtained in Step 2, calculate the Bellman function $W_{\bar{g}^*}^*(\cdot)$ as in Step 1. Let $\mathbf{F}$ denote the set of deterministic stationary strategies f providing the minimum in the Bellman equation:

$$\begin{aligned} W_{\bar{g}^*}^*(x) &= \inf_{b \in \mathbf{B}} \left\{ \bar{C}_0(x, b) + \sum_{j=1}^J g_j^* \bar{C}_j(x, b) + \int_{\mathbf{V}} W_{\bar{g}^*}^*(y) Q(dy|x, b) \right\} \\ &= \bar{C}_0(x, f(x)) + \sum_{j=1}^J g_j^* \bar{C}_j(x, f(x)) + \int_{\mathbf{V}} W_{\bar{g}^*}^*(y) Q(dy|x, f(x)), \quad x \in \mathbf{V} \end{aligned}$$

and satisfying $f(x) = f^*(x)$ for $x \in \mathbf{V}^c$. Note that the expression in the parentheses in the first line of the above equality is a lower semicontinuous function and $\mathbf{B}$ is compact, so that $\mathbf{F} \neq \emptyset$, see Proposition 7.33 of [6] and Theorem 3.3 of [12]. Moreover, according to Proposition 9.12 of [6], each $f \in \mathbf{F}$ is uniformly optimal in the sense that $\mathcal{V}(x, f, \bar{C}_0 + \sum_{j=1}^J g_j^* \bar{C}_j) = \inf_{\pi} \mathcal{V}(x, \pi, \bar{C}_0 + \sum_{j=1}^J g_j^* \bar{C}_j) = W_{\bar{g}^*}^*(x)$ for all $x \in \mathbf{X}_{\Delta}$.

Step 4. Find constants $\gamma_l \in (0, 1]$ and deterministic stationary strategies $f^l \in \mathbf{F}$ (not necessarily all distinct) for $l = 1, 2, \dots, J+1$ satisfying

- (a) $\sum_{l=1}^{J+1} \gamma_l = 1$;
- (b) for the convex combination $\mu^* = \sum_{l=1}^{J+1} \gamma_l \mu^{f^l}|_{\mathbf{V} \times \mathbf{B}}$ of the restrictions of the occupation measures of f^l , it holds for each $j = 1, 2, \dots, J$ that $\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) \leq d_j$, whereas $\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) = d_j$ if $g_j^* > 0$.

Provided that we can find such constants γ_l and strategies f^l , μ^* would solve the primal convex program (14) as well as the relaxed problem (10)-(12). This can be seen as follows. From the proof of Theorem 5.5(a), we see for each $f \in \mathbf{F}$ that $\mu^f(\mathbf{V} \times \mathbf{B}) < \infty$, and from the inequality $\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^*(dx \times db) \leq d_j$ (and keeping in mind that $\gamma_l > 0$ for all $l = 1, 2, \dots, J+1$) we deduce that, for all $j = 1, 2, \dots, J$,

$$\mathcal{V}(x_0, f^l, \bar{C}_j) = \int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^{f^l}(dx \times db) < \infty.$$

Therefore, $\mu^{f^l} \in \mathcal{D}$ for all $l = 1, 2, \dots, J+1$ and $\mu^* \in \mathcal{D}$, too. Furthermore, from the proof of Theorem 5.5(b), we see that

$$L_1(\mu^{f^l}|_{\mathbf{V} \times \mathbf{B}}, \bar{g}^*) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*) = W_{\bar{g}^*}^*(x_0) - \sum_{j=1}^J g_j^* d_j.$$

It follows that

$$L_1(\mu^*, \bar{g}^*) = \sum_{l=1}^{J+1} \gamma_l L_1(\mu^{f^l}, \bar{g}^*) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*).$$

Thus, the pair $(\mu^*, \bar{g}^*) \in \mathcal{D} \times \mathbb{R}_+^J$ satisfies the conditions in Theorem 4.13(b-ii), and the optimality of μ^* now follows from Theorem 4.13(b). Moreover, $\mu^* = \mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$ for the $(J+1)$ -mixture π^* of the strategies $\{f^l\}_{l=1}^{J+1}$.

The next result shows that one can always find such constants $\{\gamma_l\}_{l=1}^{J+1}$ and deterministic stationary strategies $\{f^l\}_{l=1}^{J+1}$ as in Step 4.

Theorem 5.6. *Suppose the induced MDP is semicontinuous, and Conditions 4.12 and 5.1 are satisfied. Then there exists a mixture $\pi^* \in \Pi$ of $J+1$ deterministic stationary strategies $\{f^l\}_{l=1}^{J+1} \subseteq \mathbf{F}$ with coefficients $\{\gamma_l\}_{l=1}^{J+1} \subseteq (0, 1]$, such that π^* is optimal in the original impulse control problem (3) and $\mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$ solves the primal convex program (14) and the relaxed problem (10)-(12).*

Proof. See Section 7. □

Consider π^* as in Theorem 5.6. Let $\mu^* := \mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$. Since μ^* solves the primal convex program (14), for any $\bar{g}^* \in \mathbb{R}_+^J$ providing the maximum of the dual functional $h(\bar{g})$, we see from Theorem 4.13(b) that condition (b) in Step 4 of the general procedure is satisfied by μ^* , and there exist constants $\{\gamma_l\}_{l=1}^{J+1} \subseteq (0, 1]$ and strategies $\{f^l\}_{l=1}^{J+1} \subset \mathbf{F}$ satisfying the both conditions (a) and (b) in Step 4.

6. Illustrative example: fluid queueing system

In this section, we illustrate the main results in the last section with an example.

Example 6.1 (fluid queueing system). Consider the impulse control problem with the following system primitives. Let the state space be $\mathbf{X} = [0, \infty)$, and the action space $\mathbf{A} = \{1\}$ be a singleton, so that we omit the action $a \in \mathbf{A}$ in the description of the model below. The flow is described by $\phi(x, t) = x + t$ for all $x \in \mathbf{X}$ and $t \in [0, \infty)$. The new state upon applying an impulse is $l(x) \equiv 0$. Let $\alpha > 0$, $x_0 = 0$. We consider the case

of a single constraint, i.e., $J = 1$, so that the index $j = 1$ will be omitted below. Let the constraint constant be $d > 0$. The impulse cost functions and running cost rates are given by

$$C_0^I(x) \equiv K > 0, C_0^g(x) \equiv 0, C_1^I(x) \equiv 0, C_1^g(x) \equiv h > 0,$$

where $h \in (0, \infty)$ and $K \in (0, \infty)$ are fixed constants.

Interpretation of the example. This impulse control problem formulated in Example 6.1 can be given the following interpretation. Consider an infinite buffer: small packets of information arriving at a high speed are served with the high intensity. In such a case, this queueing system is known to be well approximated by the so called fluid model $\dot{x} = \lambda - \mu$. Here $x \geq 0$ is the buffer filling size measured, e.g., in Mb and λ (or μ) is the arrival (resp., service) rate measured in Mb/time unit. We assume that $\lambda > \mu$ and choose the time scale such that $\lambda - \mu = 1$. Thus, the flow is described by $\phi(x, t) = x + t$ and the state space is $\mathbf{X} = [0, \infty)$. At any time moment the decision maker can use an additional very powerful server; as the result, the current state instantly changes to zero. Thus, we put $\mathbf{A} = \{1\}$. The result of applying the impulse is described by the constant function $l(x) \equiv 0$. The price for one impulse is $K > 0$, and the holding cost is $h > 0$ per unit per time unit. $\alpha > 0$ is the discount factor as usual. We will fix the initial state $x_0 = 0$ and minimize the total discounted cost of impulses under the constraint $d > 0$ on the total discounted holding cost.

Induced MDP. One can compute the cost functions and write down the transition kernel in the induced MDP as follows:

$$\begin{aligned} \bar{C}_0(x, \theta) &= \begin{cases} Ke^{-\alpha\theta} & \text{if } \theta < \infty; \\ 0, & \text{if } \theta = \infty, \end{cases} \\ \bar{C}_1(x, \theta) &= \begin{cases} h \int_0^\theta e^{-\alpha t} (x + t) dt \\ = \frac{hx}{\alpha} (1 - e^{-\alpha\theta}) + h \left[\frac{1}{\alpha^2} - \frac{1}{\alpha^2} e^{-\alpha\theta} - \frac{\theta}{\alpha} e^{-\alpha\theta} \right], & \text{if } \theta < \infty; \\ \frac{hx}{\alpha} + \frac{h}{\alpha^2}, & \text{if } \theta = \infty, \end{cases} \\ Q(dy|x, \theta) &= \begin{cases} e^{-\alpha\theta} \delta_0(dy) + (1 - e^{-\alpha\theta}) \delta_\Delta(dy) & \text{if } \theta < \infty, x \neq \Delta; \\ \delta_\Delta(dy) & \text{otherwise.} \end{cases} \end{aligned}$$

Since there is a single impulse action, which we omit for brevity, a strategy is identified with the sequence of stochastic kernels $\pi_i(d\theta|x_0, \theta_1, x_1, \dots, x_{i-1})$ on $[0, \infty]$.

Condition 4.2 is obviously satisfied, and thus the induced MDP is semicontinuous by Proposition 4.3. Condition 5.1 is obviously satisfied, too. According to the comments below Condition 5.1, $\mathbf{V} = \mathbf{X}$. In order to check Condition 4.12, consider the deterministic stationary strategy

$$\hat{f}(x) = \begin{cases} \hat{\theta} - x, & \text{if } x \leq \hat{\theta}; \\ 0, & \text{if } x > \hat{\theta}, \end{cases}$$

i.e., apply the impulse as soon as $x \geq \hat{\theta}$, where $\hat{\theta} > 0$ is fixed. Then, during one cycle formed by two consecutive visits to state 0, the holding cost equals $\bar{C}_1(0, \hat{\theta})$, and

$$\mathcal{V}(0, \hat{f}, \bar{C}_1) = \sum_{i=0}^{\infty} e^{-\alpha\hat{\theta}i} \bar{C}_1(0, \hat{\theta}) = \frac{h \left[\frac{1}{\alpha^2} - \frac{1}{\alpha^2} e^{-\alpha\hat{\theta}} - \frac{\hat{\theta}}{\alpha} e^{-\alpha\hat{\theta}} \right]}{1 - e^{-\alpha\hat{\theta}}}.$$

The above quantity depends on $\hat{\theta}$, and one can compute its limit

$$\lim_{\hat{\theta} \downarrow 0} \frac{h \left[\frac{1}{\alpha^2} - \frac{1}{\alpha^2} e^{-\alpha\hat{\theta}} - \frac{\hat{\theta}}{\alpha} e^{-\alpha\hat{\theta}} \right]}{1 - e^{-\alpha\hat{\theta}}} = \lim_{\hat{\theta} \downarrow 0} h \frac{\frac{1}{\alpha} e^{-\alpha\hat{\theta}} - \frac{1}{\alpha} e^{-\alpha\hat{\theta}} + \theta e^{-\alpha\hat{\theta}}}{\alpha e^{-\alpha\hat{\theta}}} = 0.$$

Therefore, for the constraint constant $d > 0$, when choosing small enough $\hat{\theta} > 0$, one can satisfy the constraint $\mathcal{V}(0, \hat{f}, \bar{C}_1) < d$. In the meanwhile, similarly to the above,

$$\mathcal{V}(0, \hat{f}, \bar{C}_0) = \frac{Ke^{-\alpha\hat{\theta}}}{1 - e^{-\alpha\hat{\theta}}} < \infty.$$

In view of Lemma 4.8, we see that Condition 4.12 (i.e., the Slater condition) holds for $\mu = \mu^f|_{\mathbf{V} \times \mathbf{B}}$ with a sufficiently small $\hat{\theta} > 0$; $\mu(dx \times \mathbf{B}) = \delta_0(dx) \frac{1}{1-e^{-\alpha\hat{\theta}}} \in \mathcal{D}$ is a finite measure.

Solving the impulse control problem in Example 6.1. We now solve the impulse control problem in Example 6.1 by following the procedures in Subsection 5.2. We have seen that all the conditions in that subsection are satisfied in this example. Since $J = 1$, we omit the bar in the Lagrange multiplier and write just g and g^* .

Step 1. We find $W_g^*(\cdot)$ for $g \in [0, \infty)$.

First, let $g > 0$. In view of Lemma 5.4, let us solve for the unique non-negative bounded lower semicontinuous solution to the following equation:

$$\begin{aligned} W(x) &= \inf_{\theta \in [0, \infty]} \left\{ \bar{C}_0(x, \theta) + g \bar{C}_1(x, \theta) + \int_{\mathbf{X}} W(y) Q(dy|x, \theta) \right\} \\ &= \inf_{\theta \in [0, \infty]} \tilde{G}(x, \theta), \quad x \in \mathbf{V}, \end{aligned} \quad (29)$$

where $\tilde{G}(x, \theta) = Ke^{-\alpha\theta} + \frac{ghx}{\alpha}(1 - e^{-\alpha\theta}) + gh \left\{ \frac{1}{\alpha^2} - \frac{1}{\alpha^2}e^{-\alpha\theta} - \frac{\theta}{\alpha}e^{-\alpha\theta} \right\} + e^{-\alpha\theta}W(0)$.

When $x = 0$, we have

$$W(0) = \inf_{\theta \in [0, \infty]} \left\{ Ke^{-\alpha\theta} + gh \left\{ \frac{1}{\alpha^2} - \frac{1}{\alpha^2}e^{-\alpha\theta} - \frac{\theta}{\alpha}e^{-\alpha\theta} \right\} + e^{-\alpha\theta}W(0) \right\}.$$

Let us denote $G(\theta) = Ke^{-\alpha\theta} + gh \left\{ \frac{1}{\alpha^2} - \frac{1}{\alpha^2}e^{-\alpha\theta} - \frac{\theta}{\alpha}e^{-\alpha\theta} \right\} + e^{-\alpha\theta}W(0)$. Note that $G(\cdot)$ is continuous on $[0, \infty]$, $G(0) = K + W(0) \neq W(0)$ because $K > 0$, so that $\inf_{\theta \in [0, \infty]} G(\theta) = \inf_{\theta \in (0, \infty)} G(\theta)$, and $G(\infty) = \frac{gh}{\alpha^2}$. The derivative of $G(\theta)$ in $\theta \in [0, \infty)$ is computed as follows:

$$\begin{aligned} \frac{dG(\theta)}{d\theta} &= -\alpha Ke^{-\alpha\theta} + gh \left\{ \frac{1}{\alpha}e^{-\alpha\theta} - \frac{1}{\alpha}e^{-\alpha\theta} + \theta e^{-\alpha\theta} \right\} - \alpha e^{-\alpha\theta}W(0) \\ &= -\alpha Ke^{-\alpha\theta} + gh\theta e^{-\alpha\theta} - \alpha e^{-\alpha\theta}W(0) = e^{-\alpha\theta} \{-\alpha K + gh\theta - \alpha W(0)\} = e^{-\alpha\theta} \tilde{H}(\theta), \end{aligned}$$

where $\tilde{H}(\theta) = -\alpha K + gh\theta - \alpha W(0)$. Observe that $\tilde{H}(\theta)$ is strictly increasing in θ with $\tilde{H}(0) = \lim_{\theta \rightarrow 0} \tilde{H}(\theta) = -\alpha K - \alpha W(0) < 0$ because $K > 0$ and $W(0) \geq 0$, and $\lim_{\theta \rightarrow \infty} \tilde{H}(\theta) = \infty$. Thus, $\inf_{\theta \in [0, \infty]} G(\theta) = \inf_{\theta \in (0, \infty)} G(\theta)$ and $G(\theta)$ admits a unique minimizer $\theta^* \in (0, \infty)$ given by the unique stationary point, satisfying

$$\theta^* = \frac{\alpha(W(0) + K)}{gh}; \quad W(0) = G(\theta^*). \quad (30)$$

Therefore,

$$\begin{aligned} W(0) &= Ke^{-\alpha\theta^*} + \frac{\alpha W(0) + \alpha K}{\theta^*} \left\{ \frac{1}{\alpha^2} - \frac{1}{\alpha^2}e^{-\alpha\theta^*} - \frac{\theta^*}{\alpha}e^{-\alpha\theta^*} \right\} + e^{-\alpha\theta^*}W(0) \\ &= Ke^{-\alpha\theta^*} + \frac{\alpha W(0)}{\theta^*} \left\{ (1 - e^{-\alpha\theta^*}) \frac{1}{\alpha^2} \right\} + \frac{K}{\theta^*} \left\{ \frac{1}{\alpha} - \frac{1}{\alpha}e^{-\alpha\theta^*} - \theta^* e^{-\alpha\theta^*} \right\}, \end{aligned}$$

so that

$$W(0)(\alpha\theta^* - 1 + e^{-\alpha\theta^*}) = \alpha\theta^* Ke^{-\alpha\theta^*} + K \{1 - e^{-\alpha\theta^*} - \alpha\theta^* e^{-\alpha\theta^*}\}. \quad (31)$$

Substituting $W(0) = \frac{\theta^* gh}{\alpha} - K$ (see (30)) in the above, we see that

$$\begin{aligned} &\frac{\theta^* gh}{\alpha}(\alpha\theta^* - 1 + e^{-\alpha\theta^*}) - K(\alpha\theta^* - 1 + e^{-\alpha\theta^*}) \\ &= \alpha\theta^* Ke^{-\alpha\theta^*} + K \{1 - e^{-\alpha\theta^*} - \alpha\theta^* e^{-\alpha\theta^*}\}, \end{aligned}$$

i.e., $\frac{\theta^* gh}{\alpha}(\alpha\theta^* - 1 + e^{-\alpha\theta^*}) = K\alpha\theta^*$, or, after a rearrangement,

$$\frac{gh}{\alpha} (1 - e^{-\alpha\theta^*}) = gh\theta^* - K\alpha.$$

Standard analysis shows that there is a unique solution in $(0, \infty)$ to the equation

$$\frac{gh}{\alpha}(1 - e^{-\alpha x}) = ghx - K\alpha.$$

We denote that unique solution in $(0, \infty)$ by x_g . Thus $\theta^* = x_g$. Now from (31), we see that

$$W(0) = \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}}. \quad (32)$$

Now consider equation (29), i.e., $W(x) = \inf_{\theta \in [0, \infty]} \tilde{G}(x, \theta)$. Note that $\tilde{G}(x, \theta)$ is continuous in $\theta \in [0, \infty]$. The derivative of function $\tilde{G}(x, \theta)$ is given by

$$\begin{aligned} \frac{d\tilde{G}(x, \theta)}{d\theta} &= -\alpha K e^{-\alpha\theta} + ghx e^{-\alpha\theta} + gh \left\{ \frac{1}{\alpha} e^{-\alpha\theta} - \frac{1}{\alpha} e^{-\alpha\theta} + \theta e^{-\alpha\theta} \right\} - \alpha e^{-\alpha\theta} W(0) \\ &= -\alpha K e^{-\alpha\theta} + ghx e^{-\alpha\theta} + gh\theta e^{-\alpha\theta} - \alpha e^{-\alpha\theta} W(0) \\ &= e^{-\alpha\theta} \{-\alpha K + ghx + gh\theta - \alpha W(0)\} \\ &=: F(x, \theta). \end{aligned}$$

At $\theta = 0$, $\tilde{G}(x, 0) = K + W(0) > 0$, and $F(x, 0) = -\alpha K + ghx - \alpha W(0) = gh(x - \theta^*) = gh(x - x_g)$, by (30).

If $x < x_g$, then $F(x, \theta) < 0$ for small θ . Under a fixed $x < x_g$, $F(x, \theta) = 0$ at $\theta^*(x) = \frac{\alpha(W(0)+K)-ghx}{gh} = \frac{gh\theta^* - ghx}{gh} = \theta^* - x = x_g - x$ by (30), and $F(x, \theta) > 0$ for $\theta > x_g - x$. Thus, $\theta^*(x) = x_g - x$ is the unique minimizer of $\tilde{G}(x, \theta)$, providing the minimum in

$$\begin{aligned} W(x) &= \inf_{\theta \in [0, \infty]} \tilde{G}(x, \theta) = \tilde{G}(x, \theta^*(x)) \\ &= K e^{-\alpha(x_g - x)} + \frac{ghx}{\alpha} \left(1 - e^{-\alpha(x_g - x)} \right) + gh \left\{ \frac{1}{\alpha^2} - \frac{1}{\alpha^2} e^{-\alpha(x_g - x)} - \frac{x_g - x}{\alpha} e^{-\alpha(x_g - x)} \right\} \\ &\quad + e^{-\alpha(x_g - x)} \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}}, \end{aligned}$$

where (32) is in use.

If $x = x_g$, then $F(x, 0) = 0$, and $F(x, \theta) > 0$ for all $\theta \in (0, \infty)$ because the expression in parentheses in the formula for $F(x, \theta)$ is linear in θ . Thus, $\theta^*(x) = 0 = x_g - x$ is the unique minimizer of $\tilde{G}(x, \theta)$, providing the minimum in $\inf_{\theta \in [0, \infty]} \tilde{G}(x, \theta) = W(x) = K + W(0) = K + \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}}$, where (32) is in use.

If $x > x_g$, then $F(x, \theta) > 0$ for all $\theta \in [0, \infty)$, and $\theta^*(x) = 0$ is the unique minimizer of $\tilde{G}(x, \theta)$, providing the minimum in $\inf_{\theta \in [0, \infty]} \tilde{G}(x, \theta) = W(x) = K + W(0) = K + \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}}$.

Note that we have obtained the non-negative continuous and bounded function $W(\cdot)$ satisfying (29). Thus, $W_g^*(\cdot) = W(\cdot)$. Moreover, the minimum in (29) is provided by unique minimizer $f(x) = \max\{x_g - x, 0\}$.

Second, let $g = 0$. In this case, the non-negative bounded function $W_g^*(\cdot) = W_0^*(\cdot)$ satisfies equation $W_0^*(x) = \inf_{\theta \in [0, \infty]} \{K e^{-\alpha\theta} + e^{-\alpha\theta} W_0^*(0)\} = 0$, where the minimum is provided by the unique minimizer $f(x) = \infty$.

If we extend x_g by letting $x_g = \infty$ at $g = 0$, then for all $g \in [0, \infty)$, the unique minimizer providing $W_g^*(x) = \inf_{\theta \in [0, \infty]} \{\tilde{C}_0(x, \theta + g\tilde{C}_1(x, \theta)) + \int_{\mathcal{X}} W_g^*(y) Q(dy|x, \theta)\}$ is given by the single deterministic stationary strategy $f(x) = \max\{x_g - x, 0\}$. Note that, after the extension, x_g , viewed as a function of $g \in [0, \infty)$, provides a bijection from $[0, \infty)$ onto $(0, \infty]$. When $g > 0$, $g \in (0, \infty)$ and $x_g \in (0, \infty)$ are in 1-1 correspondence because for any fixed $x > 0$ one can put

$$g = \frac{K\alpha^2}{\alpha hx - h + h e^{-\alpha x}}, \quad (33)$$

so that $x = x_g$. Here $g > 0$ because the denominator equals zero at $x = 0$, and its derivative is $\alpha h - \alpha h e^{-\alpha x} > 0$ at $x > 0$. Clearly, x_g decreases when g increases and $\lim_{g \downarrow 0} x_g = \infty$, $\lim_{g \rightarrow \infty} x_g = 0$.

We summarize the key observations in this step in the next statement.

Lemma 6.2. Consider the impulse control problem in Example 6.1. Suppose $g > 0$ and let $x_g > 0$ be the unique solution to equation

$$\frac{gh}{\alpha}(1 - e^{-\alpha x}) = ghx - K\alpha. \quad (34)$$

Then the following assertions hold

(a) If $x \leq x_g$, then

$$\begin{aligned} W_g^*(x) &= Ke^{-\alpha(x_g-x)} + \frac{ghx}{\alpha} \left(1 - e^{-\alpha(x_g-x)}\right) \\ &\quad + gh \left[\frac{1}{\alpha^2} - \frac{1}{\alpha^2} e^{-\alpha(x_g-x)} - \frac{x_g - x}{\alpha} e^{-\alpha(x_g-x)} \right] \\ &\quad + e^{-\alpha(x_g-x)} \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}}. \end{aligned}$$

(b) If $x > x_g$, then

$$W_g^*(x) = K + W_g^*(0) = K + \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}}.$$

In each case, the single uniformly optimal deterministic stationary strategy f is given by

$$f(x) = \max\{x_g - x, 0\}.$$

In words, one has to apply the impulse if and only if the current state $x \geq x_g$.

When $g = 0$, $W_g^*(x) \equiv 0$ and $f(x) \equiv \infty$ is the single uniformly optimal deterministic stationary strategy, i.e., no impulses should be applied.

The dual functional $h(\cdot)$ has the form $h(g) = W_g^*(0) - gd = \frac{K(1 - e^{-\alpha x_g})}{\alpha x_g - 1 + e^{-\alpha x_g}} - gd$.

Step 2. In this step, we find the maximizer of the dual functional $h(\cdot)$. Since x_g provides a bijection from $[0, \infty)$ onto $(0, \infty]$, as was observed in Step 1, it is convenient to express $h(g)$ in terms of x_g using formulas (32) and (33): one has to maximize function

$$H(x) = \frac{K(1 - e^{-\alpha x})}{\alpha x - 1 + e^{-\alpha x}} - \frac{K\alpha^2 d}{h(\alpha x - 1 + e^{-\alpha x})},$$

valid also for $x = x_g = \infty$ corresponding to $g = 0$. After that, g^* is calculated using (33) with x^* being the maximizer of $H(\cdot)$.

The derivative of $H(x)$ in $x \in (0, \infty)$ is computed as follows:

$$\begin{aligned} \frac{dH(x)}{dx} &= \frac{h^2 K \alpha e^{-\alpha x} (\alpha x - 1 + e^{-\alpha x}) - [hK(1 - e^{-\alpha x}) - K\alpha^2 d]h(\alpha - \alpha e^{-\alpha x})}{h^2 (\alpha x - 1 + e^{-\alpha x})^2} \\ &= \frac{hK}{h^2 (\alpha x - 1 + e^{-\alpha x})^2} \{ h\alpha^2 x e^{-\alpha x} - h\alpha e^{-\alpha x} + h\alpha e^{-2\alpha x} \\ &\quad - h\alpha + h\alpha e^{-\alpha x} + h\alpha e^{-\alpha x} - h\alpha e^{-2\alpha x} + \alpha^3 d - \alpha^3 d e^{-\alpha x} \} \\ &= \frac{K\alpha}{h(\alpha x - 1 + e^{-\alpha x})^2} \{ \alpha h x e^{-\alpha x} - (1 - e^{-\alpha x})(h - d\alpha^2) \} \\ &= \frac{K\alpha}{h(\alpha x - 1 + e^{-\alpha x})^2} D(x), \end{aligned}$$

where

$$D(x) := \alpha h x e^{-\alpha x} - (1 - e^{-\alpha x})(h - d\alpha^2) = (\alpha h x + (h - \alpha^2 d))e^{-\alpha x} - (h - \alpha^2 d).$$

Observe that $\lim_{x \rightarrow 0} D(x) = 0$. Moreover,

$$\begin{aligned} \frac{dD(x)}{dx} &= \frac{d}{dx} \{ (\alpha h x + (h - \alpha^2 d))e^{-\alpha x} - (h - \alpha^2 d) \} \\ &= \alpha h e^{-\alpha x} - \alpha^2 h x e^{-\alpha x} - \alpha(h - \alpha^2 d)e^{-\alpha x} \\ &= (h - \alpha h x - (h - \alpha^2 d))e^{-\alpha x} \alpha = \alpha^2 e^{-\alpha x} (\alpha d - h x). \end{aligned}$$

Since $\lim_{x \rightarrow 0} \frac{dD(x)}{dx} = \alpha^3 d > 0$, and function $\alpha d - hx$ strictly decreases in $x \in (0, \infty)$ with $\lim_{x \rightarrow \infty} (\alpha d - hx) = -\infty$, we see that $D(x)$ increases from $0 = \lim_{x \rightarrow 0} D(x)$ for $x > 0$ up to its unique stationary point, and, beyond that point, strictly decreases to $\alpha^2 d - h = \lim_{x \rightarrow \infty} D(x)$ as $x \rightarrow \infty$.

We now distinguish two cases.

Case (a) $d < \frac{h}{\alpha^2}$. In this case, $\alpha^2 d - h = \lim_{x \rightarrow \infty} D(x) < 0$, and we see that $\frac{dH(x)}{dx} > 0$ for small $x > 0$ up to the unique stationary point of $H(\cdot)$, say $x^* \in (0, \infty)$, and $\frac{dH(x)}{dx} < 0$ beyond that point. Thus, the unique maximizer of $H(x)$ is given by $x^* \in (0, \infty)$, i.e., the unique maximizer of the dual functional $h(g)$ is given by

$$g^* = \frac{K\alpha^2}{h(\alpha x^* - 1 + e^{-\alpha x^*})} \in (0, \infty)$$

in view of (33). Since $\frac{dH(x)}{dx} = 0$ if and only if $D(x) = 0$, and $H(\cdot)$ has a unique stationary point on $(0, \infty)$, we see that equation

$$\alpha h x e^{-\alpha x} - (1 - e^{-\alpha x})(h - d\alpha^2) = 0 \quad (35)$$

has a unique solution in $(0, \infty)$ given by x^* .

Case (b) $d \geq \frac{h}{\alpha^2}$. In this case, $\alpha^2 d - h = \lim_{x \rightarrow \infty} D(x) \geq 0$, and, from the discussion above Case (a), we see that $H(x)$ is nondecreasing, and thus ∞ is the maximizer of $H(x)$. In view of the 1-1 correspondence between $g \in [0, \infty)$ and $x_g \in (0, \infty]$, we see that $g^* = 0$ is a maximizer of the dual functional $h(g)$.

Steps 3 and 4. In Case (a) of Step 2, i.e., when $d < \frac{h}{\alpha^2}$, we have obtained $g^* \in (0, \infty)$. From Step 1, see Lemma 6.2, we see that $\mathbf{F}$ is a singleton consisting of the deterministic stationary strategy $f^*(x) = \max\{x_{g^*} - x, 0\}$, where $x_{g^*} = x^*$ comes from equation (35). According to Theorem 5.6, f^* is necessarily the optimal strategy to the original impulse control problem (3). One can also explicitly show that $\mathcal{V}(0, f^*, \bar{C}_1) = d$:

$$\begin{aligned} \mathcal{V}(0, f^*, \bar{C}_1) &= \sum_{i=0}^{\infty} e^{-\alpha x^* i} \bar{C}_1(0, x^*) = \frac{h \left[\frac{1}{\alpha^2} - \frac{1}{\alpha^2} e^{-\alpha x^*} - \frac{x^*}{\alpha} e^{-\alpha x^*} \right]}{1 - e^{-\alpha x^*}} \\ &= \frac{h}{\alpha^2} - \frac{h x^*}{\alpha} e^{-\alpha x^*} \cdot \frac{h - d\alpha^2}{\alpha h x^* e^{-\alpha x^*}} = \frac{h}{\alpha^2} - \left[\frac{h}{\alpha^2} - d \right] = d, \end{aligned}$$

where the third equality is by (35). Finally,

$$\mathcal{V}(0, f^*, \bar{C}_0) = \sum_{i=0}^{\infty} e^{-\alpha x^* i} \bar{C}_0(0, x^*) = \frac{K e^{-\alpha x^*}}{1 - e^{-\alpha x^*}}.$$

In Case (b) of Step 2, i.e., when $d \geq \frac{h}{\alpha^2}$, we have obtained $g^* = 0$. From Step 1, see Lemma 6.2, $\mathbf{F}$ is again a singleton consisting of the deterministic stationary strategy $f^*(x) \equiv \infty$. Similar argument as above shows that f^* is the solution to original impulse control problem (3) and

$$\mathcal{V}(0, f^*, \bar{C}_0) = \bar{C}_0(0, \infty) = 0, \quad \mathcal{V}(0, f^*, \bar{C}_1) = \bar{C}_1(0, \infty) = \frac{h}{\alpha^2}.$$

We now summarize the last conclusions in the following statement.

Theorem 6.3. Consider the impulse control problem in Example 6.1.

(a) If $d < \frac{h}{\alpha^2}$, then

$$g^* = \frac{K\alpha^2}{h(\alpha x^* - 1 + e^{-\alpha x^*})}$$

maximizes the dual functional $h(g)$, where $x^* > 0$ is the unique positive solution to equation

$$(1 - e^{-\alpha x})(h - d\alpha^2) = \alpha h x e^{-\alpha x}.$$

The optimal solution to problem (3) for the impulse control problem in Example 6.1 is given by the deterministic stationary strategy

$$f^*(x) = \max\{x^* - x, 0\} :$$

one has to apply the impulse if and only if the current state $x \geq x^*$, with

$$\mathcal{V}(0, f^*, \bar{C}_0) = \frac{K e^{-\alpha x^*}}{1 - e^{-\alpha x^*}}; \quad \mathcal{V}(0, f^*, \bar{C}_1) = d.$$

- (b) If $d \geq \frac{h}{\alpha^2}$, then $g^* = 0$ maximizes the dual functional $h(g)$, and the optimal solution to problem (3) for the impulse control problem in Example 6.1 is given by the deterministic stationary strategy $f^*(x) \equiv \infty$: no impulses should be applied, with

$$\mathcal{V}(0, f^*, \bar{C}_0) = 0; \quad \mathcal{V}(0, f^*, \bar{C}_1) = \frac{h}{\alpha^2}.$$

7. Proofs of the statements

Proof of Proposition 4.3. $\mathbf{B}$ is compact (in the product topology) as the product of compact sets. Thus, (a) in Definition 4.1 is verified.

If $u : \mathbf{X}_\Delta \rightarrow \mathbb{R}$ is a bounded continuous function, then

$$\int_{\mathbf{X}_\Delta} u(y)Q(dy|x, b = (\theta, a)) = \begin{cases} e^{-\alpha\theta}u(l(\phi(x, \theta), a)) \\ + (1 - e^{-\alpha\theta})u(\Delta) & \text{if } x \neq \Delta, \theta \neq +\infty; \\ u(\Delta) & \text{otherwise} \end{cases}$$

is a continuous function in $(x, b) \in \mathbf{X}_\Delta \times \mathbf{B}$. Thus, (b) in Definition 4.1 is verified.

Let us verify (c) Definition 4.1, that is, we show for a fixed $j \in \{0, \dots, J\}$ that $\bar{C}_j(\cdot, \cdot)$ is lower semicontinuous. Its lower semicontinuity on $\{\Delta\} \times \mathbf{B}$ follows from $\bar{C}_j(\Delta, b) \equiv 0$ and that Δ is isolated. Now it remains to show that $\bar{C}_j(\cdot, \cdot)$ is lower semicontinuous on $\mathbf{X} \times \mathbf{B}$, as follows. Recall that a non-negative function on $\mathbf{X} \times \mathbf{B}$ is lower semicontinuous if and only if it is the (pointwise) limit of an increasing sequence of bounded non-negative continuous functions, see [6, Lemma 7.14]. Thus, using the monotone convergence theorem, we see that the function

$$\int_0^\theta e^{-\alpha t} C_j^g(\phi(x, t)) dt = \lim_{m \rightarrow \infty} \int_0^\theta e^{-\alpha t} C_j^{g,m}(\phi(x, t)) dt$$

is non-negative and lower semicontinuous in $(x, b = (\theta, a)) \in \mathbf{X} \times \mathbf{B}$. Here $C_j^{g,m}(\cdot) \uparrow C_j^g(\cdot)$ pointwise as $m \rightarrow \infty$ for non-negative, bounded and continuous functions $C_j^{g,m}(\cdot)$ on $\mathbf{X}$, whose existence follows from the lower semicontinuity and non-negativity of the function $C_j^g(\cdot)$ and the aforementioned fact. On the other hand, the function defined by $e^{-\alpha\theta} C_j^l(\phi(x, \theta), a)$ is lower semicontinuous on $\mathbf{X} \times [0, \infty) \times \mathbf{A}$, and

$$\liminf_{n \rightarrow \infty} e^{-\alpha\theta_n} C_j^l(\phi(x_n, \theta_n), a_n) \geq 0 = e^{-\alpha\infty} C_j^l(\phi(x, \theta), a)$$

if $(x_n, (\theta_n, a_n)) \rightarrow (x, (\infty, a)) \in \mathbf{X} \times \mathbf{B}$. □

Proof of Lemma 5.2. Consider a strategy π such that $\mu^\pi(\mathbf{X} \times \mathbf{B}) = \infty$. Recall that μ^π satisfies

$$\mu^\pi(dx \times \mathbf{B}) = \delta_{x_0}(dx) + \int_{\mathbf{X}} \int_{[0, \infty)} \int_{\mathbf{A}} e^{-\alpha\theta} \delta_{l(\phi(y, \theta), a)}(dx) \mu^\pi(dy \times d\theta \times da)$$

on $(\mathbf{X}, \mathcal{B}(\mathbf{X}))$. Since $\mu^\pi(\mathbf{X} \times \mathbf{B}) = \infty$, we have

$$\infty = 1 + \int_{\mathbf{X}} \int_{[0, \infty)} \int_{\mathbf{A}} e^{-\alpha\theta} \mu^\pi(dy \times d\theta \times da) = 1 + \int_{[0, \infty)} e^{-\alpha\theta} \mu^\pi(\mathbf{X} \times d\theta \times \mathbf{A}).$$

Now,

$$\begin{aligned} \mathcal{V}(x_0, \pi, \bar{C}_0) &= \int_{\mathbf{X}_\Delta \times [0, \infty] \times \mathbf{A}} \bar{C}_0(x, (\theta, a)) \mu^\pi(dx \times d\theta \times da) \\ &\geq \int_{[0, \infty)} e^{-\alpha\theta} \delta \mu^\pi(\mathbf{X} \times d\theta \times \mathbf{A}) = \infty, \end{aligned}$$

as required. □

Proof of Lemma 5.4. (a) Consider the deterministic stationary strategy identified by $f(x) \equiv (\infty, \hat{a})$ with an arbitrarily fixed $\hat{a} \in \mathbf{A}$. Under Condition 5.1(b),

$$\mathcal{V}(x, f, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) = \mathbb{E}_x^f \left[\sum_{i=1}^\infty \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] \leq (1 + \sum_{j=1}^J g_j) C \int_0^\infty e^{-\alpha t} dt$$

for all $x \in \mathbf{X}_\Delta$. Thus, the function $\mathcal{V}^*(\cdot, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$ is non-negative and bounded by $(1 + \sum_{j=1}^J g_j)C \int_0^\infty e^{-at} dt$. Since the induced MDP is semicontinuous by assumption, according to Corollary 9.17.2 of [6], we see that $\mathcal{V}^*(\cdot, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$ is a bounded, non-negative and lower semicontinuous function on $\mathbf{X}_\Delta$. It follows that $W_g^*(\cdot)$, defined as the restriction of $\mathcal{V}^*(\cdot, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$ on $\mathbf{V}$, is bounded, non-negative and lower semicontinuous on $\mathbf{V}$. It is clear that $\mathcal{V}^*(x, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) = 0$ for all $x \in \mathbf{V}^c$. It follows from this fact and Proposition 9.8 of [6] that

$$\begin{aligned} & \mathcal{V}^*(x, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) \\ &= \inf_{b \in \mathbf{B}} \left\{ \bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) + \int_{\mathbf{X}_\Delta} \mathcal{V}^*(y, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) Q(dy|x, b) \right\} \\ &= \inf_{b \in \mathbf{B}} \left\{ \bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) + \int_{\mathbf{V}} \mathcal{V}^*(y, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) Q(dy|x, b) \right\} \end{aligned} \quad (36)$$

for all $x \in \mathbf{X}_\Delta$. Hence, $W_g^*(\cdot)$, defined as the restriction of $\mathcal{V}^*(\cdot, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$ on $\mathbf{V}$, satisfies (26), and is feasible in problem (25)-(26).

(b) According to Proposition 9.12 of [6], a deterministic stationary strategy $\hat{f}$ is uniformly optimal if and only if it attains the infimum in equality (36) for all $x \in \mathbf{X}_\Delta$. For $x \in \mathbf{V}^c$, the infimum is evidently attained at $f^*(x)$. Consequently, (b) is proved once we show that $W(\cdot) = W_g^*(\cdot)$ for the given function $W(\cdot) \in \mathbb{B}(\mathbf{V})$ satisfying equalities (28) for a measurable mapping $\hat{f}(\cdot) : \mathbf{V} \rightarrow \mathbf{B}$. Note that if $W(\cdot) \in \mathbb{B}(\mathbf{V})$ is lower semicontinuous, then such a measurable mapping $\hat{f}(\cdot) : \mathbf{V} \rightarrow \mathbf{B}$ exists, see e.g., Theorem 3.3 of [12]. It is convenient to extend $W(\cdot)$ to $\mathbf{X}_\Delta$ by putting $W(x) = 0$ for all $x \in \mathbf{V}^c$ in the rest of this proof.

First, we show that $W(x) \leq \mathcal{V}(x, \pi, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \pi, \bar{C}_j)$ for all $x \in \mathbf{V}$ and strategies π . Let $x \in \mathbf{V}$ and a strategy π be fixed. We distinguish two cases.

Case 1. $\mathcal{V}(x, \pi, \bar{C}_0) = \infty$. In this case, the desired inequality holds automatically.

Case 2. $\mathcal{V}(x, \pi, \bar{C}_0) < \infty$. Let us show in this case that $T_\infty := \lim_{i \rightarrow \infty} T_i = \infty$ $\mathbf{P}_x^\pi$ -a.s.. Suppose for contradiction that $\mathbf{P}_x^\pi(T_\infty < \infty) > 0$. Since, $\{T_\infty < \infty\} = \bigcup_{n=1}^\infty \{T_\infty < n\}$, there is some $T \in \{1, 2, \dots\}$ such that $\mathbf{P}_x^\pi(T_\infty < T) > 0$. Since $\{T_\infty < T\} \subseteq \bigcap_{i=1}^\infty \{T_i < T\}$, it holds that $\lim_{i \rightarrow \infty} \mathbf{P}_x^\pi(T_i < T) > 0$. Now

$$\begin{aligned} \mathcal{V}(x, \pi, \bar{C}_0) &\geq \sum_{i=1}^\infty \mathbf{E}_x^\pi [e^{-\alpha T_i} \delta] \geq \sum_{i=1}^\infty \mathbf{E}_x^\pi [\mathbb{I}\{T_i < T\} e^{-\alpha T_i} \delta] \\ &\geq \delta \sum_{i=1}^\infty \mathbf{E}_x^\pi [\mathbb{I}\{T_i < T\} e^{-\alpha T}] = \delta e^{-\alpha T} \sum_{i=1}^\infty \mathbf{P}_x^\pi(T_i < T) = \infty, \end{aligned}$$

where the last equality holds because, for each $i = 1, 2, \dots$,

$$\mathbf{P}_x^\pi(T_i < T) \geq \lim_{i \rightarrow \infty} \mathbf{P}_x^\pi(T_i < T) > 0.$$

This yields a contradiction, proving that $T_\infty = \lim_{i \rightarrow \infty} T_i = \infty$ $\mathbf{P}_x^\pi$ -a.s.. Now, since $\mathbf{P}_x^\pi(X_I \in \mathbf{X}) = \mathbf{E}_x^\pi[e^{-\alpha T_I}]$,

$$\begin{aligned} \mathbf{E}_x^\pi[|W(X_I)|] &= \mathbf{E}_x^\pi[\mathbb{I}\{X_I \in \mathbf{V}\} |W(X_I)|] \leq \mathbf{E}_x^\pi[\mathbb{I}\{X_I \in \mathbf{V}\}] \sup_{y \in \mathbf{V}} |W(y)| \\ &\leq \mathbf{P}_x^\pi(X_I \in \mathbf{X}) \sup_{y \in \mathbf{V}} |W(y)| = \mathbf{E}_x^\pi[e^{-\alpha T_I}] \sup_{y \in \mathbf{V}} |W(y)| \rightarrow 0 \text{ as } I \rightarrow \infty, \end{aligned}$$

where the convergence to 0 holds because of the dominated convergence theorem and the fact that $T_\infty =$

$\lim_{i \rightarrow \infty} T_i = \infty$ $\mathbf{P}_x^\pi$ -a.s.. Furthermore, for each $I = 1, 2, \dots$,

$$\begin{aligned}
0 &\leq \mathbb{E}_x^\pi \left[\sum_{i=1}^I \left\{ \bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) - W(X_{i-1}) \right. \right. \\
&\quad \left. \left. + \int_{\mathbf{V}} W(y) Q(dy | X_{i-1}, B_i) \right\} \right] \\
&= \mathbb{E}_x^\pi \left[\sum_{i=1}^I \left\{ \bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right\} \right] - \mathbb{E}_x^\pi[W(X_0)] \\
&\quad + \mathbb{E}_x^\pi[W(X_1)] - \mathbb{E}_x^\pi[W(X_1)] + \mathbb{E}_x^\pi[W(X_2)] - \dots + \mathbb{E}_x^\pi[W(X_I)],
\end{aligned} \tag{37}$$

where for the last equality we recall that $W(x) = 0$ for $x \in \mathbf{V}^c$. Note that all the terms here are finite, because the functions $\bar{C}_j(\cdot, \cdot)$ and $W(\cdot)$ are all bounded. After passing to the limit as $I \rightarrow \infty$, we obtain the claimed inequality

$$W(x) \leq \mathcal{V}(x, \pi, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \pi, \bar{C}_j),$$

too.

Secondly, we show that $W(x) = \mathcal{V}(x, \hat{f}, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \hat{f}, \bar{C}_j)$ for each $x \in \mathbf{V}$, as follows.

Let $x \in \mathbf{V}$ be fixed again. For the strategy $\pi = \hat{f}$ in the statement of this lemma, we have equalities in (37):

$$0 = \mathbb{E}_x^{\hat{f}} \left[\sum_{i=1}^I \left\{ \bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right\} \right] - \mathbb{E}_x^{\hat{f}}[W(X_0)] + \mathbb{E}_x^{\hat{f}}[W(X_I)].$$

Since $W(x)$ and $\limsup_{I \rightarrow \infty} \mathbb{E}_x^{\hat{f}}[W(X_I)]$ are finite, upon passing to the limit as $I \rightarrow \infty$, we see that $0 = \mathcal{V}(x, \hat{f}, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \hat{f}, \bar{C}_j) - W(x) + \limsup_{I \rightarrow \infty} \mathbb{E}_x^{\hat{f}}[W(X_I)]$, i.e., $W(x) = \mathcal{V}(x, \hat{f}, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \hat{f}, \bar{C}_j) + \limsup_{I \rightarrow \infty} \mathbb{E}_x^{\hat{f}}[W(X_I)]$. Again, since $W(x)$ and $\limsup_{I \rightarrow \infty} \mathbb{E}_x^{\hat{f}}[W(X_I)]$ are finite, we see that

$$\mathcal{V}(x, \hat{f}, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \hat{f}, \bar{C}_j) < \infty.$$

In particular, $\mathcal{V}(x, \hat{f}, \bar{C}_0) < \infty$, and we are in Case 2 above with $\pi = \hat{f}$. Applying the result established in Case 2 for $\pi = \hat{f}$, we see that $\lim_{I \rightarrow \infty} \mathbb{E}_x^{\hat{f}}[W(X_I)] = 0$. Hence,

$$W(x) = \mathcal{V}(x, \hat{f}, \bar{C}_0) + \sum_{j=1}^J g_j \mathcal{V}(x, \hat{f}, \bar{C}_j).$$

Finally, combining the results established in the above two steps, we see that the statement of the lemma holds. $\square$

Proof of Theorem 5.5. (a) Let $\bar{g} \in \mathbb{R}_+^J$ be arbitrarily fixed. In view of Lemma 5.4, we only need to show that the bounded, non-negative and lower semicontinuous function $W_{\bar{g}}^*(\cdot)$ defined by (27) solves problem (25)-(26) with $val(\text{problem (25)-(26)}) = W_{\bar{g}}^*(x_0)$. This is done as follows, by making use of the equivalence (up to an additive constant) between problem (25)-(26) and problem (24).

Since the induced MDP is semicontinuous by assumption, according to Corollary 9.17.2 of [6], there exists a uniformly optimal deterministic stationary strategy $\hat{f}$ satisfying $\mathcal{V}^*(x, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) = \inf_{\pi \in \Pi_G} \mathcal{V}(x, \pi, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) = \mathcal{V}(x, \hat{f}, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j)$ for all $x \in \mathbf{X}_\Delta$. Without loss of generality, we put $\hat{f}(x) = f^*(x)$ for $x \in \mathbf{V}^c$. (Recall equalities (6).)

Consider the following problem, which is dual with problem (24):

$$\text{Minimize over } \mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B}): \sup_{W(\cdot) \in \mathbb{B}(\mathbf{V})} L_2(\mu, W, \bar{g}). \quad (38)$$

We know that

$$val(\text{problem (38)}) \geq val(\text{problem (24)}). \quad (39)$$

(See, e.g., (1.8) in [23].)

On the other hand, problem (38), up to the constant $\sum_{j=1}^J g_j d_j$, is equivalent to the following one:

$$\begin{aligned} &\text{Minimize over } \mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B}) \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu(dx \times db) \\ &\text{subject to } \mu(\Gamma \times \mathbf{B}) = \delta_{x_0}(\Gamma) + \int_{\mathbf{V} \times \mathbf{B}} Q(\Gamma|y, b) \mu(dy \times db), \\ &\Gamma \in \mathcal{B}(\mathbf{V}), \end{aligned}$$

where we recall that the above constraint equality is the same as (20). Indeed, if for some $\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})$ equality (20) is not satisfied, then, similarly to Remark 5.3, $\sup_{W(\cdot) \in \mathbb{B}(\mathbf{V})} L_2(\mu, W, \bar{g}) = +\infty$; otherwise,

$$L_2(\mu, W, \bar{g}) + \sum_{j=1}^J g_j d_j = \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu(dx \times db)$$

for all $W(\cdot) \in \mathbb{B}(\mathbf{V})$.

Now consider the occupation measure $\mu^{\hat{f}}$ of a uniformly optimal deterministic stationary strategy $\hat{f}$. Recall that $\hat{f}(x) = f^*(x)$ for all $x \in \mathbf{V}^c$. Since $\mathcal{V}(x_0, \hat{f}, \bar{C}_0) \leq \mathcal{V}(x_0, \hat{f}, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) \leq (1 + \sum_{j=1}^J g_j) C \int_0^\infty e^{-at} dt$, where the last inequality follows from the uniform optimality of $\hat{f}$ and the rightmost bound was observed in the proof of Lemma 5.4(a), we see from Lemma 5.2 that $\mu^{\hat{f}}(\mathbf{V} \times \mathbf{B}) \leq \mu^{\hat{f}}(\mathbf{X} \times \mathbf{B}) < \infty$. Consider $\mu_{\bar{g}} := \mu^{\hat{f}}|_{\mathbf{V} \times \mathbf{B}}$. Then $\mu_{\bar{g}} \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})$ and satisfies equality (20) because $\hat{f}(x) = f^*(x)$ for $x \in \mathbf{V}^c$. (Recall (6).) Moreover, it holds that

$$\int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu_{\bar{g}}(dx \times db) = W_{\bar{g}}^*(x_0).$$

Therefore,

$$\begin{aligned} val(\text{problem (38)}) &\leq \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu_{\bar{g}}(dx \times db) - \sum_{j=1}^J g_j d_j \\ &= W_{\bar{g}}^*(x_0) - \sum_{j=1}^J g_j d_j. \end{aligned} \quad (40)$$

Since problem (24) is, up to the additive constant $-\sum_{j=1}^J g_j d_j$, equivalent to (25)-(26), and $W_{\bar{g}}^*(\cdot)$ is a feasible solution in problem (25)-(26), as was observed in Lemma 5.4(a), we see that

$$val(\text{problem (24)}) \geq W_{\bar{g}}^*(x_0) - \sum_{j=1}^J g_j d_j.$$

Thus, by (40),

$$val(\text{problem (24)}) \geq val(\text{problem (38)}),$$

which, together with (39), leads to

$$val(\text{problem (24)}) = val(\text{problem (38)}) = W_{\bar{g}}^*(x_0) - \sum_{j=1}^J g_j d_j.$$

Hence, $W_{\bar{g}}^*(\cdot)$ is an optimal solution in problem (24). It follows that $W_{\bar{g}}^*(\cdot)$ solves problem (25)-(26), which is equivalent to problem (24) up to an additive constant.

(b) Consider the strategy $\hat{f}$ as in the proof of part (a), and adopt the notations introduced therein. Then

$$\begin{aligned} & W_{\bar{g}}^*(x_0) - \sum_{j=1}^J g_j d_j \\ &= \inf_{\pi \in \Pi_G} \mathbb{E}_{x_0}^{\pi} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] - \sum_{j=1}^J g_j d_j \\ &= \mathcal{V}(x_0, \hat{f}, \bar{C}_0 + \sum_{j=1}^J g_j \bar{C}_j) - \sum_{j=1}^J g_j d_j \\ &= \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu^{\hat{f}}(dx \times db) - \sum_{j=1}^J g_j d_j. \end{aligned}$$

We have seen in the proof of part (a) of this theorem that $\mu^{\hat{f}}(\mathbf{X} \times \mathbf{B}) < \infty$. The non-negative functions $\bar{C}_j(\cdot, \cdot)$, $j \in \{0, \dots, J\}$, are bounded from above by $(\frac{1}{\alpha} + 1)C < \infty$. It follows that

$$\int_{\mathbf{V} \times \mathbf{B}} \bar{C}_j(x, b) \mu^{\hat{f}}(dx \times db) \leq \left(\frac{1}{\alpha} + 1\right) C \mu^{\hat{f}}(\mathbf{V} \times \mathbf{B}) < \infty$$

for each $j \in \{0, \dots, J\}$, and $\mu^{\hat{f}}|_{\mathbf{V} \times \mathbf{B}} \in \mathcal{D}$. One may write

$$\int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu^{\hat{f}}(dx \times db) - \sum_{j=1}^J g_j d_j = L_1(\mu^{\hat{f}}|_{\mathbf{V} \times \mathbf{B}}, \bar{g}).$$

Hence,

$$\begin{aligned} & \inf_{\pi \in \Pi_G} \mathbb{E}_{x_0}^{\pi} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] - \sum_{j=1}^J g_j d_j \\ & \geq \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}) = h(\bar{g}). \end{aligned}$$

Next, we prove that the opposite direction of the above inequality holds, too. For each $\mu \in \mathcal{D}$, consider its induced strategy π^{μ} . By Lemma 4.10, we see that

$$\begin{aligned} & \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j \bar{C}_j(x, b) \right) \mu(dx \times db) \\ & \geq \mathbb{E}_{x_0}^{\pi^{\mu}} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] \end{aligned}$$

and thus, $L_1(\mu, \bar{g}) \geq \mathbb{E}_{x_0}^{\pi^{\mu}} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] - \sum_{j=1}^J g_j d_j$. Since $\mu \in \mathcal{D}$ was arbitrarily fixed, it follows that

$$\begin{aligned} h(\bar{g}) &= \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}) \\ &\geq \inf_{\pi \in \Pi_G} \mathbb{E}_{x_0}^{\pi} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j \bar{C}_j(X_{i-1}, B_i) \right) \right] - \sum_{j=1}^J g_j d_j. \end{aligned}$$

The first assertion in part (b) is thus proved, whereas the second assertion follows from the first assertion, part (a) of this theorem and Theorem 4.13(a).

(c) This assertion follows from parts (a) and (b).

(d) The assertion in this part follows from the following chain of equalities. According to part (b),

$$\sup_{(W(\cdot), \bar{g}) \in \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J} \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} L_2(\mu, W, \bar{g}) = \text{val}(\text{problem (23)}) = h(\bar{g}^*).$$

By Theorem 4.13(a), $h(\bar{g}^*) = \text{val}(\text{problem (18)}) = \text{val}(\text{primal convex program (14)})$. As was noted at the beginning of Subsection 4.3,

$$\text{val}(\text{primal convex program (14)}) = \text{val}(\text{relaxed problem (10)-(12)}).$$

Finally, under the imposed conditions,

$$\begin{aligned} & \text{val}(\text{relaxed problem (10)-(12)}) = \text{val}(\text{second convex program (19)-(21)}) \\ &= \text{val}(\text{problem (22)}) = \inf_{\mu \in \mathcal{M}^{<\infty}(\mathbf{V} \times \mathbf{B})} \sup_{(W(\cdot), \bar{g}) \in \mathbb{B}(\mathbf{V}) \times \mathbb{R}_+^J} L_2(\mu, W, \bar{g}), \end{aligned}$$

where the first equality was observed in the paragraph below Lemma 5.2, and the second equality holds by Remark 5.3. $\square$

Next, we proceed with the proof of Theorem 5.6, which requires some preliminaries.

Proposition 7.1. *Suppose the induced MDP is semicontinuous, there exists a feasible strategy with a finite value, and Condition 5.1(a) is satisfied. Then there exists an optimal stationary strategy π^* in the original impulse control problem (3) in the form of a mixture of $J + 1$ deterministic stationary strategies $f^1, f^2, \dots, f^{J+1}$ with coefficients $\{\gamma_l\}_{l=1}^{J+1} \subseteq (0, 1]$ such that $f^l(x) = f^*(x)$ and $\pi^*(db|x) = \delta_{f^*(x)}(db)$ for all $x \in \mathbf{V}^c$. Recall that f^* is the fixed deterministic stationary strategy providing the infimum in (5) and satisfying (6). In particular, $\pi^* \in \Pi$.*

Proof. In view of Lemma 5.2, the conditions in Theorem 2 of [22] are satisfied. By Theorem 2 of [22], there exists an optimal stationary strategy π^* in the original impulse control problem (3) in the form of a mixture of $J + 1$ deterministic stationary strategies $f^1, f^2, \dots, f^{J+1}$ with coefficients $\{\gamma_l\}_{l=1}^{J+1} \subseteq (0, 1]$. In view of Condition 5.1(a) and the discussion below it, one can assume without loss of generality that $f^l(x) = f^*(x)$ and $\pi^*(db|x) = \delta_{f^*(x)}(db)$ for all $x \in \mathbf{V}^c$. The last assertion in the statement follows since, by assumption, there is a feasible strategy with a finite value. $\square$

Note that the strategies $\{f^l\}_{l=1}^{J+1}$ in Proposition 7.1 may not be in $\mathbf{F}$. Consequently, the key component in the proof of Theorem 5.6 is to find the suitable versions in $\mathbf{F}$ of $\{f^l\}_{l=1}^{J+1}$ in Proposition 7.1. This is based on the following result for generic MDPs, which may be of independent interest, where $\mathbf{V}$ is generic, not necessarily generated by the primitives of the impulse control model.

Lemma 7.2. *Consider a generic MDP $\{\mathbf{Y}, \mathbf{B}, Q, C\}$, where $\mathbf{Y} = \mathbf{V} \cup \{\Delta\}$ with $\Delta \notin \mathbf{V}$ being the isolated cemetery. Suppose the single cost function $C(\cdot, \cdot)$ is $[0, \infty]$ -valued, $C(\Delta, b) \equiv 0$, and this MDP is semicontinuous. Let the Bellman function*

$$\mathcal{V}^*(y, C) := \inf_{\pi \in \Pi_G} \mathbb{E}_y^\pi \left[\sum_{i=1}^{\infty} C(Y_{i-1}, B_i) \right]$$

be defined for each $y \in \mathbf{Y}$, where Π_G stands for the set of strategies in this generic MDP. Clearly, $\mathcal{V}^(\Delta, C) = 0$. Let the initial state $y_0 \in \mathbf{V}$ be fixed, and f be an (optimal) deterministic stationary strategy such that*

$$\mathbb{E}_{y_0}^f \left[\sum_{i=1}^{\infty} C(Y_{i-1}, f(Y_{i-1})) \right] = \int_{\mathbf{V}} C(y, f(y)) \mu^f(dy \times \mathbf{B}) = \mathcal{V}^*(y_0, C),$$

where μ^f is the occupation measure of f with the initial state y_0 . When $y = \Delta$, we fix $f(\Delta) = f^(\Delta)$ arbitrarily. Assume that $\mathcal{V}^*(y_0, C) < \infty$. Then $\mu^f = \mu^{\tilde{f}}$ on $\mathbf{V} \times \mathbf{B}$ for some deterministic stationary strategy $\tilde{f}$, which is uniformly optimal, i.e.,*

$$\mathbb{E}_y^{\tilde{f}} \left[\sum_{i=1}^{\infty} C(Y_{i-1}, \tilde{f}(Y_{i-1})) \right] = \mathcal{V}^*(y, C), \quad y \in \mathbf{V}. \quad (41)$$

Similarly to the above, $\tilde{f}(\Delta) := f^(\Delta)$ without loss of generality.*

Proof. By Corollary 9.17.2 of [6], the Bellman function $\mathcal{V}^*(\cdot, C)$ is lower semicontinuous and there exists a uniformly optimal deterministic stationary strategy $\check{f}$ in the sense of (41). Without loss of generality we accept that $\check{f}(\Delta) = f^*(\Delta)$. Let us show that

$$\mathcal{V}(y, f, C) := \mathbb{E}_y^f \left[\sum_{i=1}^{\infty} C(Y_{i-1}, f(Y_{i-1})) \right] = \mathcal{V}^*(y, C) \quad (42)$$

for $\mu^f(dy \times \mathbf{B})$ -almost all $y \in \mathbf{V}$.

Suppose for contradiction that there is $\Gamma_Y \in \mathcal{B}(\mathbf{V})$ such that equality (42) is violated for all $y \in \Gamma_Y$ and $\mu^f(\Gamma_Y \times \mathbf{B}) > 0$. Let I be the first moment when $\mathbb{E}_{y_0}^f[\mathbb{I}\{Y_I \in \Gamma_X\}] > 0$. Then

$$\begin{aligned} \infty > \mathbb{E}_{y_0}^f \left[\sum_{i=1}^{\infty} C(Y_{i-1}, f(Y_{i-1})) \right] &= \mathbb{E}_{y_0}^f \left[\sum_{i=1}^I C(Y_{i-1}, f(Y_{i-1})) \right] \\ &\quad + \mathbb{E}_{y_0}^f \left[\mathbb{E}_{Y_I}^f \left[\sum_{i=1}^{\infty} C(Y_{i-1}, f(Y_{i-1})) \right] \right], \end{aligned}$$

where the first inequality holds because $\mathcal{V}^*(y_0, C) < \infty$. The last term in the above equality is strictly larger than $\mathbb{E}_{y_0}^f[\mathcal{V}^*(Y_I, C)]$. At the same time, for the deterministic Markov strategy $\hat{f} = \{\hat{f}_i\}_{i=1}^{\infty}$ defined by

$$\hat{f}_i(y) := \begin{cases} f(y), & \text{if } i \leq I, \\ \check{f}(y), & \text{if } i > I \end{cases}$$

we have

$$\begin{aligned} \mathbb{E}_{y_0}^{\hat{f}} \left[\sum_{i=1}^{\infty} C(Y_{i-1}, \hat{f}(Y_{i-1})) \right] &= \mathbb{E}_{y_0}^f \left[\sum_{i=1}^I C(Y_{i-1}, f(Y_{i-1})) \right] + \mathbb{E}_{y_0}^f[\mathcal{V}^*(Y_I, C)] \\ &< \mathbb{E}_{y_0}^f \left[\sum_{i=1}^{\infty} C(Y_{i-1}, f(Y_{i-1})) \right], \end{aligned}$$

and thus the strategy f is not optimal at x_0 , yielding a contradiction. This proves that equality (42) is valid for $\mu^f(dy \times \mathbf{B})$ -almost all $y \in \mathbf{V}$.

Now we modify the strategy f on the $\mu^f(dy \times \mathbf{B})$ -null subset of $\mathbf{V}$:

$$\tilde{f}(y) := \begin{cases} f(y), & \text{if (42) holds for } y, \\ \check{f}(y), & \text{otherwise.} \end{cases}$$

The left-hand side in (42) is a measurable function of y : see Propositions C.1.1(b), B.1.31 and B.1.34 of [21]. Thus, $\tilde{f}(\cdot)$ is a measurable mapping because the set $\{y \in \mathbf{V} : (42) \text{ holds for } y\}$ is measurable.

Next, let us prove the following equality

$$\mathcal{V}^*(y, C) = C(y, \tilde{f}(y)) + \int_{\mathbf{V}} \mathcal{V}^*(z, C) Q(dz|y, \tilde{f}(y)), \quad y \in \mathbf{V}. \quad (43)$$

Let $y \in \mathbf{V}$ be fixed. If equality (42) is not valid at this $y \in \mathbf{V}$, then $\tilde{f}(y) = \check{f}(y)$, and equality (43) holds by Proposition 9.12 of [6]. Suppose equality (42) is valid for this $y \in \mathbf{V}$, so that $\tilde{f}(y) = f(y)$. Then

$$\mathcal{V}^*(y, C) = \mathcal{V}(y, f, C) = C(y, f(y)) + \int_{\mathbf{V}} \mathcal{V}(z, f, C) Q(dz|y, f(y)),$$

where

$$\mathcal{V}(z, f, C) = \mathbb{E}_z^f \left[\sum_{i=1}^{\infty} C(Y_{i-1}, f(Y_{i-1})) \right] \geq \mathcal{V}^*(z, C).$$

Thus,

$$\mathcal{V}^*(y, C) \geq C(y, f(y)) + \int_{\mathbf{V}} \mathcal{V}^*(z, C) Q(dz|y, f(y)).$$

If this inequality was strict, then we would have

$$\mathcal{V}^*(y, C) > \inf_{b \in \mathbf{B}} \left\{ C(y, b) + \int_{\mathbf{V}} \mathcal{V}^*(z, C) Q(dz|y, b) \right\},$$

which is impossible, because the Bellman function $\mathcal{V}^*(\cdot, C)$ satisfies the Bellman equation

$$W(y) = \inf_{b \in \mathbf{B}} \left\{ C(y, b) + \int_{\mathbf{V}} W(z) Q(dz|y, b) \right\}, \quad y \in \mathbf{V}.$$

Consequently, we conclude that equality (43) is valid for all $y \in \mathbf{V}$, and the deterministic stationary strategy $\tilde{f}$ is uniformly optimal, again, by Proposition 9.12 of [6].

To show that the occupation measure $\mu^{\tilde{f}}$ coincides with μ^f on $\mathbf{V} \times \mathbf{B}$, firstly note that

$$\mu^f(dy \times db) = \mu^f(dy \times \mathbf{B}) \delta_{f(y)}(db) = \mu^f(dy \times \mathbf{B}) \delta_{\tilde{f}(y)}(db) \quad (44)$$

because $f(\cdot) = \tilde{f}(\cdot)$ for $\mu^f(dy \times \mathbf{B})$ -almost all $y \in \mathbf{V}$. Next, the marginal $\mu^{\tilde{f}}(dy \times \mathbf{B})$ on $\mathbf{V}$ is the limit of the increasing (set-wise) sequence of measures

$$\mu_n(\Gamma) := \sum_{i=1}^n \mathbf{P}_{y_0}^{\tilde{f}}(Y_{i-1} \in \Gamma), \quad \Gamma \in \mathcal{B}(\mathbf{V}), \quad n = 0, 1, \dots,$$

which satisfy the following relations:

$$\begin{aligned} \mu_0(\Gamma) &= 0; \\ \mu_n(\Gamma) &= \delta_{y_0}(\Gamma) + \int_{\mathbf{V}} Q(\Gamma|y, \tilde{f}(y)) \mu_{n-1}(dy), \quad \Gamma \in \mathcal{B}(\mathbf{V}), \quad n = 1, 2, \dots \end{aligned} \quad (45)$$

The proof is by induction: (45) is valid at $n = 1$ and, if it is valid for some $n \geq 1$, then

$$\begin{aligned} \mu_{n+1}(\Gamma) &= \delta_{y_0}(\Gamma) + \sum_{i=1}^n \int_{\mathbf{V}} Q(\Gamma|y, \tilde{f}(y)) \mathbf{P}_{y_0}^{\tilde{f}}(Y_{i-1} \in dy) \\ &= \delta_{y_0}(\Gamma) + \int_{\mathbf{V}} Q(\Gamma|y, \tilde{f}(y)) \mu_n(dy), \quad \Gamma \in \mathcal{B}(\mathbf{V}). \end{aligned}$$

Since, according to (44),

$$\mu^f(\Gamma \times \mathbf{B}) = \delta_{y_0}(\Gamma) + \int_{\mathbf{V}} Q(\Gamma|y, \tilde{f}(y)) \mu^f(dy \times \mathbf{B}), \quad \Gamma \in \mathcal{B}(\mathbf{V}),$$

$\mu_n(\Gamma) \leq \mu^f(\Gamma \times \mathbf{B})$ for all $\Gamma \in \mathcal{B}(\mathbf{V})$ and $n = 0, 1, \dots$: the obvious proof is by induction w.r.t. n . Thus,

$$\mu^{\tilde{f}}(\Gamma \times \mathbf{B}) \leq \mu^f(\Gamma \times \mathbf{B}), \quad \Gamma \in \mathcal{B}(\mathbf{V}).$$

We conclude that, in turn, f is a modification of $\tilde{f}$ on a $\mu^{\tilde{f}}(dx \times \mathbf{B})$ -null subset of $\mathbf{V}$, and

$$\mu^{\tilde{f}}(dy \times db) = \mu^{\tilde{f}}(dy \times \mathbf{B}) \delta_{\tilde{f}(y)}(db) = \mu^{\tilde{f}}(dy \times \mathbf{B}) \delta_{f(y)}(db).$$

Repeating the above reasoning for f and $\tilde{f}$ being swapped, we see that $\mu^f(\Gamma \times \mathbf{B}) \leq \mu^{\tilde{f}}(\Gamma \times \mathbf{B})$, $\Gamma \in \mathcal{B}(\mathbf{V})$.

Therefore, $\mu^f(dy \times \mathbf{B}) = \mu^{\tilde{f}}(dy \times \mathbf{B})$ and hence

$$\mu^f(dy \times db) = \mu^{\tilde{f}}(dy \times \mathbf{B}) \delta_{\tilde{f}(y)}(db) = \mu^{\tilde{f}}(dy \times db)$$

on $\mathbf{V} \times \mathbf{B}$ by (44). □

Proof of Theorem 5.6. According to Proposition 7.1, there is an optimal stationary strategy $\pi^* \in \Pi$ in problem (3) in the form of a mixture of $J + 1$ deterministic stationary strategies $f^1, f^2, \dots, f^{J+1}$ with coefficients $\{\gamma_l\}_{l=1}^{J+1} \subseteq (0, 1]$ such that $f^l(x) = f^*(x)$ and $\pi^*(db|x) := \delta_{f^*(x)}(db)$ for all $x \in \mathbf{V}^c$. Now all the measures $\mu^{f^l}(dx \times \mathbf{B})$ are σ -finite on $\mathbf{V}$ by Lemma 4.8: the strategy π^* is feasible with a finite value, hence $\int_{\mathbf{X} \times \mathbf{B}} \bar{C}_j(x, b) \mu^{f^l}(dx \times db) < \infty$ for all $l = 1, 2, \dots, J + 1$ because $\gamma^l > 0$. The measures $\mu^{f^l}|_{\mathbf{V} \times \mathbf{B}}$ obviously belong to the set $\mathcal{D}$.

Problem (3) is equivalent to the relaxed problem (10)-(12) by Theorem 4.11, and the restriction of the occupation measure $\mu^* := \mu^{\pi^*}|_{\mathbf{V} \times \mathbf{B}}$ is feasible and optimal in the relaxed problem (10)-(12) and also in the primal convex program (14). Therefore, $L_1(\mu^*, \bar{g}^*) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*)$ according to Theorem 4.13(b-ii). Recall that $\bar{g}^*$ is the maximizer of the dual functional $h(\bar{g})$. Since, for each $l = 1, 2, \dots, J + 1$, the restriction of the occupation measure $\mu^{f^l}|_{\mathbf{V} \times \mathbf{B}}$ belongs to $\mathcal{D}$, $\gamma_l > 0$, $\mu^* = \sum_{l=1}^{J+1} \gamma_l \mu^{f^l}|_{\mathbf{V} \times \mathbf{B}}$, and the Lagrangian $L_1(\mu, \bar{g}^*)$ is linear in μ with non-negative functions $\bar{C}_j(\cdot, \cdot)$, $j = 0, 1, \dots, J$, we see that $L_1(\mu^{f^l}, \bar{g}^*) = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*)$ for all $l = 1, 2, \dots, J + 1$.

According to Theorem 5.5(b) and Theorem 4.13(a), for each $l = 1, 2, \dots, J + 1$,

$$\begin{aligned} h(\bar{g}^*) &= W_{\bar{g}^*}^*(x_0) - \sum_{j=1}^J g_j^* d_j = \inf_{\mu \in \mathcal{D}} L_1(\mu, \bar{g}^*) = L_1(\mu^{f^l}, \bar{g}^*) \\ &= \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j^* \bar{C}_j(x, b) \right) \mu^{f^l}(dx \times db) - \sum_{j=1}^J g_j^* d_j, \end{aligned}$$

so that

$$\begin{aligned} W_{\bar{g}^*}^*(x_0) &= \int_{\mathbf{V} \times \mathbf{B}} \left(\bar{C}_0(x, b) + \sum_{j=1}^J g_j^* \bar{C}_j(x, b) \right) \mu^{f^l}(dx \times db) \\ &= \int_{\mathbf{V}} \left(\bar{C}_0(x, f^l(x)) + \sum_{j=1}^J g_j^* \bar{C}_j(x, f^l(x)) \right) \mu^{f^l}(dx \times \mathbf{B}) : \end{aligned}$$

the deterministic stationary strategy f^l is optimal in the sense that $\mathcal{V}(x_0, f^l, \bar{C}_0 + \sum_{j=1}^J g_j^* \bar{C}_j) = W_{\bar{g}^*}^*(x_0)$. Recall that the Bellman function $W_{\bar{g}^*}^*(\cdot)$ is bounded by Theorem 5.5(a). According to Lemma 7.2, for each $l \in \{1, \dots, J + 1\}$, $\mu^{f^l}|_{\mathbf{X} \times \mathbf{B}} = \mu^{\tilde{f}^l}|_{\mathbf{X} \times \mathbf{B}}$ and hence $\mu^{f^l}|_{\mathbf{V} \times \mathbf{B}} = \mu^{\tilde{f}^l}|_{\mathbf{V} \times \mathbf{B}}$, where the deterministic stationary strategy $\tilde{f}^l$ is uniformly optimal in the sense that

$$\begin{aligned} &\mathcal{V}(x, f^l, \bar{C}_0 + \sum_{j=1}^J g_j^* \bar{C}_j) \\ &= \mathbb{E}_x^{\tilde{f}^l} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j^* \bar{C}_j(X_{i-1}, B_i) \right) \right] \\ &= \inf_{\pi \in \Pi_G} \mathbb{E}_x^{\pi} \left[\sum_{i=1}^{\infty} \left(\bar{C}_0(X_{i-1}, B_i) + \sum_{j=1}^J g_j^* \bar{C}_j(X_{i-1}, B_i) \right) \right] = W_{\bar{g}^*}^*(x), \quad x \in \mathbf{V}, \end{aligned}$$

and can be taken such that $\tilde{f}^l(x) = f^*(x)$ for all $x \in \mathbf{V}^c$. Now $\tilde{f}^l \in \mathbf{F}$ by Proposition 9.12 of [6], and π^* is a mixture of $\tilde{f}^1, \tilde{f}^2, \dots, \tilde{f}^{J+1}$. $\square$

8. Conclusion

In this paper, we studied optimal impulse control problem of a deterministic dynamical system governed by a (semi)flow. The main performance measure to be minimized is given by the total discounted running and impulse costs over an infinite horizon, subject to inequality constraints on some other similar performance measures. By exploiting the connection between this problem and MDPs, we formulate two pairs of primal and dual programs, and showed the absence of duality gaps. We also provided general procedure to find

optimal strategies in the form of mixtures over no more than $J + 1$ deterministic stationary strategies, where J is the number of constraints. Following this procedure, we solved explicitly an example of queueing system.

The impulse control model considered in this paper is positive in the sense that $C_j^g(\cdot)$ and $C_j^l(\cdot, \cdot)$ are $[0, \infty]$ -valued. In case we only consider strategies under which $t_\infty = \lim_{n \rightarrow \infty} t_n = \infty$, where $\{t_n\}_{n=1}^\infty$ are the successive time instants when impulses are applied, then the result applies to the case when $C_j^g(\cdot)$ are bounded below by a common constant say $-\kappa$, where $\kappa > 0$: one can consider new running cost rates given by $C_j^g(\cdot) + \kappa \geq 0$ and adjust the constraint constants accordingly from d_j to $d_j + \frac{\kappa}{\alpha}$.

For the future studies, one may consider to apply the obtained theoretical results in this paper to other real-life applications, partially mentioned in Section 1. It is also reasonable to develop the similar theory for piecewise deterministic processes, to diffusions and other underlying dynamical systems. The optimal impulse control problem with multi-agents (more than one decision maker) is also interesting.

Acknowledgments: Part of this work was carried out during the visit of the second author to Sun Yat-sen University (China), partially supported by the National Natural Science Foundation of China (No.72471252).

References

- [1] D. Altıntan, V. Purutcuoğlu, Ö. Uğur: *Impulsive expressions in stochastic simulation algorithms*, Intern. J. of Computational Methods 15(1) (2018) 1750075.
- [2] A. Arutyunov, D. Karamzin, F.L. Pereira: *Optimal Impulsive Control*, Springer, Cham (2019).
- [3] H. Bakan, F. Yilmaz, G.W. Weber: *An efficient algorithm for stochastic optimal control problems by means of least-squares Monte-Carlo method*, Optimization 71 (2022) 3133–3146.
- [4] G. Barles: *Deterministic impulse control problems*, SIAM J. Control Optim. 23 (1985) 419–432.
- [5] D. Bauso: *Impulsively-controlled systems and reverse dwell time: a linear programming approach*, Nonlinear Analysis: Hybrid Systems 16 (2015) 40–51.
- [6] D. Bertsekas, S. Shreve: *Stochastic Optimal Control*, Academic Press, New York (1978).
- [7] A. Blaqui re: *Impulsive optimal control with finite or infinite time horizon*, J. Optim. Theory Appl. 46 (1985) 431–439.
- [8] M. Chahim, R. Hartl, P. Kort: *A tutorial on the deterministic impulse control maximum principle: necessary and sufficient optimality conditions*, Euro. J. Oper. Res. 219 (2012) 18–26.
- [9] M. Claeys, D. Arzelier, D. Henrion, J. Lasserre: *Measures and LMIs for impulsive nonlinear optimal control*, IEEE Trans. Autom. Control. 59 (2014) 1374–1379.
- [10] F. Dufour, M. Horiguchi, A. Piunovskiy: *The expected total cost criterion for Markov decision processes under constraints: a convex analytic approach*, Adv. Appl. Probab. 44 (2012) 774–793.
- [11] E.B. Dynkin, A.A. Yushkevich: *Controlled Markov Processes*, Springer, New York (1979).
- [12] E.A. Feinberg, P.O. Kasyanov, N.V. Zadoianchuk: *Berge’s theorem for noncompact image sets*, J. Math. Anal. Appl. 397 (2013) 255–259.
- [13] D. Henrion, M. Kru  k, T. Weisser: *Optimal control problems with oscillations, concentrations and discontinuities*, Automatica 103 (2019) 159–165.
- [14] O. Hern  ndez-Lerma, J. Lasserre: *Further Topics on Discrete-Time Markov Control Processes*, Springer-Verlag, New York (1999).
- [15] E.H. Hou, K.H. Wong: *Optimal impulsive control problem with application to human immunodeficiency virus treatment*, J. Optim. Theory Appl. 151 (2011) 385–401.
- [16] R. Leander, S. Lenhart, V. Protopopescu: *Optimal control of continuous systems with impulse controls*, Optim. Control Appl. Meth. 36 (2015) 535–549.
- [17] D.G. Luenberger: *Optimization by Vector Space Methods*, John Wiley & Sons, New York (1969).
- [18] B. Miller, E. Rubinovich: *Optimal impulse control problem with constrained number of impulses*, Math. Comput. Simulat. 34 (1992) 23–49.

- [19] B. Miller, E. Rubinovich: *Impulsive Control in Continuous and Discrete-Continuous Systems*, Springer, New York (2003).
- [20] A. Piunovskiy, A. Plakhov, D. Torres, Y. Zhang: *Optimal impulse control of dynamical systems*, SIAM J. Control Optim. 57 (2019) 2720–2752.
- [21] A. Piunovskiy, Y. Zhang: *Continuous-Time Markov Decision Processes*, Springer, Cham (2020).
- [22] A. Piunovskiy, Y. Zhang: *Extreme occupation measures in Markov decision processes with an absorbing state*, SIAM J. Control Optim. 62 (2024) 65–90.
- [23] R.T. Rockafellar: *Conjugate Duality and Optimization*. SIAM, Philadelphia (1974).
- [24] T. Yang: *Impulsive Control Theory*. Springer, Berlin-Heidelberg (2001).