

# Aggregated occupation measures and linear programming approach to constrained impulse control problems

Piunovskiy, Alexey; Zhang, Yi

DOI:

[10.1016/j.jmaa.2021.125070](https://doi.org/10.1016/j.jmaa.2021.125070)

License:

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

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Piunovskiy, A & Zhang, Y 2021, 'Aggregated occupation measures and linear programming approach to constrained impulse control problems', *Journal of Mathematical Analysis and Applications*, vol. 499, no. 2, 125070. <https://doi.org/10.1016/j.jmaa.2021.125070>

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

## General rights

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

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

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

When citing, please reference the published version.

## Take down policy

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

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

# Aggregated occupation measures and linear programming approach to constrained impulse control problems\*

Alexey Piunovskiy<sup>†</sup>

Department of Mathematical Sciences, University of Liverpool, L69 7ZL, UK.  
piunov@liv.ac.uk

Yi Zhang

Department of Mathematical Sciences, University of Liverpool, L69 7ZL, UK.  
yi.zhang@liv.ac.uk

## Abstract

For a constrained optimal impulse control problem of an abstract dynamical system, we introduce the occupation measures along with the aggregated occupation measures and present two associated linear programs. We prove that the two linear programs are equivalent under appropriate conditions, and each linear program gives rise to an optimal strategy in the original impulse control problem. In particular, we show the absence of the relaxation gap. By means of an example, we also present a detailed comparison of the occupation measures and linear programs introduced here with the related notions in the literature.

**Keywords:** Dynamical System, Optimal Control, Impulse Control, Total Cost, Constraints, Linear Programming.

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

## 1 Introduction

Impulse control of dynamical systems attracts attention of many researchers. The underlying system can be described in terms of ordinary differential equations, see [3, 4, 6, 10, 23, 25, 26, 27], or by a fixed flow in an Euclidean space or in an abstract Borel space, see [15, 29]. An impulse or an intervention means an instantaneous change of the state of the system. In most of the aforementioned works, the target was to optimize a single objective functional, typically having the shape of the integral with respect to time of the running cost and the impulse costs. The popular methods of attack to such problems include dynamic programming, see [3, 4, 15, 29], and the Pontryagin maximum principle, see [6, 25, 27]. When the total number of impulses is fixed over a finite horizon, the impulse control problem can be treated as a parameter optimization problem, see [23, 26].

In this paper, we consider an impulse control problem of a dynamical system over an infinite horizon with multiple objectives. For optimal control problems with functional constraints, dynamic programming is not always convenient, and the so called convex analytic approach, also known as the linear programming approach, proved to be effective, especially in handling the solvability issues. For the linear programming approach to Markov decision processes, see [14, 20, 21, 28], and for this approach to deterministic optimal control problems without impulses, see [17, 19, 24]. In a nutshell,

---

\*Declarations of interest: none.

<sup>†</sup>Corresponding author.

the linear programming approach, if justified, reduces the original optimal control problem to a linear program in the space of suitably defined occupation measures with the same (optimal) value, and one can retrieve an optimal control strategy for the original problem from the optimal solution to the induced linear program.

For a deterministic impulse control problem over a finite horizon, a linear program formulation was presented in [10], from which, as the primitive goal of that paper, the authors established a numerical method for solving approximately the original problem. For this reason, [10] dealt with an unconstrained problem for a specific model with polynomial initial data, and did not show that the formulated linear program was equivalent to the original impulse control problem. Another, slightly different linear programming approach appeared in [11, 18], where the equivalence between the linear program and the original problem was only briefly discussed. In the aforementioned works, the flow in an Euclidean space came from an ordinary differential equation, whereas in the present paper the flow is arbitrary enough and lives in a Borel space. In [30], we introduced a different definition of occupation measures as compared to those in [10, 11, 18], and consequently a different linear program formulation was presented. Making use of the results from Markov decision processes, it was shown to be equivalent to the original impulse control problem. This type of result is often referred to as the absence of relaxation gap.

In certain sense, the present paper can be regarded as a follow-up of [30]. Indeed, we will start with recapitulating briefly the linear programming approach developed in [30], which was in the space of occupation measures, see (10),(11) and (16). Then we introduce another notion of occupation measures, referred to as the aggregated occupation measures. Accordingly, a second linear program is formulated in the space of aggregated occupation measures, see (14), (15). The main advantage of the aggregated occupation measures over the occupation measures in [30] lies in the reduction of dimensionality: see Remark 4.2. On the other hand, this paper is an independent piece of work as much effort here is made to show the equivalence between two linear programs, whose feasible solutions are measures satisfying certain characteristic equations (both arising from the impulse control problem, though). In fact, our main contributions lie in this equivalence, see Corollary 4.1, and in that we reveal how to induce an optimal strategy for the impulse control problem from an optimal solution to either of the two linear programs. Besides, by means of an example (see Section 5), we present a detailed account on the relation between the aggregated occupation measures and the associated linear program and their counterparts in [10, 11, 18].

We now summarize (and partially reiterate) the novelty of the present paper as follows:

- the dynamical system is described by a flow in an arbitrary Borel space, rather than by an ordinary differential equation in an Euclidean space;
- the optimal solution must satisfy a number of functional constraints which were absent in the cited literature;
- under suitable conditions, we rigorously prove that the optimal values of the original impulse control problem and of the introduced linear programs coincide, i.e., there is no relaxation gap;
- we show how to retrieve the optimal control strategy from the solutions to the associated linear programs.

The rest of this article is organized as follows. The problem statement is described in Section 2. In Sections 3 and 4, we formulate the preliminary observations and the main results correspondingly. In Section 5, we present an example and compare our approach with works [10, 11, 18]. The proofs of the main theorems are given in Sections 6 and 7. Some auxiliary lemmas are presented and proved in the Appendix.

**Notation.** Throughout this paper, we use the following notations:  $\mathbb{N} := \{1, 2, \dots\}$ ,  $\bar{\mathbb{R}}_+^0 := [0, \infty]$ ,  $\mathbb{R}_+^0 := [0, \infty)$ ,  $\mathbb{R}_+ := (0, \infty)$ . The term “measure” will always refer to a countably additive  $\bar{\mathbb{R}}_+^0$ -valued set function, equal to zero on the empty set. Consider two  $\sigma$ -finite measures  $\eta_1$  and  $\eta_2$  on a common measurable space  $(\Omega, \mathcal{F})$  such that  $\eta_1 \geq \eta_2$  set-wise. Then there exists a measurable decomposition  $\{\Omega_n\}_{n=1}^\infty$  of  $\Omega$  such that  $\eta_1(\Omega_n) < \infty$  and  $\eta_2(\Omega_n) < \infty$ . The difference between these two measures is defined by  $(\eta_1 - \eta_2)(d\omega) := \sum_{n=1}^\infty (\eta_1(d\omega \cap \Omega_n) - \eta_2(d\omega \cap \Omega_n))$ .  $\mathcal{P}(E)$  is the space of all probability measures on a measurable space  $(E, \mathcal{B}(E))$ . On the time axis  $\mathbb{R}_+^0$  the expression “for almost all  $u$ ” is understood with respect to the Lebesgue measure. By default, the  $\sigma$ -algebra on  $\mathbb{R}_+^0$  is just the Borel one. If  $(E, \mathcal{B}(E))$  is a measurable space then, for  $Y \in \mathcal{B}(E)$ ,  $\mathcal{B}(Y) := \{X \cap Y, X \in \mathcal{B}(E)\}$  is the restriction of the  $\sigma$ -algebra  $\mathcal{B}(E)$ . Integrals on a measure space  $(E, \mathcal{B}(E), \mu)$  are denoted as  $\int_E h(e) d\mu(e)$  or as  $\int_E h(e) \mu(de)$ . If  $b = \infty$  then the Lebesgue integrals  $\int_{[a,b]} f(u) du$  are taken over the open interval  $(a, \infty)$ . Expressions like “positive, negative, increasing, decreasing” are understood in the non-strict sense, like “nonnegative” etc. For  $I \subset \mathbb{R}$ ,  $\tau \in \mathbb{R}$ ,  $\tau + I := \{\tau + x : x \in I\}$  is the shifted set.  $\mathbb{I}\{\cdot\}$  is the indicator function;  $\delta_y(dx)$  is the Dirac measure at the point  $y$ . For  $b, c \in [-\infty, +\infty]$ ,  $b^+ := \max\{b, 0\}$ ,  $b^- := -\min\{b, 0\}$ ,  $b \wedge c := \min\{b, c\}$ ,  $b \vee c := \max\{b, c\}$ .

## 2 Problem Statement

We will deal with a control model defined through the following elements.

- $\mathbf{X}$  is the state space, which is a topological Borel space.
- $\phi(\cdot, \cdot) : \mathbf{X} \times \mathbb{R}_+^0 \rightarrow \mathbf{X}$  is the measurable flow possessing the semigroup property  $\phi(x, t + s) = \phi(\phi(x, s), t)$  for all  $x \in \mathbf{X}$  and  $(t, s) \in (\mathbb{R}_+^0)^2$ ;  $\phi(x, 0) = x$  for all  $x \in \mathbf{X}$ . Between the impulses, the state changes according to the flow.
- $\mathbf{A}$  is the action space, again a topological Borel space with a compatible metric  $\rho_A$ .
- $l(\cdot, \cdot) : \mathbf{X} \times \mathbf{A} \rightarrow \mathbf{X}$  is the mapping describing the new state after the corresponding action/impulse is applied.
- For each  $j = 0, 1, \dots, J$ , where and below  $J$  is a fixed natural number,  $C_j^g(\cdot) : \mathbf{X} \rightarrow \mathbb{R}_+^0$  is a (gradual) cost rate.
- For each  $j = 0, 1, \dots, J$ ,  $C_j^I(\cdot, \cdot) : \mathbf{X} \times \mathbf{A} \rightarrow \mathbb{R}_+^0$  is a cost function associated with the actions/impulses applied in the corresponding states.

All the mappings  $\phi, l, \{C_j^g\}_{j=0}^J$  and  $\{C_j^I\}_{j=0}^J$  are assumed to be measurable. The initial state  $x_0 \in \mathbf{X}$  is fixed.

We assume that the states  $x \in \mathbf{X}$  have the form  $x = (\tilde{x}, t)$ , where  $t \in \mathbb{R}_+^0$  equals time elapsed since the most recent impulse, and  $\tilde{x} \in \tilde{\mathbf{X}}$ , an arbitrary Borel space with a compatible metric  $\tilde{\rho}$ . In this connection,

$$\phi(x, u) = \phi((\tilde{x}, t), u) := (\tilde{\phi}(\tilde{x}, u), t + u),$$

where  $\tilde{\phi}(\cdot, \cdot) : \tilde{\mathbf{X}} \times \mathbb{R}_+^0 \rightarrow \tilde{\mathbf{X}}$  is the measurable flow in  $\tilde{\mathbf{X}}$  possessing the semigroup property. Similarly,  $l(x, a) = (\tilde{l}(\tilde{x}, a), 0)$ , where  $\tilde{l}(\cdot, \cdot) : \mathbf{X} \times \mathbf{A} \rightarrow \tilde{\mathbf{X}}$  is a measurable mapping: after each impulse, the  $t$ -component goes down to zero. Any initial state is in the form  $x_0 = (\tilde{x}_0, 0)$  and thus has the time component zero. The mappings  $\tilde{\phi}$  and  $\tilde{l}$  are assumed to be measurable.

**Remark 2.1** *If the original state space is just  $\tilde{\mathbf{X}}$ , then it is always possible to extend it by including the component  $t$ .*

We exclude from the consideration all the points from  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$  which cannot appear in the dynamical system generated by the flow  $\tilde{\phi}$ , so that

$$\mathbf{X} := \{(\tilde{y}, t) \in \tilde{\mathbf{X}} \times \mathbb{R}_+^0 : \tilde{y} = \tilde{\phi}(\tilde{x}, t) \text{ for some } \tilde{x} \in \tilde{\mathbf{X}}\}.$$

In  $\mathbb{R}_+^0$ , the standard Euclidean topology is fixed. The product space  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$  is equipped with the product topology, which is metrizable (see [1, §2.14]). The topology on  $\mathbf{X}$  is the restriction of the product topology on  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$  on it. We endow  $\mathbf{X}$  with its Borel  $\sigma$ -algebra, which is the restriction of the Borel  $\sigma$ -algebra  $\mathcal{B}(\tilde{\mathbf{X}} \times \mathbb{R}_+^0)$  on  $\mathbf{X}$ , see [5, Lem.7.4]. Since  $\mathbf{X}$  is a projection of the graph of the measurable mapping  $\tilde{\phi}$ , it is not immediately obvious whether  $\mathbf{X}$  is a Borel subset of  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$ . In this and the next section, we assume that  $\mathbf{X}$  is a Borel space. Sufficient conditions will be imposed later to guarantee this is indeed the case (see Lemma 4.1).

Let  $\mathbf{X}_\Delta := \mathbf{X} \cup \{\Delta\}$ , where  $\Delta$  is an isolated artificial point describing the case that the controlled process is over and no future costs will appear. The dynamics (trajectory) of the system can be represented as one of the following sequences

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

where  $x_0 \in \mathbf{X}$  is the initial state of the controlled process and  $\theta_i < +\infty$  for all  $i = 1, 2, \dots, n$ . For the state  $x_{i-1} \in \mathbf{X}$ ,  $i \in \mathbb{N}$ , the pair  $(\theta_i, a_i) \in \bar{\mathbb{R}}_+^0 \times \mathbf{A} =: \mathbf{B}$  is the control at the step  $i$ : after  $\theta_i$  time units, the impulsive action  $a_i$  will be applied leading to the new state

$$x_i = \begin{cases} l(\phi(x_{i-1}, \theta_i), a_i), & \text{if } \theta_i < +\infty; \\ \Delta, & \text{if } \theta_i = +\infty. \end{cases} \tag{2}$$

The state  $\Delta$  will appear forever, after it appeared for the first time, i.e., it is absorbing.

**Remark 2.2** *We underline that all the realized points  $x_i$ ,  $i = 1, 2, \dots$ , provided that they are not equal to  $\Delta$ , have the form  $(\tilde{x}_i, 0)$ . For technical needs, unless stated otherwise, we allow  $x_0$  to be an arbitrary point in  $\mathbf{X}$ .*

After each impulsive action, if  $\theta_1, \theta_2, \dots, \theta_{i-1} < +\infty$ , the decision maker has in hand the complete information about the history, that is, the sequence

$$x_0, (\theta_1, a_1), x_1, \dots, (\theta_{i-1}, a_{i-1}), x_{i-1}.$$

The selection of the next control  $(\theta_i, a_i)$  is based on this information, and we also allow the selection of the pair  $(\theta_i, a_i)$  to be randomized. Below, the control  $(\theta, a) \in \mathbf{B}$  is denoted as  $b$ .

For each  $j = 0, 1, \dots, J$ , the cost accumulated on the coming interval of length  $\theta_i$  equals

$$\int_{[0, \theta_i]} C_j^g(\phi(x_{i-1}, u)) du + \mathbb{I}\{\theta_i < +\infty\} C_j^I(\phi(x_{i-1}, \theta_i), a_i), \tag{3}$$

the last term being absent if  $\theta_i = +\infty$ . The next state  $x_i$  is given by formula (2).

In the space of all the trajectories (1)

$$\begin{aligned} \Omega &= \cup_{n=1}^{\infty} [\mathbf{X} \times ((\mathbb{R}_+^0 \times \mathbf{A}) \times \mathbf{X})^n \times (\{+\infty\} \times \mathbf{A}) \times \{\Delta\} \times ((\mathbb{R}_+^0 \times \mathbf{A}) \times \{\Delta\})^\infty] \\ &\cup [\mathbf{X} \times ((\mathbb{R}_+^0 \times \mathbf{A}) \times \mathbf{X})^\infty], \end{aligned}$$

we fix the natural  $\sigma$ -algebra  $\mathcal{F}$ . Finite sequences

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

will be called (finite) histories;  $i = 0, 1, 2, \dots$ , and the space of all such histories will be denoted as  $\mathbf{H}_i$ ;  $\mathcal{F}_i := \mathcal{B}(\mathbf{H}_i)$  is the restriction of  $\mathcal{F}$  to  $\mathbf{H}_i$ . Capital letters  $X_i, T_i, \Theta_i, A_i, B_i = (\Theta_i, A_i)$  and  $H_i$  denote the corresponding functions of  $\omega \in \Omega$ , i.e., random elements.

**Definition 2.1** A control strategy  $\pi = \{\pi_i\}_{i=1}^\infty$  is a sequence of stochastic kernels  $\pi_i$  on  $\mathbf{B} = \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  given  $\mathbf{H}_{i-1}$ . A Markov strategy is defined by stochastic kernels  $\{\pi_i(db|x_{i-1})\}_{i=1}^\infty$ . A control strategy is called stationary, and denoted as  $\tilde{\pi}$ , if there is a stochastic kernel  $\tilde{\pi}$  on  $\bar{\mathbb{R}}_+^0 \times \mathbf{A}$  given  $\mathbf{X}_\Delta$  such that  $\pi_i(db|h_{i-1}) = \tilde{\pi}(db|x_i)$  for all  $i = 1, 2, \dots$ . Every measurable mapping  $f : \mathbf{X}_\Delta \rightarrow \mathbf{B}$  defines a deterministic stationary strategy, which is given by  $\pi_i(db|h_{i-1}) := \delta_{f(x_{i-1})}(db)$ , and identified with  $f$ .

Note that every Markov strategy can be represented as

$$\pi_i(d\theta \times da|x) = p_T^i(d\theta|x)p_A^i(da|x, \theta),$$

where  $p_T^i$  and  $p_A^i$  are stochastic kernels on  $\bar{\mathbb{R}}_+^0$  given  $\mathbf{X}_\Delta$  and on  $\mathbf{A}$  given  $\mathbf{X}_\Delta \times \bar{\mathbb{R}}_+^0$ , correspondingly: see [5, Prop.7.27].

For a given initial state  $x \in \mathbf{X}$  and a strategy  $\pi$ , there is a unique probability measure  $P_x^\pi(\cdot)$  on  $\Omega$  constructed using the Ionescu-Tulcea Theorem, satisfying for all  $i \in \mathbb{N}$ ,  $\Gamma \in \mathcal{B}(\bar{\mathbb{R}}_+^0 \times \mathbf{A})$ ,  $\Gamma_X \in \mathcal{B}(\mathbf{X}_\Delta)$ ,

$$\begin{aligned} P_x^\pi(X_0 \in \Gamma_X) &= \delta_x(\Gamma_X) \text{ for } \Gamma_X \in \mathcal{B}(\mathbf{X}_\Delta); \\ P_x^\pi((\Theta_i, A_i) \in \Gamma|H_{i-1}) &= \pi_i(\Gamma|H_{i-1}); \\ P_x^\pi(X_i \in \Gamma_X|H_{i-1}, (\Theta_i, A_i)) &= \begin{cases} \delta_{l(\phi(X_{i-1}, \Theta_i), A_i)}(\Gamma_X), & \text{if } X_{i-1} \in \mathbf{X}, \Theta_i < +\infty; \\ \delta_\Delta(\Gamma_X) & \text{otherwise.} \end{cases} \end{aligned} \quad (4)$$

This is a standard definition of strategic measures in Markov Decision Processes. Let  $E_x^\pi$  be the corresponding mathematical expectation.

Let us introduce the notation

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

for each strategy  $\pi$ ,  $j = 0, 1, \dots, J$  and initial state  $x \in \mathbf{X}$ .

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

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

Here and below,  $\{d_j\}_{j=1}^J$  are fixed constraint constants and  $x_0 = (\tilde{x}_0, 0)$  is a fixed initial state, where  $\tilde{x}_0 \in \tilde{\mathbf{X}}$ .

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

We shall assume that problem (5) is consistent.

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

In what follows, we develop the linear programming approach to problem (5).

### 3 Preliminary Observations

Clearly, the control model presented in Section 2, from the formal viewpoint, is a specific constrained Markov Decision Process [2, 14, 21, 28], which is defined by the following elements. The state space is  $\mathbf{X}_\Delta := \mathbf{X} \cup \{\Delta\}$ , as before, where the state  $\Delta \notin \mathbf{X}$  is an isolated point and  $\mathbf{X}$  is assumed to be a Borel space. The action space is  $\mathbf{B} := \bar{\mathbb{R}}_+^0 \times \mathbf{A}$ , which is endowed with the product topology and the corresponding Borel  $\sigma$ -algebra. The transition kernel is defined by

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

The cost functions are given by

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

If  $\theta = \infty$ , then the above integration is understood over  $[0, \infty)$ . Below, we omit such remarks. The initial state  $(\tilde{x}_0, 0) \in \mathbf{X}$  and the constraint constants  $d_j \in \mathbb{R}_+^0$ ,  $j = 1, 2, \dots, J$  are as before.

Let us impose the next set of compactness-continuity conditions.

**Condition 3.1** (a) *The space  $\mathbf{A}$  is compact, and  $+\infty$  is the one-point compactification of the positive real line  $\mathbb{R}_+^0$ .*

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

(c) *The mapping  $(x, \theta) \in \mathbf{X} \times \mathbb{R}_+^0 \rightarrow \phi(x, \theta)$  is continuous.*

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

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

According to Theorem 1 of [29], under Condition 3.1, assuming that  $\mathbf{X}$  is a Borel space, the function on  $\mathbf{X}$  defined by  $\inf_\pi E_x^\pi \left[ \sum_{i=0}^\infty \sum_{j=0}^J \bar{C}_j(X_i, B_{i+1}) \right]$  is lower semicontinuous.

**Condition 3.2** *There exists  $\delta > 0$  such that  $\sum_{j=0}^J C_j^I(x, a) \geq \delta$  for all  $(x, a) \in \mathbf{X} \times \mathbf{A}$ .*

The above condition asserts that each impulse is costly. Below in this section, we assume that Conditions 3.1 and 3.2 are satisfied.

Consider a point  $x \in \mathbf{X}$  such that

$$\inf_\pi E_x^\pi \left[ \sum_{i=0}^\infty \sum_{j=0}^J \bar{C}_j(X_i, B_{i+1}) \right] = 0 \tag{6}$$

(provided that such a point exists). Then  $E_x^{f^*} \left[ \sum_{i=0}^\infty \sum_{j=0}^J \bar{C}_j(X_i, B_{i+1}) \right] = 0$  for the deterministic stationary strategy  $f^*(x) \equiv (\infty, \hat{a})$  with the immaterial value of  $\hat{a} \in \mathbf{A}$  being arbitrarily fixed: for all other values of  $B_1 \in \mathbf{B}$ ,  $\sum_{j=0}^J \bar{C}_j(x, B_1) \geq \delta > 0$ .

Clearly, the control  $(\infty, \hat{a})$  is optimal in problem (5) at all such states  $x \in \mathbf{X}$ , at which (6) holds. Moreover, for all such states  $x$ ,  $Q(\{\Delta\}|x, f^*(x)) = 1$  and  $X_1 = \Delta$   $P_x^{f^*}$ -almost surely, so that

$$0 = E_x^{f^*} \left[ \sum_{i=0}^\infty \sum_{j=0}^J \bar{C}_j(X_i, B_{i+1}) \right] = \sum_{j=0}^J \bar{C}_j(x, (\infty, \hat{a})) = \int_{[0, \infty)} \sum_{j=0}^J C_j^g(\phi(x, u)) du,$$

and consequently, for all  $j = 0, 1, \dots, J$ ,  $C_j^g(\phi(x, u)) = 0$  for almost all  $u \geq 0$ . Conversely, if, at some  $x \in \mathbf{X}$ , for all  $j = 0, 1, \dots, J$ ,  $C_j^g(\phi(x, u)) = 0$  for almost all  $u \geq 0$ , then (6) holds.

Below, let us denote

$$V := \left\{ x \in \mathbf{X} : \inf_{\pi} E_x^{\pi} \left[ \sum_{i=0}^{\infty} \sum_{j=0}^J \bar{C}_j(X_i, B_{i+1}) \right] > 0 \right\} = \left\{ x \in \mathbf{X} : \int_{[0, \infty)} \sum_{j=0}^J C_j^g(\phi(x, u)) du > 0 \right\}$$

and  $V^c := \mathbf{X}_{\Delta} \setminus V$ . The set  $V$ , as the preimage of an open set under a lower semicontinuous function, is open in  $\mathbf{X}$ . The set  $V^c$  can be equivalently defined as

$$V^c := \{\Delta\} \cup \left\{ x \in \mathbf{X} : C_j^g(\phi(x, u)) = 0 \text{ for all } j = 0, 1, \dots, J, \text{ for almost all } u \geq 0 \right\},$$

and it is absorbing with respect to the flow  $\phi$ : for each  $x \in \mathbf{X}$ , as soon as  $\phi(x, u) \in V^c$ ,  $\phi(x, s) \in V^c$  for all  $s \geq u$ . The case  $V = \mathbf{X}$  and  $V^c = \{\Delta\}$  is not excluded.

In view of the previous observations, under Conditions 3.1 and 3.2, it is sufficient to consider the class of reasonable strategies  $\pi = \{\pi_i\}_{i=1}^{\infty}$  defined as follows.

**Definition 3.1** *Assume  $\mathbf{X}$  is a Borel space, and suppose Conditions 3.1 and 3.2 are satisfied. A strategy  $\pi = \{\pi_i\}_{i=1}^{\infty}$  is called reasonable if  $\pi_i(db|x_0, b_1, x_1, \dots, x_{i-1}) = \delta_{f^*(x_{i-1})}(db)$  for all  $x_{i-1} \in V^c$ , and*

$$\pi_i([\tilde{\theta}^*(\tilde{x}_{i-1}), \infty) \times \mathbf{A}|x_0, b_1, x_1, \dots, x_{i-1}) = 0, \quad i = 1, 2, \dots$$

Here,  $x_{i-1} = (\tilde{x}_{i-1}, 0)$  (see Remark 2.2) and

$$\tilde{\theta}^*(\tilde{x}) := \inf\{\theta \in \mathbb{R}_+^0 : \phi((\tilde{x}, 0), \theta) \in V^c\} \quad (7)$$

is a function defined for each  $\tilde{x} \in \tilde{\mathbf{X}}$ . (As usual,  $\inf \emptyset := +\infty$ .)

Since the flow  $\phi$  is continuous, the function  $\tilde{\theta}^*(\cdot)$  is measurable: see [13, Lemma 27.1] or [16, Prop.1.5, p.154]. After we introduce notations

$$\tilde{V} := \{\tilde{x} \in \tilde{\mathbf{X}} : (\tilde{x}, 0) \in V\} \quad \text{and} \quad \tilde{V}^c := \{\tilde{x} \in \tilde{\mathbf{X}} : (\tilde{x}, 0) \in V^c\} = \tilde{\mathbf{X}} \setminus \tilde{V},$$

it is clear that, for  $\tilde{x} \in \tilde{V}$ ,  $\tilde{\theta}^*(\tilde{x}) > 0$  because the set  $V$  is open and the set  $V^c$  is closed; in case  $\tilde{\theta}^*(\tilde{x}) < \infty$ , the infimum in (7) is attained, and

$$\tilde{\theta}^*(\tilde{x}) = \sup\{t \in \mathbb{R}_+^0 : \phi((\tilde{x}, 0), t) \in V\}.$$

If  $\tilde{x} \in \tilde{V}^c$ , then  $\tilde{\theta}^*(\tilde{x}) = 0$ .

We thus concentrate on selecting actions at the states  $x \in V$  and restrict ourselves to the set of reasonable strategies.

A linear programming method was established in [30] regarding how to select actions at  $x \in V$ , and it serves the beginning of the analysis in the present paper. For this reason, let us briefly describe it: see (10), (11) below. The formulation of that linear program is related to the occupation measures  $\mu^{\pi}$  defined as follows:

$$\mu^{\pi}(\Gamma_1 \times \Gamma_2) := E_{x_0}^{\pi} \left[ \sum_{i=0}^{\infty} \mathbb{I}\{(X_i, B_{i+1}) \in \Gamma_1 \times \Gamma_2\} \right], \quad \forall \Gamma_1 \in \mathcal{B}(\mathbf{X}_{\Delta}), \Gamma_2 \in \mathcal{B}(\bar{\mathbb{R}}_+^0 \times \mathbf{A}). \quad (8)$$



Under Conditions 2.1, 3.1 and 3.2, for each reasonable  $\pi$  as in Definition 3.1,

$$\mathcal{V}_j(x_0, \pi) = \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \bar{C}_j(x, (\theta, a)) \mu^\pi(dx \times d\theta \times da).$$

It follows that the restriction on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  of any occupation measure  $\mu^\pi = \mu$  of our interest is concentrated on the measurable subset  $\mathbf{M} \times \mathbf{A}$ , where

$$\mathbf{M} := \{(y, \theta) : y = (\tilde{x}, 0) \text{ with } \tilde{x} \in \tilde{V} \text{ and } \theta \in [0, \tilde{\theta}^*(\tilde{x})] \cup \{\infty\}\}. \quad (9)$$

Moreover, there is no need to consider such occupation measures that  $\mu^\pi(\mathbf{M} \times \mathbf{A}) = \infty$ : the latter means that, with positive probability, actions from  $\bar{\mathbb{R}}_+^0 \times \mathbf{A}$  at states from  $V$  appear infinitely many times, leading to the infinite value of at least one of the objectives  $\mathcal{V}_j(x_0, \pi) = \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \bar{C}_j(x, (\theta, a)) \times \mu^\pi(dx \times d\theta \times da)$ .

The impulse control problem (5) is now equivalent to the following linear program:

$$\text{Minimize} \quad : \quad \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \bar{C}_0(x, (\theta, a)) \mu(dx \times d\theta \times da) \quad (10)$$

over finite measures  $\mu$  on  $V \times \mathbf{B} = V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  concentrated on  $\mathbf{M} \times \mathbf{A}$

$$\text{subject to } \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) = \delta_{x_0}(dx) + \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} Q(dx|y, (\theta, a)) \mu(dy \times d\theta \times da) \text{ on } \mathcal{B}(V); \quad (11)$$

$$\int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \bar{C}_j(x, (\theta, a)) \mu(dx \times d\theta \times da) \leq d_j, \quad j = 1, 2, \dots, J.$$

See Proposition 3.1 for a precise statement of this equivalence.

One can recognize that the form of this linear program is standard for the total cost Markov Decision Processes (see e.g., [2, 14, 21]). For every reasonable strategy  $\pi$ , the occupation measure  $\mu^\pi$  satisfies equality (11).

The next statement comes from Theorem 4.1 of [30].

**Proposition 3.1** *Suppose the space  $\mathbf{X}$  is Borel and Conditions 2.1, 3.1 and 3.2 are satisfied. Then the following assertions hold.*

- (a) *There exists a solution  $\mu^*$  to the program (10), (11), which gives rise to the optimal (in problem (5)) stationary strategy  $\tilde{\pi}$  coming from the decomposition*

$$\mu^*(dx \times db) = \mu^*(dx \times \mathbf{B}) \times \tilde{\pi}(db|x), \quad x \in V.$$

*On the space  $V^c$ , the optimal strategy is given by  $f^*(x) \equiv (\infty, \hat{a})$  as before; the value of  $\hat{a} \in \mathbf{A}$  is immaterial. The minimal value of the program (10), (11) is finite and coincides with the minimal value of the original problem (5).*

- (b) *If  $\pi^*$  is a reasonable strategy, whose occupation measure  $\mu^{\pi^*}$  on  $V \times \mathbf{B}$  is concentrated on  $\mathbf{M} \times \mathbf{A}$  and solves the linear program (10), (11), then the strategy  $\pi^*$  is optimal in problem (5).*

In what follows, we use the notation  $\mathbf{A}_\square = \mathbf{A} \cup \{\square\}$ , where  $\square \notin \mathbf{A}$  is an artificial isolated point.

The target of this article is to pass to the equivalent in some sense linear program in the space of measures  $\eta$  on  $V \times \mathbf{A}_\square$ . The reason is connected with the form of the objectives  $\mathcal{V}_j(x_0, \pi)$ . Since

they are linear with respect to the original functions  $C_j^g$  and  $C_j^I$  on  $V$  and  $V \times \mathbf{A}$  correspondingly, it is desirable to represent them in the form of  $\int_{V \times \mathbf{A}_\square} C_j(y, a) \eta(dy \times da)$ , where

$$C_j(y, a) := \begin{cases} C_j^g(y), & \text{if } a = \square; \\ C_j^I(y, a), & \text{if } a \in \mathbf{A}, \end{cases} \quad (12)$$

and develop the characteristic equation for the measures  $\eta$ .

Consider a finite measure  $\mu$  in the linear program (10), (11), which can be written in the form

$$\mu(dx \times d\theta \times da) = p_T(d\theta|x, a) p_A(da|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}), \quad (13)$$

where  $p_T(\cdot)$  and  $p_A(\cdot)$  are stochastic kernels on  $\bar{\mathbb{R}}_+^0$  and  $\mathbf{A}$  correspondingly: see [5, Prop.7.27]. The dependence of  $p_T$  and  $p_A$  on  $\mu$  is not explicitly indicated here. Hence, using the Tonelli Theorem (see [1, Thm.11.28]), some straightforward calculations imply that

$$\begin{aligned} & \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \int_{[0, \theta]} C_j^g(\phi(x, u)) du \right\} \mu(dx \times d\theta \times da) \\ &= \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \int_{[0, \theta]} C_j^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} du \right\} \mu(dx \times d\theta \times da) \\ &= \int_V \int_{\mathbf{A}} \int_{\bar{\mathbb{R}}_+^0} \int_{[0, \theta]} C_j^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} du p_T(d\theta|x, a) p_A(da|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &= \int_V \int_{\mathbf{A}} \int_{\bar{\mathbb{R}}_+^0} C_j^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} p_T([u, \infty]|x, a) du p_A(da|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}), \end{aligned}$$

where the first equality holds because  $C_j^g(\phi(x, u)) = 0$  for each  $x \in V$ , for almost all  $u \in \bar{\mathbb{R}}_+^0$  such that  $\phi(x, u) \in V^c$ , for all  $j = 0, 1, \dots, J$ . To put it differently,  $C_j^g(\phi(x, u)) = 0$  for almost all  $u \geq \tilde{\theta}^*(\tilde{x})$  for all  $x = (\tilde{x}, 0)$ .

After we introduce the following measure on  $V$

$$\begin{aligned} \eta(dy \times \square) &:= \int_V \int_{\bar{\mathbb{R}}_+^0} \delta_{\phi(x, u)}(dy) \mathbb{I}\{\phi(x, u) \in V\} \left( \int_{\mathbf{A}} p_T([u, \infty]|x, a) p_A(da|x) \right) du \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &= \int_{\bar{\mathbb{R}}_+^0} \left\{ \int_{V \times \mathbf{A}} \delta_{\phi(x, u)}(dy) \mu(dx \times [u, \infty] \times da) \right\} du \\ &= \int_{\bar{\mathbb{R}}_+^0} \left\{ \int_{\tilde{V}} \delta_{\phi((\tilde{x}, 0), u)}(dy) \mu(d\tilde{x} \times \{0\} \times [u, \infty] \times \mathbf{A}) \right\} du, \end{aligned} \quad (14)$$

we may write  $\int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \int_{[0, \theta]} C_j^g(\phi(x, u)) du \right\} \mu(dx \times d\theta \times da) = \int_V C_j^g(y) \eta(dy \times \square)$ .

Similarly to the above, taking into account that the measure  $\mu$  is concentrated on  $\mathbf{M} \times \mathbf{A}$ , we have that, for each  $j = 0, 1, \dots, J$ ,

$$\begin{aligned} & \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \mathbb{I}\{\theta < +\infty\} C_j^I(\phi(x, \theta), a) \right\} \mu(dx \times d\theta \times da) \\ &= \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \mathbb{I}\{\theta < +\infty\} \mathbb{I}\{\phi(x, \theta) \in V\} C_j^I(\phi(x, \theta), a) \right\} \mu(dx \times d\theta \times da) \\ &= \int_{V \times \mathbf{A}} C_j^I(y, a) \eta(dy \times da), \end{aligned}$$

where

$$\eta(dy \times da) := \int_V \int_{\mathbb{R}_+^0} \delta_{\phi(x,\theta)}(dy) \mathbb{I}\{\phi(x,\theta) \in V\} \mu(dx \times d\theta \times da) \quad (15)$$

is a finite measure on  $V \times \mathbf{A}$ , since the measure  $\mu$  is finite.

If Conditions 2.1, 3.1 and 3.2 are satisfied, and the space  $\mathbf{X}$  is Borel, then the linear program (10), (11) can now be rewritten as

$$\begin{aligned} \text{Minimize} & : \int_{V \times \mathbf{A}_\square} C_0(y, a) \eta(dy \times da) & (16) \\ & \text{over finite measures } \mu \text{ on } V \times \mathbf{B} = V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A} \text{ concentrated on } \mathbf{M} \times \mathbf{A} \\ \text{subject to} & : (11), (14), (15) \text{ and } \int_{V \times \mathbf{A}_\square} C_j(y, a) \eta(dy \times da) \leq d_j, \quad j = 1, 2, \dots, J. \end{aligned}$$

The space  $\mathbf{A}_\square := \mathbf{A} \cup \{\square\}$  and the functions  $C_j$  are as introduced above: see (12). Proposition 3.1 is valid for the linear program (16), too.

**Definition 3.2** *Suppose Conditions 2.1, 3.1 and 3.2 are satisfied, and assume that the space  $\mathbf{X}$  is a Borel space. For a finite measure  $\mu$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  satisfying equation (11), the measure  $\eta$  on  $V \times \mathbf{A}_\square$  defined by*

$$\eta(\Gamma_X \times \Gamma_A) := \eta(\Gamma_X \times (\Gamma_A \cap \mathbf{A})) + \eta(\Gamma_X \times \square) \mathbb{I}\{\square \in \Gamma_A\}, \quad \Gamma_X \in \mathcal{B}(V), \quad \Gamma_A \in (\mathbf{A}_\square), \quad (17)$$

where the measures  $\eta(dy \times \square)$  on  $V$  and  $\eta(dy \times da)$  on  $V \times \mathbf{A}$  were introduced in (14) and (15), is called the aggregated occupation measure (induced by  $\mu$ ).

In what follows, we will characterize the aggregated measures  $\eta$  without references to the measures  $\mu$ : see linear program (20), (21).

## 4 Main Results

**Definition 4.1** *We call the orbit of a point  $\tilde{x}^0 \in \tilde{\mathbf{X}}$  the following subset of  $\mathbf{X}$ :*

$$\tilde{x}^0 \mathcal{X} = \{(\tilde{\phi}(\tilde{x}^0, t), t) : t \in \mathbb{R}_+^0\} = \{\phi((\tilde{x}^0, 0), t) : t \in \mathbb{R}_+^0\}.$$

We underline that the flow  $\phi$  has no cycles and, if the flows  $\phi$  and  $\tilde{\phi}$  are continuous, then every orbit is a closed set in  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$ .

**Condition 4.1** *Two different orbits do not intersect, i.e., for any two distinct points  $\tilde{x}_1^0 \neq \tilde{x}_2^0 \in \tilde{\mathbf{X}}$ ,  $\tilde{x}_1^0 \mathcal{X} \cap \tilde{x}_2^0 \mathcal{X} = \emptyset$ .*

**Definition 4.2** *Under Condition 4.1, for each  $y = (\tilde{y}, t) \in \mathbf{X}$ , we introduce  $h(y)$  equal to the point  $\tilde{x}^0 \in \tilde{\mathbf{X}}$  such that  $\tilde{y} = \tilde{\phi}(\tilde{x}^0, t)$  and put  $\tau_y = t$ . The mappings  $F : \tilde{\mathbf{X}} \times \mathbb{R}_+^0 \rightarrow \mathbf{X}$  and  $F^{-1} : \mathbf{X} \rightarrow \tilde{\mathbf{X}} \times \mathbb{R}_+^0$  are defined as*

$$F(\tilde{x}^0, t) := (\tilde{\phi}(\tilde{x}^0, t), t) = \phi((\tilde{x}^0, 0), t), \quad \text{and} \quad F^{-1}(y) = (h(y), \tau_y). \quad (18)$$

Note that the mapping  $h : \mathbf{X} \rightarrow \tilde{\mathbf{X}}$  is well defined: if, for  $y = (\tilde{y}, t) \in \mathbf{X}$ , for two points  $\tilde{x}_1^0 \neq \tilde{x}_2^0$  from  $\tilde{\mathbf{X}}$ ,  $\tilde{y} = \tilde{\phi}(\tilde{x}_1^0, t) = \tilde{\phi}(\tilde{x}_2^0, t)$ , then the different orbits  $\tilde{x}_1^0 \mathcal{X}$  and  $\tilde{x}_2^0 \mathcal{X}$  are not disjoint having the common point  $y$ .



**Remark 4.1** Suppose 2.1, 3.1, 3.2, and 4.2 are satisfied. Then the mapping  $F$  defined in (18) (its restriction on  $\mathbf{D}$ , to say more precisely) provides a homeomorphism and thus also an isomorphism between the sets

$$\mathbf{D} := \{(\tilde{x}^0, t) : \tilde{x}^0 \in \tilde{V}, 0 \leq t < \tilde{\theta}^*(\tilde{x}^0)\} = \{(\tilde{x}^0, t) : \phi((\tilde{x}^0, 0), t) \in V\} \quad (19)$$

and  $V$ . Indeed,  $F(\tilde{x}^0, t) \in V$  if and only if the pair  $(\tilde{x}^0, t)$  belongs to the set  $\mathbf{D}$ . Thus,  $F(\mathbf{D}) = V$  and  $F^{-1}(V) = \mathbf{D}$ . Recall that  $\phi((\tilde{x}^0, 0), t) \in V^c$  for all  $t \geq \tilde{\theta}^*(\tilde{x}^0)$ .

We underline that the points  $(\tilde{x}^0, t) \in \mathbf{D}$  and  $(\tilde{y}, t) \in \mathbf{X}$  have different meanings, although the components  $\tilde{x}^0, \tilde{y} \in \tilde{\mathbf{X}}$  and  $t \in \mathbb{R}_+^0$  look the same. That is the reason to equip the first coordinates of points in  $\mathbf{D}$  with the upper index 0, to make them look different from the points in  $\mathbf{X}$ . The pair  $(\tilde{x}^0, t) \in \mathbf{D}$  is just the reference point of the orbit  $_{\tilde{x}^0}\mathcal{X}$  and the duration of movement from  $(\tilde{x}^0, 0)$ . It can easily happen that  $(\tilde{x}^0, t) \notin \mathbf{X}$ .

**Definition 4.3** Suppose Conditions 2.1, 3.1, 3.2, and 4.2 are satisfied. If  $\zeta$  is a measure on  $V$ , then  $\check{\zeta}$  denotes the image of  $\zeta$  on  $\mathbf{D}$  under the mapping  $F^{-1}$ :

$$\check{\zeta}(\Gamma) = \zeta(F(\Gamma)), \quad \Gamma \in \mathcal{B}(\mathbf{D}).$$

In case the measure  $\zeta$  is finite, we, with slight but convenient abuse of notations, introduce  $\hat{\zeta}(\tilde{\Gamma}) := \check{\zeta}(\tilde{\Gamma} \times \mathbb{R}_+^0)$  for  $\tilde{\Gamma} \in \mathcal{B}(\tilde{V})$  and  $\check{\zeta}(dt|\tilde{x}^0)$ , the stochastic kernel from  $\tilde{V}$  to  $\mathbb{R}_+^0$  such that

$$\check{\zeta}(d\tilde{x}^0 \times dt) = \hat{\zeta}(d\tilde{x}^0)\check{\zeta}(dt|\tilde{x}^0),$$

see [5, Cor.7.27.2] or [20, Prop.D.8].

Clearly,

$$\check{\zeta}_1 \leq \check{\zeta}_2 \text{ set-wise} \iff \hat{\zeta}_1 \leq \hat{\zeta}_2 \text{ set-wise.}$$

Note that if  $\zeta$  is a finite measure, then  $\check{\zeta}([0, \tilde{\theta}^*(\tilde{x}^0)]|\tilde{x}^0) = 1$  for  $\hat{\zeta}$ -almost all  $\tilde{x}^0 \in \tilde{V}$ , and we extend the kernel  $\check{\zeta}$  to  $\mathbb{R}_+^0$  by putting  $\check{\zeta}([\tilde{\theta}^*(\tilde{x}^0), \infty)|\tilde{x}^0) := 0$ . If the measure  $\zeta$  is zero outside the set  $\tilde{V} \times \{0\}$ , then  $\check{\zeta}(\Gamma) = 0$  for all measurable subsets  $\Gamma \subseteq \mathbf{D} \cap \{(\tilde{x}^0, t) \in \tilde{\mathbf{X}} \times \mathbb{R}_+^0 : t > 0\}$ ,  $\hat{\zeta}(\tilde{\Gamma}) = \zeta(\tilde{\Gamma} \times \{0\})$  for all  $\tilde{\Gamma} \in \mathcal{B}(\tilde{V})$ , and  $\check{\zeta}(dt|\tilde{x}^0) = \delta_0(dt)$  for  $\hat{\zeta}$ -almost all  $\tilde{x}^0 \in \tilde{V}$ .

**Definition 4.4** Under Conditions 2.1, 3.1, 3.2 and 4.2, a measure  $\zeta$  on  $V$  is called normal if there exist a finite measure  $L$  on  $\tilde{V}$  and a bounded measurable function  $g(\tilde{x}^0, u) : \tilde{V} \times \mathbb{R}_+^0 \rightarrow \mathbb{R}_+^0$  such that

$$\check{\zeta}(d\tilde{x}^0 \times du) = g(\tilde{x}^0, u)du L(d\tilde{x}^0).$$

Equivalently, for all  $\Gamma_X \times \Gamma_t \in \mathcal{B}(V)$ ,

$$\begin{aligned} \zeta(\Gamma_X \times \Gamma_t) &= \int_{\mathbf{D}} \mathbb{I}\{F(\tilde{x}^0, u) \in \Gamma_X \times \Gamma_t\} \check{\zeta}(d\tilde{x}^0 \times du) \\ &= \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0))} \delta_{\tilde{\phi}(\tilde{x}^0, u)}(\Gamma_X) \delta_u(\Gamma_t) g(\tilde{x}^0, u) du L(d\tilde{x}^0). \end{aligned}$$

See Remark 4.1.

A measure  $\eta$  on  $V \times \mathbf{A}_\square$  is called normal if  $\eta(V \times \mathbf{A}) < \infty$  and the measure  $\eta(dx \times \square)$  on  $V$  is normal.

Clearly, every normal measure is  $\sigma$ -finite. Similarly, a normal measure  $\zeta$  defined on some orbit  ${}_z\mathcal{X} \subseteq \mathbf{X}$  is understood:

$$\check{\zeta}(d\tilde{x}^0 \times du) = g(u)du \delta_{\tilde{z}}(d\tilde{x}^0).$$

**Lemma 4.2** *Suppose Conditions 2.1, 3.1, 3.2 and 4.2 are satisfied. Then the following assertions hold true.*

- (a) *For every finite measure  $\mu$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  satisfying equality (11), the induced aggregated occupation measure  $\eta$  on  $V \times \mathbf{A}_\square$  is normal.*
- (b) *If  $\eta^1$  and  $\eta^2$  are two normal measures on  $V \times \mathbf{A}_\square$  such that  $\eta^1 \geq \eta^2$  set-wise, and thus the difference  $\eta^1 - \eta^2$  is a (positive) measure, then  $\eta := \eta^1 - \eta^2$  is also a normal measure on  $V \times \mathbf{A}_\square$ .*

In Definition 4.5, we introduce the class of so called test functions used to characterize measures on  $\mathbf{X}$ .

**Definition 4.5**  $\mathbf{W}$  is the space of measurable bounded functions  $w$  on  $\mathbf{X}$ , absolutely continuous, either negative and increasing or positive and decreasing along the flow  $\phi$  (see Definition A.1) and satisfying conditions

- $w(y) = 0$  for all  $y \in V^c$  and
- $\lim_{t \rightarrow \infty} w(\phi(x, t)) = 0$  for all  $x \in V$  such that  $\phi(x, t) \in V$  for all  $t \in \mathbb{R}_+^0$ .

Throughout this paper,  $\chi w$  denotes a function as in Lemma A.1 (see Appendix). Without loss of generality, one can assume, for each negative (or positive) function  $w \in \mathbf{W}$ , that the function  $\chi w$  is positive (or negative), i.e., in (47) one can put  $g(\cdot) \equiv 0$ . Note that below we consider only such measures  $\zeta$  on  $V$  that the value of the integral  $\int_V \chi w(x) \zeta(dx)$  does not depend on the function  $g$  in (47).

Suppose Conditions 2.1, 3.1, 3.2, and 4.2 are satisfied and introduce the following linear program

$$\begin{aligned} & \text{Minimize over} \\ & \text{the normal measures } \eta \text{ on } V \times \mathbf{A}_\square : \int_{V \times \mathbf{A}_\square} C_0(x, a) \eta(dx \times da) \quad (20) \\ & \text{subject to} : w(x_0) + \int_V \chi w(x) \eta(dx \times \square) - \int_V w(x) \eta(dx \times \mathbf{A}) \quad (21) \\ & \quad + \int_{V \times \mathbf{A}} w(l(x, a)) \eta(dx \times da) = 0 \quad \forall w \in \mathbf{W}; \\ & \quad \int_{V \times \mathbf{A}_\square} C_j(x, a) \eta(dx \times da) \leq d_j, \quad j = 1, 2, \dots, J. \end{aligned}$$

The space  $\mathbf{A}_\square$  and functions  $C_j$  were defined in Section 3: see (12).

**Remark 4.2** *Compared with (10), (11) and (16), the dimensionality of the linear program (20),(21) is reduced in the sense that the measures  $\mu$  were on the space  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$ , and the measures  $\eta$  are on the space  $V \times (\mathbf{A} \cup \{\square\})$ . Therefore, e.g., from the computational point of view, the linear program (20),(21) is easier.*

We are ready to formulate the main results.

**Theorem 4.1** *Suppose Conditions 2.1, 3.1, 3.2 and 4.2 are satisfied. Then, for every finite measure  $\mu$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$ , concentrated on  $\mathbf{M} \times \mathbf{A}$  and satisfying equality (11), its induced aggregated occupation measure  $\eta$  on  $V \times \mathbf{A}_\square$  satisfies equation (21) for all functions  $w \in \mathbf{W}$ . All the integrals in (21) are finite.*

The proof of this statement is postponed to Section 6.

**Theorem 4.2** *Suppose Conditions 2.1, 3.1, 3.2 and 4.2 are satisfied. Then every normal measure  $\eta$  on  $V \times \mathbf{A}_\square$ , satisfying equation (21), uniquely defines a reasonable Markov strategy  $\pi^\eta$  (called “induced” by  $\eta$ ) such that, for the aggregated occupation measure  $\tilde{\eta}$  defined by (17) (recall (14) and (15)) with  $\mu$  being replaced by the occupation measure  $\mu^{\pi^\eta}$  of the strategy  $\pi^\eta$  as in (8) with  $\pi = \pi^\eta$ , the following inequalities hold:*

$$\tilde{\eta}(\Gamma) \leq \eta(\Gamma) \quad \forall \Gamma \in \mathcal{B}(V \times \mathbf{A}_\square).$$

The proofs of Theorem 4.2 and of the next corollary are postponed to Section 7.

**Corollary 4.1** *Let Conditions 2.1, 3.1, 3.2, and 4.2 be satisfied. Then linear program (16) is equivalent to linear program (20),(21).*

*To be more precise, if the finite measure  $\mu^*$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  solves linear program (16), then the measure  $\eta^*$  on  $V \times \mathbf{A}_\square$ , given by (14), (15) and (17), i.e., the aggregated occupation measure induced by  $\mu^*$ , solves linear program (20),(21). Conversely, if the measure  $\eta^*$  on  $V \times \mathbf{A}_\square$  solves linear program (20),(21), then, for the Markov strategy  $\pi^*$  induced by  $\eta^*$  as in Theorem 4.2, the corresponding occupation measure  $\mu^{\pi^*}$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$ , defined in (8), solves linear program (16).*

According to Corollary 4.1 and Section 3 (see Proposition 3.1), the minimal values of the linear programs (10), (11) and (20),(21) coincide and equal the minimal value of the original problem (5). As soon as the optimal solution  $\eta^*$  to the linear program (20),(21) is obtained, the induced Markov strategy  $\pi^*$ , solves the original optimal impulsive control problem (5): see Proposition 3.1 and remember that the linear programs (10), (11) and (16) are equivalent. Recall that linear program (16) has an optimal solution by Proposition 3.1; hence the linear program (20),(21) is also solvable. Note also that, having in hand the Markov strategy  $\pi^*$ , one can compute the corresponding occupation measure  $\mu^{\pi^*}$  (8), and after that the stationary strategy as in Proposition 3.1 also solves the optimal impulsive control problem (5).

For the discussions in the rest of this section, we suppose all the mappings and functions  $l, C_j^g$ , and  $C_j^I$  do not depend on the component  $t$  of the state  $x = (\tilde{x}, t)$ . Then the linear program (20),(21) is actually in terms of (marginal) measures  $\tilde{\eta}(d\tilde{x} \times da)$  on  $\tilde{V} \times \mathbf{A}_\square$  defined by

$$\tilde{\eta}(\Gamma_X \times \Gamma_A) := \eta([( \Gamma_X \times [0, \infty) ) \cap V] \times \Gamma_A) = \int_{(\Gamma_X \times [0, \infty)) \cap V} \eta(d\tilde{x} \times dt \times \Gamma_A).$$

The marginals  $\tilde{\eta}$  of normal measures  $\eta$  (naturally called normal on  $\tilde{V} \times \mathbf{A}_\square$ ) are characterized as follows:  $\tilde{\eta}(\tilde{V} \times \mathbf{A}) < \infty$  and there exist a finite measure  $L$  on  $\tilde{V}$  and a bounded non-negative measurable function  $g$  on  $\tilde{V} \times \mathbb{R}_+^0$  such that

$$\tilde{\eta}(\Gamma_X \times \square) = \eta([( \Gamma_X \times [0, \infty) ) \cap V] \times \square) = \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x})]} \delta_{\tilde{\phi}(\tilde{x}, u)}(\Gamma_X) g(\tilde{x}, u) du L(d\tilde{x}) \quad (22)$$

(see Definition 4.4). The test functions  $\tilde{w}$  on  $\tilde{\mathbf{X}}$  are measurable bounded, absolutely continuous, either negative and increasing or positive and decreasing along the flow, and such that  $\tilde{w}(y) = 0$  for all  $y \in \tilde{V}^c$  and  $\lim_{t \rightarrow \infty} \tilde{w}(\tilde{\phi}(\tilde{x}, t)) = 0$  for all  $\tilde{x} \in \tilde{V}$  such that  $\tilde{\phi}(\tilde{x}, t) \in \tilde{V}$  for all  $t \in \mathbb{R}_+^0$ .

This linear program, in terms of marginal measures  $\tilde{\eta}$ , is solvable under Conditions 2.1, 3.1, 3.2, and 4.2. (The last condition is for the model with the extended state space  $\mathbf{X} \subset \tilde{\mathbf{X}} \times \mathbb{R}_+^0$ .) The minimal value of this program coincides with the minimal value of the original problem (5). Therefore, when reformulating the optimal impulsive control problem in terms of aggregated occupation measures, the extension of the state space, as in Remark 2.1, is not needed.

On the other hand, the construction of the optimal control strategy  $\pi^*$ , induced by the optimal solution  $\eta^*$  to the linear program (20),(21), is essentially based on the analysis in the extended state space: see the proof of Theorem 4.2. Note that, in the case of the extended state space, the (full) orbits in  $\mathbf{X}$ , as in Definition 4.1, form a Borel space because they are characterized by the starting points  $\tilde{x}^0$ . If one manages to describe the space of orbits in  $\tilde{\mathbf{X}}$  as a Borel space, then one can avoid such an extension of the basic state space  $\tilde{\mathbf{X}}$ .

## 5 Example and Comparison with Other Works

Consider the following simple but not trivial optimal impulse control problem in the space  $\mathbb{R}_+^0$ .

$$\left. \begin{aligned} d\tilde{x} &:= G(\tilde{x})dt + dW(t), & \tilde{x}(0-) = \tilde{x}_0 > 0; \\ & \int_0^\infty C^g(\tilde{x}(u))du + \int_0^\infty dW(u) \rightarrow \inf_W, \end{aligned} \right\} \quad (23)$$

where

$$W(u) := \sum_{j=1}^{\infty} \left( \sum_{i=1}^{j-1} a_i \right) I\{T_{j-1} \leq u < T_j\};$$

$$0 = T_0 \leq T_1 \leq T_2 \leq \dots, \quad T_j \in [0, \infty]; \quad T_{j-1} = T_j \text{ only if } T_{j-1} = \infty; \quad \lim_{j \rightarrow \infty} T_j = \infty.$$

The impulse control strategy  $W$ , represented by  $\{T_j, a_j\}_{j=1}^\infty$ , can be arbitrary, satisfying the condition  $a_j \geq \delta > 0$ . Here and below,  $\delta > 0$  is a fixed number. The measurable functions  $G(\cdot) > \delta \geq 0$  and  $C^g(\cdot) \geq 0$  are fixed and smooth enough, such that  $H := \int_0^\infty C^g(\tilde{X}(u))du < \infty$ . Here  $\tilde{X}(\cdot)$  is the solution to (23) when  $W(u) \equiv 0$ , i.e., when  $T_1 = \infty$ .

Clearly, one can restrict himself to the strategies (i.e., functions  $W$ ) having only a finite number of impulses, that is,  $T_1, T_2, \dots, T_{N-1} < \infty$  and  $T_N = \infty$  for some  $N \in \mathbb{N}$ : for other strategies the objective in (23) equals  $+\infty$ .

This problem can be easily reformulated in terms of Section 2. The flow  $\tilde{\phi}$  on  $\tilde{\mathbf{X}} = [\tilde{x}_0, \infty)$  comes from the differential equation (23) at  $W(\cdot) \equiv 0$ ;  $\mathbf{A} = [\delta, H]$  (no reason to apply the impulses  $a > H$ );  $\mathbf{B} = \mathbb{R}_+^0 \times \mathbf{A}$ ;  $x_0 = (\tilde{x}_0, 0)$ ;  $\tilde{l}(\tilde{x}, t, a) = \tilde{l}(\tilde{x}, a) = \tilde{x} + a$ ; as usual,  $t$ , the second components of the state, is the time elapsed since the most recent impulse. We consider the unconstrained case with  $J = 0$ , and the  $j$  index in the functions like  $C_j^g(\cdot)$  and  $C_j^I(\cdot)$  is omitted. The gradual cost rate is  $C^g(\cdot)$ , and the cost of the impulse  $a \in \mathbf{A}$  equals  $C^I(x, a) = a$ . Simultaneous impulses of the sizes  $a, b, \dots$  can be considered as one impulse of the size  $a + b + \dots$ , and consequently, for  $j > 0$ , if  $T_j < \infty$  then  $T_{j+1} > T_j$ . We assume that, for some  $K \in (\tilde{x}_0, \infty)$ ,  $C^g(\tilde{x}) = 0$  for  $\tilde{x} \geq K$  and  $C^g(\tilde{x}) > 0$  for  $\tilde{x} < K$ , so that  $\tilde{V} = [\tilde{x}_0, K)$ , and  $\tilde{V}^c = [K, \infty)$ . Now, all the Conditions 2.1, 3.1, 3.2, and 4.2 are satisfied, and hence the minimal value of the impulse control problem (23) coincides with the minimal value of the linear program (20),(21).

Let us illustrate the definitions introduced in Sections 2 and 3. To be specific, below, unless stated otherwise, we put  $\tilde{x}_0 > \delta$ ,  $G(\tilde{x}) = \tilde{x}$ , so that  $\tilde{\phi}(\tilde{x}, t) = \tilde{x}e^t$  is the solution to (23) starting from  $\tilde{x} \in \tilde{\mathbf{X}}$ , when  $W(u) \equiv 0$ , i.e.,  $T_1 = \infty$ ;

$$\mathbf{X} = \{(\tilde{y}, t) : \tilde{y} \in \tilde{\mathbf{X}} = [\tilde{x}_0, \infty), t \in [0, \ln \frac{\tilde{y}}{\tilde{x}_0}]\} :$$



a point  $\tilde{y}$  cannot appear later than  $\ln \frac{\tilde{y}}{\tilde{x}_0}$  time units after any one impulse;  $\mathbf{X}_\Delta = \mathbf{X} \cup \{\Delta\}$ . See Figure 2. Now

$$\begin{aligned} V &= \{(\tilde{y}, t) : \tilde{y} \in \tilde{V} = [\tilde{x}_0, K), t \in [0, \ln \frac{\tilde{y}}{\tilde{x}_0}]\}; \\ \tilde{\theta}^*(\tilde{x}) &= \ln \frac{K}{\tilde{x}} \text{ for } \tilde{x} \in \tilde{V}; \\ \mathbf{M} &= \{(y, \theta) : y = (\tilde{x}, 0) \text{ with } \tilde{x} \in \tilde{V} = [0, K) \text{ and } \theta \in [0, \tilde{\theta}^*(\tilde{x}) = \ln \frac{K}{\tilde{x}}) \cup \{\infty\}\} : \end{aligned}$$

there is no reason to apply impulses after the variable  $\tilde{x}$  reaches the region  $\tilde{V}^c = [K, \infty)$ .

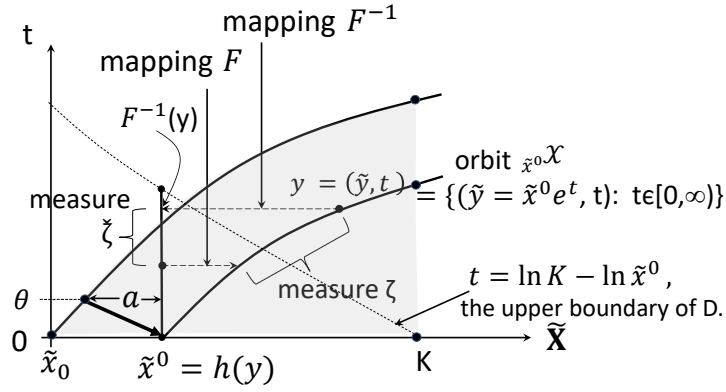


Figure 2:  $\mathbf{X}$  is the area below the orbit starting from  $\tilde{x}_0$ ; the grey area is  $V$ . The bold arrow leading to the point  $\tilde{x}^0$  represents the impulse of the size  $a$  applied at the time moment  $T_1 = \theta$ .

The model (23) is deterministic and unconstrained, so that we fix an arbitrary reasonable deterministic control strategy  $\pi$  with  $\mu^\pi(\mathbf{M} \times \mathbf{A}) < \infty$ . This means that

$$\begin{aligned} \pi_i(d\theta \times da | x_{i-1}) &= \delta_{\theta_i}(d\theta) \delta_{a_i}(da); \quad \theta_i = T_i - T_{i-1} < \infty \text{ for } i = 1, 2, \dots, N-1, \text{ and } \theta_N = \infty; \\ x_i &= (\tilde{x}_i = \tilde{x}_{i-1} e^{\theta_i} + a_i, 0) \text{ and } \tilde{x}_{i-1} e^{\theta_i} \in \tilde{V} \text{ for } i = 1, 2, \dots, N-1; \\ x_N &= \Delta. \end{aligned}$$

All the variables  $T_i, \theta_i$  etc are here non-random;  $N-1$  is the total number of impulses. The value of  $a_N \in \mathbf{A}$  is immaterial because it is never in use (is to be applied at  $T_N = \infty$ ). For this control strategy  $\pi$ , the occupation measure  $\mu^\pi$ , restricted to  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  (equivalently, to  $\mathbf{M} \times \mathbf{A}$ ), i.e., ignoring the absorbing state  $\Delta$ , is the following combination of Dirac measures:

$$\mu^\pi(\Gamma_1 \times \Gamma_2) = \sum_{i=1}^N \delta_{(\tilde{x}_{i-1}, 0)}(\Gamma_1) \delta_{(\theta_i, a_i)}(\Gamma_2) \quad \forall \Gamma_1 \in \mathcal{B}(V), \Gamma_2 \in \mathcal{B}(\bar{\mathbb{R}}_+^0 \times \mathbf{A}).$$

For  $x = (\tilde{x}, 0) \in \mathbf{X}$  and  $(\theta, a) \in \mathbf{B}$  (no other values of  $x$  appear in the expressions below),

$$\bar{C}(x, (\theta, a)) = \int_{[0, \theta]} C^g(\tilde{x} e^u) du + \mathbb{I}\{\theta < +\infty\} a,$$

and the objective corresponding to  $\mu = \mu^\pi$  in the linear program (10),(11) (i.e., the performance of  $\pi$

in problem (5)) has the form

$$\begin{aligned}\mathcal{V}(x_0, \pi) &= \mathcal{V}((\tilde{x}_0, 0), \pi) = \int_{\mathbf{M} \times \mathbf{A}} \bar{C}(x, (\theta, a)) \mu^\pi(dx \times d\theta \times da) \\ &= \sum_{i=1}^{N-1} \left( \int_{[0, \theta_i]} C^g(\tilde{x}_{i-1} e^u) du + a_i \right) + \int_{[0, \infty)} C^g(\tilde{x}_{N-1} e^u) du = \int_0^\infty C^g(\tilde{x}(u)) du + \int_0^\infty dW(u),\end{aligned}$$

where  $\tilde{x}(u)$  is as defined in (23). Since

$$Q(dx|y, (\theta, a)) = \delta_{(\tilde{l}(\tilde{\phi}(\tilde{y}, \theta), a), 0)}(dx) = \delta_{(\tilde{y}e^\theta + a, 0)}(dx)$$

for  $y = (\tilde{y}, 0)$  and  $\theta \neq +\infty$ , the characteristic equation (11) on  $\mathcal{B}(V)$  for  $\mu = \mu^\pi$  reads

$$\begin{aligned}\mu^\pi(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) &= \delta_{(\tilde{x}_0, 0)}(dx) + \int_{\mathbf{M} \times \mathbf{A}} \delta_{(\tilde{y}e^\theta + a, 0)}(dx) \mu^\pi(d(\tilde{y}, 0) \times d\theta \times da) \\ &= \delta_{(\tilde{x}_0, 0)}(dx) + \sum_{i=1}^{N-1} \delta_{(\tilde{x}_{i-1} e^{\theta_i} + a_i, 0)}(dx)\end{aligned}$$

and is obviously satisfied because  $\tilde{x}_{i-1} e^{\theta_i} + a_i = \tilde{x}_i$ , for all  $i = 1, 2, \dots, N-1$ .

The first part of the aggregated occupation measure on  $V$ , induced by  $\mu^\pi$ , looks as follows:

$$\begin{aligned}\eta^\pi(dy \times \square) &= \int_{[0, \infty)} \left\{ \int_{V \times \mathbf{A}} \delta_{\phi(x, u)}(dy) \mu^\pi(dx \times [u, \infty) \times da) \right\} du \\ &= \sum_{i=1}^N \int_{[0, \infty)} \delta_{(\tilde{x}_{i-1} e^u, u)}(dy) \mathbb{I}\{\theta_i \geq u\} du.\end{aligned}$$

(See (14).) The  $t$ -component of the state  $y = (\tilde{y}, t)$  is of no importance, and the marginal  $\tilde{\eta}^\pi(d\tilde{y} \times \square)$  on  $\tilde{V} = [\tilde{x}_0, K)$ , as at the end of Section 4, has the form

$$\tilde{\eta}^\pi(d\tilde{y} \times \square) = \sum_{i=1}^N \int_{[0, \theta_i]} \delta_{\tilde{x}_{i-1} e^u}(d\tilde{y}) du.$$

Note that this measure is non-atomic. It is normal in the sense of (22):  $L(d\tilde{x}) := \sum_{i=1}^N \delta_{\tilde{x}_{i-1}}(d\tilde{x})$  and  $g(\tilde{x}, u) := \mathbb{I}\{u \leq \theta_i\}$ . Recall that  $\theta_i < \tilde{\theta}^*(\tilde{x}_{i-1})$  for  $i < N$  and, for  $i = N$ , when  $\theta_N = \infty$ , the integration with respect to  $u$  is in fact up to  $\tilde{\theta}^*(\tilde{x}_{N-1})$ , because the measure  $\tilde{\eta}^\pi(d\tilde{y} \times \square)$  is defined on  $\tilde{V} = [\tilde{x}_0, K)$  and, for  $\Gamma \in \mathcal{B}(\tilde{V})$ ,  $\delta_{(\tilde{x}_{N-1} e^u)}(\Gamma) = 0$  if  $u \geq \tilde{\theta}^*(\tilde{x}_{N-1}) = \ln \frac{K}{\tilde{x}_{N-1}}$ . For each interval  $[\tilde{y}_1, \tilde{y}_2) \subset \tilde{V} = [\tilde{x}_0, K)$  on which the function  $\tilde{x}(\cdot)$  is continuous, the value  $\tilde{\eta}^\pi([\tilde{y}_1, \tilde{y}_2) \times \square)$  equals  $\ln \frac{\tilde{y}_2}{\tilde{y}_1} = t_2 - t_1$ , the time for the function  $\tilde{x}(\cdot)$  to change from  $\tilde{y}_1 = \tilde{x}(t_1)$  to  $\tilde{y}_2 = \tilde{x}(t_2)$  according to the flow  $\tilde{\phi}$ . To put it different, on the intervals  $[\tilde{x}_{i-1}, \tilde{x}_{i-1} e^{\theta_i}) \cap \tilde{V}$ ,  $i = 1, 2, \dots, N$ , the measure  $\tilde{\eta}^\pi(d\tilde{y} \times \square)$  is the image of the Lebesgue measure  $ds$  with respect to the mapping  $u \rightarrow \tilde{\phi}(\tilde{x}_{i-1}, u) = \tilde{x}_{i-1} e^u$ . Hence, for a bounded absolutely continuous function  $w(\cdot)$  on  $[\tilde{x}_{i-1}, \tilde{x}_{i-1} e^{\theta_i}) \cap \tilde{V}$ , for  $\tilde{y}_1 = \tilde{x}_{i-1} e^{t_1}$ ,  $\tilde{y}_2 = \tilde{x}_{i-1} e^{t_2}$ ,

$$w(\tilde{y}_2-) - w(\tilde{y}_1) = \int_{[0, t_2 - t_1]} \chi w(\tilde{\phi}(\tilde{y}_1, s)) ds = \int_{[\tilde{y}_1, \tilde{y}_2)} \chi w(\tilde{y}) \tilde{\eta}^\pi(d\tilde{y} \times \square). \quad (24)$$

For the first term of the objective in (23), we have:

$$\int_{[\tilde{x}_0, K)} C^g(\tilde{y}) \tilde{\eta}^\pi(d\tilde{y} \times \square) = \sum_{i=1}^N \int_{[0, \theta_i]} C^g(\tilde{x}_{i-1} e^u) du = \int_0^\infty C^g(\tilde{x}(u)) du,$$

where  $\tilde{x}(\cdot)$  is as defined in (23). Recall that  $C^g(\tilde{x}) = 0$  for  $\tilde{x} \geq K$ .

The second part of the aggregated occupation measure on  $V \times \mathbf{A}$ , induced by  $\mu^\pi$ , has the form

$$\begin{aligned} \eta^\pi(dy \times da) &= \int_V \int_{\mathbb{R}_+^0} \delta_{\phi(x,\theta)}(dy) \mathbb{I}\{\phi(x,\theta) \in V\} \mu^\pi(dx \times d\theta \times da) \\ &= \sum_{i=1}^N \delta_{(\tilde{x}_{i-1}e^{\theta_i}, \theta_i)}(dy) \mathbb{I}\left\{\tilde{x}_{i-1}e^{\theta_i} \in [\tilde{x}_0, K), \theta_i \in \left[0, \ln \frac{\tilde{x}_{i-1}e^{\theta_i}}{\tilde{x}_0}\right]\right\} \delta_{a_i}(da). \end{aligned}$$

(See (15).) Clearly, this measure is finite. Like previously, the  $t$ -component of the state  $y = (\tilde{y}, t)$  is of no importance, and the marginal  $\tilde{\eta}(d\tilde{y} \times da)$  on  $\tilde{V} \times \mathbf{A}$ , as at the end of Section 4, has the form

$$\tilde{\eta}^\pi(d\tilde{y} \times da) = \sum_{i=1}^{N-1} \delta_{\tilde{x}_{i-1}e^{\theta_i}}(d\tilde{y}) \delta_{a_i}(da). \quad (25)$$

Recall that  $\theta_N = \infty$ ,  $\tilde{x}_{i-1}e^{\theta_i} < K$  and  $\tilde{x}_{i-1} \geq \tilde{x}_0$  for  $i < N$ .

For the second term of the objective in (23), we have:

$$\int_{[\tilde{x}_0, K) \times \mathbf{A}} a \tilde{\eta}^\pi(d\tilde{y} \times da) = \sum_{i=1}^{N-1} a_i = \int_0^\infty dW(u),$$

where the function  $W(u) = \sum_{j=1}^N \left(\sum_{i=1}^{j-1} a_i\right) \mathbb{I}\{T_{j-1} \leq u < T_j\}$  (with  $T_0 = 0$ ,  $T_i = T_{i-1} + \theta_i$ ,  $i = 1, 2, \dots, N$ ) represents the control strategy  $\pi$ .

Let us illustrate the definitions introduced in Section 4 (see Figure 2):

$$h(y) = h(\tilde{y}, t) = \frac{\tilde{y}}{e^t}; \quad F(\tilde{x}^0, t) = (\tilde{x}^0 e^t, t); \quad F^{-1}(y) = F^{-1}(\tilde{y}, t) = \left(\frac{\tilde{y}}{e^t}, t\right);$$

$$\mathbf{D} = \{(\tilde{x}^0, t) : \tilde{x}^0 \in [0, K), 0 \leq t < \ln \frac{K}{\tilde{x}^0}\}.$$

With some abuse of notations, we omit the double brackets in the expressions like  $h(y) = h((\tilde{y}, t))$ . The mappings  $F$  and  $F^{-1}$  are one-to-one and continuous. According to Definition 4.4, a measure  $\zeta$  on  $V$  is normal if and only if there is a finite measure  $L$  on  $\tilde{V} = [\tilde{x}_0, K)$  such that the conditional distribution  $\check{\zeta}(dt|\tilde{x}^0)$  is ( $L$ -almost surely) absolutely continuous with respect to the Lebesgue measure.

According to the last paragraphs in Section 4, the linear program (20),(21) is formulated in terms of the (normal) marginal measures  $\tilde{\eta}(d\tilde{x} \times da)$  on  $\tilde{V} \times \mathbf{A}_\square$ , ignoring the  $t$  component, time elapsed since the most recent impulse. The unnecessary ‘tilde’ is omitted up to the end of this section, apart from  $\tilde{x}_0$  (the initial state).

$$\left. \begin{array}{l} \text{Minimize over the normal} \\ \text{measures } \eta \text{ on } [\tilde{x}_0, K) \times \mathbf{A}_\square : \int_{[\tilde{x}_0, K)} C^g(x) \eta(dx \times \square) + \int_{[\tilde{x}_0, K) \times \mathbf{A}} a \eta(dx \times da) \\ \text{subject to : } w(\tilde{x}_0) + \int_{[\tilde{x}_0, K)} \chi w(x) \eta(dx \times \square) \\ \quad + \int_{[\tilde{x}_0, K) \times \mathbf{A}} [w(x+a) - w(x)] \eta(dx \times da) = 0. \end{array} \right\} \quad (26)$$

The test functions  $w$  on  $[\tilde{x}_0, \infty)$  are bounded, measurable, absolutely continuous, either negative and increasing or positive and decreasing, and such that  $w(x) = 0$  on  $[K, \infty)$ . The measures  $\eta(dx \times \square)$

are finite on  $[\tilde{x}_0, K)$ : in the case of a general function  $G(\cdot) > \delta > 0$ , one can substitute the bounded function

$$w(x) := \int_{\tilde{x}_0}^x \frac{1}{G(y)} dy - \int_{\tilde{x}_0}^K \frac{1}{G(y)} dy$$

with  $\chi w(x) = 1$  for  $x \in [\tilde{x}_0, K)$ . The measures  $\eta(dx \times \mathbf{A})$  are finite due to the definition of a normal measure.

For the special case presented above and the measure  $\eta^\pi$ , the characteristic equation in (26) is just the version of the Newton-Leibniz formula:

$$\begin{aligned} \int_{[\tilde{x}_0, K)} \chi w(x) \eta^\pi(dx \times \square) &= [w(K-) - w(x(T_{N-1}+))] + [w(x(T_{N-1}-)) - w(x(T_{N-2}+))] \\ &+ \dots + [w(x(T_1-)) - w(\tilde{x}_0+)] \mathbb{I}\{T_1 > 0\} \end{aligned}$$

by (24), and

$$\begin{aligned} \int_{[\tilde{x}_0, K) \times \mathbf{A}} [w(x+a) - w(x)] \eta^\pi(dx \times da) &= [w(x(T_{N-1}+)) - w(x(T_{N-1}-))] + [w(x(T_{N-2}+)) \\ &- w(x(T_{N-2}-))] + \dots + [w(x(T_1+)) - w(x(T_1-))]. \end{aligned}$$

Recall that  $w(K-) = 0$  and, if  $T_1 = 0$ , then  $w(x(T_1-)) = w(\tilde{x}_0)$ .

The marginal of the occupation measure  $\mu^\pi$  on  $[\tilde{x}_0, K) \times \mathbb{R}_+^0 \times \mathbf{A}$  has the ‘higher dimensionality’ than the induced aggregated measure  $\eta^\pi$  on  $[\tilde{x}_0, K) \times (\mathbf{A} \cup \{\square\})$  (cf. Remark 4.2). One can also compare the representation of the first part of the objective (23) in terms of  $\mu^\pi$  and  $\eta^\pi$ . In the case of  $\mu^\pi$ , the integration with respect to time is hidden in the function  $\bar{C}(\cdot)$  and, in the case of  $\eta^\pi$ , such an integration appears just in the definition of  $\eta^\pi(dy \times \square)$ . As a result, the objective  $\int_0^\infty C^g(x(u)) du$  is represented as  $\int_{[\tilde{x}_0, K)} C^g(y) \eta^\pi(dy \times \square)$ , in terms of the original cost  $C^g(\cdot)$ .

Deterministic control strategies  $\pi$  (equivalently, deterministic functions  $W(\cdot)$ ) are sufficient for solving unconstrained problems like (23). In the constrained case, when  $J > 0$ , randomization and mixture (e.g., application of the control functions  $W_1(\cdot)$  and  $W_2(\cdot)$  with probabilities  $\alpha \in (0, 1)$  and  $1 - \alpha$ ) usually help to improve the objective.

It is interesting to compare the linear program (26) with the linear programs which appeared in [10, 11, 18]. In those articles, the impulse control problem was formulated on the finite time horizon  $[0, T]$ , but the constructions can be formally adjusted for  $[0, \infty)$ . In what follows, the measurable and sufficiently smooth function  $G(\cdot) > \delta \geq 0$  is arbitrarily fixed.

Following the ideas of [10], the problem (23) is replaced with the following linear program on the space of the so called occupation measures  $\Upsilon_1^1$  and  $\Upsilon_2^1$ :

$$\left. \begin{aligned} &\text{Minimize over the finite} \\ &\text{measures } \Upsilon_1^1 \text{ and } \Upsilon_2^1 \text{ on } [\tilde{x}_0, \infty) : \int_{[\tilde{x}_0, \infty)} C^g(x) \Upsilon_1^1(dx) + \int_{[\tilde{x}_0, \infty)} \Upsilon_2^1(dx) \\ &\text{subject to : } w(\tilde{x}_0) + \int_{[\tilde{x}_0, \infty)} \frac{dw(x)}{dx} G(x) \Upsilon_1^1(dx) \\ &\quad + \int_{[\tilde{x}_0, \infty)} \frac{dw(x)}{dx} \Upsilon_2^1(dx) = 0, \end{aligned} \right\} \quad (27)$$

where the test functions  $w$  are continuously differentiable on  $[\tilde{x}_0, \infty)$  and  $\lim_{x \rightarrow \infty} w(x) = 0$ .

Consider the test functions  $w$  as in (26), which are continuously differentiable on  $[\tilde{x}_0, \infty)$ . Now  $\chi w(x) = \frac{dw}{dx}G(x)$  and, for the measures

$$\Upsilon_1^1(dx) := \begin{cases} \eta(dx \times \square), & \text{if } dx \in \mathcal{B}([\tilde{x}_0, K]); \\ \text{arbitrary,} & \text{otherwise,} \end{cases}$$

$$\text{and } \Upsilon_2^1(dz) := \left[ \int_{[\tilde{x}_0, K) \times \mathbf{A}} \mathbb{I}\{x < z < x + a\} \eta(dx \times da) \right] dz,$$

all the expressions in (27), take the form of those in (26) because

$$\begin{aligned} \int_{[\tilde{x}_0, \infty)} \Upsilon_2^1(dz) &= \int_{[\tilde{x}_0, K) \times \mathbf{A}} \int_{[\tilde{x}_0, \infty)} \mathbb{I}\{x < z < x + a\} dz \eta(dx \times da) \\ &= \int_{[\tilde{x}_0, K) \times \mathbf{A}} a \eta(dx \times da); \\ \int_{[\tilde{x}_0, \infty)} \frac{dw(z)}{dz} \Upsilon_2^1(dz) &= \int_{[\tilde{x}_0, K) \times \mathbf{A}} \int_{[\tilde{x}_0, \infty)} \frac{dw(z)}{dz} \mathbb{I}\{x < z < x + a\} dz \eta(dx \times da) \\ &= \int_{[\tilde{x}_0, K) \times \mathbf{A}} [w(x + a) - w(x)] \eta(dx \times da). \end{aligned}$$

Recall also that  $C^g(x) = \frac{dw(x)}{dx} = 0$  for  $x \in [K, \infty)$ .

**Remark 5.1** We see that the measure  $\Upsilon_1^1(dx)$  coincides with  $\eta(dx \times \square)$  on  $\tilde{V}$ , and  $\Upsilon_2^1(dz)$  can be called the “further aggregated version” of  $\eta(dx \times da)$ : the argument “ $da$ ” disappears. This second step of aggregating is reasonable only in this specific example, where  $l(x, a) = x + a$  and  $C^I(x, a) = a$ . For other functions  $l(\cdot)$  and  $C^I(\cdot)$  it makes no sense. As was shown, having in hand the measures as in (26), one can construct the corresponding measures  $\Upsilon_1^1$  and  $\Upsilon_2^1$  as in (27). On the other hand, e.g., when dealing with the measures associated with the control strategies  $\pi$  as in the special case presented above,  $\eta^\pi(dx \times \square) = \Upsilon_1^1(dx)$  and one can reconstruct  $\eta^\pi(dx \times da)$  having in hand the corresponding measure  $\Upsilon_2^1(dz)$ . Indeed, the latter measure coincides with the Lebesgue measure on the intervals  $(x(T_i -) = x_{i-1}e^{\theta_i}, x(T_i +) = x(T_i -) + a_i)$  and equals zero outside. Knowing these intervals, one can recalculate the measure  $\eta^\pi(dx \times da)$ : see (25).

In the works [11, 18], the impulse control problem (23) is formulated in a different way which is briefly presented below. The generic notations of [11, 18] are changed to avoid the confusion with the notations in the present paper. Let a reasonable deterministic stationary control strategy, defined by  $\{T_j, a_j\}_{j=1}^\infty$  and denoted below as  $f : \mathbf{X}_\Delta \rightarrow \mathbb{R}_+^0 \times \mathbf{A}$ , be fixed, such that  $T_j = \infty$  if  $x(T_j -) \geq K$ . By the way, the number of finite moments  $T_j$  is finite, and the class of such strategies is sufficient in the unconstrained problem (23) by Theorem 1 in [29]. Introduce the measure

$$\sigma(dt) := dt + \sum_{j=1}^{\infty} \mathbb{I}\{T_j < \infty\} \delta_{T_j}(dt)$$

on the time scale  $[0, \infty)$ . The model (23) is represented as

$$\left. \begin{aligned} dx(t) &= \int_{\mathbf{A} \cup \{0\}} \frac{F(x(t), a)}{1 + |a|^p} \kappa(da|t) \sigma(dt); \quad x(0-) = \tilde{x}_0; \\ &\int_{[0, \infty)} \int_{\mathbf{A} \cup \{0\}} \frac{L(x(t), a)}{1 + |a|^p} \kappa(da|t) \sigma(dt) \rightarrow \inf_{\sigma, \kappa} \end{aligned} \right\} \quad (28)$$

with the following system primitives:

- $p \geq 1$  is some fixed natural number.
- $\kappa(da|t) = \delta_0(da)$  if  $t$  is different from all  $T_j$ , so that  $F(x, 0) = G(x)$ , and  $a = 0$  corresponds to the absence of impulses.
- $\kappa(da|T_j) = \delta_{a_j}(da)$  for  $T_j < \infty$ , and at that time moment the following fictitious process is introduced:

$$\frac{dy_{T_j}^{a_j}(u)}{du} = \frac{F(y_{T_j}^{a_j}(u), a_j)}{1 + |a_j|^p}, \quad y_{T_j}^{a_j}(0) = x(T_j-), \quad u \in [0, 1].$$

The form of the function  $F$  is seen in the next item.

- $x(T_j+) = x(T_j-) + [y_{T_j}^{a_j}(1) - y_{T_j}^{a_j}(0)] = y_{T_j}^{a_j}(1)$ . To be consistent with the model (23), we should have  $x(T_j+) = x(T_j-) + a_j$ , so that for  $a \in \mathbf{A}$  we put  $F(y, a) := a(1 + |a|^p)$ .
- Similarly, for consistency, we put  $L(x, 0) = C^g(x)$  and  $L(y, a) := a(1 + |a|^p)$  for  $a \in \mathbf{A}$ .

The occupation measure on  $[0, \infty) \times [\tilde{x}_0, \infty) \times [\mathbf{A} \cup \{0\}]$  as in [11, 18], corresponding to the strategy  $f$  (equivalently, to the pair  $(\sigma, \kappa)$ ), equals

$$\Upsilon^f(dt \times dy \times da) := \Phi(dy|a, t)\kappa(da|t)\sigma(dt) = \Upsilon_1^f(dt \times dy)\delta_0(da) + \Upsilon_2^f(dt \times dy \times da),$$

where

$$\Phi(dy|a, t) := \begin{cases} \delta_{x(t)}(dy), & \text{if } t \neq T_j \text{ for all } j = 1, 2, \dots; \\ \int_0^1 \delta_{y_t^a(u)}(dy)du = \int_0^1 \delta_{x(T_j-) + au}(dy)du, & \text{if } t = T_j; \end{cases}$$

$x(\cdot)$  is the trajectory of the system (28) (equivalently, of the system (23)) under the strategy  $f$ . The presentation  $\Upsilon^f = \Upsilon_1^f + \Upsilon_2^f$  corresponds to the decomposition of the measure  $\sigma$  to the absolutely continuous and discrete parts. Different control strategies  $f$  as above, that is, different pairs  $(\sigma, \kappa)$  define all different measures  $\Upsilon$  under consideration, which are denoted below as  $\Upsilon^2$ .

Below, the test functions  $w$  are as in (26) and continuously differentiable on  $[\tilde{x}_0, \infty)$ . In the linear program for the problem (28), suggested in [11, 18], all the integrated functions do not depend on time  $t$ . Thus, we immediately introduce the marginals  $\hat{\Upsilon}^2(dy \times da) := \int_{[0, \infty)} \Upsilon^2(dt \times dy \times da)$ :

$$\begin{aligned} \hat{\Upsilon}_1^2(dy) &:= \int_{[0, \infty)} \Upsilon_1^2(dt \times dy) = \int_{[0, \infty)} \delta_{x(t)}(dy) dt; \\ \hat{\Upsilon}_2^2(dy \times da) &:= \int_{[0, \infty)} \Upsilon_2^2(dt \times dy \times da) = \sum_{T_j} \mathbb{I}\{T_j < \infty\} \left[ \int_0^1 \delta_{x(T_j-) + au}(dy) du \right] \delta_{a_j}(da). \end{aligned}$$

Here the measure  $\Upsilon^2 = \Upsilon_1^2 + \Upsilon_2^2$  comes from the pair  $(\sigma, \kappa)$ , which also defines the trajectory  $x(\cdot)$  of the system (28); the measure  $\hat{\Upsilon}_2^2$  is finite and  $\hat{\Upsilon}_2^2([K + H, \infty) \times \mathbf{A}) = 0$  because  $x(T_j-) < K$  and  $a_j \leq H$ . The linear program as in [11, 18] has the form

$$\left. \begin{aligned} &\text{Minimize over} \\ &\text{the measures } \hat{\Upsilon}^2 \text{ on } [\tilde{x}_0, \infty) \times \mathbf{A} : \int_{[\tilde{x}_0, \infty) \times \mathbf{A}} \frac{L(y, a)}{1 + |a|^p} \hat{\Upsilon}^2(dy \times da) \\ &\text{subject to} : w(\tilde{x}_0) + \int_{[\tilde{x}_0, \infty) \times \mathbf{A}} \frac{dw(y)}{dy} \frac{F(y, a)}{1 + |a|^p} \hat{\Upsilon}^2(dy \times da) = 0 \end{aligned} \right\}$$

(note that  $\lim_{y \rightarrow \infty} w(y) = 0$ ), or, more explicitly,

$$\left. \begin{array}{l} \text{Minimize over} \\ \text{the finite measures } \hat{\Upsilon}_1^2 \text{ and } \hat{\Upsilon}_2^2 : \int_{[\tilde{x}_0, K]} C^g(y) \hat{\Upsilon}_1^2(dy) + \int_{[\tilde{x}_0, \infty) \times \mathbf{A}} a \hat{\Upsilon}_2^2(dy \times da) \\ \text{subject to : } w(\tilde{x}_0) + \int_{[\tilde{x}_0, K]} \frac{dw(y)}{dy} G(y) \hat{\Upsilon}_1^2(dy) \\ \quad + \int_{[\tilde{x}_0, \infty) \times \mathbf{A}} \frac{dw(y)}{dy} a \hat{\Upsilon}_2^2(dy \times da) = 0. \end{array} \right\} \quad (29)$$

We underline that the measure  $\hat{\Upsilon}_1^2$  is of no importance on  $[K, \infty)$  because there  $C^g(y) = \frac{dw(y)}{dy} = 0$ ; it is finite on  $[\tilde{x}_0, K)$  because  $\int_{[0, \infty)} \delta_{x(t)}([\tilde{x}_0, K)) dt \leq \theta^*(\tilde{x}_0) < \infty$ .

The measures  $\hat{\Upsilon}_1^2$  and  $\hat{\Upsilon}_2^2$  can be calculated based on the measures  $\eta$  in (26), so that all the expressions in (29) become equal to those in (26). Indeed, we put  $\hat{\Upsilon}_1^2(dx) := \Upsilon_1^1(dx) = \eta(dx \times \square)$  on  $\tilde{V} = [\tilde{x}_0, K)$  and

$$\hat{\Upsilon}_2^2(dy \times da) := \frac{1}{a} \left[ \int_{[\tilde{x}_0, K)} \mathbb{I}\{x < y < x + a\} \eta(dx \times da) \right] dy \text{ on } [\tilde{x}_0, \infty).$$

Now  $\int_{\mathbf{A}} a \hat{\Upsilon}_2^2(dy \times da) = \Upsilon_2^1(dy)$  and all the expressions in (29) coincide with those in (27) and, as shown above, are equal to those in (26). For the measures  $\eta^\pi$  as in the special case presented above, they are in a one-to-one correspondence with the measures  $\hat{\Upsilon}_1^2$  and  $\hat{\Upsilon}_2^2$ : see Remark 5.1.

## 6 Proof of Theorem 4.1

Proof of Theorem 4.1. Note that, for each function  $w \in \mathbf{W}$ , for each fixed  $x \in V$ , the function  $w(\phi(x, \cdot))$  is bounded on  $\mathbb{R}_+^0$ .

According to Lemma A.1, for each fixed  $x \in V$ ,

$$w(\phi(x, \theta)) = w(x) + \int_{[0, \theta]} \chi w(\phi(x, s)) ds,$$

where the function  $\chi w$  is given by (47). After we integrate this equation over  $V \times \mathbb{R}_+^0$  with respect to the measure

$$\int_{\mathbf{A}} \mathbb{I}\{\phi(x, \theta) \in V\} p_T(d\theta|x, a) p_A(da|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}),$$

on  $V \times \mathbb{R}_+^0$ , where the stochastic kernels  $p_T$  and  $p_A$  are as in (13), we obtain the equality

$$\begin{aligned} \int_V w(y) \eta(dy \times \mathbf{A}) &= \int_V \int_{\mathbb{R}_+^0} w(\phi(x, \theta)) \mathbb{I}\{\phi(x, \theta) \in V\} \hat{p}(d\theta|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &= \int_V w(x) \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi(x, \theta) \in V\} \hat{p}(d\theta|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &\quad + \int_V \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi(x, \theta) \in V\} \int_{[0, \theta]} \chi w(\phi(x, s)) ds \hat{p}(d\theta|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}), \end{aligned}$$

where  $\hat{p}(d\theta|x) := \int_{\mathbf{A}} p_T(d\theta|x, a) p_A(da|x)$ . Note that all the integrals here are finite because the function  $w(\cdot)$  is bounded and the measures  $\mu$  and  $\eta(dy \times \mathbf{A})$  are finite. For each  $x \in V$ , let us denote

$$\theta^*(x) := \inf\{\theta \in \mathbb{R}_+^0 : \phi(x, \theta) \in V^c\}.$$

As usual,  $\inf \emptyset := +\infty$ . Since the flow  $\phi$  is continuous, the function  $\theta^*(\cdot)$  is measurable: see [13, Lemma 27.1] or [16, Prop.1.5, p.154]. Besides,  $\theta^*(x) > 0$  because the set  $V$  is open and the set  $V^c$  is closed.

Since the set  $V^c$  is closed and the flow  $\phi$  is continuous, in case  $\theta^*(x) < +\infty$ ,  $\phi(x, \theta^*(x)) \in V^c \cap \mathbf{X}$  and the infimum is attained. Moreover, as mentioned above Definition 3.1,  $\phi(x, s) \in V^c$  for all  $s \geq \theta^*(x)$ . Therefore,

$$\begin{aligned} \int_V w(y) \eta(dy \times \mathbf{A}) &= \int_V w(x) \hat{p}(\bar{\mathbb{R}}_+^0|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &\quad - \int_V w(x) \hat{p}([\theta^*(x), \infty]|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &\quad + \int_V \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi(x, \theta) \in V\} \int_{[0, \theta]} \chi w(\phi(x, s)) ds \hat{p}(d\theta|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}). \end{aligned}$$

Recall, the measure  $\eta(dy \times \mathbf{A})$  is finite and the function  $\chi w(\phi(x, s))$  is integrable on  $[0, \theta]$  with  $\theta < \infty$ .

After we apply the Tonelli Theorem [1, Thm.11.28] to the last term, we obtain:

$$\begin{aligned} \int_V w(y) \eta(dy \times \mathbf{A}) &= \int_V w(x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) - \int_V w(x) \hat{p}([\theta^*(x), \infty]|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &\quad + \int_V \int_{\mathbb{R}_+^0} \int_{[s, \infty)} \mathbb{I}\{\phi(x, \theta) \in V\} \chi w(\phi(x, s)) \hat{p}(d\theta|x) ds \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &= \int_V w(x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) - \int_V w(x) \hat{p}([\theta^*(x), \infty]|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &\quad + \int_V \int_{\mathbb{R}_+^0} \chi w(\phi(x, s)) \mathbb{I}\{\phi(x, s) \in V\} \hat{p}([s, \theta^*(x)]|x) ds \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}). \end{aligned}$$

Note that

$$\begin{aligned} &\int_V \int_{\mathbb{R}_+^0} \chi w(\phi(x, s)) \mathbb{I}\{\phi(x, s) \in V\} \hat{p}([s, \theta^*(x)]|x) ds \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &= \int_V \int_{\mathbb{R}_+^0} \chi w(\phi(x, s)) \hat{p}([s, \theta^*(x)]|x) ds \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \end{aligned}$$

as  $\chi w(\phi(x, s)) = 0$  for  $\phi(x, s) \in V^c$ . (See (47), where, in our case,  $V^c \subseteq D$  and  $W(y) = 0$  for all  $y \in V^c$ .) Now

$$\begin{aligned} &\int_V w(y) \eta(dy \times \mathbf{A}) \\ &= \int_V w(x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) - \int_V w(x) \hat{p}([\theta^*(x), \infty]|x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) \\ &\quad + \int_V \chi w(y) \eta(dy \times \square) - \int_V \int_{\mathbb{R}_+^0} \chi w(\phi(x, s)) \hat{p}([\theta^*(x), \infty]|x) ds \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}). \end{aligned}$$

All the integrals here are finite because, no matter whether  $\theta^*(x)$  is finite or not,

$$\lim_{t \rightarrow \infty} w(\phi(x, t)) = w(x) + \int_{\mathbb{R}_+^0} \chi w(\phi(x, s)) ds = 0, \quad \forall x \in V$$



and thus

$$\int_V w(x) \hat{p}([\theta^*(x), \infty] | x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) + \int_V \int_{\mathbb{R}_+^0} \chi w(\phi(x, s)) \hat{p}([\theta^*(x), \infty] | x) ds \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) = 0.$$

This also leads to

$$\begin{aligned} \int_V w(y) \eta(dy \times \mathbf{A}) &= \int_V w(x) \mu(dx \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}) + \int_V \chi w(y) \eta(dy \times \square) \\ &= w(x_0) + \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} w(l(\phi(y, \theta), a)) \mathbb{I}\{\phi(y, \theta) \in V\} \mu(dy \times d\theta \times da) \\ &\quad + \int_V \chi w(y) \eta(dy \times \square) \end{aligned}$$

by (11), and the required formula (21) follows from the definition (15).  $\square$

## 7 Proof of Theorem 4.2 and Corollary 4.1

Below, we assume that Conditions 2.1, 3.1, 3.2 and 4.2 are satisfied. The proofs will be based on a series of lemmas.

**Lemma 7.1** *Let  $\pi = \{\pi_i\}_{i=1}^\infty$  be a reasonable Markov strategy as in Definition 3.1, defined on  $V$  by stochastic kernels  $\pi_i(d\theta \times da | x) = p_T^i(d\theta | x) p_A^i(da | x, \theta)$ . Suppose  $\eta$  is the corresponding aggregated occupation measure (17) coming from the occupation measure  $\mu^\pi$  as in (8). Introduce the (partial) aggregated occupation measures*

$$\eta^i(\Gamma_X \times \Gamma_A) := \eta^i(\Gamma_X \times (\Gamma_A \cap \mathbf{A})) + \eta^i(\Gamma_X \times \square) \mathbb{I}\{\square \in \Gamma_A\}$$

on  $V \times \mathbf{A}_\square$ , defined recursively:

$$\begin{aligned} \eta^0(\Gamma_X \times \Gamma_A) &\equiv 0; \\ \eta^{i+1}(\Gamma_X \times \square) &= \eta^i(\Gamma_X \times \square) + \int_V \int_{\bar{\mathbb{R}}_+^0} \delta_{\phi(x, u)}(\Gamma_X) p_T^{i+1}([u, \infty] | x) du \nu^i(dx), \quad \Gamma_X \in \mathcal{B}(V); \\ \eta^{i+1}(\Gamma_X \times \Gamma_A) &= \eta^i(\Gamma_X \times \Gamma_A) + \int_V \int_{\bar{\mathbb{R}}_+^0} \delta_{\phi(x, \theta)}(\Gamma_X) p_A^{i+1}(\Gamma_A | x, \theta) p_T^{i+1}(d\theta | x) \nu^i(dx), \\ &\quad \Gamma_X \in \mathcal{B}(V), \Gamma_A \in \mathcal{B}(\mathbf{A}), \end{aligned}$$

where  $\nu^i(dx) = P_{x_0}^\pi(X_i \in dx)$  is the measure on  $V$ ,  $i \geq 0$ .

Then  $\eta^i \uparrow \eta$  on  $V \times \mathbf{A}_\square$  set-wise as  $i \rightarrow \infty$ . Every measure  $\eta^i$  is normal.

Proof. We will need the (partial) occupation measure on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$

$$\mu^n(dx \times d\theta \times da) := E_{x_0}^\pi \left[ \sum_{i=1}^n \mathbb{I}\{X_{i-1} \in dx, \Theta_i \in d\theta, A_i \in da\} \right], \quad n = 0, 1, 2, \dots$$

Clearly,  $\mu^n \uparrow \mu^\pi$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  set-wise as  $n \rightarrow \infty$ . Therefore, according to the definition of the measure  $\eta$ , for each positive measurable function  $C^g$  on  $V$ ,

$$\begin{aligned} I^n &:= \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \int_{[0, \theta]} C^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} du \right\} \mu^n(dx \times d\theta \times da) \\ &\uparrow \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \left\{ \int_{[0, \theta]} C^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} du \right\} \mu^\pi(dx \times d\theta \times da) = \int_V C^g(y) \eta(dy \times \square) \end{aligned}$$

and, for each positive measurable function  $C^I$  on  $V \times \mathbf{A}$ ,

$$\begin{aligned} J^n &:= \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \mathbb{I}\{\theta < +\infty\} \mathbb{I}\{\phi(x, \theta) \in V\} C^I(\phi(x, \theta), a) \mu^n(dx \times d\theta \times da) \\ &\uparrow \int_{V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}} \mathbb{I}\{\theta < +\infty\} \mathbb{I}\{\phi(x, \theta) \in V\} C^I(\phi(x, \theta), a) \mu^\pi(dx \times d\theta \times da) \\ &= \int_{V \times \mathbf{A}} C^I(y, a) \eta(dy \times da). \end{aligned}$$

We will prove by induction the following assertions:

$$I^n = \int_V C^g(y) \eta^n(dy \times \square) \quad \text{and} \quad J^n = \int_{V \times \mathbf{A}} C^I(y, a) \eta^n(dy \times da).$$

If  $n = 0$ , then  $\mu^0 = 0$ ,  $\eta^0 = 0$ ,  $I^0 = 0$ , and  $J^0 = 0$ .

Suppose the above assertions are valid for some  $n \geq 0$ . Then

$$\begin{aligned} I^{n+1} &= I^n + \int_V \int_{\bar{\mathbb{R}}_+^0} \int_{\mathbf{A}} \left\{ \int_{[0, \theta]} C^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} du \right\} p_A^{n+1}(da|x, \theta) p_T^{n+1}(d\theta|x) \nu^n(dx) \\ &\quad \text{and} \\ J^{n+1} &= J^n + \int_V \int_{\bar{\mathbb{R}}_+^0} \int_{\mathbf{A}} \mathbb{I}\{\theta < \infty\} \mathbb{I}\{\phi(x, \theta) \in V\} C^I(\phi(x, \theta), a) p_A^{n+1}(da|x, \theta) p_T^{n+1}(d\theta|x) \nu^n(dx) \end{aligned}$$

because on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  we have equality

$$\mu^{n+1}(dx \times d\theta \times da) = \mu^n(dx \times d\theta \times da) + p_A^{n+1}(da|x, \theta) p_T^{n+1}(d\theta|x) \nu^n(dx).$$

Recall that  $Q(\{\Delta\}|x, f^*(x)) = 1$ . Using the Tonelli Theorem (see [1, Thm.11.28]), we obtain:

$$\begin{aligned} &\int_V \int_{\bar{\mathbb{R}}_+^0} \left\{ \int_{[0, \theta]} C^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} du \right\} p_T^{n+1}(d\theta|x) \nu^n(dx) \\ &= \int_V \int_{\bar{\mathbb{R}}_+^0} \int_{[u, \infty]} C^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} p_T^{n+1}(d\theta|x) du \nu^n(dx) \\ &= \int_V \int_{\bar{\mathbb{R}}_+^0} C^g(\phi(x, u)) \mathbb{I}\{\phi(x, u) \in V\} p_T^{n+1}([u, \infty]|x) du \nu^n(dx) \\ &= \int_V C^g(y) \left\{ \int_V \int_{\bar{\mathbb{R}}_+^0} \delta_{\phi(x, u)}(dy) p_T^{n+1}([u, \infty]|x) du \nu^n(dx) \right\}, \end{aligned}$$

and, by induction and the definition of the measure  $\eta^{n+1}(\Gamma_X \times \square)$ ,

$$I^{n+1} = \int_V C^g(y) \eta^{n+1}(dy \times \square).$$

Similarly,

$$\begin{aligned} &\int_V \int_{\bar{\mathbb{R}}_+^0} \int_{\mathbf{A}} \mathbb{I}\{\phi(x, \theta) \in V\} C^I(\phi(x, \theta), a) p_A^{n+1}(da|x, \theta) p_T^{n+1}(d\theta|x) \nu^n(dx) \\ &= \int_V \int_{\mathbf{A}} C^I(y, a) \left\{ \int_V \int_{\bar{\mathbb{R}}_+^0} \delta_{\phi(x, \theta)}(dy) p_A^{n+1}(da|x, \theta) p_T^{n+1}(d\theta|x) \nu^n(dx) \right\}, \end{aligned}$$

and, by induction and the definition of the measure  $\eta^{n+1}(dx \times da)$  on  $V \times \mathbf{A}$ ,

$$J^{n+1} = \int_{V \times \mathbf{A}} C^I(y, a) \eta^{n+1}(dy \times da).$$

Since, for all positive measurable functions  $C^g$  on  $V$  and  $C^I$  on  $V \times \mathbf{A}$ ,

$$\begin{aligned} \int_V C^g(y) \eta^n(dy \times \square) \uparrow \int_V C^g(y) \eta(dy \times \square) \quad \text{and} \\ \int_{V \times \mathbf{A}} C^I(y, a) \eta^n(dy \times da) \uparrow \int_{V \times \mathbf{A}} C^I(y, a) \eta(dy \times da), \end{aligned}$$

we conclude that  $\eta^n \uparrow \eta$  on  $V \times \mathbf{A}_\square$  set-wise as  $n \rightarrow \infty$ . The last assertion is obvious.  $\square$

**Lemma 7.2** (a) Suppose  $\zeta$  is a finite measure on  $V$ , and the measure  $\hat{\zeta}(d\tilde{x}^0)$  and the stochastic kernel  $\check{\zeta}(dt|\tilde{x}^0)$  are as in Definition 4.3. Then, for each bounded (or positive, or negative) measurable function  $g$  on  $V$ ,

$$\int_V g(y) \zeta(dy) = \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} g(\phi((\tilde{x}^0, 0), u)) \check{\zeta}(du|\tilde{x}^0) \hat{\zeta}(d\tilde{x}^0) = \int_{\mathbf{D}} g(\phi((\tilde{x}^0, 0), u)) \check{\zeta}(d\tilde{x}^0 \times du).$$

(b) Suppose  $\zeta$  is a normal measure on  $V$ , and the measure  $\check{\zeta}(d\tilde{x}^0 \times dt)$  is as in Definition 4.3. Then, for each positive (or negative) measurable function  $g$  on  $V$ ,

$$\int_V g(y) \zeta(dy) = \int_{\mathbf{D}} g(\phi((\tilde{x}^0, 0), u)) \check{\zeta}(d\tilde{x}^0 \times du).$$

(c) Suppose  $\zeta$  is a normal (or finite) measure on the orbit

$$\tilde{z}\mathcal{X} \cap V = \{\phi((\tilde{z}, 0), t) : t \in [0, \tilde{\theta}^*(\tilde{z})]\}$$

and

$$m(I) := \zeta(\{\phi((\tilde{z}, 0), t) : t \in I\})$$

is the  $\sigma$ -finite (or finite) measure on  $[0, \tilde{\theta}^*(\tilde{z})]$ . (The set  $\{\phi((\tilde{z}, 0), t) : t \in I\}$  is measurable because if  $\tilde{z} \in \tilde{V}$ , then  $\phi((\tilde{z}, 0), \cdot)$  is a homeomorphism between  $[0, \tilde{\theta}^*(\tilde{z})]$  and  $\tilde{z}\mathcal{X} \cap V$ .) Then, for each positive or negative measurable function  $g$  on  $\tilde{z}\mathcal{X} \cap V$ ,

$$\int_{\tilde{z}\mathcal{X} \cap V} g(y) \zeta(dy) = \int_{[0, \tilde{\theta}^*(\tilde{z})]} g(\phi((\tilde{z}, 0), t)) m(dt).$$

Proof. (a) For the case of bounded functions  $g$ , it is sufficient to check the required formula for  $g(y) = \mathbb{I}\{y \in Y\}$ , where  $Y \in \mathcal{B}(V)$  is an arbitrary set. According to the definition of the mappings  $F$  and  $F^{-1}$ ,

$$(\tilde{x}^0, u) \in F^{-1}(Y) \iff F(\tilde{x}^0, u) \in Y \iff (\tilde{\phi}(\tilde{x}^0, u), u) = \phi((\tilde{x}^0, 0), u) \in Y.$$

Hence

$$\begin{aligned} \int_V g(y) \zeta(dy) &= \zeta(Y) = \check{\zeta}(F^{-1}(Y)) = \int_{\tilde{V} \times \mathbb{R}_+^0} \mathbb{I}\{(\tilde{x}^0, u) \in F^{-1}(Y)\} \check{\zeta}(d\tilde{x}^0 \times du) \\ &= \int_{\tilde{V}} \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in Y\} \check{\zeta}(du|\tilde{x}^0) \hat{\zeta}(d\tilde{x}^0). \end{aligned}$$

Moreover, for  $u \geq \tilde{\theta}^*(\tilde{x}^0)$ ,  $\phi((\tilde{x}^0, 0), u) \in V^c$  and thus  $\phi((\tilde{x}^0, 0), u)$  cannot belong to  $Y$ . The desired formula

$$\int_V g(y)\zeta(dy) = \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in Y\} \check{\zeta}(du|\tilde{x}^0) \hat{\zeta}(d\tilde{x}^0) = \int_{\mathbf{D}} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in Y\} \check{\zeta}(d\tilde{x}^0 \times du)$$

is proved.

For the case of positive functions  $g$ , one should apply the monotone convergence theorem to the sequence  $g \wedge N \uparrow g$ . Negative functions  $g$  can be treated similarly.

(b) The required formula is justified after we represent the function  $g$  as  $g(x) = \sum_{t=1}^{\infty} g_t(x)$  with  $g_t((\tilde{x}, u)) = \mathbb{I}\{u \in [t-1, t)\}g((\tilde{x}, u))$  and use the statement (a) separately for all  $g_t$ , where one can legitimately use the (finite) restriction of  $\zeta$  to the set  $\{x = (\tilde{x}, u) \in V : t-1 \leq u < t\}$ .

(c) Without loss of generality, we assume that  $\zeta(\tilde{z}\mathcal{X} \cap V) > 0$ . This implies  $\tilde{z} \in \tilde{V}$  in particular.

If  $\tilde{\theta}^*(\tilde{z}) < \infty$  then the measure  $\zeta$  is finite and can be extended to  $V$  by putting  $\zeta(V \setminus \tilde{z}\mathcal{X}) := 0$ . Now

$$\begin{aligned} \check{\zeta}(d\tilde{x}^0 \times dt) &= m(dt)\delta_{\tilde{z}}(d\tilde{x}^0); \\ \hat{\zeta}(d\tilde{x}^0) &= \zeta(\tilde{z}\mathcal{X} \cap V)\delta_{\tilde{z}}(d\tilde{x}^0); \\ \check{\zeta}(dt|\tilde{x}^0) &= \begin{cases} m(dt)/\zeta(\tilde{z}\mathcal{X} \cap V), & \text{if } \tilde{x}^0 = \tilde{z}; \\ \text{arbitrarily fixed probability measure,} & \text{if } \tilde{x}^0 \neq \tilde{z}, \end{cases} \end{aligned}$$

and the required equality follows from Item (a). The same reasoning applies if  $\tilde{\theta}^*(\tilde{z}) = \infty$  and the measure  $\zeta$  is finite.

Suppose  $\tilde{\theta}^*(\tilde{z}) = \infty$ , so that  $\tilde{z}\mathcal{X} \cap V = \tilde{z}\mathcal{X}$ , and the measure  $\zeta$  is not finite, but normal. It is sufficient to check the required formula for  $g_t(y) = \mathbb{I}\{y \in Y_t\}$ , where

$$Y_t = \{\phi((\tilde{z}, 0), u) : u \in I_t \in \mathcal{B}([t-1, t))\}, \quad t = 1, 2, \dots$$

As mentioned in the statement of this lemma, the mapping  $[G(u) := \phi((\tilde{z}, 0), u)]$  is a homeomorphism between  $\mathbb{R}_+^0$  and  $\tilde{z}\mathcal{X}$  (see Lemma 4.1), and all different subsets  $I_t \in \mathcal{B}([t-1, t))$  produce all possible subsets  $Y_t \in \mathcal{B}(\{\phi((\tilde{z}, 0), u) : u \in [t-1, t)\})$ . Thus, for an arbitrary set  $Y \in \mathcal{B}(\tilde{z}\mathcal{X})$ , we have  $Y = \cup_{t=1}^{\infty} Y_t$  with

$$Y_t := Y \cap \{\phi((\tilde{z}, 0), u) : u \in [t-1, t)\} \in \mathcal{B}(\{\phi((\tilde{z}, 0), u) : u \in [t-1, t)\}),$$

and the proof will be completed by applying the monotone convergence theorem.

Now

$$\begin{aligned} \int_{\tilde{z}\mathcal{X}} g_t(y)\zeta(dy) &= \zeta(Y_t); \\ \int_{\mathbb{R}_+^0} g_t(\phi((\tilde{z}, 0), u))m(du) &= \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{z}, 0), u) \in Y_t\}m(du) = m(I_t), \end{aligned}$$

and  $m(I_t) = \zeta(Y_t)$  by the definition of the measure  $m$ . □

**Lemma 7.3** *Suppose an orbit*

$$\tilde{x}\mathcal{X} \cap V = \{\phi((\tilde{x}, 0), t) : t \in [0, \tilde{\theta}^*(\tilde{x}))\}$$

*is fixed and  $p^*$  is a probability measure on  $\mathbb{R}_+^0$  such that  $p^*([\tilde{\theta}^*(\tilde{x}), \infty)) = 0$ .*

Then the measures  $\tilde{\eta}_\square^*$  and  $\tilde{\eta}_A^*$  on  $\tilde{x}\mathcal{X} \cap V$ , defined as

$$\begin{aligned}\eta_\square^*(\Gamma) &:= \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}, 0), u) \in \Gamma\} p^*([u, \infty]) du = \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}, 0), u) \in \Gamma\} (1 - p^*([0, u])) du, \\ \eta_A^*(\Gamma) &:= \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}, 0), u) \in \Gamma\} p^*(du), \quad \Gamma \in \mathcal{B}(\tilde{x}\mathcal{X} \cap V),\end{aligned}$$

satisfy equation

$$0 = w((\tilde{x}, 0)) + \int_{\tilde{x}\mathcal{X} \cap V} \chi w(x) \eta_\square^*(dx) - \int_{\tilde{x}\mathcal{X} \cap V} w(x) \eta_A^*(dx) \quad (30)$$

for all functions  $w \in \mathbf{W}$ . The measure  $\eta_A^*$  is finite, and the measure  $\eta_\square^*$  is normal on that orbit.

Proof. The properties of the measures  $\eta_A^*$  and  $\eta_\square^*$  formulated in the last sentence of this lemma are obvious, c.f. the reasoning in the proof of Lemma 4.2(a).

Now let  $w \in \mathbf{W}$  be fixed. We verify the rest of the statement of this lemma by distinguishing the following two cases.

(i) Suppose that  $u^* := \inf\{u \in \bar{\mathbb{R}}_+^0 : p^*([0, u]) = 1\} \geq \tilde{\theta}^*(\tilde{x})$ . The expression

$$I := w((\tilde{x}, 0)) + \int_{\tilde{x}\mathcal{X} \cap V} \chi w(x) \eta_\square^*(dx) - \int_{\tilde{x}\mathcal{X} \cap V} w(x) \eta_A^*(dx)$$

is well defined because the measure  $\eta_\square^*$  is normal, the integral  $\int_{\tilde{x}\mathcal{X} \cap V} \chi w(x) \eta_\square^*(dx)$  is positive or negative, the function  $w$  is bounded and the measure  $\eta_A^*$  is finite. According to Lemma 7.2(c),

$$\begin{aligned}I &= w((\tilde{x}, 0)) + \int_{[0, \tilde{\theta}^*(\tilde{x})]} \chi w(\phi((\tilde{x}, 0), t)) [1 - p^*([0, t])] dt - \int_{[0, \tilde{\theta}^*(\tilde{x})]} w(\phi((\tilde{x}, 0), t)) p^*(dt) \\ &= - \left[ \int_{[0, \tilde{\theta}^*(\tilde{x})]} \chi w(\phi((\tilde{x}, 0), t)) p^*([0, t]) dt + \int_{[0, \tilde{\theta}^*(\tilde{x})]} w(\phi((\tilde{x}, 0), t)) p^*(dt) \right].\end{aligned}$$

The last equality is by Lemma A.1 and Definition 4.5 of the space  $\mathbf{W}$ :

$$w((\tilde{x}, 0)) + \lim_{T \rightarrow \tilde{\theta}^*(\tilde{x})} \int_{[0, T]} \chi w(\phi((\tilde{x}, 0), t)) dt = \lim_{T \rightarrow \tilde{\theta}^*(\tilde{x})} w(\phi((\tilde{x}, 0), T)) = 0.$$

We apply the Tonelli Theorem [1, Thm.11.28] to the first integral in the square brackets and again use Lemma A.1:

$$\begin{aligned}\int_{[0, \tilde{\theta}^*(\tilde{x})]} \chi w(\phi((\tilde{x}, 0), t)) \int_{[0, t]} p^*(du) dt &= \int_{[0, \tilde{\theta}^*(\tilde{x})]} \int_{(u, \tilde{\theta}^*(\tilde{x}))} \chi w(\phi((\tilde{x}, 0), t)) dt p^*(du) \\ &= \int_{[0, \tilde{\theta}^*(\tilde{x})]} [-w(\phi((\tilde{x}, 0), u))] p^*(du).\end{aligned}$$

Thus  $I = 0$ .

(ii) Suppose that  $u^* := \inf\{u \in \bar{\mathbb{R}}_+^0 : p^*([0, u]) = 1\} < \tilde{\theta}^*(\tilde{x})$ . Since measures  $\tilde{\eta}_A^*$  and  $\tilde{\eta}_\square^*$  both equal zero on the set  $\{\phi((\tilde{x}, 0), t) : t > u^*\}$ , it is sufficient to show that

$$I := w((\tilde{x}, 0)) + \int_{\mathcal{X}_0^{u^*}} \chi w(x) \eta_\square^*(dx) - \int_{\mathcal{X}_0^{u^*}} w(x) \eta_A^*(dx) = 0,$$

where

$$\mathcal{X}_0^{u^*} := \{\phi((\tilde{x}, 0), t) : 0 \leq t \leq u^*\}.$$

This expression is well defined because the measure  $\eta_{\square}^*$  is normal, the integral  $\int_{\mathcal{X}_0^{u^*}} \chi w(x) \eta_{\square}^*(dx)$  is positive or negative, the function  $w$  is bounded and the measure  $\eta_A^*$  is finite. The measure  $\eta_{\square}^*$  is non-atomic, and the first integral can be calculated over

$$\mathcal{X}_0^{u^*-} := \{\phi((\tilde{x}, 0), t) : 0 \leq t < u^*\},$$

so that, by Lemma 7.2(c),

$$\begin{aligned} I &= w((\tilde{x}, 0)) + \int_{[0, u^*)} \chi w(\phi((\tilde{x}, 0), t)) [1 - p^*([0, t))] dt \\ &\quad - \int_{[0, u^*)} w(\phi((\tilde{x}, 0), t)) p^*(dt) - w(\phi((\tilde{x}, 0), u^*)) [1 - p^*([0, u^*))]. \end{aligned}$$

In the last term,  $[1 - p^*([0, u^*))] = p^*({u^*})$ . Since

$$w((\tilde{x}, 0)) + \int_{[0, u^*)} \chi w(\phi((\tilde{x}, 0), t)) dt - w(\phi((\tilde{x}, 0), u^*)) = 0$$

(see Lemma A.1), after we subtract this equality from  $I$ , we obtain

$$I = - \int_{[0, u^*)} \chi w(\phi((\tilde{x}, 0), t)) p^*([0, t]) dt - \int_{[0, u^*)} w(\phi((\tilde{x}, 0), t)) p^*(dt) + w(\phi((\tilde{x}, 0), u^*)) p^*([0, u^*)).$$

Finally, apply the Tonelli Theorem (see [1, Thm.11.28]) to the first term and again use Lemma A.1:

$$\begin{aligned} \int_{[0, u^*)} \int_{[0, t)} \chi w(\phi((\tilde{x}, 0), t)) p^*(du) dt &= \int_{[0, u^*)} \int_{(u, u^*)} \chi w(\phi((\tilde{x}, 0), t)) dt p^*(du) \\ &= \int_{[0, u^*)} [w(\phi((\tilde{x}, 0), u^*)) - w(\phi((\tilde{x}, 0), u))] p^*(du) \\ &= w(\phi((\tilde{x}, 0), u^*)) p^*([0, u^*)) - \int_{[0, u^*)} w(\phi((\tilde{x}, 0), u)) p^*(du). \end{aligned}$$

Therefore,  $I = 0$ .

The proof is completed.  $\square$

**Lemma 7.4** *Suppose  $\nu$  is a finite measure on  $V$  such that  $\nu(V \cap (\tilde{V} \times \{t : t > 0\})) = 0$ ,  $\tilde{\eta}$  is a finite measure on  $V \times \mathbf{A}$ ,  $\tilde{\eta}_{\square}$  is a normal measure on  $V$  and  $\tilde{\eta}_A$  is a finite measure on  $V$  which satisfy equation*

$$0 = \int_V w(x) \nu(dx) + \int_V \chi w(x) \tilde{\eta}_{\square}(dx) - \int_V w(x) \tilde{\eta}_A(dx) + \int_{V \times \mathbf{A}} w(l(x, a)) \tilde{\eta}(dx \times da) \quad (31)$$

for all functions  $w \in \mathbf{W}$ . Then there is a stochastic kernel  $\tilde{p}(dt|x)$  on  $\bar{\mathbb{R}}_+^0$  given  $V$  such that, for  $\theta^*$  given by

$$\theta^*(x) := \inf\{\theta \in \mathbb{R}_+^0 : \phi(x, \theta) \in V^c\}, \quad (32)$$

$\tilde{p}([\theta^*(x), \infty)|x) = 0$  for all  $x \in V$  and the measures

$$\begin{aligned} \tilde{\eta}'_A(\Gamma) &:= \int_V \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi(x, u) \in \Gamma\} \tilde{p}(du|x) \nu(dx) \\ \text{and } \tilde{\eta}'_{\square}(\Gamma) &:= \int_V \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi(x, u) \in \Gamma\} \tilde{p}([u, \infty]|x) du \nu(dx), \quad \Gamma \in \mathcal{B}(V) \end{aligned}$$

satisfy equation

$$0 = \int_V w(x)\nu(dx) + \int_V \chi w(x)\tilde{\eta}'_{\square}(dx) - \int_V w(x)\tilde{\eta}'_A(dx) \quad (33)$$

for all functions  $w \in \mathbf{W}$ . Moreover, the set functions  $\tilde{\eta}_{\square}(\Gamma) - \tilde{\eta}'_{\square}(\Gamma)$  and  $\tilde{\eta}_A(\Gamma) - \tilde{\eta}'_A(\Gamma)$  on  $\mathcal{B}(V)$  are again normal and finite measures, correspondingly.

Proof. (i) Firstly, we introduce several functions, measures and sets, describe their properties and define explicitly the stochastic kernel  $\tilde{p}$ .

The necessary properties of the function  $\theta^*$  were established during the proof of Theorem 4.1. Note that, for each  $\Gamma \in \mathcal{B}(V)$ , the function  $\mathbb{I}\{\phi(x, u) \in \Gamma\}$  is measurable since the flow  $\phi$  is continuous. Below,  $\hat{\nu}(\tilde{\Gamma}) := \nu(\tilde{\Gamma} \times \{0\})$  for  $\tilde{\Gamma} \in \mathcal{B}(\tilde{V})$ .

In accordance with Definition 4.3, we introduce the finite measure  $\hat{\eta}_A(d\tilde{x}^0)$  and stochastic kernel  $\check{\eta}_A(dt|\tilde{x}^0)$  coming from  $\tilde{\eta}_A(dx)$ . Next, introduce the finite measure

$$K := \hat{\nu} + \hat{\eta}_A$$

on  $\tilde{V}$  and the Radon-Nikodym derivatives

$$n(\tilde{x}^0) := \frac{d\hat{\nu}}{dK}(\tilde{x}^0), \quad \text{and} \quad a(\tilde{x}^0) := \frac{d\hat{\eta}_A}{dK}(\tilde{x}^0).$$

Below, we fix one specific version of the derivative  $n$  and of the derivative  $a$ . On the set

$$\tilde{\mathbf{V}}_{\nu} := \{\tilde{x}^0 \in \tilde{V} : n(\tilde{x}^0) > 0\},$$

we have

$$\hat{\eta}_A(\tilde{\Gamma}) = \int_{\tilde{\Gamma}} a(\tilde{x}^0)K(d\tilde{x}^0) = \int_{\tilde{\Gamma}} \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)}\hat{\nu}(d\tilde{x}^0)$$

for all  $\tilde{\Gamma} \in \mathcal{B}(\tilde{\mathbf{V}}_{\nu})$ . See Figure 1. Note that  $\hat{\nu}(\tilde{V} \setminus \tilde{\mathbf{V}}_{\nu}) = 0$ .

Since the function  $\mathbb{I}\{u \leq t\}$  of  $(u, t)$  is measurable, the integral  $\int_{\mathbb{R}_+^0} \mathbb{I}\{u \leq t\}\check{\eta}_A(du|\tilde{x}^0)$  is a measurable function of  $(\tilde{x}^0, t)$  (see [5, Prop.7.29]), and hence the function

$$G(\tilde{x}^0, t) := \check{\eta}_A([0, t]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} = \int_{\mathbb{R}_+^0} \mathbb{I}\{u \leq t\}\check{\eta}_A(du|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)}, \quad \tilde{x}^0 \in \tilde{\mathbf{V}}_{\nu}, \quad t \in \mathbb{R}_+^0$$

is measurable. For all  $\tilde{x}^0 \in \tilde{\mathbf{V}}_{\nu}$ , the function  $G(\tilde{x}^0, \cdot)$  clearly increases and is right-continuous: it is constant for  $t \geq \tilde{\theta}^*(\tilde{x}^0)$  and, if  $t_i \downarrow t \in [0, \tilde{\theta}^*(\tilde{x}^0))$  then  $\check{\eta}_A([0, t_i]|\tilde{x}^0) \downarrow \check{\eta}_A([0, t]|\tilde{x}^0)$ .

Let us introduce the function

$$u^*(\tilde{x}^0) := \inf\{t \in \mathbb{R}_+^0 : G(\tilde{x}^0, t) \geq 1\} \in \bar{\mathbb{R}}_+^0, \quad \tilde{x}^0 \in \tilde{\mathbf{V}}_{\nu}.$$

When  $u^*(\tilde{x}^0) < \infty$ , this infimum is attained because the function  $G(\tilde{x}^0, \cdot)$  is right-continuous; and  $G(\tilde{x}^0, u^*(\tilde{x}^0)-) \leq 1$ . To show that the function  $u^*(\cdot)$  is measurable, note that the function

$$f(\tilde{x}^0, t) := \infty \times \mathbb{I}\{G(\tilde{x}^0, t) < 1\} + t \times \mathbb{I}\{G(\tilde{x}^0, t) \geq 1\}$$

is measurable and the function  $t \rightarrow f(\tilde{x}^0, t)$  is lower semicontinuous for each  $\tilde{x}^0 \in \tilde{\mathbf{V}}_{\nu}$ . Now, the function  $u^*(\tilde{x}^0) = \inf_{t \in \bar{\mathbb{R}}_+^0} f(\tilde{x}^0, t)$  is measurable by [22, Thm.2]; see also Corollary 1 and Remark 1 of [9]. Note also that if  $u^*(\tilde{x}^0) > \tilde{\theta}^*(\tilde{x}^0)$ , then  $u^*(\tilde{x}^0) = \infty$ . Figure 3 can serve as an illustration.

For  $\tilde{x}^0 \in \tilde{\mathbf{V}}_\nu$ , we put

$$\begin{aligned} \tilde{p}(I|\tilde{x}^0, 0) &:= \check{\eta}_A(I \cap [0, u^*(\tilde{x}^0) \wedge \tilde{\theta}^*(\tilde{x}^0)]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \\ &\quad + \mathbb{I}\{u^*(\tilde{x}^0) < \tilde{\theta}^*(\tilde{x}^0)\} \mathbb{I}\{u^*(\tilde{x}^0) \in I\} \left[ 1 - \check{\eta}_A([0, u^*(\tilde{x}^0)]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \right] \end{aligned}$$

for all  $I \in \mathcal{B}(\mathbb{R}_+^0)$ ,

$$\text{and } \tilde{p}(\{\infty\}|\tilde{x}^0, 0) := 1 - \tilde{p}(\mathbb{R}_+^0|\tilde{x}^0, 0).$$

For all other points  $x \in V$ , we put  $\tilde{p}(\{\infty\}|x) = 1$  and  $\tilde{p}(I|x) \equiv 0$  for  $I \in \mathcal{B}(\mathbb{R}_+^0)$ . Clearly,  $\tilde{p}([\theta^*(x), \infty)|x) = 0$  for all  $x \in V$ . The possible shapes of the distribution function  $\tilde{p}([0, t]|\tilde{x}^0, 0)$  are shown on Figure 3.

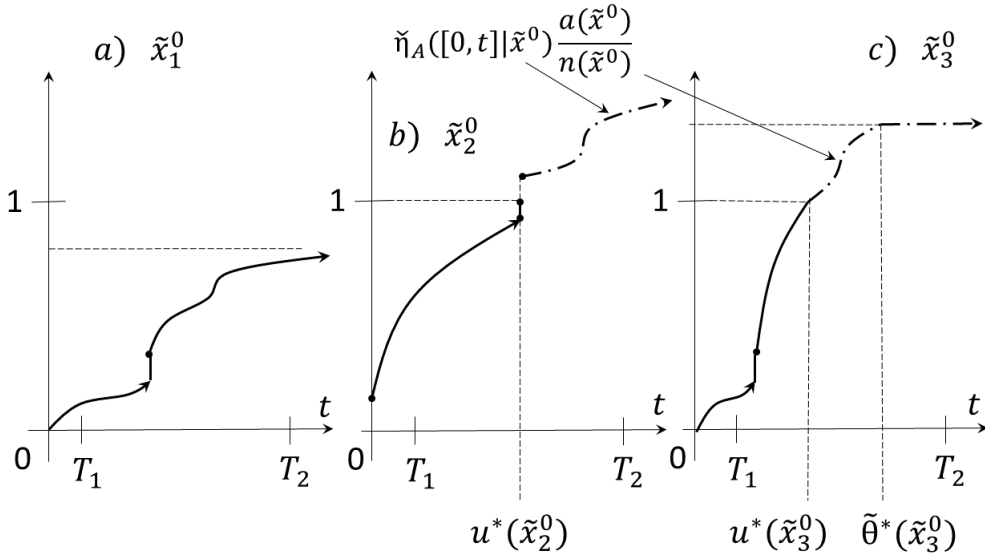


Figure 3: Graphs of the function  $\tilde{p}([0, t]|\tilde{x}^0, 0)$ , see also Figure 4. In case a),  $\tilde{p}([0, t]|\tilde{x}_1^0, 0) = \check{\eta}_A([0, t]|\tilde{x}_1^0) \frac{a(\tilde{x}_1^0)}{n(\tilde{x}_1^0)}$  for all  $t \in \mathbb{R}_+^0$ ,  $u^*(\tilde{x}_1^0) = \tilde{\theta}^*(\tilde{x}_1^0) = \infty$  and  $\tilde{p}(\mathbb{R}_+^0|\tilde{x}_1^0, 0) < 1$ . In case b),  $0 < u^*(\tilde{x}_2^0) < \tilde{\theta}^*(\tilde{x}_2^0)$ ,  $\check{\eta}_A(\{u^*(\tilde{x}_2^0)\}|\tilde{x}_2^0) > 0$ . In case c),  $0 < u^*(\tilde{x}_3^0) < \tilde{\theta}^*(\tilde{x}_3^0) < \infty$ ,  $\check{\eta}_A(\{u^*(\tilde{x}_3^0)\}|\tilde{x}_3^0) = 0$ .

(ii) Let us prove that equation (33) holds. Since  $\nu(\tilde{\mathbf{V}} \times \{t : t > 0\}) = 0$ ,

$$\tilde{\eta}'_A(\Gamma) = \int_{\tilde{\mathbf{V}}} \eta_A^*(\Gamma|\tilde{x}^0, 0) \hat{\nu}(d\tilde{x}^0); \quad \text{and} \quad \tilde{\eta}'_{\square}(\Gamma) = \int_{\tilde{\mathbf{V}}} \eta_{\square}^*(\Gamma|\tilde{x}^0, 0) \hat{\nu}(d\tilde{x}^0)$$

for all  $\Gamma \in \mathcal{B}(V)$ , where

$$\begin{aligned} \eta_A^*(\Gamma|\tilde{x}^0, 0) &:= \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \tilde{p}(du|\tilde{x}^0, 0); \\ \eta_{\square}^*(\Gamma|\tilde{x}^0, 0) &:= \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \tilde{p}([u, \infty]|\tilde{x}^0, 0) du. \end{aligned}$$

The introduced measures  $\eta_A^*$  and  $\eta_{\square}^*$  are concentrated on  $_{\tilde{x}^0}\mathcal{X} \cap V$  for each  $\tilde{x}^0 \in \tilde{\mathbf{V}}$ . By the way,  $\eta_A^*$  and  $\eta_{\square}^*$  are measurable kernels because the flow  $\phi$  is continuous and  $\tilde{p}$  is a (measurable) stochastic



kernel. Now

$$\begin{aligned}
& \int_V w(x)\nu(dx) + \int_V \chi w(x)\tilde{\eta}'_{\square}(dx) - \int_V w(x)\tilde{\eta}'_A(dx) \\
&= \int_{\tilde{V}} w((\tilde{x}^0, 0))\hat{\nu}(d\tilde{x}^0) + \int_{\tilde{V}} \int_{\tilde{x}^0\mathcal{X}\cap V} \chi w(x)\eta_{\square}^*(dx|(\tilde{x}^0, 0))\hat{\nu}(d\tilde{x}^0) \\
&\quad - \int_{\tilde{V}} \int_{\tilde{x}^0\mathcal{X}\cap V} w(x)\eta_A^*(dx|(\tilde{x}^0, 0))\hat{\nu}(d\tilde{x}^0) \\
&= \int_{\tilde{V}} \left[ w((\tilde{x}^0, 0)) + \int_{\tilde{x}^0\mathcal{X}\cap V} \chi w(x)\eta_{\square}^*(dx|(\tilde{x}^0, 0)) - \int_{\tilde{x}^0\mathcal{X}\cap V} w(x)\eta_A^*(dz|(\tilde{x}^0, 0)) \right] \hat{\nu}(d\tilde{x}^0).
\end{aligned}$$

The re-arrangement is legal because the function  $w$  is bounded, the function  $\chi w$  is positive (or negative), the measure  $\tilde{\eta}'_{\square}$  is normal, and the measures  $\hat{\nu}$  and  $\eta_A^*(dx|(\tilde{x}^0, 0))$  are finite (for all  $\tilde{x}^0 \in \tilde{V}$ ). Equation (33) follows from Lemma 7.3.

(iii) Let us show that  $\tilde{\eta}_A - \tilde{\eta}'_A$  is a finite measure. In case  $\tilde{x}^0 \in \tilde{\mathbf{V}}_{\nu}$  and  $u^*(\tilde{x}^0) < \tilde{\theta}^*(\tilde{x}^0)$ ,

$$\tilde{p}(\{u^*(\tilde{x}^0)\}|\tilde{x}^0, 0) = 1 - \check{\eta}_A([0, u^*(\tilde{x}^0)]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \leq \check{\eta}_A(\{u^*(\tilde{x}^0)\}|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)}$$

because

$$\check{\eta}_A([0, u^*(\tilde{x}^0)]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} = G(\tilde{x}^0, u^*(\tilde{x}^0)) \geq 1.$$

Therefore, whether  $u^*(\tilde{x}^0) < \tilde{\theta}^*(\tilde{x}^0)$  or  $u^*(\tilde{x}^0) \geq \tilde{\theta}^*(\tilde{x}^0)$ ,

$$\tilde{p}(I|\tilde{x}^0, 0) \leq \check{\eta}_A(I|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)}$$

for all  $I \in \mathcal{B}(\mathbb{R}_+^0)$  and for all  $\tilde{x}^0 \in \tilde{\mathbf{V}}_{\nu}$ . Now, for each measurable subset  $\Gamma \subset V$ ,

$$\begin{aligned}
\tilde{\eta}'_A(\Gamma) &= \int_{\tilde{\mathbf{V}}_{\nu}} \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \tilde{p}(du|\tilde{x}^0, 0) \hat{\nu}(d\tilde{x}^0) \\
&= \int_{\tilde{\mathbf{V}}_{\nu}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \tilde{p}(du|\tilde{x}^0, 0) \hat{\nu}(d\tilde{x}^0) \\
&\leq \int_{\tilde{\mathbf{V}}_{\nu}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \check{\eta}_A(du|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\
&= \int_{\tilde{\mathbf{V}}_{\nu}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \check{\eta}_A(du|\tilde{x}^0) \hat{\eta}_A(d\tilde{x}^0) \\
&\leq \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \Gamma\} \check{\eta}_A(du|\tilde{x}^0) \hat{\eta}_A(d\tilde{x}^0) \\
&= \int_V \mathbb{I}\{y \in \Gamma\} \tilde{\eta}_A(dy) = \tilde{\eta}_A(\Gamma).
\end{aligned}$$

The last but one equality is by Lemma 7.2(a). Hence,  $\tilde{\eta}_A - \tilde{\eta}'_A$  is a finite measure.

(iv) Let us show that  $\tilde{\eta}_{\square} \geq \tilde{\eta}'_{\square}$  set-wise. Recall that the measure  $\tilde{\eta}'_{\square}$  is normal. It is convenient to consider, with some abuse of notations, the images  $\check{\eta}_{\square}$ ,  $\check{\eta}'_{\square}$  and  $\check{\eta}$  of the measures  $\tilde{\eta}_{\square}$ ,  $\tilde{\eta}'_{\square}$  and  $\tilde{\eta}(\cdot \times \mathbf{A})$  as in Definition 4.3. Recall that  $\tilde{\eta}_{\square} \geq \tilde{\eta}'_{\square} \Leftrightarrow \check{\eta}_{\square} \geq \check{\eta}'_{\square}$ . Now, according to Lemma 7.2(a,b), equation (31) takes the form:

$$\begin{aligned}
0 &= \int_{\tilde{V}} w((\tilde{x}^0, 0))\hat{\nu}(d\tilde{x}^0) + \int_{\mathbf{D}} \chi w(\phi((\tilde{x}^0, 0), u))\check{\eta}_{\square}(d\tilde{x}^0 \times du) \\
&\quad - \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} w(\phi((\tilde{x}^0, 0), u))\check{\eta}_A(du|\tilde{x}^0)\hat{\eta}_A(d\tilde{x}^0) + \int_{\mathbf{D}} w^A(\phi((\tilde{x}^0, 0), u))\check{\eta}(d\tilde{x}^0 \times du),
\end{aligned} \tag{34}$$

where

$$w^A(y) := \int_{\mathbf{A}} w(l(y, a)) \tilde{\eta}^A(da|y) \quad (35)$$

and the stochastic kernel  $\tilde{\eta}^A(da|y)$  comes from the decomposition

$$\tilde{\eta}(dy \times da) = \tilde{\eta}^A(da|y) \tilde{\eta}(dy \times \mathbf{A}).$$

According to [7, V.1;Thm.1.5.6], it suffices to show that the value of the measure  $\check{\eta}_\square$  is greater or equal to the value of  $\check{\eta}'_\square$  on each set of the form

$$Y_{T_1, T_2, \tilde{\Gamma}} := \{(\tilde{x}^0, u) : \tilde{x}^0 \in \tilde{\Gamma}, T_1 \leq u < T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)\}, \quad \tilde{\Gamma} \in \mathcal{B}(\tilde{V}), \quad 0 \leq T_1 < T_2 < \infty.$$

See Figure 4 and also Figure 3 for illustration.

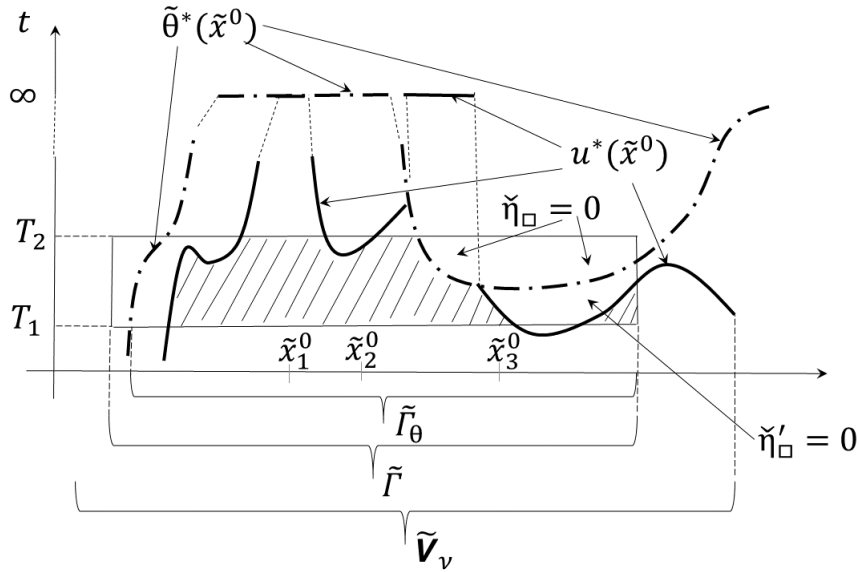


Figure 4: Space  $\mathbf{D} = \{(\tilde{x}^0, t) : \tilde{\phi}(\tilde{x}^0, t) \in \tilde{V}\}$  and “rectangle”  $Y_{T_1, T_2, \tilde{\Gamma}}$ . The points  $\tilde{x}_1^0, \tilde{x}_2^0, \tilde{x}_3^0$  belong to  $\tilde{V}_\nu \subset \tilde{V}$ . The dashed area is the part of  $Y_{T_1, T_2, \tilde{\Gamma}}$  where  $\check{\eta}'_\square$  might be positive.

Note that, in case  $\tilde{\Gamma} \subset \tilde{V} \setminus \tilde{V}_\nu$ , since  $\nu((\tilde{V} \setminus \tilde{V}_\nu) \times \{0\}) = 0$ ,  $\check{\eta}'_\square(Y_{T_1, T_2, \tilde{\Gamma}}) = 0$  and hence  $\check{\eta}_\square(Y_{T_1, T_2, \tilde{\Gamma}}) - \check{\eta}'_\square(Y_{T_1, T_2, \tilde{\Gamma}}) \geq 0$  for all  $T_1, T_2$ . Therefore, below in this proof, we assume that  $\tilde{\Gamma} \subset \tilde{V}_\nu$ .

To use equality (34) for calculating  $\check{\eta}_\square(Y_{T_1, T_2, \tilde{\Gamma}})$ , we put

$$\chi w_{T_1, T_2, \tilde{\Gamma}}(\phi((\tilde{x}^0, 0), u)) := -\mathbb{I}\{(\tilde{x}^0, u) \in Y_{T_1, T_2, \tilde{\Gamma}}\},$$

and consider the following positive function decreasing along the flow:

$$\begin{aligned} w_{T_1, T_2, \tilde{\Gamma}}(y) &= w_{T_1, T_2, \tilde{\Gamma}}(\tilde{y}, t) \\ &:= \mathbb{I}\{h(y) \in \tilde{\Gamma}\} \times \begin{cases} T_2 \wedge \tilde{\theta}^*(h(y)) - T_1 \wedge \tilde{\theta}^*(h(y)), & \text{if } 0 \leq t \leq T_1; \\ T_2 \wedge \tilde{\theta}^*(h(y)) - t, & \text{if } T_1 < t \leq T_2 \wedge \tilde{\theta}^*(h(y)); \\ 0, & \text{if } t > T_2 \wedge \tilde{\theta}^*(h(y)). \end{cases} \end{aligned} \quad (36)$$

See Figure 5. The function  $h$  was introduced in Definition 4.2. Clearly,  $w_{T_1, T_2, \tilde{\Gamma}} \in \mathbf{W}$  for all  $0 \leq T_1 < T_2 < \infty$ ,  $\tilde{\Gamma} \in \mathcal{B}(\tilde{V})$ , and

$$w_{T_1, T_2, \tilde{\Gamma}}((\tilde{x}^0, 0)) = \mathbb{I}\{\tilde{x}^0 \in \tilde{\Gamma}\} \left( T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \wedge \tilde{\theta}^*(\tilde{x}^0) \right);$$

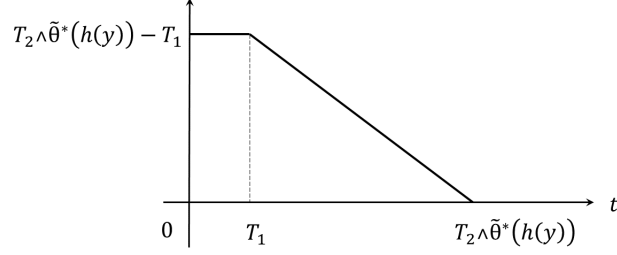


Figure 5: Graph of the function  $w_{T_1, T_2, \tilde{\Gamma}}(\tilde{y}, t) = w_{T_1, T_2, \tilde{\Gamma}}(\tilde{\phi}(\tilde{x}^0, t), t)$  for a fixed value of  $h(y) = \tilde{x}^0 \in \tilde{\Gamma}$  and  $\tilde{\theta}^*(h(y)) > T_1$ .

$$w_{T_1, T_2, \tilde{\Gamma}}(\phi((\tilde{x}^0, 0), u)) = \mathbb{I}\{\tilde{x}^0 \in \tilde{\Gamma}\} \times \begin{cases} T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \wedge \tilde{\theta}^*(\tilde{x}^0), & \text{if } u \leq T_1; \\ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - u, & \text{if } T_1 < u \leq T_2 \wedge \tilde{\theta}^*(\tilde{x}^0); \\ 0, & \text{if } u > T_2 \wedge \tilde{\theta}^*(\tilde{x}^0). \end{cases}$$

The expression (35) takes the form

$$w_{T_1, T_2, \tilde{\Gamma}}^A(y) := \int_{\mathbf{A}} w_{T_1, T_2, \tilde{\Gamma}}(l(y, a)) \tilde{\eta}^A(da|y).$$

From equality (34), using the expression  $\hat{\eta}_A(d\tilde{x}^0) = \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0)$ , we have for

$$\tilde{\Gamma}_\theta := \tilde{\Gamma} \cap \{\tilde{x}^0 : \tilde{\theta}^*(\tilde{x}^0) \geq T_1\} :$$

$$\begin{aligned} \check{\eta}_\square(Y_{T_1, T_2, \tilde{\Gamma}}) &= \int_{\tilde{\Gamma}_\theta} [T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1] \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}_\theta} [T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1] \check{\eta}_A([0, T_1]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\ &\quad - \int_{\tilde{\Gamma}_\theta} \int_{(T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} [T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - u] \check{\eta}_A(du|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\ &\quad + \int_{\mathbf{D}} w_{T_1, T_2, \tilde{\Gamma}}^A(\phi((\tilde{x}^0, 0), u)) \check{\eta}(d\tilde{x}^0 \times du). \end{aligned}$$

For the last but one integral, note that  $\check{\eta}_A(\{\tilde{\theta}^*(\tilde{x}^0)\}|\tilde{x}^0) = 0$  for  $\hat{\eta}_A$ -almost all  $\tilde{x}^0$  since  $\tilde{\eta}_A$  is concentrated on  $V$ . The corresponding integrals over  $\tilde{\Gamma} \setminus \tilde{\Gamma}_\theta$  equal zero and hence are omitted; the last term above, denoted below as  $\mathbf{J}(\tilde{\Gamma})$ , is positive. According to Lemma A.2, for  $\hat{\eta}_A$ -almost all  $\tilde{x}^0 \in \tilde{\Gamma}_\theta$ ,

$$\begin{aligned} &\int_{(T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} [u - T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)] \check{\eta}_A(du|\tilde{x}^0) = \int_{(T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} [u - T_1] \check{\eta}_A(du|\tilde{x}^0) \\ &\quad - [T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1] \check{\eta}_A((T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]|\tilde{x}^0) \\ &= [T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1] \check{\eta}_A([T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]|\tilde{x}^0) - \int_{(0, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1]} \check{\eta}_A([T_1, T_1 + s]|\tilde{x}^0) ds \\ &\quad - [T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1] \check{\eta}_A((T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]|\tilde{x}^0), \end{aligned}$$

so that

$$\begin{aligned}
\check{\eta}_{\square}(Y_{T_1, T_2, \tilde{\Gamma}}) &= \int_{\tilde{\Gamma}_{\theta}} \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}_{\theta}} \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \check{\eta}_A([0, T_1]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\
&\quad + \int_{\tilde{\Gamma}_{\theta}} \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \check{\eta}_A(\{T_1\}|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\
&\quad - \int_{\tilde{\Gamma}_{\theta}} \int_{(0, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1]} \check{\eta}_A([T_1, T_1 + s]|\tilde{x}^0) ds \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) + \mathbf{J}(\tilde{\Gamma}) \\
&= \int_{\tilde{\Gamma}_{\theta}} \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}_{\theta}} \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \check{\eta}_A([0, T_1]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\
&\quad - \int_{\tilde{\Gamma}_{\theta}} \int_{(T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} \check{\eta}_A([T_1, u]|\tilde{x}^0) du \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) + \mathbf{J}(\tilde{\Gamma}). \tag{37}
\end{aligned}$$

According to the definitions of the measures  $\check{\eta}'_{\square}$  and  $\check{\eta}'_{\square}$ ,

$$\begin{aligned}
\check{\eta}'_{\square}(Y_{T_1, T_2, \tilde{\Gamma}}) &= \check{\eta}'_{\square}(F(Y_{T_1, T_2, \tilde{\Gamma}})) = \int_{\tilde{V}} \int_{\mathbb{R}_+^0} \mathbb{I}\{(\tilde{x}^0, u) \in Y_{T_1, T_2, \tilde{\Gamma}}\} (1 - \tilde{p}([0, u]|(\tilde{x}^0, 0))) du \hat{\nu}(d\tilde{x}^0) \\
&= \int_{\tilde{V}_{\nu}} \int_{[T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} \mathbb{I}\{\tilde{x}^0 \in \tilde{\Gamma}\} (1 - \tilde{p}([0, u]|(\tilde{x}^0, 0))) du \hat{\nu}(d\tilde{x}^0) \tag{38} \\
&= \int_{\tilde{\Gamma}} \int_{[T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} du \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}} \int_{[T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} (\tilde{p}([0, T_1]|(\tilde{x}^0, 0)) \\
&\quad + \tilde{p}([T_1, u]|(\tilde{x}^0, 0))) du \hat{\nu}(d\tilde{x}^0) \\
&= \int_{\tilde{\Gamma}_{\theta}} \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}_{\theta}} \tilde{p}([0, T_1]|(\tilde{x}^0, 0)) \left[ T_2 \wedge \tilde{\theta}^*(\tilde{x}^0) - T_1 \right] \hat{\nu}(d\tilde{x}^0) \\
&\quad - \int_{\tilde{\Gamma}_{\theta}} \int_{(T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} \tilde{p}([T_1, u]|(\tilde{x}^0, 0)) du \hat{\nu}(d\tilde{x}^0).
\end{aligned}$$

Since  $Y_{T_1, T_2, \tilde{\Gamma} \setminus \tilde{\Gamma}_{\theta}} = \emptyset$ ,

$$\check{\eta}_{\square}(Y_{T_1, T_2, \tilde{\Gamma} \setminus \tilde{\Gamma}_{\theta}}) - \check{\eta}'_{\square}(Y_{T_1, T_2, \tilde{\Gamma} \setminus \tilde{\Gamma}_{\theta}}) = 0.$$

It remains to consider the set  $\tilde{\Gamma}_{\theta}$ . Below, we split it into three measurable subsets:

$$\begin{aligned}
\tilde{\Gamma}_1 &:= \tilde{\Gamma}_{\theta} \cap \{\tilde{x}^0 : u^*(\tilde{x}^0) < T_1\}, \\
\tilde{\Gamma}_2 &:= \tilde{\Gamma}_{\theta} \cap \{\tilde{x}^0 : u^*(\tilde{x}^0) \geq T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)\}, \\
\text{and } \tilde{\Gamma}_3 &:= \tilde{\Gamma}_{\theta} \cap \{\tilde{x}^0 : T_1 \leq u^*(\tilde{x}^0) < T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)\}.
\end{aligned}$$

For each  $\tilde{x}^0 \in \tilde{\Gamma}_1$ ,  $\tilde{p}([0, u]|(\tilde{x}^0, 0)) = 1$  for all  $u \in [T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]$ . Hence, according to (38) with  $\tilde{\Gamma} = \tilde{\Gamma}_1$ ,  $\check{\eta}'_{\square}(Y_{T_1, T_2, \tilde{\Gamma}_1}) = 0$  and

$$\check{\eta}_{\square}(Y_{T_1, T_2, \tilde{\Gamma}_1}) - \check{\eta}'_{\square}(Y_{T_1, T_2, \tilde{\Gamma}_1}) \geq 0.$$

For each  $\tilde{x}^0 \in \tilde{\Gamma}_2$  (see the point  $\tilde{x}_1^0$  on Figure 4),

$$\tilde{p}([0, T_1]|(\tilde{x}^0, 0)) = \check{\eta}_A([0, T_1]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \quad \text{and} \quad \tilde{p}([T_1, u]|(\tilde{x}^0, 0)) = \check{\eta}_A([T_1, u]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)}$$

for all  $u \in (T_1, T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]$ . Therefore, by (37) and (38),

$$\check{\eta}_{\square}(Y_{T_1, T_2, \tilde{\Gamma}_2}) - \check{\eta}'_{\square}(Y_{T_1, T_2, \tilde{\Gamma}_2}) = \mathbf{J}(\tilde{\Gamma}_2) \geq 0.$$

For the set  $\tilde{\Gamma}_3$  (the typical points in  $\tilde{\Gamma}_3$  are  $\tilde{x}_2^0$  and  $\tilde{x}_3^0$  on Figure 4), we compute  $\check{\eta}_\square(Y_{T_1, T_2, \tilde{\Gamma}_3})$  and  $\check{\eta}'_\square(Y_{T_1, T_2, \tilde{\Gamma}_3})$  using the representation  $Y_{T_1, T_2, \tilde{\Gamma}_3} = Y^1 \cup Y^2$ , where

$$Y^1 := \{(\tilde{x}^0, u) : \tilde{x}^0 \in \tilde{\Gamma}_3, T_1 \leq u < u^*(\tilde{x}^0)\}; \quad Y^2 := \{(\tilde{x}^0, u) : \tilde{x}^0 \in \tilde{\Gamma}_3, u^*(\tilde{x}^0) \leq u < T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)\}.$$

To compute  $\check{\eta}_\square(Y^1)$ , we introduce the function

$$w(y) = w((\tilde{y}, t)) = \mathbb{I}\{h(y) \in \tilde{\Gamma}_3\} \times \begin{cases} u^*(h(y)) - T_1, & \text{if } t \leq T_1; \\ u^*(h(y)) - t, & \text{if } T_1 < t \leq u^*(h(y)); \\ 0 & \text{if } t > u^*(h(y)) \end{cases}$$

(cf (36)). Calculations similar to those presented above, lead to the following version of expression (37):

$$\begin{aligned} \check{\eta}_\square(Y^1) &= \int_{\tilde{\Gamma}_3} [u^*(\tilde{x}^0) - T_1] \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}_3} [u^*(\tilde{x}^0) - T_1] \check{\eta}_A([0, T_1]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) \\ &\quad - \int_{\tilde{\Gamma}_3} \int_{(T_1, u^*(\tilde{x}^0)]} \check{\eta}_A([T_1, u]|\tilde{x}^0) du \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \hat{\nu}(d\tilde{x}^0) + \mathbf{J}^1. \end{aligned}$$

The last term is similar to  $\mathbf{J}(\tilde{\Gamma})$ , its calculation is based on the function similar to  $w_{T_1, T_2, \tilde{\Gamma}}^A$ : one only has to replace  $\tilde{\Gamma}$  with  $\tilde{\Gamma}_3$  and  $\tilde{\theta}^*(\cdot)$  with  $u^*(\cdot)$ . Like previously,  $\mathbf{J}^1 \geq 0$ . Again, similarly to (38), we have

$$\begin{aligned} \check{\eta}'_\square(Y^1) &= \int_{\tilde{\Gamma}_3} [u^*(\tilde{x}^0) - T_1] \hat{\nu}(d\tilde{x}^0) - \int_{\tilde{\Gamma}_3} [u^*(\tilde{x}^0) - T_1] \tilde{p}([0, T_1]|(\tilde{x}^0, 0)) \hat{\nu}(d\tilde{x}^0) \\ &\quad - \int_{\tilde{\Gamma}_3} \int_{(T_1, u^*(\tilde{x}^0)]} \tilde{p}([T_1, u]|(\tilde{x}^0, 0)) du \hat{\nu}(d\tilde{x}^0) \end{aligned}$$

and, like in the case of  $\tilde{\Gamma}_2$ , for each  $\tilde{x}_0 \in \tilde{\Gamma}_3$

$$\tilde{p}([0, T_1]|(\tilde{x}^0, 0)) = \check{\eta}_A([0, T_1]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)} \quad \text{and} \quad \tilde{p}([T_1, u]|(\tilde{x}^0, 0)) = \check{\eta}_A([T_1, u]|\tilde{x}^0) \frac{a(\tilde{x}^0)}{n(\tilde{x}^0)}$$

for all  $u \in (T_1, u^*(\tilde{x}^0)]$ . Therefore,

$$\check{\eta}_\square(Y^1) - \check{\eta}'_\square(Y^1) = \mathbf{J}^1 \geq 0.$$

Finally, similarly to (38),

$$\check{\eta}'_\square(Y^2) = \int_{\tilde{\Gamma}_3} \int_{[u^*(\tilde{x}^0), T_2 \wedge \tilde{\theta}^*(\tilde{x}^0)]} (1 - \tilde{p}([0, u]|(\tilde{x}^0, 0))) du \hat{\nu}(d\tilde{x}^0) = 0$$

because for each  $\tilde{x}^0 \in \tilde{\Gamma}_3$ ,  $\tilde{p}([0, u]|(\tilde{x}^0, 0)) = 1$  for all  $u > u^*(\tilde{x}^0)$ . Hence,

$$\check{\eta}_\square(Y^2) - \check{\eta}'_\square(Y^2) \geq 0.$$

To summarize,  $\check{\eta}_\square(Y_{T_1, T_2, \tilde{\Gamma}_3}) - \check{\eta}'_\square(Y_{T_1, T_2, \tilde{\Gamma}_3}) \geq 0$ , and thus  $\check{\eta}_\square(Y_{T_1, T_2, \tilde{\Gamma}}) - \check{\eta}'_\square(Y_{T_1, T_2, \tilde{\Gamma}}) \geq 0$  for all  $\tilde{\Gamma} \in \mathcal{B}(\tilde{V})$  and  $0 \leq T_1 < T_2 < \infty$ .

Therefore,  $\check{\eta}_\square \geq \check{\eta}'_\square$  set-wise on  $\mathbf{D}$ , and hence  $\check{\eta}_\square \geq \check{\eta}'_\square$  on  $V$ . Since the measures  $\check{\eta}_\square$  and  $\check{\eta}'_\square$  are both normal, the difference  $\check{\eta}_\square - \check{\eta}'_\square$  is a normal measure on  $V$  by Lemma 4.2(b).

The proof is completed.  $\square$

Proof of Theorem 4.2. When  $x \in V^c$ , we fix  $\pi_i(d\theta \times da) := \delta_{f^*(x)}(d\theta \times da)$ , where  $f^*(x) = (\infty, \hat{a})$  as usual. Below, for two finite or normal measures  $\zeta^1$  and  $\zeta^2$  on  $V$ , the inequality  $\zeta^1(dx) \leq \zeta^2(dx)$  is understood set-wise. The same concerns measures on  $V \times \mathbf{A}$ .

Let  $p'_A(da|x)$  be the stochastic kernel on  $\mathbf{A}$  given  $V$  coming from the decomposition  $\eta(dx \times da) = p'_A(da|x)\eta(dx \times \mathbf{A})$ . For all  $i \geq 1$ , we put

$$p_A^i(da|x, \theta) \equiv p'_A(da|\phi(x, \theta))$$

for  $x \in V$ ,  $\theta < \theta^*(x)$ , and  $p_A^i(da|x, \theta)$  is an arbitrarily fixed stochastic kernel on  $\mathbf{A}$  for  $x \in V$ ,  $\theta \geq \theta^*(x)$ .

We will prove by induction the following statement.

For each  $i \geq 1$ , there is a stochastic kernel  $\pi_i$  on  $\mathbf{B} = \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  given  $V$ , having the form

$$\pi_i(d\theta \times da|x) = p_T^i(d\theta|x)p_A^i(da|x, \theta),$$

such that, for each  $n \geq 1$  and the sequence  $\{\pi_i\}_{i=1}^n$ , the following assertions are fulfilled.

(i)  $p_T^i([\theta^*(x), \infty)|x) = 0$  for  $x \in V$ ,  $i = 1, 2, \dots, n$ , and the (partial) aggregated occupation measures  $\{\tilde{\eta}^i\}_{i=0}^n$ , defined as in Lemma 7.1, exhibit the following properties:

$$\begin{aligned} \tilde{\eta}^n(dx \times \square) &\leq \eta(dx \times \square) \quad \text{and} \\ \tilde{\eta}^n(dx \times da) = \tilde{\eta}^n(dx \times \mathbf{A})p'_A(da|x) &\leq \eta(dx \times \mathbf{A})p'_A(da|x) = \eta(dx \times da) \text{ on } \mathcal{B}(V \times \mathbf{A}). \end{aligned}$$

(ii) The measure  $\nu^n(dx) := P_{x_0}^\pi(X_n \in dx)$  on  $V$  is such that, for each function  $w \in \mathbf{W}$ ,

$$\begin{aligned} 0 &= \int_V w(x)\nu^n(dx) + \int_V \chi w(x)[\eta - \tilde{\eta}^n](dx \times \square) - \int_V w(x)[\eta - \tilde{\eta}^n](dx \times \mathbf{A}) \\ &\quad + \int_{V \times \mathbf{A}} w(l(x, a))[\eta - \tilde{\eta}^n](dx \times da), \end{aligned} \quad (39)$$

and all the integrals here are finite. Note that  $\nu^n$  is uniquely defined by the finite sequence  $\{\pi_i\}_{i=1}^n$ : see (4); moreover,  $\nu^n(\tilde{V} \times \{t : t > 0\}) = 0$ .

After that,  $\pi^n := \{\pi_i\}_{i=1}^\infty$  will be the desired Markov strategy.

When  $n = 0$ ,  $\tilde{\eta}^0(dy \times \square) \equiv 0$ ,  $\tilde{\eta}^0(dy \times da) \equiv 0$ , and  $\nu^0(dx) = \delta_{x_0}(dx)$ . Assertions (i) and (ii) are obviously fulfilled because the normal measure  $\eta$  satisfies equation (21).

Suppose assertions (i) and (ii) hold true for  $i = 0, 1, 2, \dots, n \geq 0$ . We apply Lemma 7.4 to the measures  $\nu := \nu^n$ ,  $\tilde{\eta}_\square := (\eta - \tilde{\eta}^n)(dx \times \square)$ ,  $\tilde{\eta}_A := (\eta - \tilde{\eta}^n)(dx \times \mathbf{A})$ , and  $\tilde{\eta} := (\eta - \tilde{\eta}^n)(dx \times da)$  satisfying equation (39). All of them are finite, maybe apart from  $\tilde{\eta}_\square$ , which is normal by Lemma 4.2(b) and Lemma 7.1. As a result, we have the stochastic kernel  $\tilde{p}(dt|x)$  on  $\bar{\mathbb{R}}_+^0$  given  $V$  and the measures

$$\tilde{\eta}'_A(dx) \leq (\eta - \tilde{\eta}^n)(dx \times \mathbf{A}) \quad \text{and} \quad \tilde{\eta}'_\square(dx) \leq (\eta - \tilde{\eta}^n)(dx \times \square) \quad (40)$$

on  $V$ , which satisfy equation (33):

$$0 = \int_V w(x)\nu^n(dx) + \int_V \chi w(x)\tilde{\eta}'_\square(dx) - \int_V w(x)\tilde{\eta}'_A(dx), \quad w \in \mathbf{W}. \quad (41)$$

All the integrals here are finite.

For  $x \in V$ , we put

$$p_T^{n+1}(d\theta|x) := \tilde{p}(d\theta|x).$$

Then by Lemma 7.4,  $p_T^{n+1}([\theta^*(x), \infty]|x) = 0$  for all  $x \in V$ . All the kernels  $\{\pi_i\}_{i=1}^n$  were built on the previous steps of the induction. According to the definition of the measure  $\tilde{\eta}^{n+1}$ ,

$$\begin{aligned}\tilde{\eta}^{n+1}(\Gamma \times \square) &= \tilde{\eta}^n(\Gamma \times \square) + \int_V \int_{\mathbb{R}_+^0} \delta_{\phi(x,u)}(\Gamma) \tilde{p}([u, \infty]|x) du \nu^n(dx) \\ &= \tilde{\eta}^n(\Gamma \times \square) + \tilde{\eta}'_{\square}(\Gamma) \leq \eta(\Gamma \times \square), \quad \Gamma \in \mathcal{B}(V); \end{aligned} \quad (42)$$

$$\begin{aligned}\tilde{\eta}^{n+1}(\Gamma \times \mathbf{A}) &= \tilde{\eta}^n(\Gamma \times \mathbf{A}) + \int_V \int_{\mathbb{R}_+^0} \delta_{\phi(x,u)}(\Gamma) \tilde{p}(du|x) \nu^n(dx) = \tilde{\eta}^n(\Gamma \times \mathbf{A}) + \tilde{\eta}'_{\mathbf{A}}(\Gamma) \leq \eta(\Gamma \times \mathbf{A}), \\ &\Gamma \in \mathcal{B}(V). \end{aligned} \quad (43)$$

Inequalities are valid according to the basic properties of the measures  $\tilde{\eta}'_{\square}$  and  $\tilde{\eta}'_{\mathbf{A}}$  presented in (40). Recall that

$$\begin{aligned}\tilde{\eta}^{n+1}(\Gamma_X \times \Gamma_A) &= \tilde{\eta}^n(\Gamma_X \times \Gamma_A) + \int_V \int_{\mathbb{R}_+^0} \delta_{\phi(x,\theta)}(\Gamma_X) p_A^{n+1}(\Gamma_A|x, \theta) p_T^{n+1}(d\theta|x) \nu^n(dx), \\ &\Gamma_X \in \mathcal{B}(V), \Gamma_A \in \mathcal{B}(\mathbf{A}). \end{aligned}$$

Since  $\nu^n(\tilde{V} \times \{t : t > 0\}) = 0$ , the last term equals

$$I := \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \delta_{\phi((\tilde{x}^0, 0), \theta)}(\Gamma_X) p'_A(\Gamma_A|\phi((\tilde{x}^0, 0), \theta)) \tilde{p}(d\theta|(\tilde{x}^0, 0)) \hat{\nu}^n(d\tilde{x}^0), \quad (44)$$

where  $\hat{\nu}^n(\Gamma) := \nu^n(\{(\tilde{x}^0, 0), \tilde{x}^0 \in \Gamma\})$ . According to Lemma 7.4, for all  $\Gamma \in \mathcal{B}(\mathbf{D})$  and for the mapping  $F$  as in Definition 4.2,

$$\begin{aligned}\tilde{\eta}'_{\mathbf{A}}(F(\Gamma)) &= \int_{\tilde{V}} \int_{\mathbb{R}_+^0} \mathbb{I}\{\phi((\tilde{x}^0, 0), u) \in \{y = \phi((\tilde{x}^0, 0), t) : (\tilde{x}^0, t) \in \Gamma\}\} \tilde{p}(du|(\tilde{x}^0, 0)) \hat{\nu}^n(d\tilde{x}^0) \\ &= \int_{\tilde{V}} \int_{\mathbb{R}_+^0} \mathbb{I}\{\tilde{x}^0, u \in \Gamma\} \tilde{p}(du|(\tilde{x}^0, 0)) \hat{\nu}^n(d\tilde{x}^0) = \check{\eta}'_{\mathbf{A}}(\Gamma). \end{aligned}$$

Lemma 7.2(a) implies that, for each bounded measurable function  $g$  on  $V$ ,

$$\int_V g(x) \tilde{\eta}'_{\mathbf{A}}(dx) = \int_{\mathbf{D}} g(\phi((\tilde{x}^0, 0), u)) \check{\eta}'_{\mathbf{A}}(d\tilde{x}^0 \times du) = \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} g(\phi((\tilde{x}^0, 0), u)) \tilde{p}(du|(\tilde{x}^0, 0)) \hat{\nu}^n(d\tilde{x}^0). \quad (45)$$

Therefore, for each  $\Gamma_X \in \mathcal{B}(V)$ ,

$$I = \int_V \delta_x(\Gamma_X) p'_A(\Gamma_A|x) \tilde{\eta}'_{\mathbf{A}}(dx) = \int_{\Gamma_X} p'_A(\Gamma_A|x) \tilde{\eta}'_{\mathbf{A}}(dx),$$

meaning that on  $\mathcal{B}(V \times \mathbf{A})$

$$\begin{aligned}\tilde{\eta}^{n+1}(dx \times da) &= \tilde{\eta}^n(dx \times da) + p'_A(da|x) \tilde{\eta}'_{\mathbf{A}}(dx) = \tilde{\eta}^n(dx \times \mathbf{A}) p'_A(da|x) + \tilde{\eta}'_{\mathbf{A}}(dx) p'_A(da|x) \\ &= \tilde{\eta}^{n+1}(dx \times \mathbf{A}) p'_A(da|x) \\ &\leq \tilde{\eta}^n(dx \times \mathbf{A}) p'_A(da|x) + [\eta - \tilde{\eta}^n](dx \times \mathbf{A}) p'_A(da|x) = \eta(dx \times \mathbf{A}) p'_A(da|x). \end{aligned}$$

The second equality is by the inductions supposition, the third equality follows from (43), and the inequality is according to the basic property (40) of the measure  $\tilde{\eta}'_{\mathbf{A}}$ .

Property (i) for  $n + 1$  is established, recall also inequality (42).

For the proof of Item (ii), note that, by (39) at  $n$ , (41), (42), and (43), we have equation

$$\begin{aligned} 0 &= \int_V \chi w(x)[\eta - \tilde{\eta}^{n+1}](dx \times \square) - \int_V w(x)[\eta - \tilde{\eta}^{n+1}](dx \times \mathbf{A}) \\ &\quad + \int_{V \times \mathbf{A}} w(l(x, a))[\eta - \tilde{\eta}^{n+1}](dx \times da) + \int_{V \times \mathbf{A}} w(l(x, a))[\tilde{\eta}^{n+1} - \tilde{\eta}^n](dx \times da) \end{aligned}$$

valid for all functions  $w \in \mathbf{W}$ , and all the integrals here are finite. According to property (i) for  $n$  and  $n + 1$ , the stochastic kernel  $p'_A(da|x)$  is the same in the decompositions  $\tilde{\eta}^n(dx \times da) = \tilde{\eta}^n(dx \times \mathbf{A})p'_A(da|x)$  and  $\tilde{\eta}^{n+1}(dx \times da) = \tilde{\eta}^{n+1}(dx \times \mathbf{A})p'_A(da|x)$ . Thus, the last integral, according to (43), equals

$$\int_V \int_{\mathbf{A}} w(l(x, a))p'_A(da|x)\tilde{\eta}'_A(dx),$$

i.e., the function  $w$  is integrated with respect to the measure

$$m(\Gamma) = \int_V \int_{\mathbf{A}} \delta_{l(x, a)}(\Gamma)p'_A(da|x)\tilde{\eta}'_A(dx), \quad \Gamma \in \mathcal{B}(V),$$

and it remains to show that this measure coincides with  $\nu^{n+1}$  on  $V$ .

From equation (45), we have for all  $\Gamma \in \mathcal{B}(V)$ :

$$m(\Gamma) = \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \int_{\mathbf{A}} \delta_{l(\phi((\tilde{x}^0, 0), u), a)}(\Gamma)p'_A(da|\phi((\tilde{x}^0, 0), u))p_T^{n+1}(du|(\tilde{x}^0, 0))\hat{\nu}^n(d\tilde{x}^0),$$

and, keeping in mind that  $\nu^n(\tilde{V} \times \{t : t > 0\}) = 0$ , we have from (4):

$$\nu^{n+1}(\Gamma) = \int_{\tilde{V}} \int_{[0, \tilde{\theta}^*(\tilde{x}^0)]} \int_{\mathbf{A}} \delta_{l(\phi((\tilde{x}^0, 0), \theta), a)}(\Gamma)p_A^{n+1}(da|(\tilde{x}^0, 0), \theta)p_T^{n+1}(d\theta|(\tilde{x}^0, 0))\hat{\nu}^n(d\tilde{x}^0) = m(\Gamma)$$

for all  $\Gamma \in \mathcal{B}(V)$  because  $p_A^{n+1} = p'_A$ .

The proof of the induction statement for  $n + 1$  is completed.

According to Lemma 7.1, for the constructed Markov strategy  $\pi = \{\pi_i\}_{i=1}^\infty$  and for the corresponding aggregated occupation measure  $\tilde{\eta}$ , we have the convergence  $\tilde{\eta}^n \uparrow \tilde{\eta}$  set-wise as  $n \rightarrow \infty$ . Since  $\tilde{\eta}^n \leq \eta$  set-wise on  $V \times \mathbf{A}_\square$ , the desired set-wise inequality  $\tilde{\eta} \leq \eta$  follows.

All the properties enlisted in Definition 3.1 are obviously satisfied for the strategy  $\pi$ .  $\square$

Proof of Corollary 4.1. We denote by  $Val(16)$  and  $Val(20)$  the minimal values of linear programs (16) and (20),(21), respectively. Recall that linear program (16) has an optimal solution by Proposition 3.1.

Suppose the finite measure  $\mu^*$  on  $V \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$  (concentrated on  $\mathbf{M} \times \mathbf{A}$ ) solves linear program (16). Then the aggregated occupation measure  $\eta^*$ , induced by  $\mu^*$ , is normal by Lemma 4.2(a) and satisfies equation (21) according to Theorem 4.1. The constraints-inequalities are also fulfilled by  $\eta^*$ . Thus

$$\infty > Val(16) = \int_{V \times \mathbf{A}_\square} C_0(x, a)\eta^*(dx \times da) \geq Val(20).$$

In case the last inequality is strict, there exists a feasible solution  $\eta$  to linear program (20),(21) satisfying inequality

$$\int_{V \times \mathbf{A}_\square} C_0(x, a)\eta(dx \times da) < Val(16).$$



Consider the induced reasonable Markov strategy  $\pi^\eta$  as in Theorem 4.2 and the corresponding aggregated occupation measure  $\tilde{\eta}$ . Since  $C_j \geq 0$  for  $j = 0, 1, \dots, J$ , all the conditions in linear program (16) are satisfied for  $\mu^{\pi^\eta}$  and

$$\int_{V \times \mathbf{A}_\square} C_0(x, a) \tilde{\eta}(dx \times da) \leq \int_{V \times \mathbf{A}_\square} C_0(x, a) \eta(dx \times da) < Val(16) < \infty.$$

The measure  $\mu^{\pi^\eta}$  cannot take infinite value as explained above linear program (10), (11). We obtained a contradiction to the optimality of the measure  $\mu^*$ . Hence,  $Val(16) = Val(20)$ , and the measure  $\eta^*$  solves linear program (20),(21).

Suppose now that the measure  $\eta^*$  on  $V \times \mathbf{A}_\square$  solves linear program (20),(21) and consider the reasonable Markov strategy  $\pi^* = \pi^{\eta^*}$  as in Theorem 4.2. The corresponding occupation measure  $\mu^{\pi^*}$  is feasible in linear program (16). More detailed reasoning is similar to that presented above. Therefore, for the aggregated occupation measure  $\tilde{\eta}$  induced by  $\mu^{\pi^*}$ , we have relations

$$Val(16) \leq \int_{V \times \mathbf{A}_\square} C_0(x, a) \tilde{\eta}(dx \times da) \leq \int_{V \times \mathbf{A}_\square} C_0(x, a) \eta^*(dx \times da) = Val(20).$$

But we have shown that  $Val(20) = Val(16)$ , so that

$$\int_{V \times \mathbf{A}_\square} C_0(x, a) \tilde{\eta}(dx \times da) = Val(16)$$

meaning that the measure  $\mu^{\pi^*}$  solves linear program (16). The proof is completed.  $\square$

## 8 Acknowledgement

This research was supported by the Royal Society International Exchanges award IE160503. We would like to thank Prof.A.Plakhov for his initial participation in this work and for his proof of Lemma A.1.

## A Appendix

Lemma A.1 and its proof presented below are similar to Lemma 2.2 in [12], where the authors assumed that  $E$  was a subset of an Euclidean space.

Let  $E$  be an arbitrary set and  $\phi : E \times \mathbb{R}_+^0 \rightarrow E$  be a flow in  $E$  possessing the semigroup property.

**Definition A.1** *A function  $w : E \rightarrow \mathbb{R}$  is said to be absolutely continuous along the flow if for all  $x \in E$  the function  $t \mapsto w(\phi(x, t))$ ,  $t \in \mathbb{R}_+^0$  is absolutely continuous. It is called increasing (decreasing) along the flow if so is the function  $t \mapsto w(\phi(x, t))$ ,  $t \in \mathbb{R}_+^0$  for all  $x \in E$ .*

**Lemma A.1** *Suppose function  $w$  is absolutely continuous along the flow  $\phi$ . Then the following assertions are valid.*

(a) *There exists a function  $\chi w : E \rightarrow \mathbb{R}$  such that, for any  $x \in E$ , the function  $\chi w(\phi(x, s))$  is Lebesgue integrable with respect to  $s$  on any finite interval  $[0, t] \subset \mathbb{R}_+^0$  and*

$$w(\phi(x, t)) - w(x) = \int_{[0, t]} \chi w(\phi(x, s)) ds \quad (46)$$

for all  $x \in E$  and  $t \geq 0$ .

(b) *If, additionally,  $E$  is a measurable space (that is, is equipped with a  $\sigma$ -algebra of subsets),  $w$  is measurable, and the functions  $\phi(\cdot, t) : E \rightarrow E$  are measurable for all  $t \geq 0$ , then the function  $\chi w$  satisfying (a) can be chosen measurable.*

Proof. We provide one common proof for (a) and (b) underlining the measurability properties as soon as they appear.

Define the functions

$$\overline{W}(x) := \overline{\lim}_{n \rightarrow \infty} \frac{w(\phi(x, \frac{1}{n})) - w(x)}{1/n}, \quad \underline{W}(x) := \underline{\lim}_{n \rightarrow \infty} \frac{w(\phi(x, \frac{1}{n})) - w(x)}{1/n}$$

and the set  $D := \{x \in E : \overline{W}(x) = \underline{W}(x) \neq \pm\infty\}$ . Let us additionally define the function  $W : D \rightarrow \mathbb{R}$  by  $W(x) := \overline{W}(x)$ ; that is,  $W(x)$  coincides with the limit  $\lim_{n \rightarrow \infty} n [w(\phi(x, \frac{1}{n})) - w(x)]$ , if it exists and is finite.

If  $w$  and  $\phi(\cdot, t)$  are measurable, then  $w(\phi(x, \frac{1}{n}))$  is also measurable. Hence the functions  $\overline{W}$  and  $\underline{W}$  are measurable as the upper and lower limits of the sequence of measurable functions  $n [w(\phi(x, 1/n)) - w(x)]$ . Consequently, the set  $D$  is also measurable.

Define the function  $\chi w$  on  $E$  by

$$\chi w(x) := \begin{cases} W(x), & \text{if } x \in D; \\ g(x), & \text{otherwise,} \end{cases} \quad (47)$$

where  $g$  is any function. In the measurable case we take  $g$  to be measurable and readily get that  $\chi w$  is also measurable.

Since  $w$  is absolutely continuous along the flow then for any  $x \in E$  there exists a subset of full measure  $T_x \subset \mathbb{R}_+$  such that the derivative  $\frac{d}{dt} w(\phi(x, t))$  exists and is finite for all values  $t \in T_x$ . For any such value (let it now be denoted by  $s \in T_x$ ) we can write down the following (below we denote  $x' = \phi(x, s)$  and use the semigroup property of the flow)

$$\left. \frac{dw(\phi(x, t))}{dt} \right|_{t=s} = \lim_{\varepsilon \rightarrow 0} \frac{w(\phi(x, s + \varepsilon)) - w(\phi(x, s))}{\varepsilon} = \lim_{n \rightarrow \infty} \frac{w(\phi(x', \frac{1}{n})) - w(x')}{1/n}.$$

The latter value exists and is finite, and therefore coincides with  $W(x')$ . This argument also shows that  $\phi(x, T_x) \subset D$ .

Since  $w$  is absolutely continuous along the flow, one can write down

$$w(\phi(x, t)) - w(x) = \int_{[0, t] \cap T_x} \left. \frac{dw(\phi(x, \tau))}{d\tau} \right|_{\tau=s} ds = \int_{[0, t] \cap T_x} W(\phi(x, s)) ds.$$

Now taking into account that  $[0, t] \setminus T_x$  has Lebesgue measure zero and  $\chi w$  is an extension of  $W$  to  $E$ , we conclude that the latter integral coincides with  $\int_{[0, t]} \chi w(\phi(x, s)) ds$ , and so, formula (46) is proved.  $\square$

Proof of Lemma 4.1. In this proof, let us denote by  $\rho$  and  $\tilde{\rho}$  the compatible metrics on  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$  and  $\tilde{\mathbf{X}}$ . If  $y_n \rightarrow y$ , where  $y_n = (\tilde{y}_n, t_n)$ ,  $y = (\tilde{y}, t) \in \mathbf{X}$ , then the sequence  $\{t_n\}_{n=1}^\infty$  is bounded:  $t, t_n \in [0, T]$  for some  $T < \infty$ . Now  $\tilde{\rho}(h(y_n), h(y)) \leq \sup_{t \in [0, T]} d(t)\rho(y_n, y) \rightarrow 0$ . Thus,  $h$  is continuous. The continuity of the mapping  $h$  and of the original flow  $\tilde{\phi}$  immediately implies that the flows  $\tilde{\phi}$  and  $\phi$  in the reverse time are continuous.

The mapping  $F$  is continuous because the flow  $\tilde{\phi}$  is continuous. It is a bijection from  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$  to  $\mathbf{X}$ , and the inverse mapping  $F^{-1}(y) = (h(y), \tau_y)$  is continuous, as has been proved above. (For  $y = (\tilde{y}, t) \in \mathbf{X}$ ,  $\tau_y = t$  is obviously a continuous function of  $y$ .) Thus,  $F$  is a homeomorphism, and  $\mathbf{X}$  is a Borel space, being the homeomorphic image of the Borel space  $\tilde{\mathbf{X}} \times \mathbb{R}_+^0$ . See also [5, Prop.7.15].  $\square$

Proof of Lemma 4.2. (a) The measure  $\eta$  is finite on  $V \times \mathbf{A}$  because the measure  $\mu$  is finite. Recall that the measure  $\mu$  is concentrated on  $\tilde{V} \times \{0\} \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}$ . For the measure  $\eta(dx \times \square)$  on  $V$ , we have

$$\begin{aligned} \check{\eta}(\Gamma) &= \eta(F(\Gamma) \times \square) = \int_{\mathbb{R}_+^0} \left\{ \int_{\tilde{V}} \mathbb{I}\{\phi((\tilde{x}, 0), u) \in F(\Gamma)\} \mu(d\tilde{x} \times \{0\} \times [u, \infty] \times \mathbf{A}) \right\} du \\ &= \int_{\mathbb{R}_+^0} \left\{ \int_{\tilde{V}} \mathbb{I}\{(\tilde{x}, u) \in \Gamma\} \mu(d\tilde{x} \times \{0\} \times [u, \infty] \times \mathbf{A}) \right\} du \\ &= \int_{\tilde{V}} \int_{\mathbb{R}_+^0} \mathbb{I}\{(\tilde{x}, u) \in \Gamma\} \int_{\mathbf{A}} p_T([u, \infty] | (\tilde{x}, 0), a) p_A(da | (\tilde{x}, 0)) du \mu(d\tilde{x} \times \{0\} \times \bar{\mathbb{R}}_+^0 \times \mathbf{A}), \end{aligned}$$

for all  $\Gamma \in \mathcal{B}(\mathbf{D})$ : see (13) and (14). Thus,  $\eta$  on  $V \times \mathbf{A}_\square$  is normal.

(b) As mentioned in the proof of part (a),  $\eta(V \times \mathbf{A}) < \infty$ .

Consider the measures  $\zeta^1(dy) := \eta^1(dy \times \square)$ ,  $\zeta^2(dy) := \eta^2(dy \times \square)$ , and  $\zeta(dy) := \eta(dy \times \square)$ . Since  $\zeta \geq 0$ , we have  $\check{\zeta} = \zeta^1 - \zeta^2 \geq 0$  as well. If

$$\check{\zeta}^1(d\tilde{x}^0 \times du) = g^1(\tilde{x}^0, u) du L^1(d\tilde{x}^0) \quad \text{and} \quad \check{\zeta}^2(d\tilde{x}^0 \times du) = g^2(\tilde{x}^0, u) du L^2(d\tilde{x}^0),$$

then we put  $L := L^1 + L^2$  and

$$g(\tilde{x}^0, u) := \left[ g^1(\tilde{x}^0, u) \frac{dL^1}{dL}(\tilde{x}^0) - g^2(\tilde{x}^0, u) \frac{dL^2}{dL}(\tilde{x}^0) \right] \mathbb{I} \left\{ g^1(\tilde{x}^0, u) \frac{dL^1}{dL}(\tilde{x}^0) - g^2(\tilde{x}^0, u) \frac{dL^2}{dL}(\tilde{x}^0) \geq 0 \right\}.$$

The measurable set

$$\Gamma := \left\{ (\tilde{x}^0, u) : \tilde{x}^0 \in \tilde{V}, 0 \leq u < \tilde{\theta}^*(\tilde{x}^0), g^1(\tilde{x}^0, u) \frac{dL^1}{dL}(\tilde{x}^0) - g^2(\tilde{x}^0, u) \frac{dL^2}{dL}(\tilde{x}^0) < 0 \right\} \subset \mathbf{D}$$

is null with respect to the measure  $L(d\tilde{x}^0) \times du$  because, otherwise, we would have for some  $t < \infty$ ,  $\int_{\Gamma_t} L(d\tilde{x}^0) \times du > 0$  for the set

$$\Gamma_t := \Gamma \cap \{(\tilde{x}^0, u) : \tilde{x}^0 \in \tilde{V}, u \leq t\},$$

and yield a desired contradiction:

$$0 > \int_{\Gamma_t} \left[ g^1(\tilde{x}^0, u) \frac{dL^1}{dL}(\tilde{x}^0) - g^2(\tilde{x}^0, u) \frac{dL^2}{dL}(\tilde{x}^0) \right] L(d\tilde{x}^0) \times du = \check{\zeta}^1(\Gamma_t) - \check{\zeta}^2(\Gamma_t) = \check{\zeta}(\Gamma_t) \geq 0$$

with all the terms being finite. Now, for  $\zeta = \zeta^1 - \zeta^2$ , we have

$$\check{\zeta}(d\tilde{x}^0 \times du) = \check{\zeta}^1(d\tilde{x}^0 \times du) - \check{\zeta}^2(d\tilde{x}^0 \times du) = g(\tilde{x}^0, u) du L(d\tilde{x}^0),$$

and the proof is completed. □

**Lemma A.2** *Suppose  $m$  is a finite measure on  $\mathbb{R}$ . Then, for each  $\tau, t \in \mathbb{R}$ ,*

$$tm([\tau, \tau + t]) = \int_{(0, t]} m([\tau, \tau + s]) ds + \int_{(\tau, \tau + t]} (s - \tau) dm(s)$$

and

$$tm([\tau, \tau + t]) = \int_{[0, t)} m([\tau, \tau + s]) ds + \int_{[\tau, \tau + t)} (s - \tau) dm(s).$$

Proof. For all cadlag (i.e., right-continuous with left limits) real-valued functions  $U$  and  $V$  on  $\mathbb{R}$  with finite variation (on finite intervals),

$$U(t_2)V(t_2) = U(t_1)V(t_1) + \int_{(t_1, t_2]} U(s-) dV(s) + \int_{(t_1, t_2]} V(s) dU(s) \quad (48)$$

for any  $-\infty < t_1 < t_2 < \infty$ . (See [8, Appendix A4, §2].) Equivalently, in the symmetric form:

$$U(t_2)V(t_2) = U(t_1)V(t_1) + \int_{(t_1, t_2]} U(s-)dV(s) + \int_{(t_1, t_2]} V(s-)dU(s) + \sum_{u \in (t_1, t_2]} \Delta U_u \Delta V_u.$$

Introduce cadlag functions of finite variation (on finite intervals):

$$U(s) := m([\tau, s]) \quad \text{and} \quad V(s) := s - \tau, \quad s \in \mathbb{R}.$$

Then, for  $t > 0$ , applying the previous formulae to  $t_1 = \tau$  and  $t_2 = \tau + t$ , we see that

$$\begin{aligned} tm([\tau, \tau + t]) &= \int_{(\tau, \tau+t]} m([\tau, s])ds + \int_{(\tau, \tau+t]} (s - \tau) dm(s) = \int_{(0, t]} m([\tau, \tau + s])ds + \int_{(\tau, \tau+t]} (s - \tau) dm(s) \\ &= \int_{[0, t]} m([\tau, \tau + s])ds + \int_{[\tau, \tau+t]} (s - \tau) dm(s). \end{aligned} \quad (49)$$

For the last equality to be proved, it is sufficient to consider a strictly increasing sequence  $t_i \uparrow t > 0$  and pass to the limit in (49). The case  $t \leq 0$  is trivial.  $\square$

## References

- [1] Aliprantis, Ch. and Border, K.C. (2006). *Infinite Dimensional Analysis*. Springer-Verlag, New York.
- [2] Altman, E. (1999). *Constrained Markov Decision Processes*. Chapman and Hall/CRC, Boca Raton.
- [3] Avrachenkov, K., Habachi, O., Piunovskiy, A. and Zhang, Y. (2015). Infinite horizon optimal impulsive control with applications to Internet congestion control, *Intern. J. of Control*, **88**, 703–716.
- [4] Barles, G. (1985). Deterministic impulse control problems. *SIAM J. Control Optim.* **23**, 419–432.
- [5] Bertsekas, D. and Shreve, S. (1978). *Stochastic Optimal Control*. Academic Press, New York.
- [6] Blaquière, A. (1985). Impulsive optimal control with finite or infinite time horizon. *J. Optim. Theory Appl.* **46**, 431–439.
- [7] Bogachev, V.I. (2007). *Measure Theory (Volumes 1 and 2)*. Springer-Verlag, Berlin.
- [8] Bremaud, P. (1981). *Point Processes and Queues*. Springer-Verlag, New York.
- [9] Brown, L. and Purves, R. (1973). Measurable selections of extrema. *Ann. Statist.* **1**, 902–912.
- [10] Clayes, M., Arzelier, D., Henrion, D. and Lasserre, J-B. (2014). Measures and LMIs for impulsive nonlinear optimal control, *IEEE Trans. on Automatic Control* **59**, 1374–1379.

- [11] Clayes, M., Henrion, D. and Kružík, M. (2017). Semi-definite relaxations for optimal control problems with oscillation and concentration effects. *ESAIM COCV* **23**, 95–117.
- [12] Costa, O. and Dufour, F. (2013). *Continuous Average Control of Piecewise Deterministic Markov Processes*. Springer Briefs in Mathematics.
- [13] Davis, M.H.A. (1993). *Markov Models and Optimization*. Chapman and Hall / CRC, Boca Raton.
- [14] Dufour, F., Horiguchi, M. and Piunovskiy, A. (2012). The expected total cost criterion for Markov decision processes under constraints: a convex analytic approach. *Adv. Appl. Probab.* **44**, 774–793.
- [15] Dufour, F., Horiguchi, M. and Piunovskiy, A. (2016). Optimal impulsive control of piecewise deterministic Markov processes. *Stochastics* **88**, 1073–1098.
- [16] Ethier, S. and Kurtz, T. (1986). *Markov Processes*. Wiley, New York.
- [17] Gaitsgory, V. and Quincampoix, M. (2009). Linear programming approach to deterministic infinite horizon optimal control problems with discounting. *SIAM J. Control Optim.* **48**, 2480–2512.
- [18] Henrion, D., Kružík, M. and Weisser, T. (2019). Optimal control problems with oscillations, concentrations and discontinuities. *Automatica* **103**, 159–165.
- [19] Hernández-Hernández, D., Hernández-Lerma, O. and Taksar, M. (1996). The linear programming approach to deterministic optimal control problems. *Applicationes Mathematicae* **24**, 17–33.
- [20] Hernández-Lerma, O. and Lasserre, J. (1996). *Discrete-Time Markov Control Processes*, Springer-Verlag, New York.
- [21] Hernández-Lerma, O. and Lasserre, J.B. (1999). *Further Topics on Discrete-Time Markov Control Processes*. Springer-Verlag, New York.
- [22] Himmelberg, C. and Parthasarathy, T. and Van Vleck, F. (1976). Optimal plans for dynamic programming problems, *Math. Oper. Res.* **1**, 390–394.
- [23] Hou, S.H. and Wong, K.H. (2011). Optimal impulsive control problem with application to human immunodeficiency virus treatment, *J. Optim. Theory Appl.* **151**, 385–401.
- [24] Lasserre, J., Henrion, D. Prieur, C. and Trélat, E. (2008). Nonlinear optimal control via occupation measures and LMI-relaxations. *SIAM J. Control Optim.* **47**, 1643–1666.
- [25] Leander, R., Lenhart, S. and Protopopescu, V. (2015). Optimal control of continuous systems with impulse controls, *Optim. Control Appl. Meth.* **36**, 535–549.
- [26] Liu, Y., Teo, K., Jennings, L. and Wang, S. (1998). On a class of optimal control problems with state jumps. *J. Optim. Theory Appl.* **98**, 65–82.
- [27] Miller, B. and Rubinovich, E. (2003). *Impulsive Control in Continuous and Discrete-Continuous Systems*. Springer, New York.
- [28] Piunovskiy, A. (1997). *Optimal Control of Random Sequences in Problems with Constraints*. Kluwer, Dordrecht.
- [29] Piunovskiy, A., Plakhov, A., Torres, D. and Zhang, Y. (2019). Optimal impulse control of dynamical systems. *SIAM J. Control Optim.* **57**, 2720–2752.

- [30] Piunovskiy, A. and Zhang, Y. (2021). Linear programming approach to optimal impulse control problems with functional constraints. *J. Math. Anal Appl.*, in press. <https://doi.org/10.1016/j.jmaa.2020.124817>.