



**IRWIN AND JOAN JACOBS**  
**CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGIES**

# **A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime**

**Rami Atar and Asaf Cohen**

**CCIT Report #875**  
**December 2014**

 Electronics  
Computers  
Communications

*DEPARTMENT OF ELECTRICAL ENGINEERING*  
*TECHNION - ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA 32000, ISRAEL*



# A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime\*

Rami Atar<sup>†</sup>      Asaf Cohen<sup>‡</sup>

October 14, 2014

## Abstract

We study a differential game that governs the moderate-deviation heavy-traffic asymptotics of a multiclass single-server queueing control problem with a risk-sensitive cost. We consider a cost set on a finite but sufficiently large time horizon, and show that this formulation leads to stationary feedback policies for the game. Several aspects of the game are explored, including its characterization via a (one-dimensional) free boundary problem, the semi-explicit solution of an optimal strategy, and the specification of a saddle point. We emphasize the analogy to the well-known Harrison-Taksar free boundary problem which plays a similar role in the diffusion-scale heavy-traffic literature.

**AMS subject classifications:** 49N70, 60F10, 60K25, 93E20

**Keywords:** multi-class single-server queue, moderate deviations, heavy traffic, risk-sensitive cost, the Harrison-Taksar free boundary problem

## 1 Introduction

This paper is concerned with the moderate-deviation (MD) scale heavy-traffic analysis of a queueing control problem. The model that is treated consists of a server that devotes its time to customers from a number of classes, where a decision maker (DM) controls the allocation of the server's effort to each of the classes. Customers of each class are kept in a buffer of finite length and so new arrivals of customers of a given class are lost when the corresponding buffer is full. The DM may reject arrivals even when buffers are not full. The cost, that is of risk-sensitive type and is rescaled at the MD regime, accounts for holding of customers in the buffers as well as for rejections. The term 'heavy traffic' refers to the imposition of a critical load condition, of the traffic intensity being close to one.

Heavy traffic analysis is traditionally carried out under the regime of *diffusion-scale deviations* (sometimes also referred to as *ordinary deviations*), where a vast variety of queueing control problems have been considered. However, there are very few results on the corresponding MD regime. This paper aims at continuing the direction started in [1] to develop an

---

\*Research supported in part by the ISF (Grant 1315/12) and the Technion fund for promotion of research

<sup>†</sup>Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel

<sup>‡</sup>Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA

approach to queueing control problems at the MD scale. While [1] finds various similarities to the diffusion limit theory, a major dissimilarity between the two formulations is at the heart of the present paper, as described below.

In [3] it has been shown that the *large deviation* asymptotics of risk-sensitive control problems associated with queueing models are governed by certain zero-sum differential games (DG) (this approach was inspired by and built on analogous treatments of small noise controlled diffusion models, such as [10]). In these games, the dynamics are given by a controlled ordinary differential equation (ODE), and the players are the DM and nature, that attempt to minimize and, respectively, maximize a cost. The DM again controls server's effort allocation and admission/rejection (except that the model no longer accounts for customers and so terms such as effort and admission/rejection are interpreted as suitable elements of the ODE), and nature may perturb the arrival and service rates. We will sometimes refer to the DM and nature as the minimizing player and maximizing player, respectively, or simply as minimizer and maximizer. The game's cost consists of the original holding and rejection penalties, and an additional term associated with the perturbation of the aforementioned rates. The latter term is determined by the underlying large deviation action functional. The treatment of a queueing control problem in the MD scale in [1] is similar to that of large deviations as far as the tools are concerned, but it is argued in that paper that the games one obtains in the MD regime tend to be solvable explicitly or semi-explicitly more often than under large deviation scaling. In fact, the specific DG identified in [1] turned out to be solvable due to a property of the model referred to in the heavy traffic literature as *pathwise minimality*, that enables approaching the game in a straightforward manner without appealing to dynamic programming methods. However, this property is quite special and it is therefore desired to find methods that apply in greater generality. In the present paper we aim at models that do not possess the pathwise minimality property, for which the solution methods must rely on tools such as dynamic programming. The model presented above is a prototype for models where a dynamic programming approach is required; specifically, the Bellman equation for the model identifies a *free boundary point* that is a crucial ingredient of the game's solution. The paper is devoted to the derivation and study of the DG itself, whereas the convergence of the risk-sensitive value to that of the DG is the subject of a work in progress [2].

As mentioned above, it was argued in [1] that unlike treatment at the large deviations regime, MD analysis parallels that of the diffusion regime in several ways. The present paper's contribution is best explained in the context of this relation. While the asymptotics of the MD scaled risk-sensitive control problem leads to a DG, for diffusion scaled (risk-neutral) control problems, they are typically described in terms of a *Brownian control problem* (BCP), namely a control problem involving diffusion processes. The BCP and the DG are similar in their structure, where the difference lies in that in the BCP, uncertainty is captured by the stochasticity of the Brownian motion term, whereas in the DG the role of uncertainty is played by an adversarial disturbance, in the form of the arrival and service rates being controlled by nature. The structural similarity is manifested by the fact that the mapping from the Brownian motion paths to the optimal control for the BCP is exactly the one used for the DG's optimal strategy. Moreover, a phenomenon known as *state space collapse*, where a BCP formulated in dimension  $d$  reduces to a 1-dimensional control problem, occurs in the corresponding DG as well (see [1] for details).

In the situation of the present paper, an analogy is less obvious, for reasons having to do

with a disadvantage of working with a discount factor in a risk-sensitive cost formulation. In the case of risk-neutral cost, a standard Markov decision process formulation based on discounted cost has the advantage over a finite time horizon cost, that optimal controls are given by stationary feedback. (For a typical queueing model this would often mean that decisions such as admission, server effort allocation and routing depend on the collection of queue lengths of the system in way that does not change over time.) However, it is well understood that a discounted version of a risk-sensitive cost leads to optimal feedback that is *time dependent* (this has been noticed first in [6]; several aspects of this issue were treated in [13]). Thus, although it is desirable to aim at a formulation that leads to stationary feedback because of the elegant structure of the solution, there does not seem to be a way to do so based on a discounted risk-sensitive cost. We address this issue by considering a risk-sensitive cost over a finite, but sufficiently large time horizon, in a formulation that does lead to stationary optimal feedback controls as far as the DG is concerned.

It is instructive to notice that thanks to the desired property alluded to above, the aforementioned analogy goes a step further. The *Harrison-Taksar free boundary problem* addresses a 1-dimensional BCP with an additive singular control term, for which the solution is given as a reflecting Brownian motion on an interval  $[a, b]$ , where  $a$  and  $b$  form a free boundary in a suitable Bellman equation [11]. The significance of the free boundary point  $b$  in applications to diffusion scaled queueing models is that it determines a threshold, in queue length terms, for rejection of arrivals. A (distinct) free boundary point identified in the present paper, denoted by  $\beta_0$ , plays the same role in the MD scale treatment: jobs are to be rejected once this point is exceeded. As far as the DG is concerned, an optimal strategy for the DM is one that applies the Skorohod map on the interval  $[0, \beta_0]$  to the dynamics of the game, regardless of the behavior of nature. Note the similar role played by the Skorohod map in both scenarios.

As far as analyzing the DG is concerned, the main contribution is the identification of the optimal strategy alluded to above. Apart from that, we state and prove several additional results, some of which form the basis for the result regarding the optimal strategy. Specifically, we find an equivalent game that is easier to analyze, characterize the DG's value function as the unique viscosity solution to the underlying Bellman equation, solve the equation explicitly, and identify a saddle point for the game as well as the optimal behavior of nature under the optimal play of the DM.

To summarize the contribution of this paper, we have

- identified a queueing control problem that is prototypical to situations where a dynamic programming approach is required to treat the DG governing the MD heavy traffic limit,
- proposed a risk-sensitive cost that leads to optimal stationary feedback for the DG,
- solved the Bellman equation, found an optimal strategy for the DM and studied properties of the DG.

Let us finally mention that the model studied here was recently treated in diffusion-scale heavy-traffic [4].

The organization of the paper is as follows. In Section 2 we describe the queueing model, introduce the MD scaling, present our approach to formulation of the cost and introduce the DG. In Section 3 we analyze the DG by introducing the hitting time game, present the Bellman equation, characterize the DG's value function as its unique viscosity solution, present an explicit optimal strategy for the DM, as well as a saddle point result.

We use the following notation. For a positive integer  $k$  and  $a, b \in \mathbb{R}^k$ ,  $a \cdot b$  denotes the usual scalar product. We denote  $[0, \infty)$  by  $\mathbb{R}_+$ . Denote by  $\mathcal{AC}([0, T], \mathbb{R}^k)$ ,  $\mathcal{C}([0, T], \mathbb{R}^k)$  and  $\mathcal{D}([0, T], \mathbb{R}^k)$  the spaces of absolutely continuous functions [resp., continuous functions, functions that are right-continuous with finite left limits (RCLL)] mapping  $[0, T] \rightarrow \mathbb{R}^k$ . Write  $\mathcal{AC}_0([0, T], \mathbb{R}^k)$  and  $\mathcal{C}_0([0, T], \mathbb{R}^k)$  for the subsets of the corresponding function spaces, of functions that start at zero. Endow the space  $\mathcal{D}([0, T], \mathbb{R}^k)$  with the usual Skorohod topology.

## 2 Model and results

### 2.1 The single-class model

We describe the probabilistic model that provides the main motivation for our DG formulation. The DG is derived from this model by considering scaling limits of the latter. Because the focus of the present paper is on the analysis of the DG itself, the derivation will only be carried out in a heuristic manner. In a work in progress [2] this relation is explored further and a rigorous relation is established.

We first present a basic model, then an extension of it. In the basic model, that we will call the *single-class model*, customers arrive into a buffer with finite room and are served by a single server in the order of arrival. The decision maker may reject arrivals even if the buffer is not full, but in order to keep the buffer length constraint it must always reject arrivals that occur when the buffer is full. The cost functional will involve penalties based on queue-length as well as rejection.

Scaling is introduced by considering a sequence of systems, indexed by  $n \in \mathbb{N}$ . In the  $n$ -th system, arrivals occur according to a renewal process,  $A^n$ , for which the mean inter-arrival time is given by  $1/\lambda^n$ . Denote by  $S^n(t)$  the number of service completions by the time the server has worked for  $t$  units of time. Then the inter-jump times of this process correspond to service time durations, and it is assumed that  $S^n$  is a renewal process (that is, the service times are i.i.d.). It is also assumed that  $A^n$  and  $S^n$  are mutually independent. The mean service time is denoted by  $1/\mu^n$ . It is assumed that the squared coefficient of variation for inter-arrival and service time distribution are fixed, and these constants are denoted by  $\sigma_{IA}^2 \in (0, \infty)$  and  $\sigma_{ST}^2 \in (0, \infty)$ , respectively.

Denoting the number of customers in the  $n$ -th system at time  $t$  by  $X^n(t)$ , the number of rejections until time  $t$  by  $R^n(t)$ , and the cumulative business time till  $t$  by  $T^n(t)$ , we have the balance equation

$$X^n(t) = X^n(0) + A^n(t) - S^n(T^n(t)) - R^n(t), \quad t \geq 0. \quad (1)$$

The process  $U^n = (T^n, R^n)$  is regarded as a control. The finite buffer constraint reads  $X^n(t) \leq D^n$  for all  $t$ , where  $D^n$  are given constants.

We consider the *moderate deviations rate parameters*  $\{b_n\}$ , that form a sequence, fixed throughout, with the property that  $\lim b_n = \infty$  while  $\lim n^{-1/2}b_n = 0$ , as  $n \rightarrow \infty$ . We assume that as  $n \rightarrow \infty$ ,

$$\lambda^n = \lambda n + \tilde{\lambda} b_n \sqrt{n} + o(b_n \sqrt{n}), \quad \mu^n = \mu n + \tilde{\mu} b_n \sqrt{n} + o(b_n \sqrt{n}),$$

where  $\lambda \in (0, \infty)$ ,  $\mu \in (0, \infty)$ ,  $\tilde{\lambda} \in \mathbb{R}$  and  $\tilde{\mu} \in \mathbb{R}$  are fixed. Moreover, it is assumed that the system is critically loaded, namely  $\lambda = \mu$ .

For  $T > 0$ , let  $\mathbb{I}_i(T, \cdot)$ ,  $i = 1, 2$ , be functions mapping  $\mathcal{D}([0, T], \mathbb{R})$  to  $\mathbb{R}_+ \cup \{+\infty\}$  defined as

$$\mathbb{I}_i(T, \psi) = \begin{cases} c_i \int_0^T \dot{\psi}^2(s) ds & \text{if } \psi \in \mathcal{AC}_0([0, T], \mathbb{R}), \\ +\infty & \text{otherwise,} \end{cases} \quad (2)$$

where

$$c_1 = \frac{1}{2\mu\sigma_{IA}^2} \text{ and } c_2 = \frac{1}{2\mu\sigma_{ST}^2}.$$

In the sequel we will also use the notation

$$c = (c_1^{-1} + c_2^{-1})^{-1} = \frac{1}{2\mu(\sigma_{ST}^2 + \sigma_{AI}^2)}.$$

Let  $\mathbb{I}(T, \psi) := \mathbb{I}_1(T, \psi^1) + \mathbb{I}_2(T, \psi^2)$ . Under suitable exponential moment conditions, the centered, scaled versions,

$$\tilde{A}^n(t) = \frac{1}{b_n\sqrt{n}}(A^n(t) - \lambda^n t), \quad \tilde{S}^n(t) = \frac{1}{b_n\sqrt{n}}(S^n(t) - \mu^n t), \quad t \geq 0,$$

of the renewal processes alluded to above satisfy a *moderate deviation principle* [14], [15]. That is, given  $T > 0$  and a bounded, continuous function  $H : \mathcal{D}([0, T], \mathbb{R})^2 \rightarrow \mathbb{R}$ , one has

$$\lim_{n \rightarrow \infty} \frac{1}{b_n^2} \log E e^{b_n^2 H(\tilde{A}^n, \tilde{S}^n)} = \sup\{H(\psi) - \mathbb{I}(\psi) : \psi \in \mathcal{AC}_0([0, T], \mathbb{R})^2\}. \quad (3)$$

Introduce scaled versions of the remaining two processes, namely

$$\tilde{R}^n(t) := \frac{1}{b_n\sqrt{n}}R^n(t), \quad \tilde{X}^n(t) := \frac{1}{b_n\sqrt{n}}X^n(t). \quad (4)$$

The buffer size is assumed to be given by  $D^n = b_n\sqrt{n}D$  for a constant  $D > 0$ , and so the buffer constraint can now be written as

$$\tilde{X}^n(t) \in [0, D], \quad t \geq 0. \quad (5)$$

To define the cost, let  $C^n(t) = \int_0^t h(\tilde{X}^n(s)) ds + r\tilde{R}^n(t)$ , where, throughout,  $h$  is a continuous increasing function from  $[0, D]$  to  $\mathbb{R}_+$  that satisfies  $h(0) = 0$  and accounts for holding cost, and  $r > 0$  is a constant representing per-customer rejection penalty. One may consider several versions of risk-sensitive costs based on the process  $C^n$ , the most natural ones being a cost defined over a finite time horizon,

$$J_1 = \frac{1}{b_n^2} \log E e^{b_n^2 \int_0^T C^n(s) ds},$$

a long-run average cost,

$$J_2 = \limsup_{T \rightarrow \infty} \frac{1}{T} \frac{1}{b_n^2} \log E e^{b_n^2 \int_0^T C^n(t) dt},$$

and two versions based on discount ( $\kappa > 0$  being the discount rate)

$$J_3 = \frac{1}{b_n^2} \log E e^{b_n^2 \int_0^\infty e^{-\kappa t} C^n(t) dt}, \quad J_4 = \frac{1}{b_n^2} \log E \int_0^\infty e^{-\kappa t} e^{b_n^2 C^n(t)} dt.$$

Because our setting is not Markovian, dynamic programming methods are not applicable in a direct fashion for the  $n$ th system. They are only applicable as far as the limit behavior is concerned. Thus the discussion that follows regards only the limiting behavior as well as the Markovian setting, that occurs in the special case when the driving renewal processes are Poisson. Working with a cost defined over a finite time horizon, such as  $J_1$ , suffers from the drawback that it leads to optimal feedback policies that are non-stationary. Long-run average cost of risk-sensitive type has been studied in several papers starting from [9]. It leads to stationary optimal feedback policies. However, working with this cost under scaling limits might be quite demanding technically, and therefore we will not focus on this cost here. Regarding the version  $J_3$  of a discounted cost, it is known to lead to non-stationary optimal feedback [6]. The reader is referred to [13] for discussion and several contributions regarding this issue. The formulation we propose is close in spirit to  $J_4$ . Specifically, given  $n$  and  $T$ , we consider the following cost associated with a control  $U^n$ , namely

$$\begin{aligned} J^n(U^n, T) &:= \frac{1}{b_n^2} \log \mathbb{E} \left[ \int_0^T e^{b_n^2 C^n(t)} dt \right] \\ &= \frac{1}{b_n^2} \log \mathbb{E} \left[ \int_0^T e^{b_n^2 [\int_0^t h(\tilde{X}^n(s)) ds + r \tilde{R}^n(t)]} dt \right], \end{aligned} \quad (6)$$

where  $(\tilde{X}^n, \tilde{R}^n)$  are the scaled versions of the processes  $(X^n, R^n)$  corresponding to the control  $U^n$ . The value is given by

$$V^n(T) = \inf_{U^n} J^n(U^n, T), \quad (7)$$

and our aim will be to understand the limit

$$\limsup_{T \rightarrow \infty} \limsup_{n \rightarrow \infty} V^n(T). \quad (8)$$

## 2.2 Differential game setting

We provide a heuristic argument for the claim that the limit of  $V^n$  is given by a game, before turning to a precise definition of the game. Note first that if one could write  $C^n$  as function  $H$  of the data  $(\tilde{A}^n, \tilde{S}^n)$  and the control  $U^n$ , namely  $C^n = H(U^n, \tilde{A}^n, \tilde{S}^n)$ , then based on (3), and ignoring questions of boundedness and continuity, one would obtain in the limit a game, defined in terms of the cost  $H(u, \psi) - \mathbb{I}(\psi)$ , maximized over  $\psi$  (as in (3)) and minimized over  $u$  (as in (7)). Although this conclusion is not solid due to questions of continuity, one can argue by appealing to specific aspects of the dependence of  $C^n$  on the data and control. Specifically, one notes from (1) that

$$\tilde{X}^n(t) = \tilde{X}^n(0) + \tilde{y}^n t + \tilde{A}^n(t) - \tilde{S}^n(T^n(t)) + \tilde{Z}^n(t) - \tilde{R}^n(t), \quad (9)$$

where we denote

$$\tilde{Z}^n(t) := \frac{\mu^n \sqrt{n}}{n b_n} (t - T^n(t)), \quad \tilde{y}^n := \tilde{\lambda}^n - \tilde{\mu}^n. \quad (10)$$

Moreover, the processes  $\tilde{Z}^n$  and  $\tilde{R}^n$  are nondecreasing, whereas, by assumption,  $\tilde{X}^n(0) \rightarrow x$  and  $\tilde{y}^n \rightarrow y$ . It is thus reasonable to consider a DG with dynamics  $\varphi(t) = x + yt + \psi^1(t) - \psi^2(t) + \zeta(t) - \varrho(t)$  where  $\zeta$  and  $\varrho$  are nondecreasing (corresponding to the processes  $Z^n$  and  $R^n$ ),  $\varphi$  takes values in  $[0, D]$  (because of (5)), and the cost takes the form

$$\sup_{T \in [0, \infty)} \left\{ \int_0^T h(\varphi(t)) dt + r\varrho(T) - \mathbb{I}(T, \psi) \right\}.$$

It is maximized over  $\psi$  and minimized over  $u = (\varrho, \zeta)$ . The precise formulation along these lines is given next.

Let  $y = \tilde{\lambda} - \tilde{\mu}$ . Denote by  $\mathcal{P} = \mathcal{C}_0([0, \infty), \mathbb{R})$  and

$$\mathcal{E} = \{\xi \in \mathcal{D}([0, \infty), \mathbb{R}_+) : \xi \text{ is nondecreasing}\}.$$

Endow both spaces with the uniform-on-compacts topology. Given  $\psi^1, \psi^2 \in \mathcal{P}$  and  $\zeta, \varrho \in \mathcal{E}$ , the *dynamics associated with the initial condition  $x$  and the data  $\psi, \zeta, \varrho$*  is given by

$$\varphi(t) = x + yt + \psi^1(t) - \psi^2(t) + \zeta(t) - \varrho(t), \quad t \geq 0. \quad (11)$$

We sometimes write the dependence of the dynamics on the data as  $\varphi[\psi, \zeta, \varrho]$ . Note the analogy between the above equation and equation (9), and between the condition that  $\zeta$  and  $\varrho$  are nondecreasing and property (10). The control  $\zeta$  stands for the scaled idle time process  $\tilde{Z}^n$  and  $\varrho$  stands for the scaled rejection process  $\tilde{R}^n$ . The following condition, analogous to property (5), will also be in use, namely

$$\varphi(t) \in [0, D], \quad t \geq 0. \quad (12)$$

The game is defined in the sense of Elliott and Kalton [7], for which we need the notion of strategies. A measurable mapping  $\alpha : \mathcal{P}^2 \rightarrow \mathcal{E}^2$  is called a *strategy for the minimizing player* if it satisfies a causality property. Namely, for every  $\psi = (\psi^1, \psi^2), \tilde{\psi} = (\tilde{\psi}^1, \tilde{\psi}^2) \in \mathcal{P}^2$  and  $t \in [0, \infty)$ ,

$$(\psi^1, \psi^2)(s) = (\tilde{\psi}^1, \tilde{\psi}^2)(s) \text{ for every } s \in [0, t] \quad \text{implies} \quad \alpha[\psi](s) = \alpha[\tilde{\psi}](s) \text{ for every } s \in [0, t]. \quad (13)$$

Given an initial condition  $x$ , a strategy  $\alpha$  is said to be *admissible* if, whenever  $(\psi^1, \psi^2) \in \mathcal{P}^2$  and  $(\zeta, \varrho) = \alpha[\psi]$ , the corresponding dynamics (11) satisfies the buffer size constraint (12). The set of all admissible strategies for the minimizing player is denoted by  $\mathcal{A}$  (or, when the dependence on the initial condition is important,  $\mathcal{A}_x$ ). Given  $T \in \mathbb{R}_+$ ,  $x \in [0, D]$ ,  $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$ , and  $(\zeta, \varrho) \in \mathcal{E}^2$ , we define the cost until time  $T$  by

$$c(x, T, \psi, \zeta, \varrho) := \int_0^T h(\varphi(t)) dt + r\varrho(T) - \mathbb{I}(T, \psi),$$

where  $\varphi$  is the corresponding dynamics. The value of the game is defined by

$$V(x) := \inf_{\alpha \in \mathcal{A}_x} \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi]).$$

We call  $\psi$  the *path control* and the  $T$  a *time control*, or sometimes the *termination time*. Note that both are controlled by the maximizer.



### 2.3 The multi-class model

In a similar manner to the single-class model, we describe a model with  $I \geq 1$  classes of customers, where a buffer of finite length is dedicated to each class. The various primitives, parameters and processes are denoted with an additional index  $i \in \mathcal{I} := \{1, \dots, I\}$ , representing the class. As far as the control is concerned, an important difference from the single-class model is that the DM has to determine the allocation of effort to the various classes, on top of making admission decisions. Thus, with  $\mathbb{S} = \{x = (x_1, \dots, x_I) \in [0, 1]^I : \sum x_i \leq 1\}$ , let  $B^n$  be a process taking values in  $\mathbb{S}$ , whose  $i$ -th component represents the fraction of effort devoted by the server to the class- $i$  customer at the head of the line. Then the number of service completions of class- $i$  jobs during the time interval  $[0, t]$  is given by  $S_i^n(T_i^n(t))$ , where

$$T_i^n(t) := \int_0^t B_i^n(s) ds \quad (14)$$

is the time devoted to class- $i$  customers by time  $t$ . We have the balance equation

$$X_i^n(t) = X_i^n(0) + A_i^n(t) - S_i^n(T_i^n(t)) - R_i^n(t). \quad (15)$$

The process  $U^n := (B^n, R^n)$  is now regarded as a control. For  $i \in \mathcal{I}$  fixed constants  $D_i > 0$  serve in defining the buffer sizes, and so the definition of an admissible control now requires that for every  $i \in \mathcal{I}$  one has

$$X_i^n(t) \in [0, b_n \sqrt{n} D_i], \quad t \geq 0.$$

Scaling that is applied in a similar fashion to that in the single-class case gives rise to the equation

$$\tilde{X}_i^n(t) = \tilde{X}_i^n(0) + y_i^n t + \tilde{A}_i^n(t) - \tilde{S}_i^n(T_i^n(t)) + Z_i^n(t) - \tilde{R}_i^n(t), \quad (16)$$

where we denote  $\rho_i = \lambda_i / \mu_i$  and

$$Z_i^n(t) := \frac{\mu_i^n}{n} \frac{\sqrt{n}}{b_n} (\rho_i t - T_i^n(t)), \quad y_i^n := \tilde{\lambda}_i^n - \rho_i \tilde{\mu}_i^n.$$

The critical load condition now takes the form  $\sum_i \rho_i = 1$ .

An important role is played by the property

$$\sum_i \frac{n}{\mu_i^n} Z_i^n \quad \text{starts from zero and is nondecreasing,} \quad (17)$$

that holds thanks to the fact that  $\sum_i B_i^n \leq 1$  while  $\sum_i \rho_i = 1$ . The probabilistic assumptions on the primitives are analogous to those previously imposed, where in addition, the  $2I$  primitive processes are assumed mutually independent. Finally, the cost is assumed to be of the form

$$J^n(U^n, T) := \frac{1}{b_n^2} \log \mathbb{E} \left[ \int_0^T e^{b_n^2 [\int_0^t \underline{h} \cdot \tilde{X}^n(s) ds + \underline{r} \cdot \tilde{R}^n(t)]} dt \right],$$

where  $\underline{h}$  and  $\underline{r}$  are given positive vectors and the value is given by

$$V^n(T) = \inf_{U^n} J^n(U^n, T).$$

Without going into details, we present the form of the DG that is obtained from the balance equation (16), the condition (17) and the cost structure; the derivation is very similar to that presented earlier. Let  $\underline{\theta} = (\frac{1}{\mu_1}, \dots, \frac{1}{\mu_I})$  and  $\underline{y} = (y_1, \dots, y_I)$  where  $y_i = \tilde{\lambda}_i - \rho_i \tilde{\mu}_i$ . Denote

$$\mathcal{E}_{\underline{\theta}} = \{\underline{\zeta} \in \mathcal{D}([0, T], \mathbb{R}_+^I) : \underline{\theta} \cdot \underline{\zeta} \text{ is nondecreasing}\}.$$

Given  $\underline{\psi} = (\psi^1, \psi^2) \in \mathcal{P}^{2I}$ ,  $\underline{\zeta} \in \mathcal{E}_{\underline{\theta}}$ , and  $\underline{\varrho} \in \mathcal{E}^I$ , the *dynamics*,  $\underline{\varphi} = (\varphi_1, \dots, \varphi_I)$ , associated with the initial state  $\underline{x} \in \prod_{i=1}^I [0, D_i]$  and the data  $\underline{\psi}, \underline{\zeta}$ , and  $\underline{\varrho}$  is given by

$$\varphi_i(t) = x_i + y_i t + \psi_i^1(t) - \psi_i^2(t) + \zeta_i(t) - \varrho_i(t), \quad t \geq 0, \quad i \in \mathcal{I}. \quad (18)$$

Note the analogy between the above equation and equation (16), and between the condition that  $\underline{\theta} \cdot \underline{\zeta}$  is nondecreasing and property (17). The buffer constraint is now translated into the condition

$$\varphi_i(t) \in [0, D_i], \quad t \geq 0, \quad i \in \mathcal{I}.$$

A measurable mapping  $\underline{\alpha} : \mathcal{P}^{2I} \rightarrow \mathcal{E}_{\underline{\theta}} \times \mathcal{E}^I$  is called a *strategy for the minimizing player* if it satisfies a causality property. Namely, for every  $\underline{\psi}, \tilde{\psi} \in \mathcal{P}^{2I}$  and  $t \in [0, T]$ ,

$$\underline{\psi}(s) = \tilde{\psi}(s) \text{ for all } s \in [0, t] \quad \text{implies} \quad \underline{\alpha}[\underline{\psi}](s) = \underline{\alpha}[\tilde{\psi}](s) \text{ for all } s \in [0, t].$$

As before, we denote by  $\mathcal{A}_{\underline{x}}$  the set of all admissible strategies given the initial state  $\underline{x}$ . The cost is given by

$$c(\underline{x}, T, \underline{\psi}, \underline{\zeta}, \underline{\varrho}) := \int_0^T \underline{h} \cdot \underline{\varphi}(t) dt + r \cdot \underline{\varrho}(T) - \mathbb{J}(T, \underline{\psi}),$$

where  $\underline{\varphi}$  is the corresponding dynamics given in (18) and  $\mathbb{J}$  the corresponding rate function (specified in (93) in the appendix). The value of the game is defined by

$$V_{MD}(x) := \inf_{\underline{\alpha} \in \mathcal{A}_{\underline{x}}} \sup_{\underline{\psi} \in \mathcal{P}^{2I}, T \in \mathbb{R}_+} c(\underline{x}, T, \underline{\psi}, \underline{\alpha}[\underline{\psi}]).$$

Our result concerning this DG is that *it can be reduced to the one introduced earlier* for the single-class model. In the heavy traffic literature, this type of result is often called *state space collapse*. The precise statement of this reduction, along with its proof are provided in the appendix. Since this multidimensional version of the game can be reduced to a one-dimensional one, the rest of this paper is concerned with the latter.

### 3 Solution and analysis of the game

In this section we provide a solution of the game. After proving some basic properties in Section 3.1, we present, in Section 3.2, an equivalent *hitting time* game which is easier to analyze. In Section 3.3 we present the Bellman equation associated with the game. In Section 3.4 we characterize the value function as a viscosity solution of the Bellman equation. An explicit expression for the minimizer's optimal strategy appears in Section 3.5. Finally, a saddle point result appears in Section 3.6.

The minimizer's strategy that is shown to be optimal is of a  $\beta$ -barrier form. Informally, this is a strategy that uses the minimal control  $(\zeta, \varrho)$  so as to keep the dynamics  $\varphi$  in  $[0, \beta]$

at all times. The definition is based on the *Skorohod map on an interval*. To introduce this map, fix  $a < b$ . The Skorohod map on the interval  $[a, b]$ , denoted by  $\Gamma_{[a,b]}$ , is a map  $\mathcal{D}([0, \infty), \mathbb{R}) \rightarrow \mathcal{D}([0, \infty), \mathbb{R}^3)$ . It is characterized as the solution map  $\omega \rightarrow (\varphi, \eta_1, \eta_2)$  to the so called *Skorohod Problem*, namely the problem of finding, for a given  $\omega$ , a triplet  $(\varphi, \eta_1, \eta_2)$ , such that

$$\varphi = \omega + \eta_1 - \eta_2, \quad \varphi(t) \in [a, b] \text{ for all } t, \quad (19)$$

$\eta_i$  are nonnegative and nondecreasing,  $\eta_i(0-) = 0$ , and

$$\int_{[0, \infty)} 1_{(a,b)}(\varphi) d\eta_1 = \int_{[0, \infty)} 1_{[a,b)}(\varphi) d\eta_2 = 0. \quad (20)$$

By writing  $\eta_i(0-) = 0$  we adopt the convention that  $\eta_i(0) > 0$  is regarded a jump at zero. This convention, in conjunction with  $\int_{[0, \infty)} 1_{(a,b]}(\varphi) d\eta_1 = 0$  (resp.,  $\int_{[0, \infty)} 1_{[a,b)}(\varphi) d\eta_2 = 0$ ), means that if  $\omega(0) < a$  (resp.,  $\omega(0) > b$ ) then  $\varphi(0) = a$  (resp.,  $b$ ). If, however,  $\omega(0) \in [a, b]$  then  $\varphi(0) = \omega(0)$ , and  $\eta_i$  have no jump at zero.

See [12] for existence and uniqueness of solutions, and continuity and further properties of the map. In particular, it is well-known that  $\Gamma_{[a,b]}$  is continuous in the uniform-on-compacts topology.

**Definition 3.1** Fix  $(x, \beta) \in [0, D]^2$ . The strategy  $\alpha_\beta = (\alpha_{\beta,1}, \alpha_{\beta,2}) \in \mathcal{A}_x$  is called a  $\beta$ -barrier strategy if for every  $\psi \in \mathcal{P}^2$  one has  $(\varphi, \alpha_{\beta,1}, \alpha_{\beta,2})[\psi] = \Gamma_{[0,\beta]}(\psi)$ .

The specific value  $\beta_0$  of  $\beta$  that gives rise to an optimal strategy will be determined later, based on the free boundary associated with the Bellman equation. For the moment we only comment that it is possible to select  $\beta = 0$ , and in fact this choice turns out to be useful in Section 3.1 below.

### 3.1 Basic properties

We begin by providing a necessary and sufficient condition for the value function to be finite and presenting further properties that are used in the sequel.

**Lemma 3.1** *i. If  $-y < r/(4c)$  then for every  $x \in [0, D]$  one has  $V(x) = \infty$ .  
ii. If  $-y \geq r/(4c)$  then for every  $x \in [0, D]$  one has  $0 \leq V(x) \leq rx$ , in particular  $V(0) = 0$ . Moreover,  $V$  is nondecreasing. Finally,  $V$  is continuous.*

**Proof of Lemma 3.1:** i. Fix  $x \in [0, D]$ . Consider a maximizer's control given by  $T^\# \in \mathbb{R}_+$  and  $\psi^{\#,1}(t) = rt/(2c_1)$ ,  $\psi^{\#,2}(t) = -rt/(2c_2)$ ,  $t \geq 0$ . Let  $\alpha \in \mathcal{A}_x$ , and denote by  $(\zeta^\#, \varrho^\#) = \alpha[\psi^\#]$ . The dynamics under  $(\psi^{\#,1}, \psi^{\#,2})$  and  $(\zeta^\#, \varrho^\#)$  is denoted by  $\varphi^\#$ . We show that the cost is bounded below by a quantity that converges to infinity as  $T^\# \rightarrow \infty$ . Since the maximizer is free to

choose  $T^\sharp \in \mathbb{R}_+$ , the supremum of the cost function over  $T$  is infinity. Indeed,

$$c(x, T^\sharp, \psi^\sharp, \alpha^\sharp[\psi^\sharp]) = \int_0^{T^\sharp} h(\varphi^\sharp(t))dt + r\varrho^\sharp(T) - \mathbb{I}(T^\sharp, \psi^\sharp) \quad (21)$$

$$\geq r\varrho^\sharp(T^\sharp) - \mathbb{I}(T^\sharp, \psi^\sharp) = r\varrho^\sharp(T^\sharp) - \frac{1}{4c}r^2T^\sharp \quad (22)$$

$$\geq r(x + \frac{1}{2c}rT^\sharp + yT^\sharp + \zeta^\sharp(T) - D) - \frac{1}{4c}r^2T^\sharp \quad (23)$$

$$= r(x - D) + (\frac{1}{4c}r^2 + yr)T^\sharp + r\zeta(T^\sharp). \quad (24)$$

Inequality (22) follows since  $h$  is nonnegative. Inequality (23) follows by the dynamics of the game (see (11)) and constraint (12). Equation (24) is merely a rearrangement of the terms.

ii. Fix  $x \in [0, D]$ . To see why  $0 \leq V(x) \leq rx$  we show that the maximizer (resp., minimizer) can force a cost of at least 0 (resp., at most  $rx$ ). The maximizer may choose  $T = 0$ , in which case for any strategy  $\alpha \in \mathcal{A}_x$ , the cost is bounded below by 0. This shows  $V(x) \geq 0$ . The minimizer can choose a 0-barrier strategy so that the dynamics remain at zero. That is, the minimizer may choose a strategy  $\alpha_0$  that, given  $T \in \mathbb{R}_+$  and  $\psi \in \mathcal{P}^2$  responds with  $(\zeta_0, \varrho_0)$  where

$$\begin{aligned} \dot{\zeta}_0(t) &= (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)^-, & 0 < t \leq T, \\ \dot{\varrho}_0(t) &= (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)^+, & 0 < t \leq T, \end{aligned}$$

with  $\zeta_0(0) = 0$  and  $\varrho_0(0) = x$ . Fix  $T \in \mathbb{R}_+$  and  $\psi \in \mathcal{P}^2$ . From the definition of  $(\zeta_0, \varrho_0)$ , and since for every  $t > 0$  one has  $h(\varphi(t)) = h(0) = 0$  it follows that

$$\begin{aligned} c(x, T, \psi, \alpha_0[\psi]) &= \int_0^T h(\varphi(t))dt + r\varrho_0(T) - \mathbb{I}(T, \psi) \quad (25) \\ &= r\varrho_0(0) + \int_0^T [h(\varphi(t)) + r(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)^+ \\ &\quad - c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t)]dt \\ &= r\varrho_0(0) + \int_0^T [r(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)^+ - c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t)]dt. \end{aligned}$$

One can easily verify that for every  $\dot{\psi}^1(t), \dot{\psi}^2(t) \in \mathbb{R}$  the expression given in the integrand is bounded above by  $r^2/(4c) + ry$ , which is nonpositive by assumption. Therefore the cost given in (25) is bounded above by  $r\varrho_0(0) = rx$ . This shows  $V(x) \leq rx$ .

From the inequality  $0 \leq V(x) \leq rx$  applied to  $x = 0$  it follows that  $V(0) = 0$ .

To see why  $V$  is nondecreasing fix  $0 \leq x_1 < x_2 \leq D$ . Given any strategy  $\alpha_2$  associated with the initial state  $x_2$ , the minimizer may choose for the initial state  $x_1$  a strategy  $\alpha_1$  obtained from  $\alpha_2$  by adding to the  $\zeta$  component the constant  $x_2 - x_1$  (without incurring any additional cost), leaving the  $\varrho$  component as under  $\alpha_2$ . Therefore, the optimal cost associated with the initial state  $x_1$  is less than or equal to that associated with  $x_2$ . That is,  $V(x_1) \leq V(x_2)$ .

Since  $V$  is nondecreasing its continuity follows by showing that it has no upwards jumps. Assume to the contrary that  $V$  has a jump of size  $M > 0$  at  $x \in [0, D]$ . In case,  $x \in (0, D)$

then for every sufficiently small  $\delta > 0$  one has  $V(x - \delta) \leq V(x + \delta) - M$ . W.l.o.g. assume that  $\delta < \min\{x, M/(4r)\}$ . Therefore, by the definition of  $V$  applied to  $x - \delta$  it follows that there is a strategy  $\alpha = (\alpha_1, \alpha_2) \in \mathcal{A}_{x-\delta}$  such that

$$\sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x - \delta, T, \psi, \alpha[\psi]) < V(x - \delta) + M/2. \quad (26)$$

Define a new strategy  $\hat{\alpha} = (\hat{\alpha}_1, \hat{\alpha}_2) \in \mathcal{A}_{x+\delta}$  by  $(\hat{\alpha}_1, \hat{\alpha}_2) = (\alpha_1, \alpha_2 - 2\delta)$  then clearly,

$$\sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x + \delta, T, \psi, \hat{\alpha}[\psi]) = 2\delta + \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x - \delta, T, \psi, \alpha[\psi]). \quad (27)$$

From (26) and (27) and the choice of  $\delta$  it follows that  $\sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x + \delta, T, \psi, \hat{\alpha}[\psi]) < V(x + \delta)$ , which contradicts the definition of  $V$ . In case  $x = 0$  or  $x = D$  the proof requires only minor modifications and therefore omitted.  $\square$

In the rest of the paper, our standing assumption will be that  $-y \geq r/(4c)$  holds; thus by Lemma 3.1 the value function is finite.

### 3.2 The hitting time game

In the proof of Lemma 3.1 we saw that if the initial state of the dynamics,  $\varphi(0) = x$ , equals zero then the maximizer prefers to terminate the game immediately. Therefore, it seems plausible that, for a general initial condition, if both players play in an optimal way then as soon as the dynamics hits zero, the game is terminated. Hence the minimizer has no interest in ever activating the  $\zeta$  component, since doing so only increases the running cost  $h(\varphi(t))$  and postpones the termination of the game. In this section we validate this intuition by formulating an equivalent game that admits the following modifications: first, the  $\zeta$  component of the minimizer's control vanishes. Second, the game terminates when the dynamics hit zero.

Given  $\psi \in \mathcal{P}^2$  and  $\varrho \in \mathcal{E}$ , the dynamics of the new game are given by

$$\varphi = \varphi_{x,\psi,\varrho}(t) := x + yt + \psi^1(t) - \psi^2(t) - \varrho(t), \quad t \geq 0.$$

A strategy is a map  $\bar{\alpha} : \mathcal{P}^2 \rightarrow \mathcal{E}$  that satisfies the following version of the causality property:

$$\psi(s) = \tilde{\psi}(s) \text{ for every } s \in [0, t] \text{ implies } \bar{\alpha}[\psi](s) = \bar{\alpha}[\tilde{\psi}](s) \text{ for every } s \in [0, t].$$

A strategy is said to be admissible if for every  $t \geq 0$  the dynamics satisfies  $\varphi(t) \in [0, D]$ . Denote the set of all such strategies by  $\bar{\mathcal{A}}_x$ . Let

$$\tau = \tau(x, \psi, \varrho) := \inf\{t \geq 0 : \varphi_{x,\psi,\varrho}(t) = 0\}, \quad (28)$$

where throughout we use the convention  $\inf \emptyset = \infty$ . For  $x \in [0, D]$  and  $\bar{\alpha} \in \bar{\mathcal{A}}_x$ , define

$$\mathcal{Q}[x, \bar{\alpha}] := \{\psi \in \mathcal{P}^2 : \tau(x, \psi, \bar{\alpha}[\psi]) < \infty\}.$$

This is the set of all the controls for which the dynamics of the game hits zero in a finite time. Given  $x \in [0, D]$ ,  $\bar{\alpha} \in \bar{\mathcal{A}}_x$ , and  $\psi \in \mathcal{Q}[x, \bar{\alpha}]$ , we define the cost by

$$\bar{c}(x, \psi, \bar{\alpha}[\psi]) := \int_0^\tau h(\varphi(t))dt + r\bar{\alpha}[\psi](\tau) - \mathbb{I}(\tau, \psi),$$

where  $\tau = \tau(x, \psi, \bar{\alpha}[\psi])$ . The value of the hitting time game is defined by

$$U(x) = \inf_{\bar{\alpha} \in \bar{\mathcal{A}}_x} \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}]} \bar{c}(x, \psi, \bar{\alpha}[\psi]).$$

The following proposition states that the value functions of both games are equal to each other.

**Proposition 3.1** *If  $-y \geq r/(4c)$  holds then for every  $x \in [0, D]$  one has  $V(x) = U(x)$ .*

**Proof:** For every  $x \in [0, D]$ , every  $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$ , and every  $\alpha = (\alpha_1, \alpha_2) \in \mathcal{A}_x$  let

$$\pi = \pi(x, \psi, \alpha) := \inf\{t \geq 0 : x + yt + \psi^1(t) - \psi^2(t) + \alpha_1[\psi](t) - \alpha_2[\psi](t) = 0\}.$$

This is the first time that the dynamics in the original game hit zero. Let  $\mathcal{A}_{x,0}$  be the set of all admissible strategies  $\alpha : \mathcal{P}^2 \rightarrow \mathcal{E}^2$  such that  $\alpha^1[\psi] = 0$  on  $[0, \pi)$  and thereafter  $\alpha$  acts as a 0-barrier strategy.

We prove the proposition in two steps. First, we show that the infimum in the definition of  $V$  can be taken over the set  $\mathcal{A}_{x,0}$ . That is,

$$V(x) = V_0(x) := \inf_{\alpha \in \mathcal{A}_{x,0}} \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi]). \quad (29)$$

On the second step we show that we can replace the set of controls  $\mathcal{P}^2 \times \mathbb{R}_+$  for the maximizer by  $\mathcal{Q}[x, \alpha_2]$ . That is,

$$V_0(x) = V_1(x) := \inf_{\alpha \in \mathcal{A}_{x,0}} \sup_{\psi \in \mathcal{Q}[x, \alpha_2]} \bar{c}(x, \psi, \alpha_2[\psi]). \quad (30)$$

Once this is achieved, the result follows because it is obvious that  $V_1 = U$ .

**Step 1:** Fix  $x \in [0, D]$ . Since  $\mathcal{A}_{x,0} \subseteq \mathcal{A}_x$  it is clear that  $V(x) \leq V_0(x)$ .

To prove that  $V(x) \geq V_0(x)$  we show that for every strategy  $\alpha \in \mathcal{A}_x$  there exists a strategy  $\alpha^* \in \mathcal{A}_{x,0}$  such that for every  $\psi \in \mathcal{P}^2$  one has

$$\sup_{T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi]) \geq \sup_{T \in \mathbb{R}_+} c(x, T, \psi, \alpha^*[\psi]). \quad (31)$$

By taking  $\sup_{\psi \in \mathcal{P}^2}$  and then  $\inf_{\alpha \in \mathcal{A}_x}$  it then follows that  $V(x) \geq V_0(x)$ . To show (31), fix a strategy  $\alpha = (\alpha_1, \alpha_2) \in \mathcal{A}_x$ . Given  $\psi \in \mathcal{P}^2$ , let

$$\pi^* = \pi^*(x, \psi, \alpha) := \inf\{t \geq 0 : x + yt + \psi^1(t) - \psi^2(t) - \alpha_2[\psi](t) \leq 0\}.$$

Notice that we omitted  $\alpha_1[\psi]$  from the dynamics. Define  $\alpha^*[\psi](t) = (0, \alpha_2[\psi](t))$ ,  $0 \leq t < \pi^*$ , and thereafter let  $\alpha^*$  act as a 0-barrier strategy. One has that  $\pi^*(x, \psi, \alpha) = \tau(x, \psi, \alpha_2[\psi])$ . By construction,  $\alpha_2^*[\psi](t) \leq \alpha_2[\psi](t)$  for every  $t \in [0, \pi^*]$ . Let  $\varphi$  (resp.,  $\varphi^*$ ) denote the dynamics associated with  $\psi$  and  $\alpha$  (resp.,  $\psi$  and  $\alpha^*$ ). Obviously,

$$\begin{aligned} \sup_{T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi]) &= \sup_{T \in \mathbb{R}_+} \left\{ \int_0^T h(\varphi(t)) dt + r\alpha_2[\psi](\pi^*) - \mathbb{I}(T, \psi) \right\} \\ &\geq \sup_{T \in \mathbb{R}_+} \left\{ \int_0^{T \wedge \pi^*} h(\varphi(t)) dt + r\alpha_2[\psi](T \wedge \pi^*) - \mathbb{I}(T \wedge \pi^*, \psi) \right\}. \end{aligned} \quad (32)$$

Since over the time interval  $[0, \pi^*]$  one has  $\alpha_2^*[\psi] \leq \alpha_2[\psi]$  and  $\varphi^* \leq \varphi$ , and since  $h$  is increasing, it follows that the r.h.s. of (32) will not increase if we replace  $(\alpha, \varphi)$  by  $(\alpha^*, \varphi^*)$ . Therefore, for every  $\psi \in \mathcal{P}^2$  one has

$$\sup_{T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi]) \geq \sup_{T \in \mathbb{R}_+} c(x, T \wedge \pi^*, \psi, \alpha^*[\psi]).$$

Since starting at time  $\pi^*$  the minimizer uses 0-barrier strategy, it follows by the same arguments as in part (ii) of the proof of Lemma 3.1 that the maximizer terminates the game at time  $\pi^*$  or earlier. Inequality (31) follows.

**Step 2:** Recall that  $-y \geq r/(4c)$  and therefore  $V(x) < \infty$ . Therefore, in order to show that (30) holds, it suffices to show that for every  $\alpha \in \mathcal{A}_{x,0}$  for which  $\sup_{\psi \in \mathcal{P}^2} c(x, T, \psi, \alpha[\psi]) < \infty$ , one has

$$\sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi]) = \sup_{\psi \in \mathcal{Q}[x, \alpha_2]} \bar{c}(x, \psi, \alpha_2[\psi]). \quad (33)$$

We do that by showing two sided inequalities. The inequality  $\geq$  is obvious since

$$\sup_{\psi \in \mathcal{Q}[x, \alpha_2]} \bar{c}(x, \psi, \alpha_2[\psi]) = \sup_{\psi \in \mathcal{Q}[x, \alpha_2]} c(x, \pi, \psi, \alpha[\psi]) \quad (34)$$

and the supremum on the r.h.s. of (34) is taken over a smaller set than the supremum on the l.h.s. of (33).

To prove the reverse inequality we show that for every  $\psi \in \mathcal{P}^2$  and every  $T \in \mathbb{R}_+$  there is  $\psi^* \in \mathcal{Q}[x, \alpha_2]$  such that

$$c(x, T, \psi, \alpha[\psi]) \leq \bar{c}(x, \psi^*, \alpha_2[\psi^*]). \quad (35)$$

Let  $\pi = \pi(x, \psi, \alpha)$  be the first time the dynamics under  $\psi$  and  $\alpha$  hit zero, possibly infinity. We argue separately for  $T$  that is greater than  $\pi$  and for  $T$  that is smaller than or equal to  $\pi$ . For  $T > \pi$ , let  $\psi^* = \psi$ . By the definition of  $\mathcal{A}_{x,0}$   $\pi(x, \psi, \alpha) = \tau(x, \psi^*, \alpha_2[\psi^*])$ . Denote this time by  $\tau$ . The cost  $c(x, T, \psi, \alpha[\psi])$  can be expressed as the sum of the cost until time  $\tau$  and the cost between times  $\tau$  and  $T$ . The first cost is  $c(x, \tau, \psi, \alpha[\psi]) = \bar{c}(x, \psi^*, \alpha_2[\psi^*])$ , and the second cost is non positive since the minimizer uses a 0-barrier strategy after time  $\tau$  and in the proof of Lemma 3.1 we showed that the maximizer optimal cost under this strategy is zero.

For  $T \leq \tau$  define  $\psi^* = \psi(\cdot \wedge T)$ . The dynamics associated with  $\psi^* = (\psi^{*,1}, \psi^{*,2})$  and  $\alpha$  until time  $\tau^* = \tau(x, \psi^*, \alpha_2)$  is  $\varphi^*(t) = x + yt + \psi^{*,1}(t) - \psi^{*,2}(t) - \alpha_2[\psi^*](t)$ . Since  $y < 0$  and  $\psi^*$  is constant on the time interval  $[T, \infty)$  it follows that  $\varphi^*$  hits zero in a finite time, and therefore  $\psi^* \in \mathcal{Q}[x, \alpha_2]$ . Thus

$$\begin{aligned} \bar{c}(x, \psi^*, \alpha_2[\psi^*]) &= c(x, \tau^*, \psi^*, \alpha_2[\psi^*]) \\ &= c(x, T, \psi^*, \alpha_2[\psi^*]) + \int_T^{\tau^*} h(\varphi^*(t)) dt + r(\alpha_2[\psi^*](\tau^*) - \alpha_2[\psi^*](T)) \\ &\quad - (\mathbb{I}(\tau^*, \psi^*) - \mathbb{I}(T, \psi^*)) \\ &\leq c(x, T, \psi^*, \alpha_2[\psi^*]) = c(x, T, \psi, \alpha_2[\psi^*]). \end{aligned}$$

The second equality expresses merely the sum of the cost until time  $T$  and the cost between times  $T$  and  $\tau^*$ . The inequality follows since  $\alpha_2$  and  $h$  are nondecreasing and since  $\dot{\psi}^1 = \dot{\psi}^2 = 0$  on the time interval  $[T, \tau^*]$ . The last equality follows since on the time interval  $[0, T]$  we chose  $\psi^*$  to agree with  $\psi$ . □

### 3.3 Derivation of the Bellman equation

Here we give an intuitive approach towards the solution of the game, present the corresponding Bellman equation and discuss its solvability. The derivation of the Bellman equation uses the hitting time game.

We first write the chain rule under the dynamics

$$\varphi(t) = x + yt + \psi^1(t) - \psi^2(t) - \varrho(t), \quad 0 \leq t \leq T,$$

where  $x \in [0, D]$ ,  $T \in \mathbb{R}_+$ ,  $\psi \in \mathcal{P}^2$ , and  $\varrho \in \mathcal{E}$ . Namely, for any function  $f \in \mathcal{C}^1([0, D], \mathbb{R})$  one has

$$f(\varphi(T)) = f(x) + \int_0^T f'(\varphi(t))(d\psi^1(t) - d\psi^2(t) + ydt - d\varrho^c(t)) + \sum_{0 \leq t \leq T} \Delta f(\varphi)(t), \quad (36)$$

where  $\varrho^c$  is the continuous part of  $\varrho$ . Thus

$$\begin{aligned} & \int_0^T \left[ -c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) \right] dt + r\varrho(T) + f(\varphi(T)) \\ &= f(x) + \int_0^T \left[ -c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + f'(\varphi(t))(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y) \right] dt \\ & \quad + \int_0^T (r - f'(\varphi(t)))d\varrho(t). \end{aligned} \quad (37)$$

Assuming that there exists  $\bar{\alpha} \in \bar{\mathcal{A}}_x$  that is an optimal strategy, by the definition of  $U$  one has

$$\sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}]} \left[ \int_0^\tau \left[ -c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) \right] dt + r\bar{\alpha}[\psi](\tau) \right] = U(x). \quad (38)$$

Noting that  $U(\varphi(\tau)) = U(0) = 0$ , we have by equations (38) and (37) applied to  $f = U$  and  $T = \tau$ ,

$$\begin{aligned} \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}]} \left\{ \int_0^T \left[ -c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + U'(\varphi(t))(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y) \right] dt \right. \\ \left. + \int_0^T (r - U'(\varphi(t)))d\bar{\alpha}[\psi](t) \right\} = 0. \end{aligned} \quad (39)$$

This gives rise to the following equation in the state variable, namely

$$\max_{p, q \in \mathbb{R}} \min \left\{ -c_1p^2 - c_2q^2 + h(x) + f'(x)(p - q + y), r - f'(x) \right\} = 0. \quad (40)$$



The maximum above is achieved by  $(p, q) = (rf'(x)/(2c_1), rf'(x)/(2c_2))$  and thus we obtain

$$[\text{The Bellman equation}] \quad \max \left\{ \mathcal{L}f'(x) - h(x), \mathcal{H}f'(x) \right\} = 0, \quad (41)$$

where  $\mathcal{L}$  and  $\mathcal{H}$  denote the operators

$$\mathcal{L}p = -\frac{1}{4c}p^2 - yp, \quad \mathcal{H}p = p - r, \quad p \in \mathbb{R}.$$

As suggested by the way that (41) is derived from the dynamic description (39), rejections should take place at times when the state belongs to the part of the space where  $\mathcal{H}U' = 0$ . Heuristically, it is plausible that rejections should occur only when the state exceeds a threshold, and so it is to be expected that first and second arguments on the l.h.s. of equation (41) vanish for  $x < \beta$  and  $x > \beta$ , respectively, where  $\beta$  is the free boundary point. This also suggests that the minimizer plays optimally by selecting a  $\beta$ -barrier strategy.

To get information on the free boundary point of (41), note that  $\mathcal{L}U' - h = \mathcal{H}U' = 0$  has to hold at this point (at least in case when the solution is  $\mathcal{C}^1$ ). We thus let the free boundary point,  $\beta_0$ , be defined as the unique solution  $\beta$  of the equation

$$\frac{r^2}{4c} + ry + h(\beta) = 0, \quad (42)$$

in case it exists.

Note that there exists a unique solution in case  $r^2/(4c) + ry + h(0) \leq 0$  and  $r^2/(4c) + ry + h(D) \geq 0$ , as follows from the assumption that  $h$  is strictly increasing. Otherwise, equation (42) has no solution. When  $r^2/(4c) + ry + h(D) < 0$  we define  $\beta_0 = D$ . In case  $r^2/(4c) + ry + h(0) = r^2/(4c) + ry > 0$  Lemma 3.1 tells us that the value function is infinite and there is no need to define  $\beta_0$ . Thus

$$\beta_0 = \begin{cases} h^{-1}\left(\frac{-r^2}{4c} - ry\right), & -h(D) \leq \frac{r^2}{4c} + ry \leq -h(0), \\ D, & \frac{r^2}{4c} + ry < -h(D). \end{cases} \quad (43)$$

As for a solution to the Bellman equation, we let  $g : [0, D] \rightarrow \mathbb{R}$  be defined by

$$g(x) = \begin{cases} \int_0^x 2c \left( -y - \sqrt{y^2 - \frac{h(u)}{c}} \right) du, & 0 \leq x \leq \beta_0, \\ g(\beta_0) + r(x - \beta_0), & \beta_0 < x \leq D, \end{cases}$$

where in the case  $\beta_0 = D$ ,  $g$  is defined on  $[0, D]$  through the first line alone. Then the expression under the root is always positive. To see this notice that from the discussion above and the definition of  $\beta_0$  it follows that in case  $-y \geq r/(4c)$  one has  $r^2/(4c) + ry + h(\beta_0) \leq 0$ . By simple algebraic manipulation it follows that

$$(r/(2\sqrt{c}) + y\sqrt{c})^2 - y^2c + h(\beta_0) = r^2/(4c) + ry + h(\beta_0) \leq 0,$$

which yields  $y^2 - h(\beta_0)/c \geq 0$ . By the monotonicity of  $h$  it follows that for every  $u \in [0, \beta_0]$  one has  $y^2 - h(u)/c \geq 0$ .

Moreover,  $g$  satisfies the boundary condition  $g(0) = 0$  and it is easy to check that it is continuously differentiable on  $(0, \beta_0)$  and on  $(\beta_0, D)$  (and continuous on all of  $[0, D]$ ) and satisfies on these two intervals the two parts of the Bellman equation, namely  $\mathcal{L}g' - h = 0$  and  $\mathcal{H}g' = 0$ , respectively. Thus if  $\beta_0 = D$  then  $g$  solves the Bellman equation in a classical sense. However, if  $\beta_0 < D$  it may occur that  $g$  is not differentiable at  $\beta_0$ . Specifically, when  $\beta_0 < D$ ,  $g$  is not differentiable at  $\beta_0$  if and only if

$$\frac{r}{4c} \leq -y < \left( \frac{r}{4c} + \frac{h(D)}{r} \right) \wedge \frac{r}{2c} \quad (44)$$

Moreover, under (44), the left- and right-derivative at  $\beta_0$  satisfy  $g'_L(\beta_0) < g'_R(\beta_0)$ . Although  $g$  is not always a classical solution to the Bellman equation, we will prove in the next section that it is a viscosity solution (suitably defined) to that equation.

### 3.4 Viscosity solutions

In this section we consider the Bellman equation (41) in viscosity sense and characterize the value function  $U$  of the hitting time game as the unique solution of that equation, with suitable boundary conditions. The equality  $U = g$  is finally obtained when we prove, in the next section, that  $g$  also solves the Bellman equation.

For a set  $B \subset \mathbb{R}$  denote by  $\bar{B}$  its closure.

**Definition 3.2** [*Viscosity solution*]

*i. For  $B = (0, D)$  or  $B = (0, D]$ , a continuous function  $f : [0, D] \rightarrow \mathbb{R}$  is said to be a viscosity supersolution (resp., subsolution) of (41) on  $B$  if for every  $x \in B$  and every  $\phi \in \mathcal{C}^1([0, D], \mathbb{R})$  for which  $f - \phi$  has a global minimum (maximum) on  $B$  at  $x$  one has*

$$\max \left\{ \mathcal{L}\phi'(x) - h(x), \mathcal{H}\phi'(x) \right\} \geq 0 \quad (\leq 0).$$

*ii. A function  $f : [0, D] \rightarrow \mathbb{R}$  is said to be a viscosity solution of (41) on  $[0, D]$  if it is a viscosity subsolution of (41) on  $(0, D)$ , a viscosity supersolution of (41) on  $(0, D]$ , and  $f(0) = 0$ .*

The definition accounts for a Dirichlet boundary condition at 0 as well as a state constraint boundary condition at  $D$ , consisting of the requirement that  $f$  is a supersolution up to the boundary point  $D$  (see Section IV.5 of [5]).

**Theorem 3.1** *The value function  $U$  is the unique viscosity solution of (41) on  $[0, D]$ .*

The proof is performed in two steps: we first show that  $U$  satisfies the Bellman equation and then prove uniqueness.

**Proof that  $U$  is a viscosity solution:** By Lemma 3.1 and the equality  $U = V$  established in Proposition 3.1,  $U$  is continuous and null at zero. It remains to prove that  $U$  is a subsolution on  $(0, D)$  and a supersolution on  $(0, D]$ .

**Subsolution on  $(0, D)$ :** Fix  $x \in (0, D)$  and let  $\phi \in \mathcal{C}^1([0, D], \mathbb{R})$  be such that  $U - \phi$  has a global maximum at  $x$ . We can assume w.l.o.g. that  $U(x) = \phi(x)$ . Thus  $U \leq \phi$  on  $[0, D]$ . We need to show that

$$\mathcal{L}\phi'(x) - h(x) \leq 0 \quad (45)$$

and

$$\mathcal{H}\phi'(x) \leq 0. \quad (46)$$

Consider  $\delta > 0$  such that  $x - \delta \in [0, D]$ . The proof of (46) is based on the inequality  $U(x) - U(x - \delta) \leq r\delta$ , that is an extension of the two-sided inequality stated in Lemma 3.1. Let us show this inequality. Given a strategy  $\bar{\alpha} \in \bar{\mathcal{A}}_{x-\delta}$  let  $\tilde{\alpha} \in \bar{\mathcal{A}}_x$  be defined as  $\tilde{\alpha}[\psi] = \delta + \bar{\alpha}[\psi]$ . Then the responses of the two strategies to  $\psi$  are related to each other by  $\tilde{\varphi} = \bar{\varphi}$  and  $\tilde{\varrho} = \bar{\varrho} + \delta$ . As a result,

$$\bar{c}(x, \psi, \tilde{\alpha}[\psi]) = \bar{c}(x - \delta, \psi, \bar{\alpha}[\psi]) + r\delta.$$

Hence  $U(x) \leq \sup_{\psi} \bar{c}(x - \delta, \psi, \bar{\alpha}[\psi]) + r\delta$  and since  $\bar{\alpha}$  is arbitrary, we have  $U(x) \leq U(x - \delta) + r\delta$ . Therefore  $\phi(x) - \phi(x - \delta) \leq U(x) - U(x - \delta) \leq r\delta$ , establishing (46).

To prove (45) we let  $\delta > 0$  be such that  $0 < x - \delta < x + \delta < D$ , define

$$\tau_\delta := \delta \wedge \inf\{t \geq 0 : |\varphi(t) - x| \geq \delta\} \quad (47)$$

and use the dynamic programming principle (DPP), that states

$$U(x) = \inf_{\bar{\alpha} \in \bar{\mathcal{A}}_x} \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}^\delta]} \left[ \int_0^{\tau_\delta} (-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))) dt + r\bar{\alpha}[\psi](\tau_\delta) + U(\varphi(\tau_\delta)) \right].$$

Consider a strategy  $\bar{\alpha}$  with the property that for every  $\psi \in \mathcal{Q}[x, \bar{\alpha}^\delta]$  and every  $t \in [0, \tau_\delta]$  one has  $\bar{\alpha}[\psi](t) = 0$ . By the chain rule,

$$U(x) = \phi(x) = - \int_0^{\tau_\delta} \phi'(\varphi(t))(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y) dt + \phi(\varphi(\tau_\delta)).$$

Using the DPP, the inequality  $U \leq \phi$ , and the above display, we obtain

$$\begin{aligned} U(x) &\leq \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}^\delta]} \left[ \int_0^{\tau_\delta} (-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))) dt + U(\varphi(\tau_\delta)) \right] \\ &\leq \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}^\delta]} \left[ \int_0^{\tau_\delta} (-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))) dt + \phi(\varphi(\tau_\delta)) \right] \\ &= \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}^\delta]} \left[ \phi(x) + \int_0^{\tau_\delta} (-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) \right. \\ &\quad \left. + \phi'(\varphi(t))(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)) dt \right]. \end{aligned}$$

Recalling that  $U(x) = \phi(x)$ , we have

$$\sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}^\delta]} \int_0^{\tau_\delta} (-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + \phi'(\varphi(t))(\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)) dt \geq 0$$

for every  $\delta > 0$ . Hence

$$\sup_{\dot{\psi}^1(0), \dot{\psi}^2(0) \in \mathbb{R}} \left[ -c_1(\dot{\psi}^1)^2(0) - c_2(\dot{\psi}^2)^2(0) + h(x) + \phi'(x)(\dot{\psi}^1(0) - \dot{\psi}^2(0) + y) \right] \geq 0. \quad (48)$$

The maximum in (48) is achieved by  $(\dot{\psi}^1(0), \dot{\psi}^2(0))^* = (\phi'(x)/(2c_1), \phi'(x)/(2c_1))$  and equals  $-\mathcal{L}\phi'(x) + h(x)$ , and (45) follows.

**Supersolution on  $(0, D]$ :** Fix  $x \in (0, D]$  and let  $\phi \in \mathcal{C}^1([0, D], \mathbb{R})$  be such that  $U - \phi$  has a global minimum at  $x$ . Assume w.l.o.g. that  $U(x) = \phi(x)$ . Thus  $U \geq \phi$  on  $[0, D]$ . We need to show that either

$$\mathcal{L}\phi'(x) - h(x) \geq 0 \quad (49)$$

or

$$\mathcal{H}\phi'(x) \geq 0. \quad (50)$$

Arguing by contradiction, assume that neither of the above inequalities hold. Then one can find  $\varepsilon > 0$  such that

$$\mathcal{H}\phi' \leq -2\varepsilon \quad \text{and} \quad \mathcal{L}\phi' - h \leq -2\varepsilon \quad (51)$$

hold on  $\overline{B}_\delta(x) := \{\xi \in [0, D] : |x - \xi| \leq \delta\}$ . Denote  $\psi_x(t) = (\phi'(x)t/(2c_1), \phi'(x)t/(2c_1))$ ,  $t \geq 0$ . Recall the definition of  $\tau_\delta$  from (47). Let  $\overline{\mathcal{A}}_x(\delta)$  denote the collection of strategies  $\bar{\alpha} \in \overline{\mathcal{A}}_x$  that have the property  $\varphi(\tau_\delta) \in \overline{B}_\delta(x)$  for all  $\psi$ , where  $\varphi$  denotes the dynamics under  $\psi$  and  $\bar{\alpha}$ . Pick an arbitrary  $\bar{\alpha}_x \in \overline{\mathcal{A}}_x(\delta)$  and denote  $\varrho_x = \bar{\alpha}_x[\psi_x]$ . The associated dynamics are given by  $\varphi_x = x + y \cdot \psi_x^1 - \psi_x^1 - \varrho_x$ . Let  $\tau_\delta$  be as in (47) with  $\varphi = \varphi_x$ . An application of the chain rule gives

$$\begin{aligned} & \phi(\varphi_x(\tau_\delta)) - \phi(x) \\ &= - \int_0^{\tau_\delta} (-c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x(t))) dt - r\varrho_x(\tau_\delta) \\ & \quad + \int_0^{\tau_\delta} (-c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x(t)) + \phi'(\varphi_x(t))(\dot{\psi}_x^1(t) - \dot{\psi}_x^1(t) + y)) dt \\ & \quad + \int_{[0, \tau_\delta]} (r - \phi'(\varphi_x(t))) d\varrho_x(t). \end{aligned} \quad (52)$$

We will bound from below the last three terms on the r.h.s. of (52). By (51), the definition of  $\psi_x$ , and the continuity of the functions  $\phi'$  and  $h$  it follows that there is sufficiently small  $\delta > 0$  such that for every  $t \in [0, \tau_\delta]$  one has

$$\begin{aligned} & -c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x(t)) + \phi'(\varphi_x(t))(\dot{\psi}_x^1(t) - \dot{\psi}_x^1(t) + y) \\ & \geq -c_1(\dot{\psi}_x^1)^2(0) - c_2(\dot{\psi}_x^2)^2(0) + h(\varphi_x(0)) + \phi'(\varphi_x(0))(\dot{\psi}_x^1(0) - \dot{\psi}_x^1(0) + y) - \varepsilon \\ & = -\mathcal{L}\phi'(\varphi_x(0)) + h(\varphi_x(0)) - \varepsilon = -\mathcal{L}\phi'(x) + h(x) - \varepsilon \geq \varepsilon, \end{aligned}$$

where the first equality follows by the definition of  $\psi_x$  and by (40). Therefore the second term in (52) is bounded below by  $\varepsilon\tau_\delta$ . By using similar arguments it follows from (51) that for sufficiently small  $\delta$

$$\int_{[0, \tau_\delta]} (r - \phi'(\varphi_x(t))) d\varrho_x(t) \geq \varepsilon\varrho_x(\tau_\delta).$$

Using the last two inequalities in (52), we obtain

$$\begin{aligned} & \phi(\varphi_x(\tau_\delta)) - \phi(x) \\ & \geq - \int_0^{\tau_\delta} (-c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x(t))) dt - r\varrho_x(\tau_\delta) + \varepsilon(\tau_\delta + \varrho_x(\tau_\delta)). \end{aligned} \quad (53)$$

We now show that the sum  $\tau_\delta + \varrho_x(\tau_\delta)$  is bounded below by  $C(x)\delta$ , where  $C(x)$  is a positive constant that depends solely on  $x$ . To this end, assume that  $\tau_\delta < \delta$  (otherwise one may simply take  $C(x) = 1$ ). Therefore, by the definition of  $\bar{\mathcal{A}}_x(\delta)$   $|\varphi_x(\tau_\delta) - x| = \delta$ . Notice that if  $\varrho_x(\tau_\delta) \leq \delta/2$  then in case  $y + \dot{\psi}_x^1(t) - \dot{\psi}_x^2(t) = y + \phi'(x)/(2c) < 0$  (resp.,  $> 0$ ), the function  $yt + \psi_x^1(t) - \psi_x^2(t)$  must pass  $-\delta/2$  (resp.,  $3\delta/2$ ) units over the time interval  $[0, \tau_\delta]$ . Therefore  $|yt + \frac{1}{2c}\phi'(x)t|$  must pass at least  $\delta/2$  units. Hence  $\tau_\delta \geq \{2|\frac{1}{2c}\phi'(x) + y|\}^{-1}\delta$ . Overall,<sup>1</sup>

$$\tau_\delta + \varrho_x(\tau_\delta) \geq \min \left\{ \frac{\delta}{2}, \frac{1}{2|\frac{1}{2c}\phi'(x) + y|} \delta \right\} =: C(x)\delta,$$

and

$$\phi(\varphi_x(\tau_\delta)) - \phi(x) \geq - \int_0^{\tau_\delta} (-c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x(t))) dt - r\varrho_x(\tau_\delta) + \varepsilon C(x)\delta. \quad (54)$$

By the dynamic programming principle, using the fact  $\tau_\delta \leq \tau$ , it follows that

$$\begin{aligned} U(x) &= \inf_{\bar{\alpha} \in \bar{\mathcal{A}}_x(\delta)} \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}]} \left[ \int_0^{\tau_\delta} [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))] dt + r\bar{\alpha}[\psi](\tau_\delta) + U(\varphi(\tau_\delta)) \right] \\ &\geq \inf_{\bar{\alpha} \in \bar{\mathcal{A}}_x} \left[ \int_0^{\tau_\delta} [-c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x(t))] dt + r\varrho_x(\tau_\delta) + U(\varphi_x(\tau_\delta)) \right]. \end{aligned}$$

Therefore, there exists a strategy  $\alpha^* \in \bar{\mathcal{A}}_x(\delta)$  such that, denoting  $\varrho_x^* = \alpha^*[\psi_x]$ ,

$$U(x) \geq \int_0^{\tau_\delta} [-c_1(\dot{\psi}_x^1)^2(t) - c_2(\dot{\psi}_x^2)^2(t) + h(\varphi_x^*(t))] dt + r\varrho_x^*(\tau_\delta) + U(\varphi_x^*(\tau_\delta)) - \frac{1}{2}\varepsilon C(x)\delta, \quad (55)$$

where  $\varphi_x^* = x + t \cdot \psi_x^1 - \psi_x^2 - \varrho_x^*$ . Recalling that  $\bar{\alpha}_x \in \bar{\mathcal{A}}_x(\delta)$  is arbitrary, letting it be  $\alpha^*$ , and substituting  $\varphi_x = \varphi_x^*$  in (54), it follows from inequalities (54) and (55) that

$$\phi(\varphi_x^*(\tau_\delta)) - U(\varphi_x^*(\tau_\delta)) \geq \frac{1}{2}\varepsilon C(x)\delta,$$

which contradicts the fact that  $U \geq \phi$ .

<sup>1</sup>Note that the choice of  $C(x)$  is independent of the strategy  $\bar{\alpha}_x$ .

**Proof of uniqueness of viscosity solutions:** We prove uniqueness by the comparison principle. Assume that  $v$  is a supersolution on  $(0, D]$  and  $u$  a subsolution on  $(0, D)$ . We show that  $u \leq v$ .

Arguing by contradiction, let us assume to the contrary, that  $u \not\leq v$ . Then there exist  $x \in [0, D]$  and  $\delta \in (0, 1)$  such that  $u(x) - (1 + \delta)v(x) > 0$ . Fix such  $\delta$ . Fix also  $z \in \arg \max_{[0, D]} [u - (1 + \delta)v]$ . Since  $u(0) = v(0) = 0$  it follows that  $z > 0$ .

For  $\varepsilon > 0$  define the following test function

$$\xi^\varepsilon(x, y) = u(x) - (1 + \delta)v(y) - \left( \frac{x - y}{\varepsilon} - (-1 - z) \right)^2 - (y - z)^2 \quad x, y \in [0, D].$$

Note that  $z + \varepsilon(-1 - z)$  is in  $[0, D]$  for every small enough  $\varepsilon$ . Also,

$$\xi^\varepsilon(z + \varepsilon(-1 - z), z) \geq u(z) - (1 + \delta)v(z) - \text{mod}(C\varepsilon), \quad (56)$$

where here and in what follows  $\text{mod}$  is the modulus of continuity of  $u$  and  $v$ , and  $C$  is the constant  $1 + D$ . We also have

$$\xi^\varepsilon(x, y) \leq u(x) - (1 + \delta)v(x) + 2\text{mod}(|x - y|) - \left( \frac{x - y}{\varepsilon} - (-1 - z) \right)^2 - (y - z)^2. \quad (57)$$

For each  $\varepsilon > 0$  pick  $(x_\varepsilon, y_\varepsilon) \in \arg \max_{[0, D]^2} [\xi^\varepsilon(x, y)]$ . From equations (56) and (57), substituting  $(x, y) = (x_\varepsilon, y_\varepsilon)$  and using  $u(z) - (1 + \delta)v(z) \geq u(x_\varepsilon) - (1 + \delta)v(x_\varepsilon)$ , it follows that

$$\left( \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} - (-1 - z) \right)^2 + (y_\varepsilon - z)^2 \leq \text{mod}(C\varepsilon) + 2\text{mod}(|x_\varepsilon - y_\varepsilon|). \quad (58)$$

The above yields two properties that are being used in the sequel.

(a)  $\lim_{\varepsilon \rightarrow \infty} x_\varepsilon = \lim_{\varepsilon \rightarrow \infty} y_\varepsilon = z$ .

(b) For every  $\varepsilon$  small enough one has  $x_\varepsilon < y_\varepsilon$ .

Property (a) follows since the r.h.s. of equation (58) is bounded above and therefore there is  $m > 0$  such that  $|x_\varepsilon - y_\varepsilon| < m\varepsilon$ . This also yields that for sufficiently small  $\varepsilon$  the r.h.s. of (58) is bounded above by  $z^2$ . Therefore,  $|x_\varepsilon - y_\varepsilon - \varepsilon(-1 - z)| \leq \varepsilon z$ . By these two inequalities it follows that

$$x_\varepsilon \leq y_\varepsilon + \varepsilon(-1 - z) + \varepsilon z \leq y_\varepsilon - \varepsilon < y_\varepsilon,$$

which proves property (b).

Now consider the maps

$$\phi_u(x) := (1 + \delta)v(y_\varepsilon) + \left( \frac{x - y_\varepsilon}{\varepsilon} - (-1 - z) \right)^2 + (y_\varepsilon - z)^2, \quad (59)$$

$$\phi_v(y) := \frac{1}{1 + \delta} \left[ u(x_\varepsilon) - \left( \frac{x_\varepsilon - y}{\varepsilon} - (-1 - z) \right)^2 - (y - z)^2 \right]. \quad (60)$$

Then  $u - \phi_u$  has a maximum at  $x_\varepsilon$  and  $v - \phi_v$  has a minimum at  $y_\varepsilon$ . Note that by property (a) and the fact  $z > 0$ ,  $\min(x_\varepsilon, y_\varepsilon) > 0$  provided  $\varepsilon$  is small, and that by property (b),  $x_\varepsilon < D$ . As a result, the definition of a viscosity subsolution on  $(0, D)$  and supersolution on  $(0, D]$  yields

$$\max \left\{ \mathcal{L}\phi'_u(x_\varepsilon) - h(x_\varepsilon), \mathcal{H}\phi'_u(x_\varepsilon) \right\} \leq 0 \quad (61)$$

$$\max \left\{ \mathcal{L}\phi'_v(y_\varepsilon) - h(y_\varepsilon), \mathcal{H}\phi'_v(y_\varepsilon) \right\} \geq 0. \quad (62)$$

Now,

$$\phi'_u(x_\varepsilon) = \frac{2}{\varepsilon} \left( \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} - (-1 - z) \right) \quad (63)$$

$$\begin{aligned} \phi'_v(y_\varepsilon) &= \frac{1}{1 + \delta} \left[ \frac{2}{\varepsilon} \left( \frac{x_\varepsilon - y_\varepsilon}{\varepsilon} - (-1 - z) \right) - 2(y_\varepsilon - z) \right] \\ &= \frac{1}{1 + \delta} \phi'_u(x_\varepsilon) - \frac{2}{1 + \delta} \cdot (y_\varepsilon - z). \end{aligned} \quad (64)$$

By (61) it follows that  $\mathcal{H}\phi'_u(x_\varepsilon) \leq 0$ , that is,  $\phi'_u(x_\varepsilon) \leq r$ . Together with property (a) above and equation (64) it follows that  $\mathcal{H}\phi'_v(y_\varepsilon) < 0$ . Therefore, (62) yields that

$$\mathcal{L}\phi'_v(y_\varepsilon) - h(y_\varepsilon) \geq 0 \quad (65)$$

We now claim that the sequence  $\{\phi'_v(y_\varepsilon)\}_\varepsilon$  is positive and bounded away from zero. Denote by  $q = \lim \phi'_v(y_\varepsilon) = \phi'_v(z)$ , where property (a) is used. Using again property (a) and (65) it follows that  $\mathcal{L}(q) - h(z) \geq 0$ . Since  $z > 0$ , we get  $h(z) > 0$ , and therefore  $\mathcal{L}(q) > 0$ , which in turn yields  $q > 0$ . By considering a subsequent we may assume w.l.o.g. that  $\{\phi'_v(y_\varepsilon)\}$  is bounded away from zero. Hence, for sufficiently small  $\varepsilon$  one has

$$\left| \frac{2}{1 + \delta} \cdot (y_\varepsilon - z) \right| \leq \frac{\delta}{2(1 + \delta)} \phi'_v(y_\varepsilon),$$

thanks to the fact that the l.h.s. converges to zero as  $\varepsilon \rightarrow 0$  by property (a). Together with (64) we get that for sufficiently small  $\varepsilon$

$$\phi'_v(y_\varepsilon) \leq \frac{1}{1 + \delta} \phi'_u(x_\varepsilon) + \frac{\delta}{2(1 + \delta)} \phi'_v(y_\varepsilon). \quad (66)$$

Together with the inequality  $\phi'_u(x_\varepsilon) \leq r$  we get that

$$\left(1 + \frac{\delta}{2}\right) \phi'_v(y_\varepsilon) \leq \phi'_u(x_\varepsilon) \leq r. \quad (67)$$

If we show that for sufficiently small parameters  $\varepsilon$  and  $\delta$  one has

$$\mathcal{L}(r) - h(x_\varepsilon) > 0 \quad (68)$$

and

$$\mathcal{L}\left(\left(1 + \frac{\delta}{2}\right) \phi'_v(y_\varepsilon)\right) - h(x_\varepsilon) > 0 \quad (69)$$

then by the concavity of  $p \mapsto \mathcal{L}(p)$  and (67)–(69) it follows that  $\mathcal{L}(\phi'_u(x_\varepsilon)) - h(x_\varepsilon) > 0$  in contradiction to (61). We now prove that inequalities (68) and (69) hold.

**Proof that inequality (68) holds:** By property (b) and by the monotonicity of  $h$  it follows that  $\mathcal{L}(r) - h(x_\varepsilon) > \mathcal{L}(r) - h(y_\varepsilon)$ . Therefore, it is sufficient to show that  $\mathcal{L}(r) - h(y_\varepsilon) \geq 0$ . To this end, assume to the contrary that  $\mathcal{L}(r) - h(y_\varepsilon) < 0$ . Recall that from (67) one has  $\phi'_v(y_\varepsilon) < r$ . Fix  $p_\varepsilon \in (\phi'_v(y_\varepsilon), r)$  sufficiently close to  $r$  such that  $\mathcal{L}(p_\varepsilon) - h(y_\varepsilon) < 0$ . Then

$$\max \left\{ \mathcal{L}(p_\varepsilon) - h(y_\varepsilon), \mathcal{H}(p_\varepsilon) \right\} < 0 \quad (70)$$

Denote  $b = \min_{\zeta \in [y_\varepsilon, D]} (v(\zeta) - p_\varepsilon \zeta)$  and consider the function  $\phi(\zeta) = p_\varepsilon \zeta + b$ ,  $\zeta \in [0, D]$ . Pick  $z_\varepsilon \in [y_\varepsilon, D]$  for which  $v(z_\varepsilon) = \phi(z_\varepsilon)$ . By the definition of  $b$  one has  $v \geq \phi$  on  $[y_\varepsilon, D]$ . Moreover, since  $p_\varepsilon > \phi'_v(y_\varepsilon)$  it follows that  $v \geq \phi$  in fact holds on a larger interval, namely on  $(y_\varepsilon - c, D]$  for some  $c > 0$ . Thus we may consider  $\phi$  as a test function in the definition of  $v$  as a supersolution on  $(0, D]$ , by which

$$\max \left\{ \mathcal{L}(p_\varepsilon) - h(z_\varepsilon), \mathcal{H}(p_\varepsilon) \right\} \geq 0.$$

This, the fact that  $z_\varepsilon \geq y_\varepsilon$ , and the monotonicity of  $h$  give a contradiction to (70).

**Proof that inequality (69) holds:** From property (a), together with  $z > 0$  we may assume w.l.o.g. that for every  $\varepsilon$  one has  $y_\varepsilon > 0$ . Also recall that we assumed that  $\phi'_v(y_\varepsilon) > 0$ . Similarly to (67), one may take sufficiently small  $\varepsilon$  such that

$$0 < \phi'_v(y_\varepsilon) < (1 + \frac{\delta}{2})\phi'_v(y_\varepsilon) < (1 + \frac{2\delta}{3})\phi'_v(y_\varepsilon) < r \quad (71)$$

and consequently

$$0 < q < (1 + \frac{\delta}{2})q < (1 + \frac{2\delta}{3})q \leq r, \quad (72)$$

where the fact that  $q > 0$  has been proved earlier. By taking  $\varepsilon \rightarrow 0$  in (68) and (65) it follows from the limits  $\lim x_\varepsilon = \lim y_\varepsilon = z$  (property (a)) and  $\lim \phi'_v(y_\varepsilon) = q$  that

$$\mathcal{L}(q) - h(z) \geq 0 \quad \text{and} \quad \mathcal{L}(r) - h(z) \geq 0. \quad (73)$$

From (73) and (72) and the strict concavity of  $p \mapsto \mathcal{L}(p)$  it follows that  $\mathcal{L}((1 + \frac{\delta}{2})q) - h(z) > 0$ . By the limits  $\lim x_\varepsilon = z$  and  $\lim \phi'_v(y_\varepsilon) = q$  it follows that for sufficiently small  $\varepsilon$  inequality (69) holds.  $\square$

### Explicit viscosity solution of the Bellman equation

We now prove that the function  $g$  is the value function. We do that by showing that  $g$  is a viscosity solution of the Bellman equation, which admits a unique solution.

**Theorem 3.2** *The function  $g$  is a viscosity solution of (41) on  $[0, D]$ .*

Note that this result combined with Theorem 3.1 shows  $U = g$ .

**Proof:** One verifies directly that the function  $g$  satisfies, in the classical sense,  $\mathcal{L}g' - h = 0$  and  $\mathcal{H}g' \leq 0$  on  $(0, \beta_0)$ , and (in case  $\beta_0 < D$ ) also  $\mathcal{L}g' - h \leq 0$  and  $\mathcal{H}g' = 0$  on  $(\beta_0, D)$ . Also,



if  $r/(2c) \leq -y < r/(4c) + h(D)/r$  then as follows from Section 3.3  $\beta_0 < D$  and the function  $g$  is differentiable at  $\beta_0$  and one can easily verify that  $\mathcal{L}g'(\beta_0) - h(\beta_0) = \mathcal{H}g'(\beta_0) = 0$ . Moreover,  $g(0) = 0$ . Hence it remains to verify the viscosity solution definition at the points  $\beta_0$  and  $D$  only.

**Subsolution for  $\beta_0$ :** The case  $r/(2c) \leq -y < r/(4c) + h(D)/r$  was discussed before. We analyze the two remaining cases  $-y \geq r/(4c) + h(D)/r$  and  $r/(4c) \leq -y < (r/(4c) + h(D)/r) \wedge r/(2c)$ . If  $-y \geq r/(4c) + h(D)/r$  then as we saw in Section 3.3,  $\beta_0 = D$ , and the subsolution property need not be verified since by definition the viscosity subsolution property should be verified only on the interval  $(0, D)$ . If  $r/(4c) \leq -y < (r/(4c) + h(D)/r) \wedge r/(2c)$  then as we saw in Section 3.3,  $g'_L(\beta_0) < g'_R(\beta_0)$ . Therefore, there is no function  $\phi \in \mathcal{C}^1([0, D], \mathbb{R})$  such that  $g - \phi$  has a global maximum at  $\beta_0$ , and the claim holds vacuously.

**Supersolution for  $\beta_0$  and  $D$ :** Fix  $x \in \{\beta_0, D\}$ . Let  $\phi \in \mathcal{C}^1([0, D], \mathbb{R})$  be such that  $g - \phi$  has a global minimum at  $x$ . We can assume w.l.o.g. that  $g(x) = \phi(x)$ . Thus,  $g \geq \phi$  on  $[0, D]$ . We need to show that

$$\mathcal{L}\phi'(x) - h(x) \geq 0 \quad (74)$$

or

$$\mathcal{H}\phi'(x) \geq 0. \quad (75)$$

As before, we analyze the two cases  $-y \geq r/(4c) + h(D)/r$  and  $r/(4c) \leq -y < (r/(4c) + h(D)/r) \wedge r/(2c)$ . If  $r/(4c) \leq -y < (r/(4c) + h(D)/r) \wedge r/(2c)$  then as was shown in Section 3.3  $\beta_0 < D$ ,  $g'_L(\beta_0) < g'_R(\beta_0)$ , and  $g'(D) = r$ . For  $x = D$  (75) holds since  $\phi'(D) \geq g'(D)$ . For  $x = \beta_0$ . Since  $g'_L(\beta_0) \leq \phi'(\beta_0) \leq g'_R(\beta_0)$  it follows that  $2c(-y - \sqrt{y^2 - h(\beta_0)/c}) \leq \phi'(\beta_0) \leq r$ . One can verify that

$$\mathcal{L}(2c(-y - \sqrt{y^2 - h(\beta_0)/c}) - h(\beta_0)) \geq 0 \quad \text{and} \quad \mathcal{L}r - h(\beta_0) \geq 0. \quad (76)$$

From the above and by the concavity of  $\mathcal{L}$  it follows that (74) holds.

If  $-y \geq r/(4c) + h(D)/r$  then as was shown in Section 3.3  $\beta_0 = D$ . If  $\phi'(D) \geq r$  then (75) holds. If  $\phi'(D) \leq r$  then  $2c(-y - \sqrt{y^2 - h(\beta_0)/c}) = g'_L(\beta_0) \leq \phi'(\beta_0) \leq r$  and from (76) as before it follows that (74) holds. □

### 3.5 Optimal strategy

The following theorem shows that the  $\beta_0$ -barrier strategy is optimal. We defined the barrier strategies only for the original game. The definition for the hitting time game is similar.

**Definition 3.3** Fix  $(x, \beta) \in [0, D]^2$ . The strategy  $\bar{\alpha}_\beta \in \bar{\mathcal{A}}_x$  is called a  $\beta$ -barrier strategy if for every  $\psi \in \mathcal{Q}[x, \bar{\alpha}_\beta]$  one has  $(\varphi, 0, \bar{\alpha}_\beta)(\psi) = \Gamma_{[0, \gamma]}(\psi)$  on the time interval  $[0, \tau]$ .

To state the result let us add a piece of notation. Denote by  $\bar{J}(x, \bar{\alpha})$  the optimal cost achieved by the maximizer in the hitting time game when the initial state is  $x$  and the minimizer uses the strategy  $\bar{\alpha} \in \bar{\mathcal{A}}_x$ . That is,

$$\bar{J}(x, \bar{\alpha}) := \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}]} \bar{c}(x, \psi, \bar{\alpha}[\psi]), \quad x \in [0, D], \quad \bar{\alpha} \in \bar{\mathcal{A}}_x. \quad (77)$$

**Theorem 3.3** Let  $\bar{\alpha}_{\beta_0}$  be the  $\beta_0$ -barrier strategy. For every  $x \in [0, D]$  one has  $g(x) = \bar{J}(x, \bar{\alpha}_{\beta_0})$ .

**Proof:** Recall that  $U = \inf_{\bar{\mathcal{A}}} \bar{J}(\cdot, \bar{\alpha})$ . Therefore,  $J(\cdot, \bar{\alpha}_{\beta_0}) \geq U$ . If we show that  $g \geq \bar{J}(\cdot, \bar{\alpha}_{\beta_0})$  then together with Theorem 3.2 it will follow that  $U = g \geq \bar{J}(\cdot, \bar{\alpha}_{\beta_0})$ , and so  $U = g = \bar{J}(\cdot, \bar{\alpha}_{\beta_0})$ . Hence in what follows we shall prove that  $g \geq \bar{J}(\cdot, \bar{\alpha}_{\beta_0})$ . By the definition of the  $\beta_0$ -barrier strategy and the function  $g$  it follows that for every  $x \in [\beta_0, D]$  one has  $g(x) = g(\beta_0) + r(x - \beta_0)$  and  $\bar{J}(x, \bar{\alpha}_{\beta_0}) = \bar{J}(\beta_0, \bar{\alpha}_{\beta_0}) + r(x - \beta_0)$ . Therefore, it is sufficient to prove that  $g(x) \geq \bar{J}(x, \bar{\alpha}_{\beta_0})$  only for  $x \in [0, \beta_0]$ . Apply the definition of  $\bar{J}$  for  $\bar{\alpha} = \bar{\alpha}_{\beta_0}$ , that is,

$$\bar{J}(x, \bar{\alpha}_{\beta_0}) = \sup_{\psi \in \mathcal{Q}[x, \bar{\alpha}_{\beta_0}]} \bar{c}(x, \psi, \bar{\alpha}_{\beta_0}[\psi]).$$

Then it is sufficient to show that for every  $x \in [0, \beta_0]$  and every  $\psi \in \mathcal{Q}[x, \bar{\alpha}_{\beta_0}]$  one has

$$\bar{c}(x, \psi, \bar{\alpha}_{\beta_0}[\psi]) = \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) \leq g(x),$$

where  $\varrho = \bar{\alpha}_{\beta_0}[\psi]$  and  $\varphi(t) = x + yt + \psi^1(t) - \psi^2(t) - \varrho(t)$ ,  $0 \leq t \leq \tau$ .

Fix such  $x$  and  $\psi$ . Applying the chain rule to  $g$  gives<sup>2</sup>

$$\begin{aligned} & \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) \\ & \leq \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) + g(\varphi(\tau)) \\ & = g(x) + \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)g'(\varphi(t))]dt \\ & \quad + \int_0^\tau [r - g'(\varphi(t))]d\varrho(t) \\ & \leq g(x) + \int_0^\tau [r - g'(\beta_0)]d\varrho(t), \end{aligned} \tag{78}$$

where we write  $g'(\beta_0)$  for  $g'_L(\beta_0)$ . The last inequality above holds due to the fact that for every  $0 \leq t \leq \tau$  one has  $\varphi(t) \in [0, \beta_0]$  and

$$\max_{p, q \in \mathbb{R}} [-c_1p^2(t) - c_2q^2 + h(\varphi(t)) + (p - q + y)g'(\varphi(t))] = -\mathcal{L}(g'(\varphi(t)) + h(\varphi(t))),$$

and the r.h.s. above equals zero by the choice of  $g$ .

Next, recall that the measure  $d\varrho$  charges only times  $t$  at which  $\{\varphi(t) = \beta_0\}$ . We analyze the last term in (78) separately for  $r/(2c) + y \leq 0$  and  $r/(2c) + y > 0$ . Recall by Section 3.3 that in case  $r/(2c) + y \leq 0$ , the function  $g$  is differentiable at  $\beta_0$  and satisfies  $g'(\beta_0) = r$ . Therefore,  $\int_0^\tau [r - g'(\beta_0)]d\varrho(t) = 0$  and from (78) it follows that  $g(x) \geq \bar{J}(x, \bar{\alpha}_{\beta_0})$ .

In case  $r/(2c) + y > 0$  the above argument does not hold since  $g'_L(\beta_0) < g'_R(\beta_0) = r$ . Hence we refine the proof. We will construct a sequence of functions  $\{g_\delta\}_\delta \subset \mathcal{C}^1([0, D], \mathbb{R}_+)$ ,

<sup>2</sup>Notice that since  $x \leq \beta_0$ , and by the definition of the  $\beta_0$ -barrier strategy the dynamics admit no jumps.

that converges uniformly to  $g$  as  $\delta \rightarrow 0$  with the property that for every  $x \in [0, \beta_0]$  and every  $\psi \in \mathcal{Q}[x, \bar{\alpha}_{\beta_0}]$  one has

$$\int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) \leq g_\delta(x), \quad (79)$$

by which the result thus follows upon taking  $\delta \rightarrow 0$ .

**Lemma 3.2** *There exists a sequence of functions  $\{g_\delta\}_\delta$  that satisfies the following conditions*

- (c1)  $\{g_\delta\}_\delta \subset \mathcal{C}^1([0, D], \mathbb{R}_+)$ .
- (c2)  $g'_\delta(\beta_0) = g'_R(\beta_0) = r$ .
- (c3) For every  $x \in [0, \beta_0]$  one has  $\mathcal{L}(g'_\delta(x)) - h(x) \geq 0$ .
- (c4)  $\lim_{\delta \rightarrow 0} \sup_{x \in [0, D]} |g_\delta(x) - g(x)| = 0$ .

The proof appears below. To complete the proof of the theorem, note by condition (c1) that one can apply the chain rule to  $g_\delta$ , and as in (78),

$$\begin{aligned} & \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) \\ & \leq \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) + g_\delta(\varphi(\tau)) \\ & = g_\delta(x) + \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)g'_\delta(\varphi(t))]dt \\ & \quad + \int_0^\tau [r - g'_\delta(\varphi(t))]d\varrho(t) \\ & \leq g_\delta(x) + \int_0^\tau [-\mathcal{L}(g'_\delta(\varphi(t))) + h(\varphi(t))]dt + \int_0^\tau [r - g'_\delta(\beta_0)]d\varrho(t). \end{aligned}$$

From conditions (c3) and (c2) we get that the second term on the r.h.s. is nonpositive and the third one equals zero, and therefore (79) holds.  $\square$

**Proof of Lemma 3.2:** In this proof we write  $g'(\beta_0)$  for  $g'_L(\beta_0)$ . Fix  $0 < \delta < \beta_0$ . Recall that the function  $g$  is not differentiable (only) at  $\beta_0$  and satisfies  $g'_L(\beta_0) < g'_R(\beta_0) = r$ . The idea of the proof is to construct a smooth function  $g_\delta$  that is close to  $g$  and such that  $g_\delta = g$  on  $[0, x_\delta]$  where  $x_\delta$  is a little smaller than  $\beta_0$ . The function  $g_\delta$  increases on the interval  $[x_\delta, \beta_0]$  and satisfies  $g'_\delta(\beta_0) = g'_R(\beta_0) = r$ . On  $[\beta_0, D]$  it is affine with slope  $r$ . We start with a linear interpolation. Let  $l_\delta$  be the linear interpolation between the two-dimensional points  $(\beta_0 - \delta, g'(\beta_0 - \delta))$  and  $(\beta_0, g'_R(\beta_0))$ . That is  $l_\delta(x) = g'(\beta_0 - \delta) + (x - (\beta_0 - \delta)) \frac{g'_R(\beta_0) - g'(\beta_0 - \delta)}{\delta}$ . Define  $x_\delta := \arg \max\{x \in [\beta_0 - \delta, \beta_0] : l_\delta(x) = g'(x)\}$ . Since  $l_\delta$  and  $g'$  are continuous and since  $l_\delta(\beta_0 - \delta) = g'(\beta_0 - \delta)$  and  $l_\delta(\beta_0) > g'_L(\beta_0)$  it follows that  $\beta_0 - \delta \leq x_\delta < \beta_0$ . Define

$$g_\delta(x) := \begin{cases} g(x) & 0 \leq x \leq x_\delta, \\ g(x_\delta) + \int_{x_\delta}^x l_\delta(u)du & x_\delta < x \leq \beta_0, \\ g_\delta(\beta_0) + r(x - \beta_0) & \beta_0 < x \leq D. \end{cases}$$

One can easily verify that (c1) and (c2) hold. Condition (c3) holds trivially on  $[0, x_\delta]$  since on this interval  $g_\delta = g$  and we chose  $g$  to satisfy  $\mathcal{L}(g'(x)) - h(x) = 0$  on  $[0, \beta_0]$ . We now

turn to the interval  $(x_\delta, \beta_0]$ . Fix  $x \in (x_\delta, \beta_0]$ . According to the definition of  $x_\delta$  one has  $g'_\delta(x) = l_\delta(x) > g'(x)$ . Also,  $g'_\delta(x) \leq g'_R(\beta_0) = r$ . Altogether,

$$g'(x) < g'_\delta(x) \leq r. \quad (80)$$

We claim that  $\mathcal{L}(g'_\delta(x)) - h(x) \geq 0$ . To this end, we show that

$$\mathcal{L}(g'(x)) - h(x) = 0 \quad \text{and} \quad \mathcal{L}(r) - h(x) \geq 0, \quad x \in (x_\delta, \beta_0] \quad (81)$$

and use inequality (80) and the concavity of  $\mathcal{L}$  to deduce that (c3) holds. We already showed that the first condition of (81) holds. As for the second, notice that since  $h$  is increasing, one has  $\mathcal{L}(r) - h(x) \geq \mathcal{L}(r) - h(\beta_0) \geq 0$ , where the last inequality follows by the definition of  $\beta_0$ .

To verify (c4) it is sufficient to show that  $\lim_{\delta \rightarrow 0} \sup_{x \in [x_\delta, \beta_0]} |g_\delta(x) - g(x)| = 0$ . Since  $g'_\delta(x) - g'(x) = l_\delta(x) - g'(x) > 0$  for every  $x \in [x_\delta, \beta_0]$  it follows that  $g_\delta(x) - g(x)$  increases on  $[x_\delta, \beta_0]$ , and therefore we need only to prove that  $\lim_{\delta \rightarrow 0} (g_\delta(\beta_0) - g(\beta_0)) = 0$ . This condition indeed holds, due to the fact that  $g_\delta(\beta_0) = g(x_\delta) + \int_{x_\delta}^{\beta_0} l_\delta(u) du$ ,  $\lim_{\delta \rightarrow 0} x_\delta = \beta_0$ , and  $l_\delta$  is bounded.  $\square$

As an immediate consequence we obtain an analogous result for the original game.

**Theorem 3.4** *Let  $\alpha_{\beta_0}$  be the  $\beta_0$ -barrier strategy in the original game. For every  $x \in [0, D]$  one has  $g(x) = J(x, \alpha_{\beta_0}) := \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, \psi, \alpha_{\beta_0}[\psi])$ .*

**Proof:** In the proof of Theorem 3.3, replace the set  $\mathcal{Q}[x, \bar{\alpha}_{\beta_0}]$  by  $\mathcal{P}^2$  and the hitting time  $\tau$  by any time  $T \in \mathbb{R}_+$  to obtain

$$g(x) = J(x, \alpha_{\beta_0}) = \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, \psi, \alpha_{\beta_0}[\psi]).$$

$\square$

### 3.6 Further properties of the game

This section is concerned with two particular controls for the maximizer, that besides being interesting by their own right, play an important role in the analysis of the asymptotic stochastic control problem in the forthcoming paper [2]. The first control regards the behavior of the maximizer when the dynamics are above the free boundary point  $\beta_0$ , and is used to show that the property of the  $\beta_0$ -barrier policy, by which the dynamics must initially jump from  $x > \beta_0$  to  $\beta_0$ , is necessary for optimality of the strategy (Subsection 3.6.1). The second control presented (Subsection 3.6.2) is the optimal response to the  $\beta_0$ -barrier strategy. Finally, in the case when  $h$  is convex, the latter control is shown (in Subsection 3.6.3) to correspond to a saddle point for the game. Specifically, by choosing this control, the maximizer can force a payoff of at least  $U(x)$  regardless of the strategy used by the minimizer.

#### 3.6.1 Dynamics outside the free boundary

For initial condition  $x > \beta_0$ , the optimal strategy that we have identified dictates that the dynamics make a jump to the point  $\beta_0$  at time zero. The initial (i.e., at time zero) cost associated with this jump is  $r(x - \beta_0)$ . We will show that any strategy that does not make

such an initial jump is suboptimal. We do so by identifying a control that, if used by the maximizing player, assures that under any strategy that does not perform such a jump, a larger cost is incurred up to the first time  $\beta_0$  is reached.

The precise formulation of this requires an additional parameter,  $\delta > 0$ , that can be made arbitrarily small. Fix  $x > \beta_0$ . Let  $\delta > 0$  be such that  $x > \beta_0 + \delta$ . Fix a strategy  $\bar{\alpha} \in \bar{\mathcal{A}}_x$ . Consider the control<sup>3</sup>

$$\psi(t) = \psi_\delta(t) := \begin{cases} (rt/(2c_1), -rt/(2c_2)) & 0 \leq t < T, \\ (rT/(2c_2), rT/(2c_2)) & t \geq T, \end{cases} \quad (82)$$

where  $T = T_\delta := \frac{r(D-\beta_0)}{h(\beta_0+\delta)-h(\beta_0)}$ . Denote also  $\varphi_{\bar{\alpha}} := \varphi[x, \psi, \bar{\alpha}[\psi]]$  and let

$$\tau = \tau_{\bar{\alpha}, \delta} := \inf\{t \geq 0 : \varphi_{\bar{\alpha}}(t) \leq \beta_0 + \delta\}$$

be the first time that the dynamics cross  $\beta_0 + \delta$ .

**Proposition 3.2** *If  $\tau > 0$  then one has*

$$\int_0^\tau h(\varphi_{\bar{\alpha}}(t))dt + r\bar{\alpha}[\psi](\tau) - \mathbb{I}(\tau, \psi) > r(x - (\beta_0 + \delta)).$$

*Notice that the l.h.s. of the above is the cost incurred up to time  $\tau$ , whereas the r.h.s. corresponds to cost of rejecting at time zero.*

**Proof of Proposition 3.2:** We analyze separately the cases  $0 < \tau \leq T$  and  $\tau > T$ . In case  $\tau < T$  the cost until time  $\tau$  is given by

$$\begin{aligned} & \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}_\delta^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) \\ & \geq (h(\beta_0 + \delta) - \frac{r^2}{4c})\tau + r\varrho(\tau) \end{aligned} \quad (83)$$

$$\geq r(x - \beta_0 - \delta) + (\frac{r^2}{4c} + ry + h(\beta_0))\tau + (h(\beta_0 + \delta) - h(\beta_0))\tau \quad (84)$$

$$= r(x - \beta_0 - \delta) + (h(\beta_0 + \delta) - h(\beta_0))\tau. \quad (85)$$

Above, inequality (83) follows since on the time interval  $[0, \tau]$  one has  $\varphi > \beta_0 + \delta$  and since  $h$  is increasing. Inequality (84) follows since  $\beta_0 + \delta \geq \varphi(\tau) = x + (\frac{r}{2c} + y)\tau - \varrho(\tau)$ . Equation (85) follows since in case  $\beta_0 < D$  one has  $r^2/(4c) + ry + h(\beta_0) = 0$  by (43). Since  $h$  is increasing and  $\tau > 0$ , the expression in (85) is strictly greater than  $r(x - (\beta_0 + \delta))$ .

In case  $\tau > T$  one has  $\dot{\psi}(t) = (0, 0)$  on the time interval  $[T, \tau]$ . Therefore, the cost incurred until  $\tau$  is at least  $\int_0^T [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}_\delta^2)^2(t) + h(\varphi(t))]dt + r\varrho(T)$ . By similar arguments one has that

$$\begin{aligned} & \int_0^T [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(T) \geq (h(\beta_0 + \delta) - \frac{r^2}{4c})T + r\varrho(T) \\ & \geq r(x - \beta_0 - \delta) - r(D - \beta_0 - \delta) + (\frac{r^2}{4c} + ry + h(\beta_0))T + (h(\beta_0 + \delta) - h(\beta_0))T \\ & = r(x - \beta_0 - \delta) + r\delta > r(x - \beta_0 - \delta). \end{aligned}$$

□

---

<sup>3</sup>Notice that  $\psi_\delta \in \mathcal{Q}[x, \bar{\alpha}]$ .

### 3.6.2 Optimal response to the barrier strategy

Recall that the  $\beta_0$ -barrier strategy is optimal under our standing assumption  $-y \geq r/(4c)$ . We now identify a control for the maximizer that is optimal when the minimizer uses the  $\beta_0$ -barrier strategy. More precisely, in case  $y^2 - h(\beta_0)/c > 0$  or  $x < \beta_0$  (resp.,  $y^2 - h(\beta_0)/c = 0$  and  $x = \beta_0$ ) we find an explicit expression for the maximizer's optimal (resp.,  $\delta$ -optimal) control and for the termination time.<sup>4</sup>

Throughout, it is assumed that  $x \in [0, \beta_0]$  and that the maximizer uses the  $\beta_0$ -barrier strategy,  $\bar{\alpha}_{\beta_0}$ . We use a constant  $\delta$  that in case  $y^2 - h(\beta_0)/c > 0$  or  $x < \beta_0$  is set to zero, and in case  $y^2 - h(\beta_0) = 0$  and  $x = \beta_0$  is assumed to satisfy  $0 < \delta \leq x$ .

In this subsection we assume that  $h$  is also Lipschitz continuous on  $[0, D]$ . Denote  $\tilde{\tau} = \delta + \int_0^{x+y\delta} 1/\sqrt{y^2 - h(\xi)/c} d\xi$  and let  $\omega \in C_0([0, \tilde{\tau}], \mathbb{R}) \cap C^1([\delta, \tilde{\tau}], \mathbb{R})$  be defined as 0 on  $[0, \delta]$  and let it be the unique solution of

$$\dot{\omega}(t) = -y - \sqrt{y^2 - h((x + yt + \omega(t)) \wedge (x + y\delta))/c}, \quad t \in [\delta, \tilde{\tau}]. \quad (86)$$

Indeed, existence and uniqueness will follow from the Picard-Lindelöf theorem once we show that the r.h.s. of (86) is Lipschitz with respect to  $\omega$ . This follows since  $h$  is assumed Lipschitz, and since the expression under the root sign is bounded away from zero. For the last statement, recall that we assumed that  $\delta = 0$  in case  $y^2 - h(\beta_0)/c > 0$  or  $x < \beta_0$  and  $0 < \delta < x$  in case  $y^2 - h(\beta_0)/c = 0$  and  $x = \beta_0$ . Hence,  $y^2 - h(x + y\delta)/c > 0$ . Using moreover that  $h$  is increasing we get that  $y^2 - h((x + yt + \omega(t)) \wedge (x + y\delta))/c$  is bounded away from zero.

Set

$$\psi(t) := \left( \frac{c}{c_1} \omega(t), \frac{-c}{c_2} \omega(t) \right), \quad t \in [0, \tilde{\tau}]. \quad (87)$$

**Theorem 3.5** *If  $y^2 - h(\beta_0)/c > 0$  or  $x < \beta_0$  then  $\bar{c}(x, \psi, \bar{\alpha}_{\beta_0}[\psi]) = U(x)$ . If  $y^2 - h(\beta_0)/c = 0$  and  $x = \beta_0$  then there is a positive constant  $C$  such that for every  $0 < \delta \leq x$  one has  $|\bar{c}(x, \psi, \bar{\alpha}_{\beta_0}[\psi]) - U(x)| \leq C\delta$ . Moreover, for every  $\delta \geq 0$  the termination time of the game is  $\tau = \tilde{\tau}$ .*

**Proof:** Notice that by the optimality of  $\bar{\alpha}_{\beta_0}$  it follows that for every  $\delta \geq 0$  one has  $U(x) = \sup_{\tilde{\psi} \in \mathcal{Q}[x, \bar{\alpha}_{\beta_0}]} \bar{c}(x, \tilde{\psi}, \bar{\alpha}_{\beta_0}[\tilde{\psi}]) \geq \bar{c}(x, \psi, \bar{\alpha}_{\beta_0}[\psi])$ . Therefore, in order to prove the theorem it is sufficient to prove that  $\bar{c}(x, \psi, \bar{\alpha}_{\beta_0}[\psi]) \geq U(x) - C\delta$  where  $\delta = 0$  in case  $y^2 - h(\beta_0)/c > 0$  or  $x < \beta_0$  and  $0 < \delta < x$  in case  $y^2 - h(\beta_0)/c = 0$  and  $x = \beta_0$ . To this end, it is sufficient to show that  $\psi \in \mathcal{Q}^2[x, \bar{\alpha}_{\beta_0}]$  and that the cost until the first hitting time,  $\tau$ , is bounded below by  $g$ . That is

$$\bar{c}(x, \psi, \bar{\alpha}[\psi]) = \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))] dt + r\varrho(\tau) \geq g(x), \quad (88)$$

where  $\varrho = \bar{\alpha}_{\beta_0}[\psi]$  and  $\varphi(t) = x + yt + \psi^1(t) - \psi^2(t) - \varrho(t)$ ,  $0 \leq t \leq \tau$ . We do that by showing that  $\psi$  satisfies the following three conditions.

(a) For every  $0 \leq t \leq \tau$  one has  $\varphi(t) \in [0, \beta_0]$ .

<sup>4</sup>One can show that  $y^2 - h(\beta_0)/c \geq 0$  and that equality holds if and only if  $-y = r/(2c)$  and  $-y \leq r/c + h(D)/r$ .

(b)  $\psi \in \mathcal{Q}^2[x, \bar{\alpha}_{\beta_0}]$ .

(c) For every  $\delta \leq t \leq \tau$  one has

$$-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)g'(\varphi(t)) = 0.$$

Before proving (a)–(c) let us show that they are sufficient for (88). Condition (a) is needed for  $g'(\varphi(\cdot))$  to be well-defined, where one uses the convention  $g'(\beta_0) = g'_L(\beta_0)$ . From condition (b) it follows that  $g(\varphi(\tau)) = 0$ . As a result, we can apply the chain rule to  $g$ , as in (37). Thus

$$\begin{aligned} & \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) + g(\varphi(\tau)) \\ &= g(x) + \int_0^\delta [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)g'(\varphi(t))]dt \\ & \quad + \int_\delta^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)g'(\varphi(t))]dt \\ & \quad + \int_0^\tau [r - g'(\varphi(t))]d\varrho(t). \end{aligned}$$

From condition (c) together with the fact that  $r \geq g'$  it follows that

$$\begin{aligned} & \int_0^\tau [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t))]dt + r\varrho(\tau) \\ & \geq g(x) + \int_0^\delta [-c_1(\dot{\psi}^1)^2(t) - c_2(\dot{\psi}^2)^2(t) + h(\varphi(t)) + (\dot{\psi}^1(t) - \dot{\psi}^2(t) + y)g'(\varphi(t))]dt. \end{aligned} \tag{89}$$

Recall that for every  $t \in [0, \delta]$  one has  $\dot{\omega}(t) = 0$  and therefore  $\psi(\cdot) = (0, 0)$  on that time interval. Moreover, the functions  $h$  and  $g$  are both bounded. Therefore, the integral on the r.h.s. of (89) is bounded below by  $-C\delta$ , where  $C$  is a positive constant independent of  $\delta$ . Hence, we proved that  $c(x, \psi, \bar{\alpha}_{\beta_0}[\psi]) \geq U(x) - C\delta$ .

We now show that  $\psi$  satisfies conditions (a)–(c). Hereafter, we assume that  $t \geq \delta$ . Since  $y + \dot{\omega}(t) \leq 0$  it follows that  $x + yt + \omega(t)$  is nonincreasing in  $t$  and therefore  $x + yt + \omega(t) \leq x + y\delta \leq \beta_0$ , since moreover,  $\omega = \psi^1 - \psi^2$  it follows that  $\varphi(t) = x + yt + \omega(t)$  and condition (a) holds. Since  $h$  is increasing we get that

$$y + \dot{\omega}(t) \leq -\sqrt{y^2 - h(x + y\delta)/c} < 0.$$

That is,  $\varphi(\cdot) = x + y \cdot + \omega(\cdot)$  is (strictly) decreasing with a slope that is bounded above by a negative number and therefore  $\varphi$  reaches zero at a finite time and (b) holds. Condition (c) follows by substitution of  $\psi$  and  $g$ .

We now turn to the termination time of the game. Recall that for every  $0 \leq t \leq \tau$  one has  $\varphi(t) \in [0, \beta_0]$ . Also, recall that  $\varphi(\cdot) = x + y \cdot + \omega(\cdot)$  and therefore,

$$\tau = \inf\{t \geq 0 : \varphi(t) = 0\} = \inf\{t \geq 0 : x + yt + \omega(t) = 0\}.$$

As mentioned earlier, the map  $\varphi : [0, \tau] \rightarrow [0, x]$  is strictly decreasing. Therefore,  $\varphi$  has an inverse function  $\varphi^{-1} : [0, x] \rightarrow [0, \tau]$ . Its derivative satisfies

$$(\varphi^{-1})'(\xi) = \frac{1}{\varphi'(\varphi^{-1}(\xi))},$$

and more explicitly

$$(\varphi^{-1})'(\xi) := \begin{cases} -1/\sqrt{y^2 - h(\xi)/c} & 0 \leq \xi \leq x + y\delta, \\ 1/y & x + y\delta \leq \xi \leq x. \end{cases} \quad (90)$$

Since  $\varphi^{-1}(0) = \tau$  and  $\varphi^{-1}(x) = 0$  it follows that

$$\tau = \varphi^{-1}(0) - \varphi^{-1}(x) = - \int_0^x (\varphi^{-1})'(\xi) d\xi = \delta + \int_0^{x+y\delta} 1/\sqrt{y^2 - h(\xi)/c} d\xi.$$

□

### 3.6.3 The case of convex $h$ : a saddle point property

We now present a saddle point property of the game in case when  $h$  is a convex function and twice continuously differentiable, namely that the maximizer can force a payoff of at least  $U(x)$  by choosing the control  $\psi$  defined in the previous subsection. The result is only concerned with the case  $x < \beta_0$ . Note that by the convention regarding the constant  $\delta$ , we take here  $\delta = 0$ .

**Proposition 3.3** *Fix  $x \in [0, \beta_0)$  and assume  $h$  is a convex function. Then for every  $\varrho \in \mathcal{E}$  one has  $\bar{c}(x, \psi, \varrho) \geq U(x)$ .*

**Proof:** We prove the proposition in two steps. First, we show that if the minimizer uses a control  $\varrho$  that has  $p$  rejections in total then it is optimal for it to reject the amount  $p$  at time zero. Then we argue that, for any amount  $p$  rejected at time zero, the cost is bounded below by  $U(x)$ .

**Step 1.** Let  $\varrho \in \mathcal{E}$ . Denote by  $p := \varrho(\tau)$ , where  $\tau = \tau(x, \psi, \varrho)$ . Define the modified path  $\varrho^p$  as  $\varrho^p(t) = p$  for all  $t \geq 0$ . We will show that

$$\bar{c}(x, \psi, \varrho) \geq \bar{c}(x, \psi, \varrho^p). \quad (91)$$

To this end, recall the definition of  $\psi$  from (87), which yields that  $x + y \cdot \psi^1(\cdot) - \psi^2(\cdot)$  is strictly decreasing and therefore  $\tau(x, \psi, \varrho) = \tau(x, \psi, \varrho^p) = T$ , where

$$T := \inf\{t \geq 0 : x + yt + \psi^1(t) - \psi^2(t) = p\}. \quad (92)$$

Therefore

$$\bar{c}(x, \psi, \varrho) = \int_0^T [-c_1(\psi^1)^2(t) - c_1(\psi^2)^2(t) + h(x + yt + \psi^1(t) - \psi^2(t) - \varrho(t))] dt + rp,$$

and  $\bar{c}(x, \psi, \varrho^p)$  is given by a similar expression with  $\varrho(t)$  replaced by  $p$ . Since  $h$  is increasing, (91) follows.

**Step 2.** Denote  $\varphi(t) = x + yt + \psi^1(t) - \psi^2(t)$ . Note that this is the dynamics under the control  $\psi$  and the  $\beta_0$ -barrier strategy, because  $\varrho = 0$  under that strategy for  $x < \beta_0$ . Consider any  $p \in [0, x]$  and recall the notation  $\varrho^p$  from the previous paragraph. Notice that  $\varphi[x, \psi, \varrho^p] = \varphi - p$ .



Denote  $F(p) = \bar{c}(x, \psi, \rho^p)$ . The result will follow once we show that  $F(p) \geq F(0) = g(x)$ . To this end, write  $T(p)$  for the quantity  $T$  of (92), to emphasize the  $p$ -dependence. Then

$$\begin{aligned} F(p) &= \int_0^{T(p)} [-c_1(\psi^1)^2(t) - c_1(\psi^2)^2(t) + h(\varphi(t) - p)]dt + rp \\ &= \int_0^{T(p)} [-c_1(\psi^1)^2(t) - c_1(\psi^2)^2(t) + h(\varphi(t))]dt \\ &\quad - \int_0^{T(p)} [h(\varphi(t)) - h(\varphi(t) - p)]dt + rp \\ &= g(x) - g(p) - \int_0^{T(p)} [h(\varphi(t)) - h(\varphi(t) - p)]dt + rp. \end{aligned}$$

The last equality follows since

$$g(x) = \int_0^{T(p)} [-c_1(\psi^1)^2(t) - c_1(\psi^2)^2(t) + h(x + yt + \psi^1(t) - \psi^2(t))]dt + g(p),$$

which in turn follows since both sides of the last equation represents the cost when both players play optimally where we use a DPP argument. Indeed, the integral is the cost under the dynamics hits  $p$  and  $g(p)$  is the cost when the initial state is  $p$ .

Note that  $F(0) = g(x)$ . Thus to prove that  $F(p) \geq g(x)$  for every  $p \in [0, x]$  it suffices to show that  $F' \geq 0$  on  $[0, x]$ . To this end, first notice that  $T = \varphi^{-1}$  and that  $T'(\cdot) = -1/\sqrt{y^2 - h(\cdot)}/c$ , see (90). By direct calculations it follows that

$$F'(p) = -g'(p) - T'(p)h(p) - \int_0^{T(p)} h'(\varphi(t) - p)dt + r.$$

By substituting  $q = \varphi(t)$  on the integral on the r.h.s. of the we get

$$\begin{aligned} \int_0^{T(p)} h'(\varphi(t) - p)dt &= \int_x^p h'(q - p)/(\varphi)'(\varphi^{-1}(q)) dq = \int_x^p h'(q - p)T'(q) dq \\ &= \int_p^x h'(q - p)/\sqrt{y^2 - h(q)}/c dq. \end{aligned}$$

Using moreover the convexity of  $h$  it follows that

$$F'(p) \geq -g'(p) - T'(p)h(p) - \int_p^x h'(q)/\sqrt{y^2 - h(q)}/c dq + r = -T'(p)h(p) - g'(x) + r \geq 0.$$

The equality above follows since  $h'(q)/\sqrt{y^2 - h(q)}/c = g''(q)$ . The last inequality follows since  $r \geq g'(x)$  and since  $T'(p)$  is negative.  $\square$

## A Appendix

### State space collapse for the multidimensional DG

The scaled processes  $(\tilde{A}^n, \tilde{S}^n)$  are assumed to satisfy a *moderate deviation principle*. To express this assumption, let  $\mathbb{J}_k, k = 1, 2$ , be functions on  $\mathcal{D}([0, T], \mathbb{R}^I)$  defined as follows. For  $\underline{\psi}^k =$

$(\psi_1^k, \dots, \psi_I^k) \in \mathcal{D}([0, T], \mathbb{R}^I)$ ,

$$\mathbb{J}_k(T, \underline{\psi}^k) := \begin{cases} \sum_{i=1}^I c_{i,k} \int_0^T (\psi_i^k)^2(s) ds & \text{if all } \psi_i^k \in \mathcal{AC}_0([0, T], \mathbb{R}), \\ \infty & \text{otherwise,} \end{cases} \quad (93)$$

where

$$c_{i,1} = \frac{1}{2\lambda_i \sigma_{i,IA}^2} \quad \text{and} \quad c_{i,2} = \frac{1}{2\mu_i \sigma_{i,ST}^2}, \quad i \in \mathcal{I}.$$

Let  $\mathbb{J}(T, \underline{\psi}) = \mathbb{J}_1(T, \underline{\psi}^1) + \mathbb{J}_2(T, \underline{\psi}^2)$  for  $\underline{\psi} = (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{D}([0, T], \mathbb{R}^{2I})$ .

We show that the multidimensional game can be reduced into a one dimensional game in the sense that they both share the same value. Towards this, we will need to define the components of the one dimensional cost function: the holding cost, rejection cost, and the rate function. For  $w \in \mathbb{R}_+$ , denote

$$h(w) := \inf \{ \underline{h} \cdot \underline{\xi} : \underline{\xi} \in \prod_{i=1}^I [0, D_i], \underline{\theta} \cdot \underline{\xi} = w \}. \quad (94)$$

By the convexity of the set  $\prod_{i=1}^I [0, D_i]$  it follows that  $h$  is convex. Moreover,  $h(w) \geq 0$  for  $w \geq 0$  and equality holds if and only if  $w = 0$ . Therefore,  $h$  is strictly increasing. Let

$$r := \min \{ \underline{r} \cdot \underline{q} : \underline{q} \in \mathbb{R}_+^I, \underline{\theta} \cdot \underline{q} = 1 \}. \quad (95)$$

As was shown in [4],

$$r = r_{i^*} \mu_{i^*} := \min \{ r_i \mu_i : i \in \mathcal{I} \}.$$

For every  $T \in \mathbb{R}_+$  and  $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$  set  $\mathbb{I}(T, \psi) = \mathbb{I}_1(T, \psi^1) + \mathbb{I}_2(T, \psi^2)$ , where

$$\mathbb{I}_k(T, \psi^k) := \inf \{ \mathbb{J}_k(T, \underline{\psi}^k) : \underline{\psi}^k \in \mathcal{P}^I, \underline{\theta} \cdot \underline{\psi}^k = \psi^k \}, \quad k = 1, 2. \quad (96)$$

One can verify that

$$\mathbb{I}_k(t, \psi^k) = \begin{cases} c_k \int_0^t (\psi^k)^2(s) ds & \text{if } \psi^k \in \mathcal{AC}_0([0, T], \mathbb{R}), \\ \infty & \text{otherwise,} \end{cases}$$

where

$$c_1 := \left( \sum_{i=1}^I \frac{2\rho_i \sigma_{i,IA}^2}{\mu_i} \right)^{-1}, \quad \text{and} \quad c_2 := \left( \sum_{i=1}^I \frac{2\sigma_{i,ST}^2}{\mu_i} \right)^{-1}.$$

Moreover, for every  $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$  define the multidimensional path  $\underline{\psi}_\psi = (\underline{\psi}_{\psi^1}^1, \underline{\psi}_{\psi^2}^2) \in \mathcal{P}^{2I}$  by

$$\underline{\psi}_{\psi^1}^1 := \left( \frac{2\rho_1 \sigma_{1,IA}^2}{\mu_1}, \dots, \frac{2\rho_n \sigma_{n,IA}^2}{\mu_n} \right) c_1 \psi^1$$

and

$$\underline{\psi}_{\psi^2}^2 := \left( \frac{2\sigma_{1,ST}^2}{\mu_1}, \dots, \frac{2\sigma_{n,ST}^2}{\mu_n} \right) c_2 \psi^2.$$

Simple calculation yields that

$$\underline{\theta} \cdot \underline{\psi}_{\psi^k}^k = \psi^k, \quad k = 1, 2,$$

and that

$$\mathbb{J}(T, \underline{\psi}_{\psi}) = \mathbb{I}(T, \psi). \quad (97)$$

Given  $\underline{x} \in \prod_{i=1}^I [0, D_i]$ , let  $x = \underline{\theta} \cdot \underline{x}$ . The one-dimensional value function is given by

$$V_{OD}(x) = \inf_{\alpha \in \mathcal{A}_x} \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} \left( \int_0^T h(\varphi(t)) dt + r\alpha_1[\psi](T) - \mathbb{J}(T, \psi) \right).$$

**Theorem A.1** For every  $\underline{x} \in \prod_{i=1}^I [0, D_i]$  one has  $V_{MD}(\underline{x}) = V_{OD}(x)$ .

**Proof of Theorem A.1** Fix  $\underline{x}$ . We prove the theorem by showing

$$V_{OD}(x) \leq V_{MD}(\underline{x}) \quad (98)$$

and

$$V_{MD}(\underline{x}) \geq V_{OD}(x). \quad (99)$$

We start by showing (98). Let  $\underline{\alpha} \in \mathcal{A}_{\underline{x}}$ . Define  $\alpha \in \mathcal{A}_x$  by

$$\alpha_k[(\psi^1, \psi^2)] := \underline{\theta} \cdot \underline{\alpha}_k[(\underline{\psi}_{\psi^1}^1, \underline{\psi}_{\psi^2}^2)], \quad (\psi^1, \psi^2) \in \mathcal{P}^2, \quad k = 1, 2.$$

The proof that  $\underline{\alpha}$  is admissible is straightforward and therefore omitted. Since  $(\psi^1, \psi^2) = (\underline{\theta} \cdot \underline{\psi}_{\psi^1}^1, \underline{\theta} \cdot \underline{\psi}_{\psi^2}^2)$  we get that

$$c(x, T, \psi, \alpha[\psi]) = c(x, T, (\underline{\theta} \cdot \underline{\psi}_{\psi^1}^1, \underline{\theta} \cdot \underline{\psi}_{\psi^2}^2), \underline{\alpha}[(\underline{\theta} \cdot \underline{\psi}_{\psi^1}^1, \underline{\theta} \cdot \underline{\psi}_{\psi^2}^2)]) \quad (100)$$

Fix  $\psi \in \mathcal{P}^2$  and let

$$\underline{\varphi}(t) = \underline{x} + \underline{y}t + \underline{\psi}_{\psi^1}^1(t) - \underline{\psi}_{\psi^2}^2(t) + \underline{\alpha}_1[\underline{\psi}_{\psi}](t) - \underline{\alpha}_2[\underline{\psi}_{\psi}](t), \quad t \geq 0$$

be the dynamics associated with  $\underline{x}, \underline{\psi}_{\psi}$ , and  $\underline{\alpha}[\underline{\psi}_{\psi}]$ . By the definitions of  $h$  and  $r$  (see (94), (95)) and by (97) it follows that

$$\underline{h} \cdot \underline{\varphi}(t) \geq h(\underline{\theta} \cdot \underline{\varphi}(t)), \quad t \geq 0,$$

$$\underline{r} \cdot \underline{\alpha}_2[\underline{\psi}_{\psi}](t) \geq r\underline{\theta} \cdot \alpha_2[\psi_{\psi}](t), \quad t \geq 0,$$

and

$$\mathbb{I}(T, \psi) = \mathbb{J}(T, \underline{\psi}_{\psi}), \quad T \geq 0.$$

Therefore,

$$\begin{aligned}
& c(x, T, (\underline{\theta} \cdot \underline{\psi}_{\psi^1}^1, \underline{\theta} \cdot \underline{\psi}_{\psi^2}^2), \alpha[(\underline{\theta} \cdot \underline{\psi}_{\psi^1}^1, \underline{\theta} \cdot \underline{\psi}_{\psi^2}^2)]) \\
&= \int_0^T h(\underline{\theta} \cdot \underline{\varphi}(t)) dt + r \underline{\theta} \cdot \underline{\alpha}_2[\underline{\psi}_{\psi}](T) - \mathbb{I}(T, \underline{\psi}_{\psi}) \\
&\leq \int_0^T \underline{h} \cdot \underline{\varphi}(t) dt + r \cdot \underline{\alpha}_2[\underline{\psi}_{\psi}](T) - \mathbb{J}(T, \underline{\psi}) \\
&= \underline{c}(x, T, \underline{\psi}_{\psi}, \underline{\alpha}[\underline{\psi}_{\psi}]).
\end{aligned}$$

Together with (100) we get that

$$c(x, T, (\psi), \alpha[\psi]) \leq \underline{c}(x, T, \underline{\psi}_{\psi}, \underline{\alpha}[\underline{\psi}_{\psi}])$$

and therefore,

$$\sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, (\psi), \alpha[\psi]) \leq \sup_{\underline{\psi} \in \mathcal{P}^{2I}, T \in \mathbb{R}_+} \underline{c}(x, T, \underline{\psi}, \underline{\alpha}[\underline{\psi}]).$$

This shows (98).

We next show (99). Let  $\alpha \in \mathcal{A}_x$ . Toward constructing  $\underline{\alpha} \in \mathcal{A}_x$  let  $\gamma : [0, D] \rightarrow \prod_{i=1}^I [0, D_i]$  be Borel measurable, satisfying

$$\gamma(w) \in \arg \min_{\underline{\xi}} \{ \underline{h} \cdot \underline{\xi} : \underline{\xi} \prod_{i=1}^I [0, D_i], \underline{\theta} \cdot \underline{\xi} = w \}, \quad w \in [0, D].$$

The existence follows from a measurable selection argument (such as Corollary 10.3 in the appendix of [8]). Set

$$\varphi(t) = \underline{\theta} \cdot \underline{x} + \underline{\theta} \cdot \underline{y} t + \underline{\theta} \cdot \underline{\psi}^1(t) - \underline{\theta} \cdot \underline{\psi}^2(t) + \underline{\alpha}_1[(\underline{\psi}^1, \underline{\psi}^2)](t) - \underline{\alpha}_2[(\underline{\psi}^1, \underline{\psi}^2)](t), \quad t \geq 0, \quad (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{P}^{2I}$$

and

$$\underline{\varphi}(t) = \gamma(\varphi(t)), \quad t \geq 0.$$

Define

$$\underline{\alpha}_2[(\underline{\psi}^1, \underline{\psi}^2)](t) := \alpha_2[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)](t) \mu_{i^*} e_{i^*}, \quad t \geq 0, \quad (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{P}^{2I},$$

and

$$\underline{\alpha}_1[(\underline{\psi}^1, \underline{\psi}^2)](t) := \varphi(t) - \underline{x} - \underline{y} t - \underline{\psi}^1(t) + \underline{\psi}^2(t) + \underline{\alpha}_2[(\underline{\psi}^1, \underline{\psi}^2)](t), \quad t \geq 0, \quad (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{P}^{2I}.$$

The strategy  $\underline{\alpha}$  is admissible. Indeed, causality follows since  $\alpha$  satisfies this property. Monotonicity of each coordinate of  $\underline{\alpha}_2[(\underline{\psi}^1, \underline{\psi}^2)](\cdot)$  follows by the monotonicity of  $\alpha_2[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)](\cdot)$ , and monotonicity of  $\underline{\theta} \cdot \underline{\alpha}_1[(\underline{\psi}^1, \underline{\psi}^2)](\cdot)$  follows since the latter equals to  $\alpha_1[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)](\cdot)$  which is monotone by admissibility of  $\alpha_1$ . Finally, for every  $w \in [0, D]$  one has  $\gamma(w) \in \prod_{i=1}^I [0, D_i]$  and therefore  $\underline{\varphi}$  takes values in  $\prod_{i=1}^I [0, D_i]$ .

Next, note that

$$\underline{h} \cdot \underline{\varphi}(t) = h(\varphi(t)), \quad t \geq 0,$$

$$r \cdot \underline{\alpha}_2[(\underline{\psi}^1, \underline{\psi}^2)](t) = r_{i^*} \mu_{i^*} \alpha_2[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)](t) = r \alpha_2[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)](t), \quad t \geq 0, \quad (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{P}^{2I}$$

and

$$\mathbb{J}(T, (\underline{\psi}^1, \underline{\psi}^2)) \geq \mathbb{I}(T, (\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)), \quad T \geq 0, \quad (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{P}^{2I}.$$

The last inequality follows by the definition of  $\mathbb{I}$ , see (96). Therefore for every  $\underline{\psi} = (\underline{\psi}^1, \underline{\psi}^2) \in \mathcal{P}^{2I}$  one has

$$\begin{aligned} & \underline{c}(x, T, \underline{\psi}, \underline{\alpha}[\underline{\psi}]) \\ &= \int_0^T \underline{h} \cdot \underline{\varphi}(t) dt + r \cdot \underline{\alpha}_2[(\underline{\psi}^1, \underline{\psi}^2)](T) - \mathbb{J}(T, (\underline{\psi}^1, \underline{\psi}^2)) \\ &\leq \int_0^T h(\varphi(t)) dt + r \alpha_2[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)](T) - \mathbb{I}(T, (\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)) \\ &= c(x, T, (\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2), \alpha[(\underline{\theta} \cdot \underline{\psi}^1, \underline{\theta} \cdot \underline{\psi}^2)]) \end{aligned}$$

Hence,

$$\sup_{\underline{\psi} \in \mathcal{P}^{2I}, T \in \mathbb{R}_+} \underline{c}(x, T, \underline{\psi}, \underline{\alpha}[\underline{\psi}]) \leq \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c(x, T, \psi, \alpha[\psi])$$

and (99) follows.

## References

- [1] R. Atar and A. Biswas. Control of the multiclass G/G/1 queue in the moderate deviation regime. *AAP*.
- [2] R. Atar and A. Cohen. *Work in progress*, 2014.
- [3] R. Atar, P. Dupuis, and A. Shwartz. An escape-time criterion for queueing networks: asymptotic risk-sensitive control via differential games. *Math. Oper. Res.*, 28(4):801–835, 2003.
- [4] R. Atar and M. Shifrin. An asymptotic optimality result for the multiclass queue with finite buffers in heavy traffic. *Preprint*.
- [5] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Springer, 2008.
- [6] K.-J. Chung and M. J. Sobel. Discounted MDP's: Distribution functions and exponential utility maximization. *SIAM Journal on Control and Optimization*, 25(1):49–62, 1987.
- [7] R. J. Elliott and N. J. Kalton. *The existence of value in differential games*. American Mathematical Society, Providence, R.I., 1972. iv+67 pp. *Memoirs of the American Mathematical Society*, No. 126.
- [8] S. N. Ethier and T. G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. ISBN 0-471-08186-8. x+534 pp. Characterization and convergence.
- [9] W. H. Fleming and W. M. McEneaney. Risk-sensitive control on an infinite time horizon. *SIAM J. Control Optim.*, 33(6):1881–1915, 1995.
- [10] W. H. Fleming and P. E. Souganidis. PDE-viscosity solution approach to some problems of large deviations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 13(2):171–192, 1986.
- [11] J. M. Harrison and M. I. Taksar. Instantaneous control of Brownian motion. *Math. Oper. Res.*, 8(3): 439–453, 1983.
- [12] L. Kruk, J. Lehoczky, K. Ramanan, and S. Shreve. An explicit formula for the Skorokhod map on  $[0, a]$ . *Ann. Probab.*, 35(5):1740–1768, 2007.

- [13] S. I. Marcus, E. Fernández-Gaucherand, D. Hernández-Hernández, S. Coraluppi, and P. Fard. Risk sensitive Markov decision processes. In *Systems and control in the twenty-first century (St. Louis, MO, 1996)*, volume 22 of *Progr. Systems Control Theory*, pages 263–279. Birkhäuser Boston, Boston, MA, 1997.
- [14] A. A. Puhalskii. Moderate deviations for queues in critical loading. *Queueing Systems Theory Appl.*, 31 (3-4):359–392, 1999.
- [15] A. A. Puhalskii and W. Whitt. Functional large deviation principles for first-passage-time processes. *Ann. Appl. Probab.*, 7(2):362–381, 1997.