# Brownian control problems for a multiclass M/M/1 queueing problem with model uncertainty\*

Asaf Cohen<sup>†</sup>

August 27, 2018

#### Abstract

We consider a multidimentional Brownian control problem (BCP) with model uncertainty that formally emerges from a multiclass M/M/1 queueing control problem under heavy-traffic with model uncertainty. The BCP is formulated as a multidimensional stochastic differential game with two players: a minimizer that has an equivalent role to the decision maker in the queueing control problem and a maximizer whose role is to set up the uncertainty of the model. The dynamics are driven by a Brownian motion. We show that a state-space collapse propery holds. That is, the multidimensional BCP can be reduced to a one-dimensional BCP with model uncertainty that also takes the form of a two-player stochastic differential game. Then, the value function of both games is characterized as the unique solution to a free-boundary problem from which we extract equilibria for both games. Finally, we analyze the dependence of the value function and the equilibria on the ambiguity parameters.

**AMS Classification:** Primary: 93E20, 60K25, 91A15, 60J60; secondary: 49J15, 35R35. **Keywords:** Brownian control problem; ambiguity aversion; model uncertainty; multiclass M/M/1; heavy-traffic; the Harrison–Taksar free-boundary problem.

# 1 Introduction

Typically, heavy-traffic queueing control problems (QCPs) in the diffusion scale are treated by defining a limiting control problem associated with Brownian motion, called *Brownian control* problems (BCPs), first introduced by [18]; for further reading on BCPs see e.g., [8, 11, 12] and the references therein. In this paper we study two BCPs with model uncertainty that formally emerge from a multiclass M/M/1 QCP with finitely many buffers with finite capacity under heavy-traffic with model uncertainty. The asymptotic relationship between the QCP and the BCPs is the subject of a separate paper, see [14]. Both the QCP and the BCP were studied in the classical case under the framework of G/G/1 without ambiguity about the model in [6] and in [28], where the latter considers time constraints instead of the finite buffers constraints. These problems are referred to as *risk-neutral* problems.

<sup>\*</sup>To appear in Mathematics of Operations Research

<sup>&</sup>lt;sup>†</sup>Department of Statistics, University of Haifa, Haifa, 31905, Israel, shloshim@gmail.com web: https://sites.google.com/site/asafcohentau/

We are formulating a decision maker (DM) that has a reference model in mind, which, up to some degree, describes the situation she is facing. Since the DM is uncertain about the true model (either because the parameters are not fully known, can change over time, etc.), she takes into account other models and penalizes them based on their deviation from a specific reference model. The penalization depends then on how averse the DM is to ambiguity. Such ambiguity models are sometimes referred to as *model uncertainty* or *Knightian uncertainty*, see e.g., [27, 17, 16, 7] and in the context of queueing systems see [24, 10, 26].

The first BCP with model uncertainty introduced is a multidimensional stochastic differential game (MSDG) with two players: a DM and the nature, which according to their goals are referred to as the minimizer and the maximizer, respectively. Borrowing terminology and the roles of the processes from the QCP to the MSDG (here and in the sequel), the state-space of the game is a product set of intervals of the form  $[0, \hat{b}_i]$ , where  $\hat{b}_i$  is the capacity of buffer *i*. The minimizer in this game controls the server's effort allocation among the buffers and the admission/rejection to each buffer. The maximizer chooses the underlying probability measure; this is shown to be equivalent to stochastically perturbing the drift of the Brownian motion (possibly, differently for each coordinate). The game's cost consists of holding and rejection penalties and a variant of the Kullback–Leibler divergence with respect to (w.r.t.) the relevant measures in this setup. The latter component stands for a penalty for the maximizer for changing the drift.

We show that a *state-space collapse* property holds. That is, we provide a one-dimensional BCP with uncertainty that also takes the form of a stochastic game, called the *reduced stochastic* differential game (RSDG), whose state emerges from the workload process in the QCP. The roles of the two players remain the same as in the MSDG and the dynamics and the cost functions have similar components. We show that the games are equivalent in the sense that given any strategy of the minimizer in either one of the games, we construct a strategy for the minimizer in the other game that performs at least as well, and therefore, also the value functions are the same (Proposition 2.1). For further reading about workload reduction, see [18, 22, 19, 20, 23]. The advantage of such a reduction is that the dynamics live in a lower dimension and have only two components of singular controls that represent idleness and workload rejection. Therefore, most of the analysis is performed in the  $RSDG^1$  setup. We characterize the value function of the RSDG as the unique classical solution of a Hamilton-Jacobi-Bellman (HJB) equation (Theorem 3.1). Doing so, we extend the relationship between the reduced BCP and a relevant HJB equation studied in [21, Equation (1.2)] and in [6, 1, 2]Equation (41) to a similar relationship in a stochastic game setup with a different HJB; unlike in [6], due to the existence of a maximal player in our model, the HJB is not linear. Therefore, one cannot use the results given in these papers and rather needs to establish the relationship between the two. As a first step, we reduce the problem of solving the HJB equation to a freeboundary problem. (To the best of my knowledge, this is the first model uncertainty control problem that leads to a free-boundary problem with Neumann boundary conditions). Then we use the *shooting method* to solve the latter problem. In short, this is a method for solving boundary value problems using initial value problems. We take it one step forward in the

<sup>&</sup>lt;sup>1</sup>Throughout the paper we alternatively refer to the BCPs as the 'stochastic differential games', or as 'the games', or explicitly by MSDG and RSDG.

free-boundary setup; see [33, Section 7.3] for further reading about the method. Moreover, we supply equilibria in both games and refer to the equilibrium strategies of the minimizer (in both games) as the *optimal strategy* of the minimizer (Theorem 4.1). Starting with the RSDG, we show that the optimal strategy for the minimizer in this game is a reflecting strategy. Namely, the minimizer should use minimal idleness and minimal amount of workload rejections in order to keep the workload in a fixed interval. The equilibrium strategy of the maximizer in this game is also provided. From the RSDG equilibrium we construct an equilibrium in the MSDG. According to the minimizer's optimal policy in the MSDG, the queue length processes evolve along a certain curve in the state-space. The minimizers' optimal strategies in our games are shown to have the same structures as the optimal strategies in the risk-neutral BCPs in [6]. Such a result is not so obvious for reasons having to do with the non-stationarity structure of the problem caused by the existence of the maximizer player. The difference between the policy of the minimizer in the multidimensional setup here as compared to [6] is only the cut-off level of the reflecting strategy in the RSDG, which affects the point of reflection on the multidimensional curve.

Aside from studying the games, we also analyze the dependence of the games on the ambiguity parameters. We show continuity of the value function and the optimal reflecting strategies w.r.t. the ambiguity parameters and that as the ambiguity vanishes, the problem converges to the risk-neutral problem studied in [6] (Theorems 5.1 and 5.2).

We now discuss about the position of the current work between the risk-neutral BCPs and the deterministic differential games (DDGs) studied in [2] and provide future outlook. Recall that the risk-neutral BCPs from [6] governs the limiting behavior of a diffusion scaled multiclass G/G/1 QCP with linear utility function. On the other side, the DDGs from [2] approximate the same type of QCP with the following differences: the moderate-deviation replaces the diffusion scaling and the utility function is exponential, scaled with the moderate-deviation parameters. While in the risk-neutral case the probability of buffer overflow is approximated by the probability that a Brownian motion with drift hits a positive level, which is of order  $\mathcal{O}(1)$ , in the moderate-deviation scaling this probability vanishes with the scaling parameter. As a compensation, the criteria considered is a risk-sensitive one, which gives an overwhelming scaled exponential cost for such event. This means that the DM has a large ambiguity about the model and she gives high weight even for events that are 'very' unlikely to happen. The intuition behind the relationship between the QCP in the moderate-deviation heavy-traffic regime and the limiting DDG goes back to the classical risk-sensitive control problem with small noise diffusion. As argued in [15, Ch. VI.2], consider a cost function  $\mathcal{J}$  and a positive constant  $\rho$ , then the risk-sensitive cost can be expressed as,

$$\alpha^{-1}\log\left(E^{\mathbb{P}}[e^{\alpha\mathcal{J}}]\right) = \sup_{\mathbb{Q}} E^{\mathbb{Q}}\left[\alpha\left(\mathcal{J} - \log\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right)\right],\tag{1.1}$$

where the supremum is taken over an appropriate set of measures. When the noise coefficient is  $\alpha^{-1/2}$  and  $\alpha \to \infty$ , the limiting problem is a DDG. What we consider is the criteria given in (1.1), without sending  $\alpha \to \infty$ . In the QCP, it means that we still consider the classical diffusion scaling, yet our criteria gives space for ambiguity about the true underlying probability measure. Since the higher  $\alpha$  is, the more weight the DM gives to unlikely events, we say that

the DM has more ambiguity in the moderate-deviation risk sensitive QCP then in the one related to our work, and thus we position our model between the two models described above. This work does not aim to establish a deeper connection between the model uncertainty's BCPs and the DDGs. At this point it worth mentioning that the relation between the Kullback–Leibler-constrained formulation and the exponential functional (i.e., (1.1)) is well known in other contexts, see e.g., [13] for an example in the context of uncertainty quantification.

In [2] it was shown that the optimal strategies of the minimizers in the DDGs have the same structure as the ones in the BCPs from [6]. Together with the observation stated earlier, we get that all the three models ([6], [2], and the present one) share the same optimal policy, where the different lies in the cut-off point for rejections. Moreover, in a recent line of research of queueing systems under the moderate-deviation heavy-traffic regime with risk-sensitive performance criteria it was shown that classical results from the theory of risk-neutral QCPs and BCPs such as state-space collapse and generalyzed  $c\mu$  rule hold in these models as well, see e.g., [1, 9, 2, 5, 3, 4]. The current paper together with [14] represents a line of research of QCPs and their associated BCPs under model uncertainty.

In this work we assume that the buffers are finite. This property arises naturally in private cloud computing, which are limited in data and capacity. For a more detailed application, see [31]. Intuitively, the problem with unlimited capacities and without rejections seems to be simpler. We expect that similar tools can be used to show that the optimal strategy would be to assign *fixed* lowest priority to the class with the smallest  $\hat{h}_i \mu_i$  value. Consequently, the limiting queueing sizes would vanish with the scaling parameter for the rest of the buffers. This is left for future work.

In summary, our main contributions are as follows. We

- provide and solve for the first time a BCP with model uncertainty that emerges from a multiclass M/M/1 QCP with model uncertainty;
- show that a state-space collapse property holds for this game (Corollary 2.1);
- show that the reduced game solves uniquely a relevant HJB equation, which is a nonlinear free-boundary problem and that there is an optimal reflecting strategy (Theorem 3.1);
- provide equilibria for the two games considered (Theorem 4.1);
- analyze the dependence of the value function and the equilibria on the ambiguity parameters (Theorems 5.1 and 5.2).

The paper is organized as follows. In Section 2 we motivate and present the stochastic differential games and study the relationship between the two. Next, In Section 3 we study the RSDG. We provide the HJB equation and prove that the value function of the game is the unique smooth solution of the HJB. Moreover, we show that the minimizer has an optimal reflecting strategy. In Section 4 we discuss the uniqueness of the optimal reflecting strategy and find equilibria in both games. Finally, in Section 5 we study the dependency on the ambiguity parameters.

### 1.1 Notation

We use the following notation. For  $a, b \in \mathbb{R}$ ,  $a \wedge b = \min\{a, b\}$  and  $a \vee b = \max\{a, b\}$ . For a positive integer k and  $c, d \in \mathbb{R}^k$ ,  $c \cdot d$  denotes the usual scalar product and  $||c|| = (c \cdot c)^{1/2}$ . We denote  $[0, \infty)$  by  $\mathbb{R}_+$ . For subintervals  $I_1, I_2 \subseteq \mathbb{R}$  and  $m \in \{1, 2\}$  we denote by  $\mathcal{C}(I_1, I_2)$ ,  $\mathcal{C}^m(I_1, I_2)$ , and  $\mathcal{D}(I_1, I_2)$  the space of continuous functions [resp., functions with continuous derivatives of order m, functions that are right-continuous with finite left limits (RCLL)] mapping  $I_1 \to I_2$ . The space  $\mathcal{D}(I_1, I_2)$  is endowed with the usual Skorohod topology.

# 2 The BCPs

We start this section with a motivative QCP. The two BCPs are presented in Section 2.2: in Section 2.2.1 we formally derive the MSDG and in Section 2.2.2 its reduction. Then in Section 2.3 we show that the two games share the same value and that given any strategy (for the minimizer) in either one of the games, one can construct a strategy in the second game that performs at least as well.

## 2.1 Motivative QCP

Consider a model that consists of I customer classes and a single server. Each class has its own finite buffer and upon arrival, customers are queued in the corresponding buffer or rejected. Within each class, customers are served at the order of their arrivals. Processor sharing is allowed and the server may serve up to I customers at a time, where two customers from the same class cannot be served simultaneously. The system under consideration is in heavy-traffic. For this, we consider a sequence of systems, indexed by the scaling parameter  $n \in \mathbb{N}$ . For every n we consider a reference probability space that supports independent Poisson processes  $A_i^n$  and  $S_i^n$ ,  $i \in [I] := \{1, \ldots, I\}$  with rates  $\lambda_i^n$  and  $\mu_i^n$ , representively. The value  $A_i^n(t)$  stands for the number of customers of class i that arrived to the system until time  $t \in \mathbb{R}_+$ , and  $S_i^n(t)$  is the number of service completions of class i customers had the server dedicated all of its effort to class i during the time interval [0, t].

Denote by  $T_i^n(t)$  the units of time that the server devoted to class i until time t. For every  $t \in \mathbb{R}_+$  and  $i \in [I]$ ,  $S_i^n(T_i^n(t))$  is the number of service completions of class i customers until time t. This is a Cox process with the infinitesimal intensity  $\mu_i^n dT_i^n(t)$ . Rejections of customers are allowed upon arrival and a rejected customer will never return to the system. The number of customers from class i that were rejected by time t is denoted by  $R_i^n(t)$ . For every  $i \in [I]$ , the balance equation is given by,

$$X_i^n(t) = X_i^n(0) + A_i^n(t) - S_i^n(T_i^n(t)) - R_i^n(t), \qquad t \in \mathbb{R}_+,$$
(2.1)

where  $X_i^n(t)$  stands for the number of class *i* customers in the system at time *t*. We use the notation  $L^n = (L_i^n)_{i=1}^I$  for  $\{A, S, X, R, T\}$ .

We assume that

$$\lambda_i^n := \lambda_i n + \hat{\lambda}_i n^{1/2} + o(n^{1/2}), \qquad \mu_i^n := \mu_i n + \hat{\mu}_i n^{1/2} + o(n^{1/2}), \tag{2.2}$$

where  $\lambda_i, \mu_i \in (0, \infty)$  and  $\hat{\lambda}_i, \hat{\mu}_i \in \mathbb{R}$  are fixed. Moreover, the system is assumed to be *critically* loaded, that is,  $\sum_{i=1}^{I} \rho_i = 1$ , where  $\rho_i := \lambda_i / \mu_i, i \in [I]$ .

The scaled version of (2.1) is given by,

$$\hat{X}_{i}^{n}(t) = \hat{X}_{i}^{n}(0) + \hat{m}_{i}^{n}t + \hat{A}_{i}^{n}(t) - \hat{S}_{i}^{n}(T_{i}^{n}(t)) + \hat{Y}_{i}^{n}(t) - \hat{R}_{i}^{n}(t), \qquad t \in \mathbb{R}_{+},$$
(2.3)

where

$$\hat{X}_{i}^{n}(t) := n^{-1/2} X_{i}^{n}(t), \quad \hat{A}_{i}^{n}(t) := n^{-1/2} (A_{i}^{n}(t) - \lambda_{i}^{n}t), \quad \hat{S}_{i}^{n}(t) := n^{-1/2} (S_{i}^{n}(t) - \mu_{i}^{n}t), \quad (2.4)$$

$$\hat{Y}_{i}^{n}(t) := \mu_{i}^{n} n^{-1/2} (\rho_{i}t - T_{i}^{n}(t)), \qquad \hat{R}^{n}(t) := n^{-1/2} R^{n}(t),$$

and

$$\hat{m}_i^n := n^{-1/2} (\lambda_i^n - \rho_i \mu_i^n).$$

As previously, we use the notation  $\hat{L}^n = (\hat{L}^n_i)_{i=1}^I$  for  $L \in \{A, S, X, R, Y, T, m\}$ . The capacity of buffer *i* is given by  $\hat{b}^n_i := \hat{b}_i n^{1/2}$  for some constant  $\hat{b}_i \in (0, \infty), i \in [I]$ . We assume that  $\hat{X}^n(0) \in \mathcal{X} := \prod_{i=1}^I [0, \hat{b}_i]$ , and the rejection mechanism assures that

$$\ddot{X}_i^n(t) \in \mathcal{X}, \qquad t \in \mathbb{R}_+, \ \mathbb{P}^n$$
-a.s. (2.5)

We now present the optimization criteria. The intuition behind it is as follows. The DM, also referred to as the *minimizer*, chooses a control based on the past observations. He minimizes a cost that takes into account a possible deviation from the reference model. For this, we consider an adverse 'player', also referred to as the maximizer, who has access to the policy chosen by the minimizer and to the history. This player is penalized for deviating from the reference model.

In details, the vectors  $\hat{h}, \hat{r} \in (0, \infty)^I$  stand for the holding and rejection costs, respectively. The DM is uncertain about the underlying reference probability measure, or in other words, she suspects with some level of uncertainty that the rates of the processes  $\{A_i^n\}_{i=1}^I$  and  $\{S_i^n\}_{i=1}^I$ are not exactly  $\{\lambda_i^n\}_{i=1}^I$  and  $\{\mu_i^n\}_{i=1}^I$  and may be unspecified or may even change over the time. Therefore, she considers a set of candidate measures and penalizes their deviation from the reference probability measure. The penalization is done by using discounted variants of the Kullback–Leibler divergence, given by

$$L_{1}^{\varrho}(\hat{\mathbb{Q}}_{1,i}^{n} \| \mathbb{P}_{1,i}^{n}) := \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^{n}} \left[ \int_{0}^{\infty} \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_{1,i}^{n}(t)}{d\mathbb{P}_{1,i}^{n}(t)} dt \right], \qquad n \in \mathbb{N}, \ i \in [I],$$

$$L_{2}^{\varrho}(\hat{\mathbb{Q}}_{2,i}^{n} \| \mathbb{P}_{2,i}^{n}) := \mathbb{E}^{\hat{\mathbb{Q}}_{2,i}^{n}} \left[ \int_{0}^{\infty} \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_{2,i}^{n}(t)}{d\mathbb{P}_{2,i}^{n}(t)} dT_{i}^{n}(t) \right], \qquad n \in \mathbb{N}, \ i \in [I],$$
(2.6)

where  $\mathbb{P}_{1,i}^n$  and  $\mathbb{P}_{2,i}^n$  are the reference measures under which  $A_i^n$  and  $S_i^n$  are Poisson processes with rates  $\lambda_i^n$  and  $\mu_i^n$ , respectively. To establish the level of ambiguity, for every  $i \in [I]$ , we consider the (finite and) positive parameters  $\kappa_{1,i}$  and  $\kappa_{2,i}$  that quantify the amount of ambiguity that the DM has regarding the rates  $\lambda_i^n$  and  $\mu_i^n$ , or in other words, the measures  $\mathbb{P}_{1,i}^n$  and  $\mathbb{P}_{2,i}^n$ , respectively. Set  $\kappa := (\kappa_{1,i}, \kappa_{2,i})_{i=1}^I$ . The DM is facing the following optimization problem:

$$V^{n}(X^{n}(0);\kappa) = \inf_{(T^{n},R^{n})} \sup_{\hat{\mathbb{Q}}^{n}\in\hat{\mathcal{Q}}^{n}(\hat{X}^{n}(0))} \hat{J}^{n}(\hat{X}^{n}(0),U^{n},R^{n},\hat{\mathbb{Q}}^{n};\kappa),$$

where

$$\hat{J}^{n}(\hat{X}^{n}(0), U^{n}, R^{n}, \hat{\mathbb{Q}}^{n}; \kappa) =$$

$$\mathbb{E}^{\hat{\mathbb{Q}}^{n}} \left[ \int_{0}^{\infty} e^{-\varrho t} \left( \hat{h} \cdot \hat{X}^{n}(t) dt + \hat{r} \cdot d\hat{R}^{n}(t) \right) \right] - \sum_{i=1}^{I} \frac{1}{\kappa_{1,i}} L_{1}^{\varrho}(\hat{\mathbb{Q}}_{1,i}^{n} \| \mathbb{P}_{1,i}^{n}) - \sum_{i=1}^{I} \frac{1}{\kappa_{2,i}} L_{2}^{\varrho}(\hat{\mathbb{Q}}_{2,i}^{n} \| \mathbb{P}_{2,i}^{n}),$$
(2.7)

 $\hat{\mathbb{Q}}^n = \prod_{i=1}^{I} (\hat{\mathbb{Q}}_{1,i}^n \times \hat{\mathbb{Q}}_{2,i}^n)$ , and the set of candidate measures  $\hat{\mathcal{Q}}^n(\hat{X}^n(0))$  is described at the end of this paragraph. When  $\kappa_{j,i}$  is 'small' (resp., 'large') we say that there is a weak (resp., strong) ambiguity about the rates of the processes  $A_i^n$  and  $S_i^n(T_i^n) := S_i^n(T_i^n(\cdot))$ . The idea is that for small  $\kappa_{j,i}$ 's there is a big punishment per unit of deviation from the reference measure and therefore, the measures  $\hat{\mathbb{Q}}_{j,i}^n$  and  $\mathbb{P}_{j,i}^n$  should be close to each other and as a consequence also the relevant expectations. We now turn to define the set of candidate measures. A probability measure  $\hat{\mathbb{Q}}^n$  belongs to  $\hat{\mathcal{Q}}^n(\hat{X}^n(0))$  if for every  $i \in [I]$  and  $t \in \mathbb{R}_+$  it satisfies

$$\frac{d\hat{\mathbb{Q}}_{1,i}^n(t)}{d\mathbb{P}_{1,i}^n(t)} = \exp\Big\{\int_0^t \log\left(\frac{\psi_{1,i}^n(s)}{\lambda_i^n}\right) dA_i^n(s) - \int_0^t (\psi_{1,i}^n(s) - \lambda_i^n) ds\Big\},\tag{2.8}$$

$$\frac{d\hat{\mathbb{Q}}_{2,i}^{n}(t)}{d\mathbb{P}_{2,i}^{n}(t)} = \exp\Big\{\int_{0}^{t}\log\left(\frac{\psi_{2,i}^{n}(s)}{\mu_{i}^{n}}\right)dS_{i}^{n}(T_{i}^{n}(s)) - \int_{0}^{t}(\psi_{2,i}^{n}(s) - \mu_{i}^{n})dT_{i}^{n}(s)\Big\},\tag{2.9}$$

for measurable and positive processes  $\psi_{j,i}^n$  that are predictable w.r.t. the filtration generated by the arrival and service completions processes, satisfying  $\int_0^t \psi_{j,i}^n(s) ds < \infty \mathbb{P}_{j,i}^n$ -a.s., for every  $t \in \mathbb{R}_+$ . These conditions guarantee that the right-hand sides in (2.8) are  $\mathbb{P}_{j,i}^n$ -martingales, and that under the measure  $\hat{\mathbb{Q}}_{1,i}^n$  (resp.,  $\hat{\mathbb{Q}}_{2,i}^n$ ), the processes  $A_i^n$  (resp.,  $S_i^n(T_i^n)$ ) is a counting process with infinitesimal intensity  $\psi_{1,i}^n(t) dt$  (resp.,  $\psi_{2,i}^n dT_i^n(t)$ ). Notice that we do not force the critically loaded condition under the measures  $\hat{\mathbb{Q}}_{j,i}^n$ . As we argue in [14, Section 4] such changes of measures are 'too costly' and will be avoided by the maximizer who would choose  $\psi_{1,i}^n(t) = \lambda_i^n + \mathcal{O}(n^{1/2})$  and  $\psi_{2,i}^n(t) = \mu_i^n + \mathcal{O}(n^{1/2})$ .

We now provide an approximation to the change of measure penalty that will be useful as a motivation in the next section. Since  $A_i^n(\cdot) - \int_0^{\cdot} \psi_{1,i}^n(s) ds$  is a martingale, we have

$$\begin{split} L_{1}^{\varrho}(\hat{\mathbb{Q}}_{1,i}^{n} \| \mathbb{P}_{1,i}^{n}) \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^{n}} \Big[ \int_{0}^{\infty} \rho e^{-\varrho t} \Big( \int_{0}^{t} \log \Big( \frac{\psi_{1,i}^{n}(s)}{\lambda_{i}^{n}} \Big) dA_{i}^{n}(s) - \int_{0}^{t} (\psi_{1,i}^{n}(s) - \lambda_{i}^{n}) ds \Big) dt \Big] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^{n}} \Big[ \int_{0}^{\infty} \rho e^{-\varrho t} \Big( \int_{0}^{t} \log \Big( \frac{\psi_{1,i}^{n}(s)}{\lambda_{i}^{n}} \Big) (dA_{i}^{n}(s) - \psi_{1,i}^{n}(s) ds) \\ &\quad + \int_{0}^{t} \Big\{ \psi_{1,i}^{n}(s) \log \Big( \frac{\psi_{1,i}^{n}(s)}{\lambda_{i}^{n}} \Big) - \psi_{1,i}^{n}(s) + \lambda_{i}^{n} \Big\} ds \Big) dt \Big] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^{n}} \Big[ \int_{0}^{\infty} \rho e^{-\varrho t} \Big( \int_{0}^{t} \Big\{ \psi_{1,i}^{n}(s) \log \Big( \frac{\psi_{1,i}^{n}(s)}{\lambda_{i}^{n}} \Big) - \psi_{1,i}^{n}(s) + \lambda_{i}^{n} \Big\} ds \Big) dt \Big] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^{n}} \Big[ \int_{0}^{\infty} \rho e^{-\varrho t} \Big\{ \psi_{1,i}^{n}(t) \log \Big( \frac{\psi_{1,i}^{n}(t)}{\lambda_{i}^{n}} \Big) - \psi_{1,i}^{n}(t) + \lambda_{i}^{n} \Big\} dt \Big], \end{split}$$

where a change of variables is used to obtain the last equality. The same analysis applies for  $S_i^n$  as well. Consider the processes

$$\hat{\psi}_{1,i}^n(t) := (\lambda_i n)^{-1/2} (\psi_{1,i}^n(t) - \lambda_i^n), \qquad \hat{\psi}_{2,i}^n(t) := (\mu_i n)^{-1/2} (\psi_{2,i}^n(t) - \mu_i^n).$$

The Taylor expansion of  $\log(1+x)$ , gives<sup>2</sup>

$$\frac{1}{\kappa_{1,i}} L_1^{\varrho}(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) + \frac{1}{\kappa_{2,i}} L_2^{\varrho}(\hat{\mathbb{Q}}_{2,i}^n \| \mathbb{P}_{2,i}^n) \tag{2.10}$$

$$\approx \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^n \times \hat{\mathbb{Q}}_{2,i}^n} \left[ \int_0^\infty e^{-\varrho t} \left\{ \frac{1}{2\kappa_{1,i}} (\hat{\psi}_{1,i}^n(t))^2 + \frac{1}{2\kappa_{2,i}} \rho_i (\hat{\psi}_{2,i}^n(t))^2 \right\} dt \right].$$

The term  $\rho_i$  is due to the convergence  $T_i^n(t) \to \rho_i t$ ,  $t \in \mathbb{R}_+$ . Now, since the maximizer is free to choose  $\hat{\psi}_{1,i}^n$  and  $\hat{\psi}_{2,i}^n$ , she faces the two steps optimization problem. First, to choose  $\hat{\psi}_i^n(t)$  and then to solve

$$\min_{(\hat{\psi}_{1,i}^n(t),\hat{\psi}_{2,i}^n(t))} \left\{ \frac{1}{2\kappa_{1,i}} (\hat{\psi}_{1,i}^n(t))^2 + \frac{1}{2\kappa_{2,i}} \rho_i (\hat{\psi}_{2,i}^n(t))^2 : \lambda_i^{1/2} \hat{\psi}_{1,i}^n(t) - \rho_i \mu_i^{1/2} \hat{\psi}_{2,i}^n(t) = (2\lambda_i)^{1/2} \hat{\psi}_i^n(t) \right\}.$$

The minimal value of the above equals  $(2\hat{\varepsilon}_i)^{-1}(\hat{\psi}_i^n(t))^2$ , where

$$\hat{\varepsilon}_i := \frac{1}{2} (\kappa_{1,i} + \kappa_{2,i}).$$
(2.11)

Therefore, (2.10) can be approximated by  $(2\hat{\varepsilon}_i)^{-1}(\hat{\psi}_i^n(t))^2$ . These arguments are rigorously justified in [14, Section 4].

### 2.2 Two stochastic differential games

### 2.2.1 The multidimensional stochastic differential game (MSDG)

The sequence of the scaled and centered 2*I*-dimensional Poisson processes  $(\hat{A}^n, \hat{S}^n)$  weakly converges to a 2*I*-dimensional Brownian motion starting at zero, with zero mean and the covariance matrix Diag  $(\lambda_1^{1/2}, \ldots, \lambda_I^{1/2}, \mu_1^{1/2}, \ldots, \mu_I^{1/2})$ . Formally speaking, if the process  $\hat{Y}^n$  is of order one as  $n \to \infty$ , which is rigorously proven in [14, Section 4], we get from its definition in (2.4) that  $T^n(t) \to (\rho_1 t, \ldots, \rho_I t), t \in \mathbb{R}$ , and therefore, under  $\mathbb{P}^n$ ,  $(\hat{A}^n_i - \hat{S}^n_i(T^n_i))_{i=1}^I$  weakly converges to an *I*-dimensional Brownian motion starting at zero, with zero mean and the covariance matrix

$$\hat{\sigma} = (\hat{\sigma}_{ij}) := \operatorname{Diag}\left((2\lambda_1)^{1/2}, \dots, (2\lambda_I)^{1/2}\right),$$

where  $\hat{S}_i^n(T_i^n) := \hat{S}_i^n(T_i^n(\cdot)), \ i \in [I].$ 

Recall that in the QCP, an admissible control is of the form  $(U^n, R^n)$ . Notice that  $(\hat{Y}^n(t), \hat{R}^n(t))$  is uniquely determined by  $(U^n(s), R^n(s))_{0 \le s \le t}$ . In the MSDG we consider two *I*-dimensional processes,  $\hat{R}$  and  $\hat{Y}$  that play the roles of instantaneous controls, which stand for the scaled rejection process  $\hat{R}^n$  and the scaled idle time process  $\hat{Y}^n$ , respectively. Moreover,

<sup>&</sup>lt;sup>2</sup>This is rigorously justified in [14] by another level of approximation to the processes  $\{\psi_{i,i}^n\}$ .

the ambiguity about the true underlying probability measure in the QCP, which is formulated by a penalty for deviating from the reference measure is translated to the limiting problem as well. In the definition below we refer to two players by their roles as a minimizer and a maximizer even though the roles can only be derived from the cost function, which is presented afterwards.

Definition 2.1 (admissible controls, MSDG) An admissible control for the minimizer for any initial state  $\hat{x}_0 \in \mathcal{X}$  is a filtered probability space

$$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}) := \Big(\prod_{i=1}^{I} \Omega^i, \mathcal{F}^1 \otimes \ldots \otimes \mathcal{F}^I, \{\mathcal{F}_t\}, \prod_{i=1}^{I} \mathbb{P}_i\Big),$$

that supports a process  $(\hat{Y}, \hat{R})$  taking values in  $(\mathbb{R}^{I}_{+})^{2}$  with RCLL sample paths adapted to the filtration  $\{\mathcal{F}_t\}$ , where  $(\Omega^i, \mathcal{F}^i, \{\mathcal{F}^i_t\}, \mathbb{P}_i)$  supports a one-dimensional standard Brownian motion  $\hat{B}_i$  adapted to the filtration  $\{\mathcal{F}_t^i\}, i \in [I]$ . Moreover, assume that the following properties hold: (i)

for every  $i \in [I]$  and  $0 \leq s < t$ ,  $\hat{B}_i(t) - \hat{B}_i(s)$  is independent of  $\mathcal{F}_s^i$  under  $\mathbb{P}_i$ ; (2.12)

(ii)

 $\theta \cdot \hat{Y}$  and  $\hat{R}_i, i \in [I]$  are nonnegative and nondecreasing, where  $\theta := (\mu_1^{-1}, \dots, \mu_I^{-1});$  (2.13) (iii)

$$\hat{X}(t) = \hat{x}_0 + \hat{m}t + \hat{\sigma}\hat{B}(t) + \hat{Y}(t) - \hat{R}(t), \quad t \in \mathbb{R}_+,$$
(2.14)

such that

$$\ddot{X}(t) \in \mathcal{X}, \quad t \in \mathbb{R}_+, \ \mathbb{P}\text{-}a.s.,$$
(2.15)

where  $\hat{m} := \lim_{n} \hat{m}^{n} = (\hat{\lambda}_{i} - \rho_{i}\hat{\mu}_{i})_{i=1}^{I}$  and  $\hat{B} = (\hat{B}_{i})_{i=1}^{I}$ . An admissible control for the maximizer is a product measure  $\hat{\mathbb{Q}} = \prod_{i=1}^{I} \hat{\mathbb{Q}}_{i}$ , where each  $\hat{\mathbb{Q}}_{i}$ is defined on  $(\Omega^i, \mathcal{F}^i, \{\mathcal{F}^i_t\})$ , such that

$$\frac{d\hat{\mathbb{Q}}_i(t)}{d\mathbb{P}_i(t)} = \exp\Big\{\int_0^t \hat{\psi}_i(s)d\hat{B}_i(s) - \frac{1}{2}\int_0^t \hat{\psi}_i^2(s)ds\Big\}, \quad t \in \mathbb{R}_+,$$
(2.16)

for an  $\{\mathcal{F}_t\}$ -progressively measurable process  $\hat{\psi} = (\hat{\psi}_1, \dots, \hat{\psi}_I)$  satisfying

$$\mathbb{E}^{\mathbb{P}}\left[\int_{0}^{\infty} e^{-\varrho s} \hat{\psi}_{i}^{2}(s) ds\right] < \infty \quad and \quad \mathbb{E}^{\mathbb{P}}\left[e^{\frac{1}{2}\int_{0}^{t} \hat{\psi}_{i}^{2}(s) ds}\right] < \infty \quad t \in \mathbb{R}_{+}, \ i \in [I].$$
(2.17)

We consider a probability space that is constructed from I small probability spaces, where each one supports the processes associated with one of the classes (in the QCP studied in [14] there are 2I small probability spaces as can be inferred from the structure of  $\mathbb{Q}^n$ ). The Brownian motion approximates the difference  $(\hat{A}_i^n - \hat{S}_i^n(T_i^n))_{i=1}^I$  up to the deterministic covariance matrix, and (2.14) follows by (2.1). The emphasis of Condition (2.12) is that the independent condition for  $\hat{B}_i$  is given w.r.t. the filtration  $\mathcal{F}_t^i$  and not simply w.r.t. its own filtration, which merely follows by being a Brownian motion. It is necessary for the approximation procedure, see the details in [14, Section 4]. Condition (2.13) follows since the rejection process (in the QCP) and also  $\theta^n \cdot \hat{Y}^n$  are nondecreasing. Occasionally, we refer to  $\hat{R}$  as the rejection process in the MSDG. The buffer constraint is imposed in (2.15). Pay attention that we consolidate the processes  $\hat{A}_i^n$  and  $\hat{S}_i^n(T_i^n)$  into one Brownian motion. Hence, we consider only I changes of measures instead of 2I.

We now explain the change of measure structure. Recall the definition of  $\hat{\psi}_{j,i}^n$  and  $\hat{\psi}_i$  from the previous section. Then,

$$\hat{A}_{i}^{n}(t) = n^{-1/2} \Big( A_{i}^{n}(t) - \int_{0}^{t} \psi_{1,i}^{n}(s) ds \Big) + \lambda_{i}^{1/2} \int_{0}^{t} \hat{\psi}_{1,i}^{n}(s) ds,$$

$$\hat{S}_{i}^{n}(T_{i}^{n}(t)) = n^{-1/2} \Big( S_{i}^{n}(T_{i}^{n}(t)) - \int_{0}^{t} \psi_{2,i}^{n}(s) dT_{i}^{n}(s) \Big) + \mu_{i}^{1/2} \int_{0}^{t} \hat{\psi}_{2,i}^{n}(s) dT_{i}^{n}(s).$$
(2.18)

Informally speaking, in both lines above, under  $\hat{\mathbb{Q}}_{1,i}^n \times \hat{\mathbb{Q}}_{2,i}^n$ , the first term is approximately a standard Brownian motion and the second term approximates a drift (in case it converges). This claim is handelled rigorously in [14, Section 4]. As a result, under  $\hat{\mathbb{Q}}_i^n$ ,  $\hat{A}_i^n - \hat{S}_i^n(T_i^n)$  is approximately a diffusion process with drift  $(2\lambda_i)^{1/2}\hat{\psi}_i^n := \lambda_i^{1/2}\hat{\psi}_{1,i}^n - \rho_i\mu_i^{1/2}\hat{\psi}_{2,i}^n$  and a diffusion coefficient  $\sigma_{ii} = (2\lambda_i)^{1/2}$ ; the term  $\rho_i$  is due to the convergence  $T_i^n(t) \to \rho_i t, t \in \mathbb{R}_+$ .

**Remark 2.1** (i) Notice that the structure of the information in the game is consistent with the one in the QCP. The minimizer chooses a strategy and the maximizer, which is penalized for deviating from the reference measure, responds to this strategy by choosing a worst case scenario. For further reading about the structure of the information in control problems with model uncertainty, the reader is referred to [32].

(ii) Given any  $\{\mathcal{F}_t\}$ -progressively measurable process  $\hat{\psi}$  that satisfies the conditions in (2.17), the right-hand side (r.h.s.) of (2.16) is a martingale, and therefore, there exists a probability measure  $\hat{\mathbb{Q}}$  such that  $\hat{\mathbb{Q}}|_{\mathcal{F}_t}$  satisfies (2.16) for all  $t \in \mathbb{R}_+$ .

(iii) Equation (2.14) can alternatively be written as

$$\hat{X}(t) = \hat{x}_0 + \hat{m}t + \int_0^t \hat{\sigma}\hat{\psi}(s)ds + \hat{\sigma}\hat{B}^{\hat{\mathbb{Q}}}(t) + \hat{Y}(t) - \hat{R}(t), \quad t \in \mathbb{R}_+,$$
(2.19)

where  $\hat{B}^{\hat{\mathbb{Q}}}(t) := \hat{B}(t) - \int_0^t \hat{\psi}(s) ds$ ,  $t \in \mathbb{R}_+$ , is an  $\{\mathcal{F}_t\}$ -I-dimensional standard Brownian motion under  $\hat{\mathbb{Q}}$ .

Denote by  $\hat{\mathcal{A}}(\hat{x}_0)$  the set of all admissible controls for the minimizer, given the initial condition  $\hat{x}_0$ . We often abuse notation and denote  $(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)$ , keeping in mind that the control includes a filtered probability space. The set of all admissible controls for the maximizer is denoted by  $\hat{\mathcal{Q}}(\hat{x}_0)$ .

The cost function (MSDG). Recall the discussion about the approximation of the cost function of the QCP from the previous section and (2.11). Set  $\hat{\varepsilon} = (\hat{\varepsilon}_i)_{i=1}^I$ . The cost associated

with the initial condition  $\hat{x}_0$  and the strategies  $(\hat{Y}, \hat{R})$  and  $\hat{\mathbb{Q}}$  is given by

$$\hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \hat{\varepsilon}) := \mathbb{E}^{\hat{\mathbb{Q}}} \Big[ \int_0^\infty e^{-\varrho t} \left( \hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) \right) \Big] - \sum_{i=1}^I \frac{1}{\hat{\varepsilon}_i} L^{\varrho}(\hat{\mathbb{Q}}_i || \mathbb{P}_i), \qquad (2.20)$$

where

$$L^{\varrho}(\hat{\mathbb{Q}}_{i}||\mathbb{P}_{i}) := \mathbb{E}^{\hat{\mathbb{Q}}_{i}}\left[\int_{0}^{\infty} \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_{i}(t)}{d\mathbb{P}_{i}(t)} dt\right]$$
(2.21)

The cost function can alternatively be expressed by

$$\hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \hat{\varepsilon}) = \mathbb{E}^{\hat{\mathbb{Q}}} \Big[ \int_0^\infty e^{-\varrho t} \Big( \hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) - \sum_{i=1}^I \frac{1}{2\hat{\varepsilon}_i} \hat{\psi}_i^2(t) dt \Big) \Big],$$
(2.22)

with  $\hat{\psi}$  satisfying (2.16)–(2.17) above. Indeed,

$$L^{\varrho}(\hat{\mathbb{Q}}_{i}||\mathbb{P}_{i}) = \mathbb{E}^{\hat{\mathbb{Q}}_{i}}\left[\int_{0}^{\infty} \varrho e^{-\varrho t} \left(-\frac{1}{2} \int_{0}^{t} \hat{\psi}_{i}^{2}(s) ds + \int_{0}^{t} \hat{\psi}_{i}(s) d\hat{B}_{i}(s)\right) dt\right]$$

$$= \mathbb{E}^{\hat{\mathbb{Q}}_{i}}\left[\int_{0}^{\infty} \varrho e^{-\varrho t} \left(-\frac{1}{2} \int_{0}^{t} |\hat{\psi}_{i}(s)|^{2} ds + \int_{0}^{t} \hat{\psi}_{i}(s) \cdot (d\hat{B}_{i}^{\hat{\mathbb{Q}}}(s) + \hat{\psi}_{i}(s) ds)\right) dt\right]$$

$$= \mathbb{E}^{\hat{\mathbb{Q}}_{i}}\left[\int_{0}^{\infty} \varrho e^{-\varrho t} \left(\frac{1}{2} \int_{0}^{t} \hat{\psi}_{i}^{2}(s) ds\right) dt\right] = \mathbb{E}^{\hat{\mathbb{Q}}_{i}}\left[\frac{1}{2} \int_{0}^{\infty} \hat{\psi}_{i}^{2}(s) \int_{s}^{\infty} \varrho e^{-\varrho t} dt ds\right]$$

$$= \mathbb{E}^{\hat{\mathbb{Q}}_{i}}\left[\frac{1}{2} \int_{0}^{\infty} e^{-\varrho t} \hat{\psi}_{i}^{2}(t) dt\right] < \infty.$$

$$(2.23)$$

Compare this structure with the approximated penalty in the QCP given at the end of the previous section.

While the form of the cost function given in (2.20) captures better the ambiguity aversion, the form of the cost given in (2.22) is more useful from a technical point of view. Moreover, the dynamics in (2.19) together with the cost function given in (2.22) are similar in their structure to their correspondences in Equation (11) and the display below (13) together with (2) in [2]. The DM is faced the following optimization problem

$$\hat{V}(\hat{x}_0;\hat{\varepsilon}) = \inf_{(\hat{Y},\hat{R})\in\hat{\mathcal{A}}(\hat{x}_0)} \sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0,\hat{Y},\hat{R},\hat{\mathbb{Q}};\hat{\varepsilon})$$

### 2.2.2 The reduced stochastic differential game (RSDG)

We now present the RSDG. This game is one-dimensional and obtained by projecting the processes from (2.14) in the  $\theta$  direction, which is given in (2.13). For this we need the following notation,

$$x_0 := \theta \cdot \hat{x}_0, \quad m := \theta \cdot \hat{m}, \quad \sigma := \|\theta\hat{\sigma}\|, \tag{2.24}$$

$$\varepsilon := \frac{1}{\sigma^2} \sum_{i=1}^{I} (\theta \hat{\sigma})_i^2 \hat{\varepsilon}_i, \qquad (2.25)$$

 $b := \max\{\theta \cdot \hat{\xi} : \hat{\xi} \in \mathcal{X}\} = \theta \cdot \hat{b},$ 

where  $\hat{b} = (\hat{b}_i)_{i=1}^I$ .

**Definition 2.2 (admissible controls, RSDG)** An admissible control for the minimizer for any initial state  $x_0 \in [0, b]$  is a filtered probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$  that supports a onedimensional standard Brownian motion B and a process (Y, R) taking values in  $\mathbb{R}^2_+$  with RCLL sample paths, both adapted to the filtration  $\{\mathcal{F}_t\}$  and satisfy the following properties: (i) for every  $0 \leq s < t$ , B(t) - B(s) is independent of  $\mathcal{F}_s$  under  $\mathbb{P}$ ; (ii) Y and R are nonnegative and nondecreasing; (iii)

$$X(t) = x_0 + mt + \sigma B(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+,$$
(2.26)

such that

$$X(t) \in [0, b], \quad t \in \mathbb{R}_+, \ \mathbb{P}$$
-a.s.

An admissible control for the maximizer is a meausre  $\mathbb{Q}$  defined on  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\})$  such that

$$\frac{d\mathbb{Q}(t)}{d\mathbb{P}(t)} = \exp\Big\{\int_0^t \psi(s)dB(s) - \frac{1}{2}\int_0^t \psi^2(s)ds\Big\}, \quad t \in \mathbb{R}_+,$$
(2.27)

for an  $\{\mathcal{F}_t\}$ -progressively measurable process  $\psi$  satisfying

$$\mathbb{E}^{\mathbb{P}}\left[\int_{0}^{\infty} e^{-\varrho s} \psi^{2}(s) ds\right] < \infty \quad and \quad \mathbb{E}^{\mathbb{P}}\left[e^{\frac{1}{2}\int_{0}^{t} \psi^{2}(s) ds}\right] < \infty \quad for \ every \ t \in \mathbb{R}_{+}.$$
(2.28)

The statements given in Remark 2.1 also hold for the RSDG as well. For completeness of the presentation and for later references we provide an alternative form of the dynamics given in (2.26),

$$X(t) = x_0 + mt + \int_0^t \sigma \psi(s) ds + \sigma B^{\mathbb{Q}}(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+,$$
(2.29)

where  $B^{\mathbb{Q}}(t) := B(t) - \int_0^t \psi(s) ds$ ,  $t \in \mathbb{R}_+$ , is an  $\{\mathcal{F}_t\}$ -one-dimensional standard Brownian motion under  $\mathbb{Q}$ .

Denote by  $\mathcal{A}(x_0)$  the set of all admissible controls for the minimizer, given the initial condition  $x_0$ . As before, we often abuse notation and denote  $(Y, R) \in \mathcal{A}(x_0)$ , keeping in mind that the control includes a filtered probability space. The set of all admissible controls for the maximizer is denoted by  $\mathcal{Q}(x_0)$ .

The cost function (RSDG). The expected cost associated with the initial condition x and the controls (Y, R) and  $\mathbb{Q}$  is given by

$$J(x_0, Y, R, \mathbb{Q}; \varepsilon) := \mathbb{E}^{\mathbb{Q}} \Big[ \int_0^\infty e^{-\varrho t} \left( h(X(t)) dt + r dR(t) \right) \Big] - \frac{1}{\varepsilon} L^{\varrho}(\mathbb{Q} || \mathbb{P}),$$

and

where

$$h(x) := \min\{\hat{h} \cdot \hat{\xi} : \hat{\xi} \in \mathcal{X}, \ \theta \cdot \hat{\xi} = x\},$$
(2.30)

$$r := \min\{\hat{r} \cdot q : q \in \mathbb{R}^I_+, \ \theta \cdot q = 1\},\tag{2.31}$$

and  $L^{\varrho}(\mathbb{Q}||\mathbb{P})$  is given by (2.21) with  $(\mathbb{Q}, \mathbb{P})$  replacing  $(\hat{\mathbb{Q}}_i, \mathbb{P}_i)$ . By the convexity of  $\mathcal{X}$  it follows that h is convex. In fact, h is piecewise linear and Lipschitz continuous. Moreover,  $h(x) \geq 0$ for  $x \geq 0$  and equality holds if and only if x = 0. Therefore, h is strictly increasing. In [6, page 568] it is shown that there is  $i^* \in [I]$  such that,

$$r = r_{i^*} \mu_{i^*} := \min\{r_i \mu_i : i \in [I]\}.$$
(2.32)

The index  $i^*$  stands for the class with the smallest rejection cost, weighted with the mean service rate. In fact, as we discuss in Section 2.3 and prove in Theorem 4.1, under optimality of both players in the MSDG, rejections are performed only from this class.

By the same arguments that lead to (2.22), the cost function of the RSDG can alternatively be expressed by the technically more convenient form,

$$J(x_0, Y, R, \mathbb{Q}; \varepsilon) = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^\infty e^{-\varrho t} \left( h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi^2(t) dt \right) \right],$$
(2.33)

with  $\psi$  satisfying (2.28) above. The value function is given by

$$V(x_0;\varepsilon) = \inf_{(Y,R)\in\mathcal{A}(x_0)} \sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} J(x_0,Y,R,\mathbb{Q};\varepsilon).$$
(2.34)

**Remark 2.2** In case that there is no ambiguity, we define the cost and the value functions by

$$J_{NA}(x_0, Y, R) := \mathbb{E}^{\mathbb{P}} \Big[ \int_0^\infty e^{-\varrho t} (h(X(t))dt + rdR(t)) \Big],$$
(2.35)  
$$V(x_0; 0) := \inf_{\substack{(Y,R) \in \mathcal{A}(x_0)}} J_{NA}(x_0, Y, R).$$

This problem was studied by Harrison and Taksar [21] and later on was used by Atar and Shifrin [6]. In Theorem 5.1 we show that this problem is obtained when the ambiguity vanishes, that is,  $\lim_{\varepsilon \to 0} V(x_0; \varepsilon) = V(x_0; 0)$ .

#### 2.3 The relationship between the games

We now show that the last two games share the same value and moreover, that given any admissible control for the minimizer in either one of the games, one can construct an admissible control in the other game that performs at least as well. For this, we define a function  $\gamma$ , taken from [6, Equations (48)–(49)], that sends any workload value  $x_0$  to the cheapest state of the MSDG (from the holding cost perspective) among all the states whose workload levels are  $x_0$ . Using this function and an optimal strategy for the minimizer in the RSDG, we construct an optimal strategy for the minimizer in the MSDG, see Theorem 4.1. In order to define the function it is convenient to assume without loss of generality that

$$h_1\mu_1 \ge h_2\mu_2 \ge \cdots \ge h_I\mu_I.$$



Figure 1: The graphs refer to the case I = 3,  $\hat{h} = (1, 5/2, 3/2)$ ,  $\mu = (3, 1, 3/2)$ , and  $(\hat{b}_1, \hat{b}_2, \hat{b}_3) = (4, 7, 6)$ . The graph to the left stands for the workload levels. The curve of the function  $\gamma$  is in bold in the graph to the right. The workload levels with the lower case letters are j = 0,  $k = \hat{b}_3/\mu_3 = 4$ ,  $l = \hat{b}_3/\mu_3 + \hat{b}_2/\mu_2 = 11$ ,  $m = \hat{b}_3/\mu_3 + \hat{b}_2/\mu_2 + 1$ , and  $n = \hat{b}_3/\mu_3 + \hat{b}_2/\mu_2 + \hat{b}_1/\mu_1 = b = 37/3$ . They respectively correspond to the upper case letters:  $J = (0, 0, 0), K = (0, 0, \hat{b}_3/\mu_3) = (0, 0, 4), L = (0, \hat{b}_2/\mu_2, \hat{b}_3/\mu_3) = (0, 7, 4), M = (1, \hat{b}_2/\mu_2, \hat{b}_3/\mu_3) = (1, 7, 4)$ , and  $N = (\hat{b}_1/\mu_1, \hat{b}_2/\mu_2, \hat{b}_3/\mu_3) = (4/3, 7, 4)$ .

Recall that  $b = \theta \cdot (\hat{b}_1, \ldots, \hat{b}_I)$ . Given  $x \in [0, b)$ , let (j, v) be the unique pair that is determined by

$$x = \sum_{i=j+1}^{I} \theta_i \hat{b}_i + \theta_j \upsilon, \qquad j \in [I], \quad \upsilon \in [0, \hat{b}_j),$$

and for x = b, take  $(j, v) = (1, \hat{b}_1)$ . Let  $\gamma : [0, b] \to \mathcal{X}$  be the function given by

$$\gamma(x) = \sum_{i=j+1}^{I} \hat{b}_i e_i + v e_j, \qquad (2.36)$$

where  $\{e_1, \ldots, e_I\}$  is the standard basis of  $\mathbb{R}^I$ . By filling up the cheaper buffers first (w.r.t. the holding costs) we get that,

$$\gamma(x) \in \arg\min\{\hat{h} \cdot \hat{\xi} : \hat{\xi} \in \mathcal{X}, \theta \cdot \hat{\xi} = x\}.$$
(2.37)

The curve  $\gamma(x), x \in [0, b]$  is continuous and located on the edges of  $\mathcal{X}$ , see Figure 1. The idea is as follows, recall that the components of  $\hat{Y} = (\hat{Y}_i)_{i=1}^I$  can be positive or negative, as long as  $\theta \cdot \hat{Y}$  is nonnegative and nondecreasing. Now, as the workload changes in the interval [0, b), the DM can use only the process  $\hat{Y}$ , without the need of the rejection process  $\hat{R}$ , so that  $\hat{X}$  moves along the curve of  $\gamma$ . As will be shown in Theorem 4.1, under optimality, the rejection process is used only to reduce the workload, and only from the class which has the cheapest weighted rejections cost, denoted by  $i^*$ . We discuss more about the minimizer's optimal strategy in the MSDG in Remark 2.3 and in the paragraph that comes before Theorem 4.1.

We state the proposition below for an arbitrary initial point  $\hat{x}_0 \in \mathcal{X}$  and show that  $\hat{V}(\hat{x}_0; \hat{\varepsilon}) = V(x_0; \varepsilon)$ , where recall that  $x_0 = \theta \cdot \hat{x}_0$ .

**Proposition 2.1** Fix  $\hat{\varepsilon} = (\hat{\varepsilon}_i)_{i=1}^I$  and  $\hat{x}_0 \in \mathcal{X}$ . Let  $\varepsilon$  be given by (2.25).

(i) Given an admissible control in the MSDG,  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, \hat{B}, \hat{Y}, \hat{R})$  and an admissible measure  $\mathbb{Q}_*$  associated with  $\psi_*$  in the RSDG that satisfies (2.27)–(2.28), set  $B = \frac{1}{\sigma}\theta\hat{\sigma}\cdot\hat{B}$ ,  $(X, Y, R) = (\theta \cdot \hat{X}, \theta \cdot \hat{Y}, \theta \cdot \hat{R})$ , and  $\hat{\psi}_* = (\hat{\psi}_{*,1}, \dots, \hat{\psi}_{*,I})$ , by

$$\hat{\psi}_{*,i}(t) := \frac{\sigma \psi_*(t)(\theta \hat{\sigma})_i \hat{\varepsilon}_i}{\sum_{j=1}^I (\theta \hat{\sigma})_j^2 \hat{\varepsilon}_j}, \qquad t \in \mathbb{R}_+, \ i \in [I].$$
(2.38)

Let  $\hat{\mathbb{Q}}_*$  be the associated measure defined through (2.16). Then  $(Y, R) \in \mathcal{A}(x_0), \ \hat{\mathbb{Q}}_* \in \hat{\mathcal{Q}}(\hat{x}_0),$ and

$$J(x_0, Y, R, \mathbb{Q}_*; \varepsilon) \le \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_*; \hat{\varepsilon}).$$

$$(2.39)$$

As a consequence,

$$\sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \varepsilon) \le \sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \hat{\varepsilon}).$$

(ii) Conversely, consider an admissible control in the RSDG,  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B, Y, R)$ , which is assumed to support an I-dimensional standard Brownian motion  $\hat{B}$ . Consider also an admissible  $\hat{\mathbb{Q}}_{\sharp}$  associated with  $\hat{\psi}_{\sharp}$  in the MSDG that satisfies (2.16)–(2.17). Assume that  $\hat{B}$  is  $\{\mathcal{F}_t\}$ -adapted and satisfies  $\frac{1}{\sigma}\theta\hat{\sigma}\cdot\hat{B} = B$  and (2.12). Define  $(\hat{X}, \hat{Y}, \hat{R})$  by

$$\hat{X}(t) := \gamma(X(t)), \quad \hat{R}(t) := R(t)\mu_{i^*}e_{i^*},$$
(2.40)

and

$$\hat{Y}(t) := \hat{X}(t) - \hat{x}_0 - \hat{m}t - \hat{\sigma}\hat{B}(t) + \hat{R}(t).$$
(2.41)

Also, let

$$\psi_{\sharp}(t) := \frac{1}{\sigma} \theta \hat{\sigma} \cdot \hat{\psi}_{\sharp}(t), \quad t \in \mathbb{R}_{+},$$
(2.42)

and let  $\mathbb{Q}_{\sharp}$  be the associated measure defined through (2.27). Then  $(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0), \mathbb{Q}_{\sharp} \in \mathcal{Q}(x_0),$ and

$$\hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_{\sharp}; \hat{\varepsilon}) \le J(x_0, Y, R, \mathbb{Q}_{\sharp}; \varepsilon).$$

As a consequence,

$$\sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}; \hat{\varepsilon}) \le \sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} J(x_0, Y, R, \mathbb{Q}; \varepsilon).$$
(2.43)

Notice that in part (i) we define the dynamics (X, Y, R) in the RSDG from the dynamics  $(\hat{X}, \hat{Y}, \hat{R})$  in the MSDG by projecting in the  $\theta$  direction. In part (ii) we consider an auxiliary *I*-dimensional standard Brownian motion for technical reasons. The proof that the probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}, B, Y, R)$  supports such a Brownian motion is merely technical and is fully given in [6, Proposition 2.1.(iii)]. Therefore, it is omitted. Also, in (2.40)–(2.41) we construct first  $\hat{X}$  from the one-dimensional process X using the well-defined function  $\gamma$  that maps [0, b] to  $\mathcal{X}$ . Then, we define  $\hat{R}$  as a process with degeneracy in its  $[I] \setminus \{i^*\}$  coordinates. Only then we define  $\hat{Y}$  in such a way that (2.14) holds.

**Corollary 2.1** Fix  $\hat{\varepsilon} = (\hat{\varepsilon}_i)_{i=1}^I$  and  $\hat{x}_0 \in \mathcal{X}$ . Let  $\varepsilon$  be given by (2.25). Then,  $\hat{V}(\hat{x}_0; \hat{\varepsilon}) = V(x_0; \varepsilon)$ .

**Remark 2.3** In Theorem 3.1 we show that the minimizer in the RSDG has a simple optimal strategy, which uses minimal idleness and minimal amount of rejections in order to keep the workload in some subinterval  $[0,\beta] \subseteq [0,b]$ . In Theorem 4.1 we use the function  $\gamma$  given in (2.36) and the relations given in (2.40)–(2.41) to construct an optimal strategy for the minimizer in the MSDG. An optimal strategy for the maximizer emerges from (2.38) and (2.42).

**Proof of Proposition 2.1:** Some of the arguments in the proof are given in [6, Proposition 2.1]. However, the ones that involve changes of the probability measure are new. For completeness of the proof we provide all the details.

(i) The proof that  $(Y, R) \in \mathcal{A}(x_0)$  is straightforward and therefore omitted. Notice that  $\|\hat{\psi}_*(\cdot)\| \leq |\psi_*(\cdot)|$  and therefore,  $\hat{\mathbb{Q}}_* \in \hat{\mathcal{Q}}(\hat{x}_0)$  follows since  $\mathbb{Q}_* \in \mathcal{Q}(x_0)$ .

We now show that the distribution of X under  $\mathbb{Q}_*$  is the same as under  $\hat{\mathbb{Q}}_*$ . Replacing  $\hat{\mathbb{Q}}$  by  $\hat{\mathbb{Q}}_*$  in equation (2.19) and multiplying its both sides by  $\theta$  yield

$$X(t) = x_0 + mt + \int_0^t \theta \hat{\sigma} \hat{\psi}_*(s) ds + \theta \hat{\sigma} \hat{B}^{\hat{\mathbb{Q}}_*}(t) + Y(t) - R(t)$$
(2.44)  
=  $x_0 + mt + \int_0^t \sigma \psi_*(s) ds + \sigma B^{\mathbb{Q}_*}(t) + Y(t) - R(t).$ 

The equality between the integrals follows by the definitions of  $\hat{\psi}_*$  and  $\sigma$ . The equality between the Brownian motion terms follows since

$$\theta\hat{\sigma}\hat{B}^{\hat{\mathbb{Q}}_*}(t) = \theta\hat{\sigma}\hat{B}(t) - \int_0^t \theta\hat{\sigma}\hat{\psi}_*(s)ds = \sigma B(t) - \sigma \int_0^t \psi_*(s)ds = \sigma B^{\mathbb{Q}_*},$$

which in turn follows by using the definitions of  $\hat{B}^{\hat{\mathbb{Q}}_*}$  and  $B^{\mathbb{Q}_*}$  ((2.19) and (2.29)) together with the definition of B from the proposition and once again the definition of  $\hat{\psi}_*$ . Recall that  $\hat{B}^{\hat{\mathbb{Q}}_*}$ is an I-dimensional standard Brownian motion under  $\hat{\mathbb{Q}}_*$  and that  $\theta\hat{\sigma}$  is a deterministic vector with norm  $\sigma$ . Then from the above, we get that under  $\hat{\mathbb{Q}}_*$ , the process  $B^{\mathbb{Q}_*}$  is a one-dimensional standard Brownian motion. Since  $B^{\mathbb{Q}_*}$  is a one-dimensional standard Brownian motion also under  $\mathbb{Q}_*$ , we get from (2.44) that X admits the same distribution under  $\hat{\mathbb{Q}}_*$  and under  $\mathbb{Q}_*$ .

Next, by the definitions of h and r (see (2.30) and (2.31)) it follows that

$$h(X(t)) \le \hat{h} \cdot \hat{X}(t), \qquad t \in \mathbb{R}_+, \tag{2.45}$$

$$\int_0^\infty e^{-\varrho t} r dR(t) \le \int_0^\infty e^{-\varrho t} \hat{r} \cdot d\hat{R}(t).$$
(2.46)

Moreover, using the fact that the distribution of X under  $\hat{\mathbb{Q}}_*$  is the same as under  $\mathbb{Q}_*$  and the equality

$$\sum_{i=1}^{I} \frac{1}{2\hat{\varepsilon}_i} \hat{\psi}_{*,i}^2(t) = \frac{1}{2\varepsilon} \psi_*^2(t), \qquad t \in \mathbb{R}_+,$$

we get,

$$\begin{split} J(x_0, Y, R, \mathbb{Q}_*; \varepsilon) &= \mathbb{E}^{\mathbb{Q}_*} \Big[ \int_0^\infty e^{-\varrho t} \Big( h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi_*^2(t) \Big) dt \Big] \\ &= \mathbb{E}^{\hat{\mathbb{Q}}_*} \Big[ \int_0^\infty e^{-\varrho t} \Big( h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi_*^2(t) \Big) dt \Big] \\ &\leq \mathbb{E}^{\hat{\mathbb{Q}}_*} \Big[ \int_0^\infty e^{-\varrho t} \Big( \hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) - \sum_{i=1}^I \frac{1}{2\hat{\varepsilon}_i} \hat{\psi}_{*,i}^2(t) \Big) dt \Big] \\ &= \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_*; \hat{\varepsilon}). \end{split}$$

Since  $\mathbb{Q}_* \in \mathcal{Q}(x_0)$  is arbitrary, it follows that

$$\sup_{\mathbb{Q}\in\mathcal{Q}(x_0)}J(x_0,Y,R,\mathbb{Q};\varepsilon)\leq \sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)}\hat{J}(\hat{x}_0,\hat{Y},\hat{R},\hat{\mathbb{Q}};\varepsilon)$$

(ii) We start by showing that  $(\hat{Y}, \hat{R}) \in \hat{\mathcal{A}}(\hat{x}_0)$ . The processes  $\hat{X}, \hat{R}$ , and  $\hat{Y}$  are  $\{\mathcal{F}_t\}$ -adapted since X, R, and  $\hat{B}$  are. Now,

$$\theta \cdot \hat{Y}(t) = \theta \cdot \hat{X}(t) - \theta \cdot x_0 - \theta \cdot \hat{m}t - \theta \cdot \hat{\sigma}\hat{B}(t) + \theta \cdot \hat{R}(t) = Y(t),$$

and since Y is admissible it follows that  $\theta \cdot \hat{Y} = Y$  is nonnegative and nondecreasing. The processes  $\{\hat{R}_i\}_{i=1}^{I}$  are clearly nonnegative and nondecreasing since so is R. Thus, property (2.13) holds. Finally, properties (2.14) and (2.15) follow by the definition of  $\gamma$  and the construction of  $(\hat{X}, \hat{Y}, \hat{R})$ .

Notice that,  $\|\hat{\psi}_{\sharp}(\cdot)\| = |\psi_{\sharp}(\cdot)|$  and therefore,  $\mathbb{Q}_{\sharp} \in \mathcal{Q}(x_0)$  follows since  $\hat{\mathbb{Q}}_{\sharp} \in \hat{\mathcal{Q}}(\hat{x}_0)$ . From Cauchy-Schwartz inequality,

$$\sum_{i=1}^{I} \frac{1}{2\hat{\varepsilon}_i} \hat{\psi}_{\sharp,i}^2(t) \ge \frac{1}{2\varepsilon} \psi_{\sharp}^2(t).$$

By the definitions of  $h, r, and \gamma$  and (2.32), (2.37), and (2.40) it follows that

$$\hat{h} \cdot \hat{X}(t) = h(X(t)), \qquad t \in \mathbb{R}_+,$$
$$\int_0^\infty e^{-\varrho t} \hat{r} \cdot d\hat{R}(t) = \int_0^\infty e^{-\varrho t} r d(\theta \cdot \hat{R}(t)) = \int_0^\infty e^{-\varrho t} r dR(t).$$

By the relationship between the *I*-dimensional processes  $\hat{B}$  and  $\hat{\psi}_{\sharp}$  and the one-dimensional processes B and  $\psi_{\sharp}$ , one can show as was done in the previous part that the distribution of X under  $\hat{\mathbb{Q}}_{\sharp}$  is the same as under  $\mathbb{Q}_{\sharp}$ . Therefore,

$$J(x_{0}, Y, R, \mathbb{Q}_{\sharp}; \varepsilon) = \mathbb{E}^{\mathbb{Q}_{\sharp}} \left[ \int_{0}^{\infty} e^{-\varrho t} \left( h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi_{\sharp}^{2}(t) \right) dt \right]$$

$$= \mathbb{E}^{\mathbb{Q}_{\sharp}} \left[ \int_{0}^{\infty} e^{-\varrho t} \left( h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi_{\sharp}^{2}(t) \right) dt \right]$$

$$\geq \mathbb{E}^{\mathbb{Q}_{\sharp}} \left[ \int_{0}^{\infty} e^{-\varrho t} \left( \hat{h} \cdot \hat{X}(t) dt + \hat{r} \cdot d\hat{R}(t) - \sum_{i=1}^{I} \frac{1}{2\varepsilon_{i}} \psi_{\sharp,i}^{2}(t) \right) dt \right]$$

$$= \hat{J}(\hat{x}_{0}, \hat{Y}, \hat{R}, \mathbb{Q}_{\sharp}; \hat{\varepsilon}).$$

$$(2.47)$$

Since  $\hat{\mathbb{Q}}_{\sharp} \in \hat{\mathcal{Q}}(\hat{x}_0)$  is arbitrary, it follows that

$$\sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0,\hat{Y},\hat{R},\hat{\mathbb{Q}};\hat{\varepsilon}) \leq \sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} J(x_0,Y,R,\mathbb{Q};\varepsilon).$$

# 3 Solution of the RSDG

In this section we provide a solution to the RSDG. In Section 3.1 we present the notion of a reflection strategy. Then in Section 3.2 we provide the HJB equation associated with the RSDG. We prove that the value function is the unique smooth solution of the HJB equation and show that the minimizer has an optimal reflecting strategy.

#### 3.1 Reflecting strategies

The optimal strategy of the minimizer is shown to be one that enforces the workload to stay in a specific interval of the form  $[0, \beta]$  with minimal effort. To rigorously define such a strategy we make use of the *Skorokhod map on an interval*. Fix  $\alpha < \beta$ . For any  $\eta \in \mathcal{D}(\mathbb{R}_+, \mathbb{R})$  there exists a unique triplet of functions  $(\chi, \zeta_1, \zeta_2) \in \mathcal{D}(\mathbb{R}_+, \mathbb{R}^3)$  that satisfies the following properties: (i) for every  $t \in \mathbb{R}_+$ ,  $\chi(t) = \eta(t) + \zeta_1(t) - \zeta_2(t)$ ;

(ii)  $\zeta_1$  and  $\zeta_2$  are nondecreasing,  $\zeta_1(0-) = \zeta_2(0-) = 0$ , and

$$\int_0^\infty \mathbf{1}_{(\alpha,\beta]}(\chi(t))d\zeta_1(t) = \int_0^\infty \mathbf{1}_{[\alpha,\beta)}(\chi(t))d\zeta_2(t) = 0.$$

We denote by  $\Gamma_{[\alpha,\beta]}(\eta) = (\Gamma^1_{[\alpha,\beta]}, \Gamma^2_{[\alpha,\beta]}, \Gamma^3_{[\alpha,\beta]})(\eta) = (\chi, \zeta_1, \zeta_2)$ . See [25] for existence and uniqueness of solutions, and continuity and further properties of the map. In particular, we have the following.

**Lemma 3.1** There exists a constant  $c_S > 0$  such that for every t > 0,  $\alpha < \beta$  and  $\omega, \tilde{\omega} \in \mathcal{D}(\mathbb{R}_+, \mathbb{R})$ ,

$$\sup_{s \in [0,t]} \|\Gamma_{[\alpha,\beta]}(\omega)(t) - \Gamma_{[\alpha,\beta]}(\tilde{\omega})(t)\| \le c_S \sup_{s \in [0,t]} |\omega(s) - \tilde{\omega}(s)|$$

**Definition 3.1** Fix  $x_0, \beta \in [0, b]$ . The strategy (Y, R) is called a  $\beta$ -reflecting strategy if for every  $\eta \in C(\mathbb{R}_+, \mathbb{R})$  one has  $(X, Y, R)(\eta) = \Gamma_{[0,\beta]}(\eta)$ .

One can easily verify that any  $\beta$ -reflecting strategy is admissible.

#### 3.2 The HJB equation and the value function

In case that there is no ambiguity (see Remark 2.2), the problem was analyzed by Harrison and Taksar [21]. It is shown there (see Proposition 5.11) that if  $k_{\alpha\beta}$  is a  $\mathcal{C}^1([0,b],\mathbb{R})$  is twice continuously differentiable on  $[\alpha,\beta]$  and satisfies

$$l + k'_{\alpha\beta}(x) = 0, \qquad 0 \le x \le \alpha, \tag{3.1}$$

$$\frac{1}{2}\sigma^2 k_{\alpha\beta}''(x) + mk_{\alpha\beta}'(x) - \varrho k_{\alpha\beta}(x) + h(x) = 0, \qquad \alpha \le x \le \beta,$$
(3.2)

$$r - k'_{\alpha\beta}(x) = 0, \qquad \beta \le x \le b, \tag{3.3}$$

then

$$k_{\alpha\beta}(x) = \mathbb{E}^{\mathbb{P}}\left[\int_{0}^{\infty} e^{-\varrho t} [h(X(t))dt + rdR(t) + ldY(t)]\right], \quad x \in [0,b],$$

with  $(X, Y, R)(t) = \Gamma_{[\alpha,\beta]}(x_0 + m \cdot +\sigma B(\cdot))(t)$ ,  $t \in \mathbb{R}_+$ . The rationale behind this argument is as follows. Consider  $x \in (\beta, b)$ , then in order to keep the process X between  $\alpha$  and  $\beta$ , there is an instantaneous reflection from above that contributes the cost  $r(x - \beta)$ . This explains (3.3). Similar arguments yield (3.1). When no control is taking action, standard arguments imply (3.2). The HJB in this case takes the form

$$\begin{cases} \left[\frac{1}{2}\sigma^2 f''(x) + mf'(x) - \varrho f(x) + h(x)\right] \wedge f'(x) \wedge [r - f'(x)] = 0, & x \in (0, b), \\ f'(0) = 0, & f'(b) = r. \end{cases}$$

However, the uniqueness of a solution of the HJB is not argued in [21] (see the paragraph before Section 7) but rather in [6, Proposition 2.2], using viscosity solutions.

**Remark 3.1** Notice that in [21] there is a cost associated with the reflections from both sides, unlike our case that does not consider a cost component for reflections from below, that is l = 0. It simply follows since in our QCP there is no penalty/reward for idleness. Henceforth, under optimality,  $\alpha = 0$  in (3.1)–(3.2).

Motivated by these results and the structure of the cost function given in (2.33) together with the inf-sup structure of the value function given in (2.34), we consider the following HJB,

$$\begin{cases} \left[ \sup_{p \in \mathbb{R}} \left\{ \frac{1}{2} \sigma^2 f''(x) + (m + \sigma p) f'(x) - \varrho f(x) + h(x) - \frac{1}{2\varepsilon} p^2 \right\} \right] \wedge f'(x) \wedge [r - f'(x)] = 0, \quad x \in (0, b), \\ f'(0) = 0, \quad f'(b) = r. \end{cases}$$

Or equivalently, by substituting the optimal solution of the  $\sup_{p \in \mathbb{R}}$  above,  $p^* = \varepsilon \sigma f'(x)$ ,

$$\begin{cases} [f''(x) + H(x, f(x), f'(x))] \land f'(x) \land [r - f'(x)] = 0, & x \in (0, b), \\ f'(0) = 0, & f'(b) = r, \end{cases}$$
(HJB( $\varepsilon$ ))

where hereafter,

$$H(x, y, z) := \frac{2}{\sigma^2} \left( mz + \frac{1}{2}\sigma^2 \varepsilon z^2 - \varrho y + h(x) \right).$$

Notice that when  $\varepsilon = 0$ , HJB( $\varepsilon$ ) coincides with the HJB given in Harrison and Taksar [21, Equation (1.2)] and by Atar and Shifrin [6, Equation (41)].

The idea is that we may think of the change of measure done by the maximizer as a change of the drift term from m to  $m + \sigma p$ , which costs her  $\frac{1}{2\varepsilon}p^2$ , where p depends on x and is chosen in order to maximize the cost. As can be seen, due to the term  $\frac{1}{2}\sigma^2\varepsilon(f'(x))^2$  in H(x, f(x), f'(x)), the HJB is not linear, a fact that raises some technical difficulties in proving Proposition 3.3 below.

We now state the main theorem of the paper, which is given also for  $\varepsilon = 0$ , see Remark 2.2.

**Theorem 3.1** Fix  $\varepsilon \in [0, \infty)$ . The value function  $V(\cdot; \varepsilon)$  is the unique  $C^2([0, b], \mathbb{R})$  solution of  $HJB(\varepsilon)$ . Moreover, set

$$\beta_{\varepsilon} = \inf \left\{ x \in (0, b] : V'(x; \varepsilon) = r \right\}, \tag{3.4}$$

where  $V'(x;\varepsilon)$  is the derivative of  $V(\cdot;\varepsilon)$  w.r.t. x. Then the  $\beta_{\varepsilon}$ -reflecting strategy is optimal for the minimizer and  $V = V(\cdot;\varepsilon)$  satisfies,

$$\begin{cases} V''(x) + H(x, V(x), V'(x)) = 0, & 0 \le x \le \beta_{\varepsilon}, \\ r - V'(x) = 0, & \beta_{\varepsilon} \le x \le b, \\ V'(0) = 0. \end{cases}$$
(3.5)

From the definition of  $HJB(\varepsilon)$  and the theorem above we get the following corollary, which is given for reference purposes.

**Corollary 3.1** For any  $\varepsilon \in (0, \infty)$ ,  $0 \le V'(\cdot; \varepsilon) \le r$ .

The proof of the theorem is done in several steps and follows from the next three propositions. Before stating them we define a measure  $\mathbb{Q}_f$  associated with a process  $\psi_f$ , which in turn is driven by a function  $f \in \mathcal{C}^1([0, b], \mathbb{R})$  through (2.27). This measure serves us in the sequel, especially in Propositions 3.1 and 3.2 and Theorem 4.1. In the latter, we show that the measure  $\mathbb{Q}_V = \mathbb{Q}_{V(\cdot;\varepsilon)}$  is the optimal strategy of the maximizer, where V is the value function. For any  $f \in \mathcal{C}^1([0, b], \mathbb{R})$  and  $t \in \mathbb{R}_+$  set

$$\psi_f(t) := \operatorname*{arg\,max}_{p \in \mathbb{R}} \left\{ \frac{1}{2} \sigma^2 f''(X(t)) + (m + \sigma p) f'(X(t)) - \varrho f(X(t)) + h(X(t)) - \frac{1}{2\varepsilon} p^2 \right\}$$
(3.6)  
=  $\varepsilon \sigma f'(X(t)).$ 

This is an  $\{\mathcal{F}_t\}$ -progressively measurable process since Y, R, and B are, see Definition 2.2. Let  $\mathbb{Q}_f$  be the measure associated with  $\psi_f$  through (2.27). Note that  $\psi_f$  is bounded since  $f \in \mathcal{C}^1([0, b], \mathbb{R})$  and therefore (2.28) holds trivially.

The proof of the propositions below are deferred to after the proof of Theorem 3.1.

**Proposition 3.1** Fix  $\varepsilon \in (0,\infty)$ . Assume that  $HJB(\varepsilon)$  admits a  $C^2([0,b],\mathbb{R})$  solution f. Then

$$f(x) \le \inf_{(Y,R)\in\mathcal{A}(x)} J(x,Y,R,\mathbb{Q}_f;\varepsilon), \qquad x\in[0,b],$$

and as a consequence  $f \leq V$ .

The next proposition characterizes the solution of the following ordinary differential equation,

$$\begin{cases} k_{\beta}''(x) + H(x, k_{\beta}(x), k_{\beta}'(x)) = 0, & 0 \le x \le \beta, \\ r - k_{\beta}'(x) = 0, & \beta \le x \le b, \\ k_{\beta}'(0) = 0. \end{cases}$$
(3.7)

**Proposition 3.2** Fix  $\varepsilon \in (0, \infty)$ . Assume that there is a function  $k_{\beta} = k_{\beta,\varepsilon} \in C^1([0,b],\mathbb{R}) \cap C^2([a,b] \setminus \{\beta\},\mathbb{R})$  that solves (3.7). Let  $(Y_{\beta}, R_{\beta})$  be a  $\beta$ -reflecting strategy, that is  $(X, Y_{\beta}, R_{\beta})(t) = \Gamma_{[0,\beta]}(x + m \cdot + \sigma B(\cdot))(t)$ ,  $t \in \mathbb{R}_+$ . Then, for every  $x \in [0,b]$ ,

$$k_{\beta}(x) = \sup_{\mathbb{Q}\in\mathcal{Q}(x)} J(x, Y_{\beta}, R_{\beta}, \mathbb{Q}; \varepsilon) = J(x, Y_{\beta}, R_{\beta}, \mathbb{Q}_{k}; \varepsilon),$$
(3.8)

where  $\mathbb{Q}_k = \mathbb{Q}_{k_\beta}$  is the measure associated with  $\psi_k = \psi_{k_\beta}$ , defined in (3.6).

We now claim that for every  $\varepsilon \in (0, \infty)$ , HJB( $\varepsilon$ ) admits a unique smooth solution. Harrison and Taksar provided in [21, page 450] explicit functions from which a smooth solution can be constructed. The construction is provided and the smoothness is claimed to be straightforward yet tedious and therefore omitted, see the paragraph below (6.8) there. The situation is more subtle in our case since HJB( $\varepsilon$ ) is nonlinear and therefore explicit solutions are out of reach. We choose a different path and use the *shooting method* to prove that a smooth solution of HJB( $\varepsilon$ ) uniquely exists and that it solves the free-boundary problem (3.7). In short, the shooting method is used to solve boundary value problems by reducing them to initial value problems; see [33, Section 7.3] for further reading about the method. We take it one step forward and use it in the free-boundary setup with Neumann boundary conditions. Specifically, we consider a set of Cauchy problems on the interval [0, b], indexed<sup>3</sup> by s:

$$\begin{cases} (k^{(s)})''(x) + H(x, k^{(s)}(x), (k^{(s)})'(x)) = 0, & 0 \le x \le b, \\ k^{(s)}(0) = 0, & (k^{(s)})'(0) = 0. \end{cases}$$

We prove that there is a parameter s for which there is  $\beta_{\varepsilon} \in (0, b]$  at which we can smoothly paste the linear functions  $k^{(s)}$  and  $x \mapsto k^{(s)}(\beta_{\varepsilon}) + r(x - \beta_{\varepsilon})$ ; namely,  $(k^{(s)})'(\beta_{\varepsilon}) = r$  and in case that  $\beta_{\varepsilon} < b$  also  $(k^{(s)})''(\beta_{\varepsilon}) = 0$ . This implies a smooth solution to (3.5). From the technical point of view, the proof of Proposition 3.3 is the most demanding in this section.

**Proposition 3.3** For every  $\varepsilon \in (0, \infty)$ ,  $HJB(\varepsilon)$  admits a unique  $C^2([0, b], \mathbb{R})$  solution. Moreover, the solution has the form  $k_{\beta_{\varepsilon}}$  for some parameter  $\beta_{\varepsilon} \in (0, b]$  for which  $(k_{\beta_{\varepsilon}})' < r$  on  $[0, \beta_{\varepsilon})$ , where  $k_{\beta_{\varepsilon}}$  is taken from Proposition 3.2.

**Proof of Theorem 3.1:** Recall that when  $\varepsilon = 0$ , HJB( $\varepsilon$ ) coincides with the HJB given in [21, 6]. Thus, we focus only on positive  $\varepsilon$ 's.

From Proposition 3.3, the HJB( $\varepsilon$ ) admits a unique  $C^2([0, b], \mathbb{R})$  solution that also solves (3.7) for some  $\beta_{\varepsilon} \in (0, b]$ . Denote it by  $k_{\beta_{\varepsilon}}$ . From Proposition 3.2 this is the cost of the  $\beta_{\varepsilon}$ reflecting strategy. Therefore, the value function, which is the infimum over all the strategies satisfies,  $V(\cdot; \varepsilon) \leq k_{\beta_{\varepsilon}}(\cdot)$ . Together with Proposition 3.1,  $V(\cdot; \varepsilon) = k_{\beta_{\varepsilon}}(\cdot)$ . Recalling again that  $k_{\beta_{\varepsilon}}$  is the cost of the  $\beta_{\varepsilon}$ -reflecting strategy, we obtain its optimality. The relation in (3.4) follows now from Proposition 3.3.

The rest of the section is devoted to the proofs of Propositions 3.1–3.3.

<sup>&</sup>lt;sup>3</sup>For technical reasons, in the proof we use a modification of H.

**Proof of Proposition 3.1:** Fix  $f \in C^2([0,b],\mathbb{R})$  and an arbitrary  $(Y,R) \in \mathcal{A}(x)$  with an associated standard Brownian motion B. Set

$$X(t) = x + mt + \sigma B(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+.$$

Recalling (2.29), Itô's lemma implies that for every t > 0 and every  $\mathbb{Q} \in \mathcal{Q}(x)$ ,

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\varrho t}f(X(t))\right] = f(x) + \mathbb{E}^{\mathbb{Q}}\left[\int_{0}^{t} e^{-\varrho s} \left(\frac{1}{2}\sigma^{2}f''(X(s)) + (m+\sigma\psi(s))f'(X(s)) - \varrho f(X(s))\right) ds\right] \\ + \mathbb{E}^{\mathbb{Q}}\left[\int_{0}^{t} e^{-\varrho s}f'(X(s))(dY^{c}(s) - dR^{c}(s))\right] + \mathbb{E}^{\mathbb{Q}}\left[\sum_{0\leq s\leq t} e^{-\varrho s}\Delta f(X)(s)\right],$$
(3.9)

where  $\mathbb{Q}$  and  $\psi$  are related to each other through (2.27)–(2.28). We used the fact that  $B^{\mathbb{Q}}(t) := B(t) - \int_0^t \psi(s) ds, t \in \mathbb{R}_+$  is a  $\mathbb{Q}$  standard Brownian motion. The processes

$$Y^{c}(t) := Y(t) - \sum_{0 \le s \le t} \Delta Y(s), \qquad R^{c}(t) := R(t) - \sum_{0 \le s \le t} \Delta R(s), \qquad t \in \mathbb{R}_{+},$$
(3.10)

are the continuous parts of Y and R, respectively. Consider now (3.9) with  $\psi = \psi_f$  defined in (3.6) and with the measure  $\mathbb{Q} = \mathbb{Q}_f$ . Recalling the definition of  $\psi_f$  and that f solves  $\mathrm{HJB}(\varepsilon)$ , we get that

$$\mathbb{E}^{\mathbb{Q}_f}\left[e^{-\varrho t}f(X(t))\right] \ge f(x) - \mathbb{E}^{\mathbb{Q}_f}\left[\int_0^t e^{-\varrho s}\left(h(X(s))ds + rdR(s) - \frac{1}{2\varepsilon}\psi_f^2(s)ds\right)\right] + \mathbb{E}^{\mathbb{Q}_f}\left[\sum_{0\le s\le t} e^{-\varrho s}(\Delta f(X)(s) + r\Delta R(s))\right].$$
(3.11)

Since  $\Delta X(s) = \Delta Y(s) - \Delta R(s)$ , we get that

$$\Delta f(X)(s) + r\Delta R(s) = f(X(s)) - f(X(s) - \Delta Y(s))$$

$$- [f(X(s) + \Delta R(s) - \Delta Y(s)) - f(X(s) - \Delta Y(s))] + r\Delta R(s)$$

$$= \int_{X(s) - \Delta Y(s)}^{X(s)} f'(u) du + \int_{X(s) - \Delta Y(s)}^{X(s) + \Delta R(s) - \Delta Y(s)} (r - f'(u)) du$$

$$\geq 0.$$
(3.12)

Combining (3.11)–(3.12), we get that

$$\mathbb{E}^{\mathbb{Q}_f}\left[e^{-\varrho t}f(X(t))\right] + \mathbb{E}^{\mathbb{Q}_f}\left[\int_0^t e^{-\varrho s}\left(h(X(s))ds + rdR(s) - \frac{1}{2\varepsilon}\psi_f^2(s)ds\right)\right] \ge f(x).$$

Recalling the definition of the cost function J in (2.33) and noting that the function f is bounded as a continuous function on [0, b], by taking  $t \to \infty$ , we get that

$$\sup_{\mathbb{Q}\in\mathcal{Q}(x)}J(x,Y,R,\mathbb{Q};\varepsilon)\geq J(x,Y,R,\mathbb{Q}_f;\varepsilon)\geq f(x).$$

Since x and (Y, R) are arbitrary, it follows that  $V(\cdot; \varepsilon) \ge f(\cdot)$ .

**Proof of Proposition 3.2:** We split the proof into two cases  $x \in [0, \beta]$  and  $x \in (\beta, b)$ . Fix  $x \in [0, \beta]$  and an arbitrary  $\mathbb{Q} \in \mathcal{Q}(x)$  with an associated process  $\psi$ . Recall the notation  $\psi_k = \psi_{k_\beta}$ . By (3.6), we get

$$\frac{1}{2}\sigma^{2}k_{\beta}''(X(t)) + (m + \sigma\psi(t))k_{\beta}'(X(t)) - \varrho k_{\beta}(X(t)) + h(X(t)) - \frac{1}{2\varepsilon}(\psi(t))^{2} \qquad (3.13)$$

$$\leq \frac{1}{2}\sigma^{2}k_{\beta}''(X(t)) + (m + \sigma\psi_{k}(t))k_{\beta}'(X(t)) - \varrho k_{\beta}(X(t)) + h(X(t)) - \frac{1}{2\varepsilon}(\psi_{k}(t))^{2}$$

$$= \frac{1}{2}\sigma^{2}\left(k_{\beta}''(x) + H(x,k_{\beta}(x),k_{\beta}'(x))\right) = 0.$$

Equations (3.9) and (3.10) are given for general admissible strategies  $(Y, R) \in \mathcal{A}(x)$  and  $\mathbb{Q} \in \mathcal{Q}(x)$ . Thus, they hold here as well. Notice that since  $x \in [0, \beta]$  and  $(Y_{\beta}, R_{\beta})$  is a  $\beta$ -reflecting strategy, the processes  $Y_{\beta}$  and  $R_{\beta}$  have no jumps and so,  $Y_{\beta}^{c} = Y_{\beta}$  and  $R_{\beta}^{c} = R_{\beta}$ . From (3.9), (3.10), and (3.13), one has,

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\varrho t}k_{\beta}(X(t))\right] \leq k_{\beta}(x) - \mathbb{E}^{\mathbb{Q}}\left[\int_{0}^{t} e^{-\varrho s}\left(h(X(s)ds - \frac{1}{2\varepsilon}\psi^{2}(s)ds\right)ds\right] + \mathbb{E}^{\mathbb{Q}}\left[\int_{0}^{t} e^{-\varrho s}k_{\beta}'(X(s))(dY_{\beta}(s) - dR_{\beta}(s))\right].$$
(3.14)

Using now the equalities  $k'_{\beta}(0) = 0$  and  $k'_{\beta}(\beta) = r$ , driven from (3.7), we get

$$\mathbb{E}^{\mathbb{Q}}\left[e^{-\varrho t}k_{\beta}(X(t))\right] \le k_{\beta}(x) - \mathbb{E}^{\mathbb{Q}}\left[\int_{0}^{t} e^{-\varrho s}\left(h(X(s))ds + rdR_{\beta}(s) - \frac{1}{2\varepsilon}\psi^{2}(s)ds\right)\right].$$
 (3.15)

Recalling the definition of J and noting that the function  $k_{\beta}$  is bounded as a continuous function on [0, b], by taking  $t \to \infty$ , we get that

$$k_{\beta}(x) \ge \sup_{\mathbb{Q}\in\mathcal{Q}(x)} J(x, Y_{\beta}, R_{\beta}, \mathbb{Q}; \varepsilon).$$
(3.16)

Notice that all the inequalities in (3.13)–(3.15) hold with equality in case that  $\psi = \psi_k$  and  $\mathbb{Q} = \mathbb{Q}_k$ . Then, together with (3.16), we get that (3.8) holds.

Consider now the case that  $\beta < b$  and  $x \in (\beta, b]$ . From (3.7),

$$k_{\beta}(x) = r(x - \beta) + k_{\beta}(\beta)$$

Since the strategy  $(Y_{\beta}, R_{\beta})$  starts with an instantaneous rejection of  $x - \beta_0$ , there is an immediate cost of  $r(x - \beta)$  and hence for any  $\mathbb{Q} \in \mathcal{Q}(x)$ ,

$$J(x, Y_{\beta}, R_{\beta}, \mathbb{Q}; \varepsilon) = r(x - \beta) + J(\beta, Y_{\beta}, R_{\beta}, \mathbb{Q}; \varepsilon).$$

Recall that we just showed that

$$k_{\beta}(\beta) = \sup_{\mathbb{Q}\in\mathcal{Q}(x)} J(\beta, Y_{\beta}, R_{\beta}, \mathbb{Q}; \varepsilon) = J(\beta, Y_{\beta}, R_{\beta}, \mathbb{Q}_{f}; \varepsilon).$$

From the last three equalities we get that (3.8) holds for  $x \in (\beta, b]$  as well.

Before getting to the proof of Proposition 3.2 we provide an auxiliary parametrized ordinary differential equation in order to solve the free-boundary one. Fix a parameter  $s \in \mathbb{R}$  and consider the following Cauchy problem

$$\begin{cases} (k^{(s)})''(x) + H_F(x, k^{(s)}(x), (k^{(s)})'(x)) = 0, & x \in [0, b], \\ (k^{(s)})'(0) = 0, & k^{(s)}(0) = s, \end{cases}$$
(3.17)

where

$$H_F(x, y, z) := H(x, y, F(z))$$
 (3.18)

and F is a  $\mathcal{C}^1(\mathbb{R},\mathbb{R})$  function that satisfies the following properties: F(z) = z on [-r,r],  $|F| \leq 2r$ , and  $|F'| \leq 1$  and therefore Lipschitz continuous. For example,

$$F(z) = \begin{cases} -3/2r, & z < -2r, \\ (1/2)r + 2z + z^2/(2r), & -2r \le z < -r, \\ z, & -r \le z \le r, \\ -(1/2)r + 2z - z^2/(2r), & r \le z < 2r, \\ 3/2r, & 2r \le z. \end{cases}$$

Because the function F and its derivative are bounded, and since the function h is Lipschitz (see the paragraph below (2.31)),  $H_F$  is uniformly Lipschitz. Namely, there is a constant  $c_L$  such that for every  $(x, y, z), (x', y', z') \in [0, b] \times \mathbb{R} \times \mathbb{R}$ , one has

$$|H_F(x, y, z) - H_F(x', y', z')| \le L(|x - x'| + |y - y'| + |z - z'|).$$
(3.19)

From [29, Section 0.3.1], (3.17) admits a unique  $C^2([0, b], \mathbb{R})$  solution.

 $\operatorname{Set}$ 

$$\beta^{(s)} := \inf\{x \in (0, b] : (k^{(s)})'(x) \ge r\} \land b,$$
(3.20)

where we use the convention that  $\inf \emptyset = \infty$ . The smoothness of  $k^{(s)}$  implies that

if 
$$\beta^{(s)} < b$$
 then  $(k^{(s)})'(\beta^{(s)}) = r.$  (3.21)

The following lemma provides some continuity properties that serve us in the proof of Proposition 3.3.

**Lemma 3.2** The function  $s \mapsto (k^{(s)}, (k^{(s)})', (k^{(s)})'')$  is continuous in the uniform norm topology taken on the interval [0, b]. Moreover, the mapping  $s \mapsto \beta^{(s)}$  is continuous for every s for which either  $\beta^{(s)} = b$  or the following two conditions hold  $\beta^{(s)} < b$  and  $(k^{(s)})''(\beta^{(s)}) \neq 0$ . In these cases we conclude that the mapping  $s \mapsto (k^{(s)}(\beta^{(s)}), (k^{(s)})'(\beta^{(s)}), (k^{(s)})''(\beta^{(s)}))$  is also continuous.

**Proof:** Fix  $s \in \mathbb{R}$ . One can easily verify that the conditions of [30, Theorem 23] are satisfied for  $H_F$ . Therefore, for any  $\delta_1 \in \mathbb{R}$ 

$$\sup_{x \in [0,b]} \left| k^{(s+\delta_1)}(x) - k^{(s)}(x) \right| \le |\delta_1|, \tag{3.22}$$

and the continuity of  $s \mapsto k^{(s)}$  is established.

We now turn to showing that  $s \mapsto (k^{(s)})'$  is continuous. From (3.17) it follows that for every  $x \in [0, b]$ 

$$(k^{(s)})'(x) = 0 - \int_0^x H_F(y, k^{(s)}(y), (k^{(s)})'(y)) dy,$$
  
$$(k^{(s+\delta_1)})'(x) = 0 - \int_0^x H_F(y, k^{(s+\delta_1)}(y), (k^{(s+\delta_1)})'(y)) dy$$

Set  $g_{s,\delta_1}(x) := \left| (k^{(s+\delta_1)})'(x) - (k^{(s)})'(x) \right|, x \in [0, b]$ . From (3.19), there exists a constant L > 0 independent of s and  $\delta_1$  such that

$$g_{s,\delta_1}(x) \le L \int_0^x \left( \left| k^{(s+\delta_1)}(y) - k^{(s)}(y) \right| + g_{s,\delta_1}(y) \right) dy \le Lb|\delta_1| + L \int_0^x g_{s,\delta_1}(y) dy$$

where the second inequality follows by (3.22). Now, Grönwall's inequality implies that  $\sup_{x \in [0,b]} g_{s,\delta_1}(x) \leq |\delta_1| Lbe^{Lb}$ , and the continuity of  $s \mapsto (k^{(s)})'$  is established.

Finally, the continuity of  $s \mapsto (k^{(s)})''$  follows by the relation  $(k^{(s)})''(x) = -H_F(x, k^{(s)}(x), (k^{(s)})'(x))$ , the continuity of  $s \mapsto (k^{(s)}, (k^{(s)})')$ , and the Lipschitz continuity of  $H_F$  stated in (3.19).

The rest of the proof is dedicated to the continuity of the function  $s \mapsto \beta^{(s)}$  under the conditions mentioned in the lemma. To this end, we fix  $s \in \mathbb{R}$  and show that if  $\beta^{(s)} = b$  or if  $\beta^{(s)} < b$  and  $(k^{(s)})'(\beta^{(s)}) < r$ , then

$$\limsup_{\delta \to 0} \beta^{(s+\delta)} \le \beta^{(s)} \le \liminf_{\delta \to 0} \beta^{(s+\delta)}.$$
(3.23)

We start with the first inequality. If  $\beta^{(s)} = b$  then it is obvious, since all the  $\beta^{(u)}$ 's are less or equal to b. If  $\beta^{(s)} < b$  and  $(k^{(s)})''(\beta^{(s)}) \neq 0$ , we necessarily have  $(k^{(s)})''(\beta^{(s)}) > 0$ . Otherwise,  $(k^{(s)})''(\beta^{(s)}) < 0$  and from (3.21), we get that  $(k^{(s)})'(\beta^{(s)} - \nu) > r$  for sufficiently small  $\nu > 0$ , a contradiction to the definition of  $\beta^{(s)}$ .

Using now  $(k^{(s)})''(\beta^{(s)}) > 0$ , we get that for sufficiently small  $\nu > 0$ ,  $(k^{(s)})'(\beta^{(s)} + \nu) > r$ . By the continuity of  $s \mapsto (k^{(s)})'$ , we get that for every  $\delta_2$  with sufficiently small absolute value, one has  $(k^{(s+\delta_2)})'(\beta^{(s)} + \nu) > r$ . Therefore,  $\beta^{(s+\delta_2)} < \beta^{(s)} + \nu$  and  $\limsup_{\delta \to 0} \beta^{(s+\delta)} \leq \beta^{(s)} + \nu$ . Since  $\nu > 0$  can be arbitrary small we get the first inequality on (3.23).

We now turn to proving the second inequality in (3.23). Set  $\gamma_1 > 0$  and  $\hat{\beta}^{(s)} := \liminf_{\delta \to 0} \beta^{(s+\delta)}$ . Consider a sequence  $\{\delta_j\}_j \to 0$  such that  $\{\beta^{(s+\delta_j)}\}_j \to \hat{\beta}^{(s)}$  and  $\beta^{(s+\delta_j)} < b$  for every j. If such a subsequence does not exist it means that for every  $\delta$  with sufficiently small absolute value,  $\beta^{(s+\delta)} = b$  and the claim is trivial. Since  $k^{(s)} \in \mathcal{C}^2([0, b], \mathbb{R})$ , we get that for sufficiently large j,

$$\left| (k^{(s)})'(\beta^{(s+\delta_j)}) - (k^{(s)})'(\hat{\beta}^{(s)}) \right| < \gamma_1.$$
(3.24)

Together with the continuity of  $s \mapsto (k^{(s)})'$ , we get that for sufficiently large j, one has

$$\left| (k^{(s+\delta_j)})'(\beta^{(s+\delta_j)}) - (k^{(s)})'(\beta^{(s+\delta_j)}) \right| < \gamma_1.$$
(3.25)

Recall that  $\beta^{(s+\delta_j)} < b$ . Thus,  $(k^{(s+\delta_j)})'(\beta^{(s+\delta_j)}) = r$ . From (3.24)–(3.25) we get that

$$\left| (k^{(s)})'(\hat{\beta}^{(s)}) - r \right| < 2\gamma_1$$

Since  $\gamma_1 > 0$  can be arbitrary small, we get that  $(k^{(s)})'(\hat{\beta}^{(s)}) = r$  and therefore,  $\beta^{(s)} \leq \hat{\beta}^{(s)}$ .

**Proof of Proposition 3.3:** Fix  $\varepsilon \in (0, \infty)$ . We start the proof by arguing uniqueness of the solution and then we move on to showing that a solution indeed exists.

Uniqueness: We first argue that there cannot be two different  $C^2([0, b], \mathbb{R})$  solutions of  $HJB(\varepsilon)$ . Notice that the non-linearity of our differential equation prevents us from using the same proof given in [6, Proposition 2.2]. Arguing by contradiction, assume that there are two  $C^2([0, b], \mathbb{R})$  solutions  $f_1 \neq f_2$ . Without loss of generality, assume that there exists  $x \in [0, b]$  such that  $f_1(x) > f_2(x)$ . Take  $y_0 \in \arg \max_{x \in [0,b]} \{f_1(x) - f_2(x)\}$ . Clearly,  $f_1(y_0) > f_2(y_0)$ . Also,  $f'_1(0) = f'_2(0) = 0$  and  $f'_1(b) = f'_2(b) = r$  since they are both solutions to  $HJB(\varepsilon)$ . Now, if  $y_0 \in (0, b)$ , then first order condition for  $f_1 - f_2$  implies that  $f'_1(y_0) = f'_2(y_0)$ , and therefore it holds whether  $y_0$  is an internal point or an endpoint of [0, b].

By the structure of  $HJB(\varepsilon)$  applied to  $f_1$ , we get

$$f_1''(y_0) + H(y_0, f_1(y_0), f_1'(y_0)) \ge 0.$$
(3.26)

The same inequality holds for  $f_2$ . Now, if

$$f_2''(y_0) + H(y_0, f_2(y_0), f_2'(y_0)) = 0, (3.27)$$

then by subtracting (3.27) form (3.26), and recalling that  $f_1(y_0) > f_2(y_0)$  and  $f'_1(y_0) = f'_2(y_0)$ , we get that  $f''_1(y_0) > f''_2(y_0)$ . Therefore, in a small neighborhood of  $y_0$  (in case that  $y_0 = 0$  or  $y_0 = b$ , we take right- and left-neighborhoods, respectively),  $f_1(x) - f_2(x) > f_1(y_0) - f_2(y_0)$ , which contradicts the definition of  $y_0$ .

Consider now the case

$$f_2''(y_0) + H(y_0, f_2(y_0), f_2'(y_0)) > 0.$$

Since  $f_2$  is a  $\mathcal{C}^2([0,b],\mathbb{R})$  solution of  $HJB(\varepsilon)$ , it follows that  $0 \leq y_1 < y_0 < y_2 \leq b$ , where

$$y_1 := \sup \left\{ x \in [0, y_0] : f_2''(x) + H(x, f_2(x), f_2'(x)) = 0 \right\} \lor 0,$$
  
$$y_2 := \inf \left\{ x \in [y_0, b] : f_2''(x) + H(x, f_2(x), f_2'(x)) = 0 \right\} \land b.$$

We use the convention that  $\sup \emptyset = -\inf \emptyset = -\infty$ . Notice that on the interval  $(y_1, y_2)$ , which includes  $y_0, f_2''(x) + H(x, f_2(x), f_2'(x)) > 0$ . Since  $f_2$  solves  $\operatorname{HJB}(\varepsilon)$ , it follows that  $f_2' = 0$  on  $[y_1, y_2]$  or  $f_2' = r$  on  $[y_1, y_2]$ . In the former case, we get  $f_2(y_0) = f_2(y_2)$  and since  $f_2'(b) = r$ ,  $y_2 < b$ . Since  $f_1' \ge 0$  as a solution of  $\operatorname{HJB}(\varepsilon)$ , we get that  $f_1(y_2) - f_2(y_2) \ge f_1(y_0) - f_2(y_0)$ , which by the definition of  $y_0$  is in fact an equality. Therefore,  $y_2 \in \arg \max_{x \in [0,b]} \{f_1(x) - f_2(x)\}$ .

Since  $y_2 < b$  it follows by the smoothness of  $f_2$  and the definition of  $y_2$  that (3.26) and (3.27) hold with  $y_2$  replacing  $y_0$ . Repeating the same argument proceeding (3.27), with  $y_2$  replacing  $y_0$  yields a contradiction. Now if,  $f'_2 = r$  on  $[y_1, y_2]$ , a similar argument hold by using  $f'_1 \leq r$ and the point  $y_1$  instead of  $y_2$ .

**Existence:** We now construct a  $C^2([0, b], \mathbb{R})$  solution of (3.7) with some  $\beta_{\varepsilon} \in (0, b]$  that also solves  $HJB(\varepsilon)$ . The solution is shown to satisfy also  $k'_{\beta_{\varepsilon}} \leq r$ . Repeating the same arguments given in [21, Proposition 6.1], we get that  $k''_{\beta_{\varepsilon}}(x) + H(x, k_{\beta_{\varepsilon}}(x), k'_{\beta_{\varepsilon}}(x)) \geq 0$  on  $[\beta, b]$  and therefore,  $k_{\beta_{\varepsilon}}$  satisfies  $HJB(\varepsilon)$  and we are done.

To solve (3.7), we hereby consider the Cauchy problem given in (3.17). Since eventually we find a solution to (3.17) that satisfies  $0 \leq (k^{(s)})' \leq r$ , it also solves the same ordinary differential equation from (3.17) with H replacing  $H_F$ . The rest of the proof is performed in two steps. First, we prove the existence of  $s^* \in \mathbb{R}$  for which the parameter  $\beta^{(s^*)} \in (0, b]$  from (3.20) satisfies the following

$$(k^{(s^*)})'(x) < r \text{ on } [0, \beta^{(s^*)}),$$
(3.28)

$$(k^{(s^*)})'(\beta^{(s^*)}) = r, (3.29)$$

and in case that  $\beta^{(s^*)} < b$ , also

$$(k^{(s^*)})''(\beta^{(s^*)}) = 0. (3.30)$$

In the second step we show that  $(k^{(s)})' \ge 0$  on  $[0, \beta^{(s^*)}]$ . Therefore, the function

$$k(x) = \begin{cases} k^{(s^*)}(x), & 0 \le x < \beta^{(s^*)}, \\ k^{(s^*)}(\beta^{(s^*)}) + r(x - \beta^{(s^*)}), & \beta^{(s^*)} \le x \le b, \end{cases}$$

satisfies (3.7), and the proof is done by setting

$$\beta_{\varepsilon} := \beta^{(s^*)}$$
 and  $k_{\beta_{\varepsilon}} := k^{(s^*)}$ . (3.31)

As a conclusion, we get that

$$s^* = k_{\beta_{\varepsilon}}(0) = V(0;\varepsilon). \tag{3.32}$$

**Step 1:** Notice that (3.28) holds trivially for every  $s^*$  by the definition of  $\beta^{(s)}$ . Therefore, we only need to check that (3.29) and (3.30) hold.

The first observation is that for sufficiently large s,  $\beta^{(s)} < b$  and  $(k^{(s)})''(\beta^{(s)}) > 0$ . Also, for sufficiently small s,  $\beta^{(s_0)} = b$  and  $(k^{(s_0)})'(b) < r$ . We prove the first part, the second one follows by the same lines and is therefore omitted. Fix  $s_1 > (2|m|r+2\sigma^2\varepsilon r^2+h(b)+r\sigma^2/b)/\rho$ . Now, the function  $k^{(s_1)}$  is nondecreasing on [0, b]. Indeed, arguing by contradiction, assume that there is  $x \in (0, b)$  such that  $(k^{(s_1)})'(x) < 0$ , then consider  $y_3 := \inf\{x \in [0, b] : (k^{(s_1)})'(x) < 0\}$ . The smoothness of  $k^{(s_1)}$  on [0, b] and the initial condition  $(k^{(s_1)})'(0) = 0$  imply that  $(k^{(s_1)})'(y_3) = 0$ . So,

$$\frac{1}{2}\sigma^2(k^{(s_1)})''(y_3) = \rho k^{(s_1)}(y_3) - h(y_3) \ge \rho k^{(s_1)}(0) - h(b) = \rho s_1 - h(b) > 0.$$
(3.33)

The first inequality follows since  $k^{(s_1)}$  is nondecreasing on  $[0, y_3]$ . This is a contradiction to the definition of  $y_3$ , since for every sufficiently small  $\delta > 0$  we have  $(k^{(s_1)})'(y_3 + \delta) > 0$ . Therefore,  $(k^{(s_1)})' \ge 0$  on [0, b]. Combining it with (3.17),  $|F| \le 2r$ ,  $k^{(s_1)}(0) = s_1$ , the choice of  $s_1$ , and that h is increasing (see the paragraph below (2.31)), we get that for every  $x \in [0, b]$ ,

$$(k^{(s_1)})''(x) = -\frac{2}{\sigma^2} \left( mF((k^{(s_1)})'(x)) + \frac{1}{2}\sigma^2 \varepsilon F^2((k^{(s_1)})'(x)) - \varrho k^{(s_1)}(x) + h(x) \right)$$
  
>  $\frac{2}{\sigma^2} \left( \varrho k^{(s_1)}(0) - 2|m|r - \frac{1}{2}\sigma^2 r^2 \varepsilon - h(b) \right) = \frac{2r}{b}.$ 

In particular  $(k^{(s_1)})''(\beta^{(s_1)}) > 0$ . Also  $(k^{(s_1)})'(b/2) = \int_0^{b/2} (k^{(s_1)})''(u) du \ge r$  and therefore,  $\beta^{(s_1)} \le b/2 < b$ .

By the choices of  $s_0$  and  $s_1$ , the following infimum is attained. Set

$$s^* := \inf\{s \in (-s_0, s_1) : \forall u \in (s, s_1), \ \beta^{(u)} < b \text{ and } (k^{(u)})''(\beta^{(u)}) > 0\}.$$

In case that  $\beta^{(s^*)} = b$ , then by the definition of  $s^*$ , for every  $u \in (s^*, s_1)$ ,  $\beta^{(u)} < b$  and therefore, from (3.21) we have  $(k^{(u)})'(\beta^{(u)}) = r$ . Now, the continuity of the mapping  $s \mapsto (k^{(s)})'(\beta^{(s)})$  at  $s = s^*$  in this case, provided in Lemma 3.2 implies that  $(k^{(s^*)})'(\beta^{(s^*)}) = r$  and (3.29) holds in this case.

In case that  $\beta^{(s^*)} < b$  then (3.21) implies that (3.29) holds. We now claim that (3.30) holds, that is,  $(k^{(s^*)})''(\beta^{(s^*)}) = 0$ . Otherwise, if  $(k^{(s^*)})''(\beta^{(s^*)}) < 0$ , then by the continuity of the mapping  $s \mapsto (\beta^{(s)}, (k^{(s)})''(\beta^{(s)}))$  in this case (see Lemma 3.2), we get that for every sufficiently small  $\nu > 0$ ,  $\beta^{(s^*+\nu)} < b$  and  $(k^{(s^*+\nu)})''(\beta^{(s^*+\nu)}) < 0$ , which contradicts the definition of  $s^*$ . If however,  $(k^{(s^*)})''(\beta^{(s^*)}) > 0$  then Lemma 3.2 again implies that for every sufficiently small  $\nu > 0$ , one has  $\beta^{(s^*-\nu)} < b$  and  $(k^{(s^*-\nu)})''(\beta^{(s^*-\nu)}) > 0$ , in contradiction to the definition of  $s^*$ . Therefore,  $(k^{(s^*)})''(\beta^{(s^*)}) = 0$  and (3.30) holds.

Step 2: In this step we show that

$$(k^{(s^*)})'(x) \ge 0, \qquad x \in [0, \beta^{(s^*)}].$$
 (3.34)

Arguing by contradiction, assume that there is  $y_5 \in (0, \beta^{(s^*)})$  with  $(k^{(s^*)})'(y_5) < 0$ . We omitted the endpoints where the derivatives are 0 and r. The smoothness of the function  $k^{(s^*)}$  together with  $(k^{(s^*)})'(0) = 0$  and  $(k^{(s^*)})'(\beta^{(s^*)}) = r$  (from the previous step) imply that the following supremum and infimum are attained

$$y_4 := \inf\{x \in [0, y_5) : \forall y \in (x, y_5), \ (k^{(s^*)})'(y) \le 0\},\$$
  
$$y_6 := \sup\{x \in (y_5, \beta^{(s^*)}) : \forall y \in (x, y_5), \ (k^{(s^*)})'(y) \le 0\}$$

Also,  $(k^{(s^*)})'(y_4) = (k^{(s^*)})'(y_6) = 0$  and  $(k^{(s^*)})''(y_4) \le 0 \le (k^{(s^*)})''(y_6)$ . Substituting these relations in (3.17), we obtain,

$$\varrho k^{(s^*)}(y_4) - h(y_4) \le 0 \le \varrho k^{(s^*)}(y_6) - h(y_6).$$
(3.35)

Since  $(k^{(s^*)})'(x) \leq 0$  on  $[y_4, y_6]$ , and by the smoothness of  $k^{(s^*)}$ , there is a subinterval on which  $(k^{(s^*)})'(x) < 0$ , one has  $k^{(s^*)}(y_4) > k^{(s^*)}(y_6)$ . Recall that the function h is increasing (see the paragraph below (2.31)) and therefore,  $h(y_4) < h(y_6)$ . The last two inequalities contradict (3.35) and therefore, (3.34) holds.

# 4 Optimal strategies and equilibria in the games

In Theorem 3.1 we claimed that the minimizer has an optimal strategy, which is a  $\beta_{\varepsilon}$ -reflecting strategy with  $\beta_{\varepsilon}$  given in (3.4). However, we did not argue uniqueness in the sense that for every  $\beta \neq \beta_{\varepsilon}$ , the  $\beta$ -reflecting strategy is strictly sub-optimal. Neither did Harrison and Taksar [21] nor Atar and Shifrin [6]. Hence, in the following discussion, which evolves around this issue, we allow  $\varepsilon = 0$  and consider  $\varepsilon \in [0, \infty)$ .

Recall the definition of  $\beta_{\varepsilon}$  from (3.4). Define now,

$$\hat{\beta}_{\varepsilon} := \sup \left\{ x \in (0, b] : \forall y \le x, \ V''(y; \varepsilon) + H(y, V(y; \varepsilon), V'(y; \varepsilon)) = 0 \right\}.$$

$$(4.1)$$

From (3.5) it follows that  $\hat{\beta}_{\varepsilon} \geq \beta_{\varepsilon}$ . Since on the interval  $[\beta_{\varepsilon}, \hat{\beta}_{\varepsilon}]$  both of the conditions

$$V''(x;\varepsilon) + H(x, V(y;\varepsilon), V'(x;\varepsilon)) = 0 \quad \text{and} \quad V'(x;\varepsilon) = r \quad (4.2)$$

hold, it follows that V solves (3.7) for any  $\beta \in [\beta_{\varepsilon}, \hat{\beta}_{\varepsilon}]$ . Proposition 3.2 implies that every such  $\beta$ -reflecting strategy is optimal. Thus, the non-uniqueness of the optimal reflecting strategy is equivalent to the existence of non-degenerate interval  $[\beta_{\varepsilon}, \hat{\beta}_{\varepsilon}]$ , on which the equations in (4.1) hold. Combining both of them, we get that  $V(x; \varepsilon) = (mr + \sigma^2 r^2 \varepsilon/2 + h(x))/\rho$ . Using again  $V'(x; \varepsilon) = r$ , we get that  $h'(x) = \rho r$ . Recall that h is piecewise linear with slopes in the set  $\{h_1\mu_1, \ldots, h_I\mu_I\}$  (see the paragraph below (2.31)). Thus, if  $\rho r$  does not belong to this set, we get uniqueness of the optimal reflecting strategy. Another sufficient condition would be that m is sufficiently negative such that  $(mr + \sigma^2 r^2 \varepsilon/2 + h(b))/\rho \leq 0$ , since the value function cannot be non-positive. The arguments above are summarized in the following proposition.

**Proposition 4.1** Fix  $\varepsilon \in [0, \infty)$ . A  $\beta$ -reflecting strategy is optimal if and only if  $\beta \in [\beta_{\varepsilon}, \hat{\beta}_{\varepsilon}]$ . If

for every 
$$i \in [I]$$
,  $\hat{h}_i \mu_i \neq \varrho r$  or  $mr + \frac{1}{2}\sigma^2 \varepsilon r^2 + h(b) \le 0$ , (4.3)

then  $\beta_{\varepsilon} = \hat{\beta}_{\varepsilon}$  and there is a unique optimal reflecting strategy.

Although there might be weaker assumptions under which uniqueness hold, we do not aim to find them, and remain contented with the not so restrictive conditions given in (4.3).

We already know that the minimizer has an optimal reflecting strategy. We now show that for every  $\hat{\varepsilon} \in (0, \infty)^I$  and  $\varepsilon$  given through (2.25), both the RSDG and the MSDG have equilibria. Specifically, each player has an optimal strategy that is good against any strategy of the other player. Recall that Proposition 2.1 connects between the costs in the games. Fix  $\hat{x}_0 \in \mathcal{X}, \ \hat{\varepsilon} \in (0, \infty)^I, \ \bar{\beta}_{\varepsilon} \in [\beta_{\varepsilon}, \hat{\beta}_{\varepsilon}]$ , and an *I*-dimensional standard Brownian motion  $\hat{B}$ . Set  $B = \sigma^{-1}\theta\hat{\sigma} \cdot \hat{B}$  and recall (2.24) and (2.25). Let  $(Y_{\bar{\beta}_{\varepsilon}}, R_{\bar{\beta}_{\varepsilon}})$  be an optimal reflecting strategy for the minimizer. Also, let  $\mathbb{Q}_V = \mathbb{Q}_{V(\cdot;\varepsilon)}$  be the measure driven by  $\psi_V = \psi_{V(\cdot;\varepsilon)}$  (see (3.6)). Recall the definition of the function  $\gamma$  from (2.36) and the relations given in (2.40)–(2.41). Set the strategy  $(\hat{Y}_{\bar{\beta}_{\varepsilon}}, \hat{R}_{\bar{\beta}_{\varepsilon}})$  by

$$\hat{R}_{\bar{\beta}_{\varepsilon}}(t) := R_{\bar{\beta}_{\varepsilon}}(t)\mu_{i^{*}}e_{i^{*}} \quad \text{and} \quad \hat{Y}_{\bar{\beta}_{\varepsilon}}(t) := \hat{X}_{\bar{\beta}_{\varepsilon}}(t) - \hat{x}_{0} - \hat{m}t - \hat{\sigma}\hat{B}(t) + \hat{R}_{\bar{\beta}_{\varepsilon}}(t), \quad t \in \mathbb{R}_{+},$$

where  $i^*$  is given in (2.32) and for any  $t \in \mathbb{R}_+$ ,

$$\begin{split} \hat{X}_{\bar{\beta}_{\varepsilon}}(t) &:= \gamma(X_{\bar{\beta}_{\varepsilon}}(t)), \\ X_{\bar{\beta}_{\varepsilon}}(t) &= x_0 + mt + \sigma B(t) + Y_{\bar{\beta}_{\varepsilon}}(t) - R_{\bar{\beta}_{\varepsilon}}(t). \end{split}$$

Moreover, let  $\hat{\mathbb{Q}}_V$  be the measure associated with  $\hat{\psi}_V(t) = (\hat{\psi}_{V,1}(t), \dots, \hat{\psi}_{V,I}(t))$ , given by

$$\hat{\psi}_{V,i}(t) := \frac{\sigma \psi_V(t)(\theta \hat{\sigma})_i \hat{\varepsilon}_i}{\sum_{j=1}^I (\theta \hat{\sigma})_j^2 \hat{\varepsilon}_j}, \qquad i \in [I], \quad t \in \mathbb{R}_+.$$

$$(4.4)$$

We now claim that the mentioned strategies form equilibria in the games. Using the  $\beta_{\varepsilon}$ -reflecting strategy in the RSDG it follows that the associated workload dynamics  $X_{\bar{\beta}_{\varepsilon}}$  is a reflected diffusion that moves continuously along the interval  $[0, \bar{\beta}_{\varepsilon}]$ . The function  $\gamma$  continuously maps the workload process to the bold curve from Figure 1. If for example  $m = \bar{\beta}_{\varepsilon}$  in figure 1, then the process  $\hat{X}_{\beta_{\varepsilon}}$  moves continuously along the bold curve between the points J and M, where it is reflected at these two points.

The next theorem generalizes Corollary 2.1 and Theorem 3.1 and provides equilibria in the games.

**Theorem 4.1** Fix  $\varepsilon \in (0,\infty)$ . Using the notation above, the triplets  $(Y_{\bar{\beta}_{\varepsilon}}, R_{\bar{\beta}_{\varepsilon}}, \mathbb{Q}_V)$  and  $(\hat{Y}_{\bar{\beta}_{\varepsilon}}, \hat{R}_{\bar{\beta}_{\varepsilon}}, \hat{\mathbb{Q}}_V)$  form equilibria in the RSDG and the MSDG, respectively. That is,

$$\begin{split} V(x_0;\varepsilon) &= \sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} J(x_0, Y_{\bar{\beta}_{\varepsilon}}, R_{\bar{\beta}_{\varepsilon}}, \mathbb{Q}; \varepsilon) = J(x_0, Y_{\bar{\beta}_{\varepsilon}}, R_{\bar{\beta}_{\varepsilon}}, \mathbb{Q}_V; \varepsilon) = \inf_{(Y,R)\in\mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_V; \varepsilon) \\ &= \hat{V}(\hat{x}_0; \hat{\varepsilon}) = \sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(x_0, \hat{Y}_{\bar{\beta}_{\varepsilon}}, \hat{R}_{\bar{\beta}_{\varepsilon}}, \hat{\mathbb{Q}}; \hat{\varepsilon}) = \hat{J}(\hat{x}_0, \hat{Y}_{\bar{\beta}_{\varepsilon}}, \hat{R}_{\bar{\beta}_{\varepsilon}}, \hat{\mathbb{Q}}_V; \hat{\varepsilon}) = \inf_{(\hat{Y}, \hat{R})\in\hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0, \hat{Y}, \hat{R}, \hat{\mathbb{Q}}_V; \hat{\varepsilon}) \end{split}$$

**Remark 4.1** Notice that the optimal policies for the minimizer in the MSDG has the same structure as the optimal policy in the BCP given in [6]. The only difference emerge from the cut-off point  $\beta_{\varepsilon}$ , which affects the point of reflection on the curve  $\gamma$ . Such a result is not obvious due to the non-stationarity structure of the problem caused by the existence of the maximizer player. Furthermore, the structures of the equilibria share similarities with the optimal policies in the differential game given in [2]. More specifically, [2, Theorem 3.3] states that the minimizer's optimal strategy in the one-dimensional deterministic differential game is a reflecting strategy (called there a 'barrier strategy'). Under optimality, the maximal player in the same game uses a drift change that is illustrated more compactly in [3, Section 3, (30)– (31)]. The relationship between the multidimensional and the one-dimensional games is given in [2, Appendix A].

**Proof of Theorem 4.1:** In this proof we do not check admissibility. It can be verified easily the same way as in Proposition 2.1 using Corollary 3.1. The first two equalities follow from the optimality of  $(Y_{\bar{\beta}_{\varepsilon}}, R_{\bar{\beta}_{\varepsilon}})$  and from (3.8). The third equality follows since on the one hand, by the definition of V,

$$V(x_0;\varepsilon) \ge \inf_{(Y,R)\in\mathcal{A}(x_0)} J(x_0,Y,R,\mathbb{Q}_V;\varepsilon)$$

and on the other hand, the reversed inequality follows from Proposition 3.1 since by Theorem 3.1,  $V \in C^2([0,b],\mathbb{R})$ . The forth equality follows by Corollary 2.1. For the fifth and sixth equalities notice that  $\hat{V}(\hat{x}_0;\hat{\varepsilon}) = V(x_0;\varepsilon)$  and that

$$\hat{V}(\hat{x}_0;\hat{\varepsilon}) \leq \sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(\hat{x}_0,\hat{Y}_{\bar{\beta}_{\varepsilon}},\hat{R}_{\bar{\beta}_{\varepsilon}},\hat{\mathbb{Q}};\hat{\varepsilon}) \leq \sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} J(x_0,Y_{\bar{\beta}_{\varepsilon}},R_{\bar{\beta}_{\varepsilon}},\mathbb{Q};\varepsilon) = V(x_0;\varepsilon),$$

where the second inequality follows by (2.43). For the last equality, notice that we already established

$$\hat{V}(\hat{x}_0;\hat{\varepsilon}) = \sup_{\hat{\mathbb{Q}}\in\hat{\mathcal{Q}}(\hat{x}_0)} \hat{J}(x_0,\hat{Y}_{\bar{\beta}_{\varepsilon}},\hat{R}_{\bar{\beta}_{\varepsilon}},\hat{\mathbb{Q}};\hat{\varepsilon}) \ge \hat{J}(\hat{x}_0,\hat{Y}_{\bar{\beta}_{\varepsilon}},\hat{R}_{\bar{\beta}_{\varepsilon}},\hat{\mathbb{Q}}_V;\hat{\varepsilon}) \ge \inf_{(\hat{Y},\hat{R})\in\hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0,\hat{Y},\hat{R},\hat{\mathbb{Q}}_V;\hat{\varepsilon}).$$

Therefore, it is sufficient to show that  $\inf_{(\hat{Y},\hat{R})\in\hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0,\hat{Y},\hat{R},\hat{\mathbb{Q}}_V;\hat{\varepsilon}) \geq V(x_0;\varepsilon)$ , which holds by (2.39) (replace the subindex \* with V) as follows,

$$\inf_{(\hat{Y},\hat{R})\in\hat{\mathcal{A}}(\hat{x}_0)} \hat{J}(\hat{x}_0,\hat{Y},\hat{R},\hat{\mathbb{Q}}_V;\hat{\varepsilon}) \ge \inf_{(Y,R)\in\mathcal{A}(x_0)} J(x_0,Y,R,\mathbb{Q}_V;\varepsilon) = V(x_0;\varepsilon).$$

# 5 Dependency on the ambiguity parameters

In this section we study the dependence of the value functions and the optimal cut-offs on the ambiguity parameters. We show continuity and that as  $\varepsilon \to 0$  our model converges to the risk-neutral model, studied by Harrison and Taksar [21] and by Atar and Shifrin [6]. For this, recall the definition of  $V(\cdot; 0)$  given in Remark 2.2.

**Theorem 5.1** The mapping  $[0,\infty) \ni \varepsilon \mapsto (V(\cdot;\varepsilon), V'(\cdot;\varepsilon))$  is increasing and continuous in the uniform norm topology taken on the interval [0,b]. Moreover, there is a constant C > 0such that for every  $\varepsilon \in (0,\infty)$ ,  $\sup_{x\in[0,b]} |V(x;\varepsilon) - V(x;0)| \le C\varepsilon$ . Also,  $\lim_{\varepsilon\to\infty} V(\cdot;\varepsilon) = \infty$ , uniformly on [0,b]. Finally, consider the relations  $\prec$  and  $\preceq$  on  $(0,\infty)^I \times (0,\infty)^I$  given by,

$$\hat{\varepsilon} \prec \hat{\varepsilon}' \quad (resp., \preceq) \qquad if and only if \qquad \sum_{i=1}^{I} (\theta \hat{\sigma})_i^2 \hat{\varepsilon}_i < \sum_{i=1}^{I} (\theta \hat{\sigma})_i^2 \hat{\varepsilon}_i' \quad (resp., \leq).$$

Then the mapping  $(0,\infty)^I \ni \hat{\varepsilon} \mapsto \hat{V}(\cdot;\hat{\varepsilon})$  is increasing w.r.t.  $\prec$  and continuous in the uniform norm topology taken on  $\mathcal{X}$ .

**Proof:** The last part of the theorem merely follows by the first one, (2.25), and Corollary 2.1. Therefore it is omitted and we turn to proving the first part of the theorem. We start by showing the monotonicity and continuity of the mapping  $(0, \infty) \ni \varepsilon \mapsto V(\cdot; \varepsilon)$ . The proof for  $\varepsilon = 0$  is given separately. Fix  $0 < \varepsilon_2 < \varepsilon_1$ . Denote by  $\mathbb{Q}_i = \mathbb{Q}_{V(\cdot;\varepsilon_i)}$  and  $\psi_i = \psi_{V(\cdot;\varepsilon_i)}$ , i = 1, 2.

Then, for every  $x_0 \in [0, b]$  one has

$$\begin{aligned} V(x_0;\varepsilon_1) &= \inf_{(Y,R)\in\mathcal{A}(x_0)} J(x_0,Y,R,\mathbb{Q}_1;\varepsilon_1) \\ &= \inf_{(Y,R)\in\mathcal{A}(x_0)} \left[ J(x_0,Y,R,\mathbb{Q}_1;\varepsilon_2) + \frac{1}{2} \left( \frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1} \right) \int_0^\infty e^{-\varrho t} \mathbb{E}^{\mathbb{Q}_1} \left[ \psi_1^2(t) \right] dt \right] \\ &\leq \inf_{(Y,R)\in\mathcal{A}(x_0)} \left[ J(x_0,Y,R,\mathbb{Q}_1;\varepsilon_2) + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2) \right] \\ &\leq \inf_{(Y,R)\in\mathcal{A}(x_0)} \sup_{\mathbb{Q}\in\mathcal{Q}(x_0)} \left[ J(x_0,Y,R,\mathbb{Q};\varepsilon_2) \right] + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2) \\ &= V(x_0;\varepsilon_2) + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2). \end{aligned}$$

The first equality follows by Theorem 4.1. The second equality follows by (2.33). The first inequality follows since by (3.6),  $\psi_1(t) = \varepsilon \sigma V'(X(t); \varepsilon_1)$  and since  $V'(x; \varepsilon_1) \leq r$ , see Corollary 3.1. The second inequality is trivial and finally, the last equality follows by the definition of V, see (2.34).

On the other hand,

$$V(x_0;\varepsilon_1) \ge \inf_{\substack{(Y,R)\in\mathcal{A}(x_0)}} J(x_0,Y,R,\mathbb{Q}_2;\varepsilon_1)$$
  
= 
$$\inf_{\substack{(Y,R)\in\mathcal{A}(x_0)}} \left[ J(x_0,Y,R,\mathbb{Q}_2;\varepsilon_2) + \frac{1}{2} \left( \frac{1}{\varepsilon_2} - \frac{1}{\varepsilon_1} \right) \int_0^\infty e^{-\varrho t} \mathbb{E}^{\mathbb{Q}_2} \left[ \psi_2^2(t) \right] dt \right]$$
  
> 
$$\inf_{\substack{(Y,R)\in\mathcal{A}(x_0)}} J(x_0,Y,R,\mathbb{Q}_2;\varepsilon_2)$$
  
= 
$$V(x_0;\varepsilon_2).$$

The first inequality follows by the definition of V. The first equality follows by (2.33). The strict inequality follows since  $\varepsilon_1 > \varepsilon_2$  and since by (3.6),  $\psi_1(t) = \varepsilon \sigma V'(X(t); \varepsilon_1)$  and since  $V'(x; \varepsilon_1) \ge 0$ , see Corollary 3.1. Obviously, with  $\mathbb{Q}_2$ -probability zero,  $V'(X(t); \varepsilon_1) = 0$  for almost every  $t \in \mathbb{R}_+$  w.r.t. Lebesgue measure. The last equality follows by Theorem 4.1. Combining the last two sets of relations, we obtain,

$$V(x_0;\varepsilon_2) < V(x_0;\varepsilon_1) \le V(x_0;\varepsilon_2) + \frac{\varepsilon_1 \sigma^2 r^2}{2\varepsilon_2 \varrho} (\varepsilon_1 - \varepsilon_2)$$

and the monotonicity and continuity of  $\varepsilon \mapsto V(\cdot; \varepsilon)$  is proven on the interval  $(0, \infty)$ . The monotonicity at  $\varepsilon = 0$  follows by the following sequence of relations,

$$V(x_{0};0) = \inf_{\substack{(Y,R)\in\mathcal{A}(x_{0})}} \mathbb{E}^{\mathbb{P}} \left[ \int_{0}^{\infty} e^{-\varrho t} (h(X(t))dt + rdR(t)) \right]$$
(5.1)  
$$= \inf_{\substack{(Y,R)\in\mathcal{A}(x_{0})}} J(x_{0}, Y, R, \mathbb{P}; \varepsilon_{1})$$
  
$$\leq \inf_{\substack{(Y,R)\in\mathcal{A}(x_{0})}} \sup_{\mathbb{Q}\in\mathcal{Q}(x_{0})} J(x_{0}, Y, R, \mathbb{Q}; \varepsilon_{1})$$
  
$$= V(x_{0}; \varepsilon_{1}),$$

where  $x_0 \in [0, b]$  and  $\varepsilon_1 \in (0, \infty)$ . The first equality follows by the definition of the cost in the risk-neutral case, see (2.35). The second equality follows since  $L^{\varrho}(\mathbb{P}||\mathbb{P}) = 0$ .

We now turn to proving the continuity at  $\varepsilon = 0$ . Notice that in the arguments above, the inequality  $\varepsilon_2 > 0$  cannot be relaxed to  $\varepsilon_2 \ge 0$  since we divide by  $\varepsilon_2$ . Therefore, we come up with another proof for continuity at 0. Recall that the arguments in Section 4 included the case where  $\varepsilon$  is zero. Fix an optimal cut-off in the risk-neutral problem  $\bar{\beta}_0 \in [\beta_0, \hat{\beta}_0]$ . Fix also  $\varepsilon > 0$  and set  $\mathbb{Q}^{\varepsilon} := \mathbb{Q}_{V(\cdot;\varepsilon)}$  and  $\psi^{\varepsilon} := \psi_{V(\cdot;\varepsilon)}$ . From (5.1) and Theorem 4.1 one has,

$$V(x_0;0) \le V(x_0;\varepsilon) = J(x_0, Y_{\bar{\beta}_0}, R_{\bar{\beta}_0}, \mathbb{Q}^{\varepsilon};\varepsilon).$$

Recall that by (2.33) and (3.6),

$$J(x_0, Y_{\bar{\beta}_0}, R_{\bar{\beta}_0}, \mathbb{Q}^{\varepsilon}; \varepsilon)$$

$$= \mathbb{E}^{\mathbb{Q}^{\varepsilon}} \Big[ \int_0^\infty e^{-\varrho t} (h(X_{\bar{\beta}_0}(t))dt + rdR_{\bar{\beta}_0}(t)) \Big] - \varepsilon \sigma^2 \int_0^\infty e^{-\varrho t} (V'(X_{\bar{\beta}_0}(s); \varepsilon))^2 ds,$$
(5.2)

where

$$X_{\bar{\beta}_0}(t) = x_0 + mt + \sigma B(t) + Y_{\bar{\beta}_0}(t) - R_{\bar{\beta}_0}(t), \qquad t \in \mathbb{R}_+.$$
(5.3)

Since  $0 \leq V'(\cdot; \varepsilon) \leq r$ , the last term on the r.h.s. of (5.2) goes to zero as  $\varepsilon \to 0$ . Hence, in order to show the limit  $\lim_{\varepsilon \to 0} V(x_0; \varepsilon) = V(x_0; 0)$  it is sufficient to show that

$$\lim_{\varepsilon \to 0} \mathbb{E}^{\mathbb{Q}^{\varepsilon}} \left[ \int_0^\infty e^{-\varrho t} (h(X_{\bar{\beta}_0}(t))dt + rdR_{\bar{\beta}_0}(t)) \right] = \mathbb{E}^{\mathbb{P}} \left[ \int_0^\infty e^{-\varrho t} (h(X_{\bar{\beta}_0}(t))dt + rdR_{\bar{\beta}_0}(t)) \right].$$
(5.4)

The proof of (5.4) relies on the continuity property of the Skorokhod mapping in addition to a coupling argument. We consider a probability space  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$  that supports a onedimensional standard Brownian motion B, adapted to the filtration  $\{\mathcal{F}_t\}$ , such that for every  $0 \leq s < t, B(t) - B(s)$  is independent of  $\mathcal{F}_s$  under  $\mathbb{P}$ . Recall Definition 3.1 and consider the following processes,

$$\begin{aligned} (X^{\varepsilon}, Y^{\varepsilon}, R^{\varepsilon})(t) &:= \Gamma_{[0,\bar{\beta}_0]} \left( x_0 + m \cdot + \int_0^{\cdot} \varepsilon \sigma^2 V'(X^{\varepsilon}(s); \varepsilon)(s) ds + \sigma B(\cdot) \right)(t), \\ (X^0, Y^0, R^0)(t) &:= \Gamma_{[0,\bar{\beta}_0]} \left( x_0 + m \cdot + \sigma B(\cdot) \right)(t). \end{aligned}$$

Notice that by (3.6),  $(X^0, Y^0, R^0)$  (resp.,  $(X^{\varepsilon}, Y^{\varepsilon}, R^{\varepsilon})$ ) has the same distribution under the measure  $\check{\mathbb{P}}$  as  $(X_{\bar{\beta}_0}, Y_{\bar{\beta}_0}, R_{\bar{\beta}_0})$  given in (5.3) under the measure  $\mathbb{P}$  (resp.,  $\mathbb{Q}^{\varepsilon}$ , see (2.29)). Hence, (5.4) follows once we show that

$$\lim_{\varepsilon \to 0} \mathbb{E}^{\check{\mathbb{P}}} \left[ \int_0^\infty e^{-\varrho t} (h(X^{\varepsilon}(t))dt + rdR^{\varepsilon}(t)) \right] = \mathbb{E}^{\check{\mathbb{P}}} \left[ \int_0^\infty e^{-\varrho t} (h(X^0(t))dt + rdR^0(t)) \right].$$
(5.5)

Now,

$$\begin{aligned} \left| \mathbb{E}^{\check{\mathbb{P}}} \Big[ \int_{0}^{\infty} e^{-\varrho t} (h(X^{\varepsilon}(t))dt + rdR^{\varepsilon}(t)) \Big] - \mathbb{E}^{\check{\mathbb{P}}} \Big[ \int_{0}^{\infty} e^{-\varrho t} (h(X^{0}(t))dt + rdR^{0}(t)) \Big] \right| \qquad (5.6) \\ &= \left| \mathbb{E}^{\check{\mathbb{P}}} \Big[ \int_{0}^{\infty} e^{-\varrho t} [h(X^{\varepsilon}(t)) - h(X^{0}(t))]dt + \int_{0}^{\infty} \varrho r e^{-\varrho t} [R^{\varepsilon}(t) - R^{0}(t)]dt + [e^{-\varrho t} (R^{\varepsilon}(t) - R^{0}(t))] \Big|_{0}^{\infty} \Big] \\ &\leq \mathbb{E}^{\check{\mathbb{P}}} \Big[ \int_{0}^{\infty} e^{-\varrho t} [h(X^{\varepsilon}(t)) - h(X^{0}(t))]dt + \int_{0}^{\infty} \varrho r e^{-\varrho t} [R^{\varepsilon}(t) - R^{0}(t)]dt \Big]. \end{aligned}$$

The last inequality follows since  $\lim_{t\to\infty} e^{-\varrho t} R^{\varepsilon}(t) = 0$ ,  $\check{\mathbb{P}}$ -a.s. and similarly for  $R^0$ . Indeed, consider  $w(\cdot) = x_0 + m \cdot + \int_0^{\cdot} \varepsilon \sigma^2 V'(X^{\varepsilon}(s);\varepsilon)(s)ds + \sigma B(\cdot)$  and  $\tilde{w} = 0$  in Lemma 3.1. Recalling the bound  $0 \leq V'(\cdot,\varepsilon) \leq r$ , it follows that  $R^{\varepsilon}(t) \leq c_S(x_0 + (m + \varepsilon \sigma^2 r)t + \sigma \sup_{0\leq s\leq t} |B(s)|)$  and the limit follows since  $\lim_{t\to\infty} e^{-\varrho t} \sup_{0\leq s\leq t} |B(s)| = 0$ ,  $\check{\mathbb{P}}$ -a.s. Using the Lipschitz continuity of h, again  $0 \leq V'(\cdot,\varepsilon) \leq r$ , and Lemma 3.1 with  $w(\cdot) = x_0 + m \cdot + \int_0^{\cdot} \varepsilon \sigma^2 V'(X^{\varepsilon}(s);\varepsilon)(s)ds + \sigma B(\cdot)$  and  $\tilde{w}(\cdot) = x_0 + m \cdot + \sigma B(\cdot)$ , we get that for every  $t \in \mathbb{R}_+$ ,

$$\sup_{s \in [0,t]} \| (X^{\varepsilon}, Y^{\varepsilon}, R^{\varepsilon})(s) - (X^0, Y^0, R^0)(s) \| \le \varepsilon c_S \sigma^2 r t.$$

Recalling that h is Lipschitz, we get that the r.h.s. of (5.6) is bounded above by  $C\varepsilon(1/\rho + r)/\rho$ , where the constant C > 0 depends on  $c_S$ ,  $\sigma^2 r$ , and the Lipschitz constant of h. This implies (5.5). Recalling (5.2) and the estimations above, we get the bound

$$\sup_{x \in [0,b]} |V(x;\varepsilon) - V(x;0)| \le (\sigma^2 r^2 + C\varepsilon(1/\varrho + r)/\varrho)\varepsilon$$

as required.

We now turn to proving the limit  $\lim_{\varepsilon \to \infty} V(\cdot; \varepsilon) = \infty$ . For this, consider the strategy of the maximizer given by  $\mathbb{Q}_{\varepsilon}$ , which is associated with  $\psi_{\varepsilon}(s) = \varepsilon^{1/4}$  for every  $s \in \mathbb{R}_+$ . Given that and recalling the presentation of the dynamics provided in (2.29), it follows that,

$$X(t) = x_0 + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}\varepsilon}(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+,$$
(5.7)

where  $B^{\mathbb{Q}_{\varepsilon}}$  is an  $\{\mathcal{F}_t\}$ -one-dimensional standard Brownian motion under  $\mathbb{Q}_{\varepsilon}$ . His cost function is given by

$$J(x_0, Y, R, \mathbb{Q}_{\varepsilon}; \varepsilon) = \mathbb{E}^{\mathbb{Q}} \left[ \int_0^\infty e^{-\varrho t} \left( h(X(t)) dt + r dR(t) - \frac{1}{2\varepsilon} \psi^2(t) dt \right) \right]$$
$$= \mathbb{E}^{\mathbb{Q}_{\varepsilon}} \left[ \int_0^\infty e^{-\varrho t} (h(X(t)) dt + r dR(t)) \right] - \frac{1}{2\varrho^2 \varepsilon^{1/2}}.$$

Since the last term is constant, we can think as if the minimizer is facing a risk-neutral problem with the drift mt replaced by  $(m + \sigma \varepsilon^{1/4})t$ , where her cost function has an additional deterministic term, which vanishes as  $\varepsilon \to \infty$ . Therefore, we analyze only the limit of the expectation in the second line of the above. Consider an optimal  $\beta$ -reflecting strategy  $(Y_{\beta}, R_{\beta})$ , which exists in the risk-neutral problem. From the dynamics in (5.7), we get for every  $t \in \mathbb{R}_+$ ,

$$R_{\beta}(t) \ge x_0 - X(t) + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_{\varepsilon}}(t) + Y_{\beta}(t) \ge -b + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_{\varepsilon}}(t).$$

The same arguments that lead to (5.6), imply that

$$\mathbb{E}^{\mathbb{Q}_{\varepsilon}} \Big[ \int_{0}^{\infty} e^{-\varrho t} (h(X(t))dt + rdR_{\beta}(t)) \Big] = \mathbb{E}^{\mathbb{Q}_{\varepsilon}} \Big[ \int_{0}^{\infty} e^{-\varrho t} (h(X(t)) + \varrho rR_{\beta}(t))dt \Big]$$
  
$$\geq \mathbb{E}^{\mathbb{Q}_{\varepsilon}} \Big[ \int_{0}^{\infty} e^{-\varrho t} (h(X(t)) + \varrho r(-b + (m + \sigma \varepsilon^{1/4})t + \sigma B^{\mathbb{Q}_{\varepsilon}}(t)))dt \Big].$$

Standard estimates give that the last term goes to  $\infty$  together with  $\varepsilon$ .

From Theorem 4.1, Corollary 3.1, and (3.6) it is easy to see the continuity of the maximizer's optimal control w.r.t.  $\varepsilon$ . Also, it is clear that as  $\varepsilon \to 0$ ,  $\psi_{V(\cdot;\varepsilon)} \to 0$ , uniformly on [0, b] and therefore, the maximizer's optimal strategy is relatively close to keep the original measure as is. Since we do not have explicit expressions for the cut-off points of the reflecting strategies  $\beta_{\varepsilon}$  and  $\hat{\beta}_{\varepsilon}$  described in Section 4, the situation with the minimizer's optimal strategy is more subtle and is studied now. The next theorem provides a continuity property of the optimal cut-off points  $\beta_{\varepsilon}$  and  $\hat{\beta}_{\varepsilon}$  w.r.t.  $\varepsilon$ . Recall the sufficient conditions given in (4.3) for uniqueness of the reflecting strategy.

**Theorem 5.2** For any given  $\varepsilon \in [0, \infty)$ ,

$$\beta_{\varepsilon} \le \liminf_{\delta \to 0} \beta_{\varepsilon+\delta} \le \limsup_{\delta \to 0} \hat{\beta}_{\varepsilon+\delta} \le \hat{\beta}_{\varepsilon}.$$
(5.8)

Hence, in case that  $\hat{\beta}_{\varepsilon} = \beta_{\varepsilon}$  (for the given  $\varepsilon$ ), that is, if there is a unique optimal reflecting strategy, one has,

$$\lim_{\delta \to 0} \beta_{\varepsilon + \delta} = \lim_{\delta \to 0} \hat{\beta}_{\varepsilon + \delta} = \hat{\beta}_{\varepsilon}.$$
(5.9)

As a preparation for the proof, we present an auxiliary function defined for every  $\varepsilon \in [0, \infty)$ . Recall the definition of  $k^{(s)}$  given in (3.17) and (3.32). Set  $l^{(\varepsilon)}(x) = k^{(V(0;\varepsilon))}(x), x \in [0, b]$ . More explicitly,

$$\begin{cases} (l^{(\varepsilon)})''(x) + H_F(x, l^{(\varepsilon)}(x), (l^{(\varepsilon)})'(x)) = 0, & x \in [0, b], \\ (l^{(\varepsilon)})'(0) = 0, & l^{(\varepsilon)}(0) = V(0; \varepsilon), \end{cases}$$
(5.10)

with the same F given after (3.17). The next lemma provides a continuity property of  $l^{(\varepsilon)}$  as a function of  $\varepsilon$ .

**Lemma 5.1** The mapping  $[0,\infty) \ni \varepsilon \mapsto (l^{(\varepsilon)}(\cdot),(l^{(\varepsilon)})'(\cdot))$  is continuous in the uniform norm topology taken on the interval [0,b].

**Proof:** Fix  $\varepsilon_1, \varepsilon_2 \in \mathbb{R}_+$ . Set  $f_1 = l^{(\varepsilon_1)}$  and  $f_2 = l^{(\varepsilon_2)}$ . Then,

$$f_1'(x) = -\int_0^x H_F^{\varepsilon_1}(y, f_1(y), f_1'(y)) dy,$$
  
$$f_2'(x) = -\int_0^x H_F^{\varepsilon_2}(y, f_2(y), f_2'(y)) dy = -\int_0^x H_F^{\varepsilon_1}(y, f_2(y), f_2'(y)) dy + \int_0^x (\varepsilon_2 - \varepsilon_1) F(f_2'(y)),$$

where the index  $\varepsilon_i$  in  $H_F^{\varepsilon_i}$  emphasizes its dependence on  $\varepsilon_i$ , i = 1, 2. From (3.19) it follows that there exists a constant L > 0 independent of  $\varepsilon_1, \varepsilon_2$  such that

$$|f_1'(x) - f_2'(x)| \le L\left(\int_0^x \left[|f_1(y) - f_2(y)| + |f_1'(y) - f_2'(y)|\right] dy + |\varepsilon_1 - \varepsilon_2|\right).$$

Also, since  $f_i(0) = V(0; \varepsilon_i)$ ,

$$|f_1(x) - f_2(x)| \le |V(0;\varepsilon_1) - V(0;\varepsilon_2)| + \int_0^x |f_1'(y) - f_2'(y)| dy.$$

From the last two inequalities and Grönwall's inequality we get that there is a constant C > 0 independent of  $\varepsilon_1, \varepsilon_2$ , and x, such that for every  $x \in [0, b]$ 

$$|f_1(x) - f_2(x)| + |f_1'(x) - f_2'(x)| \le C \left( |V(0;\varepsilon_1) - V(0;\varepsilon_2)| + |\varepsilon_1 - \varepsilon_2| \right).$$

Recalling from Theorem 5.1 that  $\varepsilon \mapsto V(\cdot; \varepsilon)$  is continuous, we get that  $\sup_{0 \le x \le b} (|f_1(x) - f_2(x)| + |f'_1(x) - f'_2(x)|)$  is uniformly bounded by a function of  $(\varepsilon_1, \varepsilon_2)$  that goes to zero as  $(\varepsilon_1 - \varepsilon_2) \to 0$ .

**Proof of Theorem 5.2:** For any  $\varepsilon' \in [0, \infty)$ , the function  $V(\cdot; \varepsilon')$  satisfies (3.5) and therefore also (5.10) on  $x \in [0, \hat{\beta}_{\varepsilon'}]$ . Uniqueness of the solution implies that

$$V(x;\varepsilon') = l^{(\varepsilon')}(x), \qquad x \in [0,\hat{\beta}_{\varepsilon'}], \tag{5.11}$$

see [29, Section 0.3.1].

To get the first inequality in (5.8) notice that from (5.11), the definition of  $\beta_{\varepsilon}$  in (3.4), the definition of  $\beta^{(s)}$  in (3.20), and recalling that  $l^{(\varepsilon)} = k^{(V(0;\varepsilon))}$ , it follows that for any  $\varepsilon' \in [0, \infty)$ ,  $\beta_{\varepsilon'} = \beta^{(V(0;\varepsilon'))}$ . From the continuity of  $\varepsilon \mapsto V(\cdot; \varepsilon)$ , see Theorem 5.1, and the second inequality in (3.23), we get the first inequality in (5.8).

The second inequality in (5.8) is trivial and follows since  $\beta_{\varepsilon+\delta} \leq \beta_{\varepsilon+\delta}$ . We now turn to proving the last inequality in (5.8). Fix  $\varepsilon \in [0, \infty)$ . From Theorem 5.1 and Lemma 5.1 we have

$$\lim_{\delta \to 0} V(\cdot; \varepsilon + \delta) = V(\cdot; \varepsilon) \quad \text{and} \quad \lim_{\delta \to 0} l^{(\varepsilon + \delta)}(\cdot) = l^{(\varepsilon)}(\cdot),$$

uniformly on [0, b]. Together with (5.11) applied to  $\varepsilon' = \varepsilon + \delta$ , one has  $V(x; \varepsilon + \delta) = l^{(\varepsilon + \delta)}(x)$ ,  $x \in [0, \hat{\beta}_{\varepsilon + \delta})$ . Taking  $\delta \to 0$ , we get that  $V(x; \varepsilon) = l^{(\varepsilon)}(x)$  for  $x \in [0, \alpha_{\varepsilon})$ , where  $\alpha_{\varepsilon} := \lim \sup_{\delta \to 0} \hat{\beta}_{\varepsilon + \delta}$ . Since both functions are continuous, the equality holds for  $\alpha_{\varepsilon}$  as well. By the definition of  $\hat{\beta}_{\varepsilon}$ , see (4.1), we get that (5.8) holds.

# Acknowledgments

The author is greatful to an anonymous associate editor and to two anonymous referees for their careful reading of the paper and for their suggestions that improved the presentation of the paper.

# References

- R. Atar and A. Biswas. Control of the multiclass G/G/1 queue in the moderate deviation regime. Ann. Appl. Probab., 24(5):2033–2069, 2014.
- [2] R. Atar and A. Cohen. A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime. *Math. Oper. Res.*, 41(4):1354–1380, 2016.
- [3] R. Atar and A. Cohen. Asymptotically optimal control for a multiclass queueing model in the moderate deviation heavy traffic regime. Ann. Appl. Probab., 27(5):2862–2906, 10 2017.

- [4] R. Atar and G. Mendelson. On the non-Markovian multiclass queue under risk-sensitive cost. Queueing Syst., 84(3-4):265-278, 2016.
- [5] R. Atar and S. Saha. Optimality of the generalized cμ rule in the moderate deviation regime. Queueing Systems, 87(1):113–130, Oct 2017.
- [6] R. Atar and M. Shifrin. An asymptotic optimality result for the multiclass queue with finite buffers in heavy traffic. *Stochastic Systems*, 4(2):556–603, 2014.
- [7] E. Bayraktar and Y. Zhang. Minimizing the probability of lifetime ruin under ambiguity aversion. SIAM J. Control Optim., 53(1):58–90, 2015.
- [8] S. L. Bell and R. J. Williams. Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: asymptotic optimality of a threshold policy. Ann. Appl. Probab., 11(3):608–649, 2001.
- [9] A. Biswas. Risk-sensitive control for the multiclass many-server queues in the moderate deviation regime. Math. Oper. Res., 39(3):908-929, 2014.
- [10] J. Blanchet, C. Dolan, and H. Lam. Robust rare-event performance analysis with natural non-convex constraints. In *Proceedings of the 2014 Winter Simulation Conference*, pages 595–603. IEEE Press, 2014.
- [11] A. Budhiraja and A. P. Ghosh. Diffusion approximations for controlled stochastic networks: an asymptotic bound for the value function. Ann. Appl. Probab., 16(4):1962–2006, 2006.
- [12] A. Budhiraja and A. P. Ghosh. Controlled stochastic networks in heavy traffic: convergence of value functions. Ann. Appl. Probab., 22(2):734–791, 2012.
- [13] K. Chowdhary and P. Dupuis. Distinguishing and integrating aleatoric and epistemic variation in uncertainty quantification. ESAIM: Mathematical Modelling and Numerical Analysis, 47(3):635–662, 2013.
- [14] A. Cohen. Asymptotic analysis of a multiclass queueing control problem under heavy-traffic with model uncertainty. Arxiv preprint, 2017.
- [15] W. H. Fleming and H. M. Soner. Controlled Markov processes and viscosity solutions, volume 25 of Stochastic Modelling and Applied Probability. Springer, New York, second edition, 2006.
- [16] L. P. Hansen and T. J. Sargent. Robustness. Princeton University Press, Princeton, NJ, 2008.
- [17] L. P. Hansen, T. J. Sargent, G. Turmuhambetova, and N. Williams. Robust control and model misspecification. J. Econom. Theory, 128(1):45–90, 2006.
- [18] J. M. Harrison. Brownian models of queueing networks with heterogeneous customer populations. In Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986), volume 10 of IMA Vol. Math. Appl., pages 147–186. Springer, New York, 1988.
- [19] J. M. Harrison. Brownian models of open processing networks: canonical representation of workload. Ann. Appl. Probab., 10(1):75–103, 2000.
- [20] J. M. Harrison. A broader view of Brownian networks. Ann. Appl. Probab., 13(3):1119–1150, 2003.
- [21] J. M. Harrison and M. I. Taksar. Instantaneous control of Brownian motion. Math. Oper. Res., 8(3):439– 453, 1983.
- [22] J. M. Harrison and J. A. Van Mieghem. Dynamic control of Brownian networks: state space collapse and equivalent workload formulations. Ann. Appl. Probab., 7(3):747–771, 1997.
- [23] J. M. Harrison and R. J. Williams. Workload reduction of a generalized Brownian network. Ann. Appl. Probab., 15(4):2255–2295, 2005.
- [24] A. Jain, A. E. B. Lim, and J. G. Shanthikumar. On the optimality of threshold control in queues with model uncertainty. *Queueing Syst.*, 65(2):157–174, 2010.
- [25] 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.
- [26] H. Lam. Robust sensitivity analysis for stochastic systems. Math. Oper. Res., 41(4):1248–1275, 2016.
- [27] P. J. Maenhout. Robust portfolio rules and asset pricing. Review of financial studies, 17(4):951–983, 2004.
- [28] E. Plambeck, S. Kumar, and J. M. Harrison. A multiclass queue in heavy traffic with throughput time constraints: asymptotically optimal dynamic controls. *Queueing Syst.*, 39(1):23–54, 2001.

- [29] A. D. Polyanin and V. F. Zaitsev. Handbook of exact solutions for ordinary differential equations. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2003.
- [30] M. H. Protter and H. F. Weinberger. Maximum principles in differential equations. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1967.
- [31] M. Shifrin, R. Atar, and I. Cidon. Optimal scheduling in the hybrid-cloud. In 2013 IFIP/IEEE International Symposium on Integrated Network Management (IM 2013), pages 51–59, May 2013.
- [32] M. Sîrbu. A note on the strong formulation of stochastic control problems with model uncertainty. *Electron. Commun. Probab.*, 19:no. 81, 10, 2014.
- [33] J. Stoer and R. Bulirsch. Introduction to numerical analysis. Springer-Verlag, New York-Heidelberg, 1980. Translated from the German by R. Bartels, W. Gautschi and C. Witzgall.