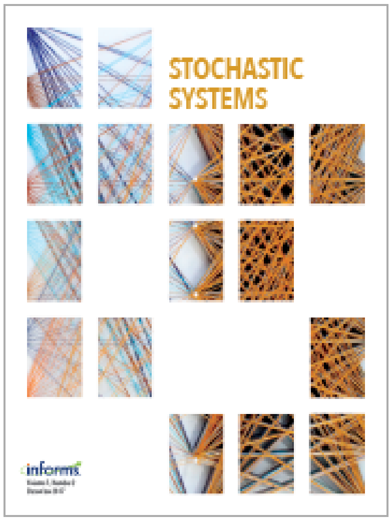


Stochastic Systems

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Asymptotic Analysis of a Multiclass Queueing Control Problem Under Heavy Traffic with Model Uncertainty

Asaf Cohen

informs
http://pubsonline.informs.org/journal/stsy

STOCHASTIC SYSTEMS
Vol. 9, No. 4, December 2019, pp. 359–391
ISSN 1946-5238 (online)

Asymptotic Analysis of a Multiclass Queueing Control Problem Under Heavy Traffic with Model Uncertainty

Asaf Cohen[a]

[a] Department of Statistics, University of Haifa, 31905 Haifa, Israel

Contact: shloshim@gmail.com, http://orcid.org/0000-0002-9211-7956 (AC)

Received: August 6, 2018
Revised: December 17, 2018; March 12, 2019
Accepted: March 12, 2019
Published Online in Articles in Advance: December 5, 2019

https://doi.org/10.1287/stsy.2019.0034

Abstract. We study a multiclass M/M/1 queueing control problem with finite buffers under heavy traffic in which the decision maker is uncertain about the rates of arrivals and service of the system and by scheduling and admission/rejection decisions acts to minimize a discounted cost that accounts for the uncertainty. The main result is the asymptotic optimality of a $c\mu$ type of policy derived via underlying stochastic differential games studied in Cohen [Cohen A (2018) Brownian control problems for a multiclass M/M/1 queueing problem with model uncertainty. *Math. Oper. Res.* 44(2):739–766.]. Under this policy, with high probability, rejections are not performed when the workload lies below some cutoff that depends on the ambiguity level. When the workload exceeds this cutoff, rejections are carried out and only from the buffer with the cheapest rejection cost weighted with the mean service rate. The allocation part of the policy is the same for all the ambiguity levels. To our knowledge, this is the first work to address a heavy-traffic queueing control problem with model uncertainty.

History: Former designation of this paper was SSY-2018-023.

Funding: This research was supported by the Israel Science Foundation [Grant 1169/18].

Keywords: multiclass M/M/1 • model uncertainty • heavy traffic • diffusion approximation • Brownian control problem • stochastic differential game • state-dependent priorities • ambiguity aversion

1. Introduction

The model under consideration consists of a multiclass M/M/1 queueing model under diffusion-scaled heavy traffic in which the decision maker (DM) is uncertain about the parameters and acts to optimize an overall cost that accounts for this uncertainty. The model consists of a server that, at any time instant, the server's effort is allocated by the DM to customers from several numbers of classes. Customers of each class are kept in a finite buffer. Apart from the scheduling control, upon arrival of a customer, the DM has to decide whether to reject it or to assign it to the buffer that corresponds to its class type. The DM has ambiguity about the rates of arrivals and the mean service times. The cost accounts for the ambiguity, the holding of customers in the buffers, and rejections of new arrivals.

The problem without ambiguity, also referred to as the *risk-neutral problem*, was analyzed by Atar and Shifrin (2014) under the framework of G/G/1. Plambeck et al. (2001) studied a similar nonrobust problem with time constraints instead of the finite buffer constraints. In these problems as well as in many other classical models of queueing control problems (QCPs), see, for example, Bell and Williams (2001) and Budhiraja and Ghosh (2006, 2012) and the references therein, there is a fixed time-homogeneous random model; that is, the DM is certain about the evolution of the system, which, moreover, does not change in time. Such an assumption is not realistic, and a robust analysis is desirable. Because of the complexity of real-world systems, lack of sufficient calibration, and inaccurate assumptions, one cannot precisely model the arrival and departure processes. In recent years there has been increasing interest in robust analysis of queueing systems. We consider uncertainty in the diffusion scale of the QCP in a way that is often referred to as *model uncertainty* or *Knightian uncertainty*, see, for example, Maenhout (2004), Hansen et al. (2006), Hansen and Sargent (2008), and Bayraktar and Zhang (2015), and in the context of queueing systems, see Jain et al. (2010), Blanchet et al. (2014), and Lam (2016); see also Petersen et al. (2000) in a discrete time setup. This is not to be confused with the terminology *robust queueing theory*, which is often referred to in optimization-based performance analysis;

see, for example, Bandi et al. (2015) and Whitt and You (2018). Uncertainty in fluid models of queueing was studied, for example, in Dupuis (2003), Whitt (2006), and Bassamboo et al. (2010).

We consider a DM who is skeptical about the validity of an underlying model and takes into account a family of models in the following way. We assume that the DM has a *reference model* in mind, which, up to some degree, describes the situation the DM is facing. To model the uncertainty about the reference model, the DM takes into account other models and penalizes them based on their deviation from the reference model. More specifically, we consider the following cost function, which the DM aims to minimize

$$\sup_{\mathbb{Q}}\left\{\mathbb{E}^{\mathbb{Q}}\left[\int_0^\infty e^{-\varrho t}\left(\hat{h}\cdot X(t)dt+\hat{r}\cdot dR(t)\right)\right]-\alpha L(\mathbb{Q}\|P)\right\},$$

where I is the number of classes. The vectors $\hat{h}$ and $\hat{r}$ stand for the holding and the rejection costs, respectively. The I-dimensional processes X and R represent the queue lengths and the rejection processes, respectively. The supremum is taken over measures equivalent to $\mathbb{P}$, and the function L is a discounted variant of the Kullback–Leibler divergence. It measures how far away $\mathbb{Q}$ is from the reference measure $\mathbb{P}$ and penalizes accordingly. We allow for different levels of model uncertainty for each of the arrival processes as well as for each of the service processes. The parameter $\alpha > 0$ measures the level of ambiguity the DM has regarding the reference model. It can also be viewed as the Lagrangian of a performance-optimization problem with an uncertainty set constraint as presented in Jain et al. (2010), equation 14, and Lam (2016), equation 2. Moreover, exploiting the Markovian structure of the problem, we argue that one may replace the supremum over measures by a supremum over (possibly time-dependent) rates, which are penalized according to their deviation from the reference rates.

To tackle QCPs under heavy traffic, one often solves a limiting control problem associated with a Brownian motion, called *Brownian control problem* (BCP) and uses its solution to construct an asymptotically optimal policy in the QCP. This concept was first introduced by Harrison (1988); for further reading on BCPs, see, for example, Bell and Williams (2001) and Budhiraja and Ghosh (2006, 2012) and the references therein. In our case, the cost function given suggests that the BCP is, in fact, a stochastic game. The players in this game are the DM and the nature, which, according to their goals, are referred to as the *minimizer* and the *maximizer*, respectively. Interpreting the roles of the processes from the QCP to a *multidimensional stochastic differential game*, the minimizer controls the server's effort allocation and the admission/rejection, and the maximizer is free to choose a probability measure and is penalized in accordance with the deviation of the chosen measure from the reference measure. This game was analyzed in Cohen (2018), in which it was also shown that a state-space collapse property holds. This is done by considering a *reduced stochastic differential game* (RSDG), which emerges from the workload process, and showing that both games share the same value and that, given equilibrium in either one of the games, one can construct an equilibrium in the other game. Further properties of the games, such as dependency on the ambiguity parameters, are also given there.

This paper is devoted to the connection between the QCP and the BCP (namely, the RSDG); we show that the value function of the QCP is approximately the value function of the RSDG and that the minimizer's optimal strategy from the multidimensional stochastic differential game leads to an asymptotically Markovian optimal policy in the QCP. Roughly speaking, this strategy suggests that the DM should use a $c\mu$ type of rule and fill in the buffers in accordance with their holding costs without using rejections unless a cutoff level of workload that depends on the ambiguity level has been reached and then to use rejections only from the buffer with the cheapest rejection cost, weighted with the mean service rate, until the workload level goes below the cutoff. More specifically, let μ_i be the mean service rate of class i customers under the reference model and recall that $\hat{h}_i$ is the holding cost per customer of class i. The DM should prioritize the classes in the order of $\{\hat{h}_i\mu_i\}_{i=1}^I$, where the lowest priority is given to the class with the lowest $\hat{h}_i\mu_i$ among the classes for which the buffers are not "almost" full. As for the admission-control part of the policy, whenever the workload level remains below the mentioned cutoff, use rejections only if there is a new arrival to a full buffer; as is shown, the probability of such an event vanishes with the scaling parameter. If the workload level exceeds the cutoff, the DM rejects all incoming customers to the class with the lowest $\hat{r}_i\mu_i$. This policy (with different cutoff level) was shown to be asymptotically optimal in the risk-neutral setup and in a QCP with the same mechanism but with the moderate-deviation heavy-traffic regime (instead of the diffusion scaling) and a risk-sensitive cost criteria. The latter QCP models a situation of a "very" risk-averse DM. The only difference between the policies in the three models is in the position of the cutoff point. Specifically, the allocation policy is the same. This shows us the usefulness of the allocation policy as it is robust to ambiguity. The optimality of the same policy (with differences only in the cutoff level) in these three models is not obvious because of the existence of the maximizer, which leads to a nonstationary problem. For further reading about the moderate-deviation heavy-traffic regime, see Atar and Biswas (2014), Biswas (2014), Atar and Cohen (2016, 2017), Atar and

Mendelson (2016), and Atar and Saha (2017) and, in the context of our paper, also the discussion in Cohen (2018). The current paper does not aim to establish a rigorous relationship between the QCP under the Knightian uncertainty and the QCP under the moderate-deviation heavy-traffic regime. Saying this, we only mention that a common feature to this model and the moderate-deviation one is that, in the limiting problem, the change of measure can be translated into a linear (in the maximizer's control variable) change in the drift and a quadratic penalty.

We now make some comments on proof techniques. The proof is divided into two parts: showing that the value function of the RSDG forms an asymptotic lower bound for the QCP and that the expected cost associated with the candidate policy is asymptotically bounded above by the value function of the RSDG. Also, recall that we consider a sequence of queueing systems. For the lower bound, we assign the maximizer a policy that is driven from the equilibrium strategy in the RSDG, and using it, we show that, for any sequence of strategies used by the minimizer, the expected cost is bounded below by the value of the RSDG. By its structure, the maximizer's strategy preserves the critical load of the system. The main technical difficulty in this part is that the sequence of strategies of the minimizer is arbitrary, and because of the nature of the control in the QCP, which is replaced by a singular control in the BCP, compactness arguments do not apply here. In the risk-neutral case, in which there is no maximizing player, Atar and Shifrin (2014) managed to bypass this issue by $\mathscr{C}$-tightness arguments applied to the integrands of the relevant processes and, later on, taking the derivatives of the implied limiting processes. In our setup, the maximizer's strategy depends on the scaled queue lengths process and not on its integrand. Therefore, the $\mathscr{C}$-tightness of the sequence of the scaled queue lengths processes is required. As a result, the integrand-derivative method cannot be applied in our case. We use the *time-stretching* method, which was introduced by Meyer and Zheng (1984) and studied in the same framework by Kurtz (1991). In the context of stochastic control, the method was first used by Martins and Kushner (1990) and Kushner and Martins (1991) and was adopted in Budhiraja and Ghosh (2006) and Budhiraja and Ross (2006, 2007). In Section 4.2.1, we set up a random time transformation for any system such that the controls are Lipschitz continuous with Lipschitz constant one. Then, in Lemma 2, we apply $\mathscr{C}$-tightness arguments for the sequence of the relevant time-stretched processes (including the scaled queue lengths process) and obtain limiting processes. Using an inverse time transformation in Section 4.2.2, we go back to the original scale. Finally, in Section 4.2.3, we connect between the costs associated with the time-rescaled limiting processes and the value of the RSDG.

To show that the expected cost associated with the candidate policy is asymptotically bounded above by the value function of the RSDG, we do the following. We consider an arbitrary sequence of strategies for the maximizer. This sequence of probability measures is not forced to satisfy the critical load condition. Therefore, at the first step, we show that it is "too costly" for the maximizer to have a big deviation from the reference measure and, thus, restrict the maximizer to strategies that, on average, do not deviate much from the reference measure (Proposition 3). Then, in Proposition 4, we adapt a state-space collapse result from the risk-neutral case to ours and show that, under the sequence of strategies chosen by the maximizer, the underlying stochastic dynamics of the scaled queue lengths lie close to a specific path, called the *minimizing curve*. At this point, one can use $\mathscr{C}$-tightness arguments to conclude the convergence of the relevant dynamics, including the holding and rejection cost parts. However, in order to estimate the change of measure penalty given by the Kullback–Leibler divergence, one shall have another reduction and "truncate" the maximizer's strategies such that the critical load condition is almost surely (a.s.) preserved. Then we show that the relevant processes and all the cost components associated with the two levels of restrictions of the maximizer's strategies are close to each other (Proposition 5). This approximation relies also on the state-space collapse. Thus, the rest of the analysis is performed in the more convenient case, in which the critical load condition holds and the expected cost is shown to be asymptotically bounded from above by the value function of the RSDG. For this, we reduce to one-dimensional dynamics and estimate from below the change of measure penalty, and together with the convergence of the holding and rejection cost components, we conclude the upper bound. Section 5 provides a future outlook, and Section 6 is dedicated to the proofs of the more technical and less innovative lemmas and propositions stated in Section 4.

The paper is organized as follows. Section 2 presents the model. Section 3 collects a few results from Cohen (2018) required for the proof. In Section 4, we provide and prove the main result (Theorem 1), which states that the QCP value converges to that of the BCP from Cohen (2018), and an asymptotically optimal policy is provided.

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 $\mathbb{R}_+$. The infimum of the empty set is taken to be ∞. For subintervals $I_1, I_2 \subseteq \mathbb{R}$ and $m \in \{1,2\}$, we denote by $\mathscr{C}(I_1, I_2)$, $\mathscr{C}^m(I_1, I_2)$, and $\mathscr{D}(I_1, I_2)$ the space of continuous functions (respectively, functions with continuous derivatives of order m and functions that are right-continuous with finite left limits [RCLL]), mapping $I_1 \to I_2$. The space $\mathscr{D}(I_1, I_2)$ is endowed with the usual Skorokhod topology. For $T, \delta > 0$ and a function $f : \mathbb{R}_+ \to \mathbb{R}^k$, $\|f\|_T = \sup_{t\in[0,T]} \|f(t)\|$ and $\mathrm{osc}_T(f, \delta) = \sup\{\|f(u) - f(t)\| : 0 \le u \le t \le (u + \delta) \wedge T\}$. For any RCLL processes X, Y, the quadratic variation of X is denoted by $[X]$, and the quadratic covariation of X and Y is denoted by $[X, Y]$.

2. The Queueing Model

We consider a QCP with customers of I different classes that arrive at the system to be serviced by a single server. Upon arrival, the customers are queued in I buffers with finite capacity in accordance to their class. 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. We study the system under heavy traffic. Hence, we consider a sequence of systems indexed by the scaling parameter $n \in \mathbb{N}$.

2.1. The Reference Model

Usually in (nonstatistical) stochastic control problems it is assumed that the probability measure is fully known. A model is fixed, and the DM finds the optimal policy given the probability measure. In a statistics-based approach, the parameters are not fully known, and the DM learns them. We consider an epistemic *reference probability space* that is driven by previous knowledge as in nonstatistical stochastic control problems. But, instead of working only with this probability measure, as detailed in Section 2.2, the DM considers a robust cost that takes into account a family of measures without statistically learning the true model.

For every $i \in [I] := \{1, \dots, I\}$ and $n \in \mathbb{N}$, we consider the probability spaces $(\Omega^n_{1,i}, \mathscr{G}^n_{1,i}\mathbb{P}^n_{1,i})$ and $(\Omega^n_{2,i}, \mathscr{G}^n_{2,i}, \mathbb{P}^n_{2,i})$ that, respectively, support a Poisson process A^n_i with a given rate λ^n_i and a Poisson process S^n_i with rate μ^n_i. The process A^n_i counts the number of arrivals to the ith buffer, and $S^n_i(t)$ stands for the number of service completions of customers of class i had service been given to class i for t units of time. Denote $A^n = (A^n_i)^I_{i=1}$ and $S^n = (S^n_i)^I_{i=1}$.

As is rigorously clarified in Section 2.2, we consider a different level of uncertainty about each of the $2I$ components of (A^n, S^n) and, thus, set the following reference probability space that supports these processes

$$(\Omega^n, \mathscr{G}^n, \mathbb{P}^n) := \left(\prod_{i=1}^{I}(\Omega^n_{1,i} \times \Omega^n_{2,i}), \otimes^I_{i=1}(\mathscr{G}^n_{1,i} \otimes \mathscr{G}^n_{2,i}), \prod_{i=1}^{I}(\mathbb{P}^n_{1,i} \times \mathbb{P}^n_{2,i})\right),$$

where $\otimes^I_{i=1}(\mathscr{G}^n_{1,i} \otimes \mathscr{G}^n_{2,i}) = (\mathscr{G}^n_{1,1} \otimes \mathscr{G}^n_{2,1}) \otimes \dots \otimes (\mathscr{G}^n_{1,I} \otimes \mathscr{G}^n_{2,I})$.

By the structure of the probability space, under the measure $\mathbb{P}^n$, the $2I$ processes $A^n_1, S^n_1, \dots, A^n_I, S^n_I$ are mutually independent. Moreover, the distribution of A^n_i (S^n_i) under $\mathbb{P}^n$ is identical to its distribution under $\mathbb{P}^n_{1,i}$ ($\mathbb{P}^n_{2,i}$).

Let $U^n = (U^n_i)^I_{i=1}$ be an RCLL process taking values in $\mathbb{U} = \{x = (x_1, \dots, x_I) \in [0,1]^I : \sum x_i \le 1\}$, where its ith component $U^n_i(t)$ represents the fraction of effort the server dedicates to class i (recall that process sharing is allowed). Then

$$T^n_i(t) := \int_0^t U^n_i(s)ds, \qquad t \in \mathbb{R}_+, \tag{1}$$

gives the cumulative effort in the interval $[0, t]$ the server dedicates to class i customers, and the number of class i job completions by time t is, thus, $S^n_i(T^n_i(t))$. This is a Cox process with rate $\mu^n_i U^n_i$.

We allow rejections of customers (only) upon arrival, and a rejected customer never returns to the system. The number of rejections from class i until time t is denoted by $R^n_i(t)$.

Denote by $X^n_i(t)$ the number of class i customers in the system at time t. Then, the balance equation is given by

$$X^n_i(t) = X^n_i(0) + A^n_i(t) - S^n_i(T^n_i(t)) - R^n_i(t), \qquad t \in \mathbb{R}_+. \tag{2}$$

For simplicity, we assume that $X^n_i(0)$ is deterministic. We use the notation $L^n = (L^n_i)^I_{i=1}$ for $L^n = X^n, R^n, T^n$. Because U^n is an RCLL process and, by construction, so are A^n and S^n, we conclude that X^n and R^n are RCLL as well. The capacity of buffer i is denoted by $\hat{b}^n_i$ and is nondecreasing with n.

We now introduce the diffusion scaling and the heavy-traffic condition. First, we assume that

$$\lambda^n_i := \lambda_i n + \hat{\lambda}_i n^{1/2} + o(n^{1/2}), \qquad \mu^n_i := \mu_i n + \hat{\mu}_i n^{1/2} + o(n^{1/2}), \tag{3}$$

for some fixed constants $\lambda_i, \mu_i \in (0,\infty)$ and $\hat{\lambda}_i, \hat{\mu}_i \in \mathbb{R}$. Moreover, the system is assumed to be *critically loaded*; that is, $\sum_{i=1}^I \rho_i = 1$, where $\rho_i := \lambda_i/\mu_i$, $i \in [I]$. Also, set up $\hat{b}_i^n := \hat{b}_i n^{1/2}$ for some constant $\hat{b}_i \in (0,\infty)$, $i \in [I]$.

The process (U^n, R^n) is regarded as a control in the nth system and is now given rigorously.

Definition 1 (Admissible Control for the Decision Maker, QCP). An admissible control for the minimizer for any initial state $X^n(0)$ is a process (U^n, R^n) taking values in $\mathbb{U} \times \mathbb{R}_+^I$ that satisfies the following:

(i) The process (U^n, R^n) is adapted to the filtration $\mathcal{G}_t^n = \mathcal{G}^n(t) := \sigma\{A_i^n(s), S_i^n(T_i^n(s)), i \in [I], s \le t\}$ and has RCLL sample paths.

(ii) The processes R^n are nondecreasing with jumps of sizes $1/\sqrt{n}$.

(iii) For each $i \in [I]$ and $t \in \mathbb{R}_+$,

$$X_i^n(t) = 0 \quad \text{implies} \quad U_i^n(t) = 0.$$

(iv) The buffer constraints $X_i^n(t) \in [0, \hat{b}_i^n]$, $t \in \mathbb{R}_+$, and $i \in [I]$, hold.

The first condition expresses the fact that the DM makes the decision based on past observations. The second condition follows because rejections are accumulated. The third condition asserts that service cannot be given to an empty buffer. We denote the set of admissible controls for the DM in the nth system by $\mathcal{A}^n(X^n(0))$.

2.2. The Optimization Problem with Model Uncertainty

Recall that we consider a DM who is uncertain about the underlying reference probability measure $\mathbb{P}^n$, or loosely speaking, the DM suspects that the rates/intensities $\{\lambda_i^n\}_{i=1}^I$ and $\{\mu_i^n\}_{i=1}^I$ may be unspecified. Therefore, instead of optimizing under a single probability space, the DM considers a reference probability measure $\mathbb{P}^n$ and a set of candidate measures (provided in the sequel) and penalizes their deviation from $\mathbb{P}^n$. The penalization is done by using a discounted variant of the Kullback–Leibler divergence. More explicitly, the QCP is set up as a stochastic game that models a type of worst-case scenario. The players are the DM who chooses a policy that minimizes a cost and an adverse player also referred to as the *nature* or *maximizer*, who has access to the policy chosen by the minimizer and to the history. The nature is penalized for deviating from the reference model.

We are interested in a cost that accounts for the scaled queue lengths and rejections in addition to the uncertainty about the model. Denote the scaled head count process and the scaled rejection count by

$$\hat{X}^n(t) := n^{-1/2} X^n(t) \qquad \text{and} \qquad \hat{R}^n(t) := n^{-1/2} R^n(t), \qquad t \in \mathbb{R}_+.$$

Fix a discount factor $\varrho > 0$, vectors of holding and rejection costs $\hat{h}, \hat{r} \in (0,\infty)^I$, and ambiguity parameters $\kappa := (\kappa_{1,i}, \kappa_{2,i})_{i=1}^I \in (0,\infty)^{2I}$. The DM chooses an admissible control (U^n, R^n) in an attempt to minimize the following robust cost:[1]

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

where the set of candidate measures $\hat{\mathfrak{Q}}^n(X^n(0))$ consists of all the product measures $\hat{\mathbb{Q}}^n = \prod_{i=1}^I (\hat{\mathbb{Q}}_{1,i}^n \times \hat{\mathbb{Q}}_{2,i}^n)$ that satisfy some integrability condition (the exact definition is provided later), and

$$\begin{aligned} &J^n(X^n(0), U^n, R^n, \hat{\mathbb{Q}}^n; \kappa) := \\ &\mathbb{E}^{\hat{\mathbb{Q}}^n}\left[\int_0^\infty e^{-\varrho t}\left(\hat{h}\cdot \hat{X}^n(t)dt + \hat{r}\cdot d\hat{R}^n(t)\right)\right] - \sum_{i=1}^I \frac{1}{\kappa_{1,i}} L_1^\varrho(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) - \sum_{i=1}^I \frac{1}{\kappa_{2,i}} L_2^\varrho(\hat{\mathbb{Q}}_{2,i}^n \| \mathbb{P}_{2,i}^n), \end{aligned} \tag{4}$$

and

$$\begin{aligned} L_1^\varrho(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) &:= \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^n}\left[\int_0^\infty \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_{1,i}^n(t)}{d\mathbb{P}_{1,i}^n(t)} dt\right], \\ L_2^\varrho(\hat{\mathbb{Q}}_{2,i}^n \| \mathbb{P}_{2,i}^n) &:= \mathbb{E}^{\hat{\mathbb{Q}}_{2,i}^n}\left[\int_0^\infty \varrho e^{-\varrho t} \log \frac{d\hat{\mathbb{Q}}_{2,i}^n(t)}{d\mathbb{P}_{2,i}^n(t)} dT_i^n(t)\right]. \end{aligned} \tag{5}$$

The first component in the cost function includes the holding and rejection costs under the measure $\hat{\mathbb{Q}}^n$. The last two sums in (4) are referred to as the *change of measure penalties*. Although the second penalty depends on the control of the DM through the integrator dT_i^n, the same asymptotic results hold if we assume a penalty with integration dt. The main reasons for choosing this form is because it simplifies the analysis.

When $\kappa_{j,i}$ is "small" ("big"), we say that there is a weak (strong) ambiguity about the rates of the processes A^n_i and $S^n_i(T^n_i) := S^n_i(T^n_i(\cdot))$. 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 maximizer would prefer to choose measures $\hat{\mathbb{Q}}^n_{j,i}$ close to $\mathbb{P}^n_{j,i}$ (in the divergence sense) and, as a consequence, also the relevant expectations. However, one needs to make sure that the total punishment given by $\frac{1}{\kappa_{j,i}} L^{\varrho}_j(\hat{\mathbb{Q}}^n_{j,i} \| \mathbb{P}^n_{j,i})$ is also small. In Cohen (2018), theorems 5.1 and 5.2, we show that, as the ambiguity parameters converge to zero, the stochastic differential games, which are provided in the same paper, converge to the risk-neutral BCP studied in Atar and Shifrin (2014). Therefore, our problem indeed models ambiguity with respect to (w.r.t.) the risk-neutral model.

The set of candidate measures $\hat{\mathfrak{Q}}^n(X^n(0))$ consists of all the product measures $\hat{\mathbb{Q}}^n = \prod_{i=1}^I (\hat{\mathbb{Q}}^n_{1,i} \times \hat{\mathbb{Q}}^n_{2,i})$ that, for every $i \in [I]$ and $t \in \mathbb{R}_+$, the Radon–Nikodym derivatives satisfy

$$\begin{aligned} \frac{d\hat{\mathbb{Q}}^n_{1,i}(t)}{d\mathbb{P}^n_{1,i}(t)} &= \exp\left\{ \int_0^t \log\left(\frac{\psi^n_{1,i}(s)}{\lambda^n_i} \right) dA^n_i(s) - \int_0^t (\psi^n_{1,i}(s) - \lambda^n_i) ds \right\}, \\ \frac{d\hat{\mathbb{Q}}^n_{2,i}(t)}{d\mathbb{P}^n_{2,i}(t)} &= \exp\left\{ \int_0^t \log\left(\frac{\psi^n_{2,i}(s)}{\mu^n_i} \right) dS^n_i(T^n_i(s)) - \int_0^t (\psi^n_{2,i}(s) - \mu^n_i) dT^n_i(s) \right\}, \end{aligned} \tag{6}$$

for $\{\mathcal{G}^n_t\}$-predictable, measurable, and positive processes $\psi^n_{j,i}$, $j \in \{1,2\}$, satisfying $\int_0^t \psi^n_{j,i}(s) ds < \infty$ $\mathbb{P}$ a.s. for every $t \in \mathbb{R}_+$. These conditions assure, first, that the right-hand sides (r.h.s.) in (6) are $\mathbb{P}^n_{j,i}$ martingales, $j \in \{1,2\}$, and second, that, under the measure $\hat{\mathbb{Q}}^n_{1,i}$ ($\hat{\mathbb{Q}}^n_{2,i}$), A^n_i ($S^n_i(T^n_i)$) is a counting process with intensity $\psi^n_{1,i}$ ($\psi^n_{2,i} U^n_i$). Notice also that, under the measures $\hat{\mathbb{Q}}^n_{j,i}$, $j \in \{1,2\}$, the critical load condition might be violated because we do not restrict the intensities $\{\hat{\psi}^n_{j,i}\}_{j,i,n}$ in such a way. However, as we detail in Section 4.4, such changes of measures are too costly and are avoided by the maximizer so that, as a consequence, on average, the critical load condition is preserved.

Finally, the value function is, thus, given by

$$V^n(x;\kappa) := \inf_{(U^n,R^n) \in \mathcal{A}^n(x)} \sup_{\hat{\mathbb{Q}}^n \in \hat{\mathfrak{Q}}^n(x)} J^n(x, U^n, R^n, \hat{\mathbb{Q}}^n; \kappa).$$

Remark 1.

(i) Notice that the r.h.s. of (6) are martingales, and therefore, there exist probability measures $\hat{\mathbb{Q}}^n_{j,i}$, $j \in \{1,2\}$, such that $\hat{\mathbb{Q}}^n_{j,i}|_{\mathcal{G}^n_t}$ satisfies (6) for all $t \in \mathbb{R}_+$, see Stroock (1987), lemma 4.2.

(ii) From the presentations in (6), it follows that one may think of the maximizer as choosing rates rather than choosing a measure. The reason is that the rates (λ^n_i, μ^n_i) under the reference measure $\mathbb{P}$ are replaced by the rates $(\psi^n_{1,i}(\cdot), \psi^n_{2,i}(\cdot))$, which are allowed to be time dependent. Including time-dependent change of measures/rates is more realistic because, in real-world systems, one has no guarantee that the rates remain fixed over time, and moreover, we show that a $c\mu$ type of policy remains optimal under this general structure.

3. The BCP

3.1. The RSDG

In Cohen (2018), we provided two equivalent games: an I-dimensional game that captures the dynamics of the buffers and a one-dimensional game the dynamics of which approximate the workload process. For the forthcoming analysis, we use only the latter. This section is devoted to the description and the relevant properties of this game.

To derive the game, we introduce some scaled processes and constants in addition to $\hat{X}^n$ and $\hat{R}^n$ introduced earlier. Set for every $i \in [I]$ and $t \in \mathbb{R}_+$

$$\begin{aligned} \hat{A}^n_i(t) &:= n^{-1/2}(A^n_i(t) - \lambda^n_i t), & \hat{S}^n_i(t) &:= n^{-1/2}(S^n_i(t) - \mu^n_i t), \\ \hat{Y}^n_i(t) &:= \mu^n_i n^{-1/2}(\rho_i t - T^n_i(t)), & \hat{m}^n_i &:= n^{-1/2}(\lambda^n_i - \rho_i \mu^n_i). \end{aligned} \tag{7}$$

We use the notation $\hat{X}^n = (\hat{X}^n_i)_{i=1}^I$ and similarly for $\hat{R}^n, \hat{A}^n, \hat{S}^n, \hat{Y}^n$, and $\hat{m}^n$. Denote also $\hat{S}^n(T^n) = (\hat{S}^n_i(T^n_i(\cdot)))_{i=1}^I$. The scaled version of (2) is given by

$$\hat{X}^n(t) = \hat{X}^n(0) + \hat{m}^n t + \hat{A}^n(t) - \hat{S}^n(T^n(t)) + \hat{Y}^n(t) - \hat{R}^n(t), \qquad t \in \mathbb{R}_+. \tag{8}$$

An admissible policy satisfies

$$\hat{X}^n(t) \in \mathscr{X} := \prod_{i=1}^{I}[0, \hat{b}_i], \qquad t \in \mathbb{R}_+, \quad \mathbb{P}^n\text{-a.s.} \tag{9}$$

Multiplying both sides of (8) by the workload vector

$$\theta^n := (n/\mu_1^n, \ldots, n/\mu_I^n). \tag{10}$$

we get

$$\theta^n \cdot \hat{X}^n(t) = \theta^n \cdot \hat{X}^n(0) + \theta^n \cdot \hat{m}^n t + \theta^n \cdot \Big(\hat{A}^n(t) - \hat{S}^n(T^n(t))\Big) + \theta^n \cdot \hat{Y}^n(t) - \theta^n \cdot \hat{R}^n(t).$$

Under the reference measure $\mathbb{P}^n$, $\{(\hat{A}^n, \hat{S}^n)\}_n$ weakly converges a $2I$-dimensional $(0, \tilde{\sigma})$-Brownian motion, where $\tilde{\sigma} := \mathrm{Diag}\,(\lambda_1^{1/2}, \ldots, \lambda_I^{1/2}, \mu_1^{1/2}, \ldots, \mu_I^{1/2})$. As we show rigorously in the proof of Lemma 5, $\hat{Y}^n$ is of order one as $n \to \infty$. Hence, its definition implies that $T^n(t) \to (\rho_1, \ldots, \rho_I)t$, $t \in \mathbb{R}_+$, and therefore, under $\mathbb{P}^n$, $\{\hat{A}^n - \hat{S}^n(T^n)\}_n$ weakly converges to an I-dimensional $(0, \hat{\sigma})$-Brownian motion, where

$$\hat{\sigma} = (\hat{\sigma}_{ij}) := \mathrm{Diag}\Big((2\lambda_1)^{1/2}, \ldots, (2\lambda_I)^{1/2}\Big).$$

Hence, $\theta^n \cdot (\hat{A}^n(t) - \hat{S}^n(T^n(t)))$ can be approximated by a one-dimensional Brownian motion. This last two approximations suggest that the $2I$ ambiguity parameters $\{\kappa_{i,j}\}_{i,j}$ collapse into I parameters $\{\varepsilon_i\}_i$, which, in turn, are folded into a single parameter ε as follows:

$$\hat{\varepsilon}_i := \frac{1}{2}(\kappa_{1,i} + \kappa_{2,i}), \qquad i \in [I], \qquad \text{and} \qquad \varepsilon := \frac{1}{\sigma^2}\sum_{i=1}^{I}(\theta\hat{\sigma})_i^2 \hat{\varepsilon}_i. \tag{11}$$

To introduce the game, we need the following notation. Set the vectors

$$\theta := (\mu_1^{-1}, \ldots, \mu_I^{-1}), \qquad \hat{b} = (\hat{b}_1, \ldots, \hat{b}_I), \tag{12}$$

and the scalars

$$x_0 := \theta \cdot \hat{x}_0, \quad m := \theta \cdot \hat{m}, \quad \sigma := \|\theta\hat{\sigma}\|, \quad b := \theta \cdot \hat{b}. \tag{13}$$

We are now ready to define the RSDG.

Definition 2 (Admissible Controls, RSDG). An admissible control for the minimizer for any initial state $x_0 \in [0, b]$ is a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, \mathbb{P})$ that supports a one-dimensional standard Brownian motion B and a process (Y, R) taking values in $\mathbb{R}_+^2$ with RCLL sample paths, both adapted to the filtration $\{\mathscr{F}_t\}$, and satisfies the following properties:

(i) For every $0 \le s < t$, $B(t) - B(s)$ is independent of $\mathscr{F}_s$ under $\mathbb{P}$.
(ii) Y and R are nonnegative and nondecreasing.
(iii) The controlled process satisfies $X(t) = x_0 + mt + \sigma B(t) + Y(t) - R(t)$, and $X(t) \in [0, b]$, $t \in \mathbb{R}_+$, $\mathbb{P}$-*a.s.*

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

$$\frac{d\mathbb{Q}(t)}{d\mathbb{P}(t)} = \exp\left\{\int_0^t \psi(s)dB(s) - \frac{1}{2}\int_0^t \psi^2(s)ds\right\}, \quad t \in \mathbb{R}_+, \tag{14}$$

for an $\{\mathscr{F}_t\}$-progressively measurable process ψ satisfying

$$\mathbb{E}^{\mathbb{P}}\left[\int_0^\infty e^{-\varrho s}\psi^2(s)ds\right] < \infty \quad \text{and} \quad \mathbb{E}^{\mathbb{P}}\left[e^{\frac{1}{2}\int_0^t \psi^2(s)ds}\right] < \infty \quad \text{for every } t \in \mathbb{R}_+. \tag{15}$$

Denote by $\mathscr{A}(x_0)$ ($\mathscr{Q}(x_0)$) the set of all admissible controls for the minimizer (maximizer) given the initial condition x_0. In Cohen (2018), equation 2.29, it is argued that the controlled process can alternatively be written as

$$X(t) = x_0 + mt + \int_0^t \sigma\psi(s)ds + \sigma B^{\mathbb{Q}}(t) + Y(t) - R(t), \quad t \in \mathbb{R}_+, \tag{16}$$

where $B^{\mathbb{Q}}(t) := B(t) - \int_0^t \psi(s)ds$, $t \in \mathbb{R}_+$, is an $\{\mathscr{F}_t\}$-one-dimensional standard Brownian motion under $\mathbb{Q}$. This form serves us in the asymptotic analysis.

The cost associated with the initial condition x_0 and the controls (Y,R) and $\mathbb{Q}$ is given by

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

where

$$h(x):=\min\{\hat{h}\cdot\xi:\xi\in\mathcal{X},\ \theta\cdot\xi=x\},\tag{17}$$

$$r:=\min\{\hat{r}\cdot q:q\in\mathbb{R}^I_+,\ \theta\cdot q=1\},\tag{18}$$

and

$$L^{\varrho}(\mathbb{Q}\|\mathbb{P}):=\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty \varrho e^{-\varrho t}\log\frac{d\hat{\mathbb{Q}}(t)}{d\mathbb{P}(t)}dt\right]=\mathbb{E}^{\mathbb{Q}}\left[\int_0^\infty e^{-\varrho t}\hat{\psi}^2(t)dt\right],\tag{19}$$

where the second inequality is given in Cohen (2018), equation 2.23. The latter form turns out to be useful in the approximation procedure, see (50) and (93). Notice that the scalar ε folds all the ambiguity parameters $\{\kappa_{i,j}\}_{i,j}$ and all the $2I$ penalties from (4) are collapsing into one penalty. 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 Atar and Shifrin (2014), it is shown that there is $i^*\in[I]$ such that,

$$r=r_{i^*}\mu_{i^*}:=\min\{\hat{r}_i\mu_i:i\in[I]\}.\tag{20}$$

The index i^* stands for the class with the smallest rejection cost, weighted with the mean service rate. In fact, under the candidate asymptotically nearly optimal policy, presented in Section 4.1, unless a buffer is overloaded, an event that happens with vanishing probability, rejections are performed only from class i^*.

The value function is given by

$$V(x_0;\varepsilon)=\inf_{(Y,R)\in\mathcal{A}(x_0)}\ \sup_{\mathbb{Q}\in\mathfrak{Q}(x_0)} J(x_0,Y,R,\mathbb{Q};\varepsilon).\tag{21}$$

3.2. Properties of the Game

The RSDG admits a simple optimal strategy for the minimizer 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 $\beta>0$. 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) ζ_1 and ζ_2 are nondecreasing, $\zeta_1(0-)=\zeta_2(0-)=0$, and

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

where $\mathbb{1}_F(x)=1$ if x belongs to the set F and zero otherwise. We denote by $\Gamma_{[0,\beta]}(\eta)=(\Gamma^1_{[0,\beta]},\Gamma^2_{[0,\beta]},\Gamma^3_{[0,\beta]})(\eta)=(\chi,\zeta_1,\zeta_2)$. See Kruk et al. (2007) for existence and uniqueness of the solution and continuity and further properties of the map. In particular, we have the following:

Lemma 1. *There exists a constant $c_S>0$ such that, for every $t>0$, $\beta,\tilde{\beta}>0$, and $\omega,\tilde{\omega}\in\mathcal{D}(\mathbb{R}_+,\mathbb{R})$,*

$$\|\Gamma_{[0,\beta]}(\omega)(t)-\Gamma_{[0,\tilde{\beta}]}(\tilde{\omega})(t)\|_t\leq c_S(\|\omega-\tilde{\omega}\|_T+|\beta-\tilde{\beta}|).$$

Definition 3. Fix $x_0,\beta\in[0,b]$. The strategy (Y,R) is called a β-reflecting strategy if, for every $\eta\in\mathcal{C}(\mathbb{R}_+,\mathbb{R})$, one has $(X,Y,R)(\eta)=\Gamma_{[0,\beta]}(\eta)$ with $X(0)=x_0$.

The next proposition summarizes the necessary results for the analysis in the sequel.

Proposition 1 (Cohen (2018, theorems 3.1 and 4.1)). *Fix $\varepsilon\in(0,\infty)$. The value function $V(\cdot):=V(\cdot;\varepsilon)$ is twice continuously differentiable and satisfies $0\leq V'\leq r$. Moreover, set the parameter*

$$\beta_\varepsilon=\inf\{x\in(0,b]:V'(x)=r\}\wedge b\tag{22}$$

and the measure $\mathbb{Q}_V = \mathbb{Q}_{V(\cdot;\varepsilon)}$ associated with $\psi_V(t) := \varepsilon\sigma V'(X(t))$ via (14). Then, the reflecting strategy $(Y_{\beta_\varepsilon}, R_{\beta_\varepsilon})$ and the measure $\mathbb{Q}_V$ form an equilibrium in the sense that

$$V(x_0) = \sup_{\mathbb{Q}\in\mathfrak{Q}(x_0)} J(x_0, Y_{\beta_\varepsilon}, R_{\beta_\varepsilon}, \mathbb{Q}; \varepsilon) = J(x_0, Y_{\beta_\varepsilon}, R_{\beta_\varepsilon}, \mathbb{Q}_V; \varepsilon) = \inf_{(Y,R)\in\mathcal{A}(x_0)} J(x_0, Y, R, \mathbb{Q}_V; \varepsilon).$$

In Cohen (2018), it is shown that the value function V is continuous and decreasing in the ambiguity parameter ε and that the reflecting barrier β_ε is continuous in ε. The meaning of the set of equations in the proposition is that, by using the reflecting strategy $(Y_{\beta_\varepsilon}, R_{\beta_\varepsilon})$, the best the maximizer can do is choosing the measure $\mathbb{Q}_V$ and vice versa. The optimality of $\mathbb{Q}_V$ serves us in Section 4.2, where we equip the maximizer in the QCP with a measure driven by $\mathbb{Q}_V$. The reflecting strategy $(Y_{\beta_\varepsilon}, R_{\beta_\varepsilon})$ is used to show the state-space collapse in Section 4.3.2.

4. Asymptotic Analysis

4.1. Nearly Optimal Policy

In this section, we present a candidate policy that is shown to be nearly optimal. The idea of the policy is to give the least priority to the buffer with the cheapest holding cost and let it fill up until it is *almost* full. If the workload increases, we keep the cheapest buffer full and start filling the next cheapest buffer by less prioritizing it and so on. This is done with almost no rejections unless the critical workload level β_ε (see Proposition 1) is reached, in which case rejections occur and only from the buffer with the cheapest rejection cost.

4.1.1. The Minimizing Curve. Before presenting the candidate policy, we provide a *minimizing curve* that the dynamics in the QCP $\hat{X}^n$ asymptotically trace. Its structure stems from the equilibrium strategy in the multidimensional stochastic differential game provided in Cohen (2018). Because of the discrete nature of the QCP, we take this curve to be bounded away from (yet close to) the boundary of $\Pi_{i=1}^{I}[0, \hat{b}_i/\mu_i]$. This curve is a key element in establishing the state-space collapse property. Without loss of generality, assume that

$$h_1\mu_1 \geq h_2\mu_2 \geq \cdots \geq h_I\mu_I, \tag{23}$$

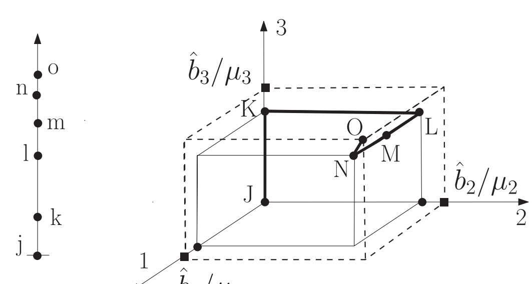
Fix $\delta_0 > 0$, and let $\hat{a} = (\hat{a}_1, \ldots, \hat{a}_I)$ be given by $\hat{a}_i = \hat{b}_i - \delta_0$, $i \in [I]$, and $a = \beta_\varepsilon \wedge (\theta \cdot \hat{a}) < b$. Note that, if δ_0 is sufficiently small, then $a = \beta_\varepsilon$ (unless $\beta_\varepsilon = b$). The difference between the intercepts on axis $i = 1, 2, 3$ in Figure 1 equals δ_0/μ_i. Hence, when $\delta_0 \to 0$, the solid box converges to the dotted one. We now rigorously define a function that maps workload values to states of the Ith dimensional system and then explain its structure. Set the function $\gamma^a : [0, b] \to \mathcal{X}$ as follows. For $x \in [0, \theta\cdot\hat{a})$, the variables $j = j(x)$ and $v = v(x)$ are determined via

$$x = \sum_{i=j+1}^{I} \theta_i\hat{a}_i + \theta_j v, \qquad j \in [I], \quad v \in [0, \hat{a}_j), \tag{24}$$

and let

$$\gamma^a(x) = \sum_{i=j+1}^{I} \hat{a}_i e_i + v e_j. \tag{25}$$

Figure 1. Case in Which $I = 3$, $\hat{h} = (1, 5/2, 3/2)$, $\mu = (3, 1, 3/2)$, and $(\hat{a}_1, \hat{a}_2, \hat{a}_3) = (4, 7, 6)$



Notes. The graph to the left stands for the workload levels. The curve of the function γ^a is in bold in the graph to the right. The workload levels with the lowercase letters are $j = 0, k = \hat{a}_3/\mu_3 = 4, l = \hat{a}_3/\mu_3 + \hat{a}_2/\mu_2 = 11, m = \hat{a}_3/\mu_3 + \hat{a}_2/\mu_2 + 1 = 12$, and $n = \hat{a}_3/\mu_3 + \hat{a}_2/\mu_2 + \hat{a}_1/\mu_1 = a = 37/3$. They, respectively, correspond to the uppercase letters $J = (0,0,0), K = (0,0,\hat{a}_3/\mu_3) = (0,0,4), L = (0,\hat{a}_2/\mu_2,\hat{a}_3/\mu_3) = (0,7,4), M = (1,\hat{a}_2/\mu_2,\hat{a}_3/\mu_3) = (1,7,4)$, and $N = (\hat{a}_1/\mu_1,\hat{a}_2/\mu_2,\hat{a}_3/\mu_3) = (4/3,7,4)$. Pay attention that, between N and O, γ^a is defined via a linear interpolation. If, for example, $m = \beta_\varepsilon = a$, the scaled queues fluctuate close to the bold curve between points J and M.

On the interval $[\theta \cdot \hat{a}, b]$, define γ^a as the linear interpolation between the points $(\theta \cdot \hat{a}, \hat{a})$ and $(b, \hat{b})$. Notice that the buffers $j+1, \ldots, I$, which are the cheapest ones (holding cost–wise), are almost full (up to δ_0), buffer j is building up, and the buffers $1, \ldots, j$ are left empty. An illustration of the curve γ^a is provided in Figure 1.

We now set up a holding cost that is associated with workload values through the curve γ^a. Recall the definition of h from (17) and let

$$h^a(x) := \min\{\hat{h} \cdot \xi : \xi \in \mathcal{X}, \theta \cdot \xi = x, \xi_i \le \gamma_i^a(x), i \in [I]\} = \hat{h} \cdot \gamma^a(x), \qquad x \in [0, \theta \cdot a].$$

Also, set

$$\omega_1(\delta_0) = \sup_{[0, \theta \cdot \hat{a}]} |h^a - h|. \tag{26}$$

By the choice of a, it is clear that $\omega_1(0+) = 0$.

4.1.2. The Candidate Policy and the Main Theorem. *4.1.2.1. Rejection Policy.*

In the case in which a class-i arrival occurs at a time t when $\hat{X}_i^n(t-) + n^{-1/2} > \hat{b}_i$, then it is rejected. Such rejections are called *forced rejections*. Whenever $\theta \cdot \hat{X}^n \ge a$, all class-$i^*$ (see the paragraph preceding (20)) arrivals are rejected, and these rejections are called *overload rejections*. Apart from that, no rejections occur from any class.

4.1.2.2. Service Policy. For each $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_I) \in \mathcal{X}$, define the class of low priority

$$\mathcal{L}(\hat{x}) = \max\{i : \hat{x}_i < \hat{a}_i\},$$

provided $\hat{x}_i < \hat{a}_i$ for some i, and set $\mathcal{L}(\hat{x}) = I$ otherwise. The complement set is the set of high-priority classes:

$$\mathcal{H}(\hat{x}) = [I] \setminus \mathcal{L}(\hat{x}).$$

When at least one class among $\mathcal{H}(\hat{x})$ is not empty, the class $\mathcal{L}(\hat{x})$ receives no service, and all classes within $\mathcal{H}(\hat{x})$ that are not empty receive service at a fraction proportional to their traffic intensities. Namely, denote $\mathcal{H}^+(\hat{x}) = \{i \in \mathcal{H}(x) : x_i > 0\}$ and define $\rho'(\hat{x}) \in \mathbb{R}^I$ as

$$\rho_i'(\hat{x}) = \begin{cases} 0, & \text{if } \hat{x} = 0, \\ \dfrac{\rho_i \mathbb{1}_{\{i \in \mathcal{H}^+(\hat{x})\}}}{\sum_{k \in \mathcal{H}^+(\hat{x})} \rho_k}, & \text{if } \mathcal{H}^+(\hat{x}) \neq \emptyset, \\ e_I, & \text{if } \hat{x}_i = 0 \text{ for all } i < I \text{ and } \hat{x}_I > 0, \end{cases} \tag{27}$$

where recall that $e_I = (0, \ldots, 0, 1) \in \mathbb{R}^I$. Note that $\mathcal{H}^+(\hat{x}) = \emptyset$ can only happen if $\hat{x}_i = 0$ for all $i < I$, which is covered by the first and last cases in the preceding display. Then, for each $t \in \mathbb{R}$,

$$U^n(t) = \rho'(\hat{X}^n(t)). \tag{28}$$

Note that, when $\mathcal{H}^+(\hat{x}) \neq \emptyset$,

$$\rho_i'(\hat{x}) > \rho_i \quad \text{for all } i \in \mathcal{H}^+(\hat{x}). \tag{29}$$

That is, all prioritized classes receive a fraction of effort strictly greater than the respective traffic intensity. Also note that $\sum_i U_i^n = 1$ whenever $\hat{X}^n$ is nonzero. This is, therefore, a work-conserving policy.

To illustrate the relationship between the candidate policy and the minimizing curve, we assume that the system starts empty. Then the low priority is given to buffer I (unless in times it is empty) and it fluctuates, and the other buffers get priority and, hence, are almost empty. This is the situation until buffer I exceeds level $\hat{a}_I$. Then, buffer I receives priority and decreases rapidly until it goes below level $\hat{a}_I$ and now again is low prioritized. This way, although buffer $I-1$ is not empty, buffers I and $I-1$ exchange priorities so that buffer I remains around the level $\hat{a}_I$ and buffer $I-1$ fluctuates.

Remark 2. Pay attention that the service policy does not depend on the ambiguity, but rather only on the prioritization of the buffers according to the holding costs. The ambiguity, which is folded into the parameter ε, plays a role in the rejection level a, which equals β_ε unless $\beta_\varepsilon = b$, in which case a is very close to β_ε. As mentioned in the introduction, the candidate policy (again, with a different rejection level) is also optimal under the moderate-deviation heavy-traffic regime. This fact illustrates the robustness of the service policy, especially because our formulation allows for time-varying rates.

Recall the definition of θ^n given in (10).

Theorem 1. *Assuming that* $x_0 := \lim_{n\to\infty} \theta^n \cdot \hat{X}^n(0)$ *exists, then*

$$\lim_{n\to\infty} V^n(X^n(0);\kappa) = V(x_0;\varepsilon). \tag{30}$$

Moreover, for every $n \in \mathbb{N}$, *denote the policy constructed by* $(U^n(a), R^n(a))$. *Then, there is a function* $w_2 : \mathbb{R}_+ \to \mathbb{R}$, *satisfying* $\omega_2(0+) = 0$ *such that*

$$\limsup_{n\to\infty} \sup_{\hat{\mathbb{Q}}^n \in \hat{\mathfrak{Q}}^n(X^n(0))} J^n(X^n(0), U^n(a), R^n(a), \hat{\mathbb{Q}}^n;\kappa) \le V(x_0;\varepsilon) + \omega_2(\beta_\varepsilon - a). \tag{31}$$

Recall that the parameter a can be chosen to be arbitrarily close to β_ε; see the paragraph following (23). Therefore, by a diagonalization argument, one can deduce an asymptotically optimal policy generated from $U^n(a)$. The proof of the theorem takes place in the next two sections. In Section 4.2, we show that the game's value function V bounds from below the liminf of the QCP's value functions. That is,

$$\liminf_{n\to\infty} V^n(X^n(0);\kappa) \ge V(x_0;\varepsilon). \tag{32}$$

In Section 4.3, we prove (31). Together, we obtain (30).

4.2. Proof of (32)

The proof follows along these lines. We equip the maximizer in the QCP with a measure in such a way that its one-dimensional folding onto the workload scale is the equilibrium measure $\mathbb{Q}_V$ from Proposition 1. The DM is equipped with an arbitrary sequence of policies $\{(\hat{Y}^n, \hat{R}^n)\}_n$. Using weak convergence arguments and projecting the I-dimensional processes on the workload vector θ^n (see (10)), we show that, along every converging subsequence, the projected dynamics converges to the dynamics given in (16) with $\psi = \psi_V$ (see Proposition 1) and establish the lower bound for the value functions.

The main difficulty of this argument is that, because $\{(\hat{Y}^n, \hat{R}^n)\}_n$ is an arbitrary sequence of singular controls, one cannot expect this sequence to be tight. Therefore, we use time stretching in order to prove the lower bound.[2] The idea is as follows. Take, for example, the process $\hat{R}^n$. In Section 4.2.1, we stretch the time (using the same transformation for all the processes) and generate a process $\tilde{R}^n$ in such a way that $\{\tilde{R}^n\}$ is Lipschitz-continuous with Lipschitz constant one (over intervals of order one) and, therefore, tight and converges to a process $\tilde{R}$. Then, in Section 4.2.2, we go back to the original scale by an inverse time transformation and get the process $\hat{R}$, which is used in Section 4.2.3 to get the value function of the RSDG.

We start with setting up the maximizer's measure in the QCP. Recall the definition of θ from (12). For every $t \in \mathbb{R}_+$ and $i \in [I]$, set

$$\hat{\psi}_i^n(t) := \frac{(\theta\hat{\sigma})_i \hat{\varepsilon}_i \sigma^2 \varepsilon}{\sum_{j=1}^I (\theta^n\hat{\sigma})_i^2 \hat{\varepsilon}_i} V'(\theta \cdot \hat{X}^n(t-);\varepsilon) \tag{33}$$

$$\hat{\psi}_{1,i}^n(t) := \frac{\kappa_{1,i}\sqrt{2}}{\kappa_{1,i} + \kappa_{2,i}} \hat{\psi}_i^n(t), \qquad \hat{\psi}_{2,i}^n(t) := -\frac{\kappa_{2,i}\sqrt{2}}{(\kappa_{1,i} + \kappa_{2,i})\rho_i^{1/2}} \hat{\psi}_i^n(t), \tag{34}$$

and also

$$\psi_{1,i}^n(t) := \lambda_i^n + \hat{\psi}_{1,i}^n(t)(\lambda_i n)^{1/2} \qquad \text{and} \qquad \psi_{2,i}^n(t) := \mu_i^n + \hat{\psi}_{2,i}^n(t)(\mu_i n)^{1/2}. \tag{35}$$

Let $\{\hat{\mathbb{Q}}_{j,i}^n\}_{j,i,n}$ be the relevant measures defined as in (6). Notice that, from Proposition 1, $0 \le V'(\cdot;\varepsilon) \le r$; hence, it follows that all the processes mentioned in (33) and (34) are uniformly bounded by some constant. Namely, there exists a constant $C_0 > 0$ such that for every $j \in \{1,2\}$, $i \in [I]$, $n \in \mathbb{N}$, and $t \in \mathbb{R}_+$,

$$|\hat{\psi}_{j,i}^n(t)| \le C_0. \tag{36}$$

We now simplify the change of measure penalty from (5). Because $A_i^n(\cdot) - \int_0^{\cdot} \hat{\psi}_{1,i}^n(s)ds$ is a martingale under $\hat{\mathbb{Q}}_{j,i}^n$, we get the following sequence of equations:

$$
\begin{aligned}
&L_1^{\varrho}(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) \\
&\quad = \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^n}\left[\int_0^\infty \rho e^{-\varrho t}\left(\int_0^t \log\left(\frac{\psi_{1,i}^n(s)}{\lambda_i^n}\right) dA_i^n(s) - \int_0^t (\psi_{1,i}^n(s) - \lambda_i^n) ds\right) dt\right] \\
&\quad = \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^n}\left[\int_0^\infty \rho e^{-\varrho t}\left(\int_0^t \log\left(\frac{\psi_{1,i}^n(s)}{\lambda_i^n}\right)(dA_i^n(s) - \psi_{1,i}^n(s)ds)\right.\right. \\
&\quad\quad \left.\left. + \int_0^t \left\{\psi_{1,i}^n(s)\log\left(\frac{\psi_{1,i}^n(s)}{\lambda_i^n}\right) - \psi_{1,i}^n(s) + \lambda_i^n\right\} ds\right) dt\right] \\
&\quad = \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^n}\left[\int_0^\infty \rho e^{-\varrho t}\left(\int_0^t \left\{\psi_{1,i}^n(s)\log\left(\frac{\psi_{1,i}^n(s)}{\lambda_i^n}\right) - \psi_{1,i}^n(s) + \lambda_i^n\right\} ds\right) dt\right] \\
&\quad = \mathbb{E}^{\hat{\mathbb{Q}}_{1,i}^n}\left[\int_0^\infty e^{-\varrho t}\left\{\psi_{1,i}^n(t)\log\left(\frac{\psi_{1,i}^n(t)}{\lambda_i^n}\right) - \psi_{1,i}^n(t) + \lambda_i^n\right\} dt\right],
\end{aligned}
\tag{37}
$$

where the last equality follows by changing the order of integration. Similar calculations apply to the change of measure penalty associated with the service time. Set $y^n = \hat{\psi}_{1,i}^n(t)(\lambda_i n)^{1/2}/\lambda_i^n$. Then, $|y^n| \le C_1 n^{-1/2} + o(n^{-1/2})$. Noticing that $|(1+y^n)\log(1+y^n) - y^n|$ is uniformly bounded over n and recalling (35), we get that there exists a constant $C_2 > 0$, independent of n and t, such that

$$
\sum_{i=1}^I \frac{1}{\kappa_{1,i}} L_1^{\varrho}(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) + \sum_{i=1}^I \frac{1}{\kappa_{2,i}} L_2^{\varrho}(\hat{\mathbb{Q}}_{2,i}^n \| \mathbb{P}_{2,i}^n) \le C_2.
$$

Fix an arbitrary sequence of controls $\{(\hat{Y}^n, \hat{R}^n)\}_n$. We focus only on those $n \in \mathbb{N}$ for which

$$
J^n(X^n(0), U^n, R^n, \hat{\mathbb{Q}}^n; \kappa) < V(x_0; \varepsilon) + 1. \tag{38}
$$

For the n's for which the reversed inequality holds, we obtain $J^n(X^n(0), U^n, R^n, \hat{\mathbb{Q}}^n; \kappa) > V(x_0; \varepsilon)$, and the lower bound holds without even taking a limit. Therefore, for every $t \in \mathbb{R}_+$,

$$
e^{-\varrho t}\mathbb{E}^{\hat{\mathbb{Q}}^n}\left[\hat{r} \cdot \hat{R}^n(t)\right] \le \mathbb{E}^{\hat{\mathbb{Q}}^n}\left[\int_0^t e^{-\varrho s}\hat{r} \cdot d\hat{R}^n(s)\right] < V(x_0; \varepsilon) + C_2 + 1, \tag{39}
$$

where $\hat{\mathbb{Q}}^n = \prod_{i=1}^I (\hat{\mathbb{Q}}_{1,i}^n \times \hat{\mathbb{Q}}_{2,i}^n)$. This property serves us in Lemma 2 when we claim tightness of a time-rescaled version of $\hat{R}^n$.

Recall that the intensity of $S_i^n(T_i^n)$ is $\psi_{2,i}^n U_i^n$ and that $dT_i^n(s) = U_i^n(s)ds$. Then, under the measure $\hat{\mathbb{Q}}^n$, the dynamics of $\hat{X}^n$ can be expressed in the following convenient form:

$$
\begin{aligned}
\hat{X}_i^n(t) = {} & \hat{X}_i^n(0) + \hat{m}_i^n t + \check{A}_i^n(t) - \check{D}_i^n(t) + \hat{Y}_i^n(t) - \hat{R}_i^n(t) \\
& + \lambda_i^{1/2}\int_0^t \hat{\psi}_{1,i}^n(s)ds - \mu_i^{1/2}\int_0^t \hat{\psi}_{2,i}^n(s)dT_i^n(s),
\end{aligned}
\tag{40}
$$

where

$$
\check{A}_i^n(t) := n^{-1/2}\left(A_i^n(t) - \int_0^t \psi_{1,i}^n(s)ds\right), \tag{41}
$$

$$
\check{D}_i^n(t) := n^{-1/2}\left(S_i^n(T_i^n(t)) - \int_0^t \psi_{2,i}^n(s)dT_i^n(s)\right). \tag{42}
$$

Set $\check{A}^n = (\check{A}_1^n, \ldots, \check{A}_I^n)$ and similarly $\check{D}^n = (\check{D}_1^n, \ldots, \check{D}_I^n)$. By standard martingale techniques, $\check{A}^n$ and $\check{D}^n$ are $\mathcal{G}_t^n$-martingales under $\hat{\mathbb{Q}}^n$, where recall that $\mathcal{G}_t^n$ is given in Definition 1(i); see, for example, the arguments given in the proof of theorem 3.4 in Kushner and Martins (1993).

4.2.1. Time Rescaling.
Recall the definition of θ^n from (10). For every $n \in \mathbb{N}$, define

$$
\tau^n(t) := t + \theta^n \cdot \hat{R}^n(t) + \theta^n \cdot \hat{Y}^n(t), \qquad t \in \mathbb{R}_+. \tag{43}
$$

Because $\hat{R}^n$ and $\theta^n \cdot \hat{Y}^n$ are nondecreasing and RCLL (see (7)), it follows that τ^n is strictly increasing and RCLL. Moreover, for every $0 \le s \le t$, $\tau^n(t) - \tau^n(s) \ge t - s$. The *time rescaled process* is given by

$$\tilde{\tau}^n(t) := \inf\{s \ge 0 : \tau^n(s) > t\}, \qquad t \in \mathbb{R}_+.$$

Notice that $\tilde{\tau}^n$ is nondecreasing and continuous. Also, because $\hat{Y}^n$ is continuous and the jumps of $\hat{R}^n$ are of size $\frac{1}{\sqrt{n}}$, we get that there $|\tau^n(\tilde{\tau}^n(t)) - t| \le \|\theta^n\|/\sqrt{n}$ and

$$0 \le \tilde{\tau}^n(t) \le s \qquad \text{if and only if} \qquad \tau^n(s) \ge t \ge 0. \tag{44}$$

Define also the following rescaled processes:

$$\tilde{L}^n(t) := \hat{L}^n(\tilde{\tau}^n(t)), \quad \tilde{A}^n(t) := \check{A}^n(\tilde{\tau}^n(t)), \quad \tilde{D}^n(t) := \check{D}^n(\tilde{\tau}^n(t)), \quad \tilde{T}^n(t) := T^n(\tilde{\tau}^n(t)),$$

for $L = X, Y$, and R.

Lemma 2. *The sequence of processes*

$$\{(\check{A}^n, \check{D}^n, \tilde{A}^n, \tilde{D}^n, \tilde{X}^n, \tilde{\tau}^n, \tilde{Y}^n, \tilde{R}^n, \tilde{T}^n)\}_n$$

is $\mathscr{C}$-tight. Let $(\check{A}, \check{D}, \tilde{A}, \tilde{D}, \tilde{X}, \tilde{\tau}, \tilde{Y}, \tilde{R}, \tilde{T})$ be a limit of a weakly convergent subsequence. Then, for every $t \in \mathbb{R}_+$ one has,

$$\tilde{X}_i(t) = \tilde{X}_i(0) + \hat{m}_i\tilde{\tau}(t) + \hat{\sigma}_{ii}(\theta\hat{\sigma})_i\hat{\varepsilon}_i\int_0^t V'(\theta\cdot\tilde{X}(s);\varepsilon)d\tilde{\tau}(s) + \hat{\sigma}_i\tilde{B}_i(t) + \tilde{Y}_i(t) - \tilde{R}_i(t), \tag{45}$$

$$\tilde{A}(t) = \check{A}(\tilde{\tau}(t)), \qquad \tilde{D}(t) = \check{D}(\tilde{\tau}(t)), \qquad \tilde{T}(t) := (\rho_1, \ldots, \rho_I)\tilde{\tau}(t), \tag{46}$$

where $\tilde{B} = (\tilde{B}_1, \ldots, \tilde{B}_I) = \hat{\sigma}^{-1}(\tilde{A} - \tilde{D})$ is a martingale w.r.t. its own filtration and with quadratic variation $\tilde{\tau}(\cdot)\mathscr{I}$, where $\mathscr{I}$ is the identity matrix of order $I \times I$.

The proof is given in Section 6.

By reducing to a subsequence and by Skorokhod's representation theorem (see Billingsley (1999), theorem 6.7), we may assume without loss of generality that there is a probability space $(\Omega^*, \mathscr{F}^*, \mathbb{Q}^*)$ that supports the sequence of processes $\{(\check{A}^n, \check{D}^n, \tilde{A}^n, \tilde{D}^n, \tilde{X}^n, \tilde{\tau}^n, \tilde{Y}^n, \tilde{R}^n, \tilde{T}^n)\}_n$ and $(\check{A}, \check{D}, \tilde{A}, \tilde{D}, \tilde{X}, \tilde{\tau}, \tilde{Y}, \tilde{R}, \tilde{T})$ such that

$$\lim_{n\to\infty}(\check{A}^n, \check{D}^n, \tilde{A}^n, \tilde{D}^n, \tilde{X}^n, \tilde{\tau}^n, \tilde{Y}^n, \tilde{R}^n, \tilde{T}^n) = (\check{A}, \check{D}, \tilde{A}, \tilde{D}, \tilde{X}, \tilde{\tau}, \tilde{Y}, \tilde{R}, \tilde{T}); \tag{47}$$

$\mathbb{Q}^*$ a.s., uniformly on compacts (u.o.c.). Throughout the rest of Section 4.2, we consider the probability space $(\Omega^*, \mathscr{F}^*, \mathbb{Q}^*)$ and without loss of generality occasionally assume that (47) is at force.

The following lemma states that the time-rescaled process $\tilde{\tau}(t)$ grows to infinity together with t. Loosely speaking, it implies that the processes $\{(\hat{R}^n, \hat{Y}^n)\}_n$ do not explode as $n \to \infty$ and that a right-inverse function $\tilde{\tau}^{-1}$ exists; see the next section. Hence, we can make the transformation back to the original time scale. The lemma plays an important role in the proof of Proposition 2, and its proof is provided in Section 6.

Lemma 3.

$$\lim_{t\to\infty}\tilde{\tau}(t) = \infty, \qquad \mathbb{Q}^*\text{-}a.s.$$

4.2.2. Back to the Original Scale in the Limiting Environment. We now define the inverse of $\tilde{\tau}$, which brings the limit processes back to the original scale. Set

$$\tau(t) := \inf\{s \ge 0 : \tilde{\tau}(s) > t\}, \qquad t \in \mathbb{R}_+.$$

One can verify that τ is right-continuous and strictly increasing. Moreover, $\lim_{t\to\infty}\tau(t) = \infty$ $\mathbb{Q}^*$ a.s. and from Lemma 3, for every $t \in \mathbb{R}_+$, $\tau(t) < \infty$ a.s. Finally, for every $t \in \mathbb{R}_+$, $\tilde{\tau}(\tau(t)) = t$, $\tau(\tilde{\tau}(t)) \ge t$, and

$$0 \le \tilde{\tau}(s) \le t \qquad \text{if and only if} \qquad \tau(t) \ge s \ge 0.$$

The time-transposed processes are defined as follows:

$$X^*(\cdot) := \tilde{X}(\tau(\cdot)), \quad A^*(\cdot) := \tilde{A}(\tau(\cdot)), \quad D^*(\cdot) := \tilde{D}(\tau(\cdot)),$$
$$Y^*(\cdot) := \tilde{Y}(\tau(\cdot)), \quad R^*(\cdot) := \tilde{R}(\tau(\cdot)), \quad T^*(\cdot) := \tilde{T}(\tau(\cdot)).$$

From (45); the equality $\tilde{\tau}(\tau(t)) = t$; and Protter (2004), theorem IV.4.5, we have, for every $t \in \mathbb{R}_+$,

$$X_i^*(t) = X_i^*(0) + m_i t + \hat{\sigma}_{ii}(\theta\hat{\sigma})_i \hat{\varepsilon}_i \int_0^t V'(\theta \cdot X^*(s); \varepsilon) ds + \hat{\sigma}_i B_i^*(t) + Y_i^*(t) - R_i^*(t), \tag{48}$$

where

$$B^* = (B_1^*, \ldots, B_I^*) := \hat{\sigma}^{-1}(A^* - D^*).$$

The relevant filtration in the limiting environment is now provided. Set $\tilde{\mathcal{G}}_t^\sharp = \tilde{\mathcal{G}}^\sharp(t) := \sigma\{(\tilde{X}(s), \tilde{A}(s), \tilde{D}(s), \tilde{Y}(s), \tilde{R}(s)),\ 0 \le s \le t\}$. Notice that, for every $0 \le s < t < \infty$, $\{\tau(s) < t\} = \{\tilde{\tau}(t) > s\} \in \tilde{\mathcal{G}}^\sharp(t)$. Therefore, $\tau(s)$ is an optional time for $\tilde{\mathcal{G}}^\sharp(t)$. From Karatzas and Shreve (1991), corollary 2.4, $\tau(s)$ is a stopping time for the complete right-continuous filtration $\tilde{\mathcal{G}}_t = \tilde{\mathcal{G}}(t) := \tilde{\mathcal{G}}^\sharp(t+) \vee \mathcal{N}$, where $\mathcal{N}$ is the collection of $\mathbb{Q}^*$-null sets. Using the monotonicity of $t \mapsto \tau(t)$, we get that $\mathcal{G}_t^* := \tilde{\mathcal{G}}(\tau(t))$ is a filtration.

Proposition 2. *The process B^* is an I-dimensional standard Brownian motion under the filtration $\mathcal{G}_t^*$.*

The proof uses classical martingale arguments and is given in Section 6.

4.2.3. Asymptotic Lower Bound. We are now ready to analyze the cost function. We start with an upper bound for the limit of the Kullback–Leibler divergences from (5). Recall (37) and consider $y^n = \hat{\psi}_{1,i}^n(t)(\lambda_i n)^{1/2}/\lambda_i^n$. By (33)–(35) and because $0 \le V' \le r$ (Proposition 1), $y^n \ge 0$. Notice that for every $y \ge 0$,

$$(1+y)\log(1+y) - y \le \frac{1}{2}y^2. \tag{49}$$

Because $\lambda_i n/\lambda_i^n \to 1$ as $n \to \infty$, we get from the last display in (37) that

$$L_1^\varrho(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) \le \mathbb{E}^{\mathbb{Q}^*}\left[\frac{1}{2}\int_0^\infty e^{-\varrho t}(\hat{\psi}_{1,i}^n(t))^2 dt\right] + o(1).$$

The same calculation yields,

$$L_2^\varrho(\hat{\mathbb{Q}}_{2,i}^n \| \mathbb{P}_{2,i}^n) \le \mathbb{E}^{\mathbb{Q}^*}\left[\frac{1}{2}\int_0^\infty e^{-\varrho t}(\hat{\psi}_{2,i}^n(t))^2 d\rho_i(t)\right] + \mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho t}(\hat{\psi}_{2,i}^n(t))^2 d(T_i^n(t) - \rho_i t)\right] + o(1).$$

The last expectation is of order $o(1)$. Indeed, by (33) and (34), $(\hat{\psi}_{2,i}^n(t))^2 = c_i\big(V'(\theta \cdot \hat{X}^n(t-)); \varepsilon)\big)^2$ for some $c_i > 0$. From Protter (2004), theorem IV.4.5,

$$\begin{aligned}\int_0^\infty e^{-\varrho t}(\hat{\psi}_{2,i}^n(t))^2 d(T_i^n(t) - \rho_i t) &= \int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}(\hat{\psi}_{2,i}^n(\tilde{\tau}^n(t)))^2 d(T_i^n(\tilde{\tau}^n(t)) - \rho_i\tilde{\tau}^n(t)) \\ &= c_i\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}(V'(\theta \cdot \tilde{X}^n(t-)))^2 d(T_i^n(\tilde{\tau}^n(t)) - \rho_i\tilde{\tau}^n(t)).\end{aligned}$$

From Lemma 2 and Dai and Williams (1996), lemma 2.4, the last integral converges to zero, $\mathbb{Q}^*$ a.s. Although the mentioned lemma is stated for a finite time interval, the discounted cost, the boundedness of V', and the bound $T^n(t) \le t$ allow us to take the integral's upper limit to be infinity. Now the uniform boundedness of the integral implies that the expectation of the last integral converges to zero as well. Using the preceding representations, together with (11), (33), and (34), it follows that

$$\sum_{i=1}^I \frac{1}{\kappa_{1,i}} L_1^\varrho(\hat{\mathbb{Q}}_{1,i}^n \| \mathbb{P}_{1,i}^n) + \sum_{i=1}^I \frac{1}{\kappa_{2,i}} L_2^\varrho(\hat{\mathbb{Q}}_{2,i}^n \| \mathbb{P}_{2,i}^n) \le \frac{1}{2\varepsilon}\mathbb{E}^{Q^*}\left[\int_0^\infty \varepsilon^2\sigma^2(V'(\theta^n \cdot \hat{X}^n(t); \varepsilon))^2 dt\right] + o(1). \tag{50}$$

Notice that $\hat{X}^n$ has only a countable number of jumps during the time interval $[0,\infty)$, and therefore, we could replace $\hat{X}^n(t-)$ by $\hat{X}^n(t)$ without affecting the integral. From (4); the preceding; and Protter (2004), theorem IV.4.5, one has

$$\begin{aligned}
&J^n(X^n(0),U^n,R^n,\hat{\mathbb{Q}}^n;\kappa)\\
&\quad\geq \mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho t}\Big\{[\hat{h}\cdot\hat{X}^n(t)-(\varepsilon\sigma V'(\theta^n\cdot\hat{X}^n(t);\varepsilon))^2/(2\varepsilon)]dt+\hat{r}\cdot d\hat{R}^n(t)\Big\}\right]+o(1)\\
&\quad= \mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}\Big\{[\hat{h}\cdot\hat{X}^n(\tilde{\tau}^n(t))-(\varepsilon\sigma V'(\theta^n\cdot\hat{X}^n(\tilde{\tau}^n(t));\varepsilon))^2/(2\varepsilon)]d\tilde{\tau}^n(t)+\hat{r}\cdot d\hat{R}^n(\tilde{\tau}^n(t))\Big\}\right]+o(1)\\
&\quad= \mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}\Big\{[\hat{h}\cdot\tilde{X}^n(t)-(\varepsilon\sigma V'(\theta^n\cdot\tilde{X}^n(t);\varepsilon))^2/(2\varepsilon)]d\tilde{\tau}^n(t)+\hat{r}\cdot d\tilde{R}^n(t)\Big\}\right]+o(1).
\end{aligned}\tag{51}$$

Recall that $\hat{h}\cdot\tilde{X}^n$ and $\tilde{R}^n$ are nonnegative; that $\tilde{X}^n$ is uniformly bounded; and also that $\tilde{R}^n$ is nondecreasing, continuous, and bounded on any compact time interval. Then, from (47); Dai and Williams (1996), lemma 2.4; and the bounded convergence theorem, we get that $\mathbb{Q}^*$ almost surely for every $s\in\mathbb{R}_+$, one has

$$\begin{aligned}
\liminf_{n\to\infty}\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}\Big\{\hat{h}\cdot\tilde{X}^n(t)d\tilde{\tau}^n(t)+\hat{r}\cdot d\tilde{R}^n(t)\Big\} &\geq \lim_{n\to\infty}\int_0^s e^{-\varrho\tilde{\tau}^n(t)}\Big\{\hat{h}\cdot\tilde{X}^n(t)d\tilde{\tau}^n(t)+\hat{r}\cdot d\tilde{R}^n(t)\Big\}\\
&=\int_0^s e^{-\varrho\tilde{\tau}(t)}\Big\{\hat{h}\cdot\tilde{X}(t)d\tilde{\tau}(t)+\hat{r}\cdot d\tilde{R}(t)\Big\}.
\end{aligned}$$

Taking $s\to\infty$ (the upper limit of the integral) first and then $\mathbb{E}^{\mathbb{Q}^*}$ on both sides, we obtain

$$\liminf_{n\to\infty}\mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}\Big\{\hat{h}\cdot\tilde{X}^n(t)d\tilde{\tau}^n(t)+\hat{r}\cdot d\tilde{R}^n(t)\Big\}\right]\geq\mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho\tilde{\tau}(t)}\Big\{\hat{h}\cdot\tilde{X}(t)d\tilde{\tau}(t)+\hat{r}\cdot d\tilde{R}(t)\Big\}\right].$$

Using now $0\leq V'\leq r$; $(\tilde{X}^n,\tilde{\tau}^n)\Rightarrow(\tilde{X},\tilde{\tau})$; Lemma 3, Dai and Williams (1996), lemma 2.4; and the bounded convergence theorem, we similarly get that

$$\liminf_{n\to\infty}\mathbb{E}^{\mathbb{Q}^*}\left[-\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}(\varepsilon\sigma V'(\theta^n\cdot\tilde{X}^n(t);\varepsilon))^2/(2\varepsilon)d\tilde{\tau}^n(t)\right]\geq\mathbb{E}^{\mathbb{Q}^*}\left[-\int_0^\infty e^{-\varrho\tilde{\tau}(t)}(\varepsilon\sigma V'(\theta^n\cdot\tilde{X}(t);\varepsilon))^2/(2\varepsilon)d\tilde{\tau}(t)\right].$$

Combining the last two bounds,

$$\begin{aligned}
&\liminf_{n\to\infty}\mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho\tilde{\tau}^n(t)}\Big\{[\hat{h}\cdot\tilde{X}^n(t)-(\varepsilon\sigma V'(\theta^n\cdot\tilde{X}^n(t);\varepsilon))^2/(2\varepsilon)]d\tilde{\tau}^n(t)+\hat{r}\cdot d\tilde{R}^n(t)\Big\}\right]\\
&\quad\geq\mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho\tilde{\tau}(t)}\Big\{[\hat{h}\cdot\tilde{X}(t)-(\varepsilon\sigma V'(\theta\cdot\tilde{X}(t);\varepsilon))^2/(2\varepsilon)]d\tilde{\tau}(t)+\hat{r}\cdot d\tilde{R}(t)\Big\}\right]\\
&\quad=\mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho t}\Big\{[\hat{h}\cdot X^*(t)-(\varepsilon\sigma V'(\theta\cdot X^*(t);\varepsilon))^2/(2\varepsilon)]dt+\hat{r}\cdot dR^*(t)\Big\}\right]\\
&\quad\geq\mathbb{E}^{\mathbb{Q}^*}\left[\int_0^\infty e^{-\varrho t}\{[h(\theta\cdot X^*(t))-(\varepsilon\sigma V'(\theta\cdot X^*(t);\varepsilon))^2/(2\varepsilon)]dt+r\cdot d(\theta\cdot R^*(t))\}\right],
\end{aligned}\tag{52}$$

where we used Protter (2004), theorem IV.4.5, to get the equality. Indeed, recall that $\tilde{\tau}(\tau(t))=t$, $X^*(t)=\tilde{X}(\tau(t))$, $R^*(t)=\tilde{R}(\tau(t))$, and Lemma 3. The last inequality follows because, by the definitions of h and r (see (17) and (18) and Cohen (2018), equations 2.45 and 2.46), $h(\theta\cdot X^*(t))\leq\hat{h}\cdot X^*(t)$ and $\int_0^\infty e^{-\varrho t}rd(\theta\cdot R^*(t))\leq\int_0^\infty e^{-\varrho t}\hat{r}\cdot dR^*(t)$.

Denote $B:=(\sum_{i=1}^I(\theta\hat{o})_iB_i^*)/\sigma$ and $(X,Y,R):=(\theta\cdot X^*,\theta\cdot Y^*,\theta\cdot R^*)$. Then, from (45), together with (11), (13), and the limit $x_0=\lim_{n\to\infty}\theta^n\cdot\hat{X}^n(0)$, we have that

$$X(t)=x_0+mt+\sigma\int_0^t\varepsilon\sigma V'(X(s);\varepsilon)ds+\sigma B(t)+Y(t)-R(t),\quad t\in\mathbb{R}_+.\tag{53}$$

From Proposition 2, we get that B is a standard one-dimensional Brownian motion w.r.t. $\mathcal{G}_t^*$. Hence, property (i) of Definition 2 holds. Also, property (ii) follows because $\theta^n\cdot\hat{Y}^n$ and $\hat{R}^n$ are nonnegative and nondecreasing;

see (7). Hence, (19) and the definition of $\mathbb{Q}_V$ in Proposition 1 imply that the last expectation in the sequence of relations in (52) equals $J(x_0, Y, R, \mathbb{Q}_V; \varepsilon)$. Together with (51), we have

$$\liminf_{n\to\infty} J^n(X^n(0), U^n, R^n, \hat{\mathbb{Q}}^n; \kappa) \geq J(x_0, Y, R, \mathbb{Q}_V; \varepsilon) \geq V(x_0; \varepsilon).$$

Because the sequence of policies $\{(U^n, R^n)\}_n$ is arbitrary, it follows that

$$\liminf_{n\to\infty} V^n(X^n(0); \kappa) \geq V(x_0; \varepsilon). \quad \square$$

4.3. Proof of (31)

In this section, we consider a sequence of arbitrary strategies for the maximizer and show that, in the limit, the candidate policy for the DM is bounded above by the value function of the RSDG. That is, asymptotically, the maximizer cannot punish the DM more than the equilibrium in the RSDG.

Recall that the maximizer's strategy can be equivalently formulated as rate perturbations. We start with considering an arbitrary sequence of strategies for the maximizer that is not too costly. That is, the rates are only moderately perturbed by the maximizer. Then, in Section 4.3.2, we prove a state-space collapse property. Specifically, we show that, by using the candidate policy, the dynamics of the buffers' sizes stay close to γ^a from (25) under the perturbed rates. In Section 4.3.3, we asymptotically bound the expected cost; in order to estimate the change of measure penalty, we truncate the rate processes $\{\psi^n_{j,i}\}_{n,j,i}$ and show that, by doing this, the penalty does not change much.

4.3.1. The Maximizer's Perspective. Consider an arbitrary sequence of measures chosen by the maximizer in the QCP, $\{\hat{\mathbb{Q}}^n\}_{n\in\mathbb{N}}$, where each $\hat{\mathbb{Q}}^n \in \hat{\mathfrak{Q}}^n(X^n(0))$. Recall that every measure $\hat{\mathbb{Q}}^n \in \hat{\mathfrak{Q}}^n(X^n(0))$ is associated with the processes $\{\psi^n_{j,i}\}_{j,i}$; see (6). These processes stand for the "new" intensity of the processes $\{A^n_i\}_{i,n}$ and $\{S^n_i(T^n_i)\}_{i,n}$. To simplify the notation and some of the arguments, we consider one probability space $(\Omega, \mathcal{G}, \hat{\mathbb{Q}})$ that supports the processes $\{(A^n_i, S^n_i(T^n_i))\}_{i,n}$ and under which, for every $n \in \mathbb{N}$, the relevant intensities are $\{\psi^n_{j,i}\}_{j,i,n}$. However, occasionally, when we want to emphasize the relevant measure, we use the measures $\{\hat{\mathbb{Q}}^n_{j,i}\}_{n,j,i}$ and $\{\hat{\mathbb{Q}}^n\}_n$.

Notice that the same calculation given in (37) is valid here as well, so

$$\begin{aligned} L^{\varrho}_1(\hat{\mathbb{Q}}^n_{1,i}\|\mathbb{P}^n_{1,i}) &= \mathbb{E}^{\hat{\mathbb{Q}}^n_{1,i}}\left[\int_0^\infty e^{-\varrho t}\left(\psi^n_{1,i}(t)\log\left(\frac{\psi^n_{1,i}(t)}{\lambda^n_i}\right) - \psi^n_{1,i}(t) + \lambda^n_i\right)dt\right],\\ L^{\varrho}_2(\hat{\mathbb{Q}}^n_{2,i}\|\mathbb{P}^n_{2,i}) &= \mathbb{E}^{\hat{\mathbb{Q}}^n_{2,i}}\left[\int_0^\infty e^{-\varrho t}\left(\psi^n_{2,i}(t)\log\left(\frac{\psi^n_{2,i}(t)}{\mu^n_i}\right) - \psi^n_{2,i}(t) + \mu^n_i\right)dT^n_i(t)\right]. \end{aligned} \tag{54}$$

We now show that without any loss, the maximizer can be restricted to measures $\hat{\mathbb{Q}}^n \in \hat{\mathfrak{Q}}^n(X^n(0))$, which are "not too far away" from the reference measure $\mathbb{P}^n$. The idea behind it, as is provided rigorously in the proof, is that, by changing the rate, the rejection cost contributes at most a linear cost, and the penalty for the change of measure is superlinear. Similarly to the bound in (38), without loss of generality, we may and do assume that

$$J^n(X^n(0), U^n(a), R^n(a), \hat{\mathbb{Q}}^n; \kappa) \geq V(x_0; \varepsilon) - 1. \tag{55}$$

Define

$$\hat{\psi}^n_{1,i}(t) := (\lambda_i n)^{-1/2}(\psi^n_{1,i}(t) - \lambda^n_i), \qquad \text{and} \qquad \hat{\psi}^n_{2,i}(t) := (\mu_i n)^{-1/2}(\psi^n_{2,i}(t) - \mu^n_i). \tag{56}$$

Recall that, in the previous section, the maximizer was given a specific strategy for which $\sup_{t,j,i}|\hat{\psi}^n_{j,i}(t)|$ was bounded. Because we consider now a sequence of arbitrary strategies for the maximizer, it does not hold in this case. However, as we show in Proposition 3 and in Section 4.3.3, this is approximately the case.

Proposition 3. *There exists $M > 0$ such that, for every $n \geq n_0$ and every $i \in [I]$,*

$$\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\left(|\hat{\psi}^n_{1,i}(t)|dt + |\hat{\psi}^n_{2,i}(t)|dT^n_i(t)\right) + \int_0^\infty e^{-\varrho t}\left((\hat{\psi}^n_{1,i}(t))^2 dt + (\hat{\psi}^n_{2,i}(t))^2 dT^n_i(t)\right)\right] \leq M$$

and

$$\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\left(\psi^n_{1,i}(t)\log\left(\frac{\psi^n_{1,i}(t)}{\lambda^n_i}\right) - \psi^n_{1,i}(t) + \lambda^n_i\right)dt + \int_0^\infty e^{-\varrho t}\left(\psi^n_{2,i}(t)\log\left(\frac{\psi^n_{2,i}(t)}{\mu^n_i}\right) - \psi^n_{2,i}(t) + \mu^n_i\right)dT^n_i(t)\right] \leq M.$$

Proof. The second bound follows from the first one together with inequality (49). Thus, we only prove the first one. As argued in Atar and Shifrin (2014), for every $t \in \mathbb{R}_+$,

$$\|\hat{R}^n(t)\| \le C(1 + t + \|\hat{A}^n\|_t + \|\hat{D}^n\|_t),$$

where, in the preceding expression and in the rest of the proof, C refers to a finite positive constant that is independent of n and t and which can change from one line to the next. Denote

$$\varphi^n(t) := \sum_{i=1}^{I} \Big(|\hat{\psi}^n_{1,i}(t)| + |\hat{\psi}^n_{2,i}(t)| U^n_i(t) \Big), \qquad t \in \mathbb{R}_+.$$

Recall the definitions of $\check{A}^n, \check{D}^n$ given in (41) and (42), then

$$\|\hat{R}^n(t)\| \le C\bigg(1 + t + \int_0^t \varphi^n(s)ds + \|\check{A}^n\|_t + \|\check{D}^n\|_t\bigg). \tag{57}$$

Applying the Burkholder–Davis–Gundy inequality to $\check{A}^n$ and $\check{D}^n$ and noticing that $n^{-1}|\psi^n_{j,i}| \le C(1 + n^{-1/2}|\hat{\psi}^n_{j,i}|)$, we have

$$\begin{aligned}
\mathbb{E}^{\hat{\mathbb{Q}}}[\|\check{A}^n\|_t^2] &\le Cn^{-1}\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^t \psi^n_{1,i}(s)ds\bigg] \le C\bigg(t + n^{-1/2}\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^t \hat{\psi}^n_{1,i}(s)ds\bigg]\bigg),\\
\mathbb{E}^{\hat{\mathbb{Q}}}[\|\check{D}^n\|_t^2] &\le Cn^{-1}\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^t \psi^n_{2,i}(s)dT^n_i(s)\bigg] \le C\bigg(t + n^{-1/2}\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^t \hat{\psi}^n_{2,i}(s)dT^n_i(s)\bigg]\bigg).
\end{aligned} \tag{58}$$

Hence,

$$\mathbb{E}^{\hat{\mathbb{Q}}}\big[\|\hat{R}^n(t)\|\big] \le C\bigg(1 + t + \mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^t \varphi^n(s)ds\bigg]\bigg). \tag{59}$$

An application of integration by parts yields

$$\int_0^\infty e^{-\varrho t}\hat{r}\cdot d\hat{R}^n(t) = [-\rho^{-1}e^{-\varrho t}\hat{r}\cdot\hat{R}^n(t)]_{t=0}^{t=\infty} + \int_0^\infty \varrho e^{-\varrho t}\hat{r}\cdot\hat{R}^n(t)dt, \tag{60}$$

and changing the order of integration implies

$$\int_0^\infty e^{-\varrho t}\int_0^t \varphi^n(s)ds = \int_0^\infty \varrho e^{-\varrho t}\varphi^n(t)dt.$$

Notice that the first term in (60) is nonpositive. Taking expectation on both sides of it and using the bound (59), we get that

$$\begin{aligned}
\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty e^{-\varrho t}\hat{r}\cdot d\hat{R}^n(t)\bigg] &\le C\bigg(1 + \mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty \varrho e^{-\varrho t}\bigg\{\int_0^t \varphi^n(s)ds\bigg\}dt\bigg]\bigg)\\
&= C\bigg(1 + \mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty \varrho e^{-\varrho t}\varphi^n(t)dt\bigg]\bigg).
\end{aligned}$$

Clearly, $\hat{h}\cdot\hat{X}^n$ is bounded. Then (55) and the last bound yield that

$$\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty e^{-\varrho t}\Big((\hat{\psi}^n_{1,i}(t))^2dt + (\hat{\psi}^n_{2,i}(t))^2dT^n_i(t)\Big)\bigg] \le C_0\bigg(1 + \mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty e^{-\varrho t}\Big(|\hat{\psi}^n_{1,i}(t)|dt + |\hat{\psi}^n_{2,i}(t)|dT^n_i(t)\Big)\bigg]\bigg), \tag{61}$$

for some $C_0 > 0$, independent of n. Because the function $z \mapsto z^2$ is superlinear, there is a constant $C_1 \in \mathbb{R}$ such that for any $z \in \mathbb{R}$, $z^2 \ge C_1 + 2C_0 z$, and thus,

$$\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty e^{-\varrho t}\Big((\hat{\psi}^n_{1,i}(t))^2dt + (\hat{\psi}^n_{2,i}(t))^2dT^n_i(t)\Big)\bigg] \ge C_1 + 2C_0\bigg(\mathbb{E}^{\hat{\mathbb{Q}}}\bigg[\int_0^\infty e^{-\varrho t}\Big(|\hat{\psi}^n_{1,i}(t)|dt + |\hat{\psi}^n_{2,i}(t)|dT^n_i(t)\Big)\bigg]\bigg).$$

Together with (61), we obtain that

$$\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\left(|\hat{\psi}^n_{1,i}(t)|dt + |\hat{\psi}^n_{2,i}(t)|dT^n_i(t)\right)\right] \le 1 - C_1/C_0,$$

and the result holds by another application of (61) and the preceding bound. □

Define

$$\hat{\Psi}^n_{1,i}(\cdot) := \int_0^\cdot \hat{\psi}^n_{1,i}(t)dt, \qquad \text{and} \qquad \hat{\Psi}^n_{2,i}(\cdot) := \int_0^\cdot \hat{\psi}^n_{2,i}(t)dT^n_i(t),$$

and set $\hat{\Psi}^n_j = (\hat{\Psi}^n_{j,i}; i \in [I])$. The next lemma assures that $\{(\hat{\Psi}^n_1, \hat{\Psi}^n_2)\}_n$ has a converging subsequence and that any of its limit points has continuous paths with probability one.

Lemma 4. *The sequence* $\{(\hat{\Psi}^n_1, \hat{\Psi}^n_2)\}_n$ *is* $\mathscr{C}$*-tight.*

The proof is given in Section 6.

4.3.2. State-Space Collapse: Staying Close to the Minimizing Curve. We consider a set of one-dimensional processes. It is obtained by multiplying the scaled processes by θ^n (see (10)). For its definition, denote

$$\hat{W}^n := \check{A}^n - \check{D}^n + \hat{m}^n \cdot, \tag{62}$$

and

$$W^{\sharp,n} := \theta^n \cdot (\hat{W}^n), \quad X^{\sharp,n} := \theta^n \cdot \hat{X}^n, \quad Y^{\sharp,n} := \theta^n \cdot \hat{Y}^n, \quad R^{\sharp,n} := \theta^n \cdot \hat{R}^n. \tag{63}$$

Moreover, set $\psi^{\sharp,n}_1 := \sigma^{-1}\sum_{i=1}^I \theta^n_i \lambda_i^{1/2}\hat{\psi}^n_{1,i}$, $\psi^{\sharp,n}_2 := \sigma^{-1}\sum_{i=1}^I \theta^n_i \mu_i^{1/2}\hat{\psi}^n_{2,i}$, and

$$\Psi^{\sharp,n}_1(\cdot) := \int_0^\cdot \psi^{\sharp,n}_1(s)ds, \qquad \Psi^{\sharp,n}_2(\cdot) := \int_0^\cdot \psi^{\sharp,n}_2(s)dT^n_i(s). \tag{64}$$

The identity from (40) is valid here as well and can be expressed as

$$\hat{X}^n_i(t) = \hat{X}^n_i(0) + \hat{m}^n_i t + \check{A}^n_i(t) - \check{D}^n_i(t) + \hat{Y}^n_i(t) - \hat{R}^n_i(t) + \lambda_i^{1/2}\hat{\Psi}^n_{1,i}(t) - \mu_i^{1/2}\hat{\Psi}^n_{2,i}(t). \tag{65}$$

As a result,

$$X^{\sharp,n}(t) = X^{\sharp,n}(0) + W^{\sharp,n}(t) + Y^{\sharp,n}(t) - R^{\sharp,n}(t) + \sigma(\Psi^{\sharp,n}_1(t) - \Psi^{\sharp,n}_2(t)), \quad t \in \mathbb{R}_+. \tag{66}$$

We now state a couple of results from Atar and Shifrin (2014) that are needed in the next section. The proofs in our case are almost identical yet require some technical modifications, and they are deferred to Section 6. The next two lemmas are needed to prove Proposition 4 and Lemma 7. The next lemma states that, under the candidate policy given in Section 4.1, the busy time of every buffer converges to its traffic intensity.

Lemma 5. *For every* $i \in [I]$, $\{T^n_i\}_n$ *converges u.o.c. to* $\bar{T}_i$, *where* $\bar{T}_i(t) := \rho_i t$, $t \in \mathbb{R}_+$. *Moreover,* $\{(\check{A}^n, \check{D}^n, \hat{W}^n)\}_n$ *is* $\mathscr{C}$*-tight.*

We now define another set of processes generated from the last set. Let τ^n be the first time a forced rejection occurred in the nth system and set

$$\begin{aligned} W^{\circ,n}(\cdot) &:= W^{\sharp,n}(\cdot \wedge \tau^n), \qquad & X^{\circ,n}(\cdot) &:= X^{\sharp,n}(\cdot \wedge \tau^n), \qquad & Y^{\circ,n}(\cdot) &:= Y^{\sharp,n}(\cdot \wedge \tau^n), \\ R^{\circ,n}(\cdot) &:= R^{\sharp,n}(\cdot \wedge \tau^n), & \Psi^{\circ,n}_1(\cdot) &:= \Psi^{\sharp,n}_1(\cdot \wedge \tau^n), & \Psi^{\circ,n}_2(\cdot) &:= \Psi^{\sharp,n}_2(\cdot \wedge \tau^n). \end{aligned} \tag{67}$$

A minor modification from Atar and Shifrin (2014) that can be observed immediately is that we defined $\hat{W}^n$ and $W^{\sharp,n}$ using $\check{A}^n$ and $\check{D}^n$ instead of $\hat{A}^n$ and $\hat{D}^n$.

We continue the proof of (31) with the case that the initial state lies close to the minimizng curve. That is,

$$\lim_{n\to\infty} \hat{X}^n(0) - \gamma^a(x_0) = 0 \qquad \text{and} \qquad \theta^n \cdot \hat{X}^n(0) \in [0, a], \quad n \in \mathbb{N}, \tag{68}$$

where recall that $x_0 := \theta^n \cdot \hat{X}^n(0)$. As argued in Atar and Shifrin (2014), step 5, this condition can be relaxed by considering a stopping time that indicates when the state is "sufficiently close" to the minimizing path. This stopping time converges to zero. Then we may continue from that stopping time in the same way. The proof follows by the same lines as in case that the initial state lies close to the minimizing curve but with heavier notation and is, therefore, omitted.

A modification of the two next results appear in Atar and Shifrin (2014), step 1.

Lemma 6. *For every $n \in \mathbb{N}$,*

$$(a \wedge X^{\circ,n}, Y^{\circ,n}, R^{\circ,n}) = \Gamma_{[0,a]}(X^{\sharp,n}(0) + W^{\circ,n} + \sigma(\Psi_1^{\circ,n} - \Psi_2^{\circ,n}) + E^n), \tag{69}$$

where

$$E^n := (a \wedge X^{\circ,n}) - X^{\circ,n} \Rightarrow 0. \tag{70}$$

Moreover, the sequence $\{(W^{\circ,n}, X^{\circ,n}, Y^{\circ,n}, R^{\circ,n}, \Psi_1^{\circ,n}, \Psi_2^{\circ,n})\}_n$ is $\mathcal{C}$-tight.

Next, we provide a state-space collapse result and claim that the multidimensional process $\hat{X}^n$ lies close to the minimizing curve. Recall the dependency of the candidate policy on the parameter δ_0.

Proposition 4. *For sufficiently small δ_0, the following limit holds u.o.c.:*

$$\hat{X}^n - \gamma^a(X^{\sharp,n}) \overset{\hat{\mathbb{Q}}}{\Longrightarrow} 0,$$

as $n \to \infty$, and moreover, for every $T > 0$, $\lim_{n\to\infty} \hat{\mathbb{Q}}(\tau^n < T) = 0$.

From the limit $\lim_n \hat{\mathbb{Q}}(\tau^n < T) \to 0$, we conclude that $\{(W^{\sharp,n}, X^{\sharp,n}, Y^{\sharp,n}, R^{\sharp,n}, \Psi_1^{\sharp,n}, \Psi_2^{\sharp,n})\}_n$ is $\mathcal{C}$-tight.

4.3.3. Asymptotic Bound for the Costs. In this section, we take advantage of the convergence of the dynamics to the minimizing curve we just argued. We start by showing that the expected holding and rejection components of the cost can be approximated by equivalent components associated with one-dimensional dynamics. Then we approximate the one-dimensional dynamics with simpler dynamics for which the intensities used by the maximizer are truncated in some sense. The difference between the expected holding and rejection cost components of these dynamics are shown to be small. The reason for this reduction is as follows. Although we know that the $\hat{\mathbb{Q}}$ a.s. absolutely continuous valued processes $(\hat{\Psi}_1^n, \hat{\Psi}_2^n)$ are $\mathcal{C}$-tight and, therefore, have a converging subsequence whose limit is denoted by $(\hat{\Psi}_1, \hat{\Psi}_2)$, it does not imply that $\hat{\mathbb{Q}}$ a.s. the paths of $(\hat{\Psi}_1, \hat{\Psi}_2)$ are absolutely continuous. Hence, we cannot argue, for example, that $\hat{\Psi}_1$ is of the form $\hat{\Psi}_{1,i} = \int_0^{\cdot} \hat{\psi}_{1,i}(t)dt$ for some $\hat{\psi}_{1,i}$. As a consequence, we cannot express the limiting measure for the maximizer nor the change of measure penalty using $\hat{\psi}_{1,i}$ as can be done, for example, in (54) by substituting $\psi_{1,i}^n = \lambda_i^n + \hat{\psi}_{1,i}^n(\lambda_i n)^{1/2}$. After this reduction, we bound the change of measure penalty. Finally, the expected cost associated with the uniformly bounded rate dynamics is approximated by the value function of the RSDG.

4.3.3.1. One-Dimensional Reduction. We start with showing the following uniform bound:

$$\limsup_{n\to\infty} \mathbb{E}^{\hat{\mathbb{Q}}} \int_0^\infty e^{-\varrho t}(\hat{R}^n(t))^2 dt < \infty. \tag{71}$$

To establish this, recall the bounds in (57) and (58). As a result,

$$\mathbb{E}^{\hat{\mathbb{Q}}}\big[\|\hat{R}^n(t)\|^2\big] \le C\Big(1 + t^2 + \mathbb{E}^{\hat{\mathbb{Q}}}\big[\|\hat{\Psi}_1^n(t)\|^2 + \|\hat{\Psi}_2^n(t)\|^2\big]\Big),$$

for some constant $C > 0$ independent of n and t. Together with Proposition 3, (71) follows.

Now turn to the holding and rejection costs. By the definitions of the rejection mechanism and of $R^{\sharp,n}$ in Section 4.3.2, on the event $\{\tau^n > T\}$, one has

$$\hat{R}^n(T) = \hat{R}_{i^*}^n(T)e_{i^*} = \mu_{i^*} R^{\sharp,n}(T)e_{i^*}.$$

Relations (59) and (60) imply that

$$\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\hat{r} \cdot d\hat{R}^n(t)\right] = \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty \varrho e^{-\varrho t}\hat{r} \cdot \hat{R}^n(t)dt\right]. \tag{72}$$

Recall that $h^a(w) = \hat{h} \cdot \gamma^a(w)$. Then, Proposition 4, the boundedness of $\hat{X}^n$ and $h^a(X^{\sharp,n})$, the bound $R^{\sharp,n} \le \|\theta^n\| \|\hat{R}^n\|$, and (71) imply that

$$\lim_{n\to\infty} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\{\hat{h}\cdot\hat{X}^n(t) + \varrho\hat{r}\cdot\hat{R}^n(t)\}dt - \int_0^\infty e^{-\varrho t}\{h^a(X^{\sharp,n}(t)) + \varrho r\hat{R}^{\sharp,n}(t)\}dt\right] = 0,$$

where we used the identity $r = r_{i^*}\mu_{i^*}$; see (20). Using the limit $\lim_{n\to\infty}\hat{\mathbb{Q}}(\tau^n < T) \to 0$ and (72), we deduce

$$\lim_{n\to\infty} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\{\hat{h}\cdot\hat{X}^n(t)dt + \hat{r}\cdot d\hat{R}^n(t)\} - \int_0^\infty e^{-\varrho t}\{h^a(X^{\circ,n}(t)) + \varrho r\hat{R}^{\circ,n}(t)\}dt\right] = 0. \tag{73}$$

4.3.3.2. Truncated Intensities. We now show that the maximizer can use probability measures for which $\{\hat{\psi}^n_{j,i}\}_{j,i,n}$ are uniformly bounded from above by a sufficiently large constant k without too much lost.

Recall that, under $\hat{\mathbb{Q}}$, $A^n_i(\cdot) - \int_0^\cdot \psi^n_{1,i}(t)dt$ and $S^n_i(T^n_i(\cdot)) - \int_0^\cdot \psi^n_{2,i}(t)dT^n_i(t)$ are martingales. We now truncate the intensities and consider the $\hat{\mathbb{Q}}$ martingales $A^{n,k}_i(\cdot) - \int_0^\cdot \psi^{n,k}_{1,i}(t)dt$ and $S^{n,k}_i(T^{n,k}_i(\cdot)) - \int_0^\cdot \psi^{n,k}_{2,i}(t)dT^{n,k}_i(t)$, where

$$\psi^{n,k}_{1,i}(t) := \psi^n_{1,i}(t) - (\lambda_i n)^{1/2}\hat{\psi}^n_{1,i}(t)\mathbb{1}_{\{|\hat{\psi}^n_{1,i}(t)|>k\}},$$
$$\psi^{n,k}_{2,i}(t) := \psi^n_{2,i}(t) - (\mu_i n)^{1/2}\hat{\psi}^n_{2,i}(t)\mathbb{1}_{\{|\hat{\psi}^n_{2,i}(t)|>k\}},$$

and $T^{n,k} = (T^{n,k}_i : i \in [I])$ is the DM's control associated with the intensities $\{\psi^{n,k}_{j,i}\}_{j,i}$. Denote $\hat{\psi}^{n,k}_{j,i}(\cdot) := \hat{\psi}^n_{j,i}(\cdot)\mathbb{1}_{\{|\hat{\psi}^n_{j,i}(\cdot)|\le k\}}$, $j \in \{1,2\}$. Clearly, $|\hat{\psi}^{n,k}_{j,i}| \le k$ and

$$\psi^{n,k}_{1,i}(t) = \lambda^n_i + \hat{\psi}^{n,k}_{1,i}(t)(\lambda_i n)^{1/2} + o(n^{1/2}),$$
$$\psi^{n,k}_{2,i}(t) = \mu^n_i + \hat{\psi}^{n,k}_{2,i}(t)(\mu_i n)^{1/2} + o(n^{1/2}).$$

The processes $(A^{n,k}, D^{n,k})$ and (A^n, D^n) are coupled such that, for every $i \in [I]$,

$$\begin{aligned}&(A^n_i - A^{n,k}_i)(\cdot) - (\lambda_i n)^{1/2}\int_0^\cdot \hat{\psi}^n_{1,i}(t)\mathbb{1}_{\{|\hat{\psi}^n_{1,i}(t)|>k\}}dt,\\ &(D^n_i - D^{n,k})_i(\cdot) - (\mu_i n)^{1/2}\int_0^\cdot \hat{\psi}^n_{2,i}(t)\mathbb{1}_{\{|\hat{\psi}^n_{2,i}(t)|>k\}}dT^{n,k}_i(t)\end{aligned} \tag{74}$$

are martingales.

Define $\check{A}^{n,k} := (\check{A}^{n,k}_i : i \in [I])$ with $\check{A}^{n,k}_i := n^{-1/2}(A^{n,k}_i(\cdot) - \int_0^\cdot \psi^{n,k}_{1,i}(t)dt)$ and similarly define $\check{D}^{n,k}$. For every $L \in \{W, X, Y, R, \Psi_1, \Psi_2\}$, let $\hat{L}^{n,k}$ and $L^{\sharp,n,k}$ be defined as $\hat{L}^n$ and $L^{\sharp,n}$, where the intensities $(\psi^{n,k}_{1,i}, \psi^{n,k}_{2,i})$ are replacing $(\psi^n_{i,i}, \psi^n_{2,i})$, $i \in [I]$ and let $T^{n,k}_i$ be the equivalence of T^n_i in this setup. Also, let $\tau^{n,k}$ be the first time a forced rejection occurs in this setup and set $L^{\circ,n,k} := L^{\sharp,n,k}(\cdot \wedge \tau^{n,k})$. Clearly, Lemmas 5 and 6 and Proposition 4 hold in this case as well, where the superscript n is replaced by n,k. For the sake of exposition, we state all the necessary results here.

As a private case of Lemma 5, we get that $\{T^{n,k}\}_n$ converges u.o.c. to $\bar{T} = (\bar{T}_1, \ldots, \bar{T}_I)$, where recall that $\bar{T}_i(t) = \rho_i t$, $t \in \mathbb{R}_+$. Proposition 4 implies that, for sufficiently small δ_0, the following limits hold

$$\lim_{n\to\infty} \hat{X}^{n,k} - \gamma^a(X^{\sharp,n,k}) \overset{\hat{\mathbb{Q}}}{\Longrightarrow} 0,$$

and for every $k, T > 0$,

$$\hat{\mathbb{Q}}(\tau^{n,k} < T) = 0. \tag{75}$$

Lemma 7. *For every $n \in \mathbb{N}$,*

$$(a \wedge X^{\circ,n,k}, Y^{\circ,n,k}, R^{\circ,n,k}) = \Gamma_{[0,a]}(X^{\sharp,n}(0) + W^{\circ,n,k} + \sigma(\Psi^{\circ,n,k}_1 - \Psi^{\circ,n,k}_2) + E^{n,k}), \tag{76}$$

where

$$E^{n,k} := (a \wedge X^{\circ,n,k}) - X^{\circ,n,k} \Rightarrow 0. \tag{77}$$

Moreover, the sequence $\{(W^{\circ,n,k}, X^{\circ,n,k}, Y^{\circ,n,k}, R^{\circ,n,k}, \Psi_1^{\circ,n,k}, \Psi_2^{\circ,n,k}, \hat{\Psi}_1^{n,k}, \hat{\Psi}_2^{n,k})\}_n$ *is* $\mathscr{C}$*-tight and any subsequential limit of it* $(\bar{W}^{\circ,k}, \bar{X}^{\circ,k}, \bar{Y}^{\circ,k}, \bar{R}^{\circ,k}, \bar{\Psi}_1^{\circ,k}, \bar{\Psi}_2^{\circ,k}, \hat{\Phi}_1^k, \hat{\Phi}_2^k)$ *satisfies* $\hat{\mathbb{Q}}$ *a.s.*

$$(\bar{X}^{\circ,k}, \bar{Y}^{\circ,k}, \bar{R}^{\circ,k}) = \Gamma_{[0,a]}(\bar{X}^{\circ,k}(0) + \bar{W}^{\circ,k} + \sigma(\bar{\Psi}_1^{\circ,k} - \bar{\Psi}_2^{\circ,k})), \tag{78}$$

where $\bar{X}^{\circ,k}(0) = X^{\sharp,n}(0) = x_0$*,* $\bar{W}^{\circ,k}$ *is an* (m,σ)*-Brownian motion w.r.t. the filtration* $\mathscr{F}_t := \sigma\{\bar{W}^{\circ,k}(s), \bar{X}^{\circ,k}(s), \bar{Y}^{\circ,k}(s), \bar{R}^{\circ,k}(s), \bar{\Psi}_1^{\circ,k}(s), \bar{\Psi}_2^{\circ,k}(s), \hat{\Phi}_1^k(s), \hat{\Phi}_2^k(s) : 0 \le s \le t\}$*, and*

$$\bar{\Psi}_1^{\circ,k} = \sigma^{-1}\sum_{i=1}^I \theta_i \lambda_i^{1/2} \hat{\Phi}_{1,i}^k, \qquad \bar{\Psi}_2^{\circ,k} = \sigma^{-1}\sum_{i=1}^I \theta_i \mu_i^{1/2} \hat{\Phi}_{2,i}^k. \tag{79}$$

Moreover, there are processes $\{\hat{\phi}_{j,i}^k\}_{j,i}$*, progressively measurable w.r.t.* $\mathscr{G}_t^n$ *such that*

$$\hat{\Phi}_{1,i}^k(\cdot) = \int_0^{\cdot} \hat{\phi}_{1,i}^k(t)dt, \qquad \hat{\Phi}_{2,i}^k(\cdot) = \int_0^{\cdot} \hat{\phi}_{2,i}^k(t)\rho_i dt, \tag{80}$$

which satisfy $\max_{j,i}\|\hat{\phi}_{j,i}^k\|_\infty \le k$*. Finally,* $\hat{\mathbb{Q}}$ *a.s.,*

$$\begin{aligned} \liminf_{n\to\infty} \int_0^\infty e^{-\varrho t}(\hat{\psi}_{1,i}^{n,k}(t))^2 dt &\ge \int_0^\infty e^{-\varrho t}(\hat{\phi}_{1,i}^k(t))^2 dt, \\ \liminf_{n\to\infty} \int_0^\infty e^{-\varrho t}(\hat{\psi}_{2,i}^{n,k}(t))^2 dT_i^n(t) &\ge \int_0^\infty e^{-\varrho t}(\hat{\phi}_{2,i}^k(t))^2 \rho_i dt. \end{aligned} \tag{81}$$

Proof. The Skorokhod mapping formulation of the prelimit processes, the limit in (77), and the $\mathscr{C}$-tightness are private cases of Lemma 6. These results imply the Skorokhod formulation of the limiting process in (78).

The expressions in (79) follow because $\theta^n \to \theta$, and from (64),

$$\Psi_1^{\circ,n,k}(\cdot) := \sigma^{-1}\sum_{i=1}^I \theta_i^n \lambda_i^{1/2} \hat{\Psi}_{1,i}^{n,k}(\tau^{n,k} \wedge \cdot), \qquad \Psi_2^{\circ,n,k}(\cdot) := \sigma^{-1}\sum_{i=1}^I \theta_i^n \mu_i^{1/2} \hat{\Psi}_{2,i}^{n,k}(\tau^{n,k} \wedge \cdot). \tag{82}$$

By definition, for every $i \in [I]$ and $j \in \{1,2\}$, $|\hat{\psi}_{j,i}^{n,k}| \le k$, and therefore, $\hat{\mathbb{Q}}$ a.s., $\{\hat{\Psi}_{j,i}^{n,k}\}_{j,i,n}$ are all Lipschitz-continuous with the same Lipschitz constant k. Because Lipschitz-continuity implies absolutely continuity, from section IV.17 of Dellacherie and Meyer (1978), we get the existence of progressively measurable processes $\{\hat{\phi}_{j,i}^k\}_{j,i}$ such that (80) holds.

The bounds in (81) follow immediately from Cohen (2017), lemma A.3, noticing that, by the Skorokhod representation theorem, the convergence $(T_i^{n,k}, \hat{\Psi}_{2,i}^{n,k}) \Rightarrow (\bar{T}_i, \hat{\Phi}_{2,i}^k)$ can be replaced by u.o.c. $\hat{\mathbb{Q}}$ a.s. convergence.

Finally, we prove that $\bar{W}^{\circ,k}$ is an (m,σ)-Brownian motion. The martingale central limit theorem (Ethier and Kurtz (1986), theorem 7.1.4) implies that $\{(\check{A}^{n,k}, \check{S}^{n,k})\}_n$ converges to a $2I$-dimensional $(0,\tilde{\sigma})$-Brownian motion, where $\check{S}^n = (\check{S}_i^n : i \in [I])$ and

$$\check{S}_i^{n,k}(t) := n^{-1/2}\left(S_i^n(t) - \int_0^t \psi_{2,i}^n(s)ds\right), \quad \tilde{\sigma} = \mathrm{Diag}(\lambda_1^{1/2}, \ldots, \lambda_I^{1/2}, \mu_1^{1/2}, \ldots, \mu_I^{1/2}).$$

Using a lemma regarding random change of time from Billingsley (1999) with $T_i^n \to \bar{T}_i$, gives that $\{(\check{A}^{n,k}, \check{D}^{n,k})\}$ converges to a $(0,\check{\sigma})$-Brownian motion, where

$$\check{\sigma} := (\lambda_1^{1/2}, \ldots, \lambda_I^{1/2}, \lambda_1^{1/2}, \ldots, \lambda_I^{1/2}).$$

As a result, from the definitions of $\hat{W}^n$, $W^{\sharp,n,k}$, and $W^{\circ,n,k}$ (see their equivalences in (62), (63), and (67)), we get that $\bar{W}^{\circ,k}$ is an (m,σ)-Brownian motion w.r.t. its own filtration. The proof that its filtration can be replaced by $\mathscr{F}_t$ follows by the same lines of Proposition 2 and is, therefore, omitted. □

The next proposition tells us that, by truncating the intensities, the expected cost does not change much.

Proposition 5. *The following limit holds:*

$$\lim_{k\to\infty}\limsup_{n\to\infty}\left|\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\{h^a(X^{\circ,n}(t)) + \varrho r R^{\circ,n}(t)\}dt - \int_0^\infty e^{-\varrho t}\{h^a(X^{\circ,n,k}(t)) + \varrho r R^{\circ,n,k}(t)\}dt\right]\right| = 0 \tag{83}$$

and for every $i \in [I]$,

$$\lim_{k\to\infty}\limsup_{n\to\infty}\left(\left|L_1^{\varrho}(\hat{\mathbb{Q}}_{1,i}^n\|\mathbb{P}_{1,i}^n) - L_1^{\varrho}(\hat{\mathbb{Q}}_{1,i}^{n,k}\|\mathbb{P}_{1,i}^n)\right| + \left|L_2^{\varrho}(\hat{\mathbb{Q}}_{2,i}^n\|\mathbb{P}_{1,i}^n) - L_2^{\varrho}(\hat{\mathbb{Q}}_{2,i}^{n,k}\|\mathbb{P}_{1,i}^n)\right|\right) = 0, \tag{84}$$

where $\hat{\mathbb{Q}}_{1,i}^{n,k}$ *and* $\hat{\mathbb{Q}}_{2,i}^{n,k}$ *are the measures associated with* $\psi_{1,i}^n(t) = \lambda_i^n t + (\lambda_i n)^{1/2}\hat{\psi}_{1,i}^{n,k}(t)$ *and* $\psi_{2,i}^n(t) = \mu_i^n t + (\mu_i n)^{1/2}\hat{\psi}_{2,i}^{n,k}(t)$.

Proof. From the representations in (69) and (76); the limits (70) and (77); the definitions (64), (67), and (82); and Lemma 1, it follows that in order to obtain (83), it is sufficient to show that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\|W^{\circ,n,k} - W^{\circ,n}\|_t dt\right] = 0, \tag{85}$$

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\|\hat{\Psi}_{1,i}^{n,k} - \hat{\Psi}_{1,i}^n\|_t dt\right] = 0, \tag{86}$$

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\|\hat{\Psi}_{2,i}^{n,k} - \hat{\Psi}_{2,i}^n\|_t dT_i^n(t)\right] = 0. \tag{87}$$

Set the functions $f_{1,i}^n : (-\lambda_i^n(\lambda_i n)^{-1/2}, \infty) \to \mathbb{R}$ and $f_{2,i}^n : (-\mu_i^n(\mu_i n)^{-1/2}, \infty) \to \mathbb{R}$ given by

$$f_{1,i}^n(x) := \left(\lambda_i^n + (\lambda_i n)^{1/2}x\right)\log\left(1 + (\lambda_i n)^{1/2}x/\lambda_i^n\right) - (\lambda_i n)^{1/2}x,$$

$$f_{2,i}^n(x) := \left(\mu_i^n + (\mu_i n)^{1/2}x\right)\log\left(1 + (\mu_i n)^{1/2}x/\mu_i^n\right) - (\mu_i n)^{1/2}x.$$

To obtain (84), recall (54). Then it is sufficient to show that

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}|f_{1,i}^n(\hat{\psi}_{1,i}^{n,k}(t)) - f_{1,i}^n(\hat{\psi}_{1,i}^n(t))|dt\right] = 0, \tag{88}$$

$$\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}|f_{2,i}^n(\hat{\psi}_{2,i}^{n,k}(t)) - f_{2,i}^n(\hat{\psi}_{2,i}^n(t))|dT_i^n(t)\right] = 0. \tag{89}$$

We start with the limit in (85). Simple algebraic manipulation gives the bound

$$\|W^{\circ,n,k} - W^{\circ,n}\|_t \le C(\|\check{A}^{n,k} - \check{A}^n\|_t + \|\check{D}^{n,k} - \check{D}^n\|_t),$$

where C refers to a finite positive constant that is independent of n and t and which can change from one line to the next. From (74) and the Burkholder–Davis–Gundy inequality, we get that, for every $i \in [I]$,

$$\mathbb{E}^{\hat{\mathbb{Q}}}[\|\check{A}_i^{n,k} - \check{A}_i^n\|_t^2] \le Cn^{-1/2}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^t |\hat{\psi}_{1,i}^n(s)|\mathbb{1}_{\{|\hat{\psi}_{1,i}^n(s)|>k\}}ds\right].$$

Now, changing the order of integration gives

$$\begin{aligned}\lim_{k\to\infty}\limsup_{n\to\infty}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\|\check{A}_i^{n,k} - \check{A}_i^n\|_t^2 dt\right] &\le \lim_{k\to\infty}\limsup_{n\to\infty} Cn^{-1/2}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\left\{\int_0^t |\hat{\psi}_{1,i}^n(s)|\mathbb{1}_{\{|\hat{\psi}_{1,i}^n(s)|>k\}}ds\right\}dt\right] \\ &\le \lim_{k\to\infty}\limsup_{n\to\infty} Cn^{-1/2}\mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty \varrho e^{-\varrho t}|\hat{\psi}_{1,i}^n(t)|\mathbb{1}_{\{|\hat{\psi}_{1,i}^n(t)|>k\}}dt\right] = 0.\end{aligned} \tag{90}$$

Pay attention to that, by Proposition 3, the last limit holds even without the $n^{-1/2}$ term.

The rest of the sufficient limits, namely (86)–(89), are treated similarly: notice that

$$\|\hat{\Psi}_{1,i}^{n,k} - \hat{\Psi}_{1,i}^n\|_t \le C\sum_{i=1}^I \int_0^t |\hat{\psi}_{1,i}^n(s)|\mathbb{1}_{\{|\hat{\psi}_{1,i}^n(s)|>k\}}ds,$$

$$|f_{1,i}^n(\hat{\psi}_{1,i}^{n,k}(t)) - f_{1,i}^n(\hat{\psi}_{1,i}^n(t))| \le C\sum_{i=1}^I f_{1,i}^n(\hat{\psi}_{1,i}^n(t))\mathbb{1}_{\{|f_{1,i}^n(\hat{\psi}_{1,i}^n(t))|>k\}}.$$

and similar bounds apply for $j = 2$ as well. We may continue in the same way as we did in (90), where now the n^{1-2} term is absent, using the two bounds in Proposition 3. □

We now show that the change of measure penalty in the truncated case can be approximated by a quadratic form. From (54), the bound $\sup_{t\in\mathbb{R}_+}|\psi^{n,k}_{1,i}(t)|\le k$, and Taylor's expansion of $\log(1+x)$, one has

$$L^{\varrho}_1(\hat{\mathbb{Q}}^{n,k}_{1,i}\|\hat{\mathbb{P}}^n_{1,i}) = \mathbb{E}^{\hat{\mathbb{Q}}}\biggl[\frac{1}{2}\int_0^\infty e^{-\varrho t}(\hat{\psi}^{n,k}_{1,i}(t))^2dt\biggr] + o(1).$$

Similarly,

$$L^{\varrho}_2(\hat{\mathbb{Q}}^{n,k}_{2,i}\|\hat{\mathbb{P}}^n_{2,i}) = \mathbb{E}^{\hat{\mathbb{Q}}}\biggl[\frac{1}{2}\int_0^\infty e^{-\varrho t}(\hat{\psi}^{n,k}_{2,i}(t))^2dT^n_i(t)\biggr] + o(1).$$

From (81), it follows that

$$\begin{aligned}&\liminf_{n\to\infty}\left[\sum_{i=1}^I\frac{1}{\kappa_{1,i}}L^{\varrho}_1(\hat{\mathbb{Q}}^{n,k}_{1,i}\|\hat{\mathbb{P}}^n_{1,i}) + \sum_{i=1}^I\frac{1}{\kappa_{2,i}}L^{\varrho}_2(\hat{\mathbb{Q}}^{n,k}_{2,i}\|\hat{\mathbb{P}}^n_{2,i})\right]\\ &\quad\ge \sum_{i=1}^I\frac{1}{\kappa_{1,i}}\mathbb{E}^{\hat{\mathbb{Q}}}\biggl[\frac{1}{2}\int_0^\infty e^{-\varrho t}(\hat{\phi}^k_{1,i}(t))^2dt\biggr] + \sum_{i=1}^I\frac{1}{\kappa_{2,i}}\mathbb{E}^{\hat{\mathbb{Q}}}\biggl[\frac{1}{2}\int_0^\infty e^{-\varrho t}(\hat{\phi}^k_{2,i}(t))^2\rho_i dt\biggr] + o(1).\end{aligned} \tag{91}$$

The identity

$$\frac{1}{k_1}\alpha_1^2 + \frac{1}{k_2}\alpha_2^2 \ge \frac{1}{k_1+k_2}(\alpha_1-\alpha_2)^2, \qquad k_1,k_2>0, \quad \alpha_1,\alpha_2\in\mathbb{R},$$

implies that

$$\frac{1}{2\kappa_{1,i}}(\hat{\phi}^k_{1,i}(t))^2 + \frac{1}{2\kappa_{2,i}}\rho_i(\hat{\phi}^k_{2,i}(t))^2 \ge \frac{1}{4\hat{\varepsilon}_i}\left(\hat{\phi}^k_{1,i}(t) - \rho_i^{1/2}\hat{\phi}^k_{2,i}(t)\right)^2, \tag{92}$$

where recall that $\hat{\varepsilon}_i = \frac{1}{2}(\kappa_{1,i}+\kappa_{2,i})$. Set

$$\phi^{\sharp,k} := \sigma^{-1}\sum_{i=1}^I\theta_i\left(\lambda_i^{1/2}\hat{\phi}^k_{1,i} - \mu_i^{1/2}\hat{\phi}^k_{2,i}\rho_i\right).$$

By the Cauchy–Schwartz inequality,

$$\left[\sigma^{-2}\sum_{i=1}^I(\theta\hat{\sigma})^2_i\hat{\varepsilon}_i\right]\times\sum_{i=1}^I\frac{1}{2\hat{\varepsilon}_i}\left(\hat{\phi}^k_{1,i}(t) - \rho_i^{1/2}\hat{\phi}^k_{2,i}(t)\right)^2 \ge (\phi^{\sharp,k}(t))^2.$$

Then,

$$\sum_{i=1}^I\frac{1}{4\hat{\varepsilon}_i}\left(\hat{\phi}^k_{1,i}(t) - \rho_i^{1/2}\hat{\phi}^k_{2,i}(t)\right)^2 \ge \frac{1}{2\varepsilon}(\phi^{\sharp,k}(t))^2,$$

where recall that $\varepsilon = \sigma^{-2}\sum_{i=1}^I(\theta\hat{\sigma})^2_i\hat{\varepsilon}_i$.

From (91), (92), and the last bound, we obtain that

$$\liminf_{n\to\infty}\left[\sum_{i=1}^I\frac{1}{\kappa_{1,i}}L^{\varrho}_1(\hat{\mathbb{Q}}^{n,k}_{1,i}\|\hat{\mathbb{P}}^n_{1,i}) + \sum_{i=1}^I\frac{1}{\kappa_{2,i}}L^{\varrho}_2(\hat{\mathbb{Q}}^{n,k}_{2,i}\|\hat{\mathbb{P}}^n_{2,i})\right] \ge \frac{1}{2\varepsilon}\mathbb{E}^{\hat{\mathbb{Q}}}\biggl[\int_0^\infty e^{-\varrho t}(\phi^{\sharp,k}(t))^2dt\biggr]. \tag{93}$$

Fix $\delta_1>0$. Combining the last limit with the ones from (73) and (83) and recalling (71) and the bound $\|R^{\circ,n}\|\le\|\theta^n\|\|\hat{R}^n\|$, we get that there is $k_{\delta_1}>0$ such that for every $k\ge k_{\delta_1}$,

$$\begin{aligned}&\limsup_{n\to\infty}\ J^n(X^n(0),U^n(a),R^n(a),\hat{\mathbb{Q}}^n;\kappa)\\ &\quad\le\mathbb{E}^{\hat{\mathbb{Q}}}\biggl[\int_0^\infty e^{-\varrho t}\{h^a(\bar{X}^{\circ,k}(t)) + \varrho r\bar{R}^{\circ,k}(t) - \frac{1}{2\varepsilon}(\phi^{\sharp,k}(s))^2\}dt\biggr] + \delta_1/2,\end{aligned} \tag{94}$$

where notice that (78)–(80) and the definition of $\phi^{\#,k}$ gives

$$(\bar{X}^{\circ,k}, \bar{Y}^{\circ,k}, \bar{R}^{\circ,k}) = \Gamma_{[0,a]}\left(x_0 + \bar{W}^{\circ,k} + \sigma\int_0^{\cdot} \phi^{\#,k}(s)ds\right).$$

Fix such k. Because $a \to \beta_\varepsilon$ as $\delta_0 \to 0$ (see the paragraph before (24)), Lemma 1 and (26) imply that

$$\int_0^\infty e^{-\varrho t}\{h^a(\bar{X}^{\circ,k}(t)) + \varrho r\bar{R}^{\circ,k}(t) - \frac{1}{2\varepsilon}(\phi^{\#,k}(s))^2\}dt \overset{\hat{\mathbb{Q}}}{\Longrightarrow} \int_0^\infty e^{-\varrho t}\{h(\bar{X}(t)) + \varrho r\bar{R}(t) - \frac{1}{2\varepsilon}(\phi^{\#,k}(s))^2\}dt,$$

as $\delta_0 \to 0$, where

$$(\bar{X}, \bar{Y}, \bar{R}) = \Gamma_{[0,\beta_\varepsilon]}\left(x_0 + \bar{W}^{\circ,k} + \sigma\int_0^{\cdot} \phi^{\#,k}(s)ds\right).$$

4.3.3.3. Obtaining the Upper Bound. As argued earlier, to conclude the convergence in expectation of the preceding, it is sufficient to show that

$$\limsup_{a\to\beta_\varepsilon} \mathbb{E}\left[\int_0^\infty e^{-\varrho t}\|\bar{R}^{\circ,k}(t)\|^2 dt\right] < \infty. \tag{95}$$

The proof is borrowed from Atar and Shifrin (2014), equation 85, and adapted to our case with the additional stochastic drift $\sigma\int_0^{\cdot}\phi^{\#,k}(s)ds$. Apply Itô's formula to $(\bar{X}^{\circ,k})^2$; use the reflection conditions $\int_0^t \bar{X}^{\circ,k}(s)d\bar{Y}^{\circ,k}(s) = 0$ and $\int_0^t \bar{X}^{\circ,k} d\bar{R}^{\circ,k}(s) = a\bar{R}^{\circ,k}(t)$ to get

$$\bar{R}^{\circ,k}(t) = \frac{1}{2a}\left\{(\bar{X}^{\circ,k}(0))^2 - (\bar{X}^{\circ,k}(t))^2 + 2\int_0^t \bar{X}^{\circ,k}(s)[d\bar{W}^{\circ,k}(s) + \sigma\phi^{\#,k}(s)ds] + \sigma^2 t\right\}.$$

Because $\bar{X}^{\circ,k}$ and $\sigma\phi^{\#,k}(s)$ are bounded, (95) follows easily.

Recall the definition of J and V from Section 3.1 and that $\bar{X}^{\circ,k}(0) = x_0$; then

$$\begin{aligned}
&\lim_{\delta_0\to 0} \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\{h^a(\bar{X}^{\circ,k}(t)) + \varrho r\bar{R}^{\circ,k}(t) - \frac{1}{2\varepsilon}(\phi^{\#,k}(t))^2\}dt\right] \\
&= \mathbb{E}^{\hat{\mathbb{Q}}}\left[\int_0^\infty e^{-\varrho t}\{h(\bar{X}(t)) + \varrho r\bar{R}(t) - \frac{1}{2\varepsilon}(\phi^{\#,k}(t))^2\}dt\right] \\
&= J(\bar{X}^{\circ,k}(0), \bar{Y}, \bar{R}, \mathbb{Q}; \varepsilon) \le V(x_0; \varepsilon),
\end{aligned}$$

where $\mathbb{Q}$ is the measure associated with the rate $\phi^{\#,k}$. The last inequality follows by the optimality of the minimizer's reflected strategy; see Proposition 1.

Together with (94), we obtain that, for sufficiently small $\delta_0 > 0$ and large n,

$$J^n(X^n(0), U^n(a), R^n(a), \hat{\mathbb{Q}}^n; \kappa) \le V(x_0; \varepsilon) + \delta_1.$$

Because the constant $\delta_1(>0)$ and the measures $\{\hat{\mathbb{Q}}^n\}_n$ were chosen arbitrarily, we get (31).

4.4. The Maximizer's Asymptotic Behavior

Recall that, under the measure $\hat{\mathbb{Q}}^n_{1,i}$ ($\hat{\mathbb{Q}}^n_{2,i}$), A^n_i ($S^n_i(T^n_i)$) is a counting process with intensity $\psi^n_{1,i}$ ($\psi^n_{2,i}U^n_i$). Notice also that, under the measures $\hat{\mathbb{Q}}^n_{j,i}$, $j \in \{1,2\}$, the critical load condition might be violated because we do not restrict the intensities $\{\hat{\psi}^n_{j,i}\}_{j,i,n}$ in such a way. However, as follows from Sections 4.2 and 4.3, such changes of measures are too costly and are avoided by the maximizer so that, as a consequence, on average, $\psi^n_{1,i}(t) = \lambda^n_i + \mathbb{O}(n^{1/2})$ and $\psi^n_{2,i}(t) = \mu^n_i + \mathbb{O}(n^{1/2})$. Indeed, in the proof of the lower bound, we used the equilibrium control for the maximizer from the RSDG and set up the maximizer's control in the QCP to be given by (33)–(35). Therefore, the critical load condition holds.

In the proof of the upper bound, we show in Section 4.3.1 and, more specifically, in Proposition 3 that $\psi^n_{1,i}(t) = \lambda^n_i + (\lambda_i n)^{1/2}\hat{\psi}^n_{i,1}(t)$, where the expectations $\mathbb{E}^{\hat{\mathbb{Q}}}[\int_0^\infty e^{-\varrho t}|\hat{\psi}^n_{1,i}(t)|dt]$ are uniformly bounded in n, and similarly for $\psi^n_{2,i}$.

The main reasoning for this phenomenon lies in the fact that we are studying a diffusion-scaled problem, and as such, the uncertainty is expressed on the scaled system. Now, by observing the diffusion-scaled system, the DM can rule out measures under which $|\psi^n_{i,1} - \lambda^n_i| + |\psi^n_{i,2} - \mu^n_i|$ is of higher order than $n^{1/2}$ (for sufficiently large n). This is because the diffusion coefficient in the limiting dynamics under the new measure would be different from the one that emerges from the reference measure, and in the limiting problem, this could be identified very quickly. Another reason is, of course, the structure of the penalty using the Kullback–Leibler divergence, which also leads to a limiting problem with linear drift and quadratic penalty in the maximizer's control; see (19).

5. Future Outlook

1. Recall Remark 1(ii). Much as in the Markovian case, the heavy traffic approximation for G/G/1 also scales (2) to (8), whose form uses the mean arrival and service rates λ^n_i and μ^n_i. In the change of measure given by (6) appear the actual arrival and service rates under the measures $\mathbb{P}^n_{j,i}$ and $\hat{\mathbb{Q}}^n_{j,i}$, namely (λ^n_i, μ^n_i) and $(\psi^n_{1,i}(\cdot), \psi^n_{2,i}(\cdot))$, respectively. In the Markovian case (M/M/1), the mean and the actual arrival and service rates are identical. Our technique heavily uses this identification. It is desirable to extend the main result of this paper beyond the Markovian setting to general service time distributions and renewal arrival distributions under second moment conditions.

2. Our model considers finite buffers and linear cost. The author is currently testing the robustness of the *generalized* $c\mu$ policy in the unconstrained case. I expect to benefit from the pathwise minimality property of the Skorokhod mapping (see Atar and Biswas 2014) in three aspects: (1) replacing the Kullback–Leibler divergence with a more general divergence satisfying basic properties, such as growth, continuity, and convexity; (2) attaining the lower bound without needing the smoothness of the value function—this allows us to skip the differential equation analysis done in Cohen (2018)—and (3) proving the lower bound without the time-rescaling procedure. The state-space collapse part in the proof of the upper bound in this case should be simpler again because rejections are not allowed. However, using a general divergence leads to a more subtle analysis of the maximizer's penalty part of the cost.

3. The present model assumes the uncertainty in the diffusion scale, and as discussed in Section 4.4, the system stays critically loaded. The author studies a similar type of uncertainty in the fluid scale. In this case, we do not assume a reference model, but rather the maximizer chooses the worst set of parameters (possibly stochastic and time-dependent) within a given set of possible parameters without a penalty component. In this case, it is less trivial that an index policy is asymptotically optimal.

4. Recently, Blanchet and Murthy (2019) quantified the impact of model misspecification when computing general expected values using an optimal transport cost instead of divergences. It would be interesting to tackle the QCP presented here and challenging the $c\mu$ rule using optimal transport tools.

6. Proofs

Proof of Lemma 2. The $\mathscr{C}$-tightness of $\{\tilde{\tau}^n, \tilde{Y}^n, \tilde{R}^n\}_n$ follows by the following observation, which relies on the estimate $|\tau^n(\tilde{\tau}^n(t)) - t| \le \|\theta^n\|/\sqrt{n}$ and the definition of τ^n.

$$\begin{aligned}\frac{2\|\theta^n\|}{\sqrt{n}} + t - s &\ge \tau^n(\tilde{\tau}^n(t)) - \tau^n(\tilde{\tau}^n(s)) \\ &= \tilde{\tau}^n(t) - \tilde{\tau}^n(s) + \theta^n \cdot \tilde{R}^n(t) - \theta^n \cdot \tilde{R}^n(s) + \theta^n \cdot \tilde{Y}^n(t) - \theta^n \cdot \tilde{Y}^n(s).\end{aligned} \tag{96}$$

Next, because the sequence $\{(\tilde{\tau}^n, \tilde{Y}^n)\}_n$ is $\mathscr{C}$-tight, we get, by (7) and the limit $\mu^n_i n^{-1/2} \to \infty$, that $\tilde{T}^n \overset{\hat{\mathbb{Q}}^n}{\Longrightarrow} \tilde{T}$.

From (40), the $\mathscr{C}$-tightness of $\{\tilde{\tau}^n, \tilde{Y}^n, \tilde{R}^n\}_n$, and (36), the $\mathscr{C}$-tightness of $\{\tilde{X}^n\}_n$ follows once we show that $\{(\tilde{A}^n, \tilde{D}^n)\}_n$ is $\mathscr{C}$-tight. Recalling that $\{\tilde{\tau}^n\}_n$ is $\mathscr{C}$-tight, then in order to prove the latter statement, it is sufficient to show the $\mathscr{C}$-tightness of $\{(\check{A}^n, \check{D}^n)\}$.

For every $i \in [I]$ and $n \in \mathbb{N}$, the processes $\psi^n_{1,i}$ and $\psi^n_{2,i} U^n_i$ are the intensities of A^n_i and $D^n_i := S^n_i(T^n_i)$, respectively. Denote

$$\begin{aligned}(W^n_1, \ldots, W^n_{2I}) &:= (A^n_1, \ldots, A^n_I, D^n_1, \ldots, D^n_I),\\ \check{W}^n = (\check{W}^n_1, \ldots, \check{W}^n_{2I}) &:= (\check{A}^n_1, \ldots, \check{A}^n_I, \check{D}^n_1, \ldots, \check{D}^n_I),\\ \tilde{W}^n = (\tilde{W}^n_1, \ldots, \tilde{W}^n_{2I}) &:= (\tilde{A}^n_1, \ldots, \tilde{A}^n_I, \tilde{D}^n_1, \ldots, \tilde{D}^n_I).\end{aligned}$$

Therefore, for every $j \in \{1, \ldots, 2I\}$, $\check{W}^n_j$ is a martingale w.r.t. its own filtration under the measure $\hat{\mathbb{Q}}^n$. Notice that the quadratic variation of $\check{W}^n_j$ satisfies

$$[\check{W}^n_j] = \frac{1}{n} W^n_j, \tag{97}$$

which is of order t thanks to (35), (36), and (3). Fix an arbitrary $T > 0$ and a stopping time π. Then, by the Burkholder–Davis–Gundy inequality (see Protter (2004), theorem 48) and (97), we get

$$\left(\mathbb{E}^{\hat{\mathbb{Q}}^n} \|\check{W}^n_j(\pi + \cdot) - \check{W}^n_j(\pi)\|_\delta\right)^2 \le \mathbb{E}^{\hat{\mathbb{Q}}^n} \|\check{W}^n_j(\pi + \cdot) - \check{W}^n_j(\pi)\|^2_\delta \le C_3 \mathbb{E}^{\hat{\mathbb{Q}}^n}[\check{W}^n_j(\pi + \cdot)]_\delta \le C_4 \delta$$

for some constants $C_3, C_4 > 0$, independent of n, δ, and π. Therefore, Aldous criterion for tightness holds (see, e.g., Billingsley (1999), theorem 16.10). Because the jumps of these processes are of order $O(n^{-1/2})$, any limit process has continuous paths with probability one and $\mathscr{C}$-tightness of $\{\check{W}^n\}_n$ is proved.

The first two identities in (46) follow by the convergence $(\check{W}^n, \tilde{W}^n, \tilde{\tau}^n) \Rightarrow (\check{W}, \tilde{W}, \tilde{\tau})$, which, in turn, follows by the tightness of $\{(\check{W}^n, \tilde{W}^n, \tilde{\tau}^n)\}_n$. Finally, the quadratic variation of $\tilde{B}$ follows by (35), (46), and the martingale central limit theorem (see Ethier and Kurtz (1986), theorem 7.1.4).

As mentioned after Lemma 2, by Skorokhod's representation theorem, we may assume that along a subsequence (47) holds a.s., u.o.c. under some probability space that supports all the relevant processes. Now, from Dai and Williams (1996), lemma 2.4; (33)–(35); and (40), we obtain (45). □

Proof of Lemma 3. Fix $0 < s < t$. From (44), it follows that

$$\begin{aligned}\mathbb{Q}^*(\tilde{\tau}^n(t) \le s) = \mathbb{Q}^*(\tau^n(s) \ge t) &= \mathbb{Q}^*(s + \theta^n \cdot \hat{R}^n(s) + \theta^n \cdot \hat{Y}^n(s) > t)\\ &\le \frac{1}{t-s} \mathbb{E}^{\mathbb{Q}^*}[\theta^n \cdot \hat{R}^n(s) + \theta^n \cdot \hat{Y}^n(s)].\end{aligned} \tag{98}$$

From (36), (40), and the inequality $T^n(u) \le u$, $u \in \mathbb{R}_+$, we get that there is a constant $C_5 > 0$ independent of n and t and such that

$$\hat{Y}^n_i(s) \le C_5(1 + \hat{R}^n_i(s) + \check{A}^n_i(s) - \check{D}^n_i(s)).$$

Recalling that $\check{A}^n_i - \check{D}^n_i$ is a martingale and (39), we get that the last display in (98) is bounded above by $\frac{C_6}{t-s}$ for some $C_6 > 0$, independent of n and t. Now, because, for every $s \in \mathbb{R}_+$ and $0 \le t \le t'$, $\{\tilde{\tau}(t) > s\} \supseteq \{\tilde{\tau}(t') > s\}$, we get that

$$\mathbb{Q}^*\left(\lim_{t\to\infty} \tilde{\tau}(t) < s\right) = \lim_{t\to\infty} \mathbb{Q}^*(\tilde{\tau}(t) < s).$$

Using the convergence in law $\tilde{\tau}^n \overset{d}{\Rightarrow} \tilde{\tau}$ and because $\mathbb{Q}^*(\tilde{\tau}^n(t) < s) \le \frac{C_6}{t-s}$, we conclude that

$$\mathbb{Q}^*\left(\lim_{t\to\infty} \tilde{\tau}(t) < s\right) \le \lim_{t\to\infty} \limsup_{n\to\infty} \mathbb{Q}^*(\tilde{\tau}^n(t) < s) = 0. \quad \square$$

Proof of Proposition 2. We use the Lévy characterization for Brownian motions. Namely, we show that $B^*(t)$ and $B^*(t)(B^*(t))^\top - \mathscr{I}t$ are continuous local martingales w.r.t. $\mathscr{G}^*$, where recall that $\mathscr{I}$ is the identity matrix of order $I \times I$.

To simplify the presentation, set

$$\begin{aligned}\check{B}^n &= (\check{B}^n_1, \ldots, \check{B}^n_I) = \hat{\sigma}^{-1}(\check{A}^n - \check{D}^n), & \check{B} &= (\check{B}_1, \ldots, \check{B}_I) = \hat{\sigma}^{-1}(\check{A} - \check{D}),\\ \tilde{B}^n &= (\tilde{B}^n_1, \ldots, \tilde{B}^n_I) = \hat{\sigma}^{-1}(\tilde{A}^n - \tilde{D}^n), & \tilde{B} &= (\tilde{B}_1, \ldots, \tilde{B}_I) = \hat{\sigma}^{-1}(\tilde{A} - \tilde{D}).\end{aligned}$$

We start by arguing the continuity. From (47), $(\check{B}^n, \tilde{\tau}^n, \tilde{B}^n) \to (\check{B}, \tilde{\tau}, \tilde{B})$ $\mathbb{Q}^*$ a.s., u.o.c. and, therefore, $\tilde{B}(\cdot) = \check{B}(\tilde{\tau}(\cdot))$, $\mathbb{Q}^*$ a.s., u.o.c. Thus, $B^*(\cdot) = \tilde{B}(\tau(\cdot)) = \check{B}(\tilde{\tau}(\tau(\cdot))) = \check{B}(\cdot)$, which is continuous $\mathbb{Q}^*$ a.s.; see Lemma 2. The proof that $B^*(t)(B^*(t))^\top - \mathcal{I}t$ is a local martingale follows by the same lines of the proof that $B^*(t)$ is a local martingale. We start with the proof of the latter and then add the missing details for the former. Recall the definition of $\mathcal{G}^n$ from Definition 1(i).

Fix $t, s \in \mathbb{R}_+$ and $n \in \mathbb{N}$. Notice that $\{\tilde{\tau}^n(s) \le t\} = \{\tau^n(t) \ge s\} = \{t + \theta \cdot \hat{R}^n(t) + \theta \cdot \hat{Y}^n(t) \ge s \in \mathcal{G}_t^n$. Thus, $\tilde{\tau}^n(s)$ is a $\mathcal{G}_t^n$-stopping time. Recall that $\check{A}^n$ and $\check{D}^n$ are $\mathcal{G}_t^n$-martingales; then the optional sampling theorem (see problem 1.3.24 in Cohen (2018)) yields that $\tilde{B}^n(t+\cdot) = \check{B}^n(\tilde{\tau}(t+\cdot))$ is a $\mathcal{G}^n(\tilde{\tau}^n(t))$ martingale. As a consequence, for every $i \in [I]$ and every $\mathcal{G}^n(\tilde{\tau}^n(t))$-measurable random variable ζ^n,

$$\mathbb{E}^{\mathbb{Q}^*}[(\tilde{B}_i^n(t+s) - \tilde{B}_i^n(t))\zeta^n] = 0.$$

Recall that $t \mapsto \tilde{\tau}^n(t)$ is nondecreasing, and therefore, $\mathcal{G}^n(\tilde{\tau}^n(t))$ is a filtration. Now, for every bounded continuous function g,

$$\begin{aligned} &g(\tilde{A}^n(s_m), \tilde{D}^n(s_m), \tilde{Y}^n(s_m), \tilde{R}^n(s_m), \tilde{\tau}^n(s_m);\ 0 \le s_m \le t, m = 1, \dots, k) \\ &\quad = g(\check{A}^n(\tilde{\tau}^n(s_m)), \check{D}^n(\tilde{\tau}^n(s_m)), \hat{Y}(\tilde{\tau}^n(s_m)), \hat{R}(\tilde{\tau}^n(s_m)), \tilde{\tau}^n(s_m);\ 0 \le s_m \le t, m = 1, \dots, k) \end{aligned}$$

is $\mathcal{G}^n(\tilde{\tau}^n(t))$-measurable, where $k \in \mathbb{N}$, and $\{s_m\}$ is an arbitrary sequence. Combining the last two displays, one gets

$$\mathbb{E}^{\mathbb{Q}^*}\big[g(\tilde{A}^n(s_m), \tilde{D}^n(s_m), \tilde{Y}^n(s_m), \tilde{R}^n(s_m), \tilde{\tau}^n(s_m);\ 0 \le s_m \le t, m = 1, \dots, k) \times \big(\tilde{B}_i^n(t+s) - \tilde{B}_i^n(t)\big)\big] = 0. \tag{99}$$

From (40),

$$\begin{aligned} \tilde{X}_i^n(t) &= \tilde{X}_i^n(0) + \hat{m}_i^n \tilde{\tau}^n(t) + \tilde{A}_i^n(t) - \tilde{D}_i^n(t) + \tilde{Y}_i^n(t) - \tilde{R}_i^n(t) \\ &\quad + \lambda_i^{1/2} \int_0^{\tilde{\tau}^n(t)} \hat{\psi}_{1,i}^n(s) ds - \mu_i^{1/2} \int_0^{\tilde{\tau}^n(t)} \hat{\psi}_{2,i}^n(s) dT_i^n(s). \end{aligned} \tag{100}$$

From (36), (100), and the bound $T^n(u) \le u$, $u \in \mathbb{R}_+$, we get that there are constants $C_7, C_8 > 0$, independent of t, such that, for sufficiently large n, one has

$$|\tilde{B}_i^n(t)| \le C_7(\theta^n \cdot \tilde{R}_i^n(t) + \theta^n \cdot \tilde{Y}_i^n(t) + \tilde{\tau}_i^n(t) + 1) \le C_8(t+1), \quad t \in \mathbb{R}_+,$$

where the last inequality follows by the same arguments leading to (96). From (47), (99), and the bounded convergence theorem, we get that

$$\mathbb{E}^{\mathbb{Q}^*}\big[g(\tilde{A}(s_m), \tilde{D}(s_m), \tilde{Y}(s_m), \tilde{R}(s_m), \tilde{\tau}(s_m);\ 0 \le s_m \le t, m = 1, \dots, k) \times \big(\tilde{B}_i(t+s) - \tilde{B}_i(t)\big)\big] = 0, \tag{101}$$

which implies that $\tilde{B}$ is a $\tilde{\mathcal{G}}_t$ martingale.

Next, for every $i \in [I]$, we show that the composition $B_i^*(t) = \tilde{B}_i(\tau(t))$ is a $\mathcal{G}_t^*$-local martingale. For this, we need to define the following stopping times. Fix $M > 0$ and set,

$$\begin{aligned} \tilde{\pi}_{i,M} &:= \inf\{t \ge 0 : \tilde{B}_i(t) > M\}, \\ \pi_{i,M} &:= \inf\{t \ge 0 : \check{B}_i(t) > M\}. \end{aligned}$$

One can verify that $\tau(\pi_{i,M}) = \tilde{\pi}_{i,M}$.

We now show that $B_i^*(t \wedge \pi_{i,M})$, which, by definition, equals $\tilde{B}_i(\tau(t \wedge \pi_{i,M}))$ and is a $\mathcal{G}_t^*$ martingale. Because $\lim_{M\to\infty} \pi_{i,M} = \infty$, $\mathbb{Q}^*$ a.s., we conclude that B^* is a $\mathcal{G}_t^*$-local martingale. For this, set, $\tilde{B}_{i,M}(\cdot) := \tilde{B}_i(\cdot \wedge \tilde{\pi}_{i,M})$. We use the optional sampling theorem given in Ethier and Kurtz (1986), theorem 2.2.13, which (in adaptation to our notation) states that, if for every $t \in \mathbb{R}_+$,

$$\mathbb{E}^{\mathbb{Q}^*}|\tilde{B}_{i,M}(\tau(t \wedge \pi_{i,M}))| < \infty \tag{102}$$

and

$$\lim_{T\to\infty} \mathbb{E}^{\mathbb{Q}^*}\big[|\tilde{B}_{i,M}(T)| \mathbb{1}_{\{\tau(t) > T\}}\big] = 0, \tag{103}$$

then, for every $0 \le s \le t$,

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}^*}\big[\tilde{B}_{i,M}(\tau(t\wedge\pi_{i,M}))\mid \mathcal{G}^*(s\wedge\pi_{i,M})\big] &= \mathbb{E}^{\mathbb{Q}^*}\big[\tilde{B}_{i,M}(\tau(t\wedge\pi_{i,M}))\mid \tilde{\mathcal{G}}(\tau(s\wedge\pi_{i,M}))\big]\\ &= \tilde{B}_{i,M}(\tau(s\wedge\pi_{i,M})),\end{aligned}$$

where we used the identity $\mathcal{G}^*(t) = \tilde{\mathcal{G}}(\tau(t))$. Therefore, $\tilde{B}_{i,M}(\tau(t\wedge\pi_{i,M}))$ is a $\mathcal{G}^*_t$ martingale. Notice that

$$\begin{aligned}B^*_i(t\wedge\pi_{i,M}) &= \tilde{B}_i(\tau(t\wedge\pi_{i,M})) = \tilde{B}_i(\tau(t)\wedge\tau(\pi_{i,M})) = \tilde{B}_i(\tau(t)\wedge\tau(\pi_{i,M})\wedge\tilde{\pi}_{i,M})\\ &= \tilde{B}_i(\tau(t\wedge\pi_{i,M}))\wedge\tilde{\pi}_{i,M}) = \tilde{B}_{i,M}(\tau(t\wedge\pi_{i,M}))\end{aligned}$$

and, therefore, $B^*_i(t\wedge\pi_{i,M})$ is a $\mathcal{G}^*_t$ martingale. Indeed, the first and last equalities follow by the definitions of B^* and $\tilde{B}_{i,M}$, respectively. The second and fourth equalities follow by the monotonicity of the function $t\mapsto\tau(t)$. Finally, the third equality follows because $\tau(\pi_{i,M}) = \tilde{\pi}_{i,M}$.

We now prove that properties (102) and (103) hold. Property (102) follows by the definition of $\tilde{B}_{i,M}$. To prove property (103), notice that

$$\mathbb{E}^{\mathbb{Q}^*}\big[|\tilde{B}_{i,M}(T)|\mathbb{1}_{\{\tau(t)>T\}}\big] \le \mathbb{E}^{\mathbb{Q}^*}\big[|\tilde{B}_{i,M}(T)|\mathbb{1}_{\{\tilde{\tau}(T)\le t\}}\big] \le M\mathbb{Q}^*(\tilde{\tau}(T)\le t).$$

Now, from Lemma 3, the left-hand side of the preceding approaches zero as $T\to\infty$, and (103) is proven.

We end the proof by providing the missing arguments for the proof that $B^*(t)(B^*(t))^\top - \mathcal{I}t$ is a martingale. One may go over the proof and replace the $\tilde{B}_i(t)$s, $i\in[I]$ with $\tilde{N}_{ij}(t) = \tilde{B}_i(t)\tilde{B}_j(t) - \delta_{ij}t$, $i,j\in[I]$. The only difference between the proofs lies in proving that $\tilde{N}(t) = \tilde{B}(t)(\tilde{B}(t))^\top - \mathcal{I}t$ is a $\tilde{\mathcal{G}}_t$ martingale or, equivalently, in showing that (101) holds with $\tilde{N}_{ij}$, replacing $\tilde{B}_i$. To this end, recall that $\hat{\mathbb{Q}}^n = \prod_{i=1}^I(\hat{\mathbb{Q}}^n_{1,i}\times\hat{\mathbb{Q}}^n_{2,i})$. Thus, $\{A^n_1,\dots,A^n_I,D^n_1,\dots,D^n_I\}$ are mutually independent under $\hat{Q}^n$ and, therefore, also under $\mathbb{Q}^*$. Moreover, using the continuity of $\tilde{\tau}^n$ and the notation $\Delta L(t) = L(t) - L(t-)$, we get that, for any $i,j\in[I]$, $i\ne j$, the following equalities hold $\mathbb{Q}^*$-a.s.,

$$[\tilde{B}^n_i,\tilde{B}^n_j](t) = \sum_{0\le s\le t}\Delta\tilde{B}^n_i(s)\Delta\tilde{B}^n_j(s) = \frac{1}{n}\sum_{0\le s\le t}\Delta B^n_i(\tilde{\tau}^n(s))\Delta B^n_j(\tilde{\tau}^n(s)) = 0$$

and

$$\begin{aligned}[\tilde{B}^n_i,\tilde{B}^n_i](t) = [\tilde{B}^n_i](t) &= \sum_{0\le s\le t}(\Delta\tilde{B}^n_i(s))^2 = \frac{1}{n\hat{\sigma}^2_i}\sum_{0\le s\le t}(\Delta A^n_i(\tilde{\tau}^n(s)))^2 + (\Delta D^n_i(\tilde{\tau}^n(s)))^2\\ &= \frac{1}{2n\lambda_i}\big(A^n_i(\tilde{\tau}^n(t)) + D^n_i(\tilde{\tau}^n(t))\big)\xrightarrow[n\to\infty]{}\tilde{\tau}(t),\end{aligned}$$

where the limit holds by the strong law of large numbers. Because $\tilde{B}^n_i(t)\tilde{B}^n_j(t) - [\tilde{B}^n_i,\tilde{B}^n_j](t)$ is a $\mathcal{G}^n(\tilde{\tau}(t))$ martingale, by taking the limit $n\to\infty$, one deduces that (101) holds with $\tilde{N}_{ij}$ replacing $\tilde{B}_i$. □

Proof of Lemma 4. From Proposition 3, it follows immediately that, for every $T>0$,

$$\lim_{K\to\infty}\lim_{n\to\infty}\hat{\mathbb{Q}}\Big(\|(\hat{\Psi}^n_1,\hat{\Psi}^n_2)\|_T\ge K\Big) = 0.$$

Fix $\delta,\eta,K>0$, and for every $n\in\mathbb{N}$, set

$$P^n := \int_0^t e^{-\varrho t}\Big((\hat{\psi}^n_{1,i}(s))^2ds + (\hat{\psi}^n_{2,i}(s))^2dT^n_i(s)\Big). \tag{104}$$

Clearly, for any $K>0$,

$$\hat{\mathbb{Q}}\big(\mathrm{osc}_T((\hat{\Psi}^n_1,\hat{\Psi}^n_2),\delta)\ge\eta\big) = \hat{\mathbb{Q}}\big(\mathrm{osc}_T((\hat{\Psi}^n_1,\hat{\Psi}^n_2),\delta)\ge\eta, P^n>K\big) + \hat{\mathbb{Q}}\big(\mathrm{osc}_T((\hat{\Psi}^n_1,\hat{\Psi}^n_2),\delta)\ge\eta, P^n\le K\big). \tag{105}$$

Proposition 3 implies that

$$\lim_{K\to\infty}\limsup_{\delta\to 0+}\limsup_{n\to\infty}\hat{\mathbb{Q}}\big(\mathrm{osc}_T((\hat{\Psi}^n_1,\hat{\Psi}^n_2),\delta)\ge\eta, P^n>K\big)\le\lim_{K\to\infty}\limsup_{n\to\infty}\hat{\mathbb{Q}}(P^n>K) = 0. \tag{106}$$

We now examine the second term on the r.h.s. of (105). In the event $\{P^n \le K\}$, Jensen's inequality implies that there exists a constant $C_T > 0$, independent of n, such that, for every $0 \le s < t \le T$,

$$\|(\hat{\Psi}_1^n, \hat{\Psi}_2^n)(t) - (\hat{\Psi}_1^n, \hat{\Psi}_2^n)(s)\|^2 \le C_T(t-s)P^n \le C_T K(t-s).$$

As a result,

$$\lim_{\delta\to 0+}\lim_{n\to\infty} \hat{\mathbb{Q}}\Big(\mathrm{osc}_T\big((\hat{\Psi}_1^n, \hat{\Psi}_2^n), \delta\big) \ge \eta, P^n \le K\Big) = 0.$$

Combining it with (106), we get

$$\lim_{\delta\to 0+}\lim_{n\to\infty} \hat{\mathbb{Q}}\Big(\mathrm{osc}_T\big((\hat{\Psi}_1^n, \hat{\Psi}_2^n), \delta\big) \ge \eta\Big) = 0,$$

and the $\mathscr{C}$-tightness is established. □

Proof of Lemma 5. By (65), we have

$$\|\hat{Y}^n\|_T \le \|\hat{X}^n\|_T + \|\check{A}^n\|_T + \|\check{D}^n\|_T + \|\hat{m}^n\|T + C_2\|(\hat{\Psi}_1^n, \hat{\Psi}_2^n)\|_T + \|\hat{R}^n\|_T, \tag{107}$$

where C_2 depends solely on $\{(\lambda_i, \mu_i)\}_i$. From (57), it follows that there exists a constant $C_3 > 0$, independent of n and T, such that, for every $T \in \mathbb{R}_+$,

$$\|\hat{R}^n\|_T \le C_3\Big(1 + T + \|\check{A}^n\|_T + \|\check{D}^n\|_T + \big\|(\hat{\Psi}_1^n, \hat{\Psi}_2^n)\big\|_T\Big). \tag{108}$$

Recall that $\|\hat{X}^n\|_T$ is bounded, and from Lemma 4, $\|(\hat{\Psi}_1^n, \hat{\Psi}_2^n)\|_T$ is tight. Then, once we show that $\{(\check{A}^n, \check{D}^n)\}_n$ is $\mathscr{C}$-tight, from (107) and (108), we get that $\{\|\hat{Y}^n\|_T\}_n$ is tight. Now, by the definitions of $\hat{Y}^n$ and μ^n—see (7) and (3)—we get that $\{T_i^n\}_n$ converges u.o.c. to $\bar{T}_i$, where $\bar{T}_i(t) = \rho_i t$.

We now argue the $\mathscr{C}$-tightness of $\{(\check{A}^n, \check{D}^n)\}_n$. From (58) and Proposition 3, it follows that, for every $T > 0$,

$$\lim_{K\to\infty}\lim_{n\to\infty} \hat{\mathbb{Q}}\big(\|(\check{A}^n, \check{D}^n)\|_T \ge K\big) = 0.$$

As argued in the proof of Lemma 4, it is sufficient to prove that, for every $K > 0$,

$$\lim_{\delta\to 0+}\lim_{n\to\infty} \hat{\mathbb{Q}}\big(\mathrm{osc}_T((\check{A}^n, \check{D}^n), \delta) \ge \eta, P^n \le K\big) = 0,$$

where P^n is given in (104). This follows because, under the event $\{P^n < K\}$, the jumps of $(\check{A}^n, \check{D}^n)$ are of order $n^{-1/2}$, and therefore, the limit process is continuous.

Finally, the $\mathscr{C}$-tightness of $\{\hat{W}^n\}_n$ follows now by its definition. □

Proof of Lemma 6. By the definition of the candidate policy, rejections do not occur when $X^{\circ,n} < a$ and the policy is work-conserving; namely, $\sum_{i=1}^I U_i^n(t) = 1$ whenever $\hat{X}^n(t) > 0$. Therefore,

$$\int_0^{\cdot} \mathbb{1}_{\{X^{\circ,n}(t) < a\}} dR^{\circ,n}(t) = \int_0^{\cdot} \mathbb{1}_{\{X^{\circ,n}(t) > 0\}} dY^{\circ,n}(t) = 0,$$

and (69) holds.

Recall that $\{(\Psi_1^{\sharp,n}, \Psi_2^{\sharp,n}, \hat{W}^n)\}_n$ is $\mathscr{C}$-tight and, therefore, also $\{(\Psi_1^{\circ,n}, \Psi_2^{\circ,n}, W^{\circ,n})\}_n$. Hence, in order to show the $\mathscr{C}$-tightness of all the processes as mentioned in the lemma, it is sufficient to prove that $E^n \Rightarrow 0$. Fix $T > 0$. It is sufficient to show that, as $n \to \infty$,

$$\left(\sup_{t\in[0,T]} X^{\circ,n}(t) - a\right)^+ \Rightarrow 0. \tag{109}$$

For $\nu > 0$, consider the event $\Omega_1^n := \{\sup_{t\in[0,T]} X^{\circ,n}(t) > a + \nu\}$. On this event, there exist random times $0 \le \tau_1^n < \tau_2^n \le \tau^n$ such that $X^{\sharp,n}(\tau_1^n) \le a + \nu/2$, $X^{\sharp,n}(\tau_2^n) \ge a + \nu$, and $X^{\sharp,n}(t) > a$ for every $t \in [\tau_1^n, \tau_2^n]$. Using the notation $L[s,t] = L(t) - L(s)$ for every process L and $0 \le s \le t$, by (66) and the fact that $Y^{\sharp,n}$ does not increase on an interval for which the system is not empty, we have

$$\begin{aligned}(a+\nu) - (a+\nu/2) &\le X^{\sharp,n}[\tau_1^n, \tau_2^n] \\ &= (W^{\sharp,n} + \sigma(\Psi_1^{\sharp,n} - \Psi_2^{\sharp,n}))[\tau_1^n, \tau_2^n] - R^{\sharp,n}[\tau_1^n, \tau_2^n] \\ &= (W^{\sharp,n} + \sigma(\Psi_1^{\sharp,n} - \Psi_2^{\sharp,n}))[\tau_1^n, \tau_2^n] - n^{-1/2}A_{i^*}^n[\tau_1^n, \tau_2^n],\end{aligned}$$

where we used the fact that the policy rejects all class-i^* arrivals when $X^{\sharp,n} > a^*$. Fix a sequence $r^n > 0$, $r^n \to 0$, such that $n^{1/2}r^n \to \infty$. In case $\tau_2^n - \tau_1^n < r^n$, one has

$$\nu/2 \le (W^{\sharp,n} + \sigma(\Psi_1^{\sharp,n} - \Psi_2^{\sharp,n}))[\tau_1^n, \tau_2^n] \le \mathrm{osc}_T(W^{\sharp,n}, r^n) + \mathrm{osc}_T(\sigma(\Psi_1^{\sharp,n} - \Psi_2^{\sharp,n}), r^n)$$

and in case $\tau_2^n - \tau_1^n \ge r^n$, one has

$$\begin{aligned} 2\Big(\big\|\sigma\big(\Psi_1^{\sharp,n} - \Psi_2^{\sharp,n}\big)\big\|_T + \|W^{\sharp,n}\|_T\Big) &\ge A_{i^*}^n[\tau_1^n, \tau_2^n]n^{-1/2} = \check{A}_{i^*}^n[\tau_1^n, \tau_2^n] + n^{-1/2}\int_{\tau_1^n}^{\tau_2^n} \psi_{1,i^*}^n(t)dt \\ &\ge -2\|\check{A}_{i^*}^n\|_T + \bar{C}_T r^n n^{1/2}, \end{aligned}$$

for some $\bar{C}_T > 0$, independent of n. The last inequality follows because the last intergal must be greater than $\bar{C}_T r^n n^{1/2}$ for some $\bar{C}_T > 0$. Otherwise, there is a sequence of intervals $\{I^n\}_n$ with lengths r^n such that $\lim_{n\to\infty}(r^n n)^{-1}\int_{I^n} \psi_{1,i^*}^n(t)dt = 0$ (recall that $\hat{\psi}_{1,i}^n > 0$). This implies that $\limsup_n (r^n n)^{-1}\int_{I^n}(\lambda_{i^*}n)^{1/2}\hat{\psi}_{1,i^*}^n(t)dt < 0$, and therefore, for sufficiently large n, $\int_{I^n}\hat{\psi}_{1,i^*}^n(t)dt \le -C_4 r^n n^{1/2}$ for some $C_4 > 0$ independent of n. This implies that $\int_{I^n}|\hat{\psi}_{1,i^*}^n(t)|dt \ge C_4 r^n n^{1/2} \to \infty$ as $n \to \infty$ in contradiction to Proposition 3. Therefore, the last inequality holds.

The tightness of $\{\check{A}^n\}_n$ (see Lemma 4) and the $\mathscr{C}$-tightness of $\{(W^{\sharp,n}, \Psi_1^{\sharp,n}, \Psi_2^{\sharp,n})\}_n$ (Lemma 5) imply that

$$\begin{aligned} \lim_{n\to\infty}\Big[\hat{\mathbb{Q}}\Big(\mathrm{osc}_T(W^{\sharp,n}, r^n) \ge \nu/2\Big) + \hat{\mathbb{Q}}\Big(\mathrm{osc}_T\big(\sigma\big(\Psi_1^{\sharp,n} - \Psi_2^{\sharp,n}\big), r^n\big) \ge \nu/2\Big) \\ + \hat{\mathbb{Q}}\Big(2\{\|(\Psi_1^{\sharp,n}, \Psi_2^{\sharp,n})\|_T + \|W^{\sharp,n}\|_T + \|\hat{A}_{i^*}^n\|_T\} \ge \bar{C}_T r^n n^{1/2}\Big)\Big] = 0. \end{aligned}$$

and thus, $\lim_{n\to\infty}\hat{\mathbb{Q}}(\Omega_1^n) = 0$. Because $\nu > 0$ was arbitrary, we obtain (70). □

Proof of Proposition 4. This is the equivalent of equation 67 and the conclusion in step 3 in Atar and Shifrin (2014). The proof there is given in step 2. We now provide an adaptation of Atar and Shifrin (2014), step 2, to our case. For simplicity, we share most of the notation and provide the claims in the same order.

Denote by $\mathscr{G} = \{\hat{x} \in \mathscr{X} : \theta \cdot \hat{x} \le a, \hat{x} = \gamma^a(\theta \cdot \hat{x})\}$ the set of points lying on the minimizing curve, and $\partial^+\mathscr{X} = \{\hat{x} \in \mathscr{X} : \hat{x}_i = \hat{b}_i \text{ for some } i\}$. These two sets are compact and disjoint. Hence, there exists $\nu_0 > 0$ such that, for any $0 < \nu' < \nu_0$, $\mathscr{G}^{\nu'} \cap (\partial^+\mathscr{X})^{\nu'} = \emptyset$, where, for a set $F \in \mathbb{R}^I$, we denote

$$F^{\nu'} = \{\hat{x} : \mathrm{dist}(\hat{x}, F) \le \nu'\}.$$

In the rest of the proof, we consider only ν' strictly smaller than ν_0. For sufficiently large n, forced rejections occur only when $\hat{X}^n$ lies in $(\partial^+\mathscr{X})^{\nu'}$. As a result, as long as the process $\hat{X}^n$ lies in $\mathscr{G}^{\nu'}$, no forced rejections occur. Therefore, $\sigma^n \le \tau^n$, where

$$\sigma^n = \hat{\zeta}^n \wedge \zeta^n, \qquad \hat{\zeta}^n = \inf\{t : X^{\sharp,n} \ge a + \nu'\}, \qquad \zeta^n = \inf\{t : \max_{i\in[I]}|\hat{X}_i^n(t) - \gamma_i^a(X^{\sharp,n}(t))| \ge \nu'\}.$$

As a result, in order to prove the first limit in the lemma, it is sufficient to show that $\hat{\mathbb{Q}}(\sigma^n < T) \to 0$ for any small $\nu' > 0$ and any T. Fix ν' and T. Because $\sigma^n \le \tau^n$,

$$\begin{aligned} \hat{\mathbb{Q}}(\sigma^n < T) = \hat{\mathbb{Q}}(\sigma^n < T, \sigma^n \le \tau^n) &\le \hat{\mathbb{Q}}(\hat{\zeta}^n \wedge \zeta^n \le T \wedge \tau^n) \\ &\le \hat{\mathbb{Q}}(\hat{\zeta}^n \le T \wedge \tau^n) + \hat{\mathbb{Q}}(\zeta^n \le T \wedge \tau^n). \end{aligned}$$

From (109), it follows that $\hat{\mathbb{Q}}(\hat{\zeta}^n \le T \wedge \tau^n) \to 0$ as $n \to \infty$. It, therefore, suffices to prove that

$$\lim_{n\to\infty}\hat{\mathbb{Q}}(\zeta^n \le T \wedge \tau^n) \to 0. \tag{110}$$

On $\zeta^n \le T \wedge \tau^n$, let $x^n := X^{\sharp,n}(\zeta^n) = X^{\circ,n}(\zeta^n)$, and let $j = j^n$ and ξ^n be the corresponding components from the representation (j, ξ) of x^n given in (24).

Fix a positive integer $K = K(\nu')$. Consider the covering of $[0, b]$ by the $K - 1$ intervals $\Xi_k = \boldsymbol{B}(k\nu_1, \nu_1)$, $k = 1, 2, \ldots, K - 1$, where $\boldsymbol{B}(x, a)$ denotes $[x - a, x + a]$ and $\nu_1 = b/K$. Let also $\tilde{\Xi}_k = \boldsymbol{B}(k\nu_1, 2\nu_1)$.

Recall that $X^{\circ,n}$ are C-tight. Hence, given $\delta > 0$, there exists $\delta' = \delta'(\delta, T, \nu_1) > 0$ such that, for all sufficiently large n,

$$|X^{\circ,n}(s) - X^{\circ,n}(t)| \le \nu_1 \text{ for all } s, t \in [0, T], |s - t| \le \delta', \text{with probability at least } 1 - \delta. \tag{111}$$

Fix such δ and δ'. Denote by $\boldsymbol{T}^n := [(\zeta^n - \delta' \vee 0), \zeta^n]$, and set

$$\Omega^{n,k} := \{\zeta^n \le T \wedge \tau^n, x^n \in \Xi_k, X^{\sharp,n}(t) \in \tilde{\Xi}_k \text{ for all } t \in \boldsymbol{T}^n\}.$$

Then, for all large n,

$$\hat{\mathbb{Q}}(\zeta^n \le T \wedge \tau^n) \le \delta + \sum_{k=1}^{K-1} \hat{\mathbb{Q}}(\Omega^{n,k}), \tag{112}$$

where we used the identity $X^{\sharp,n} = X^{\circ,n}$ on $[0,\tau^n]$. We fix $k \in \{1,\dots,K-1\}$ and use a similar (but more advanced) argument to the one given in the proof of Lemma 6 to analyze $\Omega^{n,k}$.

The value assigned by the policy to U^n (see (27)) remains fixed as $\hat{X}^n$ varies within any of the intervals $(\hat{\alpha}_j, \hat{\alpha}_{j+1})$, where $\hat{\alpha}_i := \sum_{k=i+1}^{I} \theta_k \hat{a}_k$, $i \in [I]$. We now provide four separate cases, and under each one of them, for each k, $\lim_{n\to\infty} \Omega^{n,k} = 0$:

I. $\tilde{\Xi}_k \subset (0,a)$, and for all j, $\hat{\alpha}_j \notin \tilde{\Xi}_k$. Then we consider the following cases.
II. $\tilde{\Xi}_k \subset (0,a)$, but $\hat{\alpha}_j \in \tilde{\Xi}_k$ for some $j \in \{1,2,\dots,I-1\}$.
III. $0 \in \tilde{\Xi}_k$.
IV. $a \in \tilde{\Xi}_k$.

There may be additional intervals $\tilde{\Xi}_k$, but they are all subsets of (a,∞) and, therefore, not important for our purpose.

We analyze only case I (and afterward comment on the other ones). This means that, in the representation of γ^a, the jth component is the same for all the points $x \in \tilde{\Xi}_k$. Note that $j := j(k)$ depends on k only and, in particular, does not vary with n.

Fix $i \in \{j+1,\dots,I\}$ (unless $i = I$). We estimate the probability that, on $\Omega^{n,k}$, $\zeta^n \le T \wedge \tau^n$ occurs by having $\hat{X}^n_i(\zeta^n) - \gamma^a_i(X^{\sharp,n}(\zeta^n)) \ge \nu'$. More precisely, because $i > j$, $\gamma^a_i(x^n) = a_i$. Then we show that

$$\text{for every } \nu'' \in (0,\nu'), \quad \hat{\mathbb{Q}}(\Omega^{n,k} \cap \{\hat{X}^n_i(\zeta^n) > a_i + \nu''\}) \to 0 \quad \text{as } n \to \infty. \tag{113}$$

Recall the convergence of the initial condition given in (49) and that γ^a is continuous. Now, the jumps of $\hat{X}^n$ are of size $n^{-1/2}$, and therefore, in the event indicated in (113), there must exist $\eta^n \in [0,\zeta^n]$ with the properties that

$$\hat{X}^n_i(\eta^n) < a_i + \nu''/2, \qquad \hat{X}^n_i(t) > a_i \text{ for all } t \in [\eta^n, \zeta^n]. \tag{114}$$

In this event, during the time interval $[\eta^n,\zeta^n]$, i is always a member of $\mathscr{H}(\hat{X}^n)$, and therefore, by (28) and (29), $U^n_i(t) = \rho'_i(\hat{X}^n(t)) > \rho_i + c$ for some constant $c > 0$, independent of n. By the definition of $\hat{Y}^n_i$ from (7), $\frac{d}{dt}\hat{Y}^n_i \le -\frac{\mu^n_i}{\sqrt{n}}c$. Set $\hat{\eta}^n = \eta^n \vee (\zeta^n - \delta')$. Then, for every $t \in [\hat{\eta}^n,\zeta^n]$, one has $\hat{X}^n(t) \in \tilde{\Xi}_k \subset (0,a)$, and therefore, no rejections occur. Combining these facts with (65) and the definitions of $\check{A}^n, \check{D}^n$, and $\hat{W}^n$ given in (41), (42), and (62), we have

$$\hat{X}^n_i[\hat{\eta}^n,\zeta^n] = \left(\hat{W}^n_i + \lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}\right)[\hat{\eta}^n,\zeta^n] - c\frac{\mu^n_i}{\sqrt{n}}(\zeta^n - \hat{\eta}^n), \tag{115}$$

where[3] we used the notation $L[t,s] = L(t) - L(s)$ for any process L, and $0 \le s \le t$. As in the proof of Lemma 6, fix a sequence $r^n > 0$ with $r^n \to 0$ and $r^n\sqrt{n} \to \infty$. If $\zeta^n - \eta^n < r^n$ and n is sufficiently large, then $\hat{\eta}^n = \eta^n$; thus, by (113) and the definition of η^n, $\hat{X}^n_i[\hat{\eta}^n,\zeta^n] \ge \nu''/2$. As a result,

$$\mathrm{osc}_T\left(\left(\lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}\right), r^n\right) + \mathrm{osc}_T(\hat{W}^n_i, r^n) \ge \left(\hat{W}^n_i + \lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}\right)[\eta^n,\zeta^n] \ge \nu''/2$$

must hold. If, on the other hand, $\zeta^n - \eta^n \ge r^n$, then by (115),

$$2\left(\|\sigma(\hat{\Psi}^n_{1,i} - \hat{\Psi}^n_{2,i})\|_T + \|\hat{W}^n_i\|_T\right) \ge \left(\hat{W}^n_i + \lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}\right)[\hat{\eta}^n,\zeta^n] \ge c\frac{\mu^n_i}{\sqrt{n}}r^n \ge cr_n\sqrt{n},$$

for some constant $c > 0$. Hence, the probability in (113) is bounded above by

$$\hat{\mathbb{Q}}\left(\mathrm{osc}_T\left(\lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}, r^n\right) \ge \nu''/4\right) + \hat{\mathbb{Q}}\left(\mathrm{osc}_T(\hat{W}^n_i, r^n) \ge \nu''/4\right)$$
$$+ \hat{\mathbb{Q}}\left(2\left\{\|\lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}\|_T + \|\hat{W}^n_i\|_T\right\} \ge cr_n\sqrt{n}\right),$$

which converges to zero as $n \to \infty$, by the C-tightness of $\{\hat{W}^n\}_n$, and $\left\{\lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}\right\}_n$.

The rest of the proof follows by the same lines as in Atar and Shifrin (2014), where again $\hat{W}^n$ is replaced by $\hat{W}^n_i + \lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}$. The properties needed are the $\mathscr{C}$-tightness of $\{X^{\circ,n}\}_n$ and $\{\hat{W}^n\}_n$ (Lemmas 5 and 6); the uniform continuity of γ^a; the convergence $\theta^n \to \theta$; and of the initial condition given in (68), the boundedness of $\mathscr{X}$ and that the jumps of $\hat{X}^n$ are of size $n^{-1/2}$. □

Acknowledgments

The author is grateful 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.

Endnotes

[1] Notice that this is a stochastic game, in which the minimizer chooses a strategy and the maximizer, who knows the minimizer's choice, responds by choosing a worst-case scenario. Also, we do not consider a statistical inference learning problem.

[2] At this point, it is worth mentioning that Atar and Shifrin (2014) managed to bypass this issue by arguing the $\mathscr{C}$-tightness of the integrands of the relevant processes. Repeating the same arguments, one may show that the integrands of $\hat{X}^n_i$, $\hat{Y}^n_i$, and $\hat{R}^n_i$ are $\mathscr{C}$-tight. Because V' is bounded (see Proposition 1), we get that the sequence $\{\int_0^t(\hat{\psi}^n_{1,i}(s) - \hat{\psi}^n_{2,i}(s))ds\}_n$ is $\mathscr{C}$-tight. However, because we wish to obtain (after projecting on the workload vector) the specific integral given in (53), we still need to argue tightness of $\hat{X}^n$.

[3] This is where the proof given in Atar and Shifrin (2014) requires modification: $\hat{W}^n$ is replaced by $\hat{W}^n_i + \lambda_i^{1/2}\hat{\Psi}^n_{1,i} - \mu_i^{1/2}\hat{\Psi}^n_{2,i}$.

References

Atar R, Biswas A (2014) Control of the multiclass $G/G/1$ queue in the moderate deviation regime. *Ann. Appl. Probab.* 24(5):2033–2069.
Atar R, Cohen A (2016) A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime. *Math. Oper. Res.* 41(4):1354–1380.
Atar R, Cohen A (2017) Asymptotically optimal control for a multiclass queueing model in the moderate deviation heavy traffic regime. *Ann. Appl. Probab.* 27(5):2862–2906.
Atar R, Mendelson G (2016) On the non-Markovian multiclass queue under risk-sensitive cost. *Queueing Systems* 84(3-4):265–278.
Atar R, Saha S (2017) Optimality of the generalized cμ rule in the moderate deviation regime. *Queueing Systems* 87(1):113–130.
Atar R, Shifrin M (2014) An asymptotic optimality result for the multiclass queue with finite buffers in heavy traffic. *Stochastic Systems* 4(2): 556–603.
Bandi C, Bertsimas D, Youssef N (2015) Robust queueing theory. *Oper. Res.* 63(3):676–700.
Bassamboo A, Randhawa RS, Zeevi A (2010) Capacity sizing under parameter uncertainty: Safety staffing principles revisited. *Management Sci.* 56(10):1668–1686.
Bayraktar E, Zhang Y (2015) Minimizing the probability of lifetime ruin under ambiguity aversion. *SIAM J. Control Optim.* 53(1):58–90.
Bell SL, Williams RJ (2001) 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.
Billingsley P (1999) *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics, 2nd ed. (John Wiley & Sons Inc., New York).
Biswas A (2014) Risk-sensitive control for the multiclass many-server queues in the moderate deviation regime. *Math. Oper. Res.* 39(3):908–929.
Blanchet J, Murthy KR (2019) Quantifying distributional model risk via optimal transport. *Math. Oper. Res.* 44(2):565–600.
Blanchet J, Dolan C, Lam H (2014) Robust rare-event performance analysis with natural non-convex constraints. Tolk A, Yilmaz L, Diallo SY, Ryzhov IO, eds. *Proc. 2014 Winter Simulation Conf.* (IEEE Press, Piscataway, NJ), 595–603.
Budhiraja A, Ghosh AP (2006) Diffusion approximations for controlled stochastic networks: An asymptotic bound for the value function. *Ann. Appl. Probab.* 16(4):1962–2006.
Budhiraja A, Ghosh AP (2012) Controlled stochastic networks in heavy traffic: Convergence of value functions. *Ann. Appl. Probab.* 22(2):734–791.
Budhiraja A, Ross K (2006) Existence of optimal controls for singular control problems with state constraints. *Ann. Appl. Probab.* 16(4):2235–2255.
Budhiraja A, Ross K (2007) Convergent numerical scheme for singular stochastic control with state constraints in a portfolio selection problem. *SIAM J. Control Optim.* 45(6):2169–2206.
Cohen A (2017) Asymptotic analysis of a multiclass queueing control problem under heavy-traffic with model uncertainty. Preprint, submitted October 3, https://arxiv.org/abs/1710.00968.
Cohen A (2018) Brownian control problems for a multiclass M/M/1 queueing problem with model uncertainty. *Math. Oper. Res.* 44(2):739–766.
Dai JG, Williams RJ (1996) Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons. *Theory Probab. Appl.* 40(1):1–40.
Dellacherie C, Meyer PA (1978) *Probabilities and Potential*, North-Holland Mathematics Studies, vol. 29 (North-Holland Publishing Co., Amsterdam-New York).
Dupuis P (2003) Explicit solution to a robust queueing control problem. *SIAM J. Control Optim.* 42(5):1854–1875.
Ethier SN, Kurtz TG (1986) *Markov Processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics (John Wiley & Sons Inc., New York).
Hansen LP, Sargent TJ (2008) *Robustness* (Princeton University Press, Princeton, NJ).
Hansen LP, Sargent TJ, Turmuhambetova G, Williams N (2006) Robust control and model misspecification. *J. Econom. Theory* 128(1):45–90.
Harrison JM (1988) Brownian models of queueing networks with heterogeneous customer populations. Fleming WH, Lions P-L, eds. *Stochastic Differential Systems, Stochastic Control Theory and Applications*, IMA Volumes in Mathematics and Its Applications, vol. 10 (Springer, New York), 147–186.

Jain A, Lim AEB, Shanthikumar JG (2010) On the optimality of threshold control in queues with model uncertainty. *Queueing Systems* 65(2): 157–174.

Karatzas I, Shreve SE (1991) *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, vol. 113, 2nd ed. (Springer-Verlag, New York).

Kruk L, Lehoczky J, Ramanan K, Shreve S (2007) An explicit formula for the Skorokhod map on $[0, a]$. *Ann. Probab.* 35(5):1740–1768.

Kurtz TG (1991) Random time changes and convergence in distribution under the Meyer-Zheng conditions. *Ann. Probab.* 19(3):1010–1034.

Kushner HJ, Martins LF (1991) Numerical methods for stochastic singular control problems. *SIAM J. Control Optim.* 29(6):1443–1475.

Kushner HJ, Martins LF (1993) Limit theorems for pathwise average cost per unit time problems for controlled queues in heavy traffic. *Stochastics Stochastics Rep.* 42(1):25–51.

Lam H (2016) Robust sensitivity analysis for stochastic systems. *Math. Oper. Res.* 41(4):1248–1275.

Maenhout PJ (2004) Robust portfolio rules and asset pricing. *Rev. Financial Stud.* 17(4):951–983.

Martins LF, Kushner HJ (1990) Routing and singular control for queueing networks in heavy traffic. *SIAM J. Control Optim.* 28(5):1209–1233.

Meyer PA, Zheng WA (1984) Tightness criteria for laws of semimartingales. *Ann. Inst. H. Poincaré Probab. Statist.* 20(4):353–372.

Petersen IR, James MR, Dupuis P (2000) Minimax optimal control of stochastic uncertain systems with relative entropy constraints. *IEEE Trans. Automatic. Control* 45(3):398–412.

Plambeck E, Kumar S, Harrison JM (2001) A multiclass queue in heavy traffic with throughput time constraints: Asymptotically optimal dynamic controls. *Queueing Systems* 39(1):23–54.

Protter PE (2004) *Stochastic Integration and Differential Equations*, Applications of Mathematics (New York), vol. 21, 2nd ed. (Springer-Verlag, Berlin).

Stroock DW (1987) *Lectures on Stochastic Analysis: Diffusion Theory*, London Mathematical Society Student Texts, vol. 6 (Cambridge University Press, Cambridge, UK).

Whitt W (2006) Staffing a call center with uncertain arrival rate and absenteeism. *Production Oper. Management* 15(1):88–102.

Whitt W, You W (2018) Using robust queueing to expose the impact of dependence in single-server queues. *Oper. Res.* 66(1):184–199.