The Annals of Applied Probability
2017, Vol. 27, No. 5, 2862–2906
DOI: 10.1214/16-AAP1269

ASYMPTOTICALLY OPTIMAL CONTROL FOR A MULTICLASS QUEUEING MODEL IN THE MODERATE DEVIATION HEAVY TRAFFIC REGIME[1]

BY RAMI ATAR AND ASAF COHEN

Technion–Israel Institute of Technology and University of Michigan

A multi-class single-server queueing model with finite buffers, in which scheduling and admission of customers are subject to control, is studied in the moderate deviation heavy traffic regime. A risk-sensitive cost set over a finite time horizon $[0, T]$ is considered. The main result is the asymptotic optimality of a control policy derived via an underlying differential game. The result is the first to address a queueing control problem at the moderate deviation regime that goes beyond models having the so-called pathwise minimality property. Moreover, despite the well-known fact that an optimal control over a finite time interval is generically of a nonstationary feedback type, the proposed policy forms a stationary feedback, provided T is sufficiently large.

1. Introduction. This paper continues a line of research started in [1] that aims at analyzing queueing control problems (QCPs) at the moderate deviation (MD) heavy traffic regime. The model under consideration consists of a server that serves customers from a number of classes, where allocation of the effort among classes is dynamically controlled. Customers are kept in buffers of finite size, one buffer for each class, and those that arrive to find a full buffer are lost. It is also possible to reject arrivals when buffers are not full. This control system is considered with a risk-sensitive (RS) cost, that accounts for holding of customers in the buffers as well as for rejections. At the heart of the analysis lies a differential game (DG) that has been analyzed in [2]. This paper proves the validity of the prediction of [2] that the DG governs the scaling limit, by showing that the QCP's value converges to the DG's value, and identifying an asymptotically optimal (AO) policy for the former that is constructed based on the latter. The limit result for the model treated in [1] was built on *pathwise minimality*, a property that considerably simplifies the analysis, which does not hold in our setting. Instead, the proof here is based on the Bellman (or the dynamic programming) equation and, specifically, a *free boundary point* characterized by it governs the asymptotic behavior.

Traditionally, heavy traffic analysis of queueing models, and particularly QCPs, is carried out under the regime of *diffusion-scale deviation* (sometimes also referred to as *ordinary deviation*), but it is also relevant at the MD scale, where

Received October 2015; revised September 2016.
[1]Supported in part by the ISF (Grant 1315/12).
MSC2010 subject classifications. 49N70, 60F10, 60K25, 93E20.
Key words and phrases. Moderate deviations, heavy traffic, risk-sensitive cost, differential games.

relatively few results exist [16, 20, 21]. The roots of *large deviation* (LD) analysis of control systems go back to Fleming [11], who studies the associated Hamilton–Jacobi equations. The connection of RS cost to DG was made by Jacobson [18]. Analyzing RS control by LD tools and the formulation of the corresponding maximum principle are due to Whittle [24]. Various aspects of this approach have been studied for controlled stochastic differential equations, for example, in [10, 13, 15]. The treatment of a QCP at the MD scale is similar to that at the LD scale (in papers such as [3]) as far as the tools are concerned, but there are reasons to believe that the games obtained in the MD regime are solvable more often than in the LD regime. This statement is supported by the fact that the paper [1] solves a DG for the MD scale, whereas a solution of an analogous DG for the LD regime is not known in general (see [5] for a partial solution of the latter, and an open problem regarding its general solution). Similarly, the DG of this paper has been solved in [2], but an explicit solution for the LD analogue is not known. An additional advantage the MD regime has over LD is the invariance to the stochastic data, specifically, the arrival and service time distributions, as long as they possess exponential moments, where LD results are more sensitive. The combination of these properties provides a great deal of motivation for working at the MD scale.

The aforementioned pathwise minimality property has been the basis for solving QCPs in diffusion scale heavy traffic asymptotics in various works in the past (e.g., [17]). To describe this property, consider the simple diffusion control problem of minimizing a cost $J(\zeta)$ over all control processes ζ having $\mathbb{R}_+$-valued nondecreasing sample paths. The cost takes the form

$$J(\zeta) = \mathbb{E}\int_0^T h(\xi_t)\,dt,$$

where $h : \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing, $\xi_t = x + w_t + \zeta_t$, $x \geq 0$ and $T > 0$ are fixed, w is a standard Brownian motion and the constraint $\xi_t \geq 0$ for all $t \geq 0$ must be met. The solution is to set $\zeta_t = -\inf_{s\leq t}[(x + w_s) \wedge 0]$, making ξ a reflected Brownian motion starting from x, reflecting at 0. This follows by the well-known fact that a.s., for all t, $x + w_t + \zeta_t \leq x + w_t + \tilde{\zeta}_t$, for any control $\tilde{\zeta}$ meeting the constraint (see, e.g., Section 2 of [8]). Although this problem is simpler than typical diffusion control problems in the literature, pathwise solutions of these problems owe to this simple property (or sometimes multidimensional versions thereof).

The DG of [1], identifying the MD asymptotics of a QCP, was also solved by such a consideration. A simplified version of this game, presented with one-dimensional instead of multidimensional dynamics, is as follows. It is a zero-sum game with payoff

$$\tilde{J}(\tilde{\lambda}, \tilde{\mu}, \zeta) = \int_0^T h(\varphi_t)\,dt - \int_0^T \bigl(a\tilde{\lambda}_t^2 + b\tilde{\mu}_t^2\bigr)\,dt, \tag{1}$$

and dynamics

$$\varphi_t = x + \int_0^t (\tilde{\lambda}_s - \tilde{\mu}_s)\,ds + \zeta_t. \tag{2}$$

Here, x, T, a and b are positive constants, and h is again nondecreasing. The function ζ is a control for the minimizing player, taking values in $\mathbb{R}_+$ and is nondecreasing, while the functions $\tilde{\lambda}$ and $\tilde{\mu}$ form a control for the maximizing player, and are nonnegative. The constraint $\varphi_t \geq 0$ for all t must be satisfied. In this game, the functions φ and ζ represent MD-scaled queue length and idleness processes, while $\tilde{\lambda}$ and $\tilde{\mu}$ stand for MD-scaled perturbations of the arrival and service processes. The function h is the running cost in the underlying RS cost, while the second term in (1) corresponds to penalty associated with changes of measure, and its form originates from the LD rate function (background on the structure of DGs governing RS control asymptotics appears in [12] and [14]).

It is easy to see how pathwise minimality can be used once again to find an optimal strategy for the minimizer. Namely, for $(\tilde{\lambda}, \tilde{\mu})$ given, setting $\zeta_t = -\inf_{s \leq t}(\psi_s \wedge 0)$, where $\psi_t = x + \int_0^t (\tilde{\lambda}_s - \tilde{\mu}_s)\, ds$, results with φ that bounds from below any other dynamics adhering to the constraint. Significantly, this pathwise minimality property provides not only a solution to the game but also the basis of the AO proof in [1], as one can mimic the behavior of this strategy to come up with a policy for the queueing model that is automatically AO.

It turns out that one cannot argue along the same lines for the game obtained under LD scaling. Indeed, note carefully that the solution method just presented uses the fact that the second term in (1) does not depend on the control for the minimizing player. However, under LD scaling, the corresponding penalty term, accounting for changes of measure, involves controls of both players. This makes it impossible to obtain a pathwise solution in the same fashion. This point is explained in detail in Section 1 of [1].

Although pathwise minimality is useful when it applies, it is not generic even under the diffusion and MD regimes. A natural approach to handle more general settings, that has been used in numerous papers on diffusion scale asymptotics, is to appeal to dynamic programming methods to solve diffusion control problems and then use these solutions as a vehicle for analyzing the QCP (for a small sample of these papers, see [4, 7, 23]). This approach has not been considered before for MD asymptotics of QCPs. The model studied in this paper is indeed suitable for such an approach, and in fact constitutes a prototype for QCPs that are too complex to possess directly solvable DGs, while a solution via dynamic programming is available.

The DG for our model differs from the one presented above. Again, we present it in a slightly simplified way; the precise details appear in Section 2. The payoff (1) has an additional term ϱ_T, the dynamics (2) has an additional term $-\varrho_t$ and the constraint $\varphi_t \geq 0$ is strengthened to $\varphi_t \in [0, D]$ for all t, where D is a constant. The $\mathbb{R}_+$-valued nondecreasing function ϱ represents cumulative rejections, and is considered part of the control for the minimizing player; that is, in this game the minimizing player controls the pair (ζ, ϱ). The constraint stems from the finiteness of the buffer, and the constant D is related to the buffer size (in fact, it is the buffer size measured in units of MD-scaled workload). In [2], this game was analyzed

via a free boundary value problem, and solved for the value function and optimal strategy. The contribution of this paper is to substantiate the relation of the queueing model to the game in a rigorous manner, showing that the latter indeed governs the MD asymptotics. This is established by proving that the value of the RS QCP converges to that of the DG, and translating the DG's optimal strategy into an AO policy for the QCP.

The proof of convergence of the RS value to the DG value is performed in two main steps, namely bounding the former from below and from above by the latter. We refer to them as the lower and upper bound, respectively.

The proof of an asymptotic lower bound of a RS cost in models that do not have control is often based on the classical proof of (the lower bound in) Varadhan's lemma [9] for a sequence of processes satisfying the large deviations principle. This is the case, for example, for queueing models that are studied under a specified policy. It relies on the identification of a "behavior" that contributes most to the cost, such as when the underlying stochastic dynamics (say, the suitably normalized multidimensional queue length) lie close to a specific path. In the case of controlled dynamics, this path is formulated as the control selected by the maximizing player in the DG. To obtain the lower bound, one must consider an arbitrary sequence of policies, and then the challenge stems from the fact that different policies for the queueing model may give rise to different such paths. A brute force approach of identifying an optimal path for each arbitrary policy seems intractable.

The argument uses instead properties of the DG studied in [2]. It has been shown that this DG, specified in terms of multi-dimensional dynamics, can be reduced to one dimension. The one-dimensional state corresponds to the (suitably normalized) total workload in the system. Moreover, there is a threshold, denoted by β_0, dictating the behavior of both players. When the workload is below this threshold, there is a certain fixed path that guarantees attaining at least the game's value under any action of the minimizer (although it need not be optimal). When the threshold is exceeded, there is no such fixed path. However, the following fact can be used. As long as the workload remains above the level β_0, the minimizer encounters an accumulated loss, which is higher than the cost of an immediate rejection to the level β_0. We identify a suitable path for the maximizer that is effective until the time when the threshold is reached. We focus on this path when workload is above β_0, and switch to the path alluded to above, when it is below β_0. However, this switching time depends on the policy, and so it is random and varies with the scaling parameter. Accordingly, the argument uses time discretization, where each one of a finite collection of possible switching times is estimated separately.

The upper bound is obtained by constructing a policy for the QCP for which the cost converges to the DG's value. There is a naive way of interpreting the DG solution as a control policy for the QCP. However, the two components of this policy, corresponding to rejection and service effort allocation, impose contradictory requirements. For rejections must occur only from a specific class, and only when the total workload in the system exceeds the aforementioned threshold. On the

other hand, it stems from the game solution that the service allocation policy must cause the (suitably normalized) multidimensional queue length processes to evolve along a certain curve in state space, denoted in this paper by γ. One can express the fact that buffers are finite by requiring that these queueing processes always lie in a certain hyper-rectangular domain, denoted by $\mathcal{X}$; the curve γ happens to intersect the boundary of the domain $\mathcal{X}$. When the multidimensional queueing process is on (certain parts of) that boundary, one of the buffers must be full, and then even small stochastic fluctuations require rejections so as to meet the buffer size constraint. As a consequence, rejections will not always adhere to the rejection policy alluded to above.

We address this issue by aiming at a curve γ^a, that approximates γ but does not intersect the boundary of the domain. The main body of work is to develop estimates that show that the queueing processes evolve along this approximate curve, up to negligible probabilities, and that as a result both elements of the policy are respected with sufficiently high probability.

It is well known that an optimal control over a finite time interval is generically of a *nonstationary* feedback type (see, e.g., Sections III.8–9 in [14], where the nonstationary optimal feedback is characterized by (8.5) and the stationary optimal feedback by (9.9), corresponding to a problem set on a finite time horizon and, respectively, an infinite time horizon with a discounted cost). Despite that, it was anticipated in [2], and established in this paper, that for the setting studied here, a RS cost set over a finite time horizon $[0, T]$ gives rise to a *stationary* feedback provided that T is sufficiently large. Indeed, the policy we present for the QCP has this feature, which makes it simple as compared to policies based on time-varying characteristics. We consider this as one of the main aspects of this paper's contribution.

The organization of the paper is as follows. Section 2 presents the model, the MD scaling and the main result, which states that the MD QCP's value converges to that of the DG. Section 3 collects a few results from [2] required for the proof. Section 4 gives a lower bound on the QCP's value asymptotics in terms of the DG's value, and Section 5 finds a nearly optimal policy derived from the game's optimal strategy. Together, Sections 4 and 5 provide the proof of the main result. Some auxiliary results appear in the Appendix.

We use the following notation. For a positive integer k and $a, b \in \mathbb{R}^k$, $a \cdot b$ denotes the usual scalar product, while $\|\cdot\|$ denotes Euclidean norm. $\{e_1, \ldots, e_k\}$ is the standard basis of $\mathbb{R}^k$. We denote $[0, \infty)$ by $\mathbb{R}_+$. For $0 < T < \infty$ and a function $f : \mathbb{R}_+ \to \mathbb{R}^k$, $\|f\|_T = \sup_{[0,T]} \|f\|$, while $\mathrm{osc}_T(\delta, f) = \sup\{\|f(u) - f(t)\| : 0 \le u \le t \le (u + \delta) \wedge T\}$. Denote by $\mathcal{AC}([0, T], \mathbb{R}^k)$, $\mathcal{C}([0, T], \mathbb{R}^k)$ and $\mathcal{D}([0, T], \mathbb{R}^k)$ the spaces of absolutely continuous functions [resp., continuous functions, functions that are right-continuous with finite left limits (RCLL)] mapping $[0, T] \to \mathbb{R}^k$. Write $\mathcal{AC}_0([0, T], \mathbb{R}^k)$ and $C_0([0, T], \mathbb{R}^k)$ for the subsets of the corresponding function spaces, of functions that start at zero. Endow the space

$\mathcal{D}([0,T],\mathbb{R}^k)$ with the usual Skorohod topology. For a collection $x_i \in \mathbb{R}$ indexed by $i \in \{1,\dots,I\}$ (I being a positive integer), x denotes the vector (x_i). A similar convention holds for $\mathbb{R}$-valued random variables X_i and stochastic process $\{X_i(t), t \in \mathbb{R}_+\}$, $i \in \{1,\dots,I\}$, where X and $\{X(t), t \in \mathbb{R}_+\}$ denote the $\mathbb{R}^I$-valued random variable and process.

2. Model and results.

2.1. *Model description.* We consider a model with I customer classes and a single server. A buffer with finite room is dedicated to each customer class, and upon arrival, customers are queued in the corresponding buffers, or rejected by the system administrator. Within each class, customers are served at the order of arrival, where the server may only serve the customer at the head of each line. Processor sharing is allowed, and so the server is capable of serving up to I customers of distinct classes simultaneously. The model is defined on a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Expectation with respect to $\mathbb{P}$ is denoted by $\mathbb{E}$. The parameters and processes we introduce depend on an index $n \in \mathbb{N}$, serving as the scaling parameter. Arrivals occur according to independent renewal processes, and service times are independent and identically distributed across each class. Let $\mathcal{I} = \{1,2,\dots,I\}$. Let $\lambda_i^n > 0$, $n \in \mathbb{N}$, $i \in \mathcal{I}$ be given parameters, representing the *reciprocal mean inter-arrival times* of class-i customers. Let $\{IA_i(l) : l \in \mathbb{N}\}_{i\in\mathcal{I}}$ be I sequences of positive i.i.d. random variables with mean $\mathbb{E}[IA_i(1)] = 1$ and variance $\sigma_{i,IA}^2 = \mathrm{Var}(IA_i(1)) \in (0,\infty)$. With $\sum_1^0 = 0$, the number of arrivals of class-i customers up to time t, for the nth system, is given by

$$A_i^n(t) := \sup\left\{ l \ge 0 : \sum_{k=1}^{l} \frac{IA_i(k)}{\lambda_i^n} \le t \right\}, \qquad t \ge 0. \tag{3}$$

For a collection ξ_i, $i \in \mathcal{I}$ of stochastic processes, we will always write ξ for $(\xi_i)_{i\in\mathcal{I}}$. Thus, in particular, A^n is the I-dimensional process $(A_i^n)_{i\in\mathcal{I}}$.

Similarly, we consider another set of parameters $\mu_i^n > 0$, $n \in \mathbb{N}$, $i \in \mathcal{I}$, representing *reciprocal mean service times*. We are also given I independent sequences $\{ST_i(l) : l \in \mathbb{N}\}_{i\in\mathcal{I}}$ of positive i.i.d. random variables (independent also of the sequences $\{IA_i\}$) with mean $\mathbb{E}[ST_i(1)] = 1$ and variance $\sigma_{i,ST}^2 = \mathrm{Var}(ST_i(1)) \in (0,\infty)$. The time required to complete the service of the lth class-i customer is given by $ST_i(l)/\mu_i^n$, and the *potential service time* processes are defined as

$$S_i^n(t) := \sup\left\{ l \ge 0 : \sum_{k=1}^{l} \frac{ST_i(k)}{\mu_i^n} \le t \right\}, \qquad t \ge 0.$$

We consider the *moderate deviations rate parameters* $\{b_n\}$, that form a sequence, fixed throughout, with the property that $\lim b_n = \infty$ while $\lim b_n/\sqrt{n} = 0$, as

$n \to \infty$. The arrival and service parameters are assumed to satisfy the following conditions. As $n \to \infty$,

$$\frac{\lambda_i^n}{n} \to \lambda_i \in (0, \infty), \qquad \frac{\mu_i^n}{n} \to \mu_i \in (0, \infty), \tag{4}$$

$$\begin{aligned} \tilde{\lambda}_i^n &:= \frac{1}{b_n\sqrt{n}}\bigl(\lambda_i^n - n\lambda_i\bigr) \to \tilde{\lambda}_i \in (-\infty, \infty), \\ \tilde{\mu}_i^n &:= \frac{1}{b_n\sqrt{n}}\bigl(\mu_i^n - n\mu_i\bigr) \to \tilde{\mu}_i \in (-\infty, \infty). \end{aligned} \tag{5}$$

The system is assumed to be critically loaded in the sense that the overall traffic intensity equals 1, namely $\sum_1^I \rho_i = 1$ where $\rho_i = \lambda_i/\mu_i$ for $i \in \mathcal{I}$.

For $i \in \mathcal{I}$, let X_i^n be a process representing the number of class-i customers in the nth system. Denote the number of rejection of class-i arrivals until time t in the nth system by $R_i^n(t)$. With $\mathbb{S} = \{x = (x_1, \ldots, x_I) \in \mathbb{R}_+^I : \sum x_i \le 1\}$, let B^n be an $\mathbb{S}$-values process, whose ith component represents the fraction of effort devoted by the server to the class-i customer at the head of the line. Then the number of service completions of class-i jobs during the time interval $[0, t]$ is given by $S_i^n(T_i^n(t))$, where

$$T_i^n(t) := \int_0^t B_i^n(u)\,du \tag{6}$$

is the time devoted to class-i customers by time t. With an abuse of notation, we often write $S^n \circ T^n$ for $(S_1^n \circ T_1^n, \ldots, S_I^n \circ T_I^n)$. We have the balance equation

$$X_i^n(t) = X_i^n(0) + A_i^n(t) - S_i^n\bigl(T_i^n(t)\bigr) - R_i^n(t). \tag{7}$$

For simplicity, the initial conditions $X^n(0) = (X_1^n(0), \ldots, X_I^n(0))$ are assumed to be deterministic. We also assume

$$\tilde{X}^n(0) \to \bar{x} = (x_1, \ldots, x_I) \qquad \text{as } n \to \infty, i \in \mathcal{I}.$$

Note that, by construction, the arrival and potential service processes have RCLL paths, and accordingly, so does X^n.

The MD-scaled version of the queue length process satisfies

$$\tilde{X}^n(t) := \frac{1}{b_n\sqrt{n}} X^n(t) \in \mathcal{X} := \prod_{i=1}^I [0, D_i], \qquad t \ge 0, \tag{8}$$

where $D_i > 0$ are fixed constants. Thus, the size of buffer i is given by $b_n\sqrt{n}D_i$. Additional MD-scaled processes are

$$\begin{aligned} \tilde{A}_i^n(t) &= \frac{1}{b_n\sqrt{n}}\bigl(A_i^n(t) - \lambda_i^n t\bigr), \\ \tilde{S}_i^n(t) &= \frac{1}{b_n\sqrt{n}}\bigl(S_i^n(t) - \mu_i^n t\bigr), \qquad \tilde{R}_i^n(t) = \frac{1}{b_n\sqrt{n}} R_i^n(t). \end{aligned}$$

The process $U^n := (B^n, R^n)$ is regarded as a control, that is determined based on observations from the past events in the system. The precise definition of an admissible control is as follows. Given n, the processes U^n and X^n are said to be an *admissible control* and the *corresponding queue length process* if the sample paths of U^n lie in $\mathcal{D}([0,\infty),\mathbb{S})$, (8) holds, and:

- U^n is adapted to the filtration $\sigma\{A_i^n(u), S_i^n(T_i^n(u)), i \in \mathcal{I}, u \le t\}$, where T^n is given by (6);
- For $i \in \mathcal{I}$ and $t \ge 0$, one has

$$X_i^n(t) = 0 \quad \text{implies} \quad B_i^n(t) = 0. \tag{9}$$

Denote the class of all admissible controls U^n by $\mathfrak{U}^n$. Note that this class depends on A^n and S^n, but we consider these processes to be fixed. For $U^n \in \mathfrak{U}^n$ and the corresponding queue length process X^n, the processes $\tilde{R}^n$ and $\tilde{X}^n$ are referred as the scaled rejection and queue length process corresponding to U^n.

Throughout, we assume the finite exponential moment condition, namely

ASSUMPTION 2.1. There exists $u_0 > 0$ such that for $i \in \mathcal{I}$, $\mathbb{E}[e^{u_0 IA_i}]$ and $\mathbb{E}[e^{u_0 ST_i}]$ are finite.

As shown in [22], under this condition, the scaled processes $(\tilde{A}^n, \tilde{S}^n)$ satisfy a *moderate deviation principle*. Namely, for $k = 1, 2$, let $\mathbb{J}_k(T,\cdot)$ be functions mapping $\mathcal{D}([0,T],\mathbb{R}^I)$ to $[0,\infty]$ given by

$$\mathbb{J}_k(T,\psi) = \begin{cases} \displaystyle\sum_{i=1}^I s_{i,k} \int_0^T \dot{\psi}_i(u)^2\,du & \text{if all } \psi_i \in \mathcal{AC}_0([0,T],\mathbb{R}), \\ \infty & \text{otherwise,} \end{cases} \qquad k = 1, 2, \tag{10}$$

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

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

Let $\mathbb{J}(T,\psi) = \mathbb{J}_1(T,\psi^1) + \mathbb{J}_2(T,\psi^2)$ for $\psi = (\psi^1,\psi^2) \in \mathcal{D}([0,T],\mathbb{R}^{2I})$. Note that $\mathbb{J}$ is lower semicontinuous with compact level sets. Then one has

PROPOSITION 2.1 ([22]). *Under Assumption 2.1, for $T > 0$ fixed, the following holds. For every closed set $F \subset \mathcal{D}([0,T],\mathbb{R}^{2I})$,*

$$\limsup \frac{1}{b_n^2} \log \mathbb{P}\big((\tilde{A}^n, \tilde{S}^n) \in F\big) \le -\inf_{\psi \in F} \mathbb{J}(T,\psi),$$

and for every open set $G \subset \mathcal{D}([0,T],\mathbb{R}^{2I})$,

$$\liminf \frac{1}{b_n^2} \log \mathbb{P}\big((\tilde{A}^n, \tilde{S}^n) \in G\big) \ge -\inf_{\psi \in G} \mathbb{J}(T,\psi).$$

To present the RS control problem, fix $\bar{h}, \bar{r} \in (0,\infty)^I$. Given $T \in (0,\infty)$ and n, the cost associated with a control $U^n \in \mathfrak{U}^n$ is given by

$$J^n\big(T, \tilde{X}^n(0), U^n\big) = \frac{1}{b_n^2} \log \mathbb{E}\bigg[\int_0^T e^{b_n^2[\int_0^t \bar{h}\cdot\tilde{X}^n(u)\,du + \bar{r}\cdot\tilde{R}^n(t)]}\,dt\bigg],$$

where $\tilde{X}^n$ and $\tilde{R}^n$ are the rescaled queue length and rejection processes corresponding to U^n. In the above notation, we have emphasized the dependence on the initial state $\tilde{X}^n(0)$. For background and a motivating discussion about this type of cost, the reader is referred to [12]. The value of interest is given by

$$V^n\big(T, \tilde{X}^n(0)\big) = \inf_{U^n \in \mathfrak{U}^n} J^n\big(T, \tilde{X}^n(0), U^n\big).$$

2.2. *The differential game and main result.* Whereas the scaled queue length process $\tilde{X}^n$ is multidimensional, it is suggested in [2] that it is governed by a DG defined in terms of one-dimensional dynamics. The main result of this paper is the proof of this claim. Before presenting the formulation of this game, it is useful to draw attention to several stochastic processes that are themselves one-dimensional because the structure of the DG is closely related to them. Let

$$\theta^n = \bigg(\frac{n}{\mu_1^n}, \ldots, \frac{n}{\mu_I^n}\bigg), \qquad \theta = \bigg(\frac{1}{\mu_1}, \ldots, \frac{1}{\mu_I}\bigg),$$

and note that $\theta^n \to \theta$, by (4). Denote $\theta_{\min} = \max_i \theta_i$ and $\theta_{\max} = \max_i \theta_i$. Let also

$$D^n = \sum_i \theta_i^n D_i, \qquad D = \sum_i \theta_i D_i.$$

It follows from the balance equation (7) that

$$\tilde{X}_i^n = \tilde{Y}_i^n + \tilde{A}_i^n - \tilde{S}_i^n \circ T_i^n + \tilde{Z}_i^n - \tilde{R}_i^n, \tag{11}$$

where we denote

$$\begin{aligned} \tilde{Z}_i^n(t) &= \frac{\mu_i^n}{n}\frac{\sqrt{n}}{b_n}\big(\rho_i t - T_i^n(t)\big), \\ \tilde{y}_i^n &= \tilde{\lambda}_i^n - \rho_i \tilde{\mu}_i^n, \qquad \tilde{Y}_i^n(t) = \tilde{X}_i^n(0) + \tilde{y}_i^n t. \end{aligned} \tag{12}$$

Define

$$\begin{aligned} &\check{X}^n = \theta^n \cdot \tilde{X}^n, \qquad \check{A}^n = \theta^n \cdot \tilde{A}^n, \qquad \check{S}^n = \theta^n \cdot \tilde{S}^n, \\ &\check{R}^n = \theta^n \cdot \tilde{R}^n, \qquad \check{Y}^n = \theta^n \cdot \tilde{Y}^n, \qquad \check{Z}^n = \theta^n \cdot \tilde{Z}^n, \qquad \check{y}^n = \theta^n \cdot \tilde{y}^n. \end{aligned} \tag{13}$$

Also, let

$$\begin{aligned} \tilde{\mathbf{S}}^n(t_1, \ldots, t_I) &= \big(\tilde{S}_1^n(t_1), \ldots, \tilde{S}_I^n(t_I)\big), \\ \check{\mathbf{S}}^n(t_1, \ldots, t_I) &= \theta^n \cdot \tilde{\mathbf{S}}^n(t_1, \ldots, t_I) = \sum_i \theta_i^n \tilde{S}_i^n(t_i). \end{aligned}$$

Note that the sample paths of $\tilde{S}^n$ (resp., $\check{S}^n$, $\tilde{\boldsymbol{S}}^n$, $\check{\boldsymbol{S}}^n$) map $\mathbb{R}_+ \to \mathbb{R}^I$ (resp., $\mathbb{R}_+ \to \mathbb{R}_+$, $\mathbb{R}_+^I \to \mathbb{R}^I$, $\mathbb{R}_+^I \to \mathbb{R}$). Next, the process

$$\check{Z}^n = \sum_i \frac{n}{\mu_i^n} \tilde{Z}_i^n \text{ is nonnegative and nondecreasing,} \tag{14}$$

thanks to the fact that $\sum_i B_i^n \le 1$ while $\sum_i \rho_i = 1$. Also, for every $i \in \mathcal{I}$, the process $\tilde{R}_i^n$ is nondecreasing. Thus, by (11),

$$\check{X}^n = \check{Y}^n + \check{A}^n - \check{\boldsymbol{S}}^n \circ T^n + \check{Z}^n - \check{R}^n, \tag{15}$$

where $\check{Z}^n$ and $\check{R}^n$ are nonnegative, nondecreasing processes. Moreover,

$$\check{X}^n(t) \in \bigl[0, D^n\bigr], \qquad t \ge 0. \tag{16}$$

It follows from the contraction principle that $(\check{A}^n(t), \check{S}^n(t))$, $t \in [0, T]$, satisfy the MDP with the rate function $\mathbb{I}(T, \psi) = \mathbb{I}_1(T, \psi^1) + \mathbb{I}_2(T, \psi^2)$, $\psi = (\psi^1, \psi^2) \in \mathcal{D}([0, T], \mathbb{R}^2)$, where

$$\mathbb{I}_k\bigl(T, \psi^k\bigr) = \begin{cases} s_k \displaystyle\int_0^T \bigl(\dot{\psi}^k\bigr)^2(u)\, du & \text{if } \psi^k \in \mathcal{AC}_0\bigl([0, T], \mathbb{R}\bigr), \\ \infty & \text{otherwise,} \end{cases} \tag{17}$$

$$s_1 := \left(\sum_{i=1}^I \frac{2\rho_i \sigma_{i,IA}^2}{\mu_i}\right)^{-1} \quad \text{and} \quad s_2 := \left(\sum_{i=1}^I \frac{2\rho_i \sigma_{i,ST}^2}{\mu_i}\right)^{-1}. \tag{18}$$

See Lemma A.2.

To define the DG, denote $x = \theta \cdot \bar{x}$, $y := \lim \check{y}^n = \sum_i \theta_i(\tilde{\lambda}_i - \rho_i \tilde{\mu}_i)$, and

$$\boldsymbol{y}(t) = x + yt, \qquad t \in \mathbb{R}_+,$$

and let

$$\mathcal{P} = \mathcal{C}_0\bigl([0, \infty), \mathbb{R}\bigr), \qquad \mathcal{E} = \bigl\{\xi \in \mathcal{D}\bigl([0, \infty), \mathbb{R}_+\bigr) : \xi \text{ is nondecreasing}\bigr\}. \tag{19}$$

Endow both spaces with the topology of uniform convergence on compacts. Given $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$ and $(\zeta, \varrho) \in \mathcal{E}^2$, the *dynamics associated with the initial condition x and the data ψ, ζ, ϱ* is defined as

$$\varphi = \boldsymbol{y} + \psi^1 - \psi^2 + \zeta - \varrho. \tag{20}$$

The game is played by a maximizing player that selects $\psi = (\psi^1, \psi^2)$ and a minimizing player that selects (ζ, ϱ). We sometimes write the dependence of the dynamics on the data as $\varphi[x, \psi, (\zeta, \varrho)]$. There is an analogy between the above equation and equation (15), and between the condition that ζ and ϱ are nondecreasing and property (14). The control ζ stands for the scaled idle time process $\tilde{Z}^n$ and ϱ stands for the scaled rejection process $\tilde{R}^n$. The following condition, analogous to property (16), will also be required, namely

$$\varphi(t) \in [0, D], \qquad t \ge 0, \tag{21}$$

where

$$D = \sum_{i=1}^{I} \theta_i D_i = \lim_{n\to\infty} D^n.$$

A measurable mapping $\alpha : \mathcal{P}^2 \to \mathcal{E}^2$ is called a *strategy for the minimizing player* if it satisfies the causality property: for every $\psi, \tilde{\psi} \in \mathcal{P}^2$ and $t \in [0, \infty)$,

$$\begin{aligned} &\psi(u) = \tilde{\psi}(u) \qquad \text{for every } u \in [0,t] \quad \text{implies} \\ &\alpha[\psi](u) = \alpha[\tilde{\psi}](u) \qquad \text{for every } u \in [0,t]. \end{aligned} \tag{22}$$

Given an initial condition x, a strategy α is said to be *admissible for the initial condition* x if, whenever $(\psi^1, \psi^2) \in \mathcal{P}^2$ and $(\zeta, \varrho) = \alpha[\psi]$, the corresponding dynamics (20) satisfies the buffer constraint (21). We denote by A_x the collection of admissible strategies for the initial condition x.

We now describe the components of the cost function. For $w \in \mathbb{R}_+$, denote

$$h(w) = \inf\{\bar{h} \cdot \xi : \xi \in \mathcal{X}, \theta \cdot \xi = w\}. \tag{23}$$

By the convexity of the set $\mathcal{X}$, h is convex. Moreover, $h(w) \geq 0$ for $w \geq 0$ and equality holds if and only if $w = 0$. Therefore, h is strictly increasing on $[0, D]$. Let

$$r = \min\{\bar{r} \cdot \xi : \xi \in \mathbb{R}_+^I, \theta \cdot \xi = 1\}.$$

It is easy to see that

$$r = \min\{r_i \mu_i : i \in \mathcal{I}\} = r_{i^*} \mu_{i^*}, \tag{24}$$

where i^* is an index (fixed throughout) that minimizes $r_i \mu_i$.

The index i^* indicates the class that has lowest rejection cost per unit of workload. It plays an important role in Section 5 where our AO policy is presented; specifically, the policy is aimed at rejecting jobs from this class only.

Given $x \in [0, D]$, $T \in \mathbb{R}_+$, $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$, and $(\zeta, \varrho) \in \mathcal{E}^2$, we define the cost until time T by

$$c(x, T, \psi, \zeta, \varrho) = \int_0^T h\big(\varphi(t)\big)\, dt + r\varrho(T) - \mathbb{I}(T, \psi), \tag{25}$$

where φ is the corresponding dynamics. The value of the game is defined by

$$V(x) = \inf_{\alpha \in A_x} \sup_{\psi \in \mathcal{P}^2, T \in \mathbb{R}_+} c\big(x, T, \psi, \alpha[\psi]\big). \tag{26}$$

We call ψ the *path control* and the T a *time control*, or sometimes the *termination time*. Note that both are controlled by the maximizing player.

We sometimes use the notation $V_{h,r}$ for V when we want to emphasize the dependence on the function h and the constant r. Namely, $V_{\hat{h},\hat{r}}(x)$ is defined as $V(x)$ with $(\hat{h}, \hat{r})$ in place of (h, r) in (25).

Recall that $\lim_{n\to\infty} \tilde{X}^n(0) = x$. Our first main result is the following.

THEOREM 2.1. *Let Assumption 2.1 hold. Then for all sufficiently large T,*

$$\lim_{n\to\infty} V^n\big(T, \tilde{X}^n(0)\big) = V(x).$$

The second main result of this paper is Theorem 5.1, that constructs a policy for the QCP, which is AO.

3. Some useful properties of the game. We briefly mention some results from [2] regarding the DG, to be used in the sequel. Set $s = (s_1^{-1} + s_2^{-1})^{-1}$. The following proposition follows by Lemma 3.1, Proposition 3.1 and Theorems 3.1, 3.2 and 3.5 in [2]. The contribution of the latter is to deduce part (i) below.

PROPOSITION 3.1. (i) *For $T \in \mathbb{R}_+$, set*

$$V(T,x) = \inf_{\alpha\in A_x} \sup_{\psi\in\mathcal{P}^2, t\in[0,T]} c\big(x,t,\psi,\alpha[\psi]\big)$$

[*compare with* (26)]. *Then $V(T,x) = V(x)$ for all sufficiently large T.*

(ii) *If $-y < r/(4s)$, then for every $x \in [0, D]$ one has $V(x) = \infty$.*

(iii) *If $-y \geq r/(4s)$, then*

$$V(x) = \begin{cases} \displaystyle\int_0^x 2s\Big(-y - \sqrt{y^2 - \frac{h(u)}{s}}\Big)\,du, & 0 \leq x \leq \beta_0, \\ V(\beta_0) + r(x-\beta_0), & \beta_0 < x \leq D, \end{cases} \tag{27}$$

where[2]

$$\beta_0 = \begin{cases} h^{-1}\Big(\dfrac{-r^2}{4s} - ry\Big), & -h(D) \leq \dfrac{r^2}{4s} + ry \leq -h(0), \\ D, & \dfrac{r^2}{4s} + ry < -h(D). \end{cases} \tag{28}$$

We now present an optimal strategy for the minimizer. This strategy plays an important role in proving the upper bound and in finding an AO policy in the multidimensional stochastic problem. The minimizer's optimal strategy is of a β-barrier form. Informally, this is a strategy that uses the minimal control (ζ, ϱ) so as to keep the dynamics φ in $[0, \beta]$ at all times. In the definition that follows, and throughout the paper, we denote the Skorohod map on an interval $[a, b]$ by $\Gamma_{[a,b]}$; see Appendix A.1.

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

[2] In case $\frac{r^2}{4s} + ry \geq -h(0) = 0$, we get by (i) above that $V(x) = \infty$ and we do not define β_0.

The next proposition follows by Proposition 3.1 and Theorems 3.1, 3.2 and 3.4 in [2].

PROPOSITION 3.2. *The β_0-barrier strategy, α_{β_0}, is an optimal strategy.*

We provide two propositions that are useful in the proof of the lower bound. For this, we present two path controls, ψ^* and $\psi_x^\sharp$ associated with an initial state in the intervals $(\beta_0, D]$ and $[0, \beta_0)$, respectively.

Fix $x \in (\beta_0, D]$. Let $\Delta > 0$ be such that $x > \beta_0 + \Delta$. Fix $(\zeta, \varrho) \in \mathcal{P}^2$. Define

$$\psi^\sharp(t) = \big(rt/(2s_1), -rt/(2s_2)\big), \qquad 0 \le t \le \tau_\Delta,$$

where

$$\tau_\Delta = \tau_{(\zeta,\varrho),\Delta} := \inf\big\{t \ge 0 : \varphi\big[x, \psi^\sharp, (\zeta, \varrho)\big] \le \beta_0 + \Delta\big\}$$

is the first time that the dynamics, $\varphi := \varphi[x, \psi^\sharp, (\zeta, \varrho)]$, cross $\beta_0 + \Delta$. The following proposition is Proposition 3.2 in [2].

PROPOSITION 3.3. *For every $(\zeta, \varrho) \in \mathcal{P}^2$ such that $\varrho(0) - \zeta(0) < x - (\beta_0 + \Delta)$, one has*

$$\int_0^{\tau_\Delta} h\big(\varphi(t)\big)\,dt + r\varrho(\tau_\Delta) - \mathbb{I}\big(\tau_\Delta, \psi^\sharp\big) > r\big(x - (\beta_0 + \Delta)\big). \tag{29}$$

Note that the left-hand side is the cost associated with x, $\psi^\sharp$, and (ζ, ϱ) incurred until the time the dynamics cross $\beta_0 + \Delta$, whereas the right-hand side gives the cost of an immediate rejection of $x - (\beta_0 + \Delta)$. The result thus implies that if $x > \beta_0$ then the minimizer will reject $x - \beta_0$ units of mass at time zero.

Next, fix $x \in [0, \beta_0)$. Let

$$\psi_x^* = \left(\frac{s}{s_1}\omega_x^*, \frac{-s}{s_2}\omega_x^*\right), \tag{30}$$

where $\omega_x^* \in \mathcal{C}([0, \tau_x^*), \mathbb{R})$ is the unique solution of

$$\dot\omega_x^*(t) = \frac{\dot V(x + yt + \omega_x^*(t))}{2s}, \qquad t \ge 0, \tag{31}$$

with $\omega_x^*(0) = 0$ and

$$\tau_x^* = \int_0^x 1/\sqrt{y^2 - h(\xi)/s}\,d\xi. \tag{32}$$

Existence and uniqueness for ω_x^* over the time interval $[0, \tau_x^*]$ are shown in Section 3.6.2 in [2].

The next proposition, which follows by Proposition 3.3, Theorem 3.5, and equation (87) in [2], states that by using the path control ψ_x^* and by choosing the time control to be the first time that the actual dynamics of the game hit zero the maximizer can guarantee that the cost will be at least $V(x)$. The proposition is valid since the function h is convex.

PROPOSITION 3.4. *Fix $x \in [0, \beta_0)$. For every $\varrho \in \mathcal{P}$, one has*

$$c\big(x, \tau_x, \psi_x^*, (0, \varrho)\big) \geq V(x), \tag{33}$$

where $\tau_x := \tau[x, \psi_x^*, (0, \varrho)]$ is the first time that the dynamics $\varphi[x, \psi_x^*, (0, \varrho)]$ hits zero. Moreover,

$$\tau_x \leq \tau_x^* = \inf\big\{t \geq 0 : \boldsymbol{y}(t) + \psi_x^{*,1}(t) - \psi_x^{*,2}(t) = 0\big\}. \tag{34}$$

4. Lower bound. Recall that from Proposition 3.1(i), for sufficiently large T, $V(T, x) = V(x)$. Hence, now onwards we compare the value function of the QCP to $V(x)$.

THEOREM 4.1. *Let Assumption 2.1 hold. Then for T sufficiently large,*

$$\liminf_{n\to\infty} V^n\big(T, \tilde{X}^n(0)\big) \geq V(x).$$

We present two lemmas that together yield Theorem 4.1. The first provides a lower bound on the RS cost for an arbitrary sequence of policies in terms of an expression involving only the one-dimensional processes. The latter is further bounded by the DG value function, in the second lemma.

LEMMA 4.1. *Fix a sequence of admissible controls $\{U^n \in \mathfrak{U}^n\}$, $n \in \mathbb{N}$. Then for every $T > 0$, $\delta \in (0, T)$ and $\varepsilon > 0$ one has*

$$\begin{aligned}&\liminf_{n\to\infty} J^n\big(T, \tilde{X}^n(0), U^n\big)\\ &\quad\geq \liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{T-\delta} h(\check{X}^n(u))\,du + r(1-\varepsilon)\check{R}^n(T-\delta))}\big] - \varepsilon.\end{aligned}$$

LEMMA 4.2. *Fix $\{U^n\}$ as in Lemma 4.1. Then there exists $\bar{T} > 0$ such that for every $0 < \varepsilon < 1/2$ one has*

$$\liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar{T}} h(\check{X}^n(u))\,du + r(1-\varepsilon)\check{R}^n(\bar{T}))}\big] \geq V_{h,r(1-\varepsilon)}(x). \tag{35}$$

PROOF OF THEOREM 4.1. Combining Lemma 4.1 and Lemma 4.2, for any T and $\delta > 0$ such that $T - \delta > \bar{T}$, and any $\varepsilon \in (0, 1/2)$,

$$\liminf_{n\to\infty} J^n\big(T, \tilde{X}^n(0), U^n\big) \geq \liminf_{n\to\infty} J^n\big(\bar{T}, \tilde{X}^n(0), U^n\big) \geq V_{h,r(1-\varepsilon)}(x) - \varepsilon.$$

From (27)–(28), it follows that

$$\lim_{\varepsilon\to 0} V_{h,r(1-\varepsilon)}(x) = V_{h,r}(x) = V(x),$$

and the result follows. □

PROOF OF LEMMA 4.1. Fix $\{U^n\}$, $T > 0$, $\delta > 0$ and $\varepsilon > 0$. Then

$$\int_0^T e^{b_n^2(\int_0^t \bar{h}\cdot\tilde{X}^n(u)\,du+\bar{r}\cdot\tilde{R}^n(t))}\,dt \geq e^{b_n^2(\int_0^{T-\delta} \bar{h}\cdot\tilde{X}^n(u)\,du+\bar{r}\cdot\tilde{R}^n(T-\delta))}\delta, \tag{36}$$

where we used monotonicity of the integrand with respect to t. Next, by the definition of h and r,

$$\bar{h}\cdot\tilde{X}^n \geq h\big(\theta\cdot\tilde{X}^n\big) \quad \text{and} \quad \bar{r}\cdot\tilde{R}^n \geq r\theta\cdot\tilde{R}^n. \tag{37}$$

Since $\theta^n \to \theta$, $\tilde{X}^n$ takes values in a fixed, compact set, and h is uniformly continuous on this set, it follows that for sufficiently large n,

$$\begin{aligned} \int_0^T h\big(\theta\cdot\tilde{X}^n(u)\big)\,du &\geq \int_0^T h\big(\check{X}^n(u)\big)\,du - \varepsilon \quad \text{and} \\ \theta_i &\geq \theta_i^n(1-\varepsilon), \qquad i \in \mathcal{I}. \end{aligned} \tag{38}$$

Combining (36), (37) and (38) yields the result. □

In the rest of this section, we prove Lemma 4.2.

PROOF OF LEMMA 4.2. Rather than working with general $\varepsilon \in (0, 1/2)$, we consider a general $r > 0$. For every $r > 0$ we find $\bar{T} = \bar{T}(r)$ [in (40)] that satisfies (35) with $\varepsilon = 0$. As we will see, $\bar{T}$ is continuous w.r.t. r and finite for $r > 0$. Hence, we obtain that (35) is valid with a fixed $\bar{T}$ and all $\varepsilon \in (0, 1/2)$. We thus turn to proving

$$\liminf \frac{1}{b_n^2}\log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar{T}} h(\check{X}^n(u))\,du + r\check{R}^n(\bar{T}))}\big] \geq V(x). \tag{39}$$

Sketch of the proof: The lemma relates the stochastic control problem to the DG. The strategies in the game are analogues of the policies in the stochastic problem, whereas the controls selected by the maximizing player play a similar role to the variational problem in Varadhan's lemma (Theorem 4.3.1 of [9]). Following the spirit of the proof of Varadhan's lemma, one focuses on the event that the paths $(\tilde{A}^n, \tilde{S}^n)$, projected in the θ^n direction, are in a neighborhood of a specific $\mathcal{P}^2$-path control. The latter is referred to as the reference path. One then shows that the process $\check{X}^n$ is in a neighborhood of the game dynamics obtained when the reference path is selected by the maximizing player.

We now describe the reference path. First, recall the balance equation (15). Consider paths $(\tilde{A}^n, \tilde{S}^n)$ such that $(\check{A}^n, \check{\boldsymbol{S}}^n \circ T^n)$ are close to some path $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$. Then $\check{X}^n$ is close to

$$\boldsymbol{y} + \psi^1 - \psi^2 + \check{Z}^n - \check{R}^n.$$

By the nonnegativity of $\check{Z}^n$, we have that $\check{X}^n$ is bounded from below by

$$\varphi^n = \boldsymbol{y} + \psi^1 - \psi^2 - \check{R}^n,$$

up to a small error term. The proof proceeds by comparing the process φ^n to the game dynamics with initial state x, maximizer's path control ψ, and minimizer's control given by $(0, \check{R}^n)$. This process cannot be regarded game dynamics because of the stochasticity of the minimizer's control term, and the fact that it is not attained by a strategy in the sense of the game. Moreover, it is not guaranteed that φ^n takes values in $[0, D]$.

However, these obstacles can be treated. Let us first describe the case $x \le \beta_0$, where the treatment is least complicated. In this case, we consider the event, denoted by O^n, that the paths $(\tilde{A}^n, \tilde{S}^n)$ are close to the path ψ_x^* identified in Proposition 3.4 (in this rough sketch we do not quantify the term "close"). We then obtain

$$\begin{aligned}
&\liminf \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar{T}} h(\check{X}^n(u))\,du + r\check{R}^n(\bar{T}))}\big] \\
&\quad \ge \liminf \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar{T}} h(\check{X}^n(u))\,du + r\check{R}^n(\bar{T}))} 1_{O^n}\big] \\
&\quad \ge c\big(x, \tau_x, \psi_x^*, (0, \varrho)\big) \ge V(x),
\end{aligned}$$

where the second inequality follows using the closeness of the data to ψ_x^* and the structure of the function c, while the last inequality follows from Proposition 3.4. Thus, Proposition 3.4 allows us to focus on data that is close to one fixed path, ψ_x^*, for every $n \in \mathbb{N}$.

The situation is more subtle in the case where $x > \beta_0$, as we do not have an analogue of Proposition 3.4. That is, we are unable to identify a single path that guarantees $V(x)$ as a lower bound on the cost under all strategies. Proposition 3.3 proposes how the maximizer in the game should act until the threshold β_0 is reached, namely to use the control $\psi^\sharp$ (note that the time when the threshold is reached depends on the minimizier's strategy). Therefore, we focus on data $(\tilde{A}^n, \tilde{S}^n)$ close to $\psi^\sharp$ until the workload in the stochastic model hits β_0. The proposition then guarantees that up to this time the cost incurred is bounded below by $r(x - \beta_0)$. Once β_0 is reached, ψ_{β_0} is used as explained above, and (27) is used to obtain $V(x)$ as a lower bound.

As we already mentioned, in the game, the time when one switches from $\psi^\sharp$ to $\psi_{\beta_0}^*$ depends on the strategy. As far as the QCP is concerned, this means that one has no control over the switching time, which may be random and vary with n. The argument therefore uses time discretization (see Lemma 4.4), with which each one of a finite collection of switching times is estimated separately.

Proof in details: We first prove the lower bound for the case $x > \beta_0$. The other case, which is simpler, is addressed at the end of the proof. Fix $x > \beta_0$ and $\varepsilon_1 > 0$. Let $\Delta > 0$ be small enough so that

$$x > \beta_0 + \Delta \quad \text{and} \quad r\Delta + V(\beta_0) - V(\beta_0 - \Delta) \le \varepsilon_1.$$

Such a Δ exists since V is continuous; see (27). Let

$$T_1 = \frac{V(x) + 3 + r(2 + D - x)}{h(\beta_0 + \Delta/2) - h(\beta_0)},$$

$$T_2 = \int_0^{\beta_0 - \Delta} 1/\sqrt{y^2 - h(\xi)/s}\, d\xi + 1,$$

$$\bar{T} = T_1 + T_2. \tag{40}$$

Above, T_1 is obtained by considerations along these lines. Start with (76), and consider only those policies for which

$$V(x) + 1 > b_n^{-2} \log \mathbb{E} \exp\left\{ b_n^2 \left(\int_0^{T_1} h\big(\check{X}^n(u)\big)\, du + r \check{R}^n(T_1) \right) \right\}.$$

Obtain a further lower bound on $V(x) + 1$ by putting on the right-hand side the indicator of a certain event that assures $\int_0^{T_1} h(\check{X}^n(u))\, du \geq h(\beta_0 + \Delta/2) T_1$. Moreover, we show that $\check{R}^n(T_1) \geq x - D - 2 + (y + \frac{r}{2s}) T_1$. When combined together with the definition of β_0 in (28), we get the above formula for T_1. The formula for T_2 is based on the termination time of the game (32).

The proof proceeds in several steps. In Step 1, we analyze the process T^n. In Step 2, we consider the process

$$\varphi_1^n := \boldsymbol{y} + \psi^{\sharp,1} - \psi^{\sharp,2} - \check{R}^n \tag{41}$$

and the time τ_1^n when this process first hits $\beta_0 + \Delta$. We show that for any policy that is nearly optimal $\tau_1^n < T_1$ with probability that is significant in the MD scale.

Next, in Step 3, we modify the construction of φ_1^n on the time interval $[\tau_1^n, \infty)$, and in Step 4 show that the process thus constructed hits zero before time $\bar{T}$, with probability 1. In Step 5, we combine these results to obtain a lower bound on the cost. Finally, in Step 6 we we take limits and obtain the result.

Step 1 (*Limit property of* T^n): Without loss of generality assume that the sequence of policies $\{U^n\}$ satisfies

$$V(x) + 1 > \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2 (\int_0^{\bar{T}} \bar{h} \cdot \tilde{X}^n(u)\, du + \bar{r} \cdot \tilde{R}^n(\bar{T}))} \big]. \tag{42}$$

Denote

$$\boldsymbol{\rho}(t) = \rho t, \qquad t \in \mathbb{R}_+.$$

LEMMA 4.3. *For every* $m > 0$ *there exists* $K > 0$ *such that for every* $i \in \mathcal{I}$ *one has*

$$\limsup \frac{1}{b_n^2} \log \mathbb{P}\left(\| T_i^n - \boldsymbol{\rho}_i \|_{\bar{T}} \geq \frac{b_n}{\sqrt{n}} K \right) \leq -m. \tag{43}$$

PROOF. Fix $i \in \mathcal{I}$. By (12), the left-hand side of (43) equals $\limsup \frac{1}{b_n^2} \times \log \mathbb{P}(\frac{n}{\mu_i^n} \|\tilde{Z}_i^n\|_{\bar{T}} \geq K)$. Using (11) and the fact that $\mu_n/n \to \mu$, it suffices to prove that for every $m > 0$ there exists K such that

$$\limsup \frac{1}{b_n^2} \log \mathbb{P}\big(\|L^n\|_{\bar{T}} \geq K\big) \leq -m, \tag{44}$$

for $L^n = \tilde{X}_i^n, \tilde{A}_i^n, \tilde{S}_i^n \circ T_i^n$ and $\tilde{R}_i^n$. As far as $\tilde{X}_i^n$ is concerned, the above is immediate because the process is bounded. For $L^n = \tilde{A}_i^n$, this property follows from Proposition 2.1 and from the fact that for sufficiently large K one has

$$\begin{aligned} &\inf\big\{\mathbb{J}_1(\bar{T}, \psi) : \psi \in \mathcal{P}^I \text{ and } \|\psi \cdot e_i\|_{\bar{T}} \geq K\big\} \\ &\quad = \mathbb{J}_1\big(\bar{T}, (Kt/\bar{T})e_i\big) = \frac{1}{2\mu\sigma_{IA}^2} \cdot \frac{K^2}{\bar{T}}. \end{aligned} \tag{45}$$

Since the time change $T_i^n(t) \leq t$ for all t, a similar conclusion holds for $\tilde{S}_i^n \circ T_i^n$. Finally, for $\tilde{R}_i^n$, note that by (42)

$$V(x) + 1 \geq \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2 r_i \tilde{R}_i^n(\bar{T})}\big].$$

Hence by the Chebyshev's inequality, $\mathbb{P}(\tilde{R}_i^n(\bar{T}) \geq K) \leq e^{-b_n^2(K+V(x)+1)}$. The result follows since, by the monotonicity, $\tilde{R}_i^n(\bar{T}) = \|\tilde{R}_i^n\|_{\bar{T}}$. $\square$

Step 2 (*Estimate on the time* τ_1^n): We introduce some notation that will be needed in the remainder of the proof. Since h is uniformly continuous on $[0, D]$, one can find $\delta_1 > 0$ such that

$$\operatorname{osc}_D(2\delta_1, h) < \varepsilon_1. \tag{46}$$

One may take δ_1 so that

$$\delta_1 < \min\{\Delta/4, 1\}. \tag{47}$$

Fix $m > \mathbb{I}(\bar{T}, \psi^*) + \mathbb{I}(\bar{T}, \psi^\sharp) + 1 + 6\varepsilon_1$. Define the event

$$E^n = E^n(K) = \left\{\|T_i^n - \boldsymbol{\rho}_i\|_{\bar{T}} < \frac{b_n}{\sqrt{n}} K, \text{ for all } i \in \mathcal{I}\right\}.$$

Using Lemma 4.3, fix $K > 0$ such that for all sufficiently large n

$$\frac{1}{b_n^2} \log \mathbb{P}\big((E^n)^c\big) \leq -m. \tag{48}$$

From Lemma A.2, it follows that there exists $\bar{\psi}^\sharp = (\bar{\psi}^{\sharp,1}, \bar{\psi}^{\sharp,2}) \in \mathcal{P}^{2I}$ such that

$$\big(\theta \cdot \bar{\psi}^{\sharp,1}, \theta \cdot \bar{\psi}^{\sharp,2} \circ \boldsymbol{\rho}\big) = \big(\psi^{\sharp,1}, \psi^{\sharp,2}\big) \tag{49}$$

and

$$\mathbb{J}_k(T, \bar{\psi}^{\sharp,k}) = \mathbb{I}_k(T, \psi^{\sharp,k}), \qquad k = 1, 2. \tag{50}$$

Note that for all large n,

$$\operatorname{osc}_T\left(\frac{b_n}{\sqrt{n}}K, \psi_i^{\sharp,2}\right) < \delta_2, \qquad i \in \mathcal{I}. \tag{51}$$

For $\psi = (\psi^1, \psi^2) \in \mathcal{P}^{2I}$, and $0 < \delta, t < \infty$, let

$$\mathcal{A}_{\delta,t}(\psi) = \{\bar{\psi} \in \mathcal{D}([0, \bar{T}], \mathbb{R}^{2I}) : \|\bar{\psi} - \psi\|_t < \delta\} \tag{52}$$

and

$$\Omega^n_{\delta,t}(\psi) = \{(\tilde{A}^n, \tilde{S}^n) \in \mathcal{A}_{\delta,t}(\psi)\}. \tag{53}$$

Recall from (41) the definition of φ_1^n, and let

$$\tau_1^n = \inf\{t \geq 0 : \varphi_1^n(t) \leq \beta_0 + \Delta\}.$$

Divide the time interval $[0, \bar{T}]$ into $(\bar{T}/\nu) \in \mathbb{N}$ intervals of size ν where[3]

$$\nu \leq \min\left\{\frac{4s\varepsilon_1}{r^2}, \frac{\varepsilon_1}{sy^2}, \frac{\Delta}{|r/(2s) + y|}\right\}. \tag{54}$$

Denote the intervals by $N_j = N_j(\nu) = [\nu j, \nu(j+1))$. For every n, we define an index $0 \leq j_1^n \leq \lfloor \bar{T}_1/\nu \rfloor$ in such a way that we can estimate, from below, the probability that the time τ_1^n belongs to the interval $N_{j_1^n}$. Let

$$j_1^n = \operatorname*{arg\,max}_{j \in \{0,\ldots,\lfloor T_1/\nu \rfloor\}} \mathbb{P}(\tau_1^n \in N_j | \Omega^n_{\delta_3,T_1}(\bar{\psi}^\sharp)),$$

where $\delta_2 = \delta_1/(8\theta_{\max}\sqrt{I})$.

LEMMA 4.4. *Fix $\psi \in \mathcal{D}([0, \bar{T}], \mathbb{R}^I)$ such that $\psi = \bar{\psi}^\sharp$ on $[0, (j_1^n + 1)\nu]$. Then*

$$\liminf_{n\to\infty} \mathbb{P}(\tau_1^n \in N_{j_1^n} | \Omega^n_{\delta_3,(j_1^n+1)\nu}(\psi)) \geq -2\varepsilon_1. \tag{55}$$

The proof of this lemma is differed to the end of the section.

Step 3 (*Constructing a path beyond time* τ_1^n): In Lemma 4.4, we focused on data for which the process $(\check{A}^n, \check{S}^n \circ T^n)$ is near $\psi^\sharp$. Now we will consider data for which this process is near $\psi^\sharp$ up to time $(j_1^n + 1)\nu$, and from that time on, near $\psi^*_{\beta_0-\Delta}$. Thus, we focus on the reference path:

$$\psi^0(t) = \psi^{n,0}(t) := \begin{cases} \psi^\sharp(t), & 0 \leq t \leq (j_1^n + 1)\nu, \\ \psi^\sharp((j_1^n + 1)\nu) + \psi^*_{\beta_0-\Delta}(t - (j_1^n + 1)\nu), & t > (j_1^n + 1)\nu. \end{cases}$$

[3]We use the convention that $\Delta/0 = \infty$.

Recall that τ_1^n is defined as the first time when $\varphi_1^n \le \beta_0 + \Delta$. Since $\boldsymbol{y} + \psi^{\sharp,1} - \psi^{\sharp,2}$ is continuous, and the jumps of $\check{R}^n$ are of size $(b_n\sqrt{n})^{-1}$, one has for sufficiently large n, $\beta_0 < \varphi_1^n(\tau_1^n) \le \beta_0 + \Delta$. Moreover, since on the interval $(\tau_1^n, (j_1^n+1)\nu)$ one has $y + \dot{\psi}^{\sharp,1} - \dot{\psi}^{\sharp,2} = y + r/(2s)$ and we assumed that $\nu \le \Delta/|r/(2s) + y|$, it follows that

$$\begin{aligned}\varphi_1^n\big((j_1^n+1)\nu\big) &= \varphi_1^n(\tau_1^n) + \big(y + r/(2s)\big)\big((j_1^n+1)\nu - \tau_1^n\big)\\ &\quad - \big(\check{R}^n\big((j_1^n+1)\nu\big) - \check{R}^n(\tau_1^n)\big)\\ &> \beta_0 - \Delta - \big(\check{R}^n\big((j_1^n+1)\nu\big) - \check{R}^n(\tau_1^n)\big).\end{aligned}$$

We now define a new process that starts at time $(j_1^n+1)\nu$, having the form of the game dynamics with the initial state $\beta_0 - \Delta$, the path $\psi^* := \psi^*_{\beta_0-\Delta}$, and rejection process ϱ^n. For $t \ge (j_1^n+1)\nu$, set

$$\begin{aligned}\varphi_2^n(t) &= \beta_0 - \Delta + y\big(t - (j_1^n+1)\nu\big) + \psi^{*,1}\big(t - (j_1^n+1)\nu\big)\\ &\quad - \psi^{*,2}\big(t - (j_1^n+1)\nu\big) - \varrho^n\big(t - (j_1^n+1)\nu\big),\end{aligned}$$

where $\varrho^n(s) = \check{R}^n(s) - \check{R}^n(\tau_1^n)$, $s \ge 0$. Notice that at time $t = (j_1^n+1)\nu$ there is an initial amount of rejections $(\check{R}^n((j_1^n+1)\nu) - \check{R}^n(\tau_1^n)) \ge 0$ and, therefore,

$$\varphi_1^n\big((j_1^n+1)\nu\big) > \beta_0 - \Delta - \big(\check{R}^n\big((j_1^n+1)\nu\big) - \check{R}^n(\tau_1^n)\big) = \varphi_2^n\big((j_1^n+1)\nu\big).$$

Therefore,

$$\varphi_3^n := \boldsymbol{y} + \psi^{0,1} - \psi^{0,2} - \check{R}^n \ge \varphi_2^n \qquad \text{on the interval } \big[(j_1^n+1)\nu, \bar{T}\big]. \tag{56}$$

Let

$$\tau_2^n := \inf\big\{t \ge (j_1^n+1)\nu : \varphi_2^n(t) \le 0\big\}.$$

From (34), it follows that it takes $\int_0^{\beta_0-\Delta} 1/\sqrt{y^2 - h(\xi)/s}\,d\xi$ time units for the path $\beta_0 - \Delta + y\cdot + \psi^{*,1}(\cdot) - \psi^{*,2}(\cdot)$ to reach the level zero. Therefore,

$$\tau_2^n < T_2 \qquad \text{with probability 1.} \tag{57}$$

Define

$$\varphi^n(t) = \begin{cases}\varphi_1^n(t), & 0 \le t \le (j_1^n+1)\nu,\\ \varphi_2^n(t), & (j_1^n+1)\nu < t \le (j_2^n+1)\nu.\end{cases}$$

Step 4 (*Estimate on the time* τ_2^n): Consider the $\mathcal{P}^{2I}$-path $\bar{\psi}^0$, which is defined in a similar way to $\bar{\psi}^\sharp$. From Lemma A.2, it follows that there exists $\bar{\psi}^0 = (\bar{\psi}^{0,1}, \bar{\psi}^{0,2}) \in \mathcal{P}^{2I}$ such that

$$\big(\theta\cdot\bar{\psi}^{0,1}, \theta\cdot(\bar{\psi}^{0,2}\circ\boldsymbol{\rho})\big) = \big(\psi^{0,1}, \psi^{0,2}\big) \tag{58}$$

and

$$\mathbb{J}_k(T, \bar{\psi}^{0,k}) = \mathbb{I}_k(T, \psi^{0,k}), \qquad k = 1, 2. \tag{59}$$

Let

$$H_k^n = \{\tau_k^n \in N_{j_k^n}\}, \qquad k = 1, 2,$$

and similarly to Step 2, set

$$j_2^n = \operatorname*{arg\,max}_{j \in \{j_1^n+1, \dots, j_1^n+\lfloor T_2/\nu \rfloor + 2\}} \mathbb{P}\big(\tau_2^n \in N_j | \Omega^n_{\delta_3, (j_1^n+1)\nu + T_2}(\bar{\psi}^0) \cap H_1^n\big).$$

LEMMA 4.5. *One has*

$$\liminf_{n\to\infty} \frac{1}{b_n^2} \log \mathbb{P}\big(H_2^n | \Omega^n_{\delta_3, (j_2^n+1)\nu}(\bar{\psi}^0) \cap H_1^n\big) \geq -2\varepsilon_1.$$

PROOF. Recall that $\tau_2^n \in [(j_1^n+1)\nu, (j_1^n+1)\nu + T_2)$ with probability 1. Therefore,

$$\mathbb{P}\big(H_2^n | \Omega^n_{\delta_3, (j_1^n+1)\nu + \bar{T}_2}(\psi^0) \cap H_1^n\big) \geq \frac{1}{(\lfloor T_2/\nu \rfloor + 1)}.$$

The rest of the proof is similar to the proof of Lemma 4.4 and is therefore omitted. □

Step 5 (*Bounding the cost from below*): Let us denote

$$\Omega_2^n = \Omega^n_{\delta_3, (j_2^n+1)\nu}(\bar{\psi}^0).$$

Consider the event $\Omega_3^n := H_1^n \cap H_2^n \cap \Omega_2^n \cap E^n$. On this event, we bound from below the sum

$$\int_0^{\tau_2^n} h\big(\check{X}^n(s)\big)\, ds + r\check{R}^n(\tau_2^n).$$

By (75), $\check{X}^n \geq \varphi_1^n - 2\delta_1$ on $[0, (j_1^n+1)\nu]$, and $\check{X}^n \geq \varphi_3^n - 2\delta_1$ on $[(j_1^n+1)\nu, \tau_2^n]$ (thanks to the fact that in Lemma 4.4, ψ is arbitrary on the latter time interval). Therefore, by (77) and the definition of φ^n, on the time interval $[0, \tau_2^n)$,

$$\check{X}^n \geq \varphi^n - 2\delta_1.$$

Since we chose δ_1 such that $\operatorname{osc}_D(2\delta_1, h) < \varepsilon_1$ it follows that

$$\begin{aligned}
&\int_0^{\tau_2^n} h\big(\check{X}^n(t)\big)\, dt + r\check{R}^n(\tau_2^n) \\
&\quad \geq \int_0^{\tau_2^n} h\big(\varphi^n(t)\big)\, dt + r\check{R}^n(\tau_2^n) - \varepsilon_1 \tau_2^n \\
&\quad \geq \left[\int_0^{\tau_1^n} h\big(\varphi_1^n(t)\big)\, dt + r\check{R}^n(\tau_1^n)\right] \\
&\qquad + \left[\int_{(j_1^n+1)\nu}^{\tau_2^n} h\big(\varphi_2^n(t)\big)\, dt + r\big(\check{R}^n(\tau_2^n) - \check{R}^n(\tau_1^n)\big)\right] - \varepsilon_1 \tau_2^n,
\end{aligned} \tag{60}$$

where the last inequality follows since h is nonnegative and $\tau_1^n < (j_1^n + 1)\nu$. We now bound from below the three terms above. From inequality (29), the inequality $\tau_1^n \geq j_1^n \nu$, the definitions of ψ^0 and $\psi^\sharp$, and (54) it follows that

$$
\begin{aligned}
&\int_0^{\tau_1^n} h\big(\varphi_1^n(t)\big)\,dt + r\check{R}^n(\tau_1^n) \\
&\quad \geq r\big(x - (\beta_0 + \Delta)\big) + \mathbb{I}(\tau_1^n, \psi^\sharp) \geq r\big(x - (\beta_0 + \Delta)\big) + \mathbb{I}(j_1^n \nu, \psi^\sharp) \\
&\quad = r\big(x - (\beta_0 + \Delta)\big) + \mathbb{I}\big((j_1^n + 1)\nu, \psi^0\big) \\
&\qquad - \int_{j_1^n \nu}^{(j_1^n+1)\nu} \big[s_1\big(\dot{\psi}^{\sharp,1}\big)^2(t) + s_2\big(\dot{\psi}^{\sharp,2}\big)^2(t)\big]\,dt \\
&\quad = r\big(x - (\beta_0 + \Delta)\big) + \mathbb{I}\big((j_1^n + 1)\nu, \psi^0\big) - r^2\nu/(4s) \\
&\quad \geq r\big(x - (\beta_0 + \Delta)\big) + \mathbb{I}\big((j_1^n + 1)\nu, \psi^0\big) - \varepsilon_1.
\end{aligned} \tag{61}
$$

To bound the second term, notice that

$$
\bar{\varphi}_2^n(s) := \varphi_2^n\big(s + (j_1^n + 1)\nu\big) = \beta_0 - \Delta + ys + \psi^{*,1}(s) - \psi^{*,2}(s) - \varrho^n(s), \qquad s \geq 0.
$$

Denote $\check{\tau}_2^n = \tau_2^n - (j_1^n + 1)\nu$. This is the first time when $\bar{\varphi}_2^n$ hits zero. Therefore, from the definitions of $\mathbb{I}$, $\check{\tau}_2^n$, and ψ^0, the inequality $\tau_2^n \geq j_2^n \nu$, and from inequality (33) it follows that

$$
\begin{aligned}
&\int_{(j_1^n+1)\nu}^{\tau_2^n} h\big(\varphi_2^n(t)\big)\,dt + r\big(\check{R}^n(\tau_2^n) - \check{R}^n(\tau_1^n)\big) \\
&\quad = \int_0^{\check{\tau}_2^n} h\big(\bar{\varphi}_2^n(u)\big)\,du + r\varrho^n(\tau_2^n) \\
&\quad = c\big(\beta_0 - \Delta, \check{\tau}_2^n, \psi^*, \varrho^n\big) + \mathbb{I}\big(\check{\tau}_2^n, \psi^*\big) \\
&\quad = c\big(\beta_0 - \Delta, \check{\tau}_2^n, \psi^*, \varrho^n\big) + \mathbb{I}(\tau_2^n, \psi^0) - \mathbb{I}\big((j_1^n + 1)\nu, \psi^0\big) \\
&\quad \geq c\big(\beta_0 - \Delta, \check{\tau}_2^n, \psi^*, \varrho^n\big) + \mathbb{I}(j_2^n \nu, \psi^0) - \mathbb{I}\big((j_1^n + 1)\nu, \psi^0\big) \\
&\quad \geq V(\beta_0 - \Delta) + \mathbb{I}(j_2^n \nu, \psi^0) - \mathbb{I}\big((j_1^n + 1)\nu, \psi^0\big).
\end{aligned} \tag{62}
$$

Inequality (57) gives a lower bound on the third term. Combining inequalities (60), (61) and (62), we obtain

$$
\begin{aligned}
&\int_0^{\tau_2^n} h\big(\check{X}^n(t)\big)\,dt + r\check{R}^n(\tau_2^n) \\
&\qquad \geq r\big(x - (\beta_0 + \Delta)\big) + V(\beta_0 - \Delta) - \varepsilon_1(\bar{T} + 1) + \mathbb{I}(j_2^n \nu, \psi^0).
\end{aligned} \tag{63}
$$

Step 6 (*Bounding the limit from below*): We are now ready to prove (39). First, notice that there are only finitely many possible pairs $\{(j_1, j_2) \in \mathbb{N}^2 : 0 \leq j_1 \leq j_2 \leq \bar{T}/\nu - 1\}$. For each such pair, define

$$
\mathbb{N}_{(j_1, j_2)} = \big\{n \in \mathbb{N} : \big(j_1^n, j_2^n\big) = (j_1, j_2)\big\}.
$$

If we show that for each pair (j_1, j_2),

$$\liminf_{\mathbb{N}_{(j_1,j_2)}} \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar T} h(\check X^n(u))\,du + r\check R^n(\bar T))}\big] \geq V(x) + C_0\varepsilon_1, \tag{64}$$

where C_0 is a constant independent of n and ε_1, then (39) will follow on applying Lemma 4.1 and taking $\varepsilon_1 \to 0$. Thus, in the rest of the proof we focus on a fixed (j_1, j_2), and prove (64). Hereafter, lim inf denotes the limit inferior along the subset. Denote $\hat\Omega_2^n = \Omega^n_{\delta_3,(j_2+1)\nu}(\bar\psi^0)$. From (63), and since $\tau_2^n \leq \bar T$ it follows that

$$\begin{aligned}
&\liminf \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar T} h(\check X^n(u))\,du + r\check R^n(\bar T))} 1_{\{H_1^n \cap H_2^n \cap \hat\Omega_2^n \cap E^n\}}\big] \\
&\quad \geq r\big(x - (\beta_0 + \Delta)\big) + V(\beta_0 - \Delta) - \varepsilon_1(\bar T + 1) + \mathbb{I}\big(j_2\nu, \psi^0\big) \\
&\qquad + \liminf \frac{1}{b_n^2} \log \mathbb{P}\big(H_1^n \cap H_2^n \cap \hat\Omega_2^n \cap E^n\big).
\end{aligned} \tag{65}$$

We now estimate the last term above. By Lemmas 4.4 and 4.5, for all n sufficiently large,

$$\mathbb{P}\big(H_1^n | \hat\Omega_2^n\big)\mathbb{P}\big(H_2^n | H_1^n \cap \hat\Omega_2^n\big) \geq e^{-5\varepsilon_1 b_n^2}.$$

Hence,

$$\begin{aligned}
&\liminf \frac{1}{b_n^2} \log \mathbb{P}\big(H_1^n \cap H_2^n \cap \hat\Omega_2^n \cap E^n\big) \\
&\quad \geq \liminf \frac{1}{b_n^2} \log\big[\mathbb{P}\big(H_1^n \cap H_2^n \cap \hat\Omega_2^n\big) - \mathbb{P}\big((E^n)^c\big)\big] \\
&\quad = \liminf \frac{1}{b_n^2} \log\big[\mathbb{P}\big(H_2^n | H_1^n \cap \hat\Omega_2^n\big) \\
&\qquad \times \mathbb{P}\big(H_1^n | \hat\Omega_2^n\big)\mathbb{P}\big(\hat\Omega_2^n\big) - \mathbb{P}\big((E^n)^c\big)\big] \\
&\quad \geq \liminf \frac{1}{b_n^2} \log\big(e^{-5\varepsilon_1 b_n^2}\mathbb{P}\big(\hat\Omega_2^n\big) - \mathbb{P}\big((E^n)^c\big)\big) \\
&\quad \geq \liminf \frac{1}{b_n^2} \log\big(e^{-b_n^2[\mathbb{J}((j_2+1)\nu, \bar\psi^0) + 6\varepsilon_1]} - e^{b_n^2(-m+1)}\big) \\
&\quad = -\mathbb{I}\big((j_2+1)\nu, \psi^0\big) - 6\varepsilon_1.
\end{aligned} \tag{66}$$

Above, the third inequality follows by Proposition 2.1 and (48). The last equality uses (59) and $m - 1 > \mathbb{I}((j_2+1)\nu, \psi^0) + 6\varepsilon_1$. Substituting (66) in (65) yields

$$\begin{aligned}
&\liminf \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2(\int_0^{\bar T} h(\check X^n(u))\,du + r\check R^n(\bar T))} 1_{\{H_2^n \cap H_1^n \cap \tilde A_{\delta_3,(j_2+1)\nu}(\psi^0) \cap E^n\}}\big] \\
&\quad \geq r\big(x - (\beta_0 + \Delta)\big) + V(\beta_0 - \Delta) - \varepsilon_1(\bar T + 7) \\
&\qquad + \mathbb{I}\big(j_2\nu, \psi^0\big) - \mathbb{I}\big((j_2+1)\nu, \psi^0\big).
\end{aligned} \tag{67}$$

Using (30) and then (31) and Proposition 3.1(iii) gives

$$
\begin{aligned}
&\mathbb{I}(j_2\nu,\psi^0)-\mathbb{I}((j_2+1)\nu,\psi^0)\\
&\quad=-\int_{j_2\nu}^{(j_2+1)\nu}\big[s_1(\dot\psi^{*,1}_{\beta_0-\Delta})^2(t)+s_2(\dot\psi^{*,2}_{\beta_0-\Delta})^2(t)\big]\,dt\\
&\quad=-\int_{j_2\nu}^{(j_2+1)\nu}\big[s(\dot\omega^{*}_{\beta_0-\Delta})^2(t)\big]\,dt \qquad (68)\\
&\quad=-\int_{j_2\nu}^{(j_2+1)\nu}\big[s\big(-y-\sqrt{y^2-h(x+yt+\omega^{*}_{\beta_0-\Delta}(t))/s}\big)^2(t)\big]\,dt\\
&\quad\geq -\nu s y^2\geq-\varepsilon_1,
\end{aligned}
$$

where the last two inequalities follow from the negativity of y and (54). From (67) and (68) and by recalling that for $x>\beta_0$, one has $V(x)=r(x-\beta_0)+V(\beta_0)$, and it follows that

$$
\begin{aligned}
&\liminf\frac{1}{b_n^2}\log\mathbb{E}\big[e^{b_n^2(\int_0^{\bar T}h(\check X^n(u))\,du+r\check R^n(\bar T))}1_{\{H_1^n\cap H_2^n\cap\Omega^n_{\delta_3,(j_2^n+1)\nu}(\psi^0)\cap E^n\}}\big]\\
&\quad\geq V(x)-\varepsilon_1(\bar T+8).
\end{aligned}
$$

This proves (64). Hence, the result is proved for the case $x>\beta_0$.

Finally, consider $x\leq\beta_0$. The considerations here are simpler than in the previous case. In case that the initial state is exactly β_0, the decision maker can reject a (small) amount of Δ at time $t=0$. Then the proof that $V(x)$ is a lower bound requires the focusing only on data near $\psi^*_{\beta_0-\Delta}$, starting at time zero. In case that the initial state is lower than β_0, one uses the same arguments, with data near ψ^*_x. □

4.1. *Proof of Lemma 4.4.* The proof has two parts. On the first, we show that for sufficiently large n one has

$$
\mathbb{P}\big(\tau_1^n\in[0,T_1)|\Omega^n_{\delta_3,T_1}(\psi)\big)\geq 1/2. \tag{69}
$$

Since the interval $[0,T_1]$ is divided into at most $\lfloor T_1/\nu\rfloor+1$ subintervals, there exists an interval N_j such that the conditional probability of $\tau_1^n\in N_j$ is at least $\frac{1}{2(\lfloor T_1/\nu\rfloor+1)}$. Thus, as a result of (69),

$$
\mathbb{P}\big(\tau_1^n\in N_{j_1^n}|\Omega^n_{\delta_3,T_1}(\psi)\big)\geq\frac{1}{2(\lfloor T_1/\nu\rfloor+1)}. \tag{70}
$$

On the second part, we use this to deduce (55).

Part (a): Set

$$
E_1^n=\{\tau_1^n\geq T_1\}.
$$

Write $\Omega^n_{\delta_3,T_1}(\psi)$ as Ω^n. We analyze the event $\Omega^n_1 := \Omega^n \cap E^n \cap E^n_1$. On this event, we bound from below the process $\check{X}^n$ on the time interval $[0, T_1]$ and the total number of rejections until time T_1. By the triangle inequality it follows that for every $i \in \mathcal{I}$ and every n

$$\begin{aligned} \|\tilde{S}^n_i \circ T^n_i - \psi^2_i \circ \boldsymbol{\rho}_i\|_{\bar{T}} &\le \|\tilde{S}^n_i \circ T^n_i - \psi^2_i \circ T^n_i\|_{\bar{T}} \\ &\quad + \|\psi^2_i \circ T^n - \psi^2_i \circ \boldsymbol{\rho}_i\|_{\bar{T}} \\ &\le 2\delta_2, \end{aligned} \tag{71}$$

where we have bounded each of the terms on the right-hand side by δ_2; the bound of the first term follows by (52), and the bound of the second follows by the definition of E^n and from (51). Similarly,

$$\|\tilde{A}^n_i - \psi^1_i\|_{\bar{T}} \le \delta_2, \qquad i \in \mathcal{I}. \tag{72}$$

Since $\theta^n \to \theta$ it follows from (49), (71) and (72) that for sufficiently large n,

$$\|\check{S}^n \circ T^n - \psi^2 \circ \boldsymbol{\rho}\|_{\bar{T}} = \|\theta^n \cdot \tilde{S}^n \circ T^n - \theta \cdot \psi^2 \circ \boldsymbol{\rho}\|_{\bar{T}} < \frac{\delta_1}{2}, \tag{73}$$

$$\|\check{A} - \psi^1\|_{\bar{T}} = \|\theta^n \cdot \tilde{A}^n - \theta \cdot \psi^1\|_{\bar{T}} < \frac{\delta_1}{4}. \tag{74}$$

Moreover, for sufficiently large n one has, for $t \in [0, \bar{T}]$, $|\check{X}^n(0) + \check{y}^n t - \boldsymbol{y}(t)| \le \delta_1/4$. Using the above inequalities, it follows that for every $u \in [0, T_1]$ one has

$$\begin{aligned} \check{X}^n(u) &= \check{X}^n(0) + \check{y}^n t + \check{A}^n(u) - \check{S}^n \circ T^n(u) + \check{Z}^n(u) - \check{R}^n(u) \\ &\ge \varphi^n_1(u) - \delta_1 - \sup_{t\in[0,\bar{T}]} \|\check{X}^n(0) + \tilde{y}^n t - \boldsymbol{y}(t)\| \\ &\quad - \|\check{A}^n - \psi^1\|_{\bar{T}} - \|\check{S}^n \circ T^n - \psi^2\|_{\bar{T}} \\ &\ge \varphi^n_1(u) - 2\delta_1. \end{aligned} \tag{75}$$

By the definition of E^n_1 and the choice of δ_1 [see (47)],

$$\check{X}^n(u) \ge \beta_0 + \Delta - 2\delta_1 \ge \beta_0 + \Delta/2. \tag{76}$$

By using similar arguments and the inequality $\check{X}^n \le D^n$, one obtains

$$\begin{aligned} \check{R}^n(T_1) &\ge \boldsymbol{y}(T_1) - D^n - \delta_1 + \psi^1(T_1) - \psi^2(T_1) \\ &= x - D^n - \delta_1 + \big(y + r/(2s)\big)T_1 \\ &\ge x - D^n - 1 + \big(y + r/(2s)\big)T_1 \\ &\ge x - D - 2 + \big(y + r/(2s)\big)T_1, \end{aligned} \tag{77}$$

where the equality follows by the definition of $\psi^\sharp$, the second inequality follows by the choice of δ_1, and the last inequality follows since $\lim_{n\to\infty} D^n = D$. From (37), (42) and since $T_1 < \bar{T}$, we obtain[4] for sufficiently large n

$$V(x)+1 > \frac{1}{b_n^2}\log\mathbb{E}\big[e^{b_n^2(\int_0^{T_1} h(\check{X}^n(u))\,du + r\check{R}^n(T_1))}\big].$$

Along with (76) and (77), it follows that

$$\begin{aligned} V(x)+1 &> \frac{1}{b_n^2}\log\mathbb{E}\big[e^{b_n^2(\int_0^{T_1} h(\check{X}^n(u))\,du + r\check{R}^n(T_1))}1_{\Omega_1^n}\big] \\ &\geq \frac{1}{b_n^2}\log\mathbb{E}\big[e^{b_n^2((h(\beta_0+\Delta/2)+yr+r^2/(2s))T_1 + r(x-D-2))}1_{\Omega_1^n}\big] \\ &= r(x-D-2) + \big(h(\beta_0+\Delta/2) - h(\beta_0) + r^2/(4s)\big)T_1 \\ &\quad + \frac{1}{b_n^2}\log\mathbb{P}(\Omega_1^n). \end{aligned}$$

The above equality follows since $r^2/(4s) + ry + h(\beta_0) = 0$, which in turn follows since $x > \beta_0$ and therefore $\beta_0 < D$. Since

$$\mathbb{P}(\Omega_1^n) = \mathbb{P}(E_1^n \cap \Omega^n \cap E^n) \geq \mathbb{P}(E_1^n|\Omega^n)\mathbb{P}(\Omega^n) - \mathbb{P}\big((E^n)^c\big),$$

and using (48), it follows that

$$\begin{aligned} \mathbb{P}(E_1^n|\Omega^n) &\leq e^{b_n^2(V(x)+1+r(2+D-x)-T_1(h(\beta_0+\Delta/2)-h(\beta_0)-r^2/(4s)))}\mathbb{P}(\Omega^n)^{-1} \\ &\quad + e^{-mb_n^2}\mathbb{P}(\Omega^n)^{-1}. \end{aligned}$$

We show that for sufficiently large n, each of the terms on the right-hand side can be bounded by $1/4$. From Proposition 2.1 and (50), it follows that, for all large n,

$$\begin{aligned} \frac{1}{b_n^2}\log\mathbb{P}(\Omega^n) &\geq -\inf_{\bar{\psi}\in\mathcal{A}_{\delta_3,T_1}(\psi)} \mathbb{J}(T_1,\bar{\psi}) - 1 \geq -\mathbb{J}(T_1,\psi) - 1 \\ &= -\mathbb{I}(T_1,\psi^\sharp) - 1 = -r^2T_1/(4s) - 1. \end{aligned}$$

Hence, by the definition of T_1, the first term is bounded by $1/4$. Since $m > \mathbb{I}(\bar{T},\psi^\sharp)$ and $\bar{T} > T_1$, so is the second term. As a result, (69) holds.

Part (b): By the definition of an admissible control, R^n is adapted to the filtration $\mathcal{F}_t := \sigma\{A_i^n(u), S_i^n(T_i^n(u)), i\in\mathcal{I}, u\leq t\}$, hence so is $\check{R}^n$, and, by (41), so is φ_1^n. Since $T_i^n(t)\leq t$, $t\geq 0$, and T_i^n are themselves adapted to $\mathcal{F}_t$, it follows that the event $\{\tau_1^n \in N_{j_1^n}\}$ is measurable on $F^n_{\nu(j_1^n+1)}$, where $F_t^n = \sigma\{A_i^n(u), S_i^n(u), i\in\mathcal{I}, u\leq t\}$. Note that $F_t^n = \sigma\{\tilde{A}_i^n(u), \tilde{S}_i^n(u), i\in\mathcal{I}, u\leq t\}$.

[4]Recall the convention $\varepsilon = 0$ and generic r.

Fix $v > 0$ and a sequence v_n, with $v_n < v$ (both deterministic). We will show the following. Given a constant $c_1 > 0$ and a sequence of events $Q_n \in F^n_{v_n}$, for every $\varepsilon > 0$ there exists $\delta > 0$ and $n_1 \in \mathbb{N}$, such that

$$p^n_1 := \mathbb{P}\big(Q_n | \Omega^n_{\delta,v}(\psi)\big) \geq c_1, \qquad n \geq 1, \tag{78}$$

implies

$$p^n_2 := \mathbb{P}\big(Q_n | \Omega^n_{\delta,v_n}(\psi)\big) \geq e^{-\varepsilon b^2_n}, \qquad n \geq n_1. \tag{79}$$

Note that this will prove (55), based on (70) that has now been established by part (a).

Extending the definition of $\Omega^n_{\delta,t}(\psi)$ (53), we let

$$\Omega^n_{\delta,a,b}(\psi) = \Big\{ \sup_{s\in[a,b]} \big\|(\tilde{A}^n, \tilde{S}^n)(u) - \psi(u)\big\| < \delta \Big\},$$

for $0 \leq a \leq b$. Also, we drop ψ from the notation $\Omega^n_{\delta,a}$ and $\Omega^n_{\delta,a,b}$. Note that there is no loss of generality in proving the statement for $\tilde{A}^n$ (a collection of I independent renewal processes) in place of $(\tilde{A}^n, \tilde{S}^n)$ (a collection of $2I$ such processes). Thus, we will consider only the former.

To prove the aforementioned statement, let $\varepsilon > 0$ be given. Consider the quantities p^n_1 and p^n_2, depending on δ. Assume that (78) is valid. Then we can write

$$p^n_1 = \frac{\mathbb{P}\{Q_n \cap \Omega^n_{\delta,v}\}}{\mathbb{P}\{\Omega^n_{\delta,v}\}} = \frac{\mathbb{P}\{Q_n \cap \Omega^n_{\delta,v_n} \cap \Omega^n_{\delta,v_n,v}\}}{\mathbb{P}\{\Omega^n_{\delta,v_n} \cap \Omega^n_{\delta,v_n,v}\}}.$$

A basic independence property for a renewal process, to be used, is the following. Let A be a renewal process of the form

$$A(t) = \sup\Big\{ l \geq 0 : \sum_{k=1}^{l} U(k) \leq t \Big\}, \qquad t \geq 0,$$

where $\{U(k)\}$ are i.i.d. [compare with (3)]. Fix t and let π denote the time of the first jump at or after t, namely $\pi = \inf\{u \geq t : A(u) > A(t-)\}$ [convention: $A(0-) = 0$]. Consider an event Q measurable on $\sigma\{A(u) : 0 \leq u \leq t\}$. Then, for each $k \in \mathbb{Z}_+$, the event $Q \cap \{A_t = k\}$ is statistically independent of the sequence $\{U_{k+1}, U_{k+2}, \ldots\}$. Based on this, it is not hard to see that, if we let SA denote the shifted version $SA(u) = A(\pi + u) - A(\pi)$, $u \geq 0$, of A, we have that Q is independent of SA. For a collection of independent renewal processes, a similar statement holds if each of them is shifted according to its own first jump after t. To state this property for the processes (A^n_i), if Q is measurable on F^n_t, then it is independent of (SA^n_i). Now, let us apply this to study $(\tilde{A}^n)$. Let $\pi_n = (\pi^n_i)$ be defined by $\pi^n_i = \inf\{u \geq v_n : \tilde{A}^n_i(u) > \tilde{A}^n_i(v_n-)\}$. If $\|\pi_n\| := \max_i |\pi^n_i - v_n|$, then given any $k > 0$, $\mathbb{P}\{\|\pi_n\| > \delta\} \leq e^{-kb^2_n}$ for all sufficiently large n, as can be verified

using the exponential moment assumption and applying Chebychev's inequality. Hence, given any k, for all large n,

$$\begin{aligned} p_{1,1}^n &:= \mathbb{P}\{Q_n \cap \Omega_{\delta,v_n}^n \cap \Omega_{\delta,v_n,v}^n\} \\ &\le \mathbb{P}\{Q_n \cap \Omega_{\delta,v_n}^n \cap \Omega_{\delta,v_n,v}^n \cap \{\|\pi_n\| \le \delta\}\} + e^{-kb_n^2}. \end{aligned}$$

Let us denote by ω a modulus [by which we mean a function mapping $\mathbb{R}_+$ to itself with $\omega(0+) = 0$], that dominates the modulus of continuity of $\bar{\psi}^m$ for all $m \le M$. Then, using the definition of $\Omega_{\delta,v_n,v}^n$, adding and subtracting the shifted version $S\tilde{A}^n$ and using the triangle inequality gives

$$p_{1,1}^n \le \mathbb{P}\{Q_n \cap \Omega_{\delta,v_n}^n \cap \{\|S\tilde{A}^n - (\psi(v_n + \cdot) - \psi(v_n))\|_{v-v_n-\delta} < \delta'\}\} + e^{-kb_n^2},$$

where $\delta' = 2\delta + 2\omega(\delta)$. Using the independence alluded to above,

$$p_{1,1}^n \le \mathbb{P}\{Q_n \cap \Omega_{\delta,v_n}^n\}\mathbb{P}\{\|S\tilde{A}^n - (\psi(v_n + \cdot) - \psi(v_n))\|_{v-v_n-\delta} < \delta'\} + e^{-kb_n^2}.$$

Shifting back gives

$$p_{1,1}^n \le \mathbb{P}\{Q_n \cap \Omega_{\delta,v_n}^n\}\mathbb{P}\{\Omega_{\delta'',v_n,v-\delta}^n\} + e^{-kb_n^2}$$

for δ'' that can be made arbitrarily small by taking δ to be small. A similar argument shows that, for all large n,

$$p_{1,2}^n := \mathbb{P}\{\Omega_{\delta,v_n}^n \cap \Omega_{\delta,v_n,v}^n\} \ge \mathbb{P}\{\Omega_{\delta''',v_n}^n\}\mathbb{P}\{\Omega_{\delta''',v_n,v}^n\} - e^{-kb_n^2},$$

for suitably chosen $\delta''' > 0$, that again, can be made arbitrarily small by taking small δ. Thus,

$$\begin{aligned} p_1^n = \frac{p_{1,1}^n}{p_{1,2}^n} &\le \frac{\mathbb{P}\{Q_n \cap \Omega_{\delta,v_n}^n\}\mathbb{P}\{\Omega_{\delta'',v_n,v-\delta}^n\} + e^{-kb_n^2}}{\mathbb{P}\{\Omega_{\delta''',v_n}^n\}\mathbb{P}\{\Omega_{\delta''',v_n,v}^n\} - e^{-kb_n^2}} \\ &= \frac{\mathbb{P}\{Q_n|\Omega_{\delta,v_n}^n\}\mathbb{P}\{\Omega_{\delta,v_n}^n\}\mathbb{P}\{\Omega_{\delta'',v_n,v-\delta}^n\} + e^{-kb_n^2}}{\mathbb{P}\{\Omega_{\delta''',v_n}^n\}\mathbb{P}\{\Omega_{\delta''',v_n,v}^n\} - e^{-kb_n^2}}. \end{aligned}$$

Using the LDP and writing $\mathbb{I}[a,b]$ for $\sum_i \int_a^b s_i \dot{\psi}(u)^2\,du$ [see (10)] gives

$$\begin{aligned} p_1^n &\le \frac{p_2^n e^{b_n^2(-\mathbb{I}[0,v_n]+\varepsilon')} e^{b_n^2(-\mathbb{I}[v_n,v-\delta]+\varepsilon')} + e^{-kb_n^2}}{e^{b_n^2(-\mathbb{I}[0,v_n]-\varepsilon')} e^{b_n^2(-\mathbb{I}[v_n,v]-\varepsilon')} - e^{-kb_n^2}} \\ &= \frac{p_2^n e^{b_n^2(-\mathbb{I}[0,v-\delta]+2\varepsilon')} + e^{-kb_n^2}}{e^{b_n^2(-\mathbb{I}[0,v]-2\varepsilon')} - e^{-kb_n^2}}, \end{aligned}$$

for arbitrary $\varepsilon' > 0$, provided that δ is sufficiently small. Hence, by selecting k sufficiently large,

$$p_1^n \le \frac{p_2^n e^{b_n^2(-\mathbb{I}[0,v-\delta]+2\varepsilon')} + e^{-kb_n^2}}{\frac{1}{2}e^{b_n^2(-\mathbb{I}[0,v]-2\varepsilon')}}$$
$$= 2p_2^n e^{b_n^2(4\varepsilon'+\mathbb{I}[v-\delta,v])} + 2e^{-kb_n^2} e^{b_n^2(\mathbb{I}[0,v]+2\varepsilon')}.$$

Hence, again by selecting k large, for all large n,

$$p_2^n \ge \frac{1}{2}c_1 e^{-b_n^2(4\varepsilon'+\mathbb{I}[v-\delta,v])} - e^{b_n^2(-k+\mathbb{I}[0,v-\delta]-2\varepsilon')} \ge \frac{1}{4}c_1 e^{-b_n^2(4\varepsilon'+\mathbb{I}[v-\delta,v])}.$$

Selecting $\delta > 0$ such that $4\varepsilon' + \mathbb{I}[v-\delta, v] < \varepsilon$ gives (79) for some n_1. This completes the proof of part (b) and the lemma.

5. A nearly optimal policy. In this section, we show that the policy from [6] is AO for the present setting. While the policy is similar, the proof of AO is quite different, as the paper [6] addresses the diffusion scale, rather than the MD scale.

Let the classes be labeled so that $h_1\mu_1 \ge h_2\mu_2 \ge \cdots \ge h_I\mu_I$. Let $\gamma : [0, D] \to \mathcal{X}$ be a Borel measurable mapping satisfying

$$\gamma(w) \in \arg\min_{\xi}\{\bar{h}\cdot\xi : \xi \in \mathcal{X}, \theta\cdot\xi = w\}, \qquad w \in [0, D]. \tag{80}$$

We note on passing that, as shown in [2], Theorem A.1, one can equivalently work with one-dimensional dynamics, thanks to the fact that the minimum over queue length ξ in the above expression is a function of only of the (one-dimensional) workload w.

Since the mapping $\bar{\xi} \mapsto \bar{h}\cdot\bar{\xi}$ is linear and the domain $\mathcal{X}$ is polyhedral, it can be assumed, without loss of generality, that γ is continuous and takes values on the boundary of $\mathcal{X}$. We have, by definition, that $\theta\cdot\gamma(w) = w$, and $\bar{h}\cdot\gamma(w) = h(w) \le \bar{h}\cdot\xi$ for $\xi \in \mathcal{X}$ for which $\theta\cdot\xi = w$. A particular selection of γ is as follows. Given $w \in [0, D]$, set $(j, \xi) = (j, \xi)(w)$ by $w \in [\hat{D}_j, \hat{D}_{j-1})$ and $\xi = \xi(w) := (w - \hat{D}_j)/\theta_j$, where

$$\hat{D}_j := \sum_{i=j+1}^{I} \theta_i D_i, \qquad j \in \{0, \ldots, I\}$$

and one has $0 = \hat{D}_I < \hat{D}_{I-1} < \cdots < \hat{D}_1 < \hat{D}_0 = \theta\cdot\bar{D} = D$, $\bar{D} := (D_1, \ldots, D_I)$. Then

$$\gamma(w) = \sum_{i=j+1}^{I} D_i e_i + \xi e_j. \tag{81}$$

Approximate γ by a curve that is bounded away from the part of the boundary of $\mathcal{X}$ that corresponds to the buffer limit, namely $\partial\mathcal{X} = \{x \in \mathcal{X} : x_i = D_i \text{ for some } i\}$.

Fix $0 < \varepsilon_0 < \min_i D_i/4$. Let $a_i = D_i - 3\varepsilon_0$, $i \in \mathcal{I}$, and $a^* = \beta_0 \wedge (\theta \cdot a) < D$. Note that if ε_0 is small then $a^* = \beta_0$ (unless $\beta_0 = D$). Define $\gamma^a[0, D] \to \mathcal{X}$ first on $[0, \theta \cdot a]$ as the function obtained upon replacing the parameters (D_i) by (a_i) in (81). That is, for $w \in [0, \theta \cdot a)$, the variables $j = j(w)$ and $\xi = \xi(w)$ are determined via

$$
\begin{aligned}
w &= \sum_{i=j+1}^{I} \theta_i a_i + \theta_j \xi, \qquad j \in \{1, 2, \ldots, I\}, \xi \in [0, a_j), \\
\gamma^a(w) &= \sum_{i=j+1}^{I} a_i e_i + \xi e_j.
\end{aligned} \tag{82}
$$

Given $w \in [0, \theta \cdot a)$, we will sometimes refer to the unique pair (j, ξ) alluded to above as the *representation* (j, ξ) *of* w *via* (82). Next, on $[\theta \cdot a, \theta \cdot \bar{D}]$ define γ^a as the linear interpolation between the points $(\theta \cdot a, a)$ and $(\theta \cdot \bar{D}, \bar{D})$. Also set $\hat{a}_j = \sum_{i=j+1}^{I} \theta_i a_i$, $j \in \{0, 1, \ldots, I\}$. Let

$$
\begin{aligned}
h^a(w) &:= \min\bigl\{\bar{h} \cdot \xi : \xi \in \mathcal{X}, \theta \cdot \xi = w, \xi_i \le \gamma_i^a(w), i = 1, \ldots, I\bigr\} \\
&= \bar{h} \cdot \gamma^a(w), \qquad w \in [0, \theta \cdot a].
\end{aligned}
$$

Note the similarity to the payoff h in (23). Note that the construction depends on the parameter ε_0, and denote

$$
\omega_1(\varepsilon_0) = \sup_{[0, \theta \cdot a]} \bigl|h^a - h\bigr|. \tag{83}
$$

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

Before providing the precise construction of the policy, we explain its rationale. The solution to the DG indicates that rejections should occur when the normalized workload in the system is above the threshold β_0, and that most rejections should be from a specific class, i^*, defined in (24). The DG solution also indicates that prioritization should be according to (80) (see the proof of (100) in [2], Theorem A.1) and that, consequently, the resulting normalized queue length processes should be close to the curve γ. These two goals are contradictory, as parts of the curve γ lie on the part $\partial\mathcal{X}$ of the domain where some of the buffers are full, and so even small stochastic fluctuations cause rejections due to the buffer size constraints. Such rejections do not satisfy the requirement to reject only when the workload is above the specified threshold, nor that rejections are from class i^*. To address this issue, we have defined the curve γ^a, which approximates γ without intersecting the part $\partial\mathcal{X}$ of the boundary. The service policy is designed to keep the normalized queue length processes close to this curve.

The precise definition of the policy is provided by specifying $(B^n(t), R^n(t))$ as a function of $X^n(t)$.

Rejection policy: As under any policy, in order to meet the buffer size constraint (8), all *forced rejections* take place. That is, if a class-i arrival occurs at a time t

when $\tilde{X}_i^n(t-) + \frac{1}{b_n\sqrt{n}} > D_i$, then it is rejected. Apart from that, no rejections occur from any class except class i^*, which is defined through (24), and no rejections occur (from any class) when $\theta \cdot \tilde{X}^n < a^*$. When $\theta \cdot \tilde{X}^n \geq a^*$, all class-$i^*$ arrivals are rejected, and these rejections are called *overload rejections*.

Service policy: For each $\bar{x} \in \mathcal{X}$, define the class of low priority

$$\mathcal{L}(\bar{x}) = \max\{i : x_i < a_i\},$$

provided $x_i < a_i$ for some i, and set $\mathcal{L}(\bar{x}) = I$ otherwise. The complement set is the set of high priority classes

$$\mathcal{H}(\bar{x}) = \mathcal{I} \setminus \{\mathcal{L}(\bar{x})\}.$$

When there is at least one class among $\mathcal{H}(\bar{x})$ having at least one customer in the system, $\mathcal{L}(\bar{x})$ receives no service, and all classes within $\mathcal{H}(\bar{x})$, having at least one customer, receive service at a fraction proportional to their traffic intensities. Namely, denote $\mathcal{H}^+(\bar{x}) = \{i \in \mathcal{H}(\bar{x}) : x_i > 0\}$, and define $\rho'(\bar{x}) \in \mathbb{R}^I$ as

$$\rho_i'(\bar{x}) = \begin{cases} 0 & \text{if } \bar{x} = 0, \\ \dfrac{\rho_i 1_{\{i \in \mathcal{H}^+(\bar{x})\}}}{\sum_{k \in \mathcal{H}^+(\bar{x})} \rho_k} & \text{if } \mathcal{H}^+(\bar{x}) \neq \varnothing, \\ e_I & \text{if } x_i = 0 \text{ for all } i < I \text{ and } x_I > 0. \end{cases} \tag{84}$$

[Note that $\mathcal{H}^+(\bar{x}) = \varnothing$ can only happen if $x_i = 0$ for all $i < I$, which is covered by the first and last cases in the above display.] Then, for each t,

$$B^n(t) = \rho'\big(\tilde{X}^n(t)\big). \tag{85}$$

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

$$\rho_i'(\bar{x}) > \rho_i \qquad \text{for all } i \in \mathcal{H}^+(\bar{x}). \tag{86}$$

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

THEOREM 5.1. *Let Assumption 2.1 hold. For every $\varepsilon_0 > 0$ and $n \in \mathbb{N}$, denote the policy constructed above by $U^n(\varepsilon_0)$. Then, for all sufficiently large T,*

$$\limsup_{n\to\infty} J^n\big(T, \tilde{X}^n(0), U^n(\varepsilon_0)\big) \leq V(x) + \omega(\varepsilon_0),$$

where $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ is a function satisfying $\omega(0+) = 0$.

PROOF OF THEOREM 5.1. Introduce the notation

$$H_t^n = \int_0^t \bar{h} \cdot \tilde{X}^n(u)\, du + \bar{r} \cdot \tilde{R}^n(t).$$

Fix $T > 0$ sufficiently large for the identity $V(T, x) = V(x)$ stated in Proposition 3.1(i) to hold. First, notice that

$$\mathbb{E}\int_0^T e^{b_n^2 H_t^n}\,dt \le T\mathbb{E}\big[e^{b_n^2 H_T^n}\big]. \tag{87}$$

The argument will be based on a bound on the right-hand side of (87).

Recall the definition (10) of $\mathbb{J}$, and for $J > 0$, define

$$\mathcal{AC}_J = \big\{\psi \in \mathcal{D}\big([0,T],\mathbb{R}^{2I}\big) : \mathbb{J}(T,\psi) \le J\big\}.$$

Then $\mathcal{AC}_J$ is compact in the J_1 topology, and consists of absolutely continuous paths starting at zero. Fix $0 < \varepsilon_1 < \theta_{\min}\varepsilon_0/8$. Fix also $\delta_1 > 0$ such that

$$\delta_1 < \min\big\{\varepsilon_0/12, \varepsilon_1/(11C), \varepsilon_0/(5C), \operatorname{osc}_D\big(\varepsilon_0/(5C), h\big)\big\}, \tag{88}$$

where C is the constant from Lemma A.1. For $0 < \delta < t \le T$ and $\psi \in \mathcal{D}([0,T],\mathbb{R}^{2I})$, denote

$$\mathcal{A}_{\delta,t}(\psi) = \big\{\bar\psi \in \mathcal{D}\big([0,T],\mathbb{R}^{2I}\big) : \|\bar\psi - \psi\|_t < \delta\big\}$$

[where we slightly modified the notation (52)]. By the compactness of $\mathcal{AC}_J$ and the continuity of its members, one can find a finite number of members $\bar\psi^1, \bar\psi^2, \dots, \bar\psi^M$ of $\mathcal{AC}_J$, and positive constants $\delta^1, \dots, \delta^M$ with $\delta^m < \delta_1$, satisfying $\mathcal{AC}_J \subset \bigcup_{m=1}^M \mathcal{A}_T^m$, and

$$\inf\big\{\mathbb{J}(T,\bar\psi) : \bar\psi \in \overline{\mathcal{A}_T^m}\big\} \ge \mathbb{J}(T,\bar\psi^m) - \varepsilon_0, \qquad m = 1,2,\dots,M, \tag{89}$$

where, throughout, for $0 \le t \le T$,

$$\mathcal{A}_t^m = \mathcal{A}_{\delta^m,t}(\bar\psi^m).$$

By the continuity of each of the paths $\bar\psi^m$, one can find $\nu_1 > 0$ such that for $m = 1, \dots, M$,

$$\operatorname{osc}_T\big(\nu_1, \bar\psi_i^m\big) \le \frac{\delta_1}{2(\theta_{\max} \vee 1)\sqrt{I}}, \qquad i \in \mathcal{I}. \tag{90}$$

In this proof, $C_1, C_2, \dots$ denote positive constants that do not depend on n, ε_0, δ^m or J. Write $\bar\psi^m \in \mathcal{P}^{2I}$ as $(\bar\psi^{m,1}, \bar\psi^{m,2})$. For $m = 1, \dots, M$, let

$$\varphi^m = \boldsymbol{y} + \psi^{m,1} - \psi^{m,2} + \zeta^m - \varrho^m,$$

where, as before, $\boldsymbol{y}(t) = x + yt$,

$$\psi^{m,1} = \theta \cdot \bar\psi^{m,1}, \qquad \psi^{m,2} = \theta \cdot \bar\psi^{m,2}$$

and

$$(\varphi^n, \zeta^m, \varrho^m) = \Gamma_{[0,\beta_0]}\big[\boldsymbol{y} + \psi^{m,1} - \psi^{m,2}\big]. \tag{91}$$

Denote $\Lambda^n = \|\tilde{A}^n\|_T + \|\tilde{S}^n\|_T$. As argued in [6], at the bottom of page 595,

$$\|\tilde{R}^n(T)\| \le C_1(1+\Lambda^n). \tag{92}$$

Since $\tilde{X}^n$ is bounded, one has $H_T^n \le C_2(1+\Lambda^n)$. Hence, given any $J_1 > 0$,

$$H_T^n > J_1 \quad \text{implies} \quad \Lambda^n > C_2^{-1}J_1 - 1 =: G(J_1).$$

Therefore,

$$\begin{aligned}\mathbb{E}\big[e^{b_n^2 H_T^n}\big] &\le \mathbb{E}\big[e^{b_n^2[H_T^n \wedge J_1]}\big] + \mathbb{E}\big[e^{b_n^2 H_T^n} 1_{\{H_T^n > J_1\}}\big] \\ &\le A_1^n + A_2^n + A_3^n,\end{aligned} \tag{93}$$

where, with $\mathcal{B} = (\bigcup_{m=1}^M \mathcal{A}_T^m)^c$,

$$\begin{aligned}A_1^n &= \sum_{m=1}^M \mathbb{E}\big[e^{b_n^2[H_T^n \wedge J_1]} 1_{\Omega^{n,m}}\big], \qquad \Omega^{n,m} = \{(\tilde{A}^n, \tilde{S}^n) \in \mathcal{A}_T^m\}, \\ A_2^n &= \mathbb{E}\big[e^{b_n^2[H_T^n \wedge J_1]} 1_{\{(\tilde{A}^n, \tilde{S}^n) \in \mathcal{B}\}}\big], \\ A_3^n &= \mathbb{E}\big[e^{b_n^2 C_2(1+\Lambda^n)} 1_{\{\Lambda^n > G(J_1)\}}\big].\end{aligned}$$

An argument to be presented shortly will show that there exist $t_1, \dots, t_M \in [0,T]$ such that for large n,

$$A_1^n \le M \max_{m=1}^M e^{b_n^2[\int_0^{t^m} h(\varphi^m(u))\,du + r(1+\varepsilon_0)\varrho^m(t^m) - \mathbb{I}(t^m, \psi^m) + \omega_2(\varepsilon_0)]} + \varepsilon_0, \tag{94}$$

where $\omega_2(0+) = 0$ (this step translates the multidimensional formulation, by which H_T^n is defined, into a one-dimensional form, given by φ^m). As for A_2^n and A_3^n, note, by Proposition 2.1, that for large n,

$$\frac{1}{b_n^2}\log \mathbb{P}\big((\tilde{A}^n, \tilde{S}^n) \in \mathcal{B}\big) \le -\inf_{\psi \in \mathcal{B}} \mathbb{I}(T, \psi) + \varepsilon_0.$$

Along with the fact that $\mathcal{B} \subset \mathcal{AC}_J^c$ and the definition of $\mathcal{AC}_J$, this shows

$$A_2^n \le e^{b_n^2[J_1 - J + \varepsilon_0]}. \tag{95}$$

Also, $A_3^n \le \mathbb{E}[e^{b_n^2((C_2+1)\Lambda^n + C_2 - G(J_1))}]$. As shown in the Appendix of [1], Assumption 2.1 implies that, for any $K < \infty$,

$$\limsup_{n\to\infty} \frac{1}{b_n^2}\log \mathbb{E}\big[e^{b_n^2 K(\|\tilde{A}^n\|_T + \|\tilde{S}^n\|_T)}\big] < \infty.$$

Hence, there exists a constant C_3 such that

$$A_3^n \le e^{b_n^2 C_2 + C_3 - G(J_1)}. \tag{96}$$

Combining (93), (94), (95) and (96),

$$\begin{aligned}
&\limsup \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2 H_T^n}\big] \\
&\quad \le \max_{1\le m\le M} \bigg[\int_0^{t^m} h\big(\varphi^m(u)\big)\,du + r(1+\varepsilon_0)\varrho^m\big(t^m\big) - \mathbb{I}\big(t^m, \psi^m\big) + \omega_2(\varepsilon_0)\bigg] \\
&\qquad \vee [J_1 - J + \varepsilon_0] \vee \big[C_2 + C_3 - G(J_1)\big] \\
&\quad \le \sup_{\psi\in\mathcal{P}, t\in[0,T]} \big[c^{\varepsilon_0}\big(x, t, \psi, \alpha_{\beta_0}[\psi]\big) + \omega_2(\varepsilon_0)\big] \\
&\qquad \vee [J_1 - J + \varepsilon_0] \vee \big[C_2 + C_3 - G(J_1)\big],
\end{aligned}$$

where the cost c^{ε_0} is defined as c with the rejection cost $r(1+\varepsilon_0)$ instead of r. Now, let $\varepsilon_0 \to 0$ first, then $J \to \infty$, recalling that C_2, C_3 and G do not depend on J. Finally let $J_1 \to \infty$, so $G(J_1) \to \infty$, to obtain

$$\begin{aligned}
\limsup V^n\big(T, \tilde{X}^n(0)\big) &\le \limsup \frac{1}{b_n^2} \log \mathbb{E}\big[e^{b_n^2 H_T^n}\big] \\
&\le \sup_{\psi\in\mathcal{P}^2, t\in[0,T]} c\big(x, t, \psi, \alpha_{\beta_0}[\psi]\big) = V(x),
\end{aligned}$$

where in the first inequality we used (87), and in the equality we used the optimality of the β_0-barrier strategy in the game; see Proposition 3.2, as well as Proposition 3.1(i).

We thus turn to the proof of (94). We argue in two steps. In step 1, we show the multidimensional process $\tilde{X}^n$ lies close to the minimizing curve. Consequently, we also deduce that no forced rejections occur, provided n is sufficiently large. In step 2, we deduce (94) from step 1.

Step 1. We show that, for large n,

$$\max_i \big\|\Delta_i^n\big\|_T \le \varepsilon_0, \tag{97}$$

where we denote the difference process

$$\Delta_i^n(t) = \tilde{X}_i^n(t) - \gamma_i^a\big(\check{X}^n(t)\big), \qquad t \in [0, T]. \tag{98}$$

Denote by $\mathcal{G} = \{x \in \mathcal{X} : \theta \cdot x \le a^*, x = \gamma^a(\theta \cdot x)\}$ the set of points lying on the minimizing curve, and recall $\partial^+\mathcal{X} := \{x \in \mathcal{X} : x_i = b_i \text{ for some } i\}$, the set corresponding to the buffer limit boundary. By the choice of a and ε_0 it follows that $\mathcal{G}^{\varepsilon_0}$ and $(\partial^+\mathcal{X})^{\varepsilon_0}$ do not intersect, where for a set $A \in \mathbb{R}^I$ we denote $A^\varepsilon := \{x : \operatorname{dist}(x, A) \le \varepsilon\}$. Forced rejections occur only at times when $\tilde{X}^n$ lies in $(\partial^+\mathcal{X})^{\varepsilon_0}$ (for all n large). As a result, as long as the process $\tilde{X}^n$ lies in $\mathcal{G}^{\varepsilon_0}$, no forced rejections occur. Thus, by showing (97), one also obtains

$$\tilde{R}^n(T) = \tilde{R}_{i^*}^n(T) e_{i^*}. \tag{99}$$

Denote

$$\tau^n = \inf\Big\{t \geq 0 : \max_i |\Delta_i^n(t)| \geq \varepsilon_0\Big\}.$$

LEMMA 5.1. *For all large n, for every $m \in \{1, \dots, M\}$, one has on the event $\Omega^{n,m}$,*

$$\|\tilde{S}^n \circ T^n - \bar{\psi}^{m,2} \circ \boldsymbol{\rho}\|_T \leq 3\delta_1/2, \qquad \|\tilde{A}^n - \bar{\psi}^{m,1}\|_T \leq \delta_1 \tag{100}$$

and

$$\|\check{S}^n \circ T^n - \psi^{m,2} \circ \boldsymbol{\rho}\|_T < 2\delta_1, \qquad \|\check{A} - \psi^{m,1}\|_T < 2\delta_1. \tag{101}$$

Moreover, for all large n and all $t, u \in [0, \tau^n]$ such that $|t-u| < \nu_1$,

$$|\check{X}^n(t) - \check{X}^n(u)| \leq \varepsilon_1. \tag{102}$$

LEMMA 5.2. *For all large n, (97) holds on the event $\bigcup_{m=1}^M \Omega^{n,m}$.*

These two lemmas are proved at the end of the section.

Step 2. As mentioned earlier, for sufficiently small ε_0 one has $a^* = \beta_0$. Moreover, by Lemma 5.2 and the discussion in the beginning of step 1, no forced rejections occur on the event under consideration. Consider the balance equation (15) and recall that $\check{Z}^n$ and $\check{R}^n$ are nonnegative, nondecreasing processes. Recall $\check{Y}^n$ defined in (13). Also, these processes are flat on the set of times where $\check{X}^n > 0$ and $\check{X}^n < a^*$, respectively, where we used work conservation and the absence of forced rejections. Recalling from Section A.1 the characterization of the Skorohod map on an interval, it follows that, on the event under consideration,

$$(\check{X}^n, \check{Z}^n, \check{R}^n)(t) = \Gamma_{[0,\beta_0]}(\check{Y}^n + \check{A}^n - \check{S}^n \circ T^n). \tag{103}$$

Compare this relation with (91). Let n be sufficiently large so that $\|\check{Y}^n - \boldsymbol{y}\|_T \leq \delta_1$. Then by (101),

$$\|(\check{Y}^n + \check{A}^n - \check{S}^n \circ T^n) - (\boldsymbol{y} + \psi^{m,1} - \psi^{m,2})\|_T \leq 5\delta_1.$$

From the above, using (88), (91) and Lemma A.1 it follows that on the event $\Omega^{n.m}$ one has

$$\|\check{R}^n(T) - \varrho^m(T)\| \leq \varepsilon_0 \quad \text{and} \quad \|\check{X}^n - \varphi^m\|_T \leq \operatorname{osc}_D(\varepsilon_0, h). \tag{104}$$

Now we bound H_T^n. Let $L = \sum_i h_i$. For sufficiently large n,

$$\begin{aligned}\int_0^T \bar{h} \cdot \tilde{X}^n(u)\, du &\leq \int_0^T \bar{h} \cdot \gamma^a(\check{X}^n(u))\, du + 2L\varepsilon_0 T \\ &= \int_0^T h^a(\check{X}^n(u))\, du + 2L\varepsilon_0 T\end{aligned}$$

$$\le \int_0^T h\big(\check{X}^n(u)\big)\,du + (2L+1)\varepsilon_0 T + T\omega_1(\varepsilon_0)$$
$$\le \int_0^T h\big(\varphi^m(u)\big)\,du + (2L+2)T\varepsilon_0 + T\omega_1(\varepsilon_0),$$

where the first inequality follows by Lemma 5.2, the equality follows by the definitions of h^a and γ^a, the second inequality follows by (83), and the last inequality follows by (104). Also notice that for sufficiently large n,

$$\bar{r}\cdot\tilde{R}^n = r_{i^*}\tilde{R}^n_{i^*} = \frac{r}{\mu_{i^*}}\tilde{R}^n_{i^*} \le r(1+\varepsilon_0)\theta^n_{i^*}\tilde{R}^n_{i^*}$$
$$= r(1+\varepsilon_0)\theta^n\cdot\tilde{R}^n = r(1+\varepsilon_0)\check{R}^n,$$

where the first and third equalities follow from (99), the second equality follows since $r = r_{i^*}\mu_{i^*}$, the inequality follows since $\theta^n \to \theta$, and the last equality follows by the definition of $\check{R}^n$. Denote $\omega_2(\varepsilon_0) = ((2L+2)T + r(1+\varepsilon_0))\varepsilon_0 + T\omega_1(\varepsilon_0)$. Then from (83) one has $\omega_2(0+) = 0$. Moreover,

$$\begin{aligned}
&\mathbb{E}\big[e^{b_n^2[H_T^n\wedge J_1]}1_{\Omega^{n,m}}\big]\\
&\quad\le \mathbb{E}\big[e^{b_n^2[\int_0^T \bar{h}\cdot\tilde{X}^n(u)\,du+\bar{r}\cdot\tilde{R}^n(T)]}1_{\Omega^{n,m}}\big]\\
&\quad\le \mathbb{E}\big[e^{b_n^2[\int_0^T h(\varphi^m(u))\,du+r(1+\varepsilon_0)\varrho^m(T)+\omega_2(\varepsilon_0)]}1_{\Omega^{n,m}}\big]\\
&\quad\le \max_{0\le t\le T} e^{b_n^2[\int_0^t h(\varphi^m(u))\,du+r(1+\varepsilon_0)\varrho^m(t)+\omega_2(\varepsilon_0)]}\mathbb{P}\big[\big(\tilde{A}^n,\tilde{S}^n\big)\in\mathcal{A}^m_t\big].
\end{aligned}\tag{105}$$

Let t^m be such that

$$\begin{aligned}
&\max_{0\le t\le T}\big[e^{b_n^2[\int_0^t h(\varphi^m(u))\,du+r(1+\varepsilon_0)\varrho^m(t)+\omega_2(\varepsilon_0)]}\mathbb{P}\big[\big(\tilde{A}^n,\tilde{S}^n\big)\in\mathcal{A}^m_t\big]\big]\\
&\quad\le e^{b_n^2[\int_0^{t^m} h(\varphi^m(u))\,du+r(1+\varepsilon_0)\varrho^m(t^m)+\omega_2(\varepsilon_0)]}\mathbb{P}\big[\big(\tilde{A}^n,\tilde{S}^n\big)\in\mathcal{A}^m_{t^m}\big]+\varepsilon_0.
\end{aligned}\tag{106}$$

By Proposition 2.1 and (89),

$$\frac{1}{b_n^2}\log\mathbb{P}\big(\big(\tilde{A}^n,\tilde{S}^n\big)\in\overline{\mathcal{A}^m_{t^m}}\big) \le -\inf_{\psi\in\mathcal{A}^m_{t^m}}\mathbb{J}\big(t^m,\psi\big)+\varepsilon_0 \le -\mathbb{J}\big(t^m,\bar{\psi}^m\big)+2\varepsilon_0$$
$$\le -\mathbb{I}\big(t^m,\psi^m\big)+2\varepsilon_0.$$

This, along with (106), shows that the right-hand side of (105) is bounded by

$$e^{b_n^2[\int_0^{t^m} h(\varphi^m(u))\,du+r(1+\varepsilon_0)\varrho^m(t^m)-\mathbb{I}(t^m,\psi^m)+\omega_2(\varepsilon_0)]}+\varepsilon_0.$$

Thus, (94) follows. □

PROOF OF LEMMA 5.1. Recall relation (11) and the fact that $\tilde{X}^n(t)$ remains bounded. This, along with (92) give $\|\tilde{Z}^n\| \le C_4(1+\Lambda^n)$. Denote

$$\Lambda_J = \sup_{\psi\in\mathcal{AC}_J}\big\|\psi^1\big\|_T + \big\|\psi^2\big\|_T < \infty,\tag{107}$$

where the finiteness follows by (45) and the definition of $\mathcal{AC}_J$. Thus, on the event $\bigcup_m \Omega^{n,m}$, one has $\|\tilde{Z}^n\|_T \le \Lambda_J + 2\delta_1$. Recalling the expression (12) for $\tilde{Z}^n$, it follows that $\|\boldsymbol{\rho} - T^n\|_T < \nu_1$ for large n. Therefore, for every $m \in \{1, \dots, M\}$, $\|\bar{\psi}^{m,2} \circ T^n - \bar{\psi}^{m,2} \circ \boldsymbol{\rho}\|_T < \delta_1/2$. Moreover, on the event $\Omega^{n,m}$, $\|\tilde{S}^n \circ T^n - \bar{\psi}^{m,2} \circ T^n\|_T < \delta_1$ and, therefore,

$$\|\tilde{S}^n \circ T^n - \bar{\psi}^{m,2} \circ \boldsymbol{\rho}\|_T \le 3\delta_1/2.$$

Similarly, $\|\tilde{A}^n - \bar{\psi}^{m,1}\|_T \le \delta_1$.

By (100), (90), and since $\theta^n \to \theta$, it follows that on the event $\Omega^{n,m}$, (101) holds.

It remains to prove (102). Fix $0 \le u \le t < \tau^n$ such that $t - u < \nu_1$. By the definition of the time τ^n, an argument as that leading to (103) shows

$$(\check{X}^n, \check{Z}^n, \check{R}^n)(u) = \Gamma_{[0,\beta_0]}[\check{W}^n](u), \qquad u \in [0, \tau^n),$$

where

$$\check{W}^n(u) = \check{X}^n(0) + \check{y}^n u + \check{A}^n(u) - \check{S}^n(T^n(u)) - \check{R}^n(u).$$

If we show that $|\check{W}^n(t) - \check{W}^n(u)| < 11\delta_1$ then the result follows by (88). Using (101) along with (90) and the fact

$$\lim_{n\to\infty} \|\check{Y}^n - \boldsymbol{y}\|_T = 0$$

shows that, for large n,

$$\begin{aligned}|\check{W}^n(t) - \check{W}^n(u)| \le{}& |\check{W}^n(t) - (x + yt + \psi^{m,1}(t) - \psi^{m,2}(t))| \\ &+ |\check{W}^n(u) - (x + yu + \psi^{m,1}(u) - \psi^{m,2}(u))| \\ &+ |\psi^{m,1}(t) - \psi^{m,1}(u)| + |\psi^{m,2}(t) - \psi^{m,2}(u)| < 11\delta_1.\end{aligned}$$

This completes the proof. □

PROOF OF LEMMA 5.2. The structure of the proof borrows ideas from the proof of Lemma 4.1 of [6] (however, the content is different, as [6] addresses weak convergence). We begin with the case where the initial state lies close to the minimizing curve. That is,

$$\max_i |\Delta_i^n(0)| \le \varepsilon_0. \tag{108}$$

At the last step of the proof we relax this assumption.

By reducing to a subsequence of $\{n\}$, one then has that there exists an $m \in \{1, \dots, M\}$ such that on the event $(\tilde{A}^n, \tilde{S}^n) \in \mathcal{A}_T^m$ one has $\tau^n \le T$. Let $j = j^n$ and ξ^n be the corresponding components from the representation (j, ξ) of $\check{X}^n(\tau^n)$ [with $w = \check{X}^n(\tau^n)$]. Fix a positive integer $K = K(\varepsilon_1) = [D/\varepsilon_1]$, where $\varepsilon_1 \le \theta_{\min}\varepsilon_0/8$ as defined right before (88). Consider the covering of $[0, D]$ by the

$K-1$ intervals $\Xi_k = \boldsymbol{B}(k\varepsilon_1, \varepsilon_1)$, $k = 1, 2, \ldots, K-1$, where $\boldsymbol{B}(x, a)$ denotes $[x-a, x+a]$. Let $\tilde{\Xi}_k = \boldsymbol{B}(k\varepsilon_1, 2\varepsilon_1)$. From (102), we obtain

(109) if $\check{X}^n(\tau^n) \in \Xi_k$, then $\check{X}^n(t) \in \tilde{\Xi}^k$ for every $t \in \boldsymbol{T}^n := [((\tau^n - \nu_1) \vee 0), \tau^n]$.

By considering a further subsequence, we may assume that there exists a $k = k(m)$ such that $\check{X}^n(\tau^n) \in \Xi_k$ for all n. The value assigned by the policy to B^n [see (85)] remains fixed as $\check{X}^n$ varies within any of the intervals $(\hat{a}_j, \hat{a}_{j-1})$. Aiming at showing a contradiction for each k, we consider the following four cases:

(i) $\tilde{\Xi}_k \subset (0, a^*)$ and for all j, $\hat{a}_j \notin \tilde{\Xi}_k$.
(ii) $\tilde{\Xi}_k \subset (0, a^*)$ but $\hat{a}_j \in \tilde{\Xi}_k$ for some $j \in \{1, 2, \ldots, I-1\}$.
(iii) $0 \in \tilde{\Xi}_k$.
(iv) $a^* \in \tilde{\Xi}_k$.

(i) $\tilde{\Xi}_k \subset (0, a^*)$ and for all j, $\hat{a}_j \notin \tilde{\Xi}_k$. Then all points x in $\tilde{\Xi}_k$ lead to the same j in the representation (j, ξ) of x given by (82). This $j = j(k)$ depends on k only, and in particular does not vary with n. Fix $i > j$ (except when $j = I$). Note that $\gamma_i^a(\check{X}^n(\tau^n)) = a_i$ (because $i > j$). We show first that for sufficiently large n one has

(110) for every $i > j$, one has $\Delta_i^n(t) < \varepsilon_0$ for all $t \in [0, T]$.

This is done as follows. Assume to the contrary that $\tau^n \le T$ and that $\Delta_i^n(\tau^n) \ge \varepsilon_0$. Then, since the jumps of $\tilde{X}_i^n$ are of size $(b_n\sqrt{n})^{-1}$ it follows that there must exist $\eta^n \in [0, \tau^n]$ with the properties that

(111) $\tilde{X}_i^n(\eta^n) < a_i + \varepsilon_0/2$, $\tilde{X}_i^n(t) > a_i$ for all $t \in [\eta^n, \tau^n]$.

Therefore, during the time interval $[\eta^n, \tau^n]$, i is always a member of $\mathcal{H}^+(\tilde{X}^n)$ and, therefore, by (85)–(86), $B_i^n(t) = \rho_i'(\tilde{X}^n(t)) > \rho_i + C_4$, for some constant $C_4 > 0$. Thus, by (12),

$$\frac{d}{dt}\tilde{Z}_i^n \le -\frac{\mu_i^n}{b_n\sqrt{n}}C_4.$$

Moreover, if we define $\hat{\eta}^n = \eta^n \vee (\tau^n - \nu_1)$ then by (109) we get that for all $t \in [\hat{\eta}^n, \tau^n]$ one has $\check{X}^n(t) \in \tilde{\Xi}_k \subset (0, a^*)$ and therefore no rejections occur. Using these facts in (12), we have

$$\begin{aligned}\tilde{X}_i^n(\tau^n) - \tilde{X}_i^n(\hat{\eta}^n) \le{}& [\tilde{A}_i^n(\tau^n) - \tilde{A}_i^n(\hat{\eta}^n)] - [\tilde{S}_i^n(T_i^n(\tau^n)) - \tilde{S}_i^n(T_i^n(\hat{\eta}^n))] \\ &- \frac{\mu_i^n}{b_n\sqrt{n}}C_4(\tau^n - \hat{\eta}^n).\end{aligned} \tag{112}$$

Fix a sequence $r_n > 0$ with $r_n \to 0$ and $r_n b_n\sqrt{n} \to \infty$. If $\tau^n - \eta^n < r_n$ and n is sufficiently large then $\hat{\eta}^n = \eta^n$, thus by the definitions of τ^n and η^n, one has $\tilde{X}_i^n(\tau^n) - \tilde{X}_i^n(\hat{\eta}^n) \ge \varepsilon_0/2$. As a result,

(113) $[\tilde{A}_i^n(\tau^n) - \tilde{A}_i^n(\eta^n)] - [\tilde{S}_i^n(T_i^n(\tau^n)) - \tilde{S}_i^n(T_i^n(\eta^n))] \ge \varepsilon_0/2$.

From Lemma 5.1 and inequality (90) it follows that the left-hand side of the above is smaller than $6\delta_1$. Altogether, $6\delta_1 > \varepsilon_0/2$, which contradicts the choice of δ_1 [see (88)]. If, on the other hand, $\tau^n - \eta^n \geq r_n$ then by (112),

$$[\tilde{A}_i^n(\tau^n) - \tilde{A}_i^n(\hat{\eta}^n)] - [\tilde{S}_i^n(T_i^n(\tau^n)) - \tilde{S}_i^n(T_i^n(\hat{\eta}^n))] \geq C_4 r_n b_n \sqrt{n}. \tag{114}$$

The left-hand side of the above is bounded from above by $2\Lambda^n$, which is bounded by Lemma 5.1 and the definition of Λ_J; see (107). This contradicts the fact $C_4 r_n b_n \sqrt{n} \to \infty$. Therefore, (110) holds.

Next, fix $i < j$ (provided $j \neq 1$). Then $\gamma_i^a(\check{X}^n(\tau^n)) = 0$ and whenever $\tilde{X}_i^n > 0$, i is a member of the high priority set $\mathcal{H}^+(\tilde{X}^n)$. This is due to the fact that there must exist $l > i$ such that $\tilde{X}_l^n < a_l$; otherwise, the workload would be at least $\sum_{p=l}^I a_p \theta_p^n > \check{X}^n$. Hence, the same argument yields a contradiction. Therefore,

$$\text{for every } i < j, \qquad \Delta_i^n(t) < \varepsilon_0 \text{ for all } t \in [0, T]. \tag{115}$$

Consider now j itself. We will show, for the case $j < I$, that

$$\Delta_j^n(t) < \varepsilon_0 \qquad \text{for all } t \in [0, T]. \tag{116}$$

Suppose that we show that, for every $t \in [0, \tau^n]$ and every large n,

$$\text{if } \Delta_j^n(t) \in (\varepsilon_0/2, \varepsilon_0), \text{ then } j \in \mathcal{H}^+\big(\tilde{X}^n(t)\big). \tag{117}$$

Then, assuming to the contrary that $\tau^n \leq T$ and $\tilde{X}_j^n(\tau^n) \geq \gamma_j^a(\check{X}^n(\tau^n)) + \varepsilon_0$, implies that there exists $\eta^n \in [0, \tau^n]$ with the properties that

$$\begin{aligned} &\tilde{X}_j^n(\eta^n) - \gamma_j^a\big(\check{X}^n(\tau^n)\big) < 3\varepsilon_0/4, \\ &\tilde{X}_j^n(t) - \gamma_j^a\big(\check{X}^n(\tau^n)\big) > \varepsilon_0/2 \qquad \text{for all } t \in \big[\eta^n, \tau^n\big]. \end{aligned}$$

From (117) during the time interval $[\eta^n, \tau^n]$, j is always a member of $\mathcal{H}^+(\tilde{X}_j^n(t))$. Therefore, we still have inequality (112) valid. Arguing separately for the cases $\tau^n - \eta^n < r_n$ and $\tau^n - \eta^n \geq r_n$, leads, in analogy to (113) and (114), to a contradiction. There is a slight difference in the first case. If $\tau^n - \eta^n < r_n$ and n is sufficiently large then $\hat{\eta}^n = \eta^n$, and now $\tilde{X}_j^n(\tau^n) - \tilde{X}_j^n(\hat{\eta}^n) \geq \varepsilon_0/4 + \gamma_j^a(\check{X}^n(\tau^n)) - \gamma_j^a(\check{X}^n(\hat{\eta}^n))$. Recall that $\tau^n - \hat{\eta}^n < \nu_1$. Therefore, by taking δ_1 to be sufficiently small, one can verify that $|\gamma_j^a(\check{X}^n(\tau^n)) - \gamma_j^a(\check{X}^n(\hat{\eta}^n))|$ would be significantly smaller than $\varepsilon_0/2$. This argument follows by the continuity of γ^a and by similar arguments to the ones we used in order to prove (102).

Now we show that (117) holds (except in the case $j = I$). Since $\theta \cdot \gamma^a(\theta \cdot \bar{x}) = \theta \cdot \bar{x}$ for all $\bar{x} \in \mathcal{X}$, $\theta^n \to \theta$, γ^a is uniformly continuous, and $\mathcal{X}$ is bounded, we have

$$q_n := \sup_{\bar{x} \in \mathcal{X}} \big|\theta \cdot \gamma^a\big(\theta^n \cdot \bar{x}\big) - \theta \cdot \bar{x}\big| \to 0 \qquad \text{as } n \to \infty. \tag{118}$$

Note by (118) that $|\theta \cdot \tilde{X}^n(t) - \theta \cdot \gamma^a(\check{X}^n(t))| \leq q_n \to 0$.

Fix $t \in [0, \tau^n]$. If $\Delta_j^n(t) \geq \varepsilon_0/2$, then

$$\begin{aligned} -\theta_j \varepsilon_0/2 &\geq \sum_{i \neq j} \theta_i \big(\tilde{X}_i^n(t) - \gamma_i^a\big(\check{X}^n(t)\big)\big) - \|\theta\| q_n \\ &\geq \sum_{i > j} \theta_i \big(\tilde{X}_i^n(t) - a_i\big) - \|\theta\| q_n, \end{aligned} \tag{119}$$

where we used $\gamma_i^a(\check{X}^n(t)) = 0$ for $i < j$ and $\gamma_i^a(\check{X}^n(t)) = a_i$ for $i > j$. These two equations hold from (109). For all large n, this implies $\tilde{X}_i^n(t) < a_i$ for at least one $i > j$, by which $j \in \mathcal{H}^+(\tilde{X}^n)$.

We can now show that for every $t \in [0, T]$ one has $\max_{i \leq I} |\Delta_i^n(t)| < \varepsilon_0$. Indeed, in the case $j = I$, we have by (115), $\max_{i<I} |\Delta_i^n(t)| < \varepsilon_0$. By (118), $|\theta \cdot \Delta^n(t)| \leq q_n$. Since $\theta \in (0, \infty)^I$ and $q_n \to 0$, we obtain

$$\max_{i \leq I} \big|\Delta_i^n(t)\big| < \varepsilon_0. \tag{120}$$

In the case $j < I$, combining (110), (115), (116), we have $\max_{i \leq I} \Delta_i^n(t) < \varepsilon_0$. Using again the fact $|\theta \cdot \Delta^n(\tau^n)| \leq q_n \to 0$ shows that (120) is valid in this case as well.

(ii) $\tilde{\Xi}_k \subset (0, a^*)$ but $\hat{a}_j \in \tilde{\Xi}_k$ for some $j \in \{1, 2, \ldots, I-1\}$. Let $(j^n(t), \xi^n(t))$ denote the representation (82) for $\check{X}^n(t)$. Note that in the time window $\boldsymbol{T}^n$, j^n varies between two values, namely j and $j+1$, and so it is no longer true that $\gamma_{j+1}^a(\check{X}^n(t)) = a_{j+1}$ on that time interval. However, it is true that

$$\gamma_{j+1}^a\big(\check{X}^n(t)\big) \geq a_{j+1} - 4\varepsilon_1/\theta_{\min} \geq a_{j+1} - \varepsilon_0/2, \qquad t \in \boldsymbol{T}^n, \tag{121}$$

where the second inequality follows since $\varepsilon_1 < \theta_{\min}\varepsilon_0/8$. Indeed, we have for any $w \in \tilde{\Xi}_k$, $|w - \hat{a}_j| \leq 4\varepsilon_1$, since $\hat{a}_j$ is also in $\tilde{\Xi}_k$. Now, if $w \geq \hat{a}_j$ then $\gamma_{j+1}^a(w) = a_{j+1}$. Otherwise,

$$w = \hat{a}_{j+1} + \theta_{j+1}\xi = \hat{a}_j - \theta_{j+1}a_{j+1} + \theta_{j+1}\xi,$$

hence $|a_{j+1} - \xi| \leq 4\theta_{j+1}^{-1}\varepsilon_1$ and (121) follows.

By the same arguments as in case (i), we get that

$$\text{for every } i \neq j+1, \qquad \text{one has } \Delta_i^n(t) < \varepsilon_0 \text{ for all } t \in [0, T].$$

As for $i = j+1$, assume to the contrary that $\tau^n \leq T$ and that $\Delta_{j+1}^n(\tau^n) \geq \varepsilon_0$. Then by (121), we get that $\tilde{X}_{j+1}^n(\tau^n) \geq a_{j+1} + \varepsilon_0/2$ and the same arguments as in case (i) are valid.

Combining all the estimates gives

$$\max_{i \leq I} \Delta_i^n(t) < \varepsilon_0, \qquad t \in [0, T].$$

Along with the fact $|\theta \cdot \Delta^n(\tau^n)| \leq q_n$, this gives

$$\max_{i \leq I} \big|\Delta_i^n(t)\big| < \varepsilon_0, \qquad t \in [0, T].$$

(iii) $0 \in \tilde{\Xi}_k$. This differs from case (i) in that during $\boldsymbol{T}^n$, $\tilde{X}^n$ may hit zero and, therefore B^n might vanish. Note however that the way case (i) is handled, one focuses only on time intervals where $\tilde{X}^n \neq 0$ and, therefore, the proof is valid here as well. We thus have $\max_{i \leq I} |\Delta_i^n(t)| < \varepsilon_0$, $t \in [0, T]$.

(iv) $a^* \in \tilde{\Xi}_k$. In this case, $\theta \cdot \tilde{X}^n$ may exceed the threshold a^*, and rejections may occur. The argument provided in case (i) is then slightly changed. A negative term is added to the right-hand side of (112), but the consequences of (112) remain valid with this addition. [Note that for small ε_0, $\hat{a}_i \neq a^*$ holds for all i, hence assuming ε_0 is small, we do not need to check case (ii) here.]

Having shown that $\max_{i \leq I} |\Delta_i^n(t)| < \varepsilon_0, t \in [0, T]$ in all cases completes the proof of the lemma under (108).

The relaxation of (108) is performed by showing that within a short time t, $\max_i |\Delta_i^n(t)| \leq \varepsilon_0$. This is sufficient, because on the remaining time interval the argument provided above for the case (108) gives the result.

Fix $\delta > 0$. We will show that for each i and all sufficiently large n, there exists $t \in [0, \delta]$ such that $|\Delta_i^n(t)| \leq \varepsilon_0$. Since the proof provided above for the case (108) treats each i separately, this will assure that once $|\Delta_i^n|$ is bounded by ε_0 for some i, it remains so for the remaining time interval.

We thus fix i and prove that for all sufficiently large n, there exists $t \in [0, \delta]$ such that $|\Delta_i^n(t)| \leq \varepsilon_0$. Assume to the contrary that $|\Delta_i^n| > \varepsilon_0$ on $[0, \delta]$. Since the jumps of Δ^n are of order $n^{-1/2}$, we either have $\Delta_i^n > \varepsilon_0$ on $[0, \delta]$, or $\Delta_i^n < -\varepsilon_0$ on $[0, \delta]$, provided n is large. In the former case, i is always a member of $\mathcal{H}^+(\tilde{X}^n(t))$, for every t and, therefore by (85)–(86), $B_i^n(t) = \rho_i'(\tilde{X}^n(t)) > \rho_i + C_4$, for some constant $C_4 > 0$. Thus, by (12),

$$\frac{d}{dt}\tilde{Z}_i^n \leq -\frac{\mu_i^n}{b_n\sqrt{n}}C_4.$$

Therefore, as was argued in (112), by (12) we obtain that

$$\begin{aligned}\frac{\mu_i^n}{b_n\sqrt{n}}C_4\delta &\leq -\big[\tilde{X}_i^n(\delta) - \tilde{X}_i^n(0)\big] + \big[\tilde{A}_i^n(\delta) - \tilde{A}_i^n(0)\big] \\ &\quad - \big[\tilde{S}_i^n\big(T_i^n(\delta)\big) - \tilde{S}_i^n\big(T_i^n(0)\big)\big] - \big[\tilde{R}_i^n(\delta) - \tilde{R}_i^n(0)\big] \\ &\leq 2D_i + 2\Lambda^n.\end{aligned}$$

The right-hand side of the above bounded from above, as follows from Lemma 5.1 and the fact that $\Lambda_J < \infty$ [see (107)]. This contradicts the fact that the left-hand side goes to infinity.

In case that $\Delta_i^n < -\varepsilon_0$ on $[0, \delta]$, i is not a member of $\mathcal{H}^+(\tilde{X}^n(t))$ for any t and, therefore, by (85)–(86), $B_i^n(t) = 0$. Thus, by (12),

$$\frac{d}{dt}\tilde{Z}_i^n = \frac{\mu_i^n}{b_n\sqrt{n}}\rho_i.$$

Hence,

$$\frac{\mu_i^n}{b_n\sqrt{n}}\rho_i\delta = \big[\tilde{X}_i^n(\delta) - \tilde{X}_i^n(0)\big] - \big[\tilde{A}_i^n(\delta) - \tilde{A}_i^n(0)\big] + \big[\tilde{S}_i^n\big(T_i^n(\delta)\big) - \tilde{S}_i^n\big(T_i^n(0)\big)\big] + \big[\tilde{R}_i^n(\delta) - \tilde{R}_i^n(0)\big].$$

The right-hand side of the above is bounded, using similar considerations along with (92), whereas again, the left-hand side tends to infinity with n. $\square$

APPENDIX

A.1. The Skorohod map on a finite interval. Given a and b, $a < b$, the *Skohorod map* on the interval $[a, b]$ maps $\mathcal{D}([0,\infty),\mathbb{R})$ to $\mathcal{D}([0,\infty),\mathbb{R}^3)$. It is denoted by $\Gamma_{[a,b]}$, and is characterized as the solution map $\omega \to (\varphi, \eta_1, \eta_2)$ to the problem of finding, for a given ω, a triplet $(\varphi, \eta_1, \eta_2)$, such that

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

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

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

By writing $\eta_i(0-) = 0$, we adopt the convention that $\eta_i(0) > 0$ is regarded a jump at zero. This convention, in conjunction with $\int_{[0,\infty)} 1_{(a,b]}(\varphi)\, d\eta_1 = 0$ [resp., $\int_{[0,\infty)} 1_{[a,b)}(\varphi)\, d\eta_2 = 0$], means that if $\omega(0) < a$ [resp., $\omega(0) > b$] then $\varphi(0) = a$ (resp., b). If, however, $\omega(0) \in [a, b]$ then $\varphi(0) = \omega(0)$, and η_i have no jump at zero. See [19] for existence and uniqueness of solutions, and continuity and further properties of the map. In particular, we have the following.

LEMMA A.1. *Fix $b > 0$. Then there exists a constant C such that for every $T > 0$, $\delta > 0$ and $\omega, \tilde{\omega} \in \mathcal{D}([0,\infty),\mathbb{R})$,*

$$\big\|\Gamma_{[0,b]}(\omega) - \Gamma_{[0,b]}(\tilde{\omega})\big\|_T \le C\|\omega - \tilde{\omega}\|_T$$

and

$$\operatorname{osc}_T\big(\delta, \Gamma_{[0,b]}(\omega)\big) \le C \operatorname{osc}(\delta, \omega).$$

A.2. On the rate functions $\mathbb{I}$ and $\mathbb{J}$. For every $T \in \mathbb{R}_+$ and $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$ set $\tilde{\mathbb{I}}(T, \psi) = \tilde{\mathbb{I}}_1(T, \psi^1) + \tilde{\mathbb{I}}_2(T, \psi^2)$,

$$\tilde{\mathbb{I}}_1(T, \psi^1) = \inf\{\mathbb{J}_1(T, \bar{\psi}^1) : \bar{\psi}^1 \in \mathcal{P}^I, \theta \cdot \bar{\psi}^1 = \psi^1\} \tag{122}$$

and

$$\tilde{\mathbb{I}}_2(T, \psi^2) = \inf\{\mathbb{J}_2(T, \bar{\psi}^2) : \bar{\psi}^2 \in \mathcal{P}^I, \theta \cdot (\bar{\psi}^2 \circ \boldsymbol{\rho}) = \psi^2\}. \tag{123}$$

Recall that $\mathbb{I}$ is defined in (17) and $\mathbb{J}$ in (10).

LEMMA A.2. *For every* $\psi = (\psi^1, \psi^2) \in \mathcal{P}^2$, *there exists* $\bar{\psi} = (\bar{\psi}^1, \bar{\psi}^2) \in \mathcal{P}^I \times \mathcal{P}^I$ *such that*

$$\big(\theta \cdot \bar{\psi}^1, \theta \cdot (\bar{\psi}^2 \circ \boldsymbol{\rho})\big) = \big(\psi^1, \psi^2\big) \tag{124}$$

and

$$\tilde{\mathbb{I}}_k\big(T, \psi^k\big) = \mathbb{I}_k\big(T, \psi^k\big) = \mathbb{J}_k\big(T, \bar{\psi}^k\big), \qquad k = 1, 2, T > 0. \tag{125}$$

PROOF. Define $\mathbb{K} : [0, \infty) \times \mathcal{P}^I \to [0, \infty]$ by

$$\mathbb{K}(T, \bar{\psi}) = \begin{cases} \displaystyle\sum_{i=1}^{I} \alpha_i \int_0^T \dot{\psi}_i^2(u)\, du & \text{if all } \psi_i \in \mathcal{AC}_0\big([0, T], \mathbb{R}\big), \\ \infty & \text{otherwise,} \end{cases}$$

where $\alpha_1, \ldots, \alpha_I > 0$. Then both $\mathbb{J}_1$ and $\mathbb{J}_2$ are of the form $\mathbb{K}$. Set $l_1, \ldots, l_I \in (0, 1]$. Define $\mathbb{L} : [0, \infty) \times \mathcal{P} \to [0, \infty]$ by

$$\mathbb{L}(T, \psi) = \inf\big\{\mathbb{K}(T, \bar{\psi}) : \bar{\psi} \in \mathcal{P}^I, \theta \cdot \bar{\psi}(l) = \psi\big\}, \tag{126}$$

where $\bar{\psi}(l) = (\psi_1(l_1 t), \ldots, \psi_I(l_I t))$. Note that in case $l_1 = \cdots = l_I = 1$ (resp., $l_1 + \cdots + l_I = 1$) the function $\mathbb{L}$ gives $\tilde{\mathbb{I}}_1$ (resp., $\tilde{\mathbb{I}}_2$). We show that for every $\psi \in \mathcal{P}$ there is $\bar{\psi}$ such that

$$\theta \cdot \bar{\psi}(l) = \psi \tag{127}$$

and

$$\mathbb{L}(T, \psi) = \mathbb{K}(T, \bar{\psi}). \tag{128}$$

We now calculate $\mathbb{L}$:

$$\begin{aligned} &\inf\big\{\mathbb{K}(T, \bar{\psi}) : \bar{\psi} \in \mathcal{P}^I, \theta \cdot \bar{\psi}(l) = \psi\big\} \\ &\quad = \inf\left\{\sum_{i=1}^{I} \alpha_i \int_0^T \dot{\psi}_i^2(u)\, du : \sum_{i=1}^{I} \theta_i \psi_i(l_i u) = \psi(u), u \in [0, T]\right\} \\ &\quad = \inf\left\{\sum_{i=1}^{I} \alpha_i \int_0^{l_i T} \dot{\psi}_i^2(u)\, du : \sum_{i=1}^{I} \theta_i \psi_i(l_i u) = \psi(u), u \in [0, T]\right\} \\ &\quad = \inf\left\{\sum_{i=1}^{I} \alpha_i l_i \int_0^T \dot{\psi}_i^2(l_i t)\, dt : \sum_{i=1}^{I} \theta_i \psi_i(l_i u) = \psi(u), u \in [0, T]\right\} \\ &\quad = \inf\left\{\sum_{i=1}^{I} \frac{\alpha_i}{l_i} \int_0^T \left(\frac{d}{dt}\psi_i(l_i t)\right)^2 dt : \sum_{i=1}^{I} \theta_i \psi_i(l_i u) = \psi(u), u \in [0, T]\right\}. \end{aligned}$$

Above, the second equality follows since for every $i \in \mathcal{I}$ the constraint on ψ_i exists only on the time interval $[0, l_i T]$. Hence, we are free to choose ψ_i on the interval

$(l_i T, T]$. By simple calculation, the infimum on the right-hand side is attained by $\bar{\psi} = (\psi_1, \dots, \psi_I)$,

$$\psi_i(u) = \begin{cases} \dfrac{\theta_i l_i}{\alpha_i} \left(\dfrac{\theta_1^2 l_1}{\alpha_1} + \cdots + \dfrac{\theta_I^2 l_I}{\alpha_I} \right)^{-1} \psi(u/l_i), & 0 \le u \le l_i T, \\ \psi_i(l_i T), & u > l_i T. \end{cases}$$

By substitution,

$$\mathbb{K}(T, \bar{\psi}) = \left(\frac{\theta_1^2 l_1}{\alpha_1} + \cdots + \frac{\theta_I^2 l_I}{\alpha_I} \right)^{-1} \int_0^T \dot{\psi}^2(t)\, dt.$$

By the taking proper l_i's and α_i's for each of the cases $\mathbb{I}_1$ and $\mathbb{I}_2$, one obtains (18). This completes the proof. □

Acknowledgments. The authors thank an Associate Editor and a referee for their helpful comments that considerably helped improve the exposition.

REFERENCES

[1] ATAR, R. and BISWAS, A. (2014). Control of the multiclass $G/G/1$ queue in the moderate deviation regime. *Ann. Appl. Probab.* **24** 2033–2069. MR3226171

[2] ATAR, R. and COHEN, A. (2016). A differential game for a multiclass queueing model in the moderate-deviation heavy-traffic regime. *Math. Oper. Res.* **41** 1354–1380.

[3] ATAR, R., DUPUIS, P. and SHWARTZ, A. (2003). An escape-time criterion for queueing networks: Asymptotic risk-sensitive control via differential games. *Math. Oper. Res.* **28** 801–835. MR2015913

[4] ATAR, R., MANDELBAUM, A. and REIMAN, M. I. (2004). Scheduling a multi class queue with many exponential servers: Asymptotic optimality in heavy traffic. *Ann. Appl. Probab.* **14** 1084–1134. MR2071417

[5] ATAR, R. and MENDELSON, G. (2016). On the non-Markovian multiclass queue under risk-sensitive cost. *Queueing Syst.* **84** 265–278. MR3568183

[6] ATAR, R. and SHIFRIN, M. (2014). An asymptotic optimality result for the multiclass queue with finite buffers in heavy traffic. *Stoch. Syst.* **4** 556–603. MR3353227

[7] BUDHIRAJA, A. and GHOSH, A. P. (2012). Controlled stochastic networks in heavy traffic: Convergence of value functions. *Ann. Appl. Probab.* **22** 734–791. MR2953568

[8] CHEN, H. and MANDELBAUM, A. (1991). Leontief systems, RBVs and RBMs. In *Applied Stochastic Analysis* (*London*, 1989). *Stochastics Monogr.* **5** 1–43. Gordon and Breach, New York.

[9] DEMBO, A. and ZEITOUNI, O. (1998). *Large Deviations Techniques and Applications*, 2nd ed. *Applications of Mathematics* (*New York*) **38**. Springer, New York. MR1619036

[10] DUPUIS, P. and MCENEANEY, W. M. (1997). Risk-sensitive and robust escape criteria. *SIAM J. Control Optim.* **35** 2021–2049. MR1478651

[11] FLEMING, W. H. (1971). Stochastic control for small noise intensities. *SIAM J. Control* **9** 473–517. MR0304045

[12] FLEMING, W. H. (2006). Risk sensitive stochastic control and differential games. *Commun. Inf. Syst.* **6** 161–177.

[13] FLEMING, W. H. and MCENEANEY, W. M. (1995). Risk-sensitive control on an infinite time horizon. *SIAM J. Control Optim.* **33** 1881–1915. MR1358100

[14] FLEMING, W. H. and METE SONER, H. (2006). *Controlled Markov Processes and Viscosity Solutions*, 2nd ed. *Stochastic Modelling and Applied Probability* **25**. Springer, New York.
[15] FLEMING, W. H. and SOUGANIDIS, P. E. (1986). PDE-viscosity solution approach to some problems of large deviations. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (4) **13** 171–192. MR0876121
[16] GANESH, A., O'CONNELL, N. and WISCHIK, D. (2004). *Big Queues*. *Lecture Notes in Math.* **1838**. Springer, Berlin. MR2045489
[17] HARRISON, J. M. and LÓPEZ, M. J. (1999). Heavy traffic resource pooling in parallel-server systems. *Queueing Syst. Theory Appl.* **33** 339–368. MR1742575
[18] JACOBSON, D. H. (1973). Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Trans. Automat. Control* **AC-18** 124–131. MR0441523
[19] KRUK, L., LEHOCZKY, J., RAMANAN, K. and SHREVE, S. (2007). An explicit formula for the Skorokhod map on $[0, a]$. *Ann. Probab.* **35** 1740–1768. MR2349573
[20] MAJEWSKI, K. (2006). Sample path large deviations for multiclass feedforward queueing networks in critical loading. *Ann. Appl. Probab.* **16** 1893–1924. MR2288708
[21] PUHALSKII, A. A. (1999). Moderate deviations for queues in critical loading. *Queueing Syst. Theory Appl.* **31** 359–392.
[22] PUHALSKII, A. A. and WHITT, W. (1997). Functional large deviation principles for first-passage-time processes. *Ann. Appl. Probab.* **7** 362–381. MR1442318
[23] WARD, A. R. and KUMAR, S. (2008). Asymptotically optimal admission control of a queue with impatient customers. *Math. Oper. Res.* **33** 167–202. MR2393546
[24] WHITTLE, P. (1990). A risk-sensitive maximum principle. *Systems Control Lett.* **15** 183–192. MR1076968

DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION—ISRAEL INSTITUTE OF TECHNOLOGY
HAIFA 32000
ISRAEL

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN 48109
USA
E-MAIL: shloshim@gmail.com
URL: https://sites.google.com/site/asafcohentau/