Electron. Commun. Probab. **19** (2014), no. 11, 1–13.
DOI: 10.1214/ECP.v19-2905
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

On the risk-sensitive cost for a Markovian multiclass queue with priority *

Rami Atar† Anindya Goswami‡ Adam Shwartz§

Abstract

A multi-class M/M/1 system, with service rate $\mu_i n$ for class-i customers, is considered with the risk-sensitive cost criterion $n^{-1} \log E \exp \sum_i c_i X_i^n(T)$, where $c_i > 0$, $T > 0$ are constants, and $X_i^n(t)$ denotes the class-i queue-length at time t, assuming the system starts empty. An asymptotic upper bound (as $n \to \infty$) on the performance under a fixed priority policy is attained, implying that the policy is asymptotically optimal when c_i are sufficiently large. The analysis is based on the study of an underlying differential game.

Keywords: Multi-class M/M/1; risk-sensitive control; large deviations; differential games.
AMS MSC 2010: 60F10; 60K25; 49N70; 93E20.
Submitted to ECP on July 2, 2013, final version accepted on February 21, 2014.

1 Introduction

A Markovian queueing model consisting of a single server capable of serving jobs of k classes is considered. Job arrival rates are proportional to a (large) parameter n, and so are the processing rates for each of the class-i jobs, that, specifically, are given by $\mu_i n$. Let $X_i^n(t)$ denote the number of jobs in the ith class at time t, assuming the system starts empty at time 0, and consider the scaled version $\bar{X}^n = n^{-1} X^n$. Under specific service policies, for example, serve-the-longest-queue and certain priority policies, it is well known that $\{\bar{X}^n, n \in \mathbb{N}\}$ satisfy a sample-path large deviation principle [7]. In this note we are interested in the dynamic control problem where a service policy is sought to minimize a cost at the large deviation scale. In particular, we consider the cost

$$\frac{1}{n} \log \mathbb{E} \exp\{c \cdot X^n(T)\} = \frac{1}{n} \log \mathbb{E} \exp\{n G(\bar{X}^n)\}, \tag{1.1}$$

where $T > 0$ and $c \in (0, \infty)^k$ are fixed, and we denote $G(\xi) = c \cdot \xi(T)$ for $\xi : [0, T] \to \mathbb{R}^k$. The motivation for considering such a cost, referred to in the literature as *risk-sensitive*, for a queueing model, is that it strongly emphasizes large values of terminal queue length, and is thus natural when one seeks to prevent buffer overflow. Avoiding large waiting times so as to assure quality of service is a closely related motivation (though not directly addressed in this paper).

*Support: US-Israel BSF Grant 2008466 ; ISF Grant 1349/08 ; Technion's fund for promotion of research.
†Dep. of E. Engineering, Technion–Israel Institute of Technology, Israel. E-mail: atar@ee.technion.ac.il
‡IISER Pune, Pune 411021, India. E-mail: anindya@iiserpune.ac.in
§Julius M. and Bernice Naiman chair in Engineering, Department of Electrical Engineering, Technion–Israel Institute of Technology, Israel. E-mail: adam@ee.technion.ac.il

In an earlier paper [1] we considered a broader setting, of a model with multiple, heterogenous servers, of which the above is a special case, and a risk-sensitive cost defined similarly to (1.1), with a more general functional G of the whole path $\{\bar{X}^n(t), 0 \le t \le T\}$. The limit of the optimal cost, as $n \to \infty$, was characterized as the value of a certain two-player zero sum *differential game* (DG). In this paper a particular priority-type strategy for the DG is studied. It is shown that at for sufficiently large c_i this strategy is optimal for the DG, and that an analogous policy for the queueing control problem is asymptotically optimal. We further show that in a more general setup, the worst performance of that priority type strategy has a specific upper bound which is also obeyed by the asymptotic performance of the induced policy.

The strategy alluded to above is one that prioritizes the classes in the order of the index $(1 - e^{-c_i})\mu_i$, with highest priority given to the class with highest index. This is reminiscent of the $c\mu$ rule, where priority is given according to the index $c_i\mu_i$, known to be optimal under *linear* queue-length cost with weights c_i: note in particular that if we scale all c_i's by the same small parameter ε, then the exponential priority rule agrees with the linear one for all sufficiently small ε. This result is useful in practical implementation because the priority based resource allocation policy is simple as well as robust (note, in particular, that it is independent of the arrival rates). The proof builds on results in our earlier paper [1] and on the general large deviation upper bound of Dupuis, Ellis and Weiss [5]. In particular, the main argument consists of comparing the priority policy's performance, estimated using the results of [5], with the DG value using the connection established in [1].

The paper is organized as follows. In Section 2 we present the queueing model and the main result. Section 3 describes the connection between the control problem and the DG, obtained in [1]. An estimate of performance of DG is also obtained. In Section 4 that estimate is used to analyze the priority policy. Section 5 gives a lower bound on the DG's value, by which optimality of the priority rule for large c_i follows. The appendix establishes the existence of a strategy for the DG that acts according to the priority discipline.

2 Model and main result

The model is parameterized by $n \in \mathbb{N}$. It consists of k customer classes and one server. Arrivals into the system occur according to independent Poisson processes, with respective parameters $n\lambda_i$, where $\lambda_i > 0$ are fixed. Arriving jobs are queued in buffers, one dedicated to each class. The server is available to serve the customers at the head of the k lines, and is capable of splitting its effort among them. The service times are exponential, where a class-i customer is served at rate $n\mu_i$ if the server dedicates all its effort to it. An *allocation vector*, representing the fraction of effort dedicated to each of the classes, is any member of $U = \{u \in \mathbb{R}_+^k \mid \sum_{i\in\mathcal{K}} u_i \le 1\}$, where $\mathcal{K} = \{1, 2, \dots, k\}$. Denote e_i as the n-tuple with 1 at ith place and 0's elsewhere. For $n \in \mathbb{N}$ denote $S_n = n^{-1}\mathbb{Z}_+^k$. Given $n \in \mathbb{N}$ and $u \in U$ consider the operator (a generalization of Q-matrix of finite state case)

$$\mathcal{L}^{n,u} f(x) = \sum_{i\in\mathcal{K}} n\lambda_i (f(x + \frac{1}{n}e_i) - f(x)) + \sum_{i\in\mathcal{K}} n\mu_i u_i (f(x - \frac{1}{n}e_i) - f(x)) 1_{\{x_i \ge n^{-1}\}}, \qquad x \in S_n, \tag{2.1}$$

for $f : S_n \to \mathbb{R}$. A *control system* consists of a triplet $\mathcal{U}^n = (U^n, \bar{X}^n, (\mathcal{F}_t)_{t\in[0,T]})$, defined on a given complete probability space $(\Omega, \mathcal{F}, P)$, where U^n and $\bar{X}^n$ are processes taking values in U and S_n, having RCLL sample paths, $\mathcal{F}_t \subset \mathcal{F}$ forms a filtration to which these processes are adapted, with probability one, $\bar{X}^n(0) = 0$ and $U_i^n(t) = 0$ whenever $\bar{X}_i^n(t) = 0$, and finally, for every bounded $f : S_n \to \mathbb{R}$, the process $f(\bar{X}^n(t)) - \int_0^t \mathcal{L}^{n,U^n(s)} f(\bar{X}^n(s))ds$, $t \in [0,T]$, is a martingale w.r.t. $(\mathcal{F}_t)$. We refer to U^n and $\bar{X}^n$ as the control and controlled process, respectively. For $n \in \mathbb{N}$, the cost functional associated with a control system $\mathcal{U}$ is given by

$$C^{n,\mathcal{U}} = \frac{1}{n} \log \mathbb{E}[e^{ng(\bar{X}^{n,\mathcal{U}}(T))}] \tag{2.2}$$

where $g(x) = c \cdot x, c \in (0,\infty)^k$ and $T > 0$. The value of the control problem is given by $V^n = \inf C^{n,\mathcal{U}}$, where the infimum ranges over all control systems. It is known from [1] that the limit $V_{\lim} = \lim_{n\to\infty} V^n$ exists (see Theorem 3.1 below for more details).

We also consider a special class of control systems. Given n, a *stationary feedback control* is any mapping $\mathcal{U} : S_n \to U$ such that

$$\mathcal{U}_i(x) = 0 \text{ whenever } x_i = 0, \qquad i \in \mathcal{K}. \tag{2.3}$$

The corresponding *controlled process* is the Markov process $\bar{X}^{n,\mathcal{U}}$ on S_n, starting from zero, with infinitesimal generator $\mathcal{L}^n_{\mathcal{U}}$ given by

$$\mathcal{L}^n_{\mathcal{U}} f(x) = \mathcal{L}^{n,\mathcal{U}(x)} f(x). \tag{2.4}$$

In the queueing model, $n\bar{X}^{n,\mathcal{U}}(t)$ represents the vector of queue lengths at time t when allocation is performed according to the feedback control $\mathcal{U}$. With an abuse of notation, $\mathcal{U}$ is both a generic symbol for a control system and for a stationary feedback control. This will cause no confusion.

We will be interested in the stationary feedback control that prioritizes classes according to the index $\hat{\mu}_i = \mu_i(1 - e^{-c_i})$. Denote $\hat{\lambda}_i = \lambda_i(e^{c_i} - 1)$ and $W = \min_{u\in U} \sum_i (\hat{\lambda}_i - u_i\hat{\mu}_i)^+$. Assume throughout that the class labels are ordered so that

$$\hat{\mu}_1 \geq \hat{\mu}_2 \geq \cdots \geq \hat{\mu}_k. \tag{2.5}$$

For $n \in \mathbb{N}$, this control, denoted by $\mathcal{U}^* = \mathcal{U}^{*,n}$, is given by

$$\mathcal{U}^*_i(x) = 1_{\{x_i>0\}} \prod_{j<i} 1_{\{x_j=0\}}, \qquad i \in \mathcal{K}, x \in S_n, \tag{2.6}$$

where the product is defined as 1 when $i = 1$. Our main result is as follows.

Theorem 2.1. *The cost under the feedback controls of priority type, given by* (2.6)*, obeys the following bounds*

i.

$$V_{\lim} \leq \liminf_{n\to\infty} C^{n,\mathcal{U}^*} \leq \limsup_{n\to\infty} C^{n,\mathcal{U}^*} \leq WT. \tag{2.7}$$

ii. If $e^{c_i} \geq \frac{\mu_i}{\lambda_i}$ *for all* i *then* $V_{\lim} = WT$*. Consequently,* $\mathcal{U}^*$ *is asymptotically optimal in the sense that* $\lim_{n\to\infty} C^{n,\mathcal{U}^*} = V_{\lim}$*.*

3 Differential game setup

The limit on the l.h.s. of (2.7) can be characterized as the value of a DG, formulated as follows. Let $M = \mathbb{R}^k_+ \times \mathbb{R}^k_+$ and write generic members of M as $m = ((\bar{\lambda}_i)_{i\in\mathcal{K}}, (\bar{\mu}_i)_{i\in\mathcal{K}})$. While λ and μ denote the actual arrival and service parameters for the system, a possibly different member $m = (\bar{\lambda}, \bar{\mu})$ of M will be interpreted as a perturbed set of parameters. Due to the exponential nature of the cost functional it is natural to expect this additional control which reposes on the Laplace's principle [3]. For $u \in U$ and $m \in M$, let

$$v(u,m) = \sum_i \bar{\lambda}_i e_i - \sum_i u_i \bar{\mu}_i e_i, \qquad \rho(u,m) = \sum_i \lambda_i \, \omega\Big(\frac{\bar{\lambda}_i}{\lambda_i}\Big) + \sum_i u_i \mu_i \, \omega\Big(\frac{\bar{\mu}_i}{\mu_i}\Big), \tag{3.1}$$

where

$$\omega(r) = \begin{cases} r\log r - r + 1, & r \geq 0, \\ \\ +\infty, & r < 0, \end{cases}$$

with the convention $0\log 0 = 0$. Let $\bar{U} = \{\bar{u} : [0,T] \to U \mid \bar{u} \text{ is measurable}\}$ be the set of admissible dynamic allocations. Define the set of admissible dynamic perturbations $\bar{M} = \{m : [0,T] \to M \mid m \text{ is measurable}, \omega \circ m \text{ is locally integrable}\}$. Endow $\bar{U}$ and $\bar{M}$ with the metric $d(v_1, v_2) = \int_0^T \|v_1(t) - v_2(t)\| dt$, and with the corresponding Borel σ-fields. A mapping $\alpha : \bar{M} \to \bar{U}$ is called a *strategy* if it is measurable and if for every $m, \tilde{m} \in \bar{M}$ and $t \in [0,T]$,

$$m(r) = \tilde{m}(r) \quad \text{for a.e. } r \in [0,t] \quad \text{implies} \quad \alpha[m](r) = \alpha[\tilde{m}](r) \quad \text{for a.e. } r \in [0,t].$$

The set of all strategies is denoted by A.

Let Γ_1, the one-dimensional Skorohod map from $C([0,T] : \mathbb{R})$ to itself, be defined as

$$\Gamma_1[\psi](t) = \psi(t) - \inf_{r\in[0,t]} \psi(r) \wedge 0, \qquad t \in [0,T],$$

and let Γ, mapping $C([0,T] : \mathbb{R}^k)$ to itself, be given by $\Gamma[\psi]_i = \Gamma_1[\psi_i]$, $i \in \mathcal{K}$. The cost C associated with $u \in \bar{U}$ and $m \in \bar{M}$ is given by

$$C(u,m) = g(\varphi(T)) - \int_0^T \rho(u(r), m(r))dr, \qquad \varphi = \Gamma[\psi], \qquad \psi = \int_0^{\cdot} v(u(r), m(r))dr. \tag{3.2}$$

Thus ρ, heuristically, constitutes the cost of changing the measure and is incurred to player 2. Let

$$V = \inf_{\alpha\in A} \sup_{m\in\bar{M}} C(\alpha[m], m). \tag{3.3}$$

It is established in Theorem 2.1 and Proposition 3.3 of [1] that

Theorem 3.1. $\lim_{n\to\infty} V^n = V$. *Thus* $V_{\lim} = V$.

Now we consider a strategy α^* that prioritizes according to the indices $\hat{\mu}_i$, as in (2.5). More precisely, let α^* be the strategy that sends $m = ((\bar{\lambda}_i(t))_{i\in\mathcal{K}}, (\bar{\mu}_i(t))_{i\in\mathcal{K}})_{t\in[0,T]} \in \bar{M}$ to $u \in \bar{U}$, where, for $t \in [0,T]$, denoting $a_i(t) = \bar{\lambda}_i(t)/\bar{\mu}_i(t)$, one has

$$u_1(t) = \begin{cases} 1 & \text{if} \quad \varphi_1(t) > 0, \\ 1 \wedge a_1(t) & \text{if} \quad \varphi_1(t) = 0, \end{cases} \tag{3.4}$$

$$u_i(t) = \begin{cases} 1 - \sum_{j=1}^{i-1} u_j(t) & \text{if} \quad \varphi_i(t) > 0, \\ \left(1 - \sum_{j=1}^{i-1} u_j(t)\right) \wedge a_i(t) & \text{if} \quad \varphi_i(t) = 0, \end{cases} \qquad i \geq 2. \tag{3.5}$$

These relations give rise to a unique, well-defined strategy as proved in the appendix. We denote the performance of α^* by $V^* = \sup_{m\in\bar{M}} C(\alpha^*[m], m)$.

Proposition 3.2. *One has*

$$V^* = \sup_{m\in\bar{M}} \left(g\left(\int_0^T v(\alpha^*[m](t), m(t))dt \right) - \int_0^T \rho(\alpha^*[m](t), m(t))dt \right) \leq WT. \tag{3.6}$$

Since for every n, $C^{n,\mathcal{U}^*} \geq V^n$, it follows from Theorem 3.1 that $\liminf_{n\to\infty} C^{n,\mathcal{U}^*} \geq V$. Thus in view of Proposition 3.2, to prove Theorem 2.1(i), it suffices to show, as we do in the next section, that

$$\limsup_{n\to\infty} C^{n,\mathcal{U}^*} \leq V^*. \tag{3.7}$$

Toward proving Proposition 3.2, let us introduce some notation. Let $\omega_i(y) = \lambda_i\omega(y/\lambda_i)$ and $\tilde\omega_i(y) = \mu_i\omega(y/\mu_i)$. Let

$$C_i(u,m) = -\omega_i(\bar\lambda_i) - u_i\tilde\omega_i(\bar\mu_i) + c_i(\bar\lambda_i - u_i\bar\mu_i), \qquad u \in U,\, m \in M. \tag{3.8}$$

Given $r \geq 0$, let

$$W(r) = \min\Big\{\sum_{i=1}^k \Big(\hat\lambda_i - v_i\hat\mu_i\Big)^+ : v_i \geq 0, \sum_{i=1}^k v_i \leq r\Big\}. \tag{3.9}$$

Note that, with $\hat\rho_i = \hat\lambda_i/\hat\mu_i$, the following v is a minimizer in (3.9)

$$v_1^* = r \wedge \hat\rho_1, \qquad v_i^* = \Big(r - \sum_{m=1}^{i-1} v_m^*\Big) \wedge \hat\rho_i, \qquad i \geq 2.$$

Lemma 3.3. *Given $r \geq 0$ and $m = (\bar\lambda_i, \bar\mu_i) \in M$, one has $\sum_{i=1}^k C_i(u,m) \leq W(r)$, provided that*

$$u_1 \in \{r,\, r \wedge \bar\rho_1\}, \tag{3.10}$$

$$u_i \in \{r - u_{1,i-1},\, (r - u_{1,i-1}) \wedge \bar\rho_i\}, \qquad i \geq 2, \tag{3.11}$$

where $\bar\rho_i = \bar\lambda_i/\bar\mu_i$ (here, $r \wedge (y/0)$ is interpreted as r) and $u_{1,j} = \sum_1^j u_i$.

Before presenting the proof of the lemma, we show that the proposition follows.

Proof of Proposition 3.2. The fact that a strategy α^* exists, as well as that under this strategy one has $\psi_i(s) \geq 0$ for all s, is proved in Proposition A.1 in the appendix. Fix an arbitrary $m \in \bar M$ and set $u = \alpha^*[m]$. To prove the proposition it suffices to show that $C(u,m) \leq WT$. Since $\psi_i(s) \geq 0$ for all s, we have $\varphi(T) = \psi(T)$. Thus $C(u,m)$ is given by

$$C(u,m) = \int_0^T \sum_i C_i(u(t), m(t))dt.$$

By (3.4) and (3.5), for each t, $u(t)$ satisfies the hypotheses of Lemma 3.3, with data $m(t)$ and $r = 1$. Hence $C(u,m) \leq WT$, which completes the proof. □

Proof of Lemma 3.3. The claim is proved by induction on k. The precise statement proved by induction involves an *arbitrary* set of parameters λ_i, μ_i, c_i. Namely, given k and r, and *any* $3k$-tuple of positive numbers λ_i, μ_i, c_i, for which the parameters $\hat\mu_i = \mu_i(1 - e^{-c_i})$ are ordered as in (2.5), the statement of the lemma is valid.

Consider first $k = 1$. We will show

$$C_1(u,m) \leq \begin{cases} \hat\lambda_1 - r\hat\mu_1 & \text{if} \quad u_1 = r, \\ 0 & \text{if} \quad u_1 = \bar\rho_1. \end{cases} \tag{3.12}$$

First, the inequalities

$$-\omega_i(\bar\lambda_i) + c_i\bar\lambda_i \leq \hat\lambda_i, \quad -\tilde\omega_i(\bar\mu_i) - c_i\bar\mu_i \leq -\hat\mu_i \tag{3.13}$$

hold for every $\bar\lambda_i, \bar\mu_i$, as can be verified in the following way. By direct calculation, the concave functions on the left hand sides have maxima at $\bar\lambda_i = \lambda_i e^{c_i}$ and $\bar\mu_i = \mu_i e^{-c_i}$ respectively. Thus, their maximum values can be computed and those are $\lambda_i(e^{c_i} - 1)$ and $\mu_i(e^{-c_i} - 1)$ which are the same as $\hat\lambda$ and $-\hat\mu$ respectively. By (3.8), this gives the first line in (3.12). If $u_1 = \bar{\rho}_1$ then the last term in (3.8) is zero, hence $C_1(u,m) \le 0$. This shows (3.12), from which it follows that $C_1(u,m) \le W(r)$ in case $k=1$.

Next, assuming that the claim holds for a given k, we show that it holds for $k+1$. Let then r and m be given, and let u be as in (3.10)–(3.11). Denote $C_{a,b} = \sum_{i=a}^{b} C_i(u,m)$. Also, let $W_{a,b}(r)$ be defined as in (3.9), where the sums range from a to b. The induction assumption implies

$$C_{2,k+1} \le W_{2,k+1}(r-u_1). \tag{3.14}$$

Case 1: $u_1 < v_1^*$. Then by (3.10), $u_1 = \bar{\rho}_1$. As a result, arguing as in the induction base, $C_1(u,m) \le 0$. Thus $C_{1,k+1} \le C_{2,k+1}$. Hence by the induction assumption, $C_{1,k+1} \le W_{2,k+1}(r-u_1)$. Clearly $W(r)$ is decreasing with r. Hence

$$C_{1,k+1} \le W_{2,k+1}(r-v_1^*) \le (\hat{\lambda}_1 - v_1^*\hat{\mu}_1) + W_{2,k+1}(r-v_1^*) = W_{1,k+1}(r).$$

Case 2: $\delta := u_1 - v_1^* \ge 0$. Using again (3.13), $C_{1,1} \le \hat{\lambda}_1 - u_1\hat{\mu}_1$. Hence by (3.14),

$$C_{1,k+1} \le \hat{\lambda}_1 - u_1\hat{\mu}_1 + W_{2,k+1}(r-u_1).$$

By definition of W, it is not hard to see that $|W(r_1) - W(r_2)| \le |r_1 - r_2|\hat{\mu}_{\max}$, where $\hat{\mu}_{\max}$ is the largest parameter $\hat{\mu}_i$ involved. Thus, recalling $\hat{\mu}_2 \ge \cdots \ge \hat{\mu}_k$,

$$|W_{2,k+1}(r_1) - W_{2,k+1}(r_2)| \le |r_1 - r_2|\,\hat{\mu}_2, \qquad r_1, r_2 \ge 0.$$

As a result,

$$\begin{aligned}
C_{1,k+1} &\le \hat{\lambda}_1 - u_1\hat{\mu}_1 + W_{2,k+1}(r-u_1) \\
&= \hat{\lambda}_1 - v_1^*\hat{\mu}_1 + W_{2,k+1}(r-v_1^*) - \delta\hat{\mu}_1 + W_{2,k+1}(r-u_1) - W_{2,k+1}(r-v_1^*) \\
&\le \hat{\lambda}_1 - v_1^*\hat{\mu}_1 + W_{2,k+1}(r-v_1^*) - \delta\hat{\mu}_1 + \delta\hat{\mu}_2 \\
&\le \hat{\lambda}_1 - v_1^*\hat{\mu}_1 + W_{2,k+1}(r-v_1^*) = W_{1,k+1}(r).
\end{aligned}$$

We have thus shown that $C_{1,k+1} \le W_{1,k+1}(r)$ and completed the argument. □

4 Priority-based feedback controls

In this section we prove (3.7), based on the general large deviation upper bound of [5]. We begin by analyzing a wider class of stationary feedback controls (which, in this section we call controls, for short), and then specialize to $\mathcal{U}^*$. Recall that, given n, a control is defined as a map from S_n to U. In this section we will consider sequences $\mathcal{U}^n$ of controls that are all obtained from a single map $\mathcal{U} : \mathbb{R}_+^k \to U$ by way of restricting $\mathcal{U}$ to S_n, for each n. Given n, there will be no confusion in referring to $\mathcal{U}$ itself as the control, and we shall do so.

For $x \in \mathbb{R}_+^k$ let $\mathbf{I}(x) = \{i \in \mathcal{K} : x_i = 0\}$. Note that it is an empty set in the interior of $\mathbb{R}_+^k$ and $\mathcal{K}$ at the origin. $\mathbf{I}$ partitions $\mathbb{R}_+^k$ into sets that we will call *facets*. Let also $\bar{\mathbf{I}}(x) = 2^{\mathbf{I}(x)}$ be collection of all subsets of $\mathbf{I}(x)$. The class of controls $\mathcal{U} : \mathbb{R}_+^k \to U$ that we analyze consists of those that satisfy (2.3) and, in addition, take integer values and are constant on facets. That is,

$$\begin{aligned}
&\mathcal{U}_i(x) \in \{0,1\} && \text{for every } i \in \mathcal{K}, x \in \mathbb{R}_+^k, \\
&\mathcal{U}(x) = \mathcal{U}(y) && \text{for every } x, y \in \mathbb{R}_+^k \text{ whenever } \mathbf{I}(x) = \mathbf{I}(y).
\end{aligned} \tag{4.1}$$

Under (4.1) $\mathcal{U}$ induces a map from $2^{\mathcal{K}}$ to U, given by $J \subset \mathcal{K} \mapsto \mathcal{U}(x)$ for some x such that $J = \mathbf{I}(x)$. For the ease of notation, we identify facets (i.e., subsets of $\mathbb{R}_+^k$ on which $\mathbf{I}$ is constant) with collections of indices (the corresponding value of $\mathbf{I}$); moreover, we refer this map ($\mathcal{U} \circ \mathbf{I}^{-1}$) by the same symbol $\mathcal{U}$ throughout this section. We follow this convention for other functions whose dependence on x is via $\mathcal{U}$ only. For a given $\mathcal{U}$ as in (4.1), we define the following quantities for each $x \in \mathbb{R}^k_+$ and $p, q \in \mathbb{R}^k$ by

$$\begin{aligned} H(x,p) &= \sum_{i\in\mathcal{K}} \lambda_i(e^{p_i}-1) + \sum_{i\in\mathcal{K}} \mu_i \mathcal{U}_i(x)(e^{-p_i}-1) \\ h(x,p) &= \max_{J\in\bar{\mathbf{I}}(x)} H(J,p) \\ L(x,q) &= \sup_{p\in\mathbb{R}^k} [p\cdot q - H(x,p)] \\ l(x,q) &= \sup_{p\in\mathbb{R}^k} [p\cdot q - h(x,p)] \\ \mathcal{A} &= \{\varphi : [0,T] \to \mathbb{R}^k_+ \text{ absolutely continuous} \mid \varphi(0)=0\} \\ I(\varphi) = I^{\mathcal{U}}(\varphi) &= \begin{cases} \int_0^T l(\varphi(t),\dot{\varphi}(t))dt, & \text{if } \varphi \in \mathcal{A}; \\ +\infty, & \text{else.} \end{cases} \end{aligned}$$

Here h is the upper semi continuous regularization of H, whereas L and l are the Legendre- Fenchel transforms of H and h respectively. Exclusively for this section we consistently use φ to denote a generic element of $\mathcal{A}$ (this notation will be convenient when used in relation (4.10) below). Since, the maps H, h, L and l depend on x via $\mathcal{U}$ only, they are constants on each facets provided the other variable is fixed. Thus the naturally induced maps $H(J,p), h(J,p), L(J,q)$ and $l(J,q)$ are well defined.

Proposition 4.1. *Given a control $\mathcal{U}$ satisfying all assumptions stated in this section (in particular,* (4.1)*), the corresponding sequence $\mathcal{U}^n$ satisfies*

$$\limsup_{n\to\infty} C^{n,\mathcal{U}^n} \le A^{\mathcal{U}} := \sup_{\varphi\in\mathcal{A}} (g(\varphi(T)) - I^{\mathcal{U}}(\varphi)). \tag{4.2}$$

Proof. By Theorem 1.1 of [5], the sequence $\bar{X}^n$ of controlled processes associated with the controls $\mathcal{U}^n$, that are merely Markov processes with infinitesimal generators $\mathcal{L}^n_{\mathcal{U}^n}$ (2.4), satisfies a large deviation upper bound in $\mathbb{D}([0,T] : \mathbb{R}^k)$ with the good rate function I (see [5], [3] for this terminology; in particular, $\mathbb{D}$ is the space of RCLL functions with the Skorohod topology). The upper bound in Varadhan's lemma (Lemma 4.3.6 of [3]) can therefore be used. It is easy to verify the moment condition $\limsup n^{-1} \log \mathbb{E} e^{\gamma n g(\bar{X}^n(T))} < \infty$ (for some $\gamma > 1$) required for that lemma, by noting that $ng(\bar{X}^n(T))$ is stochastically bounded by a r.v. $\alpha\,\mathrm{Poisson}(\beta n)$ for some constants α, β. As a result,

$$\limsup_{n\to\infty} C^{n,\mathcal{U}^n} = \limsup \frac{1}{n} \log \mathbb{E}[e^{ng(\bar{X}^n)}] \le \sup_{\varphi\in\mathcal{A}} (g(\varphi(T)) - I^{\mathcal{U}}(\varphi)).$$

□

Proposition 4.2. *One has $A^{\mathcal{U}^*} \le V^*$.*

Proof of Theorem 2.1(i). By Proposition 3.2 and the discussion following it, the result follows from Propositions 4.1 and 4.2. □

In the rest of this section we prove Proposition 4.2. Given $x \in \mathbb{R}^k_+$, $q \in \mathbb{R}^k$ denote

$$\begin{aligned} \Xi &= \{\xi : 2^{\mathcal{K}} \to [0,1] \mid \sum_{J\in 2^{\mathcal{K}}} \xi^J = 1\}, \\ \Xi(x) &= \{\xi \in \Xi \mid \xi^J = 0 \ \forall J \notin \bar{\mathbf{I}}(x)\}, \\ \mathcal{S}_{(x,q)} &= \{(m,\xi) \in M \times \Xi(x) \mid \bar{\lambda}_i - \sum_{J\in\bar{\mathbf{I}}(x)} \xi^J \mathcal{U}_i(J)\bar{\mu}_i = q_i \forall i\}. \end{aligned}$$

The collection Ξ consists of all possible normalized weights on the collection of all facets. If x belongs to a particular facet J, then $\Xi(x)$ includes those members of Ξ which assign nonzero weights only to the facets whose closure include J. $\mathcal{S}_{(x,q)}$ is the collection of pairs of rates and weights such that the speed due to resulting weighted service allocation match with q.

Lemma 4.3. *Under* (4.1)*, for* $x \in \mathbb{R}_+^k$*,* $q \in \mathbb{R}^k$*,* $l(x,q) = \inf_{\mathcal{S}_{(x,q)}} \rho\big(\sum_{J \in 2^{\mathcal{K}}} \xi^J \mathcal{U}(J), m\big)$*, where* ρ *is as in* (3.1)*.*

Proof. First, by the definition of L and H and using Lemma A.2 in the appendix,

$$\begin{aligned} L(x,q) &= \sum_i \sup_{p_i \in \mathbb{R}} [p_i q_i - \big(\lambda_i(e^{p_i}-1) + \mu_i \mathcal{U}_i(x)(e^{-p_i}-1)\big)] \\ &= \sum_i \inf_{\bar\lambda_i, \bar\mu_i} \Big\{ \lambda_i\, \omega\Big(\frac{\bar\lambda_i}{\lambda_i}\Big) + \sum_i \mathcal{U}_i(x)\mu_i\, \omega\Big(\frac{\bar\mu_i}{\mu_i}\Big) \Big| \bar\lambda_i - \bar\mu_i \mathcal{U}_i(x) = q_i \Big\} \\ &= \inf_{\bar\lambda, \bar\mu} \Big\{ \sum_i \lambda_i\, \omega\Big(\frac{\bar\lambda_i}{\lambda_i}\Big) + \sum_i \mathcal{U}_i(x)\mu_i\, \omega\Big(\frac{\bar\mu_i}{\mu_i}\Big) \Big| \bar\lambda_i - \bar\mu_i \mathcal{U}_i(x) = q_i \Big\} \\ &= \inf_{\bar\lambda, \bar\mu} \big\{ \rho(\mathcal{U}(x), (\bar\lambda, \bar\mu)) | \bar\lambda_i - \bar\mu_i \mathcal{U}_i(x) = q_i \ \forall i \big\}, \end{aligned}$$

where the second equality follows by directly solving both optimization problems. We use the following representation of l, from Theorem 3.1 of [5]:

$$l(x,q) = \inf_{(q^J, \xi^J)} \left\{ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J L(J, q^J) \ \Big| \ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J q^J = q,\, \xi \in \Xi(x) \right\}, \tag{4.3}$$

where the infimum ranges over all maps $J \mapsto (q^J, \xi^J)$. Using the expression of L above,

$$\begin{aligned} l(x,q) &= \inf_{(q^J,\xi^J)} \left\{ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \inf_m \{ \rho(\mathcal{U}(J), m) \mid \bar\lambda_i - \bar\mu_i \mathcal{U}_i(J) = q_i^J \ \forall i \} \ \Big| \ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J q^J = q, \xi \in \Xi(x) \right\} \\ &= \inf_{(\xi^J)} \inf_{(m^J)} \left\{ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \rho(\mathcal{U}(J), m^J) \ \Big| \ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J (\bar\lambda_i^J - \bar\mu_i^J \mathcal{U}_i(J)) = q_i \ \forall i,\, \xi \in \Xi(x) \right\}. \end{aligned}$$

Hence, by restricting the minimizing set for variable (m^J), $l(x,q)$ is bounded above by

$$\inf_{\xi \in \Xi(x)} \inf_{m \in M} \left\{ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \rho(\mathcal{U}(J), m) \ \Big| \ \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J (\bar\lambda_i - \bar\mu_i \mathcal{U}_i(J)) = q_i \ \forall i \right\} = \inf_{\mathcal{S}_{(x,q)}} \rho \left(\sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \mathcal{U}(J), m \right). \tag{4.4}$$

In order to prove the lemma, it remains to show that $l(x,q)$ is also bounded below by the above quantity. Given $x, (\xi^J)$ and (m^J), define, with $0/0 = 0$, $\bar\lambda[x] = \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \bar\lambda^J$, $v_i(x) = \sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \mathcal{U}_i(J)$,

$$\bar\mu_i[x] = \frac{\sum_{J \in \bar{\mathbf{I}}(x)} \xi^J \mathcal{U}_i(J) \bar\mu_i^J}{v_i(x)} \quad \text{and} \quad m[x] = (\bar\lambda[x], \bar\mu[x]).$$

Then

$$\sum_{J \in \bar{\mathbf{I}}(x)} \xi^J (\bar\lambda_i^J - \bar\mu_i^J \mathcal{U}_i(J)) = \bar\lambda_i[x] - v_i(x)\bar\mu_i[x] \qquad i \in \mathcal{K}. \tag{4.5}$$

Since ω is convex, we have by changing the order of summation on the l.h.s. below and using Jensen's inequality,

$$\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\left(\sum_{i\in\mathcal{K}}\lambda_i\,\omega\Big(\frac{\bar\lambda_i^J}{\lambda_i}\Big)+\sum_{i\in\mathcal{K}}\mathcal{U}_i(J)\mu_i\,\omega\Big(\frac{\bar\mu_i^J}{\mu_i}\Big)\right)\geq\sum_{i\in\mathcal{K}}\lambda_i\,\omega\Big(\frac{\bar\lambda_i[x]}{\lambda_i}\Big)+\sum_{i\in\mathcal{K}}v_i(x)\mu_i\,\omega\Big(\frac{\bar\mu_i[x]}{\mu_i}\Big).$$

Thus $\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\rho\left(\mathcal{U}(J),m^J\right)\geq\rho\left(\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\mathcal{U}(J),m[x]\right)$. Hence given $x\in\mathbb{R}_+^k, q\in\mathbb{R}^k$ using (4.5),

$$\begin{aligned}
l(x,q)&=\inf_{(\xi^J,m^J)}\left\{\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\rho(\mathcal{U}(J),m^J)\;\Big|\;\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J(\bar\lambda_i^J-\bar\mu_i^J\mathcal{U}_i(J))=q_i\;\forall i,\;\;\xi\in\Xi(x)\right\}\\
&\geq\inf_{\xi\in\Xi(x)}\inf_{(m^J)}\left\{\rho\left(\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\mathcal{U}(J),m[x]\right)\;\Big|\bar\lambda_i[x]-\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\mathcal{U}_i(J)\bar\mu_i[x]=q_i\;\forall i\right\}\\
&\geq\inf_{\xi\in\Xi(x)}\inf_{m\in M}\left\{\rho\left(\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\mathcal{U}(J),m\right)\;\Big|\bar\lambda_i-\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\mathcal{U}_i(J)\bar\mu_i=q_i\;\forall i\right\}\\
&=\inf_{\mathcal{S}_{(x,q)}}\rho\left(\sum_{J\in\bar{\mathbf{I}}(x)}\xi^J\mathcal{U}(J),m\right).
\end{aligned}$$

Thus the result follows from the above and (4.4). □

Proof of Proposition 4.2. Consider the following system of equations, for $(m,\xi):[0,T]\to M\times\Xi$ measurable and $\varphi\in\mathcal{A}$,

$$\begin{cases}\bar\lambda_i(t)-\sum_{J\in2^{\mathcal{K}}}\xi^J(t)\mathcal{U}_i(J)\bar\mu_i(t)=\dot\varphi_i(t) & i\in\mathcal{K},\\ \xi(t)\in\Xi(\varphi(t)),\end{cases}\qquad\text{a.e. } t\in[0,T]. \tag{4.6}$$

For $\varphi\in\mathcal{A}$, and $\mathcal{U}$ as in (4.1) denote

$$\mathcal{S}_\varphi=\{(m,\xi):[0,T]\to M\times\Xi\text{ measurable }\;|\,(4.6)\text{ holds}\},\text{ and}$$
$$\mathcal{S}^*_\varphi=\mathcal{S}_\varphi\text{ with }\;\mathcal{U}=\mathcal{U}^*.$$

By a standard argument based on a measurable selection result such as [6], one can show

$$\int_0^T\inf_{\mathcal{S}_{(\varphi(t),\dot\varphi(t))}}\rho\left(\sum_{J\in2^{\mathcal{K}}}\xi^J\mathcal{U}(J),m\right)dt=\inf_{\mathcal{S}_\varphi}\int_0^T\rho\left(\sum_{J\in2^{\mathcal{K}}}\xi^J(t)\mathcal{U}(J),m(t)\right)dt.$$

Thus using Lemma 4.3,

$$I^{\mathcal{U}}(\varphi)=\inf_{\mathcal{S}_\varphi}\int_0^T\rho\left(\sum_{J\in2^{\mathcal{K}}}\xi^J(t)\mathcal{U}(J),m(t)\right)dt. \tag{4.7}$$

The inequality to be proved is $A^{\mathcal{U}^*}\leq V^*$. Using the expression (4.2) for $A^{\mathcal{U}}$, and (4.7), it will follow if we show the inequality

$$\begin{aligned}
&\sup_{\varphi\in\mathcal{A}}\sup_{(m,\xi)\in\mathcal{S}^*_\varphi}\Big(g(\varphi(T))-\int_0^T\rho(u^*(t),m(t))dt\Big)\\
&\leq\sup_{m\in\bar M}\left(g\left(\int_0^T v(\alpha^*[m](t),m(t))dt\right)-\int_0^T\rho(\alpha^*[m](t),m(t))dt\right)\equiv V^*,
\end{aligned}\tag{4.8}$$

where, for given φ and $(m,\xi)\in\mathcal{S}^*_\varphi$,

$$u^*(t)=u^*[\varphi,m,\xi](t)=\sum_{J\in 2^{\mathcal{K}}}\xi^J(t)\mathcal{U}^*(J). \tag{4.9}$$

Note that, in turn, the above will follow once we prove the statement

$$\text{if } \varphi\in\mathcal{A},\ (\xi,m)\in\mathcal{S}^*_\varphi \text{ then } u^*[\varphi,m,\xi](t)=\alpha^*[m](t) \text{ for a.e. } t. \tag{4.10}$$

Indeed, if (4.10) holds then $\varphi=\int_0^\cdot v(\alpha^*[m](t),m(t))dt$, and thus, for every $\varphi\in\mathcal{A}$, $(\xi,m)\in\mathcal{S}^*_\varphi$,

$$g(\varphi(T))-\int_0^T\rho(u^*(t),m(t))dt=g\left(\int_0^T v(\alpha^*[m](t),m(t))dt\right)-\int_0^T\rho(\alpha^*[m](t),m(t))dt\le V^*,$$

which implies (4.8).

It thus remains to prove (4.10). We do this by arguing that if $\varphi\in\mathcal{A}$, $(\xi,m)\in\mathcal{S}^*_\varphi$ then $u^*[t]$ satisfies (3.4), (3.5) for a.e. t. Recall that $\mathcal{U}^*(x)$ is defined, for $x\in S_n$, in (2.6). For facets J, $\mathcal{U}^*(J)$ is defined via the association of x with a facet to which it belongs. We can write this as

$$\mathcal{U}^*_i(J)=1_{\{i\notin J\}}\prod_{j<i}1_{\{j\in J\}},\qquad i\in\mathcal{K}, s\subset\mathcal{K}. \tag{4.11}$$

Consider the case $\varphi_1=\varphi_1(t)>0$. In this case, by the definition of $\mathbf{I}$, $1\notin\mathbf{I}(\varphi)$. Consequently, 1 is not a member of any subset of $\mathbf{I}(\varphi)$, namely it is not a member of any $J\in\bar{\mathbf{I}}(\varphi)$. Since (4.6) holds, $\xi\in\Xi(\varphi)$ (where $\xi=\xi(t)$, $\varphi=\varphi(t)$, and this is valid for a.e. t). Thus by definition of Ξ, ξ charges only facets $J\in\mathbf{I}(\varphi)$. In particular, it charges only facets J with $1\notin J$. By (4.9) and (4.11), it follows that $u^*_1=u^*_1(t)=1$. This shows that the first line of (3.4) is valid for a.e. t.

Next consider the case that, for some fixed $i>1$, $\varphi_i(t)>0$, so as to verify the first line of (3.5). In this case $i\notin J$ for all $J\in\bar{\mathbf{I}}(\varphi)$, and ξ is supported on such facets. Now,

$$\sum_{j=1}^i u^*_j=\sum_{j=1}^i\sum_J\xi^J\mathcal{U}^*_j(J)=\sum_J\xi^J\sum_{j=1}^i 1_{\{j\notin J\}}\prod_{r<j}1_{\{r\in J\}}.$$

For J in the support of ξ we have $j^J:=\min\{j:j\notin J\}\le i$. As a result, for $j\in\{1,\dots,i\}$,

$$1_{\{j\notin J\}}\prod_{r<j}1_{\{r\in J\}}=\begin{cases}0, & \text{if } j\ne j^J,\\ 1, & \text{if } j=j^J.\end{cases}$$

This shows $\sum_{j=1}^i u^*_i=\sum_J\xi^J=1$ and verifies the first line of (3.5).

For each i let $A_i=\{t\in[0,T]\mid\varphi_i(t)=0\}$ and $\dot{A}_i=\{t\in[0,T]\mid\varphi_i(t)=0,\dot\varphi_i(t)=0\}$. Then by Theorem A.6.3 of [4], $|A_i\setminus\dot{A}_i|=0$. Therefore, for a.e. $t\in[0,T]$, $\varphi_i(t)=0$ implies $\dot\varphi_i(t)=0$, thus by (4.6), $u^*_i(t)=\bar\lambda_i(t)/\bar\mu_i(t)$. Since we always have $\sum_i u^*_i\le 1$, it is also true that, for a.e. t, $\varphi_i(t)=0$ implies

$$u^*_i(t)=\Big(1-\sum_{j=1}^{i-1}u^*_j(t)\Big)\wedge\frac{\bar\lambda_i(t)}{\bar\mu_i(t)},$$

which verifies the second line of (3.4) and that of (3.5). This establishes (4.10), thus (4.8), and we conclude that $A^{\mathcal{U}^*}\le V^*$. □

5 A lower bound on V

To prove Theorem 2.1(ii), it remains to show that when $e^{c_i} \geq \frac{\mu_i}{\lambda_i}$ for all i, one has $V_{\lim} = WT$. Since we already have that $V = V_{\lim} \leq WT$, it remains to show that $V \geq WT$ in this case.

Proof of Theorem 2.1(ii). Note first that $\hat{\lambda}_i \geq \hat{\mu}_i$ for each i. Therefore, $W = \min_{u\in U}\sum_i(\hat{\lambda}_i - u_i\hat{\mu}_i) = \sum_i \hat{\lambda}_i - \max_i \hat{\mu}_i$. From (3.2) we deduce that

$$\varphi_i(T) = \sup_{r\in[0,T]}[\psi_i(T) - \psi_i(r)] = \sup_{r\in[0,T]}\int_r^T v_i(u(r'), m(r'))dr'.$$

Using this in (3.3) and interchanging the order of suprema gives

$$V = \inf_{\alpha\in A}\sup_{(r_i)_i}\sup_{m\in\bar{M}}\Big[-\sum_i\int_0^T \omega_i(\bar{\lambda}_i(r))dr - \sum_i\int_0^T u_i(r)\tilde{\omega}_i(\bar{\mu}_i(r))dr + \sum_i c_i\int_{r_i}^T v_i(u(r), m(r))dr\Big], \tag{5.1}$$

where the outer supremum ranges over $\bar{r} = (r_i)_i \in [0,T]^k$. Now, bound V below by replacing the supremum over $\bar{r}\in[0,T]^k$ by taking $\bar{r} = 0$. Further, replace the supremum over all $m \in \bar{M}$ by the following particular choice of m, $(\bar{\lambda}_i(r), \bar{\mu}_i(r)) = (\lambda_i^*, \mu_i^*) = (\lambda_i e^{c_i}, \mu_i e^{-c_i})$ $r \in [0,T]$, $i \in \mathcal{I}$. Then

$$V \geq \inf_{\alpha\in A}\sum_i\int_0^T\Big(-\omega_i(\lambda_i^*) - u_i(r)\tilde{\omega}_i(\mu_i^*) + c_i v_i(u(r), m)\Big)dr = \inf_{\alpha\in A}\sum_i\Big(\hat{\lambda}_i - \bar{u}_i\hat{\mu}_i\Big)T, \tag{5.2}$$

where $\bar{u}_i = T^{-1}\int_0^T u_i(r)dr$ and, as before, $u(\cdot) = \alpha[m]$. Since $\bar{u}$ is always a member of U, the above expression is equal to WT. This shows $V \geq WT$ and completes the proof of Theorem 2.1(ii). □

A Appendix

We argue that the relations (3.4)–(3.5) give rise to a well-defined strategy.

Proposition A.1. *There exists a strategy $\alpha^* \in A$ with the following properties. Given $m \in \bar{M}$, let $u = \alpha^*[m]$ and let ψ and φ be given by* (3.2)*. Then (ψ, φ, u) satisfy the relations* (3.4)–(3.5)*. Moreover, $\psi_i(t) \geq 0$ for all t and i, hence $\varphi = \psi$.*

Proof. We will use the following fact regarding Γ_1. Recall that if $q : \mathbb{R}_+ \to \mathbb{R}$ and $p = \Gamma_1[q]$ then $p = q + d$ where $d(t) = -\inf_{t'\leq t} q(t') \wedge 0$. In case that q is absolutely continuous and $q(0) \geq 0$, the term d is given by

$$d(t) = \int_0^t \Big(\frac{dq}{dt'}\Big)^- 1_{\{p(t')=0\}}dt'. \tag{A.1}$$

The above is an immediate consequence of a general fact that solutions p of the Skorohod problem with absolutely continuous data q solve ODE of the form $\dot{p} = \pi(p, \dot{q})$, where $\pi(x, v)$ is a certain projection map, which in the one-dimensional case is given by

$$\pi(x, v) = v1_{\{x>0\}} + v^+1_{\{x=0\}}.$$

For this fact and further details see [2]. Let $m = (\bar{\lambda}_i, \bar{\mu}_i)$ be given. We will construct ψ, φ and u satisfying relations (3.4)–(3.5) and (3.2), and then argue that the map $m \mapsto u$ is a strategy.

For $i = 1, \dots, k$, denote $\bar{\rho}_i(t) = \bar{\lambda}_i(t)/\bar{\mu}_i(t)$. Let $q_1 = \int_0^\cdot(\bar{\lambda}_1 - \bar{\mu}_1)dt$ and $p_1 = \Gamma_1[q_1]$. Then $p_1 = q_1 + d_1$, where, by (A.1),

$$d_1(t) = \int_0^t(\bar{\lambda}_1(t') - \bar{\mu}_1(t'))^- 1_{\{p_1(t')=0\}}dt'.$$

As a result, $p_1 \geq 0$ and can be written as $p_1(t) = \int_0^t (\bar{\lambda}_1(t') - u_1(t')\bar{\mu}_1(t'))dt'$, where

$$u_1(t) = \begin{cases} 1 & \text{if} \quad p_1(t) > 0, \\ 1 \wedge \bar{\rho}_1(t) & \text{if} \quad p_1(t) = 0. \end{cases}$$

Now set $\psi_1 = p_1$. Then $\psi_1 \geq 0$ hence $\varphi_1 := \Gamma_1[\psi_1] = \psi_1$, and relations (3.2) and (3.4) hold. This gives a construction of (ψ_1, φ_1, u_1).

To proceed to (ψ_i, φ_i, u_i) for $i \geq 2$, we argue recursively. Fix $i \geq 2$. Denote $u_{1,i-1} = \sum_{m=1}^{i-1} u_m$. Set $q_i = \int_0^{\cdot} (\bar{\lambda}_i - (1-u_{1,i-1})\bar{\mu}_i)dt$ and $p_i = \Gamma_1[q_i]$. Arguing as before, $p_i = q_i + d_i$ where

$$d_i(t) = \int_0^t (\bar{\lambda}_i - (1 - u_{1,i-1})\bar{\mu}_i)^- 1_{\{p_i=0\}} dt',$$

hence $p_i \geq 0$ and $p_i(t) = \int_0^t (\bar{\lambda}_i - u_i\bar{\mu}_i)dt'$, where

$$u_i(t) = \begin{cases} 1 - u_{1,i-1}(t) & \text{if} \quad p_i(t) > 0, \\ (1 - u_{1,i-1}(t)) \wedge \bar{\rho}_i(t) & \text{if} \quad p_i(t) = 0. \end{cases}$$

Setting $\psi_i = p_i$ and $\varphi_i = \Gamma_1[\psi_i]$ gives $\varphi_i = \psi_i$ and agrees with (3.2) and (3.5). This completes the construction of (ψ, φ, u). The construction has the property that for every $t \geq 0$, $m|_{[0,t]}$ uniquely defines $(\psi, \varphi, u)|_{[0,t]}$, and moreover, the map $m \mapsto u$ is measurable. Thus the map is a strategy. □

The following lemma is used in Section 4.

Lemma A.2. *Fix $\lambda, \mu > 0$. The following identity holds for every $\beta \in \mathbb{R}$:*

$$\sup_{\alpha \in \mathbb{R}} (\alpha\beta - (\lambda(e^{\alpha} - 1) + \mu(e^{-\alpha} - 1))) = \inf(\lambda\omega(\bar{\lambda}/\lambda) + \mu\omega(\bar{\mu}/\mu) \mid \bar{\lambda} - \bar{\mu} = \beta, \text{ and } \bar{\lambda}, \bar{\mu} > 0).$$

Proof. Let $f_1(\beta)$ and $f_2(\beta)$ denote the l.h.s. and the r.h.s. Fix β. Then $f_1(\beta)$ is given as the supremum of a concave function of α. The maximizer $\bar{\alpha}$ satisfies $\beta - \lambda e^{\bar{\alpha}} + \mu e^{-\bar{\alpha}} = 0$. Also, $f_2(\beta)$ is the infimum of a convex function. Let $\lambda^* = \lambda e^{\bar{\alpha}}$ and $\mu^* = \mu e^{-\bar{\alpha}}$. Then from the above equality, (λ^*, μ^*) satisfies the constraint in the infimum, and a direct calculation shows that (λ^*, μ^*) is indeed the minimizer. Hence

$$\begin{aligned} &f_1(\beta) - f_2(\beta) \\ &= \bar{\alpha}\beta - (\lambda^* - \lambda + \mu^* - \mu) - \left(\lambda^* \ln(\frac{\lambda^*}{\lambda}) - \lambda^* + \lambda + \mu^* \ln(\frac{\lambda}{\lambda^*}) - \mu^* + \mu\right) \\ &= \bar{\alpha}\beta - (\lambda^* - \lambda + \mu^* - \mu) - (\lambda^* - \mu^*)\bar{\alpha} + (\lambda^* - \lambda + \mu^* - \mu) \\ &= 0. \end{aligned}$$

□

References

[1] R. Atar, A. Goswami, and A. Shwartz. Risk-sensitive control for the parallel server model. *SIAM J. Control Optim.* 51(2013) 4363–4386. MR-3134419

[2] A. Budhiraja and P. Dupuis. Simple necessary and sufficient conditions for the stability of constrained processes. *SIAM J. Appl. Math.* 59(1999) 1686–1700. MR-1699034

[3] A. Dembo and O. Zeitouni. *Large deviations techniques and applications*, volume 38 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, second edition, 1998. ISBN 0-387-98406-2. xvi+396 pp. MR-1619036

[4] P. Dupuis and R. S. Ellis. *A weak convergence approach to the theory of large deviations.* Wiley, New York, 1997 MR-1431744

[5] P. Dupuis, R. S. Ellis, and A. Weiss. Large deviations for Markov processes with discontinuous statistics. I. General upper bounds. *Ann. Probab.*, 19(3):1280–1297, 1991. MR-1112416

[6] K. Kuratowski and C. Ryll-Nardzewski. A general theorem on selectors. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 13: 397–403, 1965.

[7] A. Shwartz and A. Weiss. *Large deviations for performance analysis*. Stochastic Modeling Series. Chapman & Hall, London, 1995. ISBN 0-412-06311-5. x+556 pp. Queues, communications, and computing, With an appendix by Robert J. Vanderbei. MR-1335456

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards
- Free for authors, free for readers
- Quick publication (no backlog)

Economical model of EJP-ECP

- Low cost, based on free software (OJS[1])
- Non profit, sponsored by IMS[2], BS[3], PKP[4]
- Purely electronic and secure (LOCKSS[5])

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)
- Submit your best articles to EJP-ECP
- Choose EJP-ECP over for-profit journals

[1] OJS: Open Journal Systems http://pkp.sfu.ca/ojs/
[2] IMS: Institute of Mathematical Statistics http://www.imstat.org/
[3] BS: Bernoulli Society http://www.bernoulli-society.org/
[4] PK: Public Knowledge Project http://pkp.sfu.ca/
[5] LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[6] IMS Open Access Fund: http://www.imstat.org/publications/open.htm