Online Companion for

“A Call-Routing Problem with Service-Level Constraints”

Operations Research
Volume 51, Number 2, March-April 2003

Noah Gans
University of Pennsylvania

and

Yong-Pin Zhou
University of Washington

Online Companion for

A Call-Routing Problem with Service-Level Constraints

Noah Gans

OPIM Department
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6366
gans@wharton.upenn.edu

Yong-Pin Zhou

Dept. of Management Science
School of Business Administration
University of Washington
Seattle, WA 98195-3200
yongpin@u.washington.edu

Proof of Lemma 1 Proof We prove the lemma in two steps. First, we show that type-H priority policies are optimal. Then we show that, among the class of type-H priority policies, type-H work conserving policies are optimal.

Step 1. For any feasible policy π we construct an alternative, priority policy π' that outperforms π. The proof relies on the following coupling mechanism for the two systems: 1) arrival times of type-H calls are the same in the two systems; 2) service times of both types of calls are drawn from a single *i.i.d.* sequence $\{S_1, S_2, \ldots\}$ in the order in which they are put into service.

Given an arbitrary policy π, we construct the new policy π' as follows: 1) at each event epoch at which π puts calls into service, π' puts exactly the same number of calls into service; 2) at the event epochs of (1), π' gives priority to type-H calls, putting type-H calls into service if there are any waiting in queue.

Because service times are drawn from $\{S_1, S_2, \ldots\}$ in both systems, without regard to type, the occupancy of the system is exactly the same under both policies. Furthermore, because π' routes type-H calls ahead of type-L calls, at any given event epoch t, $\bar{H}_{\pi'}(t) \geq \bar{H}_{\pi}(t)$ and $\bar{q}_{\pi'}(t) = \Lambda_H(t) - \bar{H}_{\pi'}(t) \leq \Lambda_H(t) - \bar{H}_{\pi}(t) = \bar{q}_{\pi}(t)$. (Here $\bar{q}_{\pi}(t) \equiv \bar{q}(t)$ under policy π, etc.) Then, $d(\bar{q}_{\pi'}(t)) \leq d(\bar{q}_{\pi}(t))$ due to part i) in Assumption 1. Because this holds for any sample path, we have $\mathsf{E}_{\pi'}\left[d(\bar{q}(t))\right] \leq \mathsf{E}_{\pi}\left[d(\bar{q}(t))\right]$. Therefore,

$$\bar{D}_{\pi'}(s) \;=\; \limsup_{n\to\infty} \frac{1}{n}\mathsf{E}_{\pi'}\left[\sum_{t=0}^{n-1} d(\bar{q}(t))\right] \;\leq\; \limsup_{n\to\infty} \frac{1}{n}\mathsf{E}_{\pi}\left[\sum_{t=0}^{n-1} d(\bar{q}(t))\right] \;=\; \bar{D}_{\pi}(s)\,. \tag{6.1}$$

In addition, at every event epoch, the total number of calls routed by both policies is exactly the same, $\bar{n}_{\pi'}(t) = \bar{H}_{\pi'}(t) + \bar{L}_{\pi'}(t) = \bar{H}_{\pi}(t) + \bar{L}_{\pi}(t) = \bar{n}_{\pi}(t)$. Then recalling from (3.2) that, for admissible π and π', $\lim_{n\to\infty} \frac{\mathsf{E}_{\pi}[\bar{H}(n)]}{n} = \lim_{n\to\infty} \frac{\mathsf{E}_{\pi'}[\bar{H}(n)]}{n} = \lambda_H$ and defining $\bar{L}_{\pi}(-1) = \bar{L}_{\pi'}(-1) \equiv 0$, we have the following set of (equivalent) equalities (note that if $\lim_{n\to\infty} b_n$ exists, then $\liminf_{n\to\infty}(a_n + b_n) = \liminf_{n\to\infty} a_n + \lim_{n\to\infty} b_n$):

$$\begin{aligned}
\liminf_{n\to\infty} \frac{\mathsf{E}_{\pi}\left[\bar{n}(n)\right]}{n} &= \liminf_{n\to\infty} \frac{\mathsf{E}_{\pi'}\left[\bar{n}(n)\right]}{n} \\
\lim_{n\to\infty} \frac{\mathsf{E}_{\pi}\left[\bar{H}(n)\right]}{n} + \liminf_{n\to\infty} \frac{\mathsf{E}_{\pi}\left[\bar{L}(n)\right]}{n} &= \lim_{n\to\infty} \frac{\mathsf{E}_{\pi'}\left[\bar{H}(n)\right]}{n} + \liminf_{n\to\infty} \frac{\mathsf{E}_{\pi'}\left[\bar{L}(n)\right]}{n} \\
\lambda_H + \liminf_{n\to\infty} \frac{\mathsf{E}_{\pi}\left[\bar{L}(n)\right]}{n} &= \lambda_H + \liminf_{n\to\infty} \frac{\mathsf{E}_{\pi'}\left[\bar{L}(n)\right]}{n} \\
\liminf_{n\to\infty} \frac{\mathsf{E}_{\pi}\left[\bar{L}(n)\right]}{n} &= \liminf_{n\to\infty} \frac{\mathsf{E}_{\pi'}\left[\bar{L}(n)\right]}{n} \\
\liminf_{n\to\infty} \frac{1}{n}\mathsf{E}_{\pi}\left[\sum_{t=0}^{n}\left(\bar{L}(t) - \bar{L}(t-1)\right)\right] &= \liminf_{n\to\infty} \frac{1}{n}\mathsf{E}_{\pi'}\left[\sum_{t=0}^{n}\left(\bar{L}(t) - \bar{L}(t-1)\right)\right] \\
\bar{R}_{\pi}(s) &= \bar{R}_{\pi'}(s)\,.
\end{aligned} \tag{6.2}$$

Thus, from (6.1) we know that whenever π is feasible, π' is feasible as well, and from (6.2) we see that $\bar{R}_\pi(s) = \bar{R}_{\pi'}(s)$. That is, π' performs at least as well as π.

Step 2. Given any feasible type-H priority policy π', we construct an alternative, type-H work conserving policy π'' as follows: 1) at each event epoch at which there exist type-H calls in the queue, π'' puts as many type-H calls as possible into service; 2) whenever π' puts one or more type-L calls into service, π'' puts the same number of type-L calls into service.

To prove that π'' does at least as well as π', we couple the two systems as follows: 1) interarrival times in both systems are the same; 2) type-L calls in both systems are sampled from a common sequence $\{S_1^L, S_2^L, \ldots\}$; and 3) type-H calls in both systems are sampled from a common sequence $\{S_1^H, S_2^H, \ldots\}$.

The proof follows an inductive argument and considers the two systems in continuous time $t \in [0, \infty)$. First, suppose at time $t_0 = 0$ two systems, one controlled by π' and the other by π'', are identical, and let $t_1 > t_0$ be the first time at which π' puts a type-L job into service. Because π' is a type-H priority policy, it must be the case that at t_1 there are no type-H jobs in queue and that there is at least one server idle.

Note that from t_0 to t_1, the number of type-H jobs in queue under π'' will be no larger than that under π'. The sequence of service times of type-H jobs that the two systems process is identical, so type-H jobs that π'' puts into service earlier than π' will finish earlier than under π', and π'' can always feasibly put at least as many type-H jobs into service as π'.

Therefore $\bar{H}_{\pi''}(t) \geq \bar{H}_{\pi'}(t)$ and $\bar{q}_{\pi''}(t) = \Lambda_H(t) - \bar{H}_{\pi''}(t) \leq \Lambda_H(t) - \bar{H}_{\pi''}(t) = \bar{q}_{\pi'}(t)$ for all $t \in [t_0, t_1]$, and $\bar{L}_{\pi'}(t) = \bar{L}_{\pi''}(t) = 0$ on $[t_0, t_1)$ as well.

The same reasoning implies that at t_1 the system under π'' will have an occupancy (type-H and type-L jobs in service plus type-H jobs in queue) that is no larger than that of π' . Furthermore, the remaining processing times of those jobs in the system under π'' will be at most that of those under π'.

Next, let $t_2 > t_1$ be the time at which the next type-H job arrives. Because π'''s original occupancy (numbers in system and remaining service times) at t_1 is not more than that under π', and because the service times of the type-L jobs put into service are exactly the same in the two systems, π'' can feasibly put type-L jobs into service at any time over $[t_1, t_2)$ that π' does. Therefore, $\bar{L}_{\pi''}(t) = \bar{L}_{\pi'}(t)$ for all $t \in [t_1, t_2)$, and, trivially, $\bar{H}_{\pi'}(t) = \bar{H}_{\pi''}(t)$ for all $t \in [t_1, t_2)$ as well.

More broadly, $\bar{L}_{\pi''}(t) = \bar{L}_{\pi'}(t)$ and $\bar{q}_{\pi''}(t) \leq \bar{q}_{\pi'}(t)$ for all $t \in [t_0, t_2)$. Furthermore, at t_2 the occupancy under π'' is no more than that under π', and at t_2, the argument for $[t_0, t_1)$ again holds. In this way we can inductively show that $\bar{L}_{\pi''}(t) = \bar{L}_{\pi'}(t)$ and $\bar{q}_{\pi''}(t) \leq \bar{q}_{\pi'}(t)$ for all $t \in [0, \infty)$, so that $\bar{R}_{\pi''}(s) = \bar{R}_{\pi'}(s)$ and $\bar{D}_{\pi''}(s) \leq \bar{D}_{\pi'}(s)$. □

Proof of Lemma 2

To prove the lemma we formulate an embedded Semi-Markov Decision Process (SMDP) with $c+1$ states – 0 to c – and a finite number of associated actions. States $s \in \{0, \ldots, c\}$ and actions $\{A_s | s = 0, \ldots, c\}$ are exactly the same in both the MDP and the embedded SMDP. When the system enters state c, however, the expected time and service-level cost associated with the SMDP are functions of the "busy period" behavior of the original MDP.

More formally, in the original MDP, suppose that at epoch t_1 an arriving type-H call drives the system to enter state c (after action) and at $t_2 = \min\{t > t_1 | s_t = c-1\}$ the system reenters state $c-1$ before action. Then as a preliminary result we show that the expectations of $\tau_{c,c-1} \stackrel{\text{def}}{=} t_2 - t_1$ and $D_{c,c-1} \stackrel{\text{def}}{=} \sum_{t=t_1}^{t_2} d(\bar{q}_t)$, the number of periods and service-level cost incurred until the next epoch at which the system enters state $(c-1)$ (before action), are finite.

The proof uses the following bound:

Lemma 11. *If $\theta \in (0, \infty)$, then there exists an integer $m^* < \infty$ and real numbers $\theta^* \in (0, \theta)$ and $C^* < \infty$ such that $m\, e^{-\theta m} \leq C^* e^{-\theta^* m}$ for all $m \geq m^*$.*

<u>Proof</u> Let $\theta^* = \theta/2$ so that $m\, e^{-\theta m} = m\, e^{-\theta^* m} e^{-\theta^* m}$. Note that $\frac{\partial m\, e^{-\theta^* m}}{\partial m} = e^{-\theta^* m}(1 - \theta^* m) < 0$ for all $m \geq 1/\theta^*$. Then letting $m^* = \lceil 1/\theta^* \rceil$ and $C^* = m^* e^{-\theta^* m^*}$ we have $m\, e^{-\theta m} = m\, e^{-\theta^* m} e^{-\theta^* m} \leq C^* e^{-\theta^* m}$ for all $m \geq m^*$. □

Lemma 12. *Suppose $\rho < 1$. If $\pi \in \Pi$ is a type-H priority, type-H work-conserving policy then $\mathsf{E}[\tau_{c,c-1}], \mathsf{E}[D_{c,c-1}] < \infty$.*

<u>Proof</u> Note that when all c servers are busy, there are no state transitions due to uniformization, so the transitions of the MDP correspond exactly to the transitions of the underlying CTMC. Furthermore, $\tau_{c,c-1}$, the first passage time from state c to state $c-1$, corresponds exactly to the number of transitions in a busy period of an M/M/1 queue with arrival rate λ_H and service rate $c\mu$ (see Figure 3).

We begin with $\mathsf{E}[\tau_{c,c-1}]$. It is straightforward to show that $\mathbf{P}\{\tau_{c,c-1} \geq m\} = \mathrm{O}(e^{-\theta m})$. First, note that $\tau_{c,c-1}$ only takes odd integer values. Next, let the sequences of service times and interar-

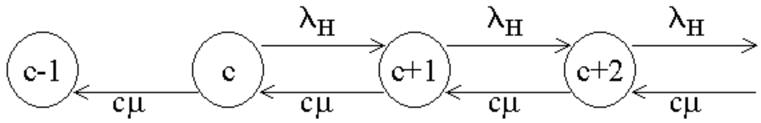

Figure 3: State Transition Diagram for First Passage Time from c to $c-1$

rival times during the busy period be $\{S_1, S_2, \ldots\}$ and $\{T_1, T_2, \ldots\}$. Then

$$\mathbf{P}\{\tau_{c,c-1} \geq 2m+1\} = \mathbf{P}\{\tau_{c,c-1} \geq 2m\} \leq \mathbf{P}\left\{\sum_{t=1}^{m} S_t > \sum_{t=1}^{m} T_t\right\} = \mathbf{P}\left\{\frac{1}{m}\sum_{t=1}^{m} S_t > \frac{1}{m}\sum_{t=1}^{m} T_t\right\} \leq C_1\, e^{-\theta m}$$

for some $\theta \in (0,\infty)$ and $C_1 < \infty$. The last inequality is the well-known Chernoff bound and holds whenever $\mathsf{E}[S] = \frac{1}{c\mu} < \frac{1}{\lambda_H} = \mathsf{E}[T]$, or $\rho < 1$ (for example, see §1.9 in Durrett [15]). Then let $\theta_1 = \theta/2$ so that $\mathbf{P}\{\tau_{c,c-1} \geq m\} \leq C_1\, e^{-\theta_1 m}$, and $\mathsf{E}[\tau_{c,c-1}] = \sum_{m=0}^{\infty} \mathbf{P}\{\tau_{c,c-1} \geq m\} \leq \sum_{m=0}^{\infty} C_1\, e^{-\theta_1 m} = C_1 \frac{1}{1-e^{\theta_1}} < \infty$ whenever $\rho < 1$.

We turn to $\mathsf{E}[D_{c,c-1}]$. First note that $\mathbf{P}\{\tau_{c,c-1} \geq m\} \leq C_1\, e^{-\theta_1 m}$ implies that $\mathbf{P}\{\tau_{c,c-1} = m\} \leq C_1\, e^{-\theta_1 m}$ as well. Second, note that, given a busy period of length m, a simple upper bound on the maximum queue length is m. Then we have

$$\begin{aligned}\mathsf{E}[D_{c,c-1}] &= \sum_{m=0}^{\infty}\left[\mathbf{P}\{\tau_{c,c-1} = m\}\sum_{t=1}^{m} d(\bar{q}_t)\right] \leq \sum_{m=0}^{\infty} C_1\, e^{-\theta_1 m}\, m\, d(m) \\ &\leq C_1 \sum_{m=0}^{\infty} C_2\, e^{-\theta_2 m}\, d(m) = C_1\, C_2\, \tilde{d}(e^{-\theta_2}) < \infty,\end{aligned}$$

where $\theta_2 \in (0,\theta_1)$ and $C_2 < \infty$. Here the second inequality follows from Lemma 11, and the equality with $\tilde{d}(\cdot)$ follows from Assumption 1.

Thus, $\mathsf{E}[D_{c,c-1}] < \infty$ as well whenever $\rho < 1$. □

Thus, the finite-state-space, finite-action-space SMDP embedded in the original MDP has finite expected durations and revenues in all states and is well defined. With this result in hand, we are ready to prove Lemma 2.

Proof (of Lemma 2). First we prove part (i) of the lemma. Then we prove the main result, that stationary, deterministic policies are optimal. Finally, we prove parts (ii)–(iv).

Part (i). Consider the original MDP with state space $S = \{0,1,2,\ldots\}$ and action space $A_s = \{0,\ldots,(c-s)^+\}$. We claim that for *any* stationary policy the class of states $\{c, c+1, \ldots\}$ is positive recurrent and that the mean absorption time into the class is finite. Then, given a particular stationary policy $\pi \in \Pi$, the states in $\{0,\ldots,c-1\}$ that communicate with c under π will be positive recurrent as well. Those that do not will be transient.

First, we note that the transition equations (3.6) imply that all states in the set $\{c, c+1, \ldots\}$ communicate. Therefore the set of states belongs to a single class.

Second, we show that the expected hitting time for the set $\{c, c+1, \ldots\}$ is finite. Note that any c transitions in a row that are type-H arrivals will drive the system to enter the set, no matter what the initial state. Let $E_n = \{\text{events } (n-1)c+1, \ldots, nc \text{ are arrivals}\}$. Then $\mathbf{P}\{E_n\} = {\lambda_H}^c$, and letting τ be the hitting time we have

$$\mathsf{E}[\tau] \;\leq\; c\sum_{n=1}^{\infty} n\left(\prod_{m=1}^{n-1}(1-\mathbf{P}\{E_m\})\right)\mathbf{P}\{E_n\} \;=\; c\sum_{n=1}^{\infty} n\,(1-{\lambda_H}^c)^{n-1}\,{\lambda_H}^c \;=\; c\,\frac{1-{\lambda_H}^c}{({\lambda_H}^c)^2} \;<\; \infty\,. \tag{6.3}$$

The right-hand-side of the inequality is c times the expected number of draws until success when sampling from a random variable which succeeds with probability $\mathbf{P}\{E_n\}$, the probability that c consecutive system events are arrivals.

Finally, we note that the positive recurrence of the class follows directly from the fact that the states $\{c, c+1, \ldots\}$ behave as an M/M/1 queue with $\rho = \frac{\lambda_H}{c\mu} < 1$ (for example, see §4.4 in Wolff [35]). In particular, suppose the system is in state $s > c$. Then the expected time until a return to state c equals the time to complete $s-c$ standard busy periods: the first standard busy period begins with s and ends with $s-1$ jobs in queue; the second begins with $s-1$ and ends with $s-2$ jobs, and so on. Then using Lemma 12, we see that the expected number of periods until the system returns to state c equals $(s-c)\mathsf{E}[\tau_{c,c-1}] < \infty$. Thus, the expected number of transitions until the system reaches state c from below is bounded by $\mathsf{E}[\tau] < \infty$, and the expected number of transitions until the system reaches state c from above is bounded by $(s-c)\mathsf{E}[\tau_{c,c-1}] < \infty$. Therefore state c is positive recurrent, as are all the states with which it communicates.

Main Result. We note three facts concerning the embedded SMDP. First, $S = \{0, \ldots, c\}$ and for each $s \in S$, $A_s = \{0, \ldots, c-s\}$, so the SMDP has finite state and action spaces. Second, part (i) of this lemma shows that the SMDP is unichain. Finally, we show, below, that on entry to each state of the SMDP, the expected time and the expected revenues earned until a transition occurs are finite. This implies that the maximand of the SMDP's value function is achieved (see Assumption 8.0.2 and Theorem 8.4.3 in Puterman [31]). Together, these three facts imply that there exists a stationary, deterministic policy that is average optimal (see Assumption 11.1.1 and Theorem 11.4.6 in Puterman [31]).

To complete the proof we therefore show that in each state of the SMDP the expected number of periods and the expected revenues accrued during each state transition are finite. For states 0 to $c-1$, each visit to a state generates a transition that lasts exactly one (MDP) period with expected revenues equal to $R(s,a) \in \{0,1,\ldots,c\}$, since there is no backlog and at most c type-L jobs may be routed at once. For state c, Lemma 12 implies that $\mathsf{E}[\tau_{c,c-1}]$, the expected sojourn time, and $-\delta\mathsf{E}[D_{c,c-1}]$, the expected revenues, are finite whenever $\rho < 1$.

Part (ii). Part (i) of the lemma shows that each stationary policy induces a unique set of positive recurrent states. Furthermore, it is not difficult to show that the class is aperiodic. If there is a state $s < c$ that is a part of the class, then because of uniformization, with probability $(c-s)\mu$ the system spends two periods in a row in state s, so the system is aperiodic. Otherwise c is the minimum recurrent state, and with probability $c\mu$ the system spends two periods in a row in state c, so it is aperiodic.

Thus each stationary policy π induces a positive recurrent and aperiodic Markov chain. This implies that the system's limiting state-action frequencies correspond to the stationary distribution of the induced Markov chain: $\lim_{n\to\infty} \frac{1}{n}\sum_{t=0}^{n-1} \mathbf{1}\{(s_t,a_t) = (s,a)\} = \xi_\pi(s,a)$ w.p. 1 $\forall s_0$. (For example see §3.1 in Karlin and Taylor [24].)

Part (iii). The fact that expected revenues in each state of the SMDP are finite implies that revenues in the MDP are uniformly integrable. Alternatively we have,

$$\begin{aligned}
\sum_{s\in S}\sum_{a\in A_s} \xi_\pi(s,a)|R(s,a) \,-\, \delta\, D(s,a)| &\leq \sum_{s=0}^{c-1}\sum_{a=0}^{c-s} \xi_\pi(s,a)a \,+\, \delta \sum_{s=c}^{\infty} \xi_\pi(s,0)d(s-c) \\
&= \sum_{s=0}^{c-1}\sum_{a=0}^{c-s} \xi_\pi(s,a)a \,+\, \delta\, \xi_\pi(c,0) \sum_{s=c}^{\infty} \rho^{s-c} d(s-c) \\
&= \sum_{s=0}^{c-1}\sum_{a=0}^{c-s} \xi_\pi(s,a)a \,+\, \xi_\pi(c,0)\,\delta\, \tilde{d}(\rho) \\
&\leq c \,+\, \delta\, \tilde{d}(\rho)\,.
\end{aligned}$$

Part (iv). Consider an arbitrary stationary, deterministic policy $\pi \in \Pi$, and let $i^* = \sup\{s \in S \,|\, a_s > 0\}$ be the maximal state in which π puts at least one type-L job into service. If $a_s = 0$ for all $s \in S$, then let $i^* = 0$. Note that, because π is stationary and deterministic, it always puts a_{i^*} type-L jobs into service whenever the system enters i^* before action, so states $\{s < i^*\}$ are transient. Furthermore, by the definition of i^*, in states $\{s > i^*\}$, the system never puts type-L jobs into service. Thus, the behavior of π is that of a generalized threshold policy with threshold i^* and number $a^* = a_{i^*}$. □

Proof of Lemma 3 Proof We consider the CTMC induced by a generalized threshold reservation policy π with arbitrary parameters $0 \leq i^* \equiv i < c$ and $1 \leq a^* \equiv a \leq c - i^*$. Let x_s denote the steady-state probability that the system state is s, as determined by the Markov chain induced by π. Then by balancing the flows across the cuts (as shown in the top half of Figure 4) we obtain the following equations:

$$\begin{aligned}
(i+2)\mu x_{i+2} &= ((i+1)\mu + \lambda_H)x_{i+1} && (6.4)\\
&\vdots \\
(i+a-1)\mu x_{i+a-1} &= (i+1)\mu x_{i+1} + \lambda_H x_{i+a-2} && (6.5)\\
(i+a)\mu x_{i+a} &= (i+1)\mu x_{i+1} + \lambda_H x_{i+a-1} && (6.6)\\
(i+a+1)\mu x_{i+a+1} &= \lambda_H x_{i+a} && (6.7)\\
&\vdots
\end{aligned}$$

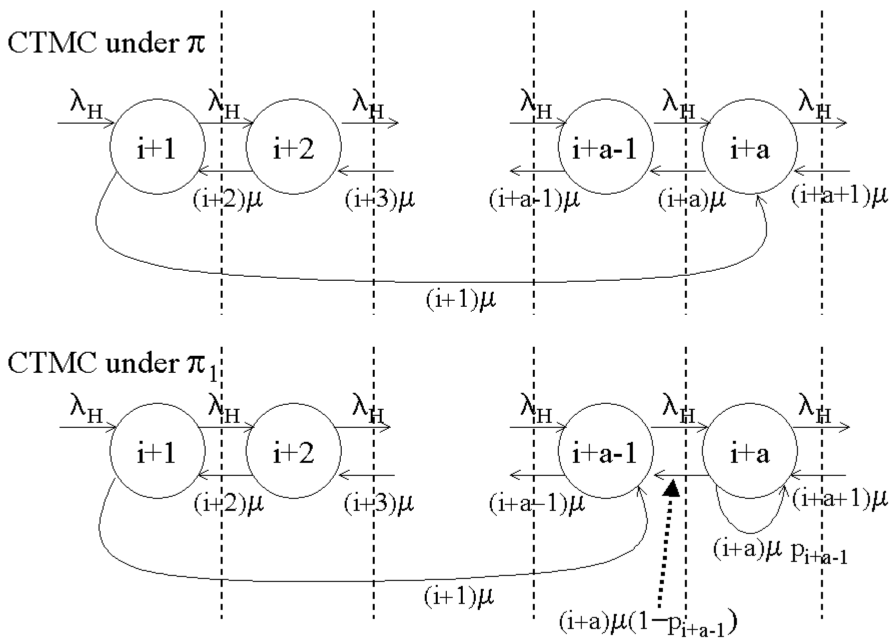

Figure 4: State Transition Diagrams of π and π_1

Now define $p_{i+a-1} = \frac{(i+1)x_{i+1}}{(i+a)x_{i+a}}$. First, note that (6.6) implies that $p_{i+a-1} \in (0,1)$. Next, we claim that the following policy, π_1, induces a CTMC that has a stationary distribution that is identical to that of the original, generalized threshold reservation policy:

$$a_s = \begin{cases} 1 & \text{for } 0 \leq s < i, \\ a-1 & \text{for } s = i, \\ 0 & \text{for } i < s < i+a-1, \\ 1 \text{ w. p. } p_{i+a-1} \text{ and } 0 \text{ w. p. } (1-p_{i+a-1}) & \text{for } s = i+a-1, \\ 0 & \text{for } x \geq i+a. \end{cases}$$

The (after-action) state transition diagram corresponding to π_1 is the bottom one in Figure 4, and inspection of Figure 4 shows that, by the definition of p_{i+a-1}, the stationary distributions of the CTMC's induced by π and π_1 are identical. In particular, under π and π_1 the flows into state $i+a-1$ match,

$$\lambda_H \, x_{i+a-2} \; + \; (i+a)\mu \, x_{i+a} \; = \; \lambda_H \, x_{i+a-2} \; + \; (i+1)\mu \, x_{i+1} \; + \; (i+a)\mu \, (1 - p_{i+a-1}) \, x_{i+a} \; ,$$

as do the flows into $i+a$,

$$\lambda_H \, x_{i+a-1} + (i{+}1)\mu \, x_{i+1} + (i{+}a{+}1)\mu \, x_{i+a+1} \; = \; \lambda_H \, x_{i+a-1} + (i{+}a)\mu \, p_{i+a-1} \, x_{i+1} + (i{+}a{+}1)\mu \, x_{i+a+1}.$$

Thus, the policy π_1, which 1) puts $(a-1)$ type-L jobs into service with probability one when it enters state i, and 2) puts one job into service with probability p_{i+a-1} when it enters state $a-1$, is equivalent to π. Furthermore, we can recursively apply this transformation to π_1, π_2, etc. to define an equivalent policy π_{a-1} that puts at most one job into service at a time. □

Proof of Lemma 6 Proof We begin by proving part (ii); then we prove part (i).

Part (ii). As the threshold increases from i to $i+1$, the denominator of (3.24) loses a positive term, so that $\xi_i(c) < \xi_{i+1}(c)$. Therefore, from (3.26) we see that $\bar{D}_i$ is increasing in i. Furthermore, $\xi_0(c)$, the steady-state probability that there are exactly c jobs in the system in a standard M/M/c queue, is strictly greater than zero, so that $\bar{D}_0 > 0$ as well. Similarly. when the threshold equals c, all c servers are always busy, and state c is the analogue of the empty state in an M/M/1 queue with arrival rate λ_H and service rate $c\mu$ (see Figure 3). In this case the sum of the probabilities, $\sum_{s=c}^{\infty} \xi_c(s) = \sum_{s=c}^{\infty} \xi_c(c)\rho^{s-c} = \xi_c(c)/(1-\rho)$ equals one, so that $\xi_c(c) = (1-\rho)$, and $\bar{D}_c = (1-\rho)\tilde{d}(\rho)$.

Part (i). The fact that $\xi_i(c) < \xi_{i+1}(c)$, together with the definition $\xi_i(s) = \xi_i(c)\rho^{s-c}$ for all $s \geq c$, implies that $\sum_{s=c}^{\infty} \xi_i(s) < \sum_{s=c}^{\infty} \xi_{i+1}(s)$ and that $\sum_{s=i}^{c-1} \xi_i(s) > \sum_{s=i+1}^{c-1} \xi_{i+1}(s)$. At the same time (3.22) implies that $\xi_i(s) < \xi_{i+1}(s)$ for $s \in \{i+1, \ldots, c-1\}$ so that $\xi_i(i) > \sum_{s=i+1}^{c-1}(\xi_{i+1}(s) - \xi_i(s))$.

Therefore, we let $x_s = \xi_{i+1}(s) - \xi_i(s)$ for $s = i+1, \ldots, c-1$ and $x_i = \xi_i(i) - \sum_{s=i+1}^{c-1} x_s$, so that

$$\begin{aligned}
\sum_{s=i}^{c-1}(c-s)\xi_i(s) &= (c-i)\xi_i(i) + \sum_{s=i+1}^{c-1}(c-s)\xi_i(s) \\
&= (c-i)x_i + \sum_{s=i+1}^{c-1}\left[(c-i)x_s \,+\, (c-s)\xi_i(s)\right] \\
&\geq \sum_{s=i+1}^{c-1}\left[(c-s)x_s \,+\, (c-s)\xi_i(s)\right] \\
&= \sum_{s=i+1}^{c-1}(c-s)\xi_{i+1}(s)\ .
\end{aligned}$$

Thus, $\sum_{s=i}^{c-1}(c-s)\xi_i(s)$ is decreasing in i, and from (3.25) we see that $\bar{R}_i$ is, in turn, increasing in i. Finally, note that when the threshold is zero, no type-L jobs are ever put into service, and $\bar{R}_0 = 0$. Similarly, when the threshold is c, no server is idle, the last term of (3.25) drops out, and $\bar{R}_c = c\mu - \lambda_H$. □

Proof of Lemma 8 The proof of the lemma parallels that of Lemma 2. In a preliminary lemma we demonstrate that the expected duration and backlog cost associated with any busy period is finite. We then apply the result to prove the main lemma.

More formally, in the original MDP, suppose that at epoch t_1 an arriving type-H call drives the system to enter state $i_1 + j_1 = c+1$ and at $t_2 = \min\{t > t_1 | i_t + j_t = c\}$ the busy period ends in state $i_2 + j_2 = c$. (Note that this differs slightly from Lemma 12 in which busy periods start with c and end with $c-1$ in the system.) Then letting $p(j_1, j_2) \stackrel{\text{def}}{=} \mathbf{P}\{j_{t_2} = j_2 \,|\, j_1\}$ be the probability that a busy period that starts in j_1 ends in j_2, and defining $\tau(j_1, j_2) \stackrel{\text{def}}{=} t_2 - t_1$ and $D(j_1, j_2) \stackrel{\text{def}}{=} \sum_{t=t_1}^{t_2} d(i_t + j_t - c)$ to be, respectively, the number of transition epochs and the service-level cost incurred during that busy period, we have the following.

Lemma 13. *Suppose that $\pi \in \Pi$ is type-H priority policy and that $\rho < 1$. Then*

i) *for all $j_2 > j_1$ $p(j_1, j_2) = 0$, and for all $0 \leq j_2 \leq j_1 \leq c$, $p(j_1, j_2) > 0$;*

ii) $\mathsf{E}\left[\tau(j_1, j_2)\right] < \infty$ *and* $\mathsf{E}\left[D(j_1, j_2)\right] < \infty$.

<u>Proof</u> We prove the two parts in turn.

Part (i). It is immediate that $p(j_1, j_2) = 0$ for $j_2 > j_1$, since under a type-H priority policy no type-L jobs will be put in service as long as $i_t + j_t \geq c$. Similarly, it is not difficult to show that

$p(j_1, j_2) > 0$ for any $0 \leq j_2 \leq j_1 \leq c$. In particular, if $j_1 - j_2 - 1$ arrivals of type-H calls are followed by $j_1 - j_2$ departures of type-L calls, then a busy period with the desired ending state is generated, and this sequence of events occurs with probability ${\lambda_H}^{j_1-j_2-1} \prod_{j=j_2}^{j_1} (j\mu_L)$, so that $p(j_1, j_2) \geq {\lambda_H}^{j_2-j_1-1} \prod_{j=j_2}^{j_1} (j\mu_L) > 0$.

Part (ii). First, we define $\tau(j_1)$ and $D(j_1)$ to be the duration and cost of the busy period that starts in state $i_1 + j_1 = c + 1$, and ends in any state $i_2 + j_2 = c$ (no matter what the particular i_2 and j_2). That is

$$\mathsf{E}\left[\tau(j_1)\right] = \sum_{j_2=0}^{j_1} p(j_1, j_2) \mathsf{E}\left[\tau(j_1, j_2)\right] \tag{6.9}$$

$$\mathsf{E}\left[D(j_1)\right] = \sum_{j_2=0}^{j_1} p(j_1, j_2) \mathsf{E}\left[D(j_1, j_2)\right] . \tag{6.10}$$

From part (i) we know that $p(j_1, j_2) > 0$ for all $0 \leq j_2 \leq j_1 \leq c$. Thus, if we can show that $\mathsf{E}\left[\tau(j_1)\right]$ and $\mathsf{E}\left[D(j_1)\right]$ are finite, then the proof will be complete.

To show that $\mathsf{E}\left[\tau(j_1)\right]$ and $\mathsf{E}\left[D(j_1)\right]$ are finite, we analyze the system under the following policy π': first, the system processes exactly one type-L job at a time and no type-H jobs; then, as soon as $j_t = 0$, the system processes any type-H jobs in the backlog using all c servers. Thus, the busy period behavior of π' is one in which: 1) first, j_1 exponentially distributed services at rate μ_L occur; 2) during those j_1 services, some random number N of type-H arrivals takes place; and 3) once the j_1 type-L services are complete an M/M/1 busy period – with arrival rate λ_H, service rate $c\mu_H$, and $(N + i_1 - c)^+ = (N - j_1)^+$ initial customers in queue – ensues.

If we couple two systems – one using a type-H priority, type-H work-conserving policy π, and the other using π' – and compare their performance, we can show that at all times the backlog under π' exceeds that in π. In both systems we let interarrival times of type-H calls be the same. Similarly, we sample service times of both types of calls in the order jobs are put into service. Then no matter what order π' puts the j_1 type-L jobs into service, at any time during the busy period, the number of type-L jobs it has processed will be no larger than the number of type-L jobs processed by π. Similarly, at any time during the busy period, the number of servers π' dedicates to type-H jobs is smaller than that dedicated by π, which implies that the number of type-H jobs processed by π' always lags that processed by π'. Thus the number of arrivals to both systems is the same, but at any time the number of service completions in π' lags that in π. Therefore, the number of jobs in the backlog under π' is always at least as large as that under π, and upper bounds on $\mathsf{E}\left[\tau(j_1)\right]$ and $\mathsf{E}\left[D(j_1)\right]$ under π' will be upper bounds under π as well.

Given the definition of π', N is the number of Poisson (λ_H) arrivals during an Erlang (j_1,μ_L) service period, so that

$$\begin{aligned}
\mathbf{P}\{N=n\} &= \int_0^\infty \frac{e^{-\lambda_H s}(\lambda_H s)^n}{n!}\,\frac{\mu_L\, e^{-\mu_L s}(\mu_L s)^{j_1-1}}{(j_1-1)!}\,ds \\
&= \frac{{\lambda_H}^n\,{\mu_L}^{j_1}}{n!\,(j_1-1)!}\int_0^\infty e^{-(\lambda_H+\mu_L)s}\,s^{n+j_1-1}\,ds \\
&= \frac{{\lambda_H}^n\,{\mu_L}^{j_1}}{n!\,(j_1-1)!}\,\frac{(n+j_1-1)!}{(\lambda_H+\mu_L)^{n+j_1}}\int_0^\infty \frac{(\lambda_H+\mu_L)\,e^{-(\lambda_H+\mu_L)s}((\lambda_H+\mu_L)s)^{n+j_1-1}}{(n+j_1-1)!}\,ds \\
&= \frac{{\lambda_H}^n\,{\mu_L}^{j_1}}{n!\,(j_1-1)!}\,\frac{(n+j_1-1)!}{(\lambda_H+\mu_L)^{n+j_1}} \\
&= \binom{n+j_1-1}{n}\left(\frac{\lambda_H}{\lambda_H+\mu_L}\right)^n\left(\frac{\mu_L}{\lambda_H+\mu_L}\right)^{j_1} \\
&\le \binom{n+j_1}{n}\left(\frac{\lambda_H}{\lambda_H+\mu_L}\right)^n\left(\frac{\mu_L}{\lambda_H+\mu_L}\right)^{j_1}.
\end{aligned}$$

Note that the right-hand-side of the last inequality is a sample realization of a binomially distributed random variable. If we define X_m to be a Bernoulli random variable with a probability $\frac{\lambda_H}{\lambda_H+\mu_L}$ of being one, then

$$\mathbf{P}\{N=n\} \;=\; \mathbf{P}\left\{\sum_{m=1}^{n+j_1} X_m = n\right\} \;\le\; \mathbf{P}\left\{\sum_{m=1}^{n+j_1} X_m \ge n\right\} \;=\; \mathbf{P}\left\{\frac{1}{n+j_1}\sum_{m=1}^{n+j_1} X_m \;\ge\; \frac{n}{n+j_1}\right\}.$$

Now $\frac{n}{n+j_1} \uparrow 1$ as $n\to\infty$, so there exists some $n^* < \infty$ such that $\frac{n}{n+j_1} > \frac{\lambda_H}{\lambda_H+\mu_L}$ for all $n \ge n^*$. Then for all $n \ge n^*$ we have

$$\mathbf{P}\left\{\frac{1}{n+j_1}\sum_{m=1}^{n+j_1} X_m \;\ge\; \frac{n}{n+j_1}\right\} \;\le\; \mathbf{P}\left\{\frac{1}{n+j_1}\sum_{m=1}^{n+j_1} X_m \;\ge\; \frac{\lambda_H}{\lambda_H+\mu_L}\right\} \;\le\; C\,e^{-\theta_1 n},$$

for some $\theta_1 \in (0,\infty)$ and $C<\infty$. The last inequality again follows from the Chernoff bounds (see §1.9 in Durrett [15]). Furthermore, we can let $C_1 = \max\{1, (Ce^{-\theta_1 n^*})^{-1}\}$ so that

$$\mathbf{P}\{N=n\} \;\le\; C_1\,e^{-\theta_1 n} \tag{6.11}$$

for all $n \ge 0$.

Given this bound on $\mathbf{P}\{N=n\}$ we can characterize upper bounds on $\mathsf{E}[\tau(j_1)]$ and $\mathsf{E}[D(j_1)]$. If $N \le j_1$, then the number of jobs in queue equals zero and the busy period has ended when the processing of the j_1 type-L jobs is complete. If, however, $N > j_1$ then there are $N+i_1 > c$ type-H jobs and $j_t = 0$ type-L jobs in the system, and the system, once again, behaves like an M/M/1 queue with arrival rate λ_H, service rate $c\mu_H$, and an initial number $N-j_1$ jobs in queue.

In this case, the busy period behaves as $N-j_1$ standard busy periods: the first standard busy period begins with $N-j_1$ and ends with $N-j_1-1$ jobs in queue; the second begins with $N-j_1-1$

and ends with $N - j_1 - 2$ jobs, and so on. Furthermore, we can use Lemma 12 to demonstrate the finiteness of each of these busy periods.

Because the uniformization rate is $\lambda_H + c\mu_H + c\mu_L = 1$, however, the expected number of transitions of the MDP per event in the M/M/1 system equals $\frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}$. Similarly, the expected number of transitions in the MDP for each of the initial j_1 type-L services in the system under π' equals $\frac{\lambda_H + c\mu_H + c\mu_L}{\mu_L}$. When we bound the expected duration and cost of these busy periods, we must adjust for the uniformization rate.

Thus, we have the following upper bound on $\mathsf{E}[\tau(j_1)]$:

$$\mathsf{E}[\tau(j_1)] \;\le\; j_1 \left(\frac{1}{\mu_L}\right) \frac{\lambda_H + c\mu_H + c\mu_L}{\mu_L} \;+\; \sum_{n=0}^{\infty} \mathbf{P}\{N = n\}\; (n - j_1)^+ \,\mathsf{E}[\tau_{c,c-1}]\; \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\,.$$

Here the first term represents the expected number of periods spent serving the j_1 type-L jobs, and the second represents the expected number of periods spent serving type-H jobs. The term $\mathsf{E}[\tau_{c,c-1}] < \infty$ is the expected length of an M/M/1 busy period as defined in Lemma 12.

Then from (6.11) and Lemma 12 we have that, whenever, $\rho < 1$,

$$\begin{aligned}
\mathsf{E}[\tau(j_1)] &\le \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2} + \sum_{n=0}^{\infty} \mathbf{P}\{N = n\}\; (n - j_1)^+ \,\mathsf{E}[\tau_{c,c-1}]\; \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H} \\
&\le \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2} + \sum_{n=0}^{\infty} \mathbf{P}\{N = n\}\; n\,\mathsf{E}[\tau_{c,c-1}]\; \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H} \\
&\le \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2} + \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\,\mathsf{E}[\tau_{c,c-1}] \sum_{n=0}^{\infty} C_1\, n\, e^{-\theta_1 n} \\
&= \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2} + \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\,\mathsf{E}[\tau_{c,c-1}]\; C_1\, \frac{e^{-\theta_1}}{(1 - e^{-\theta_1})^2}\,.
\end{aligned}$$

Thus $\mathsf{E}[\tau(j_1)] < \infty$.

The finiteness $\mathsf{E}[D(j_1)]$ is demonstrated in the same fashion. Let $D_L(j_1)$ be the backlog cost associated with the j_1 initial type-L services and $D_H(j_1)$ be the backlog cost of the ensuing M/M/1 busy period. Then $D(j_1) = D_L(j_1) + D_H(j_1)$, and from (6.11) and Assumption 1 we have

$$\begin{aligned}
\mathsf{E}[D_L(j_1)] &\le j_1 \left(\frac{1}{\mu_L}\right) \frac{\lambda_H + c\mu_H + c\mu_L}{\mu_L} \sum_{n=0}^{\infty} \mathbf{P}\{N = n\}\; d(n+1) \\
&\le \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2} \sum_{n=0}^{\infty} C_1\, e^{-\theta_1 n}\, d(n+1) \\
&\le \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2}\, C_1\, e^{-\theta_1} \sum_{n=0}^{\infty} e^{-\theta_1 n}\, d(n) \\
&= \frac{j_1(\lambda_H + c\mu_H + c\mu_L)}{{\mu_L}^2}\, C_1 e^{-\theta_1}\, \tilde{d}(e^{-\theta_1}) \quad < \quad \infty
\end{aligned}$$

whenever $\rho < 1$. Note that the terms before the summation on the right-hand-sides of the inequalities represent the expected number of periods spent serving the first j_1 jobs. The summation that follows is an upper bound on the expected backlog cost per period during that time.

Similarly, from Lemma 12 we know that whenever $\rho < 1$ there exist a $\theta_3 \in (0, \infty)$ and a $C_3 < \infty$ such that the probability of a busy period of length m is bounded above by $C_3\, e^{\theta_3 m}$ and that an upper bound on the backlog cost of a busy period of length m is $md(m)$. Then, given a busy period that begins with $(N - j_1)^+$ type-H jobs in queue, we have the following upper bound:

$$
\begin{aligned}
\mathsf{E}\left[D_H(j_1)\right] &\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H} \sum_{n=j_1+1}^{\infty} \left[\mathbf{P}\left\{N = n\right\} \sum_{l=1}^{n} \sum_{m=1}^{\infty} \mathbf{P}\left\{\tau_{c,c-1} = m\right\} md(m+l)\right] \\
&\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H} \sum_{n=0}^{\infty} \left[\mathbf{P}\left\{N = n\right\} n \sum_{m=1}^{\infty} \mathbf{P}\left\{\tau_{c,c-1} = m\right\} md(m+n)\right] \\
&\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H} \sum_{n=0}^{\infty} \left[C_1 e^{-\theta_1 n} n \sum_{m=1}^{\infty} C_3 e^{-\theta_3 m} md(m+n)\right] \\
&\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H} \sum_{n=0}^{\infty} \left[C_2 e^{-\theta_2 n} \sum_{m=1}^{\infty} C_4 e^{-\theta_4 m} d(m+n)\right] \\
&= \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\, C_2\, C_4 \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} (e^{-\theta_2 n}\, e^{-\theta_4 m}\, d(m+n)) \\
&\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\, C_5 \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} (e^{-\theta_5 (n+m)}\, d(m+n))\,, \qquad (6.12)
\end{aligned}
$$

where $C_5 = C_3 C_4$ and $\theta_5 = \min\{\theta_3, \theta_4\}$. Again, the term before the summations on the right-hand-sides accounts for the expected number of transitions of the MDP per transition in the M/M/1 queue, and the summations that follow are an upper bound on the expected cost associated with a busy period that starts with $(N - j_1)^+$ type-H jobs in queue.

Finally, we can collapse the two summations of (6.12) into one and use Lemma 11 and Assumption 1 to show that

$$
\begin{aligned}
\mathsf{E}\left[D_H(j_1)\right] &\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\, C_5 \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} (e^{-\theta_5 (n+m)}\, d(m+n)) \\
&= \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\, C_5 \sum_{n=0}^{\infty} (n\, e^{-\theta_5 n}\, d(n)) \\
&\leq \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\, C_5 \sum_{n=0}^{\infty} (C_6\, e^{-\theta_6 n}\, d(n)) \\
&= \frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}\, C_5\, C_6\, \tilde{d}(e^{-\theta_6}) \;<\; \infty\,,
\end{aligned}
$$

whenever $\rho < 1$.

Thus, whenever $\rho < 1$, $\mathsf{E}[\tau(j_1)] < \infty$ and $\mathsf{E}[D(j_1)] = \mathsf{E}[D_L(j_1)] + \mathsf{E}[D_H(j_1)] < \infty$ for all $0 \le j_1 \le c$. This completes the proof. □

With Lemma 13 at our disposal, we are ready to prove Lemma 8.

Proof (of Lemma 8). Note that we explicitly prove only parts (i) and (iii) of the lemma. Given part (i), the proof of the main result and of part (ii) follow exactly the same line of argument as that in Lemma 2.

Part (i). We show that each stationary, type-H priority policy induces a unique, positive recurrent class of policies and that the expected number of transitions until absorption into the class is finite.

Our proof follows the argument of part (i) of Lemma 2. In particular, we show that for *any* stationary, type-H priority policy the set of states $\{(c,0),(c{+}1,0),\ldots\}$ is positive recurrent and that the expected absorption time into the set is finite. Then for an arbitrary stationary, type-H priority policy π, the unique class that satisfies the lemma's requirements is the union of $\{(c,0),(c{+}1,0),\ldots\}$ and the states under π that communicate with $\{(c,0),(c+1,0),\ldots\}$.

First, we note that the transition equations (4.1) imply that all states in the set $\{(c,0),(c+1,0),\ldots\}$ communicate. Therefore the set of states belongs to a single class.

Second, we show that the expected hitting time for the set $\{(c,0),(c{+}1,0),\ldots\}$ is finite. Suppose the current state is $i{+}j \le c$. Then, as in (6.3), we can show that for any such i and j there is a finite upper bound on the expected number of transitions until the system reaches a state $i_t + j_t = c+1$. Call that bound $\mathsf{E}[\tau_1] < \infty$.

Similarly, from Lemma 13 we know the following. For any state $i_t + j_t = c + 1$ there is a probability $p(j_t, 0) > 0$ that the system will hit the state $(c, 0)$ at the end of the busy period, and $1 - p(j_t, 0)$ that it will not. Let $p \stackrel{\text{def}}{=} \min\{p(j,0) \,|\, j = 0,\ldots,c\}$. Furthermore, whenever $\rho < 1$, the expected number of periods until the system returns to a state $i_t + j_t = c$, $\mathsf{E}[\tau_2] \stackrel{\text{def}}{=} \max\{\mathsf{E}[\tau(j_1,j_2)] \,|\, 0 \le j_2 \le j_1 \le c\}$, is finite as well.

Then the expected number of periods until the system hits state $(c, 0)$ is bounded above by

$$\mathsf{E}[\tau_1] \;+\; \sum_{n=0}^{\infty} [(1-p)^n \, p \, n \, (\mathsf{E}[\tau_2] + \mathsf{E}[\tau_1])] \;+\; \mathsf{E}[\tau_2] \quad < \quad \infty \tag{6.13}$$

for any $p \in (0,1)$ and $\mathsf{E}[\tau_2] + \mathsf{E}[\tau_1] < \infty$. Here, the first term is an upper bound on the initial number of periods until a state $i + j = c + 1$ is reached. The second, summation term is an upper bound on the expected number of transitions in a geometrically distributed number of cycles in which the busy period does not end with state $(c, 0)$. The last term is an upper bound on the length of the busy period that ends with state $(c, 0)$.

Thus, if the system is already in $\{(c,0),(c+1,0),\ldots\}$, then the expected hitting time to enter the class is zero, and if the system is not in one of the states $\{(c,0),(c+1,0),\ldots\}$, then the expected hitting time to state $(c,0)$ is finite as well.

Finally, we note from (4.1) that once the system is in class $\{(c,0),(c+1,0),\ldots\}$ it behaves as an M/M/1 queue with arrival rate λ_H and service rate $c\mu_H$. Using the line of reasoning developed in part (i) of Lemma 2, we can again show that the set $\{(c,0),(c+1,0),\ldots\}$ is positive recurrent. The only adjustment we must make is for the uniformization rate of $\lambda_H + c\mu_H + c\mu_L$. This requires that we adjust the expectation $\mathsf{E}\left[\tau_{c,c-1}\right]$ upwards by a factor of $\frac{\lambda_H + c\mu_H + c\mu_L}{\lambda_H + c\mu_H}$.

Part (iii). To show that single-period rewards are uniformly integrable, we use a renewal-reward argument. Let the renewal state be $(c,0)$, and let τ be the time and $|R|$ be the sum of the absolute values of the rewards earned in a renewal cycle.

From (6.13) we know that an upper bound on $\mathsf{E}\left[\tau\right]$ is finite, and we can calculate an analogous upper bound for $\mathsf{E}\left[D\right]$. More specifically, given $i+j \le c$, an upper bound on the one-period reward is c. Similarly, given Lemma 13 we can let $\mathsf{E}\left[D\right] = \max_{j_1}\{\mathsf{E}\left[D(j_1\right]\} < \infty$ be an upper bound on the service-level cost of any single busy period.

Then we modify (6.13) as follows to bound $\mathsf{E}\left[R\right]$:

$$\mathsf{E}\left[|R|\right] \;\; \le \;\; c\mathsf{E}\left[\tau_1\right] \; + \; \sum_{n=0}^{\infty} \left[(1-p)^n \, p \, n \left(\delta \mathsf{E}\left[D\right] + c\mathsf{E}\left[\tau_1\right]\right)\right] \; + \; \delta \mathsf{E}\left[D\right] \;\; < \;\; \infty \, .$$

Note that the $c\mathsf{E}\left[\tau\right]$ terms are upper bounds on the revenues earned during the portion of each cycle spent with $i+j \le c$, and the $\delta\mathsf{E}\left[D\right]$ terms are upper bounds on the δ-cost, *absolute value* of the (negative) revenue that is incurred during the busy-period portion of each cycle.

Thus $\mathsf{E}\left[\tau\right] < \infty$ and $\mathsf{E}\left[|R|\right] < \infty$, and we can use the Renewal Reward Theorem to express the expected absolute value of the reward per period as $\frac{\mathsf{E}[|R|]}{\mathsf{E}[\tau]}$ (for example, see §2.3 in Wolff [35]). Furthermore $\mathsf{E}\left[\tau\right] \ge 1$, so that we have $\frac{\mathsf{E}[|R|]}{\mathsf{E}[\tau]} \le \mathsf{E}\left[|R|\right] < \infty$.

Finally we note that, because of uniformization at rate $\lambda_H + c\mu_H + c\mu_L$, there is a positive probability of spending two consecutive periods in any recurrent state. Together with part (i) of the lemma this implies that each stationary, type-H priority, type-H work-conserving policy π induces a positive recurrent, aperiodic Markov chain. Therefore, the process is ergodic and $\sum_{s \in S, a \in A_s} \xi_\pi(s,a)|R(s,a) - \delta D(s,a)| = \frac{\mathsf{E}[|R|]}{\mathsf{E}[\tau]} < \infty$ as well. □

Proof of Lemma 9 Proof First, we consider $0 < j < c$. For any fixed j, by definition $0 < p_1(j) + p_2(j) < 1$, and the determinant of $g(j,z)$, $1 - 4p_1(j)p_2(j) > (p_1(j)+p_2(j))^2 - 4p_1(j)p_2(j) =$

$(p_1(j) - p_2(j))^2 \geq 0$. Therefore, for $0 < j < c$, there exist two, distinct real roots of $g(j,z)$: $z_j \neq z'_j$. Moreover, because $z_j + z'_j = 1/p_2(j) > 0$ and $z_j z'_j = p_1(j)/p_2(j) > 0$, we must have $z_j, z'_j > 0$.

In addition, note that $g(j,0) > 0$ and $g(j,1) < 0$. As a quadratic polynomial with positive leading coefficient, $g(j,z)$ must have one root between 0 and 1 and the other greater than 1: $0 < z_j < 1 < z'_j$, $\forall\ 0 < j < c$.

Next, we address the boundary cases. From (4.13) we see that $z_0 = \frac{\lambda_H}{c\mu_H} < 1 = z'_0$ are the roots of $g(0)$. In the case of $j = c$, the $p_2(j)$ term vanishes from the transition equation (4.9), and $g(c,z) = -z + p_1(j)$. In this case $z_c = p_1(c) = \frac{\lambda_H}{\lambda_H + c\mu_L} \in (0,1)$ is the single root of the equation.

□

Proof of Lemma 10 Proof Note that throughout the proof of this lemma (only) we make the following three changes to reduce the notational burden. First, throughout the proof the argument j of $p_1(j)$, $p_2(j)$ and $p_3(j)$ remains fixed. Therefore, we drop the (j) and write p_1, p_2, and p_3. Second, for states in which $i + j \geq c$, the action $a = 0$ is fixed and we write $\xi(i,j)$, rather than $\xi(i,j,0)$. Furthermore, note that in these states the before-action state and after-action states are identical: $(i_t, j_t) = (\bar{\imath}_t, \bar{\jmath}_t)$.

Finally, for states in which $i + j = c$ we abuse our original convention and let

$$\xi(i,j) \stackrel{\text{def}}{=} \sum_{a=0}^{j} \xi(i, j-a, a) \tag{6.14}$$

denote the stationary probability of being in state $i + j = c$ *after-action*. In particular, this lets us write the generating function for states $\{\xi(i,j) \,|\, i \geq c - j\}$ as

$$f_j(z) = \sum_{q=0}^{\infty} \xi(c - j + q, j) z^q \,. \tag{6.15}$$

In addition, given this definition of $\xi(i,j)$ for $i + j = c$, the balance constraints (4.8)–(4.9) are well defined for all $i + j \geq c + 1$ (rather than $i + j \geq c + 2$).

Part (i). We prove part (i) in three steps.

First, we show that $z_j \neq z_k$ for all $j \neq k$. From (4.13) and the definition of z_j, we have

$$(c-j)\mu_H z_j^2 - [(c-j)\mu_H + j\mu_L + \lambda_H]\, z_j + \lambda_H = 0 \,,$$

or

$$(\mu_H z_j^2 - \mu_H z_j + \mu_L z_j) j = c\mu_H z_j^2 - (c\mu_H + \lambda_H) z_j + \lambda_H \,. \tag{6.16}$$

Note that the left-hand-side of (6.16) equals zero if and only if $z_j = 1 - \frac{\mu_L}{\mu_H}$. But from Lemma 9 we know that $z_j > 0$, so $\mu_H \leq \mu_L$ implies that the left-hand-side cannot equal zero. Similarly, the right-hand-side of (6.16) equals zero if and only if $z_j = \frac{\lambda_H}{c\mu_H}$, and equating the two expressions for z_j, we see that both sides equal zero, if and only if $\lambda_H = c(\mu_H - \mu_L)$ as well.

Thus if either $\mu_H \leq \mu_L$ or $\lambda_H \neq c(\mu_H - \mu_L)$, then neither of the two sides of (6.16) can equal zero, and we can solve (6.16) for j to show

$$j \;=\; \frac{c\mu_H z_j^2 - (c\mu_H + \lambda_H) z_j + \lambda_H}{\mu_H z_j^2 - \mu_H z_j + \mu_L z_j} \,. \tag{6.17}$$

Note that each z_j determines a j, and $z_j = z_k \;\rightarrow\; j = k$. Thus, it must be the case that $z_j \neq z_k$ for $j \neq k$.

Second, we show that for any $0 \leq j \leq c$, the generating function (6.15) satisfies the following form

$$f_j(z) = \sum_{k=j}^{c} \frac{a_{j,k}}{1 - z_k z} \tag{6.18}$$

and obtains the probabilities defined in (4.14). We shall prove that (4.14) and (6.18) hold by induction on j.

Note that the balance equations (4.8) are defined for $i + j > c$. Therefore, we let $q = i + j - c$, we multiply each of the terms in (4.8) by z^{q+1}, and we sum over $q \geq 1$ to derive $f_j(z)$.

When $j = c$ it is clear from (4.9) that $\xi(q, c) = a_{c,c} z_c^q$, where $a_{c,c} = \xi(0, c)$ and $z_c = p_1(c)$. Thus (4.14) holds. From (6.15), the definition of $f_j(z)$, it also follows that (6.18) holds.

Now suppose (4.14) and (6.18) hold for $j + 1, \ldots, c$.

For j, we multiply both sides of (4.8) by z^{q+1} and then sum over $q \geq 1$. For the left-hand-side, we obtain

$$\sum_{q=1}^{\infty} z^{q+1} \xi(c - j + q, j) \;=\; z \sum_{q=1}^{\infty} z^q \xi(c - j + q, j) \;=\; z \left[f_j(z) - \xi(c - j, j) \right] \,. \tag{6.19}$$

For the first term on the right hand size we have

$$\begin{aligned} \sum_{q=1}^{\infty} z^{q+1} p_1 \xi(c - j + q - 1, j) \;&=\; p_1 z^2 \sum_{q=1}^{\infty} z^{q-1} \xi(c - j + q - 1, j) \\ &=\; p_1 z^2 \sum_{q=0}^{\infty} z^q \xi(c - j + q, j) \;=\; p_1 z^2 f_j(z) \,. \end{aligned} \tag{6.20}$$

For the second term we have

$$\begin{aligned} \sum_{q=1}^{\infty} z^{q+1} p_2 \xi(c - j + q + 1, j) \;&=\; p_2 \sum_{q=2}^{\infty} z^q \xi(c - j + q, j) \\ &=\; p_2 \left[f_j(z) - \xi(c - j, j) - z\xi(c - j + 1, j) \right] \,. \end{aligned} \tag{6.21}$$

Finally, for the last term we have

$$\begin{aligned}
\sum_{q=1}^{\infty} z^{q+1} p_3 \xi(c-j+q, j+1) &= p_3 \sum_{q=2}^{\infty} z^q \xi(c-(j+1)+q, j+1) \\
&= p_3 \left[f_{j+1}(z) - \xi(c-(j+1), j+1) - z\xi(c-(j+1)+1, j+1) \right] \\
&= p_3 \left[f_{j+1}(z) - \xi(c-j-1, j+1) - z\xi(c-j, j+1) \right] . \qquad (6.22)
\end{aligned}$$

Setting (6.19) equal to (6.20)–(6.22) we then have

$$\begin{aligned}
z\left[f_j(z) - \xi(c-j,j)\right] &= p_1 z^2 f_j(z) \; + \; p_2 \left[f_j(z) - \xi(c-j,j) - z\xi(c-j+1,j)\right] \\
&\quad + p_3 \left[f_{j+1}(z) - \xi(c-j-1,j+1) - z\xi(c-j,j+1)\right] , \qquad (6.23)
\end{aligned}$$

and we collect $f_j(z)$ terms on the left-hand-side and terms by degree of z on the right-hand-side so that

$$\begin{aligned}
(z - p_1 z^2 - p_2) f_j(z) &= \left[\xi(c-j,j) - p_2\xi(c-j+1,j) - p_3\xi(c-j,j+1)\right] z \\
&\quad - \left[p_2 \xi(c-j,j) + p_3 \xi(c-j-1,j+1)\right] \; + \; p_3 f_{j+1}(z) \qquad (6.24)
\end{aligned}$$

or

$$\begin{aligned}
&f_j(z) \\
&= \frac{\left[p_2\xi(c-j,j) + p_3\xi(c-j-1,j+1)\right] + \left[p_2\xi(c-j+1,j) + p_3\xi(c-j,j+1) - \xi(c-j,j)\right]z - p_3 f_{j+1}(z)}{p_1 z^2 - z + p_2} , \\
&= \frac{\left[p_2\xi(c-j,j) + p_3\xi(c-j-1,j+1)\right] + \left[p_2\xi(c-j+1,j) + p_3\xi(c-j,j+1) - \xi(c-j,j)\right]z - p_3 f_{j+1}(z)}{p_2(1-z_j z)(1-z'_j z)} , \\
& \qquad (6.25)
\end{aligned}$$

where the denominator in (6.25) follows from the fact that $g(j,z)$ is quadratic so that $z_j z'_j = \frac{p_1}{p_2}$, $z_j + z'_j = \frac{1}{p_2}$, and

$$p_2(1-z_j z)(1-z'_j z) \;=\; p_2 z_j z'_j z^2 - p_2(z_j + z'_j) z + p_2 \;=\; p_2 \frac{p_1}{p_2} z^2 - p_2 \frac{1}{p_2} z + p_2 \;=\; p_1 z^2 - z - p_2 \, .$$

Note also that the last term in the numerator of (6.25), $f_{j+1}(z) = \sum_{k=j+1}^{c} \frac{a_{j+1,k}}{1-z_k z}$ by the induction assumption. Thus,

$$f_j(z) \;=\; \frac{P(z)}{(1-z_j z)\left(1 - z'_j z\right) \prod_{k=j+1}^{c} (1 - z_k z)} ,$$

where $P(z)$ is a polynomial of order $c - j + 1$ in z, and because $z_j \neq z_k$ for all $j \neq k$, the partial fraction expansion of $f_j(z)$ obtains

$$f_j(z) = \sum_{k=j}^{c} \frac{a_{j,k}}{1 - z_k z} + \frac{a'_j}{1 - z'_j z} \qquad (6.26)$$

(see §I.2 in Kleinrock [25]).

In turn, we can invert the $z-$transform (6.26) to show $\xi(c-j+q,j) = \sum_{k=j}^{c} a_{j,k} z_k^q + a'_j (z'_j)^q$. (See entry 6 of Table I.2 in Kleinrock [25].) Furthermore, because $\lim_{q\to\infty} \xi(c-j+q,j) = 0$ and $z'_j \geq 1$, we must have $a'_j = 0$. Hence (6.18) and (4.14) hold for $j+1$ as well.

Third, we explicitly calculate the coefficients found in (4.15)-(4.17). For $a_{c,c}$ we note that (4.9) and (6.14) imply (4.15). For $0 \leq j < k \leq c$, by (6.26) and (6.25), we have

$$a_{j,k} \;=\; f_j(z)(1-z_k z)|_{z=1/z_k} \;=\; \frac{-p_3 a_{j+1,k}}{p_2(1-z_j/z_k)(1-z'_j/z_k)}\,,$$

which gives us (4.17). Finally, for $0 \leq j \leq c-1$, by (6.26) and (6.25), we have

$$a_{j,j} \;=\; f_j(z)(1-z_j z)|_{z=1/z_j}$$

$$= \frac{[p_2\xi(c-j,j) + p_3\xi(c-j-1,j+1)] \;+\; [p_2\xi(c-j+1,j) - \xi(c-j,j) + p_3\xi(c-j,j+1)]/z_j \;-\; p_3 f_{j+1}(1/z_j)}{p_2(1-z'_j/z_j)}$$

or

$$\begin{aligned} a_{j,j} \;\;=\;\; & \frac{1-p_2 z_j}{p_2(z'_j-z_j)}\xi(c-j,j) \;-\; \frac{1}{z'_j-z_j}\xi(c-j+1,j) \;-\; \frac{p_3 z_j}{p_2(z'_j-z_j)}\xi(c-j-1,j+1) \\ & -\; \frac{p_3}{p_2(z'_j-z_j)}\xi(c-j,j+1) \;+\; \frac{p_3 z_j}{p_2(z'_j-z_j)} \sum_{k\geq j+1} \frac{a_{j+1,k}}{1-z_k/z_j}\,. \end{aligned} \tag{6.27}$$

Similarly, for $0 \leq j \leq c$ we have

$$a'_j \;=\; f_j(z)(1-z'_j z)\;|_{z=1/z'_j}$$

$$= \frac{[p_2\xi(c-j,j) + p_3\xi(c-j-1,j+1)] + [p_2\xi(c-j+1,j) + p_3\xi(c-j,j+1) - \xi(c-j,j)]/z'_j - p_3 f_{j+1}(1/z'_j)}{p_2(1-z_j/z'_j)},$$

or

$$\begin{aligned} a'_j \;\;=\;\; & \frac{p_2 z'_j - 1}{p_2(z'_j-z_j)}\xi(c-j,j) \;+\; \frac{1}{z'_j-z_j}\xi(c-j+1,j) + \frac{p_3 z'_j}{p_2(z'_j-z_j)}\xi(c-j-1,j+1) \\ & +\; \frac{p_3}{p_2(z'_j-z_j)}\xi(c-j,j+1) \;-\; \frac{p_3 z'_j}{p_2(z'_j-z_j)} \sum_{k\geq j+1} \frac{a_{j+1,k}}{1-z_k/z'_j}\,. \end{aligned} \tag{6.28}$$

Then because $a'_j = 0$, we can rearrange terms to show that

$$\begin{aligned} & -\frac{1}{z'_j-z_j}\xi(c-j+1,j) - \frac{p_3}{p_2(z'_j-z_j)}\xi(c-j,j+1) \\ & \quad = \frac{p_2 z'_j - 1}{p_2(z'_j-z_j)}\xi(c-j,j) + \frac{p_3 z'_j}{p_2(z'_j-z_j)}\xi(c-j-1,j+1) - \frac{p_3 z'_j}{p_2(z'_j-z_j)} \sum_{k\geq j+1} \frac{a_{j+1,k}}{1-z_k/z'_j}\,. \end{aligned}$$

We can then use this equality to substitute for the left-hand-side terms that appear in (6.27) and to simplify as follows:

$$\begin{aligned}
a_{j,j} &= \frac{1-p_2 z_j}{p_2(z'_j - z_j)}\xi(c-j,j) - \frac{p_3 z_j}{p_2(z'_j - z_j)}\xi(c-j-1,j+1) + \frac{p_3 z_j}{p_2(z'_j - z_j)}\sum_{k\geq j+1}\frac{a_{j+1,k}}{1-z_k/z_j} \\
&\quad + \frac{p_2 z'_j - 1}{p_2(z'_j - z_j)}\xi(c-j,j) + \frac{p_3 z'_j}{p_2(z'_j - z_j)}\xi(c-j-1,j+1) - \frac{p_3 z'_j}{p_2(z'_j - z_j)}\sum_{k\geq j+1}\frac{a_{j+1,k}}{1-z_k/z'_j} \\
&= \xi(c-j,j) + \frac{p_3}{p_2}\xi(c-j-1,j+1) + \frac{p_3 z_j}{p_2(z'_j - z_j)}\sum_{k\geq j+1}\frac{a_{j+1,k}}{1-z_k/z_j} - \frac{p_3 z'_j}{p_2(z'_j - z_j)}\sum_{k\geq j+1}\frac{a_{j+1,k}}{1-z_k/z'_j}.
\end{aligned}$$

Finally, using (6.14) to substitute for $\xi(c-j,j)$ and $\xi(c-j-1,j+1)$ we obtain (4.16).

Part (ii) First, note that when $\mu_H > \mu_L$ and $\lambda_H = c(\mu_H - \mu_L)$, then both sizes of (6.16) equal zero for any z_j. Therefore, we may let $z_j = 1 - \mu_L/\mu_H \in (0,1)$ for all $j = 0,\ldots,c$. Again, because $g(j,z)$ is quadratic, when $j < c$ we have $z_j z'_j = \frac{p_1}{p_2}$ so that

$$z'_j = \frac{p_1}{p_2}/z_j = \frac{\lambda_H}{(c-j)\mu_H}\cdot\frac{\mu_H}{\mu_H - \mu_L} = \frac{c(\mu_H - \mu_L)}{(c-j)\mu_H}\cdot\frac{\mu_H}{\mu_H - \mu_L} = \frac{c}{c-j}.$$

Second, we show by induction that (4.18) holds for the relevant probabilities $\xi(i,j)$. Again, for $j = c$ (4.9) implies that $\xi(q,c) = a_{c,c}z_c^q$, where $a_{c,c} = \xi(0,c)$ and $z_c = p_1(c)$. Similarly, (6.25) holds for the generating function $f_j(z)$, and

$$f_j(z) = \frac{Q(z)}{\left(1-\frac{c}{c-j}z\right)(1-z^*z)^{c-j+1}},$$

where $Q(z)$ is a polynomial of order $c-j+1$ in z. Here, the partial fraction expansion of $f_j(z)$ is

$$f_j(z) = \sum_{k=1}^{c-j+1}\frac{a_{j,k}}{(1-z^*z)^k} + \frac{b'_j}{1-\frac{c}{c-j}z}, \tag{6.29}$$

(see §I.2 in Kleinrock [25]), where $b'_j = f_j(z)\left(1-\frac{c}{c-j}z\right)\big|_{z=(c-j)/c}$, and the computation of $a_{j,k}$ involves taking the $(c-j+1-k)-$th derivative of $f_j(z)\,(1-z^*z)^{c-j+1}$ and then evaluating at $1/z^*$. It will not be detailed here.

Again, inverting the $z-$transform (6.29) we obtain: $\xi(c-j+q,j) = \sum_{k=1}^{c-j+1}a_{j,k}\binom{q+k-1}{k-1}z^{*q} + b'_j\left(\frac{c}{c-j}\right)^q$. (See entries 6 and 13 of Table I.2 in Kleinrock [25].) As before, we have $b'_j = 0$ and $\xi(c-j+q,j) = \sum_{k=1}^{c-j+1}a_{j,k}\binom{q+k-1}{k-1}z^{*q}$. □