Probability in the Engineering and Informational Sciences, **29**, 2015, 483–506.
doi:10.1017/S0269964815000157

OPTIMAL SWITCHING ON AND OFF THE ENTIRE SERVICE CAPACITY OF A PARALLEL QUEUE

Eugene A. Feinberg
Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY 11794, USA
E-mail: eugene.feinberg@stonybrook.edu

Xiaoxuan Zhang
IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA
E-mail: zhangxiaoxuan@live.com

This paper studies optimal switching on and off of the entire service capacity of an $M/M/\infty$ queue with holding, running and switching costs. The running costs depend only on whether the system is on or off, and the holding costs are linear. The goal is to minimize average costs per unit time. The main result is that an average-cost optimal policy either always runs the system or is an (M,N)-policy defined by two thresholds M and N, such that the system is switched on upon an arrival epoch when the system size accumulates to N and is switched off upon a departure epoch when the system size decreases to M. It is shown that this optimization problem can be reduced to a problem with a finite number of states and actions, and an average-cost optimal policy can be computed via linear programming. An example, in which the optimal (M,N)-policy outperforms the best $(0,N)$-policy, is provided. Thus, unlike the case of single-server queues studied in the literature, $(0,N)$-policies may not be average-cost optimal.

1. INTRODUCTION

This paper studies optimal control of a parallel $M/M/\infty$ queue with Poisson arrivals and an unlimited number of independent identical servers with exponentially distributed service times. The cost to switch the system on is s_1, and the cost to switch the system off is s_0. The other costs include the linear holding cost h for each unit of time that a customer spends in the system, the running cost c_1 per unit time when the system is on, and the idling cost c_0 per unit time when the system is off. It is assumed that $s_0, s_1 \geq 0$, $s_0 + s_1 > 0$, $h > 0$, and $c_1 > c_0$. Denote $c = c_1 - c_0$. Without loss of generality, let $c_0 = 0$ and $c_1 = c > 0$. The goal is to minimize average costs per unit time.

The main result of this paper is that either the policy that always keeps the system on is average-cost optimal or, for some integers M and N, where $N > M \geq 0$, the so-called (M,N)-policy is average-cost optimal. The (M,N)-policy switches the running system off when the number of customers in the system is not greater than M and it switches the

 0269-9648/15 $25.00

idling system on when the number of customers in the queue reaches or exceeds N. It is shown in this paper that this optimization problem can be reduced to a problem with finite number of states and actions and an average-cost optimal policy can be computed via linear programming. An example, when the best $(0, N)$-policy is not optimal, is provided.

From an intuitive point of view, the reason that (M, N)-policies can outperform $(0, N)$-policies for our problem is because the system running cost rate c does not depend on the number of occupied servers. If there are n customers in the system, then the running cost rate per customer is c/n. When n is small, it may be cost-inefficient to run the system because the running cost rate per customer is high, and the system should be switched off. Similarly to the single-server case, the non-zero switching costs cause hysteretic behavior of an optimal policy and therefore $M < N$.

Studies of control problems for queues started around 50 years ago, and one of the first papers on this topic, Yadin and Naor [38], dealt with switching on and off the server of a single-server queue. Heyman [16] showed the optimality of a $(0, N)$-policy, which is usually called an N-policy, for $M/G/1$ queues. Sobel [37] studied (M, N)-policies for $GI/G/1$ queues. The early results on switching servers in single-server queues led to two relevant research directions:

(i) Optimality of $(0, N)$-policies or their ramifications under very general assumptions such as batch arrivals, start-up and shut-down times and costs, nonlinear holding costs, known workload and so on; see Lee and Srinivasan [27],Federgruen and So [8], Altman and Nain [1], Denardo, Feinberg, and Kella [7], and Feinberg and Kella [11];

(ii) Decomposition results for queues with vacations; see Fuhrmann and Cooper [12], Hofri [17], Shanthikumar [36], Kella [21], and Kella and Whitt [22].

As for general multi-server parallel queues, switching on and off individual servers for a parallel queue is a more difficult problem. Even for an $M/M/n$ queue, there is no known description of an optimal switching policy for individual servers when $n > 2$; see Bell [2,3], Rhee and Sivazlian [34], and Igaki [19]. Studies of stationary distributions and performance evaluations for parallel queues with vacations (Levy and Yechiali [28], Huang et al. [18], Kao and Narayanan [20], Browne and Kella [4], Chao and Zhao [5] and Li and Alfa [29]) usually assume that vacations begin when the system is empty. Observe that, if vacations start when the system becomes empty and end simultaneously for all the servers, the model describes a particular case of switching the entire service capacity of the system on and off. Browne and Kella [4] studied $M/G/\infty$ queues with vacations and, for a model switching and linear holding costs, described how to compute the best $(0, N)$-policy for switching on and off the entire service capacity.

This research is motivated by two observations: (i) the problem of switching on and off the entire service capacity of the facility has an explicit solution described in this paper, while there is no known explicit solution for problems with servers that can be switched on and off individually, and (ii) with the development of internet and high-performance computing, many applications behave in the way described in this paper. For example, consider a service provider that uses cloud computing and pays for the time the cloud is used. When there are many service requests, it is worth paying for using the cloud, and when there is a small number of service requests, it may be too expensive to use the cloud. This paper analyzes such a situation and finds an optimal solution. Many papers model cloud computing facilities as multi-server queues; see Mazzucco, Dyachuk, and Deters [30] and Khazaei, Misic, and Misic [23]. Mazzucco et al. [30] studies the revenue management problem from the perspective of a cloud computing provider and investigates the resource allocation via dynamically powering the servers on or off. There can be a huge number of servers in a cloud computing center, typically of the order of hundreds or thousands; see, e.g., Greenberg et al. [13]. Given that the number of servers is large and tends to increase over time with the development of new technologies, it is natural to model controlling of the facility as an $M/M/\infty$ queue rather than an $M/M/n$ queue if this leads to analytical advantages. Here we study a model based on an $M/M/\infty$ queue and find an optimal solution.

In addition to cloud computing, another example comes from the application to software maintenance. Kulkarni et al. [26] studied the software maintenance problem as a control problem for a queue formed by software maintenance requests generated by software bugs experienced by customers. Once a customer is served and the appropriate bug is fixed in the new software release or patch, it also provides solutions to some other customers in the queue and these customers are served simultaneously. In Kulkarni et al. [26], it was assumed that the number of customers leaving the queue at a service completion time has a binomial distribution. This problem was modeled in Kulkarni et al. [26] as an optimal switching problem for an $M/G/1$ queue in which a binomially distributed number of customers depending on the queue size are served each time, and the best $(0, N)$-policy was found. Here we observe that after an appropriate scaling, the software maintenance problem with exponential service times and the optimal switching problem for an $M/M/\infty$ queue have the same fluid approximations. So, the result on average-optimality of (M, N)-policies described here provides certain insights to the software maintenance problem studied in Kulkarni et al. [26].

There are two main obstacles in the analysis of the $M/M/\infty$ switching problem compared to a single-server one. First, the service intensities are unbounded, and therefore the standard reduction of continuous-time problems to discrete time via uniformization can not be applied. Second, there are significantly more known decomposition and performance analysis results for single-server queues than for parallel queues and, in particular, we are not aware of such results for $M/M/\infty$ queues with vacations that can start when the queue is not empty. The first obstacle is resolved by reducing the discounted version of the problem to negative dynamic programming instead of to discounted dynamic programming. The second obstacle is resolved by solving a discounted problem for the system that cannot be switched off. This problem is solved by using optimal stopping, where the stopping decision corresponds to starting the servers, and its solution is used to derive useful inequalities and to reduce the problem for the original $M/M/\infty$ queue to a control problem of a semi-Markov process with finite state and action sets representing the system being always on when the number of customers exceeds a certain level.

The optimal switching problem for an $M/M/\infty$ queue is modeled in Section 2 as a Continuous-Time Markov Decision Process (CTMDP) with unbounded transition rates. Such a CTMDP cannot be reduced to discrete time via uniformization; see, e.g., Guo, Hernández-Lerma, and Prieto-Rumeau [15], Piunovskiy and Zhang [32]. We analyze the problem by studying the total expected discounted costs and applying the vanishing discount rate approach.

Section 3 studies expected total discounted costs. Such a CTMDP can be reduced to a discrete-time problem with the expected total costs; see Feinberg [10], Piunovskiy and Zhang [32]. Since transition rates are unbounded, expected total costs for the discrete-time problem cannot be presented as expected total discounted costs with the discount factor smaller than 1. However, since all the costs are nonnegative, the resulting discrete-time problem belongs to the class of negative MDPs that deal with minimizing expected total nonnegative costs, which is equivalent to maximizing expected total non-positive rewards. For this negative MDP we derive the optimality equation, show that the value function is finite, and establish the existence of stationary discounted-cost optimal policies; see Theorem 1.

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

Section 3.2 investigates the discounted total-cost problem limited to the policies that never switch the running system off. By using the fact that the number of customers in an $M/G/\infty$ queue at each time has a Poisson distribution (see Ross [35, p. 70]), for this problem we compute in Theorem 3 and in Corollary 1 the discounted-cost optimal policy and the optimal value. This is done by analyzing the optimality equation for an optimal stopping problem with stopping, in fact, corresponding to the decision to start the system. The optimal policy is defined by an explicitly computed number n_α, such that the system should be switched on as soon as the number of customers is greater than or equal to n_α, where $\alpha > 0$ is the discount rate. The function n_α is increasing in α and therefore bounded in $\alpha \in (0, \alpha^*]$ for each $\alpha^* \in (0, \infty)$. In Section 3.3, the problem with the expected discounted total costs is reduced to a problem with finite state and action sets by showing in Lemma 7 that the system should always be on, if the number of customers is greater than or equal to n_α. In Section 4, by using the vanishing discount rate arguments, we prove the existence of stationary average-cost optimal policies and describe the optimal (M, N)-policies in Theorem 5. A linear program (LP) for their computation is provided in Section 5. Section 6 deals with computing the best $(0, N)$-policies and showing that they may not be optimal.

2. PROBLEM FORMULATION

We model the above described control problem for an $M/M/\infty$ queue as a CTMDP with a countable state space and a finite number of actions; see Kitaev and Rykov [24] and Guo and Hernández-Lerma [14]. In general, such a CTMDP is defined by the tuple $\{Z, A, A(z), q, c\}$, where Z is a countable state space, A is a finite action set, $A(z)$ are sets of actions available in states $z \in Z$, and q and c are transition and cost rates, respectively. A general policy can be time-dependent, history-dependent, and at a jump epoch the action that controls the process is the action selected at the previous state; see Kitaev and Rykov [24, p. 138].

Without going into details, recall that a trajectory of a CTMDP is a sequence $\omega = \{(T_n, X_n)\}_{n=0}^{\infty}$, where $0 = T_0 < T_1 \le \cdots \le T_{n+1} \in (0, \infty]$, and $T_n < T_{n+1}$, if $T_n < \infty$, are jump epochs and $X_n \in Z$ is the state at time $t \in [T_n, T_{n+1})$, $n = 0, 1, \ldots$. In particular, for our problem T_n are either arrival or departure epochs, $n = 0, 1, \ldots$. For an $M/M/\infty$ queue, arrivals occur according to a Poisson process, and the number of departures by time t is bounded from above by the number of arrivals plus the number of customers in the system at time 0. Therefore, the number of jumps $n(t)$ up to time t is a.s. finite, where $n(t) = \sup\{n : T_n < t\}$, $0 < t < \infty$. Let $H = \cup_{n=1}^{\infty} (Z \times (0, \infty])^n$. A policy π is a function $\pi(\omega, t)$ with values in A such that for each $t > 0$

$$\pi(\omega, t) = \pi^*(X_0, T_1, X_1, \ldots, T_{n(t)}, X_{n(t)}, t - T_{n(t)}), \quad (1)$$

where $\pi^* : H \to A$ such that $\pi^*(X_0, T_1, X_1, \ldots, T_n, X_n, s) \in A(X_n)$, and the function π^* is Borel in $(T_1, \ldots, T_n, s)$ for each $n = 1, 2, \ldots$, and Borel in s for $n = 0$. As is usual in probability theory and in the theory of stochastic processes, we often omit the variable ω. For example, we write $\pi(t)$ instead of $\pi(\omega, t)$.

An initial state $z \in Z$ and a policy π define a stochastic process z_t. Expectations for this stochastic process are denoted by E_z^π. Let $C(t)$ be the cumulative costs incurred during the time interval $[0, t]$. For $\alpha > 0$, the expected total discounted cost is

$$V_\alpha^\pi(z) = E_z^\pi \int_0^\infty e^{-\alpha t} dC(t), \quad (2)$$

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

and the average cost per unit time is

$$v^\pi(z) = \limsup_{t\to\infty} t^{-1} E_z^\pi C(t). \tag{3}$$

Let

$$V_\alpha(z) = \inf_\pi V_\alpha^\pi(z), \tag{4}$$

$$v(z) = \inf_\pi v^\pi(z). \tag{5}$$

A policy π is called discounted-cost optimal if $V_\alpha^\pi(z) = V_\alpha(z)$ for all initial states $z \in Z$. A policy π is called average-cost optimal if $v^\pi(z) = v(z)$ for all initial states $z \in Z$.

For our problem the states of the system change only at arrival and departure epochs, which we call jump epochs. The state of the system at time $t \geq 0$ is $z_t = (x_t, \delta_t)$, where x_t is the number of customers in the system at time t, and δ_t is the status of the servers that an arrival or departure saw at the last jump epoch. If $\delta_t = 0$, the servers at the last jump epoch during the interval $[0, t]$ were off, and, if $\delta_t = 1$, they were on. In particular, if the last jump epoch was a departure, $\delta_t = 1$. If the last jump epoch was an arrival, then $\delta_t = 1$, if the last arrival saw the servers being on, and $\delta_t = 0$ otherwise. The initial state $z_0 = (x_0, \delta_0)$ is given, and $t = 0$ is assumed to be a jump epoch. According to the definition of δ_t, the functions x_t and δ_t are right-continuous.

The state space is $Z = \mathbb{N} \times \{0, 1\}$, where $\mathbb{N} = \{0, 1, \ldots\}$, and the action set is $A = \{0, 1\}$, with 0 meaning that the system is off and 1 meaning that the system is on. At time t the state is $z_t = (x_t, \delta_t)$ with the variables x_t and δ_t described in the previous paragraph. The action sets $A(z) = A$ for all $z \in Z$. A stationary policy chooses actions deterministically at jump epochs and follows them until the next jump. In addition, the choice of an action depends only on the state of the system $z = (x, \delta)$, where x is the number of customers in the system and $\delta \in \{0, 1\}$ is the status of the system prior to the last jump.

The transition rate from a state $z = (i, \delta)$ with an action $a \in A$ to a state $z' = (j, a)$, where $j \neq i$, is $q(z'|z, a) = q(j|i, a)$, with

$$q(j|i,a) = \begin{cases} \lambda, & \text{if } j = i + 1; \\ i\mu, & \text{if } i > 0,\ a = 1,\ j = i - 1; \\ 0, & \text{otherwise}; \end{cases} \tag{6}$$

where λ is the intensity of the arrival process and μ is the service rate of individual servers. At state $z = (i, \delta)$, define $q(z, a) = q(i, a) = \sum_{j \in \mathbb{N}\setminus\{i\}} q(j|i, a)$ and $q(z|z, a) = q(i|i, a) = -q(i, a)$.

The costs include the linear holding cost h per unit time that a customer spends in the system, the running cost c per unit time when the system is on, the start-up cost s_1, and the shut-down cost s_0, where $h, c > 0$, $s_0, s_1 \geq 0$, and $s_0 + s_1 > 0$. At state $z = (i, \delta)$, if action 1 is taken, the cost rate is $hi + c$; if action 0 is taken, the cost rate is hi. At state $z = (i, \delta)$, if action 1 is taken, the instantaneous cost $(1 - \delta)s_1$ is incurred; if action 0 is taken, the instantaneous cost δs_0 is incurred. The presence of instantaneous switching costs s_0 and s_1 complicates the situation because standard models of CTMDPs deal only with cost rates. However, since $s_0 + s_1 > 0$, the cost function $C(t)$ can be written explicitly for this problem. Let $N(t)$ be the number of times the system's status (on or off) changes during the time interval $[0, t]$.

Fix a policy π and an initial state z. If with positive probability $N(t) = \infty$ for some $t < \infty$, then with positive probability $C(t) = \infty$ since $s_0 + s_1 > 0$ and the number of switching times is infinite with positive probability. Thus, in view of (2) and (3), $V_\alpha^\pi(z) = v^\pi(z) = \infty$ for all $\alpha > 0$.

Suppose that $N(t) < \infty$ a.s. for some $t < \infty$. Let $0 \le t_1 < t_2 < \ldots < N(t)$ be the times, when the system is switched on or off during the time interval $[0, t]$. The function $\pi(u)$ is a.s. piecewise constant on $[0, t]$; here we write the argument u instead of the argument t used in (1). It is impossible that $\pi(u-) = \pi(u+) \neq \pi(u)$. Such a possibility would mean that the system is switched on and off at time u, but at most one switching is allowed at each time instance. For each trajectory ω, it does not matter whether $\pi(u) = \pi(u-)$ or $\pi(u) = \pi(u+)$ at the points u of discontinuity of $\pi(u)$ for $u \in [0, t]$. The values of $\pi(u)$ at discontinuity points do not affect the costs and future states of the system, except for the discontinuity points of $\pi(u)$ at which the process jumps (that is, arrivals or departures take place). However, the probability of an arrival or departure at any time instance $u > 0$ is 0. Therefore, without loss of generality, we define the piecewise constant function $\pi(u)$ in u to be left continuous in u. Then

$$C(t) = \int_0^t (hx_u + c\pi(u))\, du + \sum_{n=1}^{N(t)} s_{\pi(t_n+)} |\pi(t_n+) - \pi(t_n)|.$$

As explained in Section 3.1, it is sufficient to use policies that change actions only at arrival and departure times, and for such policies $N(t) < \infty$ a.s. when $t < \infty$.

3. DISCOUNTED-COST CRITERION

In this section, we study the expected total discounted cost criterion. In Section 3.1, we reduce the CTMDP to the discrete-time MDP with the expected total costs, provide the optimality equation, and prove the existence of stationary optimal policies. In Section 3.2, we explicitly solve the version of the problem when the running servers cannot be switched off. For this version of the problem, the question is: when should the servers be switched on, if they are off at time 0? Theorem 3 and Corollary 1 describe for each discount rated $\alpha > 0$ the number n_α, such that an optimal decision is to switch the inactive system on, if the number of waiting customers is greater than or equal to n_α. This solution is used in Section 3.3 to provide estimates for the discounted version of the original problem and establish the properties of its optimal policies. Theorem 4 states several properties of discounted-cost optimal policies. In particular, it is optimal to keep or turn the servers on, if the number of customers in the system is greater than or equal to n_α. This theorem reduces the discounted problem to a problem with a finite number of states. Since the function n_α is bounded in $\alpha \in (0, \alpha']$ for each $\alpha' \in (0, \infty)$, Theorem 4 is useful for the reduction of the average-cost problem to a problem with a finite number of states. Other properties of discounted-cost optimal policies stated in Theorem 4 are used in Section 4 to describe the structure of average-cost optimal policies.

3.1. Reduction to Discrete Time and Existence of Stationary Discounted-Cost Optimal Policies

In this subsection, we formulate the optimality equation and prove the existence of a stationary discounted-cost optimal policy. This is done by reducing the problem to discrete time by using the results from Feinberg [10]. After the reduction is described, Theorem 1 provides the optimality equations and the upper bound on the value function. Lemma 1 is needed to establish this bound.

When the system is on and there are i customers, the time until the next jump has an exponential distribution with intensity $q(i,1)=\lambda+i\mu\to\infty$ as $i\to\infty$. Since the jump rates are unbounded, it is impossible to present the problem as a discounted MDP in discrete-time with a discount factor smaller than 1. Thus, we shall present our problem as minimization of the expected total costs. If the decisions are chosen only at jump times, the expected total discounted sojourn time until the next jump epoch is $\tau_\alpha(z,a)=\int_0^\infty(\int_0^t e^{-\alpha u}du)q(z,a)e^{-q(z,a)t}dt=\int_0^\infty e^{-\alpha t}e^{-q(z,a)t}dt=\dfrac{1}{\alpha+q(z,a)}$, and the one-step cost is $C_\alpha(z,a)=|a-\delta|s_a+(hi+ac)\tau_\alpha(z,a)$ with $z=(i,\delta)$. For $\alpha=0$, we denote $\tau_0(z,a)$ and $C_0(z,a)$ as $\tau(z,a)$ and $C(z,a)$, respectively.

By Feinberg [10, Theorem 5.5.6], there exists a stationary discounted-cost optimal policy, the value function $V_\alpha(z)$ satisfies the optimality equation

$$V_\alpha(z)=\min_{a\in A(z)}\{C_\alpha(z,a)+\sum_{z'\in Z}\frac{q(z'|z,a)}{\alpha+q(z,a)}V_\alpha(z')\},\qquad z\in Z, \tag{7}$$

and a stationary policy ϕ is discounted-cost optimal if and only if

$$V_\alpha(z)=C_\alpha(z,\phi(z))+\sum_{z'\in Z}\frac{q(z'|z,\phi(z))}{\alpha+q(z,\phi(z))}V_\alpha(z),\quad z\in Z. \tag{8}$$

In view of Feinberg [10, Theorem 5.5.6], these conclusions also hold for the CTMDP with actions that can be changed at any time, but with switching cost rates $s(z,a)=s_a|a-\delta|(\alpha+q(z,a))$, where $z=(i,\delta)$, charged instead of instantaneous switching costs s_0 and s_1. Formulae (7) and (8) imply that the discounted version of the problem is equivalent to finding a policy that minimizes the expected total costs for the discrete-time MDP $\{Z,A,A(z),p_\alpha,C_\alpha\}$ with substochastic transition probabilities $p_\alpha(z'|z,a)=q(z'|z,a)/(\alpha+q(z,a))$ and with one-step costs $C_\alpha(z,a)$, where $\alpha>0$.

As mentioned above, classic CTMDPs do not deal with instantaneous switching costs described in the previous section. However, if we replace the instantaneous cost rates s_a, $a\in\{0,1\}$, with the cost rates $s(z,a)$ defined above, then a stationary optimal policy for the problem with switching cost rates $s(z,a)$ is also optimal for the original problem with instantaneous switching costs s_a. To see that this is true, first observe that the expected total discounted cost until the next jump are the same for models with cost rates $s(z,a)$ and instantaneous costs s_a. In both cases, this cost is $s_a|a-\delta|$, where δ is the status of servers (on or off) and a is the chosen action. For an arbitrary policy, the expected total discounted cost until the next jump can either decrease or remain unchanged, if instantaneous switching costs s_a are replaced with the switching cost rates $s(z,a)$.

Indeed, let a policy use an action a during the first s units of time it spends at state $z=(x,\delta)$, then it switches to action b, and then it either uses action b or switches between actions a and b. The total expected discounted switching cost until the first jump is $s_a|a-\delta|+s_be^{-(\alpha+q(z,a))s}$ or higher. Recall that each state has two actions: 0 and 1. Thus $\{a,b\}=\{0,1\}$.

To compute switching costs incurred until the first jump for a model with switching cost rates $s(z,a)$, let us interpret discounting as a jump intensity to the absorbing state. So, the first jump can be caused either by a transition to an absorbing state or a transition to the next state. As follows from Feinberg [9, Theorem 1], the total expected discounted switching cost until the next jump is $s_a|a-\delta|p_a+s_b|b-\delta|p_b$, where p_a (p_b) is the probability that the first jump takes place when the CTMDP is controlled by the action a (b). Since $1\geq p_a$, $e^{-(\alpha+q(z,a))s}\geq p_b$, and $|b-\delta|\in\{0,1\}$, the expected discounted costs until the next jump for the model with instantaneous switching costs are greater than or equal to the similar costs for the model with cost rates. Thus, a stationary discounted-cost optimal policy for the problem with the switching cost rates $s(z,a)$ is also discounted-cost optimal for the original problem with instantaneous switching costs, and the optimality equation (7) is also the optimality equation for the original problem with the goal to minimize the expected total discounted costs.

The following lemma introduces the formula for the expected total discounted costs under the policy that always runs the system. This formula provides an upper bound for the value function V_α and, in addition, it shows that the value function takes finite values.

LEMMA 1: *Let ϕ be a policy that always runs the system. For all $i = 0, 1, \ldots,$*

$$V_\alpha^\phi(i,\delta) = (1-\delta)s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha} < \infty. \tag{9}$$

PROOF: $V_\alpha^\phi(i,0) = s_1 + V_\alpha^\phi(i,1)$, or equivalently, $V_\alpha^\phi(i,\delta) = (1-\delta)s_1 + V_\alpha^\phi(i,1)$. Observe that

$$V_\alpha^\phi(0,1) = E\left[\int_0^\infty e^{-\alpha t}\left(hX_0(t)+c\right)\,dt\right] = hE\left[\int_0^\infty e^{-\alpha t}X_0(t)\,dt\right] + \frac{c}{\alpha} = \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}, \tag{10}$$

where $X_0(t)$ is the number of busy servers at time t if at time 0 the system is empty. The last equality in (10) holds because, according to Page 70 in Ross [35], $X_0(t)$ has a Poisson distribution with the mean $\lambda\int_0^t e^{-\mu u}du = \frac{\lambda}{\mu}\left(1-e^{-\mu t}\right)$. Thus,

$$E\left[\int_0^\infty e^{-\alpha t}X_0(t)\,dt\right] = \int_0^\infty e^{-\alpha t}\frac{\lambda}{\mu}\left(1-e^{-\mu t}\right)\,dt = \frac{\lambda}{\alpha(\mu+\alpha)}.$$

If at time 0 there are i customers in an $M/M/\infty$ queue and the servers are always on, the total discounted cost is the sum of the total discounted holding cost to serve these i customers and the total discounted cost to run the system and serve future arrivals. Thus,

$$V_\alpha^\phi(i,1) = G_\alpha(i) + V_\alpha^\phi(0,1) = iG_\alpha(1) + V_\alpha^\phi(0,1), \tag{11}$$

where $G_\alpha(i)$ is the expected total discounted holding cost incurred by i customers served in parallel. Since service times are exponential,

$$G_\alpha(1) = E\left[\int_0^\xi e^{-\alpha t}h\,dt\right] = \frac{h}{\mu+\alpha},$$

where $\xi \sim \exp(\mu)$. Formulae (10), (11), and $V_\alpha^\phi(i,0) = s_1 + V_\alpha^\phi(i,1)$ imply (9). ■

We follow the conventions that $V_\alpha(-1,\delta) = 0$, $\sum_\emptyset = 0$, and $\prod_\emptyset = 1$. The following theorem is the main result of this subsection.

THEOREM 1: *For any $\alpha > 0$ the following statements hold:*

(i) *For all $i = 0, 1, \ldots,$*

$$V_\alpha(i,\delta) \le (1-\delta)s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}; \tag{12}$$

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

(ii) *For all* $i = 0, 1, \ldots$ *and for all* $\delta = 0, 1$, *the value function* $V_\alpha(i, \delta)$ *satisfies the discounted-cost optimality equation*

$$V_\alpha(i,\delta) = \min_{a\in\{0,1\}}\left\{C_\alpha((i,\delta),a) + \frac{q(i-1|i,a)}{\alpha+q(i,a)}V_\alpha(i-1,a) + \frac{q(i+1|i,a)}{\alpha+q(i,a)}V_\alpha(i+1,a)\right\}$$
$$= \min\left\{(1-\delta)s_1 + \frac{hi+c}{\alpha+\lambda+i\mu} + \frac{\lambda}{\alpha+\lambda+i\mu}V_\alpha(i+1,1) + \frac{i\mu}{\alpha+\lambda+i\mu}V_\alpha(i-1,1), \delta s_0 + \frac{hi}{\alpha+\lambda} + \frac{\lambda}{\alpha+\lambda}V_\alpha(i+1,0)\right\}; \quad \textbf{(13)}$$

(iii) *There exists a stationary discounted-cost optimal policy, and a stationary policy* ϕ *is discounted-cost optimal if and only if for all* $i = 0, 1, \ldots$ *and for all* $\delta = 0, 1$,

$$V_\alpha(i,\delta) = \min_{\phi(i,\delta)\in\{0,1\}}\left\{C_\alpha((i,\delta),\phi(i,\delta)) + \frac{q(i-1|i,a)}{\alpha+q(i,\phi(i,\delta))}V_\alpha(i-1,\phi(i,\delta)) + \frac{q(i+1|i,a)}{\alpha+q(i,\phi(i,\delta))}V_\alpha(i+1,\phi(i,\delta))\right\}.$$

Proof: Consider the policy ϕ that always runs the system. Then $V_\alpha(i,\delta) \le V_\alpha^\phi(i,\delta)$, and (12) follows from Lemma 1. Statements (ii) and (iii) follow from (7) and (8). ∎

In view of Theorem 1(iii), we consider only stationary policies in the remaining parts of this paper, unless the opposite is specified for a particular policy.

3.2. Discounted-Cost Optimal Policies, when Running Servers Cannot be Switched off

In this subsection, we explicitly solve the problem of finding discounted-cost optimal policies within the class of policies that never turn the running system off. This solution is used later to study the original problem. If each action set $A(i,1)$ is reduced to the singleton $\{1\}$, the class of policies that never turn the system off is the set of all policies for the model with the reduced action sets. Let $U_\alpha(i,\delta)$, $i = 0, 1, \ldots$, be the optimal expected total discounted cost under the policies that never switch the system off.

Theorem 2: *For any* $\alpha > 0$ *the following statements hold:*

(i) *For all* $i = 0, 1, \ldots$,

$$U_\alpha(i,1) = \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha};$$

(ii) *For all* $i = 0, 1, \ldots$, *the value function* $U_\alpha(i,0)$ *satisfies the optimality equation*

$$U_\alpha(i,0) = \min\left\{s_1 + \frac{hi+c}{\alpha+\lambda+i\mu} + \frac{\lambda}{\alpha+\lambda+i\mu}U_\alpha(i+1,1) + \frac{i\mu}{\alpha+\lambda+i\mu}U_\alpha(i-1,1), \frac{hi}{\alpha+\lambda} + \frac{\lambda}{\alpha+\lambda}U_\alpha(i+1,0)\right\}. \quad \textbf{(14)}$$

Proof: (i) For a policy ϕ, that never switches the running system off, $U_\alpha(i,1) = V_\alpha^\phi(i,1)$, and the rest follows from Lemma 1. (ii) Since $U_\alpha(i,0)$ is the optimal discounted cost for the sub-model of the original MDP, it satisfies the discounted-cost optimality equation (7), which implies (14). ∎

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

Definition 1: *For an integer $n \geq 0$, a policy is called an n-full-service policy, if it never switches the running system off, and switches the inactive system on if and only if there are n or more customers in the system. In particular, the 0-full-service policy switches the system on at time 0, if it is off, and always keeps it on. A policy is called a full-service policy, if it is an n-full-service policy for some $n \geq 0$.*

The following theorem implies that a full-service policy is discounted-cost optimal within the class of policies that never switch the running system off.

Theorem 3: *A policy ϕ is discounted-cost optimal within the class of the policies that never switch off the running system if and only if for all $i = 0, 1, \ldots,$*

$$\phi(i,0) = \begin{cases} 1, & \text{if } i > A(\alpha); \\ 0, & \text{if } i < A(\alpha); \end{cases}$$

where

$$A(\alpha) = \frac{(\mu+\alpha)(c+\alpha s_1)}{h\mu}. \tag{15}$$

From an intuitive point of view, $A(\alpha)$ defines the discounted-cost optimal threshold for switching on the system within the class of the policies that never switch off the running system; see Corollary 1. This threshold increases with the increase of the service cost rate c, discount rate α, and cost of switching on the system s_1, and it decreases with the increase of the holding cost rate h and service rate μ. In addition, $A(\alpha) \to c/h$ as $\alpha \to 0$. The explicit computation of $A(\alpha)$ is possible because the expected total discounted costs for the policy that always runs the system has an explicit form; see Lemma 1. Before proving Theorem 3, we introduce the definition of passive policies and some lemmas. In particular, the passive policy never changes the status of the system.

Definition 2: *The policy φ, with $\varphi(i, \delta) = \delta$ for all $i = 0, 1, \ldots$ and for all $\delta = 0, 1$, is called passive.*

According to the following lemma, the passive policy is not discount-cost optimal.

Lemma 2: *For any $\alpha > 0$, the passive policy φ is not discounted-cost optimal within the class of policies that never switch off the running system. Furthermore, $V_\alpha^\varphi(i,0) > U_\alpha(i,0)$ for all $i = 0, 1, \ldots$.*

Proof: For the passive policy φ,

$$V_\alpha^\varphi(i,0) = \sum_{k=0}^{\infty} \left(\frac{\lambda}{\lambda+\alpha}\right)^k \frac{h(i+k)}{\lambda+\alpha} = \frac{hi}{\alpha} + \frac{h\lambda}{\alpha^2}.$$

For the policy ϕ that always runs the system, in view of Lemma 1,

$$V_\alpha^\phi(i,0) = s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}. \tag{16}$$

Thus

$$\begin{aligned} V_\alpha^\varphi(i,0) - V_\alpha^\phi(i,0) &= \left(\frac{hi}{\alpha} + \frac{h\lambda}{\alpha^2}\right) - \left(s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}\right) \\ &= \frac{hi\mu}{\alpha(\mu+\alpha)} + \frac{h\lambda\mu}{\alpha^2(\mu+\alpha)} - s_1 - \frac{c}{\alpha} > 0, \end{aligned}$$

when i is large enough. Let i^* be the smallest natural integer such that the last inequality holds with $i = i^*$. Let the initial state be $(i, 0)$ with $i < i^*$. Consider a policy π that keeps the system off in states $(j, 0)$, $j < i^*$, and switches to a discounted-cost optimal policy, when the number of customers in the system reaches i^*. Then $V_\alpha^\varphi(i,0) > V_\alpha^\pi(i,0) \geq U_\alpha(i,0)$, where the first inequality holds because, before the process hits the state $(i^*, 0)$, the policies φ and π coincide, and, after the process hits the state $(i^*, 0)$, the policy π, which starting from that event coincides with ϕ, incurs lower expected total discounting costs than the passive policy φ. ■

Lemma 3: *Let ψ be the policy that switches the system on at time 0 and keeps it on forever, and π be the policy that waits for one arrival and then switches the system on and keeps it on forever. Then*

$$\begin{cases} V_\alpha^\pi(i,0) > V_\alpha^\psi(i,0), & \text{if } i > A(\alpha); \\ V_\alpha^\pi(i,0) < V_\alpha^\psi(i,0), & \text{if } i < A(\alpha); \\ V_\alpha^\pi(i,0) = V_\alpha^\psi(i,0), & \text{if } i = A(\alpha); \end{cases}$$

where $A(\alpha)$ is as in (15).

Proof:

$$V_\alpha^\pi(i,0) - V_\alpha^\psi(i,0) = \left(\frac{hi}{\lambda+\alpha} + \frac{\lambda}{\lambda+\alpha}\left(s_1 + U_\alpha(i+1,1)\right)\right) - \left(s_1 + U_\alpha(i,1)\right)$$
$$= \left[\frac{hi}{\lambda+\alpha} + \frac{\lambda}{\lambda+\alpha}\left(s_1 + \frac{h(i+1)}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}\right)\right] - \left[s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}\right]$$
$$= \frac{hi}{\lambda+\alpha}\frac{\mu}{\mu+\alpha} - \frac{\alpha}{\lambda+\alpha}\left(s_1 + \frac{c}{\alpha}\right) = \frac{h\mu}{(\lambda+\alpha)(\mu+\alpha)}\left(i - A(\alpha)\right),$$

where the second equality holds in view of Theorem 2(i) and the rest is straightforward. ■

Proof of Theorem 3: Let ϕ be a stationary discounted-cost optimal policy within the class of the policies that never switch off the running system. Let ψ be the policy that switches the system on at time 0 and keeps it on forever, and π be the policy that waits for one arrival and then switches the system on and keeps it on forever. By (14),

$$V_\alpha^\phi(i,0) = \min\left\{s_1 + U_\alpha(i,1),\ \frac{hi}{\lambda+\alpha} + \frac{\lambda}{\lambda+\alpha}U_\alpha(i+1,0)\right\}. \tag{17}$$

First, consider the case $i > A(\alpha)$. Then $\phi(i,0) = 1$. Assume $\phi(i,0) = 0$ for some $i > A(\alpha)$. By Lemma 2, $\phi(j,0) = 1$ for some $j > i$. Thus, there exists an $i^* \geq i$ such that $\phi(i^*,0) = 0$ and $\phi(i^*+1,0) = 1$. This implies that $V_\alpha^\psi(i^*,0) \geq V_\alpha^\pi(i^*,0)$, where $i^* > A(\alpha)$. By Lemma 3, this is a contradiction. Thus $\phi(i,0) = 1$ for all $i > A(\alpha)$.

Second, in the case that $i < A(\alpha)$, Lemma 3 implies $V_\alpha^\pi(i,0) < V_\alpha^\psi(i,0)$. Thus $\phi(i,0) = 0$ for all $i < A(\alpha)$.

In the case that $A(\alpha) = i$, Lemma 3 implies $V_\alpha^\psi(i,0) = V_\alpha^\pi(i,0)$. From (14), $V_\alpha^\psi(i,0) = V_\alpha^\pi(i,0) = U_\alpha(i,0) = \min\left\{V_\alpha^\psi(i,0), V_\alpha^\pi(i,1)\right\}$. Thus $\phi(i,0) = 1$ or $\phi(i,0) = 0$. ■

Corollary 1: *Let*

$$n_\alpha = \lceil A(\alpha) \rceil, \tag{18}$$

where $A(\alpha)$ is as in (15). Then the following statements hold:

(i) *If $A(\alpha)$ is not an integer, then the n_α-full-service policy is the unique stationary discounted-cost optimal policy within the class of policies that never switch the running system off;*

(ii) *If $A(\alpha)$ is an integer, then there are exactly two stationary discounted-cost optimal policies within the class of policies that never switch the running system off, and these policies are n_α- and $(n_\alpha + 1)$-full-service policies;*

(iii)

$$U_\alpha(i,0) = \begin{cases} \sum_{k=0}^{n_\alpha - i - 1} \left(\frac{\lambda}{\lambda+\alpha}\right)^k \frac{h(i+k)}{\lambda+\alpha} + \left(\frac{\lambda}{\lambda+\alpha}\right)^{n_\alpha - i}, \\ \left(s_1 + \frac{hn_\alpha}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}\right), \quad \text{if } i < n_\alpha; \\ s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}, \quad \text{if } i \geq n_\alpha. \end{cases} \tag{19}$$

Proof: Statements (i) and (ii) follow directly from Theorem 3 and Definition 1. Statements (i) and (ii) imply that $V_\alpha^\phi = U_\alpha$, where ϕ is the n_α-full-service policy. The first line of (19) is the discounted cost to move from state $(i,0)$ to state $(n_\alpha, 0)$, when the system is off, plus the discounted cost $U_\alpha(n_\alpha, 0)$. The second line of (19) follows from (16). ■

Corollary 2: *Let $n = \lfloor (c/h) + 1 \rfloor$. Then there exists $\alpha^* > 0$ such that, for every discount rate $\alpha \in (0, \alpha^*]$, the n-full-service policy is discounted-cost optimal within the class of the policies that never switch the running system off.*

Proof: In view of (15), the function $A(\alpha)$ is strictly monotone when $\alpha > 0$. In addition, $A(\alpha) \searrow \dfrac{c}{h}$ when $\alpha \searrow 0$. This implies that $n_\alpha = n$ for all $\alpha \in (0, \alpha^*]$, where α^* can be found by solving the quadratic inequality $A(\alpha) \leq n$. The rest follows from Corollary 1 (i) and (ii). ■

We remark that Corollary 2 describes a policy that is Blackwell optimal within the class of policies that never switch the running server off. In general, the topic of Blackwell optimality lays outside of the scope of this paper.

3.3. Properties of Discounted-Cost Optimal Policies and Reduction to a Problem with a Finite State Space

This subsection introduces the properties of the discounted-cost optimal policies, formulated in Lemmas 4 and 5, and describes the inequalities between the major thresholds in Lemma 7 that lead to the reduction of the original infinite-state problem to a finite-state problem. This reduction essentially follows from Corollary 4. Certain structural properties of discounted-cost optimal policies are described in Theorem 4. These properties are used in Section 4 to describe the structure of average-cost optimal policies.

Lemma 4: *Let ϕ be a stationary discounted-cost optimal policy. Then $\phi(i,1) = 1$ for $i \geq (h\lambda + (c - s_0\alpha)(\mu+\alpha))/h\mu$.*

Proof: Let $\phi(i,1) = 0$. Then $V_\alpha^\phi(i,1) > s_0 + hi/\alpha$, since the number of customers in the system is always greater or equal than i and after the first arrival it is greater than i.

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

Observe that $V_\alpha^\phi(i,1) = V_\alpha(i,1) \le U_\alpha(i,1)$. From (9),

$$s_0 + \frac{hi}{\alpha} < \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}.$$

This inequality implies $i < \dfrac{h\lambda + (c - s_0\alpha)(\mu+\alpha)}{h\mu}$. Thus, $\phi(i,1) = 1$ otherwise ■

Let $V_\alpha^1(i,\delta)$ and $V_\alpha^0(i,\delta)$ be the expected discounted total costs for policies that switch the system on or off respectively at time 0, keep the system status unchanged until the next jump, and then follow the discounted-cost optimal policy,

$$V_\alpha^1(i,\delta) = (1-\delta)s_1 + \frac{hi+c}{\alpha+\lambda+i\mu} + \frac{\lambda}{\alpha+\lambda+i\mu}V_\alpha(i+1,1) + \frac{i\mu}{\alpha+\lambda+i\mu}V_\alpha(i-1,1),$$

$$V_\alpha^0(i,\delta) = \delta s_0 + \frac{hi}{\alpha+\lambda} + \frac{\lambda}{\alpha+\lambda}V_\alpha(i+1,0).$$

Let M_α^* be the largest number of customers in the system, for which it is optimal to switch the running system off,

$$M_\alpha^* = \begin{cases} \max\{0 \le i < \infty : V_\alpha^0(i,1) \le V_\alpha^1(i,1)\}, & \text{if } \{0 \le i < \infty : V_\alpha^0(i,1) \le V_\alpha^1(i,1)\} \neq \emptyset; \\ -1, & \text{otherwise.} \end{cases} \tag{20}$$

COROLLARY 3: *For all $\alpha > 0$*

$$M_\alpha^* \le \frac{\lambda}{\mu} + \frac{(c+s_0\mu)^2}{4s_0h\mu} < \infty. \tag{21}$$

PROOF: According to Lemma 4, $M_\alpha^* \le f(\alpha)$, where

$$f(\alpha) = \frac{h\lambda + (c - s_0\alpha)(\mu+\alpha)}{h\mu}.$$

For $\alpha > 0$, the maximum of $f(\alpha)$ equals to the expression in the middle of (21). ■

LEMMA 5: *Let ϕ be a stationary discounted-cost optimal policy. Then for each integer $j \ge 0$ there exists an integer $i \ge j$ such that $\phi(i,0) = 1$.*

PROOF: By contradiction. Fix an arbitrary integer $j \ge 0$. If $\phi(i,0) = 0$ for all $i \ge j$ then, by Lemma 2, $V_\alpha^\phi(j,0) > U_\alpha(j,0) \ge V_\alpha(j,0)$. This contradicts the optimality of ϕ. ■

Let N_α^* be the smallest number of customers in the system exceeding M_α^*, for which it is optimal to switch the servers on, if they are off,

$$N_\alpha^* = \min\{i > M_\alpha^* : V_\alpha^1(i,0) \le V_\alpha^0(i,0)\}. \tag{22}$$

Lemma 5 implies that N_α^* is well defined and $N_\alpha^* < \infty$ for all $\alpha > 0$.

To prove that $N_\alpha^* \le n_\alpha$, we introduce the following lemma.

LEMMA 6: *The following properties hold for the function* $V_\alpha(i,\delta)$*:*

(i) *if* $V_\alpha(i,0) = V_\alpha^1(i,0)$*, then* $V_\alpha^1(i,1) < V_\alpha^0(i,1)$*;*
(ii) *if* $V_\alpha(i,1) = V_\alpha^0(i,1)$*, then* $V_\alpha^0(i,0) < V_\alpha^1(i,0)$*;*
(iii) $-s_1 \le V_\alpha(i,1) - V_\alpha(i,0) \le s_0$.

PROOF: (i) If $V_\alpha(i,0) = V_\alpha^1(i,0)$, then $V_\alpha^1(i,0) \le V_\alpha^0(i,0)$. Hence $V_\alpha(i,1) = V_\alpha(i,0) - s_1 < V_\alpha(i,0) + s_0 = V_\alpha^0(i,1)$, where the inequality follows from the assumption that $s_0 + s_1 > 0$. This implies $V_\alpha^1(i,1) < V_\alpha^0(i,1)$.

(ii) If $V_\alpha(i,1) = V_\alpha^0(i,1)$, then $V_\alpha^0(i,1) \le V_\alpha^1(i,1)$. Hence $V_\alpha(i,0) = V_\alpha(i,1) - s_0 < V_\alpha(i,1) + s_1 = V_\alpha^1(i,0)$.

(iii) $V_\alpha(i,0) \le s_1 + V_\alpha(i,1)$ because $V_\alpha(i,0) = \min\{s_1 + V_\alpha(i,1), V_\alpha^0(i,0)\} \le s_1 + V_\alpha(i,1)$, and $V_\alpha(i,1) \le s_0 + V_\alpha(i,0)$ because $V_\alpha(i,1) = \min\{V_\alpha^1(i,1), s_0 + V_\alpha(i,0)\} \le s_0 + V_\alpha(i,0)$. ■

The following lemma shows the orders among M_α^*, N_α^* and n_α. This leads to the description of the properties of discounted-cost optimal policies in Corollary 4 that essentially reduces the problem to a finite state-space problem.

LEMMA 7: $M_\alpha^* < N_\alpha^* \le n_\alpha$ *for all* $\alpha > 0$.

PROOF: The definition (22) of N_α^* implies that $M_\alpha^* < N_\alpha^*$. Thus, we need only to prove that $N_\alpha^* \le n_\alpha$.

If $M_\alpha^* = -1$, according to (20), a discounted-cost optimal policy should never switch the running system system off and therefore $V_\alpha = U_\alpha$. In view of Corollary 1, $V_\alpha^0(i,0) < V_\alpha^1(i,0)$, when $i = 0, \ldots, n_\alpha - 1$, and $V_\alpha^0(n_\alpha,0) = V_\alpha^1(n_\alpha,0)$. Thus, in this case, $N_\alpha^* = n_\alpha$.

Let $M_\alpha^* \ge 0$. Consider a stationary discounted-cost optimal policy φ that switches the system on at state $(N_\alpha^*,0)$. Such a policy exists in view of the definition of N_α^*. It follows from the definition of M_α^* that $V_\alpha^1(i,1) < V_\alpha^0(i,1)$ for $i > M_\alpha^*$. Thus, the discounted-cost optimal policy φ always keeps running the active system at states $(i,1)$, when $i > M_\alpha^*$. Observe that

$$V_\alpha^0(N_\alpha^* - 1, 0) < V_\alpha^1(N_\alpha^* - 1, 0). \tag{23}$$

If $M_\alpha^* < N_\alpha^* - 1$, (23) follows from the definition of N_α^*. If $M_\alpha^* = N_\alpha^* - 1$, (23) follows from $V_\alpha^0(M_\alpha^*,1) \le V_\alpha^1(M_\alpha^*,1)$ and from Lemma 6 (ii). Thus, starting from the state $(N_\alpha^* - 1, 0)$, the discounted-cost optimal policy φ waits until the next arrival, then switches the system on and runs it until the number of customers in queue becomes $M_\alpha^* \le N_\alpha^* - 1$. For $i = 0, 1, \ldots$, let $F_\alpha^1(i)$ be the expected total discounted cost incurred until the first time $\theta(i)$ when the number of customers in the system is i and the system is running, if at time 0 the system is off, there are i customers in queue, and the system is switched on after the first arrival and is kept on as long as the number of customers in system is greater than i. Let $\theta = \theta(N_\alpha^* - 1)$. Since φ is the discounted-cost optimal policy, $V_\alpha(N_\alpha^* - 1, 0) = F_\alpha^1(N_\alpha^* - 1) + [Ee^{-\alpha\theta}]V_\alpha(N_\alpha^* - 1, 1)$.

Let π be a policy that switches the system on in state $(N_\alpha^* - 1, 0)$ and then follows a discounted-cost optimal policy. Then, in view of (23), the policy π is not discounted-cost optimal at the initial state $(N_\alpha^* - 1, 0)$. Thus, $V_\alpha^\pi(N_\alpha^* - 1, 0) > V_\alpha(N_\alpha^* - 1, 0)$. Since

$V_\alpha^\pi(N_\alpha^*-1,0)=s_1+V_\alpha(N_\alpha^*-1,1)$,

$$F_\alpha^1(N_\alpha^*-1)+[Ee^{-\alpha\theta}]V_\alpha(N_\alpha^*-1,1)<s_1+V_\alpha(N_\alpha^*-1,1),$$

and this is equivalent to

$$\left(1-[Ee^{-\alpha\theta}]\right)V_\alpha(N_\alpha^*-1,1)>F_\alpha^1(N_\alpha^*-1)-s_1. \quad \textbf{(24)}$$

Assume that $n_\alpha<N_\alpha^*$. Then $n_\alpha\le N_\alpha^*-1$ and, in view of Theorem 3, $\psi(N_\alpha^*-1,0)=1$ for a stationary discounted-cost optimal policy ψ within the class of policies that never switches the system off. Thus, $U_\alpha(N_\alpha^*-1,0)=V_\alpha^\psi(N_\alpha^*-1,0)=s_1+U_\alpha(N_\alpha^*-1,1)$. In addition, $U_\alpha(N_\alpha^*-1,0)\le V_\alpha^\varphi(N_\alpha^*-1,0)=F_\alpha^1(N_\alpha^*-1)+[Ee^{-\alpha\theta}]V_\alpha(N_\alpha^*-1,1)$. Thus,

$$(1-[Ee^{-\alpha\theta}])U_\alpha(N_\alpha^*-1,1)\le F_\alpha^1(N_\alpha^*-1)-s_1. \quad \textbf{(25)}$$

Since $\theta\ge0$ and $U_\alpha(N_\alpha^*-1,1)\ge V_\alpha(N_\alpha^*-1,1)$, (25) contradicts (24). Thus $N_\alpha^*\le n_\alpha$. ■

LEMMA 8: *For each $\alpha>0$, the inequality $V_\alpha^1(i,0)\le V_\alpha^0(i,0)$ holds when $i\ge n_\alpha$.*

PROOF: Fix any $\alpha>0$. Consider two cases: in case (i) the best full-service policy is discounted-cost optimal, and in case (ii) the best full-service policy is not discounted-cost optimal.

Case (i). According to Corollary 1, the n_α-full-service policy is discounted-cost optimal. This implies that $V_\alpha^1(i,0)\le V_\alpha^0(i.0)$ for all $i\ge n_\alpha$.

Case (ii). Let ϕ be a stationary discounted-cost optimal policy. Assume that there exists an integer $j\ge n_\alpha$ such that $\phi(j,0)=0$. Then, in view of Lemma 5, there is $i\ge j$ such that $\phi(i,0)=0$ and $\phi(i+1,0)=1$. As shown in Lemma 7, $n_\alpha>M_\alpha^*$ and therefore $\phi(\ell,1)=1$ for all $\ell>M_\alpha^*$. Thus, $\phi(\ell,1)=1$ for all $\ell>i$. We have

$$\begin{aligned}V_\alpha^\phi(i,0)&=F_\alpha^1(i)+[Ee^{-\alpha\theta(i)}]V_\alpha(i,1)\le s_1+V_\alpha(i,1)\Rightarrow F_\alpha^1(i)-s_1\\&\le(1-[Ee^{-\alpha\theta(i)}])V_\alpha(i,1),\end{aligned} \quad \textbf{(26)}$$

where the stopping time $\theta(i)$ and the expected total discounted cost $F_\alpha^1(i)$ are defined in the proof of Lemma 7. On the other hand, since $i\ge n_\alpha$, under n_α-full-service policy π we have

$$\begin{aligned}V_\alpha^\pi(i,0)&=s_1+U_\alpha(i,1)\le F_\alpha^1(i)+[Ee^{-\alpha\theta(i)}]U_\alpha(i,1)\\&\Rightarrow(1-[Ee^{-\alpha\theta(i)}])U_\alpha(i,1)\le F_\alpha^1(i)-s_1.\end{aligned} \quad \textbf{(27)}$$

By (26) and (27), we have $U_\alpha(i,1)\le V_\alpha(i,1)$. Since the best full-service policy is not discounted-cost optimal, $U_\alpha(i,1)>V_\alpha(i,1)$. This contradiction implies the correctness of the lemma. ■

COROLLARY 4: *Let $\alpha>0$ and $\alpha'\in(0,\alpha]$. For a stationary discounted-cost optimal policy ϕ for the discount rate α', consider the stationary policy ϕ',*

$$\phi'(i,\delta)=\begin{cases}\phi(i,\delta), & \text{if } i<n_\alpha;\\ 1, & \text{if } i\ge n_\alpha.\end{cases}$$

Then the policy ϕ' is also discounted-cost optimal for the discount rate α'.

Proof: Let $\alpha' = \alpha$. By the definition (20) of M^*_α, the inequality $V^1_\alpha(i,1) \le V^0_\alpha(i,1)$ holds for all $i > M^*_\alpha$. By Lemma 8 and by Corollary 1, $V^1_\alpha(i,0) \le V^0_\alpha(i,0)$ for all $i \ge n_\alpha$. In view of Lemma 7, $M^*_\alpha < n_\alpha$. Thus, $V^1_\alpha(i,\delta) \le V^0_\alpha(i,\delta)$ for all $i \ge n_\alpha$ and for all $\delta = 0, 1$. This implies the discounted-cost optimality of ϕ' for the discount rate $\alpha' = \alpha$. Now let $\alpha' \in (0, \alpha)$. Since $\alpha > \alpha' > 0$, then $n_{\alpha'} \le n_\alpha$, and 1 is an optimal decision for the discount rate α' at each state (i,δ) with $i \ge n_{\alpha'}$ and thus with $i \ge n_\alpha$. ■

Corollary 4 implies that it is optimal to turn and keep the system on, if there are n_α or more customers and the discount rate is not greater than α. This essentially means that, in order to find a discounted-cost optimal policy for discount rates $\alpha' \in (0, \alpha]$, the decision maker should find such a policy only for a finite set of states (i,δ) with $i = 0, 1, \ldots, n_\alpha - 1$ and $\delta = 0, 1$. Thus, Corollary 4 reduces the original problem of optimization of the total discounted costs to a finite-state problem, and for every $\alpha > 0$ this finite-state set is the same for all discount factors between 0 and α. The following theorem describes structural properties of a discounted-cost optimal policy for a fixed discount factor.

Theorem 4: *For each $\alpha > 0$, either the n_α-full-service policy is discounted-cost optimal, or there exists a stationary discounted-cost optimal policy ϕ_α with the following properties:*

$$\phi_\alpha(i,\delta) = \begin{cases} 1, & \text{if } i > M^*_\alpha,\ \delta = 1 \ \textit{or}\ i = N^*_\alpha,\ \delta = 0 \text{ or } i \ge n_\alpha,\ \delta = 0; \\ 0, & \text{if } i = M^*_\alpha,\ \delta = 1 \ \textit{or}\ M^*_\alpha \le i < N^*_\alpha,\ \delta = 0. \end{cases} \tag{28}$$

Proof: Consider a stationary discounted-cost optimal policy ψ for the discount rate $\alpha > 0$. If $M^*_\alpha = -1$, then ψ never switches the servers off. Therefore, the n_α-full-service policy is discounted-cost optimal. If $M^*_\alpha \ge 0$, change ψ to ϕ_α according to (28) on the set of states specified on the right-hand side of (28). The optimality of the new policy, denoted by ϕ_α, follows from the definitions of M^*_α and N^*_α, and from Corollary 4. ■

4. THE EXISTENCE AND STRUCTURE OF AVERAGE-COST OPTIMAL POLICIES

In this section, we study the average-cost criterion, prove the existence of average-cost optimal policies, and describe their properties.

Definition 3: *For two nonnegative integers M and N with $N > M$, a stationary policy is called an (M,N)-policy if*

$$\phi(i,\delta) = \begin{cases} 1, & \textit{if } i > M,\ \delta = 1 \text{ or } i \ge N,\ \delta = 0; \\ 0, & \textit{if } i \le M,\ \delta = 1 \text{ or } i < N,\ \delta = 0. \end{cases}$$

Theorem 5: *There exists a stationary average-cost optimal policy and, depending on the model parameters, either every full-service policy is average-cost optimal or an (M,N)-policy is average-cost optimal for some $N > M \ge 0$ and $N \le n^*$, where*

$$n^* = \lfloor \frac{c}{h} + 1 \rfloor. \tag{29}$$

In addition, the optimal average-cost value $v(i,\delta)$ is the same for all initial states (i,δ); that is, $v(i,\delta) = v$.

PROOF: We first prove that either the n^*-full-service policy is average-cost optimal or an (M, N)-policy is average-cost optimal for some $N > M \geq 0$ and $N \leq n^*$. For the initial CTMDP, consider a sequence $\alpha_k \downarrow 0$ as $k \to \infty$. Let ϕ^k be a stationary discounted-cost optimal policy for the discount rate α_k. According to Theorem 4, for each k this policy can be selected either as an n_{α_k}-full-service policy or as a ϕ_{α_k} policy satisfying (28). Since, in view of (15) and (18), $n_{\alpha_k} \leq n_{\alpha_1} < (\mu + \alpha_1)(c + \alpha_1 s_1)/h\mu + 1 < \infty$ for all $k = 1, 2, \ldots$, there exists a subsequence $\{\alpha_{k_\ell}\}$, $\ell = 1, 2, \ldots$, of the sequence $\{\alpha_k\}$, $k = 1, 2, \ldots$ such that all the policies $\phi^{k_\ell} = \phi$, where ϕ is a stationary policy such that either (i) the policy ϕ is the n^*-full-service policy for some integer n or (ii) the policy ϕ satisfies the conditions on the right-hand side of (28) with the same $M^*_\alpha = M$ and $N^*_\alpha = N$ for $\alpha = \alpha_{k_\ell}$.

Observe that the values of $v^\phi(i, \delta)$ do not depend on the initial state (i, δ). Indeed, in case (i), when the policy ϕ is the n^*-full-service policy, the stationary policy ϕ defines a Markov chain with a single positive recurrent class $\{(i, 1) \in Z : i = 0, 1, \ldots\}$, and all the states in its complement $\{(i, 0) \in Z : i = 0, 1, \ldots\}$ are transient. The same is true for case (ii) with the positive recurrent class $Z^* = \{(i, 1) \in Z : i = M, M + 1, \ldots\} \cup \{(i, 0) \in Z : i = M, M + 1, \ldots, N\}$ and with the set of transient states $Z \setminus Z^*$. In each case, the Markov chain leaves the set of transient states in a finite expected amount of time incurring a finite expected cost until the time the chain enters the single positive recurrent class. Thus, for all initial states (i, δ), in each case $v^\phi(i, \delta) = v^\phi$ does not depend on (i, δ), and

$$v^\phi = \lim_{t \to \infty} t^{-1} E^\phi_{(i,\delta)} C(t) = \lim_{\alpha \downarrow 0} \alpha V^\phi_\alpha(i, \delta),$$

where the second equality and the existence of the second limit follow from the Tauberian theorem; see, e.g., Korevaar [25]. In addition, if $\alpha > 0$ is sufficiently close to 0, then, in view of (15) and (18), $n_\alpha = \lceil c/h \rceil$, if c/h is not integer, and $n_\alpha = c/h + 1$, if c/h is integer. This explains why $n^* = \lfloor \frac{c}{h} + 1 \rfloor$ in Theorem 5. In conclusion, $v(i, \delta) = v$, since $v^\phi(i, \delta) = v^\phi$. In addition, as follows from (9), $v^\sigma = (\lambda h/\mu) + c$ for each full-service policy σ. Thus, if the n^*-full-service policy is average-cost optimal, then every full-service policy is average-cost optimal. ■

5. COMPUTATION OF AN AVERAGE-COST OPTIMAL POLICY

In this section, we show how an optimal policy can be computed via linear programming. According to Theorem 5, there is an optimal policy ϕ with $\phi(i, \delta) = 1$ when $i \geq n^* = \lfloor (c/h) + 1 \rfloor$. Thus, the goal is to find the values of $\phi(i, \delta)$ when $i = 0, 1, \ldots, n^* - 1$ and $\delta = 0, 1$. To do this, we truncate the state space Z to $Z' = \{0, 1, \ldots, n^* - 1\} \times \{0, 1\}$. If the action 1 is selected at state $(n^* - 1, 1)$, the system moves to the state $(n^* - 2, 1)$, if the next change of the number of the customers in the system is a departure and the system remains in $(n^* - 1, 1)$, if an arrival takes place. In the latter case, the number of customers increases by one at the arrival time and then it moves according to the random work until it hits the state $(n^* - 1, 1)$ again. Thus the system can jump from the state $(n^* - 1, 1)$ to itself. Furthermore, the distributions of the sojourn times at this state are not exponential. Therefore, the truncated problem cannot be described as a CTMDP. However, it can be described as a semi-Markov Decision Process (SMDP); see Mine and Osaki [31, Chapter 5] and Puterman [33, Chapter 11].

We formulate our problem as an SMDP with the state set Z' and the action set $A(z) = A = \{0, 1\}$. If an action a is selected at state $z \in Z'$, the system spends an average time τ' in this state until it moves to the next state $z' \in Z'$ with the probability $p(z'|z, a)$. During

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

this time the expected cost $C'(z,a)$ is incurred. For $z=(i,\delta)$ with $i=0,1,\dots,n^*-2$ and $\delta=0,1$, these characteristics are the same as for the original CTMDP and are given by

$$p(z'|z,a)=\begin{cases}1, & \text{if } a=0,\ z'=(i+1,0);\\ \dfrac{\lambda}{\lambda+i\mu}, & \text{if } a=1,\ z'=(i+1,1);\\ \dfrac{i\mu}{\lambda+i\mu}, & \text{if } a=1,\ z'=(i-1,1);\\ 0, & \text{otherwise};\end{cases} \tag{30}$$

$$\tau'((i,\delta),a)=\begin{cases}\dfrac{1}{\lambda}, & \text{if } a=0;\\ \dfrac{1}{\lambda+i\mu} & \text{if } a=1;\end{cases} \tag{31}$$

and $C'((i,\delta),a)=|a-\delta|s_a+(hi+ac)\tau'((i,\delta),a)$. The transition probabilities in states (n^*-1,δ) with $\delta=0,1$ are defined by $p((n^*-2,1)|(n^*-1,\delta),1)=(n^*-1)\mu/(\lambda+(n^*-1)\mu)$, $p((n^*-1,1)|(n^*-1,\delta),1)=\lambda/(\lambda+(n^*-1)\mu)$, and $p((n^*-1,1)|(n^*-1,\delta),0)=1$. In the last case, the number of customers increases by 1 to n^*, the system switches on, and eventually the number of customers becomes n^*-1.

Let T_i be the expected time between an arrival seeing i customers in an $M/M/\infty$ queue and the next time when a departure leaves i customers behind, $i=0,1,\dots$. Applying the memoryless property of the exponential distribution, $T_i=B_{i+1}-B_i$, where B_i is the expected busy period for $M/M/\infty$ starting with i customers in the system and $B_0=0$. By formula (34b) in Browne and Kella [4],

$$B_i=\frac{1}{\lambda}\left(e^\rho-1+\sum_{k=1}^{i-1}\frac{k!}{\rho^k}\left(e^\rho-\sum_{j=0}^{k}\frac{\rho^j}{j!}\right)\right), \tag{32}$$

where $\rho=(\lambda/\mu)$. Thus

$$T_{n^*-1}=B_{n^*}-B_{n^*-1}=\frac{1}{\lambda}\sum_{k=0}^{\infty}\frac{\rho^{k+1}}{n^*(n^*+1)\dots(n^*+k)}.$$

The expected time $\tau'((n^*-1,\delta),1)$, where $\delta=0,1$, is the expected time until the next arrival plus T_{n^*-1}, if the next event is an arrival. Thus, $\tau'((n^*-1,\delta),1)=(\lambda/(\lambda+(n^*-1)\mu))\left((1/\lambda)+T_{n^*-1}\right)$, $\delta=0,1$. In addition $\tau'((n^*-1,\delta),0)=(1/\lambda)+T_{n^*-1}$, $\delta=0,1$.

To compute the one-step cost $C'((n^*-1,1),1)$, we define m_i as the average number of visits to state $(i,1)$ starting from state $(n^*-1,1)$ and before revisiting state $(n^*-1,1)$, $i=n^*-1,n^*,\dots$. And define $m_{i,i+1}$ as the expected number of jumps from $(i,1)$ to $(i+1,1)$, $i=n^*-1,n^*,\dots$, and $m_{i,i-1}$ as the expected number of jumps from $(i,1)$ to $(i-1,1)$, $i=n^*,n^*+1,\dots$. Then $m_{i,i+1}=(\lambda/(\lambda+i\mu))m_i$, $m_{i,i-1}=(i\mu/(\lambda+i\mu))m_i$ and $m_{i,i+1}=m_{i+1,i}$. Since $m_{n^*-1}=1$,

$$m_i=\prod_{j=0}^{i-n^*}\frac{\lambda}{\lambda+(n^*-1+j)\mu}\frac{\lambda+(n^*+j)\mu}{(n^*+j)\mu},\qquad i=n^*,n^*+1,\dots . \tag{33}$$

https://doi.org/10.1017/S0269964815000157 Published online by Cambridge University Press

Thus,

$$C'((n^*-1,1),1) = \sum_{i=n^*-1}^{\infty} m_i C((i,1),1) = \sum_{i=n^*-1}^{\infty} m_i \frac{hi+c}{\lambda+i\mu},$$

where $C((i,1),1) = ((hi+c)/(\lambda+i\mu))$, $i = n^*-1, n^*, \ldots$ is the cost incurred in state $(i,1)$ under action 1 for the original state space model; see Section 3.1. The one-step cost $C'((n^*-1,0),1) = s_1 + C'((n^*-1,1),1)$.

Let C_{n^*} be the expected total cost incurred in an $M/M/\infty$ system until the number of customers becomes (n^*-1), if at time 0 there are n^* customers in the system and the system is running. Then

$$C'((n^*-1,1),1) = \frac{h(n^*-1)+c}{\lambda+(n^*-1)\mu} + \frac{\lambda}{\lambda+(n^*-1)\mu} C_{n^*},$$

and this implies

$$C_{n^*} = \left(1+\frac{(n^*-1)\mu}{\lambda}\right) C'((n^*-1,1),1) - \frac{h(n^*-1)+c}{\lambda}.$$

We also have $C'((n^*-1,0),0) = \dfrac{h(n^*-1)}{\lambda} + s_1 + C_{n^*}$, $C'((n^*-1,0),1) = s_1 + C'((n^*-1,1),1)$, and $C'((n^*-1,1),0) = s_0 + C'((n^*-1,0),0)$.

With the definitions of the transition mechanisms, sojourn times, and one-step costs for the SMDP, now we formulate the LP according to Section 5.5 in Mine and Osaki [31] or Theorem 11.4.2 and formula (11.4.17) in Puterman [33] as

$$\begin{aligned} &\text{Minimize} \sum_{z\in Z'}\sum_{a\in A} C'(z,a)x_{z,a} \\ \text{s.t.}\quad &\sum_{a\in A(z)} x_{z,a} - \sum_{z'\in Z'}\sum_{a\in A(z)} p(z|z',a)x_{z,a} = 0,\ z\in Z', \\ &\sum_{z\in Z'}\sum_{a\in A(z)} \tau'(z,a)x_{z,a} = 1, \\ &x_{z,a}\ge 0,\ z\in Z',\ a\in A. \end{aligned} \qquad \textbf{(34)}$$

Let x^* be the optimal basic solution of (34). According to general results on SMDPs in Denardo [6, Section III], for each $z\in Z'$, there exists at most one $a\in\{0,1\}$ such that $x^*_{z,a}>0$. If $x^*_{z,a}>0$, then for the average-cost optimal policy ϕ, $\phi(z)=a$, for $a=0,1$. If $x^*_{z,0}=x^*_{z,1}=0$, then $\phi(z)$ can be either 0 or 1. For our problem, Theorem 6 explains how $x^* := \{x^*_{z,a} : z\in Z', a\in A\}$ can be used to construct a stationary average-cost optimal policy ϕ with the properties stated in Theorem 5.

THEOREM 6: *For an optimal basic solution x^* of (34), the following statements hold:*

(i) *if $x^*_{(0,1),1}>0$, then every full-service policy is average-cost optimal;*

(ii) *If $x^*_{(0,1),0}>0$, then the $(0,N)$-policy is average-cost optimal with*

$$N = \begin{cases} n^*, & \text{if } \min\{i=1,\ldots,n^*-1 : x^*_{(i,0),1}>0\} = \emptyset; \\ \min\{i=1,\ldots,n^*-1 : x^*_{(i,0),1}>0\}, & \text{if} \min\{i=1,\ldots,n^*-1 : x^*_{(i,0),1}>0\} \ne \emptyset; \end{cases} \qquad \textbf{(35)}$$

(iii) *if* $x^*_{(0,1),0} = x^*_{(0,1),1} = 0$, *then the* (M, N)*-policy is average-cost optimal with* $M = \min\{i = 1, \ldots, n^* - 1 : x^*_{(i,1),0} > 0\} > 0$ *and* N *being the same as in* (35).

Proof: Let ϕ^* be a stationary average-cost optimal policy defined by the optimal basic solution x^* of LP (34). Since at most one of the values $\{x^*_{(0,1),0}, x^*_{(0,1),1}\}$ is positive and they both are nonnegative, cases (i)–(iii) are mutually exclusive and cover all the possibilities.

(i) If $x^*_{(0,1),1} > 0$, then the state $(0, 1)$ is recurrent under the policy ϕ^* and $\phi^*(0, 1) = 1$. Since the state $(0, 1)$ is recurrent, $\phi^*(n, 1) = 1$ for all $n = 1, 2, \ldots$. This is true because, if $\phi^*(n, 1) = 0$ for some $n = 1, 2, \ldots$, then in the long-run the number of customers will be always greater than or equal to n, and the state $(0, 1)$ cannot be recurrent. Thus, $v^{\phi}(j, 0) = v^{\phi^*}(i, 0) = v = c + h\lambda/\mu$ for every full-service policy ϕ, for all $i, j = 1, 2, \ldots$, and each full-service policy is average-cost optimal.

(ii) If $x^*_{(0,1),0} > 0$, then the state $(0, 1)$ is recurrent under the policy ϕ^*, and $\phi^*(0, 1) = 0$. Since the state $(0, 1)$ is recurrent, the policy ϕ^* always keeps the running system on as long as the system is nonempty. By Lemma 6 (ii), $\phi^*(0, 0) = 0$. The first constraint in LP (34) implies that $x^*_{(1,0),0} + x^*_{(1,0),1} > 0$. In general, if $x^*_{(i,0),0} + x^*_{(i,0),1} > 0$ for some $i = 1, \ldots, n^* - 1$, then $\phi^*(j, 0) = 0$ if $x^*_{(j,0),1} = 0$ for $j = 0, \ldots, i - 1$, and $\phi^*(i, 0) = 1$ if $x^*_{(i,0),1} > 0$. Otherwise, if $x^*_{(i,0),0} + x^*_{(i,0),1} = 0$ for all $i = 1, \ldots, n^* - 1$, $\phi^*(i, 0)$ can be arbitrary and we define $\phi^*(i, 0) = 0$ for $i = 0, 1, \ldots, n^* - 1$. Thus, formula (35) defines the minimal number N of customers in the system, at which the inactive system should be switched on by the average-cost optimal policy ϕ^*. We recall that the SMDP is defined for the LP in the way that the system always starts on in state $(n^*, 0)$. Thus, the policy ϕ^* always keep running the active system if the system is not empty, switches it off when the system becomes empty, and switches on the inactive system when the number of customers becomes N. If there are more than N customers when the system is inactive, the corresponding states are transient. The defined $(0, N)$-policy starts the system in all these states, and therefore it is average-cost optimal.

(iii) If $x^*_{(0,1),0} = x^*_{(0,1),1} = 0$, then the state $(0, 1)$ is transient under the policy ϕ^*. In transient states the average-cost optimal policy ϕ^* can be defined arbitrary. First observe that $x^*_{(i,1),0} > 0$ for some $i = 1, \ldots, n^* - 1$ and therefore M is well-defined in the theorem. Indeed, if $x^*_{(i,1),0} = 0$ for all states $i = 0, \ldots, n^* - 1$, we can set $\phi^*(i, 1) = 1$ for all these values of i. This means that in the original Markov chain, where the running system is always kept on when the number of customers in the system is greater or equal than n^*, the system is always on. Since the birth-and-death for an $M/M/\infty$ system is positive recurrent, we have a contradiction. Since the state $(M, 1)$ is recurrent for the Markov chain defined by the policy ϕ^*, this policy always keeps the running system on when the number of the customers in the system is M or more. Since $x^*_{(i,\delta),a} = 0$ for $i < M$ and for all $\delta, a = 0, 1$, we can define $\phi^*(i, \delta)$ arbitrarily when $i < M$. Let $\phi(i, \delta) = 0$, when $i < M$ and $\delta = 0, 1$. Similar to case (i), the policy ϕ^* prescribes to keep inactive system off as long as the number of customers in the system is less than N, switches it on when this number becomes N, and it can be prescribed to switch the inactive system on when the number of customers is greater than N, because all such states are transient. Thus, the defined (M, N)-policy is optimal.

■

Similar to (34), the LP can be formulated to find the discounted-cost optimal policy. However, we do not elaborate on the LP for the expected total discounted costs because they are not used in this paper either for computing or for studying average-cost optimal policies.

6. FINDING THE BEST $(0, N)$-POLICY AND ITS NON-OPTIMALITY

In this section we explain how to compute the best $(0, N)$-policy and show that it may not be average-cost optimal. To do the latter, we consider an example.

Before providing the example, we show how to find the best $(0, N)$-policy. This problem was studied by Browne and Kella [4] for an $M/G/\infty$ queue without running costs. Here we extend the solution from Browne and Kella [4] to a problem with a running cost rate $c > 0$. Let ψ_N be a $(0, N)$-policy. The average cost under ψ_N can be found by formula (26) in Browne and Kella [4] by replacing the setup cost there with the sum of switching costs and running costs $s_0 + s_1 + cB_N$, where B_N is the expected busy period for an $M/G/\infty$ queue that starts with N busy servers; see formula (6a) in Browne and Kella [4]. This implies

$$v^{\psi_N} = hl_N + \frac{s_0 + s_1 + cB_N}{N/\lambda + B_N}, \tag{36}$$

where l_N is the expected long-run average number of customers in the system under $(0, N)$-policy. By formulae (22), (23) in Browne and Kella [4],

$$l_N = \rho + \frac{N-1}{2}\frac{N}{N + \lambda B_N}. \tag{37}$$

The optimal N^* for the best $(0, N)$-policy is found by

$$N^* = \arg\min_N v^{\psi_N}. \tag{38}$$

The following theorem extends Theorem 3 on p. 874 in Browne and Kella [4] to a non-negative running cost rate $c \geq 0$.

THEOREM 7: *For an $M/G/\infty$ queue, let*

$$\tilde{N} = \min\left\{N > \frac{c}{h} : \frac{N(N+1)}{2\lambda} \geq \frac{s_0 + s_1}{h}\right\}. \tag{39}$$

Then $v^{\psi_N} < v^{\psi_{N+1}}$, when $N \geq \tilde{N}$. Thus $N^* = \arg\min_{1 \leq N \leq \tilde{N}} v^{\psi_N}$.

PROOF: Let $b_n = (1/\lambda) + T_n$, $n = 0, 1, \ldots$, where $T_n = B_{n+1} - B_n$. Fotmula (29) in Browne and Kella [4] provides an explicit expression for b_n. Note that $\sum_{i=0}^{N-1} b_i = B_N + (N/\lambda)$ is the expected duration of a cycle for an $M/G/\infty$ queue controlled by a $(0, N)$-policy ψ^N,

$N = 1, 2, \ldots$. By (36) and (37),

$$v^{\psi_N} = h\left(\rho + \frac{N-1}{2}\frac{N/\lambda}{\sum_{i=0}^{N-1} b_i}\right) + \frac{s_0 + s_1 + c\left(\sum_{i=0}^{N-1} b_i - N/\lambda\right)}{\sum_{i=0}^{N-1} b_i}.$$

Straightforward and somewhat lengthy calculations imply that $v^{\psi_N} < v^{\psi_{N+1}}$, if $hN - c > 0$ and

$$\left(\frac{h(N-1)}{2\lambda} + \frac{s_0 + s_1}{N} - \frac{c}{\lambda}\right) / \left(\frac{hN}{\lambda} - \frac{c}{\lambda}\right) < \frac{\sum_{i=0}^{N-1} b_i}{N b_N}. \tag{40}$$

As follows from (39), if $N \geq \tilde{N}$, the left hand side in (40) is not greater than 1. The right hand side is greater than 1 since the sequence $\{b_i\}_{i=0,1,\ldots}$ is decreasing. Thus $v^{\psi_N} < v^{\psi_{N+1}}$ for $N \geq \tilde{N}$. ■

We remark that, if $c = 0$, then (40) becomes

$$\frac{N-1}{2N} + \frac{\lambda}{h}\frac{(s_0 + s_1)}{N^2} < \frac{\sum_{i=0}^{N-1} b_i}{N b_N}. \tag{41}$$

Inequality (41) is equivalent to inequality (30) in Browne and Kella [4], which in notations of this paper is

$$\frac{1}{2} + \frac{\lambda}{h}\frac{(s_0 + s_1)}{N(N+1)} < \frac{\sum_{i=0}^{N} b_i}{(N+1) b_N}. \tag{42}$$

Indeed, if (41) is written as $A_N < B_N$, then (42) is $(A_N + (1/N))(N/N+1) < (B_N + (1/N))(N/N+1)$.

Theorem 7 implies that an average-cost optimal $(0, N)$-policy can also be found by solving the LP (34) with the state space $Z'' = \{(i, \delta) : \ i = 0, 1, \ldots, \tilde{N} - 1, \ \delta = 0, 1\}$ and with the new action set $A''(\cdot)$ defined as $A''(0, 1) = \{0\}$, $A''(i, 1) = \{1\}$ for $i = 1, \ldots, \tilde{N} - 1$, and $A''(i, 0) = \{0, 1\}$ for $i = 1, \ldots, \tilde{N} - 1$. The following example demonstrates that the best $(0, N)$-policy may not be optimal.

Example 1: Consider an $M/M/\infty$ queue with the arrival rate $\lambda = 2$, each server's rate $\mu = 1$, holding cost rate $h = 1$, service cost rate $c = 100$, and switching costs $s_0 = s_1 = 100$. Then $n^* = \lfloor (c/h) + 1 \rfloor = 101$ and $Z' = \{(i, \delta) : i = 0, 1, \ldots, 100, \ \delta = 0, 1\}$. We solved the LP (34) with CPLEX in MatLab. The value of the objective function is approximately equal to 43.39. For the found solution, $x^*_{(39,0),1} > 0$, $x^*_{(i,0),0} > 0$ for $i = 4, \ldots, 38$, $x^*_{(i,1),1} > 0$ for $i = 5, \ldots, 100$, $x^*_{(4,1),0} > 0$, and $x^*_{z,a} = 0$ for all the other $z \in Z', a \in A = \{0, 1\}$. In view of Theorem 6, the average-cost optimal policy ϕ is the $(4, 39)$-policy. Thus $v^{\phi} \approx 43.39$. The best average-cost $(0, N)$-policy was found by using Theorem 7, and $N^* = 47$. The corresponding average cost is $v^{\psi_{N^*}} \approx 51.03 > v^{\phi}$.

Acknowledgements

This research was partially supported by NSF grants CMMI-0928490 and CMMI-1335296. The authors thank Manasa Mandava for useful comments.

References

1. Altman, E. and Nain, P. (1993). Optimal control of the M/G/1 queue with repeated vacations of the server. *IEEE Transactions on Automatic Control* 38(12): 1766–1775.
2. Bell, C.E. (1975). Turning off a server with customers present: is this any way to run an M/M/c queue with removable servers? *Operations Research* 23(3): 571–574.
3. Bell, C.E. (1980). Optimal operation of an $M/M/2$ queue with removable servers. *Operations Research* 28(5): 1189–1204.
4. Browne, S. and Kella, O. (1995). Parallel service with vacations. *Operations Research* 43(5): 870–878.
5. Chao, X. and Zhao, Y.Q. (1998). Analysis of multi-server queues with station and server vacations. *European Journal of Operational Research* 110(2): 392–406.
6. Denardo, E. (1970). On linear programming in a Markov decision problem. *Management Science* 16(5): 281–288.
7. Denardo, E., Feinberg, E., and Kella, O. (1997). Stochastic monotonicity for stationary recurrence times of first passage heights. *The Annals of Applied Probability* 7(2): 326–339.
8. Federgruen, A. and So, K. (1991). Optimality of threshold policies in single-server queueing systems with server vacations. *Advances in Applied Probability* 23(2): 388–405.
9. Feinberg, E. (1994). A generalization of 'expectation equals reciprocal of intensity' to nonstationary distributions. *Journal of Applied Probability* 31(1): 262–267.
10. Feinberg, E. (2012). Reduction of discounted continuous-time MDPs with unbounded jump and reward rates to discrete-time total-reward MDPs. In *Optimization, Control, and Applications of Stochastic Systems*, D. Hernaández-Hernaández and J.A. Minjárez-Sosa (Eds.) New York: Birkhäuser/Springer, pp. 77–97.
11. Feinberg, E. and Kella, O. (2002). Optimality of D-policies for an $M/G/1$ queue with a removable server. *Queueing Systems* 42(4): 355–376.
12. Fuhrmann, S. and Cooper, R.B. (1985). Stochastic decompositions in the $M/G/1$ queue with generalized vacations. *Operations Research* 33(5): 1117–1129.
13. Greenberg, A., Hamilton, J., Maltz, D., and Patel, P. (2008). The cost of a cloud: research problems in data center networks. *ACM SIGCOMM Computer Communication Review* 39(1): 68–73.
14. Guo, X. and Hernández-Lerma, O. (2005). *Continuous-time Markov Decision Processes*. Berlin: Springer-Verlag.
15. Guo, X., Hernández-Lerma, O., and Prieto-Rumeau, T. (2006). A survey of recent results on continuous-time Markov decision processes. *Top* 14(2): 177–246.
16. Heyman, D. (1968). Optimal operating policies for M/G/1 queuing systems. *Operations Research* 16(2): 362–382.
17. Hofri, M. (1986). Queueing systems with a procrastinating server. *ACM SIGMETRICS Performance Evaluation Review* 14(1): 245–253.
18. Huang, C., Brumelle, S., Sawaki, K., and Vertinsky, I. (1977). Optimal control for multi-servers queueing systems under periodic review. *Naval Research Logistics Quarterly* 24(1): 127–135.
19. Igaki, N. (1992). Exponential two server queue with n-policy and general vacations. *Queueing Systems* 10(4): 279–294.
20. Kao, E.P. and Narayanan, K.S. (1991). Analyses of an M/M/N queue with server's vacations. *European Journal of Operational Research* 54(2): 256–266.
21. Kella, O. (1989). The threshold policy in the $M/G/1$ queue with server vacations. *Naval Research Logistics* 36(1): 111–123.
22. Kella, O. and Whitt, W. (1991). Queues with server vacations and lévy processes with secondary jump input. *The Annals of Applied Probability* 1(1): 104–117.
23. Khazaei, H., Misic, J., and Misic, V. (2012). Performance analysis of cloud computing centers using M/G/m/m+ r queuing systems. *IEEE Transactions on Parallel and Distributed Systems* 23(5): 936–943.
24. Kitaev, M. and Rykov, V. (1995). *Controlled queueing systems*. Boca Raton: CRC Press.
25. Korevaar, J. (2004). *Tauberian Theory: a Century of Developments*. New York: Springer.
26. Kulkarni, V., Kumar, S., Mookerjee, V., and Sethi, S. (2009). Optimal allocation of effort to software maintenance: A queuing theory approach. *Production and Operations Management* 18(5): 506–515.
27. Lee, H. and Srinivasan, M. (1989). Control policies for the $M^X/G/1$ queueing system. *Management Science* 35(6): 708–721.
28. Levy, Y. and Yechiali, U. (1976). An M/M/s queue with servers' vacations. *Information Systems and Operational Research* 14(2): 153–163.
29. Li, W. and Alfa, A.S. (2000). Optimal policies for M/M/m queue with two different kinds of (n, t)-policies. *Naval Research Logistics* 47(3): 240–258.

30. Mazzucco, M., Dyachuk, D., and Deters, R. (2010). Maximizing cloud providers' revenues via energy aware allocation policies. 2010 IEEE 3rd International Conference on Cloud Computing, pp. 131–138.
31. Mine, H. and Osaki, S. (1970). *Markovian Decision Processes*. vol. 25. New York: Elsevier Publishing Company.
32. Piunovskiy, A. and Zhang, Y. (2012). The transformation method for continuous-time Markov decision processes. *Journal of Optimization Theory and Applications* 154(2): 691–712.
33. Puterman, M. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: John Wiley and Sons, Inc..
34. Rhee, H.-K. and Sivazlian, B. (1990). Distribution of the busy period in a controllable M/M/2 queue operating under the triadic $(0, k, n, m)$ policy. *Journal of Applied Probability* 27(2): 425–432.
35. Ross, S. (1996). *Stochastic Processes*. New York: John Wiley and Sons, Inc.
36. Shanthikumar, J. (1988). On stochastic decomposition in $M/G/1$ type queues with generalized server vacations. *Operations Research* 36(4): 566–569.
37. Sobel, M. (1969). Optimal average-cost policy for a queue with start-up and shut-down costs. *Operations Research* 17(1): 145–162.
38. Yadin, M. and Naor, P. (1963). Queueing systems with a removable service station. *Operations Research* 14(4): 393–405.