2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Switching on and off the full capacity of an M/M/∞ queue

Eugene Feinberg and Xiaoxuan Zhang

Abstract— **This paper studies optimal switching on and off the full capacity of an $M/M/\infty$ queue with holding, running and switching costs. The main result is that an average-cost optimal policy either always runs the system or is defined by two thresholds M and N, such that the system is switched on at arrival epochs when the number of customers in the system accumulates to N and it is switched off at departured epoch when the number of customers in the system decreases to M.**

I. Introduction

Consider an $M/M/\infty$ queue with customers arriving according to a Poisson process with parameter λ. Service times are exponential with parameter μ and independent. The number of servers is unlimited. The system can be switched on and off any time. All occupied servers operate when the system is on, and all the servers are off when the system is off.

The costs include the linear holding cost h for a unit of time that a customer spends in the system, the start-up cost s_1, the shut-down cost s_0, and the running costs c_1 per unit time when the system is on and c_0 per unit time when the server is off. All the costs are nonnegative and $c_1 > c_0$. Let $c = c_1 - c_0$. The goal is to maximize the average cost per unit time. It is clear that without loss of generality we may assume that $c_1 = c > 0$ and $c_0 = 0$. So, throughout the paper the cost of running the system per unit time equals $c > 0$, and the cost of keeping the system off is zero. We also assume that $h > 0$, and at least one of the costs s_0 or s_1 is positive, that is $s_0 + s_1 > 0$.

To simplify the initial analysis, we assume that the server can be turned on and off only at time 0, customer arrival times, and customer departure times. These times are jump epochs for the process $X(t)$ of the number of customers in the system at time t. Let $t_0, t_1, \ldots$ be the sequence of jump epochs. We initially consider the servers can be switched on and off only at jump epochs. Switching takes place only at these times is not restrictive, and the optimal policies described in the paper are also optimal when the system can be turned on and off any time.

The main result of this paper is that either the policy that always keeps the server 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 turns the idling system on when the number of customers in the queue reaches or exceeds N.

In particular, $(0, N)$-policies are known in the literature under the name of N-policies. It is well-known [8] that for an $M/G/1$ queue either the policy that always runs the sever is average-cost optimal or, for some natural number N, the N-policy is average-cost optimal. As numerical results show, N-policies may not be average-cost optimal for $M/M/\infty$ queues. [14] studied (M, N)-policies for $GI/G/1$ queues, and [2] studied N-policies for $M/M/\infty$ queues.

The difference between the results for single-server queues (optimality of N-policies) and multi-server queues (optimality of (M, N)-policies and suboptimality of N-policies) allows the following intuitive explanations. In classic single-server queues, the same resource (the single server) is allocated to each customer during its service, and this does not depend on when the customer is served. Therefore, to minimize average switching and holding costs, the system should be always run when it is not empty. For queues with parallel servers and running costs independent of the number of customers in service, the costs to run the servers are spread among the customers in service. Therefore, if the number of customers in system is small, the cost to serve each customer is relatively high. Thus, it may be profitable to switch the system off when the number of customers in the system is small (though it is positive). Among other results, the current paper proves this concept for $M/M/\infty$ queues.

Reference [10] analyzes a queueing model of software releases, computes performance characteristics for N policies, and finds the best N-policy. In this model, the customers are software users, and each release serves each customer waiting for a software release with a probability p. Thus, if there are n customers waiting for service, on average np of them will be served by a software release. Similarly to $M/M/\infty$, for this model the number of simultaneously served customers increases with the number of customers in system, and service times are independent on the number of customers in the system. Though the queue studied in in [10] differs from $M/M/\infty$, these two models have the same fluid approximations. Therefore, the results of the current paper can provide insights regarding the optimal policies for the software maintenance model considered in [10].

Another possible application is a scalable IT system. Consider an IT application or a server that is capable of parallel processing of a large number of requests. Additional request have negligible effects on the times to process other requests. This system can be approximately modelled by a parallel queue with an unlimited number of servers. Therefore, if

This research was partially supported by NSF grants CMMI-0900206 and CMMI-0928490.

E. Finberg is with the Department of Applied Mathematics and Statistics, Stony Brook University, Ny 11794-3600, USA Eugene.Feinberg@sunysb.edu

X. Zhang is with IBM T.J. Watson Research Center, Yorktown Heights, NY 10598, USA zhangxia@us.ibm.com

the number of requests is small then it may be profitable to turn the system off or switch it to other operations, such as maintenance, if there are customers in the system, but their number is small.

While it is easy to derive performance characteristics for $M/G/1$ queues controlled by (M, N)-policies, this problem is nontrivial for $M/M/\infty$ queues. As far as we know, there are no finite formulae for expectations of busy periods and waiting times for $M/M/\infty$ queues controlled by (M, N)-policies. This complicates the analysis of the problem studied in this paper. Performance characteristics for specific multi-server queues with switching on and off certain groups of customers were recently analyzed in [19] and [16], and a genetic algorithm to find the closed-form multi-threshold control policy for a two-server case is presented in [11].

A controlled $M/M/\infty$ queue can be modeled as a Continuous-Time Markov Decision Process. Since the service rate is $i\mu$ when there are i customers in the system and i is not bounded, the transition rates in this CTMDP are not bounded. This is an additional complication. CTMDPs with unbounded transition rates were recently studied by [6]. However, it was assumed there that any stationary policy defines an ergodic continuous-time Markov chain. This condition does not hold for the problem we consider because the policy that always keeps the system off defines a transient Markov chain. Therefore, in this paper we provide a direct analysis of this problem.

First we study the expected total discounted costs. The problem with the expected total discounted costs can be reduced to a discrete-time problem with the expected total costs. Since the transition rates are unbounded, the total costs for discrete-time problem cannot be presented as total discounted costs with the discounted factor smaller than 1. However, since all the costs are nonnegative, the resulted discrete-time problem belongs to the class of negative MDPs. For this negative MDP we derive the optimality equation, show that the value function is finite, and establish the existence of stationary optimal policies; see Theorem 1.

Then, in Section III-B we investigate the discrete-time total-cost problem limited to the policies that never turn the running system off. Such policies are called full-service. We show that within the class of full-service policies there exist stationary total-cost optimal policies. These policies are computed explicitly in Theorem 3. They are defined by a number n_α such that the system should be switched on as soon as the number of customers is greater or equal than n_α, where $\alpha > 0$ is the discount rate. The important feature of the function n_α is that it is increasing in α and therefore bounded when $\alpha \in [0, \alpha^*]$ for any $\alpha \in (0, \infty)$.

In Section III-D we prove the existence of stationary Blackwell-optimal policies and describe their structure. A policy is called Blackwell-optimal, if there exists $\alpha^* > 0$ such that this policy is discount-optimal for any discount rate $\alpha \in (0, \alpha^*]$.

II. Problem Formulation

We model the above described problem as a CTMDP. The state space is $S = \mathbb{N}\times\{0, 1\}$, where $\mathbb{N} = \{0, 1, \ldots\}$. If the state of the system at the decision epoch n is $x_n = (X_n, \delta_n) \in S$, this means that the number of customers in the system is X_n and the state of the system is δ_n, with $\delta_n = 1$ if the system is on and $\delta_n = 0$ if the system is off. The action set is $A = \{0, 1\}$, where $a = 0$ means that the system should be off and $a = 1$ means that the system should be on. If the action a_n is selected at a decision epoch t_n, when the system is at a state (X_n, δ_n), the system is switched immediately, if $a_n \neq \delta_n$, and its status (on or off) remains unchanged, if $a_n = \delta_n$. In particular, if the system is off, that is $\delta_n = 0$, the decision $a_n = 1$ turns the system on, and, if the system is on, that is $\delta_n = 1$, the decision $a_n = 0$ turns it off.

If the system is off or $X_n = 0$, the time until the next jump epoch, which is an arrival, has an exponential distribution with the intensity λ. If $X_n = i > 0$ and the system is on, the time until the next jump epoch has an exponential distribution with the intensity $\Lambda_i = \lambda + i\mu$, and this jump is an arrival with the probability λ/Λ_i and a departure with the probability $i\mu/\Lambda_i$.

A history of the process up to nth jump, $n = 0, 1, \ldots$, is the sequence $t_0, x_0, a_0, \ldots, t_{n-1}, x_{n-1}, a_{n-1}, t_n, x_n$. Let H_n be the set of all histories up to n*th* decision epoch. Then $H = \cup_{0\le n<\infty} H_n$ is the set of all histories that contain a finite number of decision epochs. A (possibly randomized) policy π is defined as a transition probability from H to A such that $\pi(A|h_n) = 1$ for each $h_n \in H, n = 0, 1, \ldots$. A *stationary* policy is defined by a mapping $\pi : S \to A$ such that $\pi(x) \in A, x \in S$.

For each initial state of the system $x_0 = (i, \delta)$, and for any policy π, the policy π defines a stochastic sequence $\{x_n, a_n, t_n, n = 0, 1, \ldots\}$, where $t_0 = 0$ and $t_{n+1} \ge t_n$. We denote by $E^\pi_{x_0}$ the expectation of this process.

Now we define the cost function. If $x_n = \{i, \delta\}$, and an action a is selected then the cumulative cost during the interval $[t_n, t_{n+1}]$, where $0 \le t_n \le t_{n+1}$ is

$$c(i, \delta, a, t_n, t_{n+1}) = \int_{t_n}^{t_{n+1}} (hi + cI\{\delta = 1\})dt + s_a|\delta - a|,$$

where I is the indicator function. The cumulative cost over the interval t is

$$C(t) = \sum_{n=0}^{N(t)-1} c(X_n, \delta_n, a_n, t_n, t_{n+1}) + c(X_{N(t)}, \delta_{N(t)}, a_{N(t)}, t_{N(t)}, t),$$

where $N(t)$ is the number of jump epochs up to time t. Thus, $N(t)$ does not count the jump at $t_0 = 0$.

Let $N_1(t)$ be the number of arrivals and $N_2(t)$ be the number of departures by time t. Since $N_1(t)$ is a Poisson process then with probability 1 $N_1(t) < \infty$ for $t < \infty$ and $N_1(t) \to \infty$ as $t \to \infty$. Since $N(t) = N_1(t) + N_2(t)$ and $N_2(t) \le N_1(t) + X_0$, we have that $N_1(t) \le N(t) \le 2N_1(t) + X_0$. This implies that with probability 1 $N_1(t) < \infty$ for $t < \infty$ and $N_1(t) \to \infty$ as $t \to \infty$. Thus with probability 1 all the epochs t_n are finite and $t_n \to \infty$ as $n \to \infty$.

We observe that $C(t) = \infty$ with probability 1 when $N(t) = \infty$. We define the state of the server at time t as $\delta(t) = \delta_n$

for $t_n \le t \le t_{n+1}$, and the number of customers at time t as $X(t) = X_n$ for $t_n \le t \le t_{n+1}$. Using these notations, we can rewrite

$$C(t) = \int_0^t (hX(t) + c\delta(t))dt + \sum_{n=0}^{N(t)} s_{a_n}|a_n - \delta_n|, \qquad (1)$$

where we use that $|a - \delta| = 0$, if $a = \delta$, and $|a - \delta| = 1$, if $a \ne \delta$. Observe that $C(t)$ is a nondecreasing nonnegative function.

For any initial state of the system $x_0 = (i, \delta)$, and for any policy π, the expected total discounted cost over the infinite horizon is

$$V_\alpha^\pi(i,\delta) = E_{(i,\delta)}^\pi \int_0^\infty e^{-\alpha t} dC(t)$$

$$= E_{(i,\delta)}^\pi [\int_0^\infty e^{-\alpha t}(hX(t) + \delta(t)c)dt + \sum_{n=0}^{\infty} e^{-\alpha t_n}|a_n - \delta_n| s_{a_n}]. \quad (2)$$

The average cost per unit time is defined as

$$v^\pi(i,\delta) = \limsup_{t\to\infty} t^{-1} E_{x_0}^\pi C(t). \qquad (3)$$

Let

$$V_\alpha(i,\delta) = \inf_\pi V_\alpha^\pi(i,\delta), \qquad (4)$$

$$v = \inf_\pi v^\pi(i,\delta). \qquad (5)$$

A policy φ is called discount-optimal if $V_\alpha^\varphi(i,\delta) = V_\alpha(i,\delta)$ for any policy π and for any $x_0 = (i,\delta)$. A policy φ is called average-cost optimal if $v^\varphi(i,\delta) = v$ for any policy π and for any $x_0 = (i,\delta)$. If π is a Blackwell optimal policy and the limit in (3) exists then π is average-cost optimal. This follows from Tauberian theorems [18]. An average-cost optimal policy may not be Blackwell optimal.

III. Discounted Cost Criterion

In this section we study the expected total cost criterion.

A. Reduction to Discrete Time and Existence of Stationary Discount-Optimal Policies

In this subsection, we formulate the optimality equation, prove the existence of stationary discount-optimality equations. This is done by reduction of our problem to discrete time.

When the system is on and there are i customers in it, the time until the next jump has an exponential distributions with intensity $\Lambda_i \to \infty$ as $i \to \infty$. Since the jump rates are unlimited, it is impossible to present the problem as a discounted MDP with the discount rate smaller than 1. Thus, we shall present our problem as a negative MDP.

A discrete time MDP is called negative [15], [1], [12], [5], if the costs are nonnegative and the goal is to minimize the expected total costs. Similar to discounted MDPs, the value function for a negative MDP satisfies the optimality equation. In addition, if the action sets are finite, there exists a stationary optimal policy. Furthermore, a stationary policy is optimal if and only if it satisfies an optimality equation. This means that for an MDP with a countable state set X, action sets $A(x)$, transition probabilities $p(y|x,a)$, and nonnegative one-step costs $c(x,a)$, a stationary policy ϕ is optimal if and only if for all $x \in X$ it satisfies

$$V(x) = c(x,\phi(x)) + \sum_{y\in X} p(y|x,\phi(x))V(y), \qquad (6)$$

where $V(x)$ is the infimum of the expected total costs starting from state x. In addition, the value function $V(x)$ satisfies the optimality equation

$$V(x) = \min_{a\in A(x)} \{c(x,a) + \sum_{y\in X} p(y|x,a)V(y)\}, \qquad x \in X. \quad (7)$$

For our queueing control problem, define the following values:

$$\beta(i,a) = \begin{cases} \dfrac{\lambda}{\lambda+\alpha}, & \text{if } a = 0, \\ \dfrac{\Lambda_i}{\Lambda_i+\alpha}, & \text{if } a = 1, \end{cases} \qquad (8)$$

$$p(j|i,a) = \begin{cases} 1, & \text{if } a = 0,\ j = i+1, \\ \dfrac{\lambda}{\Lambda_i}, & \text{if } a = 1,\ j = i+1, \\ \dfrac{i\mu}{\Lambda_i}, & \text{if } i > 0,\ a = 1\ j = i-1, \\ 0, & \text{otherwise}, \end{cases} \qquad (9)$$

and $c((i,0),0) = \dfrac{hi}{\lambda+\alpha}$, $c((i,1),0) = s_0 + c((i,0),0)$, $c((i,1),1) = \dfrac{hi+c}{\Lambda_i+\alpha}$, and $c((i,0),1) = s_1 + c((i,1),1)$. Let

$$p_\alpha(j|i,a) = \beta(i,a)p(j|i,a)).$$

We follow the conventions that $p_\alpha(-1|i,a) = 0$, $V_\alpha(-1,\delta) = 0$, $\sum_\emptyset = 0$, 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 $i = 0, 1, \ldots$

$$V_\alpha(i,\delta) \le (1-\delta)s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha} < \infty. \qquad (10)$$

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

$$\begin{aligned} V_\alpha(i,\delta) = &\min_{a\in\{0,1\}} \{c((i,\delta),a) + p_\alpha(i-1|i,a)V_\alpha(i-1,a) \\ &+ p_\alpha(i+1|i,a)V_\alpha(i+1,a)\} = \min\{(1-\delta)s_1 \qquad (11) \\ &+ \frac{hi+c}{\alpha+\Lambda_i} + \frac{\lambda}{\alpha+\Lambda_i}V_\alpha(i+1,1) + \frac{i\mu}{\alpha+\Lambda_i}V_\alpha(i-1,1), \delta s_0 \\ &+ \frac{hi}{\alpha+\lambda} + \frac{\lambda}{\alpha+\lambda}V_\alpha(i+1,0)\}. \end{aligned}$$

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

$$\begin{aligned} V_\alpha(i,\delta) = \min_{\phi(i,\delta)\in[0,1]} &\{c((i,\delta),\phi(i,\delta)) \\ &+ p_\alpha(i-1|i,\phi(i,\delta))V_\alpha(i-1,\phi(i,\delta)) \qquad (12) \\ &+ p_\alpha(i+1|i,\phi(i,\delta))V_\alpha(i+1,\phi(i,\delta))\}. \end{aligned}$$

Because of Theorem 1(iii), we consider only stationary policies in the remaining part of the paper. Define $V^1_\alpha(i,\delta)$ and $V^0_\alpha(i,\delta)$ as follows:

$$V^1_\alpha(i,\delta) = (1-\delta)s_1 + \frac{hi+c}{\alpha+\Lambda_i} + \frac{\lambda}{\alpha+\Lambda_i}V_\alpha(i+1,1) + \frac{i\mu}{\alpha+\Lambda_i}V_\alpha(i-1,1),$$
$$V^0_\alpha(i,\delta) = \delta s_0 + \frac{hi}{\alpha+\lambda} + \frac{\lambda}{\alpha+\lambda}V_\alpha(i+1,0). \tag{13}$$

The following lemma follows from optimality equation (11).

Lemma 3.1: The following properties hold for the function $V_\alpha(i,\delta)$.

(a) If $V_\alpha(i,0) = V^1_\alpha(i,0)$, then $V_\alpha(i,1) = V^1_\alpha(i,1)$.
(b) If $V_\alpha(i,1) = V^0_\alpha(i,1)$, then $V_\alpha(i,0) = V^0_\alpha(i,0)$.
(c) $-s_1 \le V_\alpha(i,1) - V_\alpha(i,0) \le s_0$.

B. Full Service Policy

In subsection III-A, we defined the MDP action sets $A(i,\delta) = A = \{0,1\}$ for all $i = 0,1,\ldots$ and for all $\delta = 0,1$. The class of the policies that never turns the running server off is the class of all policies in the MDP with $A(i,0) = A$ and $A(i,1) = \{1\}$, $i = 0,1,\ldots$. This is a sub-model of our original model. Define by (4) $U_\alpha(i,\delta)$ as the optimal total discount cost for this new MDP. From (10) we have

$$V_\alpha(i,\delta) \le U_\alpha(i,\delta) \le (1-\delta)s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}. \tag{14}$$

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

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

$$U_\alpha(i,1) = \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}. \tag{15}$$

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

$$U_\alpha(i,0) = \min\{s_1 + \frac{hi+c}{\alpha+\Lambda_i} + \frac{\lambda}{\alpha+\Lambda_i}U_\alpha(i+1,1) + \frac{i\mu}{\alpha+\Lambda_i}U_\alpha(i-1,1), \frac{hi}{\alpha+\lambda} + \frac{\lambda}{\alpha+\lambda}U_\alpha(i+1,0)\}. \tag{16}$$

Proof:

(i) Let π be the policy that never turns the running system off. $U_\alpha(i,1) = V^\pi_\alpha(i,1) = \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}$.
(ii) Since $U_\alpha(i,0)$ is the optimal discount cost for the sub-model of the original MDP, it satisfies the optimality equation of the original MDP. Thus, (16) follows from (11). ■

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

The following theorem implies that an n-full service policy is discount-optimal within the class of policies that never turn the running system off.

Theorem 3: A policy ϕ is discount optimal within the class of the policies that never turn off the server 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

$$A(\alpha) = \frac{(\mu+\alpha)(c+\alpha s_1)}{h\mu}. \tag{17}$$

The following definition and lemmas are used in the proof of Theorem 3.

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

Lemma 3.2: For any $\alpha > 0$, the passive policy φ is not optimal within the class of policies that never turn off the running system. Furthermore, $V^\varphi_\alpha(i,0) > U_\alpha(i,0)$ for all $i = 0,1,\ldots$.

Lemma 3.3: Let ψ be the policy that turns the system on at time 0 and keeps it on forever, and π be the policy that waits for one arrival and then turns the system on and keeps it on forever. Then

$$\begin{cases} V^\pi_\alpha(i,0) > V^\psi_\alpha(i,0), & \text{if } i > A(\alpha), \\ V^\pi_\alpha(i,0) < V^\psi_\alpha(i,0), & \text{if } i < A(\alpha), \\ V^\pi_\alpha(i,0) = V^\psi_\alpha(i,0), & \text{if } i = A(\alpha); \end{cases}$$

where $A(\alpha)$ is as in (17).

Proof: Let ϕ be a stationary optimal policy within the class of the policies that never turn off the running system. Let ψ be the policy that turns the system on at time 0 and keeps it on forever, and π be the policy that waits for one arrival and then turns the system on and keeps it on forever. By (16),

$$V^\phi_\alpha(i,0) = \min\left\{s_1 + U_\alpha(i,1), \frac{hi}{\lambda+\alpha} + \frac{\lambda}{\lambda+\alpha}U_\alpha(i+1,0)\right\}. \tag{18}$$

We show that if $i > A(\alpha)$, then $\phi(i,0) = 1$. Indeed, let $\phi(i,0) = 0$ for some $i > A(\alpha)$. By Lemma 3.2, $\phi(j,0) = 1$ for some $j > i$. Thus, there exists an $i^* \ge i$ such that $\phi(i^*,0) = 0$ and $\phi(i^*+1,0) = 1$. This implies that $V^\psi_\alpha(i^*,0) \ge V^\pi_\alpha(i^*,0)$, where $i^* > A(\alpha)$. By Lemma 3.3, this is a contradiction. Thus $\phi(i,0) = 1$ for all $i > A(\alpha)$.

If $i < A(\alpha)$, then Lemma 3.3 implies $V^\pi_\alpha(i,0) < V^\psi_\alpha(i,0)$. Thus $\phi(i,0) = 0$ for all $i < A(\alpha)$.

Let $A(\alpha)$ be an integer and $i = A(\alpha)$. In this case, Lemma 3.3 implies $V^\psi_\alpha(i,0) = V^\pi_\alpha(i,0)$. From (16), $V^\psi_\alpha(i,0) = V^\pi_\alpha(i,0) = U_\alpha(i,0) = \min\left\{U^\psi_\alpha(i,0), U^\pi_\alpha(i,1)\right\}$. Thus $\phi(i,0) = 1$ or $\phi(i,0) = 0$. ■

Corollary 1: Let

$$n_\alpha = \lceil A(\alpha) \rceil, \tag{19}$$

where $A(\alpha)$ is as in (17). Then for $i < n_\alpha$

$$U_\alpha(i,0) = \sum_{k=0}^{n_\alpha - i - 1} \left(\frac{\lambda}{\lambda+\alpha}\right)^k \frac{h(i+k)}{\lambda+\alpha} + \left(\frac{\lambda}{\lambda+\alpha}\right)^{n_\alpha - i} \left(s_1 + \frac{hn_\alpha}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}\right), \quad (20)$$

and for $i \geq n_\alpha$

$$U_\alpha(i,0) = s_1 + \frac{hi}{\mu+\alpha} + \frac{h\lambda}{\alpha(\mu+\alpha)} + \frac{c}{\alpha}, \quad (21)$$

Proof: Theorem 3 implies that n_α-full service policy is discount-optimal within the class of policies that never turn off the running system, where n_α is as in (19). ■

C. Reduction to Finite State Space and Existence of Blackwell Optimal Policies

In this section, we explore the existence of Blackwell optimal [3], [9] policy. Define

$$N^*_\alpha = \min\{i \geq 0 : V^1_\alpha(j,0) \leq V^0_\alpha(j,0), \text{ for all } j \geq i\}. \quad (22)$$

The following lemma implies that N^*_α is well defined.

Lemma 3.4: $N^*_\alpha \leq n_\alpha$ for all $\alpha > 0$.

From Lemma 3.4, N^*_α is bounded from above by n_α for each α. Define an SMDP with finite state space $S' = \{0,1,\dots,n_\alpha\} \times \{0,1\}$.The state of this SMDP at the decision epoch n is $x_n = (X_n, \delta_n) \in S'$. The action set $A = \{0,1\}$ is the same as the original CTMDP. The time until the next decision epoch is the same as the original CTMDP for $X_n = 0,1,\dots,n_\alpha - 1$ and $\delta_n = 0,1$. When at state $(n_\alpha, 1)$, let τ be a random variable that represents the first time the system returns to $(n_\alpha, 1)$ before transition to $(n_\alpha - 1, 1)$. The transition probabilities $\tilde{p}(j|i,a) = p(j|i,a)$ from (9) for $i,j = 0,1,\dots,n_\alpha - 1$ and $a = 0,1$, except that $\tilde{p}(n_\alpha|n_\alpha,a) = 1$. The one step cost $\tilde{c}((i,\delta),a) = c((i,\delta),a)$ for $i = 0,1,\dots,n_\alpha - 1$, $\delta = 0,1$ and $a = 0,1$, except $\tilde{c}((n_\alpha,1),1) = E\left[\int_0^\tau (c + hX(t))dt\right]$, and $\tilde{c}((n_\alpha,0),1) = s_1 + \tilde{c}((n_\alpha,1),1)$, where $X(t)$ is the system size at time t. Denote by $\tilde{V}_\alpha(i,\delta)$ as the optimal total discounted cost for this SMDP. Define T'_i as the time for the number of customers in the system becomes $i-1$ if at time 0 it is $i = 1,2,\dots$ if the system is running all the time. Let $C'_{i,\alpha}$ be the total holding and serving costs during T'_i, i.e.

$$C'_{i,\alpha} = \int_0^{T'_i} (c + hX(t))dt. \quad (23)$$

We show next that this SMDP is equivalent to the original CTMDP, i.e., $\tilde{V}_\alpha(i,\delta) = V_\alpha(i,\delta)$ for all $i = 0,1,\dots$ and $\delta = 0,1$.

Lemma 3.5: If $V^\varphi_\alpha(i,\delta) = V_\alpha(i,\delta)$ for $i = 0,1,\dots,n_\alpha$ and $\varphi(i,\delta) = 1$ for all $i > n_\alpha$, then $V^\varphi_\alpha(i,\delta) = V_\alpha(i,\delta)$ for $i = n_\alpha + 1, n_\alpha + 2, \dots$.

Proof: Proof of Lemma 3.5. For an $\alpha^* > 0$, let φ be an optimal stationary policy for the SMDP defined above. Define φ^* for the original CTMPD as

$$\varphi^* = \begin{cases} \varphi(i,\delta), & \text{if } i \leq n_{\alpha^*}, \\ 1, & \text{otherwise.} \end{cases} \quad (24)$$

We show that for $\alpha \in (0,\alpha^*]$, $\tilde{V}_\alpha(i,\delta) = V_\alpha(i,\delta)$, for all $i = 0,1,\dots$ and $\delta = 0,1$. Indeed,

$$\tilde{V}_\alpha(i,\delta) = \tilde{V}^\varphi_\alpha(i,\delta) = V^{\varphi^*}_\alpha(i,\delta) \leq V_\alpha(i,\delta).$$

On the other hand, since $V_\alpha(i,\delta) = V^1_\alpha(i,\delta)$ for all $i \geq N_\alpha$, then

$$V_\alpha(i,\delta) \leq V^{\varphi^*}_\alpha(i,\delta) = \tilde{V}^\varphi_\alpha(i,\delta) = \tilde{V}_\alpha(i,\delta).$$

Thus each optimal stationary policy for the reduced SMDP is also optimal for the original CTMDP. ■

Theorem 4: There exist a Blackwell optimal policy for the original model.

Proof: Proof of Theorem 4. Consider the SMDP defined before Lemma 3.5. Since $\tilde{V}^\phi_\alpha(i,\delta) > 0$ for all $\alpha > 0$ and any for any policy π, then $\alpha = 0$ is the isolated singularity of every function $\alpha\tilde{V}^\pi_\alpha(i,\delta), \alpha > 0$. According to [3, Theorem 3], this implies that the reduced SMDP has a Blackwell optimal policy φ. Because of Lemma 3.5, the policy φ^* defined in (24) is Blackwell optimal for the original problem. ■

D. Structure of Blackwell Optimal Policies

Definition 3: A policy is called (M,N)-policy if there exists two integers M and N, with $0 \leq M < N < \infty$, such that at state $(i,0)$, leave the system off if $i < N$ and turn on the system if $i \geq N$; at state $(i,1)$, leave the system on if $i > M$ and turn off the system if $i \leq M$.

The main result of this section is Theorem 5.

Theorem 5: Let $n = \lim_{\alpha\to 0} n_\alpha = \lfloor c/h + 1 \rfloor$, where n_α is as in (19).

(i) When $c < \lambda(s_0 + s_1)/n + h(n-1)/2$, the n-full service policy is Blackwell optimal;

(ii) When $c > \lambda(s_0 + s_1)/n + h(n-1)/2$, there exist two integers M and N, with $0 \leq M < N \leq n$, such that the (M,N)-policy is Blackwell optimal.

Consider an $\alpha^* > 0$ such that a Blackwell optimal policy is discount optimal for all $\alpha \in (0,\alpha^*]$. By Definition 3, M is the threshold that we switch off the system upon a departure, i.e.

$$M = \max\{i \geq 0 : V^0_\alpha(i,1) \leq V^1_\alpha(i,1)\}, \qquad \alpha \in (0,\alpha*]. \quad (25)$$

In view of Theorem 5 (ii), M is well defined. Let N be the threshold that we turn on the system upon an arrival,, i.e.

$$N = \min\{i \geq 0 : V^1_\alpha(i,0) < V^0_\alpha(i,0)\}, \qquad \alpha \in (0,\alpha*]. \quad (26)$$

$0 \leq N \leq N^*_\alpha$ for $\alpha \in (0,\alpha^*]$, where N does not depend on α. We first provide some lemmas before proving Theorem 5.

Lemma 3.6: There exists an $\alpha^* > 0$ such that for all $\alpha \in (0,\alpha^*]$, $E[C'_{i,\alpha}] = \dfrac{h}{\mu+\alpha} + vE[T'_i] + O(\alpha)$, where v is as in (5).

The next lemma implies the monotonicity property of $E[T'_i]$. The stochastic monotonicity for stationary recurrence times in queueing control model with a removable server is considered in [4].

Lemma 3.7: $E[T'_i] - E[T'_{i+1}] > 0$ and is decreasing in i for $i = 1,2,\dots$.

Lemma 3.8: There exists some $\alpha^* > 0$ such that for $\alpha \in (0, \alpha^*]$, if $V_\alpha(i, 1) = V_\alpha^1(i, 1)$, then $V_\alpha(i+1, 1) = V_\alpha^1(i+1, 1)$.

Corollary 2: $V_\alpha(i, 1) = V_\alpha^1(i, 1)$ for all $i > M$ and $V_\alpha(i, 1) = V_\alpha^0(i, 1)$ for all $i \leq M$, where M is as in (25).

Proof: By definition of M, $V_\alpha(M, 1) = V_\alpha^0(M, 1)$ and $V_\alpha(i, 1) = V_\alpha^1(i, 0)$ for all $i > M$. Assume that there exists an $0 \leq i \leq M$ such that $V_\alpha(i, 1) = V_\alpha^1(i, 1)$. By Lemma 3.8, $V_\alpha(i+1, 1) = V_\alpha^1(i+1, 1)$ and by induction we have $V_\alpha(M, 1) = V_\alpha^1(M, 1)$. This is a contradiction. ■

Lemma 3.9: There exists some $\alpha^* > 0$ such that for $\alpha \in (0, \alpha^*]$, if $V_\alpha(i, 0) = V_\alpha^1(i, 0)$, then $V_\alpha(i+1, 0) = V_\alpha^1(i+1, 0)$.

Corollary 3: There exists some $\alpha^* > 0$ such that for $\alpha \in (0, \alpha^*]$, $V_\alpha(i, 0) = V_\alpha^0(i, 0)$ for all $i < N$, and $V_\alpha(i, 0) = V_\alpha^1(i, 0)$ for all $i \geq N$, where N is as in (26).

Proof: By definition of N, $V_\alpha(N, 0) = V_\alpha^1(N, 0)$ and $V_\alpha(i, 0) = V_\alpha^0(i, 0)$ for all $i < N$. From Lemma 3.9, $V_\alpha(i, 0) = V_\alpha^1(i, 0)$ for all $i > N$. ■

Corollary 4: There exists $\alpha^* > 0$ such that for all $\alpha \in (0, \alpha^*]$, $N = N_\alpha^*$,where N_α^* is as in (22).

Proof: By Corollary 3, $V_\alpha(i, 0) = V_\alpha^1(i, 0)$ for all $i \geq N$, thus $N \geq N^*$. On the other hand, $V_\alpha(N^*, 0) = V_\alpha^1(N^*, 0)$, thus $N^* \geq N$. ■

Corollary 5: There exists some α^*, such that for all $\alpha \in (0, \alpha^*]$, $M < N$

Proof: Assume that $N \leq M$. Let i be such that $N \leq i \leq M$. By Corollary 3, $V_\alpha(i, 1) = V_\alpha^1(i, 1)$. However, by Corollary 2, $V_\alpha(i, 1) = V_\alpha^0(i, 1)$. This is a contradiction. ■

IV. Average Cost Criterion

The theorem below follows from Theorem 5. It describes the structure of average-cost optimal policies.

Corollary 6: For average costs per unit time, let $n = \lfloor c/h + 1 \rfloor$. Then

(i) When $c \leq \lambda(s_0 + s_1)/n + h(n-1)/2$, any full-service policy is average-cost optimal.

(ii) When $c > \lambda(s_0 + s_1)/n + h(n-1)/2$, there exist two integers M and N, with $0 \leq M < N \leq n$, such that the (M, N)-policy is average-optimal.

V. Conclusions

This paper describes optimal policies for switching on and off the full service capacity of an $M/M/\infty$ queue. The cost structure consists of linear holding costs, the cost to run the systems per unit time, and the switching costs. For average costs per unit time, we prove that there is an optimal policy that either always runs the system or is an (M, N)-policy. An (M, N)-policy, where $N > M \geq 0$, switches the system on when there are N or more customers in the system and turns it off when the number of customers in the system is M or less than M. Unlike the case of single-server queues, $(0, N)$-optimal policies may not be optimal. The described structure of average-cost optimal policies has been established by describing the properties of optimal policies for the expected discounted total costs when the discount rate is close to 0 (the so-called Blackwell optimal policies).

References

[1] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York; 1978.

[2] S. Browne and O. Kella, Parallel service with vacations, *Operations Research*, **44**, 1995, pp 870-878.

[3] E.V. Denardo, E.V. Markov renewal programs with small interest rates, *The Annals of Mathematical Statistics*, **42**, 1971, pp 477-496.

[4] E.V. Denardo, E.A. Feinberg and O. Kella. Stochastic monotonicity for stationary recurrence times of first passage heights, *The Annals of Applied Probability,* **7**, 1997, pp 326-339.

[5] E.A. Feinberg, Total reward criteria, *Handbook of Markov Decision Processes: Methods and Applications*, Kluwer, Boston, 2002, pp 173-207.

[6] X. Guo and O. Hernández-Lerma. *Continuous-Time Markov Decision Processes*, Springer, Heidelberg; 2009.

[7] N. Hadidi, Busy period of queues with state dependent arrival and service rates, *Journal of Applied Probability*, **11**, 1974, 842–848.

[8] D.P. Heyman, Optimal operating policies for $M/G/1$ queuing systems, *Operations Research,* **16**, 1968, pp 362-382.

[9] A. Hordijk and A. Yushkevich, Blackwell optimality, *Handbook of Markov Decision Processes: Methods and Applications*, Kluwer, Boston, 2002, pp 231-267.

[10] V.G. Kulkarni, S. Kumar, V.S. Mookerjee, S.P. Sethi, Optimal Allocation of Effort to Software Maintenance: A Queuing Theory Approach, *Production and Operations Management,* **18**, 2009, pp 506–519.

[11] C.H. Lin and J.C. Ke, Genetic algorithm for optimal thresholds of an infinite capacity multi-server system with triadic policy, *Expert Systems with Applications*, **37**, 2010, 4276–4282.

[12] M.L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley, New York; 1994.

[13] S.M. Ross, *Stochastic processes*, Wiley, New York; 1996.

[14] M.J. Sobel, Optimal average-cost policy for a queue with start-up and shut-down costs, *Operations Research*, **17**, 1969, 145–162.

[15] R.E. Strauch, Negative dynamic programming, *The Annals of Mathematical Statistics*, **37**, 1966, 871–890.

[16] N. Tian and Z.G. Zhang, A two threshold vacation policy in multi-server queueing systems, *European Journal of Operational Oesearch*, **168**, 2006, 153–163.

[17] H.C. Tijms, *A first course in stochastic models*, Wiley, New York; 2003.

[18] D.V. Widder, *The Laplace Transform*, Princeton University Press, Princeton; 1941.

[19] Z.G. Zhang, On the three threshold policy in the multi-server queueing system with vacations, *Queueing Systems*, **51**, 2005, 173–186.