

# STOCHASTIC OPTIMAL DYNAMIC CONTROL OF $GI_m/GI_m/1_n$ QUEUES WITH TIME-VARYING WORKLOADS

YINGDONG LU, MAYANK SHARMA, MARK S. SQUILLANTE AND BO ZHANG  
*IBM Research, Yorktown Heights, NY 10598, USA*  
*E-mail: [yingdong@us.ibm.com](mailto:yingdong@us.ibm.com)*

Motivated by applications in areas such as cloud computing and information technology services, we consider  $GI/GI/1$  queueing systems under workloads (arrival and service processes) that vary according to one discrete time scale and under controls (server capacity) that vary according to another discrete time scale. We take a stochastic optimal control approach and formulate the corresponding optimal dynamic control problem as a stochastic dynamic program. Under general assumptions for the queueing system, we derive structural properties for the optimal dynamic control policy, establishing that the optimal policy can be obtained through a sequence of convex programs. We also derive fluid and diffusion approximations for the problem and propose analytical and computational approaches in these settings. Computational experiments demonstrate the benefits of our theoretical results over standard heuristics.

## 1. INTRODUCTION

We consider a general capacity management and planning problem motivated by applications in diverse fields such as cloud computing, information technology services delivery, and modern energy management. In these applications, a high volume of demand arrives at the system and it is required that the system serves this demand in a timely fashion according to some form of service level agreement (SLA) guarantee. For this purpose, resource capacity has to be allocated to meet the demand and satisfy these guarantees. In addition, there are several important common features that have been observed in these applications that will have significant impact on the performance of the system, such as

- Time-varying demand,
- Timing of control,
- Server speed selection, and
- Server failure and departure.

The combination of these problem features results in a general class of very complex resource capacity management and planning problems.

A realistic model should accommodate the statistical patterns of the demand as it varies over time. It is also quite common in many of these systems that the time-varying workloads exhibit periodic or cyclic patterns. Seasonal effects and product cycles are some of the typical examples. In many cases, different forms and sources of periodicities can be mixed to create very complex behaviors; for example, seasonality, which is usually tangled with account cycles, might be coupled with the introduction of new products, and thus with product cycles. Hence, it is necessary to consider the time between shifts of statistical patterns of the demand to be random, where there is uncertainty in both the magnitude and direction of the pattern shifts.

It is unrealistic to assume that capacity decision can be made at the same time scale as the demand arrival; namely, we should not expect the service capacity to react in real time to any anomalies and changes in the demand, which would otherwise result unnecessarily in prohibitive costs. Meanwhile, the time scale at which control is applied should be short enough to affect the performance that will be measured against the SLA guarantees. Therefore, in addition to physical constraints for applying control, the right balance between operational cost and SLA penalties should also be included in the time scale for control.

Modern servers can be tuned to operate at different speeds. Different costs, largely related to different levels of energy consumptions, will be incurred at different speeds. At the same time, the servers will have different performance outcomes. While managing a large-scale data center of servers, it will be beneficial to have the option of being able to accelerate and slow down the speed at which each server operates for both cost and performance purposes, since it is widely known that the network performance is not necessarily monotonic with respect to the service rate of individual servers.

It is not uncommon in typical large-scale data centers that servers might leave the system under consideration for various reasons [14]. This includes servers being relocated to run other applications, being temporarily unavailable due to maintenance or upgrade, or being gone purely due to physical failure. It has been observed that such exodus of resource capacity can happen at a rather stable rate and that the impact of these effects on the performance of the system is not negligible. Hence, it is reasonable to assume that the server capacity decays at a particular rate over control periods, which aggregates the different server departure and/or failure phenomenon mentioned above, and thus allows the capacity planning process to factor these effects in its decision process.

The backbone of our mathematical model is a general queueing system under a Markov-modulated demand process. As we mentioned above, the demand model needs to be tractable but versatile enough to accommodate the randomness in both the time interval between the demand pattern shifts as well as the pattern shifts themselves. A continuous-time Markov chain model is employed to capture these characterizations of the demand pattern shifts. Although we fix the time scale of our control periods, the SLA and operational cost will be factored in the objective function, which is an indirect way of determining the timing of the control. This way, the optimal control policy will include the time scale of the control. In the first part of the paper, we use a single server model with controllable service rate to capture features such as time-varying demand, control scaling, and server failure. Some of these features are modeled approximately, but we believe our mathematical model is capturing fundamental trade-offs that provide key insights while making it possible to solve a very complex stochastic dynamic program. In the second part of the paper, we discuss how our theoretical results can be extended to multi-server queues, which can model these features more precisely, and more importantly, is capable of modeling the server at different speeds.

### 1.1. Related Work

Time-varying stochastic systems have been employed quite frequently as a model for studying system performance. For queues and queueing networks with time-varying inputs, there is an extensive literature devoted to their asymptotic performance; see, for example, [10,15]. The control of time-varying systems is studied in the area of inventory management; see, for example, [2,13]. The optimal control of an Erlang loss system is studied in [1], where structural properties of the optimal solution as well as asymptotically optimal heuristic policies are obtained. It is worth emphasizing that queueing control is exercised only at discrete time periods in the present paper. We refer to [7,12] and further references therein for work in which server capacity can be adjusted at all times.

More recently, different online algorithms have been developed for a general class of resource allocation problems that are somewhat related to those considered in this paper; see, for example, [5,9]. An optimal control approach is developed in [6] for the case when different resources have to be utilized to fulfill a common demand.

### 1.2. Our Contributions

We develop a general model with some important fundamental features.

- We provide a general and unified capacity planning model, based on a stochastic optimal control approach with general assumptions on the arrival and service patterns of the demand. Our approach also considers features such as the timing of control, server speeds, and server failure and/or departure.
- Under these very general assumptions, we derive important structural properties for the stochastic optimal control problem. More specifically, we show that the value function is concave with respect to the capacity (service rate), and hence can be obtained through a sequence of convex programming problems. These structural properties allow us to compute the optimal policy for problem sizes of interest, and also lead to some critical-point threshold optimal control policies as surrogates for larger-sized problems.
- For the general model, we also derive forms of fluid and diffusion approximations for the stochastic optimal control problem, which enable us to derive easy-to-implement heuristic policies. The performance of the systems under heuristic policies are investigated via computational experiments.

The rest of the paper is organized as follows. Section 2 presents the detailed mathematical model of the queueing system. Structural properties of the general problem are derived in Section 3. Fluid and diffusion approximations together with a heuristic analysis and related computational experiments are considered in Section 4. A discussion of a generalized model is presented in Section 5, and the paper concludes with a summary in Section 6.

## 2. MODEL AND FORMULATION

Consider a  $GI/GI/1$  queue over a time horizon of length  $L$  under time-varying workloads and time-varying control of the service rate, with  $L$  finite or infinite. The arrival and service processes vary according to an exogenous discrete time scale from one workload period to the next, indexed by  $m = 1, \dots, M$ , where workload periods have length  $T_w$ . The control of the single-server service rate varies according to an independent decision-making discrete time scale from one control period to the next, indexed by  $n = 1, \dots, N$ , where control

periods have length  $T_c$ . The period lengths  $T_w$  and  $T_c$  can be deterministic or random variables; to simplify the exposition, we focus on the case where  $T_c$  is deterministic. We use the notation  $GI_m/GI_m/1_n$  to refer to this general class of queueing systems.

Let  $\{A_m(t); t \geq 0\}$  denote the arrival process for workload period  $m$  over the time interval  $[t_{m-1}^w, t_m^w)$ , where  $A_m(t) := \sup\{k : \mathcal{A}_m(k) \leq t\}$  with finite rate  $\lambda_m$ ,  $\mathcal{A}_m(k) := \sum_{\ell=1}^k a_m^{(\ell)}$ ,  $\mathcal{A}_m(0) := 0$ , the random variable  $a_m^{(k)}$  is the interarrival time between the  $(k - 1)$ st and  $k$ th customer in workload period  $m$ ,  $k \geq 1$ , and  $a_m^{(0)} := 0$ , all defined on the same probability space. Similarly, let  $\{S_m(t); t \geq 0\}$  denote the service process for workload period  $m$  over the time interval  $[t_{m-1}^w, t_m^w)$ , where  $S_m(t) := \sup\{k : \mathcal{S}_m(k) \leq t\}$ ,  $\mathcal{S}_m(k) := \sum_{\ell=1}^k s_m^{(\ell)}$ ,  $\mathcal{S}_m(0) := 0$ , and the random variable  $s_m^{(k)}$  is the service demand for the  $k$ th customer in workload period  $m$ ,  $k \geq 1$ , all defined on the same probability space. Further, let  $\mu_n$  denote the service rate of the  $GI_m/G_m/1_n$  queue in control period  $n$  over the time interval  $[t_{n-1}^c, t_n^c)$ . Then the service time of the  $k$ th customer from workload period  $m$  served in control period  $n$  is given by  $s_m^{(k)}/\mu_n$ .

The dynamics of the time-varying workload is governed by a sequence of Markov-modulated processes with transition probability matrices  $\mathbf{P}_m = [P_{m,jj'}]$  and an initial probability vector  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_J)$ , for  $m = 1, \dots, M$  and  $j, j' = 1, \dots, J < \infty$ . More specifically, the workload starts period  $m = 1$  in state  $j$  with probability  $\alpha_j$ , remains in state  $j$  for the duration of the period of length  $T_w$ , and then transitions at the end of period  $m = 1$  to state  $j'$  with probability  $P_{1,jj'}$ , for all  $j, j' = 1, \dots, J$ . These workload dynamics continue for all subsequent periods  $m = 2, \dots, M$  starting in any state  $j$  where the workload remains in state  $j$  for the duration of the period of length  $T_w$  and then transitions to state  $j'$  with probability  $P_{m,jj'}$  at the end of period  $m$ .

The objective of our stochastic optimal control problem formulation is to determine the service rates  $\mu_n$  for every control period  $n = 1, \dots, N$  that maximize net-benefit in expectation over the entire time horizon of length  $L = N \cdot T_c$ , subject to model inputs and constraints. Let  $W_n$  denote the amount of workload remaining in the queue at the beginning of control period  $n$ ,  $\{D_n(t); t \geq 0\}$  denote the workload process over the time interval  $[t_{n-1}^c, t_n^c)$ ,  $D_n$  denote the aggregate amount of workload arriving during control period  $n$ , and  $B_n$  denote the proportion of time the server is busy over the interval  $[t_{n-1}^c, t_n^c)$ . The quantity  $W_n$  and the sequence of previous service rate decisions  $\mu_{n'}$ ,  $n' = 1, \dots, n - 1$ , are known with certainty at the start of control period  $n$ . In addition, the stochastic process  $\{D_n(t); t \geq 0\}$  is probabilistically characterized by the sequence of arrival and service processes  $\{A_m(t); t \geq 0\}$  and  $\{S_m(t); t \geq 0\}$  over the time interval  $[t_{n-1}^c, t_n^c)$ , the random variable  $D_n$  is probabilistically characterized by  $\{D_n(t); t \geq 0\}$ , and the random variable  $B_n$  is probabilistically characterized by and dependent upon  $W_n$ ,  $\{D_n(t); t \geq 0\}$  and the control decision  $\mu_n$ . The dynamics of the workload over the control periods take the form

$$W_{n+1} = W_n + D_n - \mu_n B_n T_c, \tag{1}$$

where we assume there is no server idle time between any two control periods unless all the workload has been served.

Our objective function is based on the expected total discounted net-benefit over the entire time horizon, with discount factor  $\beta$ , in which rewards are gained for the completion of workload within each control period as a function  $\mathcal{R}(\mu_n, B_n, T_c)$ , costs are incurred for the service-rate capacity deployed for each control period as a function  $\mathcal{C}(\mu_n, T_c)$ , and penalties are incurred for server idleness and workload delay within each control period as a function

$\mathcal{P}_I(\mu_n, B_n, T_c)$  and  $\mathcal{P}_D(W_n, D_n, \mu_n, B_n, T_c)$ , respectively. We consider the reward and cost functions to be respectively given by

$$\mathcal{R}(\mu_n, B_n, T_c) = r\mu_n B_n T_c, \quad \mathcal{C}(\mu_n, T_c) = c\mu_n T_c,$$

with revenue rate  $r \geq 0$  and cost rate  $c \geq 0$ . Similarly, we consider the penalty functions to be given by

$$\begin{aligned} \mathcal{P}_I(\mu_n, B_n, T_c) &= p_I \mu_n (1 - B_n) T_c, \\ \mathcal{P}_D(W_n, D_n, \mu_n, B_n, T_c) &= p_D (W_n + D_n - \mu_n B_n T_c), \end{aligned}$$

with idle and delay penalty rates  $p_I \geq 0$  and  $p_D \geq 0$ , respectively.

Define the service rate vector  $\boldsymbol{\mu} := (\mu_1, \dots, \mu_N)$ . We then have the following general stochastic dynamic program formulation of our stochastic optimal control problem over the time horizon of length  $L = N \cdot T_c$ :

$$\max_{\boldsymbol{\mu}} \mathbb{E} \left[ \sum_{n=1}^N e^{-\beta(n-1)} \left( r\mu_n B_n T_c - c\mu_n T_c - p_I \mu_n (1 - B_n) T_c - p_D (W_n + D_n - \mu_n B_n T_c) \right) \right] \tag{2}$$

$$\text{s.t. } \boldsymbol{\mu} \geq \mathbf{0}, \tag{3}$$

where  $\beta$  is the discount factor and the expectation is over the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  of the  $GI_m/GI_m/1_n$  queue. The service rate vector  $\boldsymbol{\mu}$  comprises the decision variables over the time horizon that we seek to obtain, with all other variables as input parameters.

### 3. GENERAL SOLUTION

In this section, we consider our main results for the general solution of the stochastic dynamic program (2) and (3). We start with structural properties of the optimal dynamic control policy and establish that the general solution can be obtained through a sequence of convex programs. We then turn to consider a surrogate problem of the original stochastic dynamic program and establish that it is equipped with a critical-value (threshold) optimal dynamic control policy.

#### 3.1. Structural Properties

For any control period  $n$  and any non-negative service rate  $\mu_n$ , the expected discounted net-benefit can be rewritten as

$$\mathbb{E} \left[ e^{-\beta(n-1)} \left\{ (r + p_I + p_D) \mu_n B_n T_c - (c + p_I) \mu_n T_c - p_D (W_n + D_n) \right\} \right].$$

Let  $J_n(W_n)$  be the value function for maximizing the expected discounted net-benefit of the stochastic dynamic program of interest over the time horizon from control period  $n$  to control period  $N$  with  $W_n$  workload in the queue at the start of control period  $n$ .

Then we can formulate our stochastic dynamic program in terms of the Bellman optimality equations

$$J_n(W_n) := \max_{\mu_n} U_n(\mu_n, W_n), \tag{4}$$

$$U_n(\mu_n, W_n) := \mathbb{E}[(r + p_I + p_D)\mu_n B_n T_c - (c + p_I)\mu_n T_c - p_D(W_n + D_n) + e^{-\beta} J_{n+1}(W_{n+1})], \tag{5}$$

where we assume  $J_{N+1}(W_{N+1}) = 0$  and a discount factor  $\beta$ . It is important to note that the probabilistic characterization of the random variable  $D_n$  is based on the stochastic process  $\{D_n(t); t \geq 0\}$ , which in turn is probabilistically characterized by the sequence of time-varying workload processes  $\{A_m(t); t \geq 0\}$  and  $\{S_m(t); t \geq 0\}$  over the time interval  $[t_{n-1}^c, t_n^c]$ .

One of the most interesting and important quantities in the Bellman optimality equations is  $\mu_n B_n$ . We next present the following key monotonicity result for this quantity under the general stochastic dynamic control model.

LEMMA 3.1: *For each control period  $n = 1, \dots, N$ ,  $\mu_n B_n$  is increasing and concave with respect to  $\mu_n$  almost surely.*

PROOF: Note that  $\mu_n B_n T_c$  represents the amount of work that is served within each control period  $n$ . It is easy to see that  $\mu_n B_n$  is a piecewise linear function of  $\mu_n$  that consists of two elements, the first having positive slope and the second element having slope zero. Hence, as  $T_c$  is a constant, it naturally follows that  $\mu_n B_n$  is increasing and concave with respect to  $\mu_n$  almost surely. ■

Since the expectation operator preserves monotonicity and concavity, we also have the following result.

LEMMA 3.2: *For each control period  $n = 1, \dots, N$ ,  $\mathbb{E}[\mu_n B_n]$  is increasing and concave with respect to  $\mu_n$ .*

By similar arguments, we obtain another key monotonicity result.

LEMMA 3.3: *For each control period  $n = 1, \dots, N$ ,  $\mu_n B_n$  is linear with respect to  $W_n$  when  $\mu_n B_n < T_c$ .*

Recall that the random variable  $B_n$  depends on the stochastic process  $\{D_n(t); t \geq 0\}$  and the control decision  $\mu_n$ , and that the process  $\{D_n(t); t \geq 0\}$  depends on the sequence of arrival and service processes  $\{A_m(t); t \geq 0\}$  and  $\{S_m(t); t \geq 0\}$  over the time interval  $[t_{n-1}^c, t_n^c]$ . Hence, generally speaking, there are important implications with respect to both performance and cost management for the server to be busy throughout the entire period. A more detailed analysis that includes such effects is beyond the scope of this paper, and therefore we assume that the decision variable  $\mu_n$  will always be taken from the set in which  $\mu_n B_n < T_c$  is guaranteed. Then, from Lemma 3.3,  $\mu_n B_n$  is always linear with respect to  $W_n$ .

We now can present our first main result of this section for the finite-horizon version of the stochastic optimal control problem.

THEOREM 3.1: *For each control period  $n = 1, 2, \dots, N < \infty$ , there exists a finite service-rate capacity allocation that realizes the global optimal solution of the problem (4), (5) starting at*

control period  $n$  with a given initial workload  $W_n$ , and this optimal solution is the service-rate capacity planning policy that employs the service-rate capacity  $\mu_n^*$  where  $\mu_n^*$  is the smallest service-rate capacity assignment that maximizes  $U_n(\mu_n, W_n)$ .

PROOF: The proof proceeds by induction where we first consider the basis step  $n = N$ . Suppose that the remaining workload is  $W_N$  at the beginning of the last control period and that the total amount of workload to arrive over the last control period is given by  $D_N$ . Then, from Lemma 3.3, we know that  $U_N(\mu_N, W_N)$  is concave in  $\mu_N$ . In addition, it is obvious that

$$\lim_{\mu_N \rightarrow \infty} U_N(\mu_N, W_N) = -\infty.$$

We therefore conclude that a finite optimal service-rate capacity  $\mu_N$  can be obtained through the solution of a corresponding convex program. By definition, for control period  $N - 1$ , we also know that

$$U_{N-1}(\mu, W) \geq U_N(\mu, W)$$

for any pair  $(\mu, W)$ . It then follows that

$$J_{N-1}(W) \geq J_N(W) \geq 0$$

for any feasible  $W$ .

Next, as part of the induction step, suppose the above statements are true for  $n + 1$  and consider the problem starting at control period  $n$ . We then have the following four induction-step properties:

- $U_{n+1}(\mu_{n+1}, W_{n+1})$  is concave in  $\mu_{n+1}$ ;
- $\lim_{\mu_{n+1} \rightarrow \infty} U_{n+1}(\mu_{n+1}, W_{n+1}) = -\infty$ ;
- $J_{n+1}(W) \geq J_{n+2}(W)$  for any  $W \geq 0$ .

Now, for each sample path in control period  $n$ , consider  $U_{n+1}(\mu_{n+1}, W_{n+1})$  as a function of  $\mu_n B_n$ , from which it is readily verified that the decision variable  $\mu_n$  is a convex function of  $\mu_n B_n$  and thus  $-\mu_n$  is a concave function of  $\mu_n B_n$ . Meanwhile,  $W_{n+1}$  is a linear function of  $\mu_n B_n$  with negative drift. Moreover, since  $\mu_n$  is taken from the set in which  $\mu_n B_n < T_c$  is guaranteed,  $J_{n+1}(W_{n+1})$  is increasing and linear in  $W_{n+1}$ . Combining all these properties, we conclude that  $J_{n+1}(W_{n+1})$  is a concave function of  $\mu_n B_n$ . Since the expectation operator preserves concavity, a corresponding convex program can be employed to obtain the optimal solution.

Furthermore, from the perspective of the stochastic dynamic program with respect to  $\mu_n$ , we have

$$\lim_{\mu_n \rightarrow \infty} U_n(\mu_n, W_n) = -\infty.$$

We therefore conclude there is a finite optimal service-rate capacity  $\mu_n$  that corresponds to the solution of the convex program for control period  $n$ . From (5) and the fact that  $J_{n+1}(W) \geq J_{n+2}(W)$ , we have

$$U_n(\mu, W) \geq U_{n+1}(\mu, W)$$

for any pair  $(\mu, W)$  and thus we conclude

$$J_n(W) \geq J_{n+1}(W) \geq 0$$

for any  $W \geq 0$ , completing the proof. ■

We next establish our second main result of this section for the infinite-horizon version of the stochastic optimal control problem. To this end, let us denote by  $U_n^N(\mu_n, W_n)$  and  $J_n^N(W_n)$  the Bellman optimality equations in (4) and (5) with finite  $N$ .

**THEOREM 3.2:** *Letting  $N \rightarrow \infty$ , the sequences  $\{U_1^N\}$  and  $\{J_1^N\}$  converge pointwise to limits  $U_1^\infty$  and  $J_1^\infty$ , respectively. Furthermore,  $J_1^\infty(W)$  is convex in  $W$ ,  $U_1^\infty$  is concave in  $\mu$ , and*

$$\lim_{\mu \rightarrow \infty} U_1^\infty(\mu, W) = -\infty.$$

*Hence, there exists a finite service-rate capacity vector that realizes the global optimal solution of the infinite-horizon problem (4), (5), and this optimal solution is the service-rate capacity planning policy that employs the service-rate capacity  $\mu_n^*$  where  $\mu_n^*$  is the smallest service-rate capacity assignment that maximizes  $U_1^\infty(\mu, W)$ .*

**PROOF:** The statements concerning the limits  $U_1^\infty$  and  $J_1^\infty$ , as well as the properties of  $U_1^\infty$  and  $J_1^\infty$ , directly follow from Theorem 3.1 provided there exists a finite non-decreasing function  $\mathcal{J}(W)$  such that

$$J_1^N(W) \leq \mathcal{J}(W)$$

for all  $N$ . From the induction step of Theorem 3.1, we know there exists a finite non-decreasing function  $\mathcal{U}(\mu, W)$  such that

$$U_1^N(\mu, W) \leq \mathcal{U}(\mu, W)$$

for all  $N$ . Replacing the right-hand side of (4) with  $\mathcal{U}(\mu, W)$  and applying the induction arguments of Theorem 3.1 to this revised stochastic optimization problem reveals the existence of a finite non-decreasing function  $\mathcal{J}(W)$  such that  $J_1^N(W) \leq \mathcal{J}(W)$  for all  $N$ . This together with the convex programming arguments of Theorem 3.1 complete the proof. ■

In Theorems 3.1 and 3.2 we establish crucial structural properties of the optimal solution for our stochastic dynamic program (2), (3) in the case of finite and infinite time horizons, respectively. The original stochastic dynamic program is an extremely difficult problem to solve and these structural properties are important from a theoretical perspective, as well as for developing corresponding stochastic optimal control policies in practice. In particular, we have established second-order monotonicity properties for the value function of the stochastic dynamic program (4), (5), and have ensured that the stochastic optimal control policy which solves this dynamic program can be obtained through the solutions of a sequence of convex programs. Clearly, these convex program solutions take into account the probabilistic characterization and dependence of the random variable  $B_n$  through  $\{D_n(t); t \geq 0\}$  and  $\mu_n$ , and the probabilistic characterization and dependence of the stochastic process  $\{D_n(t); t \geq 0\}$  through  $\{A_m(t); t \geq 0\}$  and  $\{S_m(t); t \geq 0\}$  over  $[t_{n-1}^c, t_n^c]$ .

We note that, within the context of our general results, the stochastic optimal control policy tends to not leave  $W_n$  either very high or very low. Hence, upon solving the convex programs for appropriately selected discrete values of  $W_n$ , then only a relatively small amount of possible values would need to be considered for each control period. This in turn can be used to significantly reduce the computational effort needed for computing the sequence of convex programs as part of a stochastic optimal control policy in practice.

Since the stochastic optimal control policy for each control period  $n$  as the optimal solutions of a sequence of convex programs depends on the initial condition  $W_n$  in a non-linear fashion, one should not expect the optimal solution to be of a critical-value (threshold) type, which is often sought after in many stochastic dynamic programming problems. This further

speaks to the difficulty of the general stochastic dynamic program (2), (3), even with steps taken to reduce the computational effort required. Hence, to address these complexities, we next turn to a surrogate problem of the original stochastic dynamic program.

### 3.2. A Surrogate Problem

We consider in this section an approximation to the stochastic dynamic programming problem (2), (3) in which the dependence of the convex programming on  $W_n$  is assumed to be linear. Such a surrogate problem is identified with the goal that the corresponding stochastic optimal dynamic control policy will be of a critical-value (threshold) type. Indeed, we will establish that this surrogate problem is equipped with a threshold optimal dynamic control policy, together with establishing performance bounds on the corresponding approximate solution relative to the original stochastic dynamic programming solution.

Specifically, the approximation we consider here consists of revising the Bellman optimality equations as follows

$$\tilde{J}_n(W_n) := \max_{\mu_n} \tilde{U}_n(\mu_n, W_n), \tag{6}$$

$$\begin{aligned} \tilde{U}_n(\mu_n, W_n) := \mathbb{E}[(r + p_I + p_D)(W_n + \mu_n \tilde{B}_n T_c) \\ - (c + p_I)\mu_n T_c - p_D(W_n + D_n) + e^{-\beta n} \tilde{J}_{n+1}(W_{n+1})], \end{aligned} \tag{7}$$

where we assume  $\tilde{J}_{N+1}(W_{N+1}) = 0$  and denote by  $\mu_n \tilde{B}_n T_c$  the work that would be served assuming there is no initial workload in the system and the service-rate capacity is  $\mu_n$ ; that is, we decouple the overall workload from  $W_n$  at the start of the control period and from the arrivals within the control period. The per-period control decision of this approximation is physically equivalent to a system in which there is an infinitely fast server available to serve all of the initial workload at the beginning of the control period.

Observe that, in the case of (6), (7), the stochastic dynamic program depends on  $W_n$  only linearly, which allows us to establish that there is a critical-value (threshold) type of optimal control policy for the revised stochastic dynamic program. Namely, as a result of the first term in the expectation of (7), for each control period  $n$  we only need to solve one convex programming problem independent of  $W_n$  in order to determine the optimal service-rate capacity policy. Meanwhile, for any  $\mu_n$ , we certainly know that

$$W_n + \mu_n \tilde{B}_n T_c \geq \mu_n B_n T_c.$$

From the setup of the problem, we also have

$$\mu_n B_n T_c \geq W_n, \quad \mu_n B_n T_c \geq \mu_n \tilde{B}_n T_c.$$

It then follows that

$$\frac{1}{2}(W_n + \mu_n \tilde{B}_n T_c) \leq \mu_n B_n T_c \leq W_n + \mu_n \tilde{B}_n T_c,$$

which in turn implies

$$\frac{1}{2} \tilde{J}_n(W_n) \leq J_n(W_n) \leq \tilde{J}_n(W_n). \tag{8}$$

The above arguments directly lead to the following approximation of the solution of our general stochastic dynamic program (2), (3) and its performance bounds.

**THEOREM 3.3:** *Under mild conditions on revisions to the original Bellman optimality equations, there exists an optimal control policy of the corresponding stochastic dynamic program that is of a critical-value (threshold) type which can be computed through (6) and (7). In addition, this optimal control policy has the performance guarantee given in (8).*

**4. APPROXIMATE SOLUTIONS**

In this section we consider various approximate solutions of the stochastic dynamic program (2), (3) primarily within a single workload period  $m$ . We start with a heuristic analysis, including comparisons of system performance under heuristic policies with respect to optimal solutions. We then derive forms of fluid and diffusion approximations for the general stochastic optimal control problem. Lastly, we present the results of computational experiments that investigate some of the properties identified herein; this includes exploring a highly dynamic workload scenario.

**4.1. Heuristic Analysis**

Let us begin by conducting a heuristic analysis with the goal of obtaining some qualitative properties of the stochastic dynamic program (2), (3). For simplicity, we assume the discount factor  $\beta = 0$  throughout this section.

First, we provide an equivalent representation of the objective function (2). From (1), we directly have

$$\mu_n B_n T_c = W_n + D_n - W_{n+1}, \tag{9}$$

which allows one to express  $B_n$  in terms of  $W_n$  and  $W_{n+1}$ . Then we can rewrite the objective function (2) using (9) to obtain

$$\mathcal{P}(\{\mu_n\}) = \mathbb{E} \left[ \sum_{n=1}^N [r\mu_n B_n T_c - c\mu_n T_c - p_I \mu_n (1 - B_n) T_c - p_D W_{n+1}] \right] \tag{10}$$

$$= \mathbb{E} \left[ \sum_{n=1}^N [(r + p_I)(W_n + D_n - W_{n+1}) - (c + p_I)\mu_n T_c - p_D W_{n+1}] \right] \tag{11}$$

$$= \mathbb{E} \left[ (r + p_I) \sum_{n=1}^N D_n - (c + p_I) T_c \sum_{n=1}^N \mu_n - p_D \sum_{n=1}^N W_n + (r + p_I + p_D)(W_1 - W_{N+1}) \right]. \tag{12}$$

Note that the first term in (12) corresponds to the exogenous demand process over which we have no control. Moreover, for a well controlled system and moderately large  $N$ , the last term in (12) is of constant order and therefore does not play a significant role from a qualitative standpoint. Hence, maximizing the expected profit  $\mathcal{P}(\{\mu_n\})$  is equivalent to minimizing

$$\min_{\mu_n} \mathbb{E} \left[ (c + p_I) T_c \sum_{n=1}^N \mu_n + p_D \sum_{n=1}^N W_n \right]. \tag{13}$$

In other words, one wants to strike a balance between the sequence of controls  $\{\mu_n\}$  and the sequence of workloads  $\{W_n\}$  relative to the corresponding costs and penalties. As is clear from (9),  $W_{n+1}$  is non-increasing in  $\mu_n$  for a fixed demand process. However, even for a fixed  $D_n$  and  $\mu_n$ , there is significant variability in  $W_{n+1}$  depending on the arrival pattern

of the customers. Furthermore, the exact cost implications are complicated by the idleness cost since capacity in excess of the unknown and uncertain demand is penalized both at the nominal procurement cost rate  $c$  and also at the idleness rate  $p_I$ .

Two intuitive control policies will be analyzed to shed light on some fundamental relationships for the problem. We first consider an *on-off* policy that alternates between setting  $\mu_n = 0$  for successive periods  $n = 1, \dots, K$ , and then setting  $\mu_{K+1} = \mu$  (for some constant  $\mu$ ). In this case, no capacity or idleness costs are incurred for the  $K$  periods of inaction with the sole penalty based on accumulated work, thus rendering

$$\mathbb{E} \left[ \sum_{n=1}^K W_n \right] = K\mathbb{E}[D_1] + (K - 1)\mathbb{E}[D_2] + \dots + \mathbb{E}[D_K] \approx K^2\mathbb{E}[D_1],$$

which grows quadratically in the length of the *off*-period (assuming  $\mathbb{E}[D_n]$  is constant over the  $K$  periods). However, even if we choose the constant *on*-capacity  $\mu = W_{K+1}$  in order to clear all the backlog, the expected revenue is  $rKE[D_1]$  and the sole expected production cost (since there are no capacity and idle costs) is  $c\mu_n T = cKE[D_1]$ . This together with the expected delay penalty above results in the following expression for expected profit over the  $K + 1$  periods

$$[(r - c)K - p_D K^2]\mathbb{E}[D_1];$$

we will later show that this expected profit can be significantly improved upon. Roughly speaking, the insight is that if we make  $\mu_n$  too small so as to consistently leave a fraction of the demand unfulfilled, then the backlog grows quadratically in the number of periods, whereas any possible profit accruable by clearing that backlog at a future time will only grow linearly in the same number of periods.

Next, we consider an *off-line* policy that assumes to know the actual values taken by the random demands  $D_n$  in advance but not the arrival pattern. It is reasonable in this case to set

$$\mu_n^{\text{offline}} = \frac{D_n}{T_c}$$

with the goal of clearing the demand within the period in which it arrives; but owing to the uncertain arrival times of the demand, there is a non-trivial probability that  $W_n \neq 0$  and  $B_n < 1$  so that we incur additional backlog and idleness penalties. The performance of this off-line policy can be better understood by studying a related  $M/G/1$  queueing system where, instead of assuming the actual offered demand is known, we assume knowledge of the first two moments of the service demand random variable  $S$ . We will then show how one can determine the optimal capacity  $\mu$  that balances all revenue and costs.

Recall the objective function (10) which is the expected value of profit. Consider the case where we do not partition the planning horizon into periods and instead want to maximize profit for an  $M/G/1$  queueing system with Poisson arrivals at rate  $\lambda$ , independent and identically distributed (i.i.d.) service demands  $S_n$  drawn from a general distribution, and constant server speed/capacity set at  $\mu$ . To maintain stability, we assume the system load  $\rho = \lambda\mathbb{E}[S]/\mu < 1$ . The average workload (or work in system) is known to be [16]

$$\mathbb{E}[W] = \frac{\lambda\mathbb{E}[S^2]}{2\mu^2(1 - \rho)}.$$

We then can express the expected profit per unit time as a function of the capacity  $\mu$  by

$$\begin{aligned} \mathcal{P}(\mu) &= r\lambda\mathbb{E}[S] - c\mu - p_I(1 - \rho)\mu - p_D \frac{\lambda\mathbb{E}[S^2]}{2\mu^2(1 - \rho)} \\ &= (r + p_I)\lambda\mathbb{E}[S] - (c + p_I)\mu - p_D \frac{\lambda\mathbb{E}[S^2]}{2\mu(\mu - \lambda\mathbb{E}[S])}. \end{aligned} \tag{14}$$

Note that the backlog penalty  $p_D$  is a crucial parameter since it helps determine how much service one should provision over and above the bare minimum amount of  $\lambda\mathbb{E}[S]$ .

It is clear that the first term in (14) is a constant, and therefore in order to maximize expected profit we want to minimize the function

$$f(\mu) = (c + p_I)\mu + p_D \frac{\lambda\mathbb{E}[S^2]}{2\mu(\mu - \lambda\mathbb{E}[S])}$$

over  $\mu > \lambda\mathbb{E}[S]$ . Such a minimizer exists because  $f(\mu) \rightarrow \infty$  both as  $\mu \downarrow \lambda\mathbb{E}[S]$  and  $\mu \rightarrow \infty$ . Differentiating with respect to  $\mu$ , we derive

$$f'(\mu) = (c + p_I) - \frac{p_D\lambda\mathbb{E}[S^2]}{2} \frac{2\mu - \lambda\mathbb{E}[S]}{\mu^2(\mu - \lambda\mathbb{E}[S])^2};$$

and then upon setting the derivative to zero we obtain

$$\frac{2(c + p_I)}{p_D\lambda\mathbb{E}[S^2]} = \frac{2\mu - \lambda\mathbb{E}[S]}{\mu^2(\mu - \lambda\mathbb{E}[S])^2}. \tag{15}$$

Solving the non-linear Eq. (15) renders the optimal capacity  $\mu_{MG1}^*$ . In Section 4.4, we will consider the special case when the service times are i.i.d.  $\text{Exp}(\theta)$  whereupon  $\mathbb{E}[S] = 1/\theta$ ,  $\mathbb{E}[S^2] = 2/\theta^2$ .

Now, we turn to consider lazy processing and workload-based policies. The following heuristic provides a lower bound to the optimal profit: In each period, set the processing rate to be

$$\mu_n^{\text{lazy}} = \frac{W_n}{T_c}.$$

Such a lazy processing approach is particularly attractive because it eliminates the need to estimate the time-varying characteristics of the demand by simply processing work with a delay of one period. Within this setting, we guarantee that the server never idles and hence  $B_n = 1$ ; moreover, all the work that arrives in a period is left unprocessed so that  $W_{n+1} = D_n$ . The objective function (11) also simplifies (since there is no idling cost) to

$$\mathcal{P}^{\text{lazy}} = \mathbb{E} \left[ (r - c - p_D) \sum_{n=1}^N W_n + p_D(W_1 - W_{N+1}) \right].$$

Assuming the system starts empty we have  $W_1 = 0$  and the objective function reduces further to

$$\mathcal{P}^{\text{lazy}} = (r - c - p_D) \sum_{n=1}^{N-1} \mathbb{E}[D_n] - p_D\mathbb{E}[D_N]. \tag{16}$$

The lazy processing approach suggests the following parametric family of policies with multiplicative factor  $\alpha > 0$ :

$$\mu_n = \frac{\alpha W_n}{T_c} = \alpha\mu_n^{\text{lazy}}.$$

For  $0 \leq \alpha \leq 1$ , it is still true that  $B_n = 1$  and hence  $W_{n+1} = (1 - \alpha)W_n + D_n$ . We therefore have

$$\begin{aligned} \mathcal{P}^\alpha &= \mathbb{E} \left[ \sum_{n=1}^N [(r + p_I)(W_n + D_n - W_{n+1}) - (c + p_I)\mu_n T_c - p_D W_{n+1}] \right] \\ &= \mathbb{E} \left[ \sum_{n=1}^N (\alpha(r - c)W_n - p_D W_{n+1}) \right] \\ &= \mathbb{E} \left[ (\alpha(r - c) - p_D) \sum_{n=1}^N W_n - p_D W_{N+1} \right]. \end{aligned}$$

It is not difficult to show that  $\mathbb{E}[W_n] \approx \frac{1}{\alpha}\mathbb{E}[D_n]$  and thus we have

$$\mathcal{P}^\alpha \approx \left( r - c - \frac{p_D}{\alpha} \right) \sum_{n=1}^N \mathbb{E}[D_n] - \frac{p_D}{\alpha} \mathbb{E}[D_{N+1}]. \tag{17}$$

Upon comparing this with (16), we observe that the  $\alpha < 1$  provides worse expected profit than  $\alpha = 1$ .

It is a difficult problem to ascertain the optimal  $\alpha^*$  for the family of lazy multiplicative policies, in part because it is hard to characterize how  $\alpha^*$  depends on the model parameters such as  $c$  and  $p_D$ . We have considered so far in this section two kinds of policies: (1) a workload-based policy and (2) an  $M/G/1$ -based policy. Both have their respective benefits: The workload-based policies do not need to estimate the statistics of the demand process and are self-regulating but at the expense of an added backlog penalty; whereas the  $M/G/1$ -based policy tries to use process statistics and estimates to set the right capacity that stays fixed over a large number of consecutive periods but at the expense of reduced flexibility and an added penalty due to excess capacity when the instantaneous load is low. We believe that the optimal policy is able to use a hybrid setting for capacity that depends on both  $W_n$  and an estimate of expected demand over the next period, for example  $\mu_n = \frac{W_n}{T_c} + \alpha\mu_{MG1}^*$  with  $\alpha$  smaller than 1.

### 4.2. Fluid Model Analysis

We next investigate the more formal fluid model approximation (e.g., see [3]), again considering the stochastic dynamic program over a single workload period  $m$ . We assume the demand  $D_n$  in each control period  $n$  is a deterministic quantity, arriving at constant speed  $D_n/T_c$ . Let  $\mu_n^*(w)$  denote the optimal policy (i.e., the optimal capacity level to set) for control period  $n$  when  $w$  units of workload is observed at the beginning of the period. For the fluid model,  $\mu_n^*(w) \in [0, \bar{\mu}_n(w)]$  where  $\bar{\mu}_n(w) := (w + D_n)/T_c$ ; that is, the server speed that is needed to process the sum of the starting workload and the newly arriving workload.

Define  $I_1 := (-\infty, -(N - 1)p_D]$ ,  $I_n := (-(N - n + 1)p_D, -(N - n)p_D]$  for  $n = 2, \dots, N - 1$ , and  $I_N := (-p_D, \infty)$ . The following theorem characterizes the structure of the optimal policy in the fluid model approximation.

**THEOREM 4.1:** *If  $r - c \in I_1$ , then  $\mu_k^*(w) = 0$  for all  $k$ ; If  $r - c \in I_N$ , then  $\mu_k^*(w) = \bar{\mu}_k(w)$  for all  $k$ ; If  $r - c \in I_n$  for some  $n = 2, \dots, N - 1$ , then  $\mu_k^*(w) = \bar{\mu}_k(w)$  for all  $k \leq n$ , and  $\mu_k^*(w) = 0$  for all  $k \geq n + 1$ .*

PROOF: To simplify the presentation, we provide the details of the proof for a two control-period problem. The proof for problems with more control periods follows similarly. The one-period-to-go optimal value function reads

$$J_N(w) = -p_D(w + D_N) + \max_{\mu \in [0, \bar{\mu}_N(w)]} \mu[(r + p_D - c)T_c], \tag{18}$$

and hence

$$\mu_N^*(w) = \begin{cases} \bar{\mu}_N(w), & \text{if } r - c > -p_D, \\ 0, & \text{if } r - c \leq -p_D. \end{cases} \tag{19}$$

Upon substituting (19) into (18), we have

$$J_N(w) = \begin{cases} (r - c)(w + D_N), & \text{if } r - c > -p_D, \\ -p_D(w + D_N), & \text{if } r - c \leq -p_D. \end{cases} \tag{20}$$

Now consider the two-period-to-go optimal value function, which after some term manipulation yields

$$J_{N-1}(w) = z(D_{N-1} + D_N + w) - p_D(w + D_{N-1}) + \max_{\mu \in [0, \bar{\mu}_{N-1}(w)]} [\mu((r + p_D - c)T_c - zT_c)], \tag{21}$$

where  $z := \max\{r - c, -p_D\}$ . The optimal solution then can be broken down into three cases.

1. If  $z = r - c$  or  $r - c > -p_D$ , we have that  $\mu_{N-1}^*(w) = \bar{\mu}_{N-1}(w)$  and

$$J_{N-1}(w) = (r - c)[w + D_{N-1} + D_N]. \tag{22}$$

2. If  $z = -p_D$  (or equivalently  $r - c \leq -p_D$ ) and in the meantime  $r - c > -2p_D$ , we also have that  $\mu_{N-1}^*(w) = \bar{\mu}_{N-1}(w)$  and

$$J_{N-1}(w) = (r - c)[w + D_{N-1}] - p_D D_N. \tag{23}$$

3. If  $r - c \leq -2p_D$ , we have that  $\mu_{N-1}^*(w) = 0$  and

$$J_{N-1}(w) = -2p_D[w + D_{N-1}] - p_D D_N. \tag{24}$$

■

In words, Theorem 4.1 states that depending into which of the  $I_n$  intervals the quantity  $r - c$  falls, the optimal policy is to always deplete the workload to zero until a certain control period and then set the capacity level to zero thereafter. The key quantity  $r - c$  can be naturally viewed as a profitability index of the system.

The structure of the optimal policy in the fluid model approximation suggests that a good policy for the original stochastic dynamic program is to set the capacity level for control period  $n$  such that the expected remaining workload at the end of the period is equal to some target level  $w_n^*$ , with  $w_1^* \geq w_2^* \geq w_3^* \geq \dots \geq w_N^*$ . Note that for the fluid model approximation some of the  $w_n^*$  are simply zero, according to Theorem 4.1.

The above insights from the fluid model approximation inspires the following control policy:

1. At the beginning of control period 1, observe the current workload  $w$  and solve

$$\min_{(\mu_1, \dots, \mu_n)} (c + p_I)T_c \sum_{n=1}^N \mu_n + p_D \sum_{n=1}^N w_n + (r + p_I + p_D)w_{n+1} \tag{25}$$

$$\text{s.t. } w_n = w_{n-1} - \mu_n T_c, \quad n = 1, 2, \dots, N \tag{26}$$

$$w_0 = w \tag{27}$$

2. Set the server speed to the optimal  $\mu_1$  that results from solving the above linear program.
3. At the beginning of control period 2, repeat the above two steps by treating it as the beginning of an  $(N - 1)$ -period problem.
4. Continue the above three steps until the end.

Note that the linear program objective function (25) is motivated by the equivalent expression for the  $N$ -period objective function (12), in which the non-controllable terms are dropped from (25). Moreover, the recursion (26) is simply a linear approximation that relates the workload to the capacity level. By resolving the linear program that aims to achieve a desirable target workload path at each control period, the above policy exercises good control on the evolution of the stochastic system.

### 4.3. Diffusion Approximation

Now, since the system state at each control period over a single workload period can be viewed as the time- $T_c$  workload of a  $G/G/1$  queue, we consider the use of the diffusion approximation result for the workload process of a  $G/G/1$  queue (see [4]) to derive an explicit approximating expression for the transition probability of the stochastic dynamic program. Let  $1/\lambda$  denote the mean interarrival time,  $1/\mu$  the mean service time, and  $c_a^2$  ( $c_s^2$ ) the squared coefficient of variation for the interarrival (service) times. Define  $\rho := \lambda/\mu$ . Then the following approximation holds

$$\mathbb{P}[W_{n+1} > y | W_n = w, \mu_n = \mu] \approx \mathbb{P}[\tilde{W}(T_c(1 - \rho)^2) > y(1 - \rho)],$$

where  $\tilde{W}(\cdot)$  is a reflected Brownian motion with initial level  $w(1 - \rho)$ , drift  $-1$  and variance  $(c_a^2 + c_s^2)\mu^{-1}$ . Furthermore, using the time dependent distribution of the reflected Brownian motion (see [8]), we have

$$\begin{aligned} \mathbb{P}[W_{n+1} > y | W_n = w, \mu_n = \mu] \approx & \Phi \left( \frac{-(y - w)(1 - \rho) - T_c(1 - \rho)^2}{\sqrt{(c_a^2 + c_s^2)\mu^{-1} \cdot (1 - \rho)\sqrt{T_c}}} \right) \\ & - e^{-\frac{2y(1-\rho)}{(c_a^2 + c_s^2)\mu^{-1}}} \Phi \left( \frac{-(y + w)(1 - \rho) + T_c(1 - \rho)^2}{\sqrt{(c_a^2 + c_s^2)\mu^{-1} \cdot (1 - \rho)\sqrt{T_c}}} \right), \tag{28} \end{aligned}$$

which serves as the action-dependent transition probability function of the continuous-state-space, continuous-action-space Markov decision process at hand. Specifically, due to the

value function expression in Section 3 and relation (9), the diffusion-approximated dynamic programming recursion boils down to

$$\mu_n^*(w) = \arg \max_{\mu} \left\{ - (r + p_I + p_D) \mathbb{E}_{\mu}[W_{n+1}|W_n = w, \mu_n = \mu] - (c + p_I) T_c \mu + \mathbb{E}_{\mu}[J_{n+1}(W_{n+1})|W_n = w, \mu_n = \mu] \right\}, \quad (29)$$

subject to (28). This can be solved by first discretizing the state and/or action space and then using the value iteration method.

### 4.4. Computational Experiments

We end this section by presenting the results of computational experiments that quantitatively investigate some of the results and properties derived above. To ground the discussion, we make a few basic assumptions. First, we assume  $r > c$  since this setting represents a realistic situation where the system is capable of earning a positive profit. Second, motivated by (13) and the discussion preceding it, we fix  $c = 1$  and  $p_I = 1$  and study the performance of control policies as a function of changes in  $p_D$ ; roughly speaking, the performance of the system is a function of the ratio  $(c + p_I)/p_D$  and therefore only  $p_D$  needs to be varied. We set  $T_c = 1$  and the number of control periods to be  $N = 100$ . Further we consider a system with low arrival rate, that is,  $\lambda = \lambda_{\ell} = 1$ , and another with a high arrival rate, that is,  $\lambda = \lambda_h = 5$ . As mentioned in Section 4.1, we will solve (15) to obtain  $\mu_{MG1}^*$  by assuming that service demands are i.i.d.  $\text{Exp}(\theta)$  whereupon  $\mathbb{E}[S] = 1/\theta$ ,  $\mathbb{E}[S^2] = 2/\theta^2$ . Without loss of generality, we set  $\theta = 1$  and therefore  $\mathbb{E}[S] = 1$  and  $\mathbb{E}[S^2] = 2$ . As in Section 4.1 we set the discount factor  $\beta = 0$  for simplicity.

Before proceeding to describe our computational results, we briefly summarize some of the key insights derived earlier in this section. At the end of Section 4.1, we consider a heuristic control policy that depends on a linear combination of  $W_n$ , the backlog workload at the beginning of a period, and some estimate of the expected demand that will be realized in the period, such as  $\mu_{MG1}$ . Our fluid analysis in Section 4.2 confirms this insight. Specifically, when  $r > c$ , Theorem 4.1 recommends that the system always drain fluid at a rate

$$\mu_n^{\text{fluid}} = \frac{W_n + \mathbb{E}[D_n]}{T_c} = \frac{W_n}{T_c} + \lambda \mathbb{E}[S],$$

given that  $r - c$  always belongs to the interval  $I_N$  by assumption. In essence, our heuristic control policy is an affine decision rule that only depends on the current workload (as opposed to the whole sequence of past workloads) and an affine term that can be chosen to be proportional to some estimate of anticipated demand such as  $\lambda \mathbb{E}[S]$  or  $\mu_{MG1}^*$ .

For our first set of computational results, we focus on the low arrival rate case where  $\lambda = \lambda_{\ell} = 1$ . Figure 1 depicts the expected profit curves as a function of  $\alpha$  for three different values of  $p_D = 1, 5, 10$ . As expected, the expected profit reduces with an increased backlog penalty; further the value of  $\alpha$  that maximizes profit increases with  $p_D$ . This latter trend is depicted in Figure 2 where we consider the best value of  $\alpha$  as a function of  $p_D = 1, \dots, 10$ . Observe that  $\alpha^* = 0$  as long as  $p_D \leq c + p_I$ , which corresponds to our intuition above that in this case the capacity maintenance and idleness costs dominate the backlog penalty and hence a purely workload-based policy renders optimal expected profits within this family of linear decision rules. Note also that  $\alpha$  appears to grow linearly in  $p_D$  once  $p_D > c + p_I$ . It is worth pointing out that  $\alpha$  allows one to balance between choosing a server with high average utilization or with low average workload. Hence, we observe in Figure 3 that while

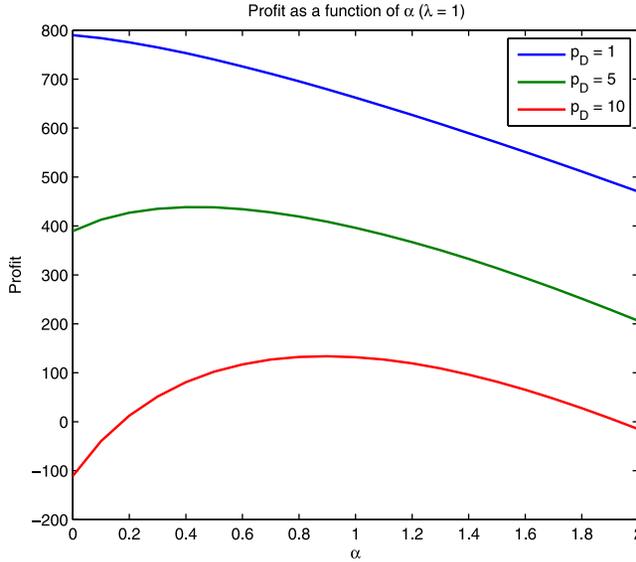


FIGURE 1. Expected profit as a function of  $\alpha$  for  $\lambda = \lambda_l = 1$ .

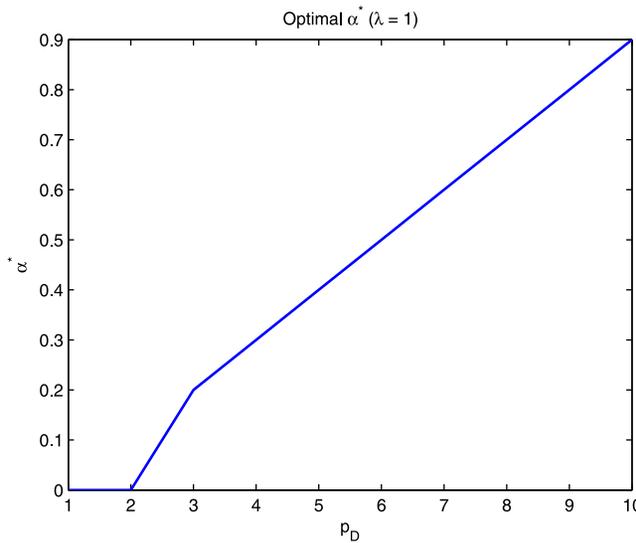


FIGURE 2. Optimal  $\alpha$  as a function of  $p_D$  for  $\lambda = \lambda_l = 1$ .

increasing  $\alpha$  leads to a system with decreasing average workload and hence a lower backlog penalty, it also leads to the server idling more and thus incurring a higher idling penalty.

Next, we consider the high arrival rate case where  $\lambda = \lambda_h = 5$ . We observe that a non-zero  $\alpha$  always appears to provide the highest expected profit as illustrated in Figures 4 and 5. The dependence between  $\alpha^*$  and  $p_D$ , however, still appears to follow a fairly linear trend as shown in Figure 5.

Finally, we consider a case where the demand arrival rate alternates between  $\lambda = \lambda_l = 1$  in every odd numbered control period and  $\lambda = \lambda_h = 5$  in every even numbered control period. We further suppose that the system operator has some indication of this highly

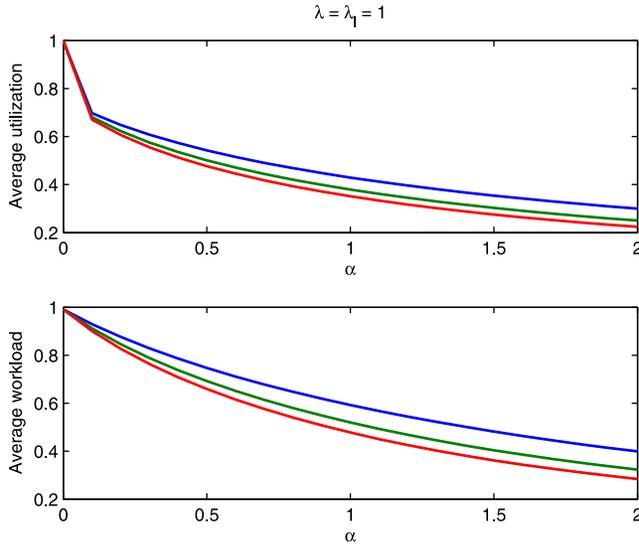


FIGURE 3. Average utilization and average workload as a function of  $\alpha$  for  $\lambda = \lambda_l = 1$ .

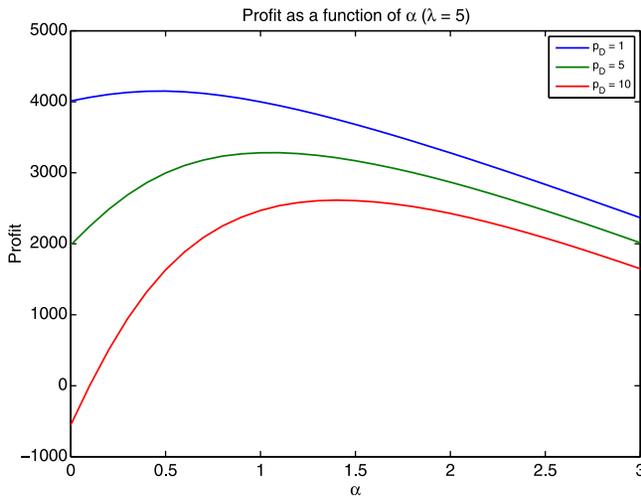


FIGURE 4. Expected profit as a function of  $\alpha$  for  $\lambda = \lambda_h = 5$ .

variable workload pattern so that a fluid-model control policy can be implemented; namely, the capacity settings

$$\mu_n = \frac{W_n}{T_c} + \alpha\lambda_l$$

are chosen for  $n$  odd, while the capacity settings

$$\mu_n = \frac{W_n}{T_c} + \alpha\lambda_h$$

are chosen for  $n$  even. We compare this fluid-model control policy with a control policy setting that depends purely on an estimate of expected demand, namely

$$\mu_n = \lambda_l, \quad \text{when } n \text{ odd}; \quad \mu_n = \lambda_h, \quad \text{when } n \text{ even.}$$

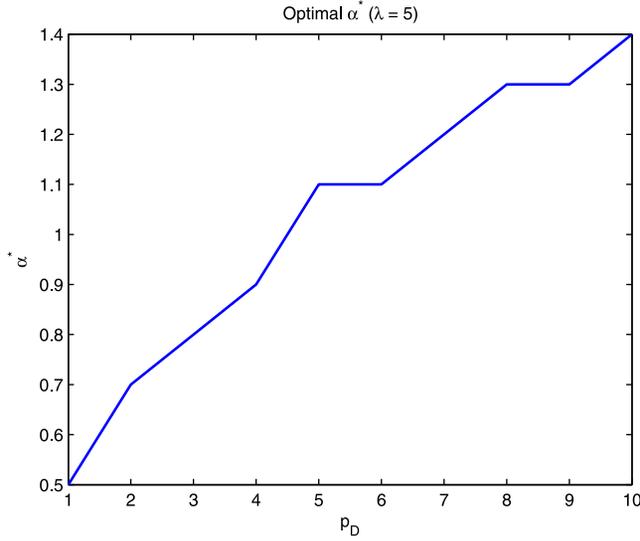


FIGURE 5. Optimal  $\alpha$  as a function of  $p_D$  for  $\lambda = \lambda_h = 5$ .

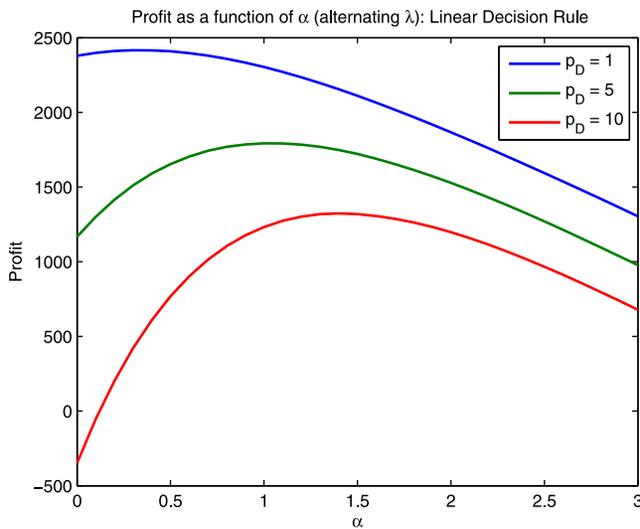


FIGURE 6. Expected profit as a function of  $\alpha$  for alternating  $\lambda$  using a linear decision rule.

Note that both of these control approaches are similar in spirit to the off-line policy discussed in Section 4.1 since we assume some knowledge of future demand.

Through this last computational experiment, we seek to investigate the importance of employing a linear decision rule wherein the capacity depends on the backlogged demand from the previous period. With such a highly variable load, it is almost futile to try and calculate the required capacity based solely on an estimate of expected demand. Figures 6 and 7 demonstrate the stark difference between the two policies: Whereas the fluid-model control policy is able to maintain expected profit within a reasonable range over all values of  $\alpha$ , expected profit under the demand-based control policy exhibits a large sensitivity to the correct choice of  $\alpha$ . Moreover, the expected profit for this latter policy is also lower than that accrued when using the linear decision rule.

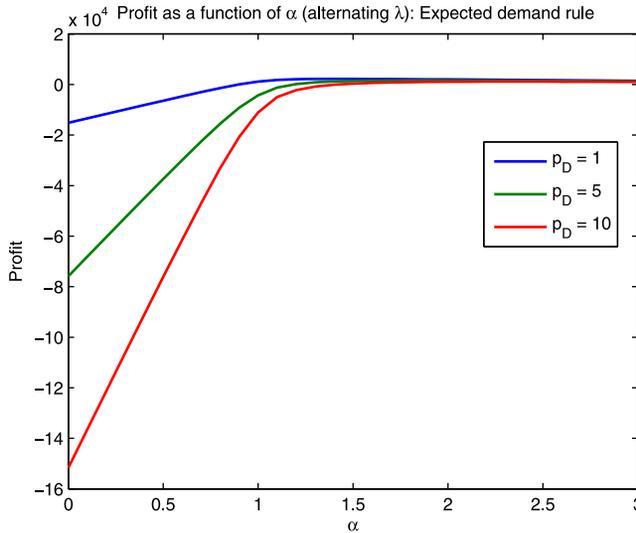


FIGURE 7. Optimal  $\alpha$  as a function of  $p_D$  for alternating  $\lambda$  with expected demand policy.

5. GENERALIZED MODELS

To elucidate the exposition in earlier sections, we chose to postpone the inclusion of certain problem features, mentioned in the introduction, from our original general stochastic model, optimization formulation and optimal solution. We now discuss how these problem features can be addressed as part of and incorporated in our analysis and results of earlier sections.

One of the features that was not included in our earlier analysis concerns the failure and/or departure of the servers. In this case, for each control period  $n$ , if the service-rate capacity is set to a certain level  $\mu_n$  at the beginning of the control period, then the realized service-rate capacity will vary over time within the interval  $[t_{n-1}^c, t_n^c)$  according to  $\mu_n(t)$ . To this end, we consider both an additive model and a multiplicative model for  $\mu_n(t)$  as follows:

- additive model:  $\mu_n(t) = \mu_n - \eta(t - (n - 1)T_c)$ , for  $t \in [t_{n-1}^c, t_n^c)$ ;
- multiplicative model:  $\mu_n(t) = \mu_n e^{-\gamma(t - (n-1)T_c)}$ , for  $t \in [t_{n-1}^c, t_n^c)$ ;

where  $\eta > 0$  and  $\gamma > 0$  are two fixed parameters. It is then readily verified that the structural properties we derived in Section 3 continue to hold and that the computational results to obtain both the general and surrogate problem solutions are easily adapted to that of the corresponding  $GI_m/GI_m/1_n$  queues with time-varying service rates.

On the other hand, to faithfully model the detailed dynamics of server failures and departures, it is essential to consider a multi-server model. Doing so also allows us to incorporate the phenomenon that a server can operate at different speeds together with addressing the corresponding cost/benefit trade-offs in such a queueing system. Suppose that, for each control period  $n$ , we need to determine a vector  $\mathbf{X}_n$  that consists of the number of servers operating at each of the speed levels  $s = 1, 2, \dots, S_L$ . The objective function of our stochastic dynamic programs can then be revised accordingly. As a specific example, we can replace the objective (13) by

$$\min_{\mathbf{X}_n} T_c \sum_{n=1}^N \sum_{s=1}^{S_L} (p_I + c_s) \mathbf{X}_{n,s} + p_D \sum_{n=1}^N W_n,$$

where the cost factor  $c$  is replaced by a vector  $\{c_s\}$  that reflects the different costs incurred for operating servers at different speed levels. Note that  $W_n$ ,  $n = 1, 2, \dots, N$ , will follow the much more complex dynamics of a multi-server queueing system with different speed levels for each server and where each server has an independent clock to follow for its failure and departure from the system. However, even in such very complex instances of our general stochastic dynamic programs, the corresponding optimal solutions continue to enjoy the general structural properties we derived earlier in the paper. We note, however, that detailed fluid and diffusion approximations of such systems require sophisticated studies of multi-server queues with time-varying service rates, which is the subject of future work.

## 6. CONCLUSIONS

In this paper, we considered a general class of stochastic dynamic control problems for a queueing system with time-varying Markov-modulated arrival and service processes (indexed by  $m$ ) and time-varying control of the service rates of the system (indexed by  $n$ ). Briefly reiterating the contributions of our study, we formulated a stochastic dynamic control problem for maximizing the expected total discounted net-benefit under very general assumptions on the customer arrivals and service distributions. For this general formulation, we established fundamentally important structural properties of the optimal solution, including the second-order monotonicity of some key performance metrics as well as results on the concavity of the objective function for our net-benefit maximization problem. These results reduced the calculation of the value function to a sequence of convex programming problems, which in turn allows the computation of the value functions for various real-world problem sizes of practical interest. Although larger-sized problems still face a prohibitively difficult task of computing the value function due to the inherent *curse of dimensionality*, the structural properties established in this paper (e.g., concavity) make it possible to employ various additional techniques, such as those from approximate dynamic programming, to compute approximate optimal solutions; refer to, for example, [11]. Furthermore, making use of the derived structural properties, we also demonstrated that there is a critical-value (threshold) type of approximate optimal control policy that can be fairly easily computed together with certain performance guarantees.

We also investigated different options for approximations within a single workload period  $m$ . The first class of approximations consists of several parameterized control policies that are commonly used in real applications, along with some observations of system behaviors under such heuristics. Both theoretical and computational results have been obtained on the performance of the system under these control policies. Furthermore, we investigate both fluid and diffusion approximations that in general have been successfully applied in the analysis and control of complex queueing systems and networks. The optimal solution for the fluid model is obtained, which suggests a certain targeted level for the expected remaining workload; a heuristic control policy based on these observations is evaluated in our computational experiments. A diffusion approximation provides a further refinement of these models and bridges the differences between the heuristics and the exact solutions. Various computational experiments were conducted, demonstrating and quantifying our results and related observations.

### References

1. Bhadra, S., Lu, Y., & Squillante, M.S. (2007). Optimal capacity planning in stochastic loss networks with time-varying workloads. *SIGMETRICS Performance Evaluation Review* 35(1): 227–238.

2. Chen, F. & Song, J.-S. (2001). Optimal policies for multiechelon inventory problems with markov-modulated demand. *Operations Research* 49(2): 226–234.
3. Chen, H. & Yao, D. (1992). A fluid model for systems with random disruptions. *Operations Research* 40(3-supplement-2): S239–S247.
4. Chen, H. & Yao D. (2001). *Fundamentals of queueing networks: performance, asymptotics, and optimization*. New York, USA: Springer.
5. Ciocan, D.F. & Farias, V.F. (2012). Model predictive control for dynamic resource allocation. *Mathematics of Operations Research* 37(3): 501–525.
6. Gao, X., Lu, Y., Sharma, M., Squillante, M.S., & Bosman, J.W. (2013). Stochastic optimal control for a general class of dynamic resource allocation problems. *SIGMETRICS Performance Evaluation Review* 41(2): 3–14.
7. George, J. & Harrison, J. (2001). Dynamic control of a queue with adjustable service rate. *Operations Research* 49(5): 720–731.
8. Harrison, J.M. (2013). *Brownian models of performance and control*. New York, USA: Cambridge University Press.
9. Lin, M., Wierman, A., Andrew, L.L.H. & Thereska, E. (2013). Dynamic right-sizing for power-proportional data centers. *IEEE/ACM Transactions on Networking* 21(5): 1378–1391.
10. Massey, W.A. & Mandelbaum, A. (1995). Strong approximations for time dependent queues. *Mathematics of Operations Research* 20(1): 33–64.
11. Powell, W.B. (2007). *Approximate dynamic programming: solving the curses of dimensionality (Wiley Series in Probability and Statistics)*. Hoboken, NJ, USA: Wiley-Interscience.
12. Sobel, M.J. (1974). Optimal operation of queues. In *Mathematical Methods in Queueing Theory*, (A.B. Clarke, Ed.) vol. 98 of *Lecture Notes in Economics and Mathematical Systems*, Berlin, Heidelberg: Springer, pp. 231–261.
13. Song, J.-S. & Zipkin, P. (1993). Inventory control in a fluctuating demand environment. *Operations Research* 41(2): 351–370.
14. Vishwanath, K.V. & Nagappan, N. (2010). Characterizing cloud computing hardware reliability. In *Proceedings of ACM Symposium on Cloud Computing*, pp. 193–204.
15. Whitt, W. (1991). The pointwise stationary approximation for  $M_t/M_t/s$  queues is asymptotically correct as the rate increases. *Management Science* 37(2): 307–314.
16. Wolff, R.W. (1989). *Stochastic modeling and the theory of queues*. Englewood Cliffs, NJ, USA: Prentice-Hall.