Dynamic Service Rate Control for a Single Server Queue with Markov Modulated Arrivals

Ravi Kumar¹, Mark E. Lewis² and Huseyin Topaloglu³

School of Operations Research and Information Engineering
Cornell University
Ithaca, NY 14853

September 13, 2011

¹rk454@cornell.edu
²mark.lewis@cornell.edu
³ht88@cornell.edu
Abstract

We consider the problem of service rate control of a single server queueing system when the arrival process is governed by a finite-state Markov-modulated Poisson process. There are two main technical contributions. First, in keeping with intuition, we show that the optimal service rate is non-decreasing in the number of customers in the system. That is, higher congestion rates warrant higher service rates. On the contrary, however, we show that the optimal service rate is not necessarily monotone in the current arrival rate. If the modulating process satisfies a stochastic monotonicity property we show that the monotonicity is recovered. Together these results imply that we have reasonable conditions for the optimal control to follow a monotone switching curve in the current state of the system.

Our numerical study also has two components. We examine several heuristics and show where those heuristics can be reasonable substitutes for the optimal control. None of the heuristics perform well in all the regimes we consider. Secondly, we discuss when the Markov-modulated Poisson process with service rate control can act as a heuristic itself. In particular, we show that it can approximate the optimal control of a system with a periodic non-homogeneous Poisson arrival process. Thus, not only is the current model of interest in the control of Internet or mobile networks with bursty traffic, but it is also useful in providing a tractable alternative for the control of service centers with non-stationary arrival rates.
1 Introduction

Queueing models with non-stationary arrivals are widely used in the telecommunications industries to study congestion problems related to Internet and mobile networks. One method for capturing the non-stationary behavior of the arrival process is via Markov-modulated Poisson processes (MMPPs). MMPPs are those in which the rate of the underlying Poisson process is influenced by another exogenous Markov process commonly known as the phase process. An alternative model to capture the non-stationary nature of arrivals is the non-homogenous Poisson process (NHPP). However, in this case control policies depend not only on the congestion level but also on the current time. This makes analysis and computation of an optimal control difficult and most often intractable. Markov-modulated queueing systems provide a simpler approach to approximate such non-stationary properties that is analytically tractable via the theory of Markov decision processes. Moreover, the MMPP arrival model can be used to capture bursty and correlated characteristics of data transmission. Whereas the NHPP model allows for random arrivals based on changes in a known rate function, the MMPP allows for random and unknown state changes; an added level of flexibility for modeling.

In view of the rapid proliferation of mobile and Internet devices relying on a growing infrastructure, power aware transmission policies have become increasingly important. Given this context, and the aforementioned usefulness of MMPPs to model nonstationary behavior, we investigate the problem of service rate control for an infinite buffer single server queue with Markov-modulated Poisson arrivals and exponentially distributed service requirements. Given that the state of the phase modulating process and state of the queue are completely known, the goal of the controller is to decide service rates that minimize either the expected discounted cost or the long-run average cost rate over an infinite horizon.

It is well known that in the case of service rate control of queues with stationary arrival and service processes, the optimal decisions are monotone in the queue length [17]. For the current study, suppose that the arrival intensities change according to a general continuous time Markov chain (CTMC) defined on the state space \{1, 2, \ldots, L\} and the arrival rates of the MMPP are ordered such that \(\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L\). One may ask two fundamental questions

- Is the optimal control policy monotone in the queue length for each state of the phase modulating process?
- Is the optimal control policy monotone in the state of the phase modulating process for each queue length value?

The main technical contributions of this paper are to show that for both the discounted cost criterion and the average cost criterion, the first question can be answered in the affirmative. However, perhaps surprisingly, that despite the ordering of the arrival rates, the answer to the

---

1 In Kendall’s notation, such systems are classified as MMPP/M/1
second question is not necessarily yes. In fact, we show that the structure of the optimal policy depends on the nature of the phase modulating process and provide sufficient conditions on the phase process under which such monoticity does hold.

We include a numerical study with two goals in mind. First, we examine when it is important to explicitly capture the nonstationary behavior of an arrival process via MMPP as opposed to using some natural heuristic like assuming the system has stationary arrivals. Second since we mentioned the alternative of using an NHPP to capture the nonstationarity in the arrival process we explore the possibility of computing an approximate control policy for a system with an NHPP with a periodic rate function using a “suitable” MMPP. We find that our preliminary results in this direction are encouraging. This is a significant diversion from previous studies since the focus here is on computing an optimal control and not solely on evaluating performance measures.

Most of the previous research related to Markov-modulated queueing systems deals with performance characteristics for systems without control (i.e., with constant service rate). Excellent overviews of this line of work can be found in the survey paper by Prabhu [20] or more recently in Gupta et al. [13]. A hierarchical scheme based on MMPP was proposed by Muscariello et al. [19] to model the data generated by Internet users. Heffess et al.[14] used MMPP to approximate a statistical multiplexer whose inputs consist of superposition of packetized voice sources and data. A two state MMPP model was proposed by Shah-Heydari [22] to model the aggregate Asynchronous Transfer Mode (ATM) traffic. For more general scenarios, Frost [10] proposed a scheme to approximate a simple NHPP using an MMPP by suitably quantizing the rate function of the NHPP into a finite number of rates. Each rate corresponds to a state in the Markov modulating process and the parameters of the MMPP model can be estimated using empirical data.

There is also a rich body of literature on the subject of monotone optimal policies for the control of a single server queue in a setting similar to the one considered here but with stationary arrivals. See for example the classic work of Crabill [6], Lippman [17] and Stidham and Weber [23]. In the context of telecommunications systems, the existing literature addresses a more closely related problem of service rate control of queues when the job service requirements are influenced by an exogenous stochastic process². Such models arise frequently in point-to-point wireless data transmission where the induced transmission rates are affected by the time varying properties of the transmission medium. Berry [3] considers a very general model for this problem under a discrete-time Markovian setting. In this work, packet arrivals follow a batch Markov process and the state of the transmission channel varies according to a secondary discrete-time Markov chain. The data buffer and transmitter are modeled using a single server queue with finite capacity. The goal of the transmitter is to minimize the average cost rate (power consumption) over an infinite time horizon subject to a constraint on packet delay (another

²Similar to the present work, the policy for this type of model depends on both the queue length and the state of the exogenous process.
case in which the constraint is on the probability of buffer overflow is also discussed in this work). He proves several results related to the monotonicity of the optimal policy. Motivated by mobile networks, Ata and Zachariadis [2] address the problem of finding optimal service rates for multiple users that are being served by a central controller. Data gets transmitted through a time varying channel that is being modulated by a two-state continuous time Markov chain. Packet data for each user arrives based on a Poisson process and gets stored in a finite capacity queue before getting transmitted. The objective is to maximize some measure of overall quality of service subject to a constraint on the long-run average power consumption. The authors show that the optimal service rates for each user depend only on his/her own queue length and the state of the transmission channel. They also present a method to explicitly characterize the optimal policy for each user. To the best of our knowledge, none of the aforementioned work considers the case of a multi-state Markov-modulated arrival process with service rate control as is discussed in this paper.

We have organized this paper in the following manner. In Section 2 we give a detailed description of the model, the associated assumptions formulate the problem as a Markov decision process. In the average cost case we provide stability conditions that guarantee convergence of the discounted cost value function to the relative value function. In Section 3 we give structural results related to the optimal policy in each case. In Section 4, we present a detailed numerical study comparing the performance of the optimal policy with heuristic policies and present an exploratory study related to the computation of a heuristic policy for non-homogeneous Poisson arrivals using the optimal policy for a MMPP/M/1 queue. We conclude the paper in Section 5 by summarizing the main results.

## 2 Model Formulation

We consider a single server queue with infinite buffer capacity and job arrivals that follow a Markov-modulated Poisson process (MMPP). Each arriving job has an exponentially distributed service requirement with mean 1. The phase transition process for arrivals is an ergodic, finite state continuous time Markov chain with generator matrix $Q$. Let the state space for this process be denoted by $S := \{1, 2, \ldots, L\}$. When the phase transition process is in phase $s \in S$, jobs arrive to the queue according to a Poisson process with rate $\lambda_s$. Without loss of generality we assume that the states are ordered such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$. Let the number of jobs in the system (buffer state) be denoted by $n \in \mathbb{Z}^+$, where $\mathbb{Z}^+$ is the set of non-negative integers. The service rate can be changed at the times of arrivals, departures or phase transitions. Together, the union of these event times and (in a moment) the added dummy transitions due to uniformization comprise the set of decision epochs. Based on the queue length, $n$, and the state of the arrival process, $s$, the controller selects a service rate $\mu_{n,s}$ from the compact set $A = [0, \bar{u}]$, $\bar{u} < \infty$. When a service rate $\mu \in A$ is chosen, the system incurs a cost at the rate of $c(\mu)$ per unit time. The cost rate function, $c(\cdot)$, is defined on $A$ and
is assumed to be strictly convex, continuously differentiable, strictly increasing and (without loss of generality) such that $c(0) = 0$. Furthermore, a holding or congestion cost is incurred at rate $h(n)$ per unit time when the buffer state is $n$. The holding cost function $h(n)$ is assumed to be convex, non-decreasing in $n$ and such that $h(0) = 0$ and $h(1) > 0$. In the average-cost case in order to guarantee stability we assume $h(\cdot)$ is linear.

Let $\Pi$ be the set of non-anticipating policies. A stationary control policy, $\pi \in \Pi$, is defined as

$$\pi = \{\mu_{n,s} | n \in \mathbb{N}, s \in \mathcal{S}\}$$

where $\mu_{n,s}$ is the service rate to be selected when the state of the system is $(n,s)$. The controller remains idle when the queue is empty i.e, for any policy, $\mu_{0,s} \equiv 0$.

Thus, given a policy $\pi$, the overall process, $X(t)$, evolves as a two dimensional continuous time Markov chain on the state space $\mathcal{X} = \{(n, s) | n \in \mathbb{Z}^+, s \in \mathcal{S}\}$. Our objective is to find a control policy that minimizes the discounted expected cost or average expected cost per unit time over an infinite time horizon.

### 2.1 The Discounted Expected Cost Formulation

For $x = (n, s)$ and service rate $\mu \in \mathcal{A}$, let

$$f(x, \mu) := c(\mu) + h(n).$$

Let $\{(X^\pi(t), D^\pi(t)), t \geq 0\}$ be the stochastic process representing the evolution of states and decisions under an admissible policy $\pi$. Given the initial state $x$ and discount factor $\alpha > 0$, the $\alpha$–discounted expected cost until time $t$ under policy $\pi$ is given by

$$v^\pi_{t, \alpha}(x) := \mathbb{E}_x^\pi \left[ \int_0^t e^{-\alpha u} f(X(u), D(u)) du \right].$$

The total discounted expected cost of a policy $\pi$ given that the initial state of the system is $x$, is

$$v^\pi(x) := \lim_{t \to \infty} v^\pi_{t, \alpha}(x).$$

The optimal total discounted expected cost is

$$v^*_\alpha(x) := \inf_{\pi \in \Pi} v^\pi_{t, \alpha}(x).$$

A policy, $\pi^*$, is total discounted expected cost optimal if $v^*_\alpha(x) = v^\pi_{t, \alpha}(x)$ for all $x \in \mathcal{X}$.

We apply uniformization in the spirit of Lippman [17] and consider the discrete time equivalent of the continuous time Markov chain described above. The uniformization rate is chosen to be $\nu := \lambda_L + \bar{\eta} + \bar{u}$, where $\bar{\eta} \geq \max\{-Q_{ss} | s \in \mathcal{S}\}$ is any finite rate larger than the maximum of the holding time parameters for the phase transition process.
Using standard arguments of Markov decision theory [4], the discrete-time finite horizon optimality equations (FHOE) for the system can be written (for each \( s \in S \)):

\[
v_{\alpha,k+1}(0,s) = \frac{1}{\alpha + \nu} \left[ h(0) + \lambda s v_{\alpha,k}(1,s) + \sum_{s' = 1}^{L} Q_{ss'} v_{\alpha,k}(0,s') + (\nu - \lambda_s) v_{\alpha,k}(0,s) \right]
\]

\[
v_{\alpha,k+1}(n,s) = \frac{1}{\alpha + \nu} \min_{\mu \in \mathcal{A}} \left\{ c(\mu) + h(n) + \mu v_{\alpha,k}(n-1,s) + \lambda_s v_{\alpha,k}(n+1,s) + \sum_{s' = 1}^{L} Q_{ss'} v_{\alpha,k}(n,s') + (\nu - \lambda_s - \mu) v_{\alpha,k}(n,s) \right\} \text{ for } n \geq 1,
\]

where \( v_{\alpha,0} \) is assumed to be zero for each state. Note that the cost function has compact level sets. That is, \( \{(n,s,\mu) | f((n,s,\mu) \leq \beta\} \) is compact for all \( \beta \in \mathbb{R} \). Since the state space is discrete, we may apply Proposition 3.1 of [9] to get \( v_{\alpha,k} \uparrow v_{\alpha} \). Moreover, \( v_{\alpha} \) satisfies (2.2) with \( v_{\alpha} \) replacing \( v_{\alpha,k} \) on the right hand side and \( v_{\alpha,k+1} \) on the left hand side. The resulting set of equations are called the discounted cost optimality equations (DCOE) and are stated next for completeness (for each \( s \in S \)).

\[
v_{\alpha}(0,s) = \frac{1}{\alpha + \nu} \left[ h(0) + \lambda s v_{\alpha}(1,s) + \sum_{s' = 1}^{L} Q_{ss'} v_{\alpha}(0,s') + (\nu - \lambda_s) v_{\alpha}(0,s) \right]
\]

\[
v_{\alpha}(n,s) = \frac{1}{\alpha + \nu} \min_{\mu \in \mathcal{A}} \left\{ c(\mu) + h(n) + \mu v_{\alpha}(n-1,s) + \lambda_s v_{\alpha}(n+1,s) + \sum_{s' = 1}^{L} Q_{ss'} v_{\alpha}(n,s') + (\nu - \lambda_s - \mu) v_{\alpha}(n,s) \right\} \text{ for } n \geq 1.
\]

2.2 The Long-Run Average Cost Formulation

In this section we provide conditions under which an average cost optimal policy exists and may be computed as a limit of discounted cost optimal policies. The long-run average or gain of a policy \( \pi \) given that the initial state of the system is \( x \), is

\[
g^{\pi}(x) := \limsup_{t \to \infty} \frac{v_{t,0}^{\pi}(x)}{t},
\]

where \( v_{t,0} \) is as defined in (2.1). The optimal expected average cost \( g^{*}(x) \) is

\[
g^{*}(x) := \inf_{\pi \in \Pi} g^{\pi}(x),
\]
and \( \pi^* \) is an average cost optimal policy if \( g^\pi(x) = g^*(x) \) for all \( x \in X \). After uniformization the average cost optimality inequalities (ACOI) (cf. [21]) are,

\[
w(n, s) \geq \nu \left[ -g + c(x) + h(n) + \lambda_s w(n + 1, s) + xw(n - 1^+, s) + \sum_{s' = 1}^L Q_{ss'} w(n, s') + (\nu - \lambda_s - x)w(n, s) \right] \quad \text{for } n \geq 0, s \in S.
\]

When the solution, \((w, g)\) to the ACOI exists, \( w \) is called a relative value function and \( g^* (x) = g \) is the optimal long-run expected average cost for any initial state \( x \).

A solution to the ACOI (2.4) exists under a necessary and sufficient stability condition which is provided in (2.5) below. This condition requires that the maximum available service rate is higher than the long-run average arrival rate and coincides with the one derived by Yechiali [24] for the stability of queue with Markov-modulated arrivals. However, since Yechiali used the balance equations to show the existence of a steady state distribution there is no guarantee of finite long-run average cost. This is required for the MDP formulation provided. Since the phase transition process is assumed to be ergodic, it has unique stationary probabilities denoted by, \( \{p_1, p_2, \ldots, p_L\} \).

**Proposition 2.1.** There exists a stationary policy under which the system is stable (steady state distribution exists) if and only if the maximum available service rate satisfies the following condition

\[
\bar{u} > \sum_{s=1}^L p_s \lambda_s.
\]

Furthermore, the long-run average cost, \( g^\pi(x) \), is finite and independent of the initial state \( x \).

**Proof.** See Appendix.

**Proposition 2.2.** The following hold

1. For \( \alpha > 0 \), \( v_\alpha(x) \) satisfies the DCOE (2.3). Moreover, any stationary policy \( \pi_\alpha \) that minimizes the right side of the DCOE (2.3) is \( \alpha \)-discounted expected cost optimal.

2. If the stability condition (2.5) holds, we have,

   (a) There exists a stationary long-run average expected cost optimal policy \( \pi^* = \{\mu^*_{n,s} \mid n \geq 1, s \in \{1, 2, \ldots, L\} \} \) that is a limit point of a sequence of discounted expected cost optimal policies say \( \{\pi_{\alpha_k}, k \geq 1\} \). That is, \( \mu^*(n, s) = \lim_{k \to \infty} \mu_{\alpha_k}(n, s) \), where \( \alpha_k \downarrow 0 \).

   (b) The long-run average expected cost associated with policy \( \pi^* \) is \( g^* = \lim_{\alpha \downarrow 0} \alpha v_\alpha(x) \) for every \( x \in X \). Moreover, there exists a subsequence \( \alpha_k \downarrow 0 \) such that \( \lim_{k \to \infty} w_{\alpha_k}(x) := v_{\alpha_k}(x) - v_{\alpha_k}(0) = w(x) \) for a distinguished state \( 0 \) such that \( (w, g^*) \) satisfy the ACOI (2.4).
3 Structural Properties of Optimal Policies

In this section we derive structural results for optimal policies for both the discounted cost and the average cost criterion. In a manner similar to \[11\], we use following definitions to simplify the optimality equations,

\[ y_{\alpha}(0, s) = 0 \text{ for } s = 1, 2, \ldots, L, \]
\[ y_{\alpha}(n, s) = v_{\alpha}(n, s) - v_{\alpha}(n - 1, s) \text{ for } n \in \mathbb{N}, s \in \mathcal{S}, \]

\[ \phi(y) = \max_{\mu \in \mathcal{A}} \{\mu y - c(\mu)\}, \text{ and} \]
\[ \psi(y) = \arg\max_{\mu \in \mathcal{A}} \{\mu y - c(\mu)\}, \]

where the argmax is a singleton by the assumptions on \(c\) (cf. Section 4.3 of \[1\]). The definitions above yield the following simplified form of the DCOE (2.3):

\[ v_{\alpha}(n, s) = \frac{1}{\alpha + \nu} \left[ h(n) - \phi(y_{\alpha}(n, s)) + \lambda_s v_{\alpha}(n + 1, s) + \sum_{s' = 1}^{L} Q_{ss'} v_{\alpha}(n, s') \right] \]
\[ + (\nu - \lambda_s) v_{\alpha}(n, s) \]

for \(n \in \mathbb{Z}^+, s \in \mathcal{S}. \) (3.2)

In order to derive structural results for an optimal discounted expected cost policy, we make use of several important properties of functions \(\phi(y) = \max_{x \in \mathcal{A}} \{yx - c(x)\}\) and its associated maximizers \(\psi(y) = \arg\max_{x \in \mathcal{A}} \{yx - c(x)\}\) that were introduced in the DCOE (2.3). Recall that the conjugate of \(c(\cdot), \phi(\cdot),\) is convex (cf. \[5\]). Moreover, \(\psi(y)\) is continuous, non-decreasing and equals \(\phi'(y)\) wherever the derivative exists. As described in \[1\], since \((c')^{-1}(\cdot)\) is well-defined, continuous and strictly increasing we have the following characterization of the function \(\psi(\cdot)\)

\[ \psi(y) = \begin{cases} 
0 & \text{if } y \leq c'(0), \\
(c')^{-1}(y) & \text{if } c'(0) < y < c'(\bar{u}), \\
\bar{u} & \text{if } y > c'(\bar{u}).
\end{cases} \] (3.3)

It may also be established (see \[11\]) that \(\phi(\cdot)\) is continuous and non-decreasing with the following characterization

\[ \phi(y) = \begin{cases} 
0 & \text{if } y < 0, \\
\int_0^y \psi(x)dx & \text{if } y \geq 0.
\end{cases} \] (3.4)

3.1 Monotone in the number of customers

We show the intuitive result that an optimal policy exists that is monotone in \(n.\)
Proposition 3.1. The following hold

1. For each \( s \in S \), the optimal discounted expected cost value function, \( v_\alpha(n, s) \), satisfies the DCOE (2.3) and is a non-decreasing, convex function of \( n \).

2. There exists a discounted expected cost optimal policy \( \{ \mu_\alpha(n, s), n \geq 1, s \in S \} \) that is non-decreasing in \( n \) for each \( s \in S \).

3. Under the assumptions that the holding cost is linear and (2.5) hold, there exists a long-run average optimal policy, \( \{ \mu(n, s), n \geq 1, s \in S \} \) that is non-decreasing in \( n \) for each \( s \in S \).

Proof. We use induction and the FHOE (2.2) to prove the first result. The result holds trivially for \( k = 0 \). For the inductive step, suppose \( v_{\alpha,k}(\cdot, s) \) is non-decreasing and convex on \( \mathbb{Z}^+ \) for each \( s \in S \). Let \( u_n = \mu_{\alpha,k}^\alpha(n, s) \) be the optimal service rate for the \((k+1)\)-stage problem when the state is \((n, s)\). Suppose we use the potentially sub-optimal decision \( u_{n-1} = \mu_{\alpha,k}^\alpha(n-1, s) \) be the optimal rate when the state is \((n-1, s)\). The FHOE (2.2) yield,

\[
v_{\alpha,k+1}(n-1, s) \leq \frac{1}{\alpha + \nu} \left[ c(u_{n-1}) + h(n-1) + u_{n}v_{\alpha,k}(n-2, s) + \lambda_s v_{\alpha,k}(n, s) + \sum_{s' \neq s} Q_{ss'} v_{\alpha,k}(n-1, s') + (\nu - \lambda_s - u_{n-1}) v_{\alpha,k}(n-1, s) \right],
\]

where the second inequality follows from the induction hypothesis. Thus, \( v_{\alpha,k} \) is non-decreasing for all \( k \).

To show convexity note that by the inductive hypothesis, \( y_{\alpha,k}(n+1, s) = v_{\alpha,k}(n+1, s) - v_{\alpha,k}(n, s) \) is a non-decreasing function of \( n \) for each \( s \in S \). Let \( u_{n+1} = \mu_{\alpha,k}^\alpha(n+1, s) \) be the optimal rate for the \((k+1)\)-stage problem when the state is \((n+1, s)\) and \( u_{n-1} = \mu_{\alpha,k}^\alpha(n-1, s) \) be the optimal rate when the state is \((n-1, s)\). The DCOE (2.3) imply (for \( n \geq 1 \))

\[
(\alpha + \nu) y_{\alpha,k+1}(n+1, s) \geq c(u_{n+1}) - c(u_{n+1}) + h(n+1) - h(n) - u_{n+1}(y_{\alpha,k}(n+1, s) - y_{\alpha,k}(n, s)) + \lambda_s y_{\alpha,k}(n+2, s) + \sum_{s' = 1}^L Q_{ss'} y_{\alpha,k}(n+1, s').
\]
Similarly for $y_{\alpha,k}(n, s) = v_{\alpha,k}(n, s) - v_{\alpha,k}(n-1, s)$,

$$(\alpha + \nu)y_{\alpha,k+1}(n, s) \leq c(u_{n-1} - c(u_{n-1}) + h_n - h(n-1) + (\nu - \lambda_s)y_{\alpha,k}(n, s)$$

$$- u_{n-1}(y_{\alpha,k}(n, s) - y_{\alpha,k}(n-1, s)) + \lambda_s y_{\alpha,k}(n+1, s)$$

$$+ \sum_{s' = 1}^{L} Q_{s,s'} y_{\alpha,k}(n, s').$$

Using the definitions of $\nu = \lambda_L + \bar{\eta} + \bar{u}$ and $Q = \bar{\eta} I + \bar{Q} \lambda$ we have,

$$(\alpha + \nu)(y_{\alpha,k+1}(n+1, s) - y_{\alpha,k+1}(n, s)) \geq h(n + 1) - 2h(n) + h(n-1)$$

$$+ (\lambda_L + \bar{u} - \lambda_s - u_{n+1}) \left( y_{\alpha,k}(n+1, s) - y_{\alpha,k}(n, s) \right)$$

$$+ u_{n-1}(y_{\alpha,k}(n, s) - y_{\alpha,k}(n-1, s))$$

$$+ \lambda_s (y_{\alpha,k}(n+2, s) - y_{\alpha,k}(n+1, s))$$

$$+ \sum_{s' = 1}^{L} Q_{s,s'} \left( y_{\alpha,k}(n+1, s') - y_{\alpha,k}(n, s') \right) \geq 0.$$ 

The second inequality follows as $h(n)$ is convex, the coefficients of $y_{\alpha,k}$ terms are non-negative and the inductive hypothesis. So $y_{\alpha,k}(\cdot, s)$ is non-decreasing on $\mathbb{Z}^+$ for all $s \in S$ as required.

Taking limits as $k \to \infty$ yields that $v_{\alpha}$ is non-decreasing and convex; the first result is proven.

Now as the function $\psi(\cdot)$ is non-decreasing, we conclude that there exists an optimal policy for the discounted cost problem $\mu_{\alpha}(\cdot, s)$ is non-decreasing on $\mathbb{Z}^+$ for each $s \in S$. This is the second result. Consider now a subsequence of discount factors such $\alpha_k \to 0$ and stationary discounted expected cost optimal policies $\mu_{\alpha_k}(\cdot, \cdot)$ that converge to an average cost optimal policy $\mu(\cdot, \cdot)$. The previous result implies that for each fixed $s \in S$, $\mu_{\alpha_k}(n, s) \leq \mu_{\alpha_k}(n+1, s)$. Thus, the same inequality holds for $\mu(\cdot, \cdot)$ and the result is proven.

We remark that we have explicitly used the fact that argmax in $\psi$ is a singleton (which follows from the strict convexity assumption on $c(\cdot)$). When the convexity is not assumed to be strict, the results still hold, but we need to take care to define $\psi$ as the minimal element of the argmax and consider a subsequence of discount factors such that $w_{\alpha_k}(x) = v_{\alpha_k}(x) - v_{\alpha_k}(0) \to w(x)$. Since $v_{\alpha_k}(n, s)$ is non-decreasing in $n$, so is $w$ and proof in the average cost case follows in the same way as the discounted cost case except that we use the ACOI instead of the DCOE.

### 3.2 Monotone in the phase process

Since the states of the phase transition process are ordered such that $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_L$, one might conjecture that the optimal policy is non-decreasing in the phase state, $s$, for each congestion level, $n$. However, we present two examples to show that depending on the transition structure of the phase process, this property may not hold. In both examples, we use value
iteration with $\alpha = 0.05$ to compute the optimal policy numerically. We consider an exponential cost rate function, $c(\mu) = e^\mu - 1$, and a linear holding cost function, $h(n) = n$. Service rates can be selected from the set $\mathbb{A} = [0, 5]$.

![Transition structure of Phase process for Examples 3.1 and 3.2.](image)

**Figure 1:** Transition structure of Phase process for Examples 3.1 and 3.2.

**Example 3.1.**

In this example, the phase process is a birth and death process on the states $\{1, 2, 3\}$. See Figure 1(a). The infinitesimal generator for the phase process is given by

$$Q = \begin{bmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix},$$

and arrival rates are $\lambda_1 = 0.5, \lambda_2 = 1$ and $\lambda_3 = 1.25$.

Figure 2 shows that the optimal policy is a non-decreasing function of $n$ for each $s$. It should also be clear that the optimal service rates are non-decreasing in $s$ for each $n$.

**Example 3.2.**

Consider a phase process with a cyclic transition structure on the set of states $\{1, 2, 3\}$. See Figure 1(b). The infinitesimal generator matrix for the phase process is

$$Q = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix},$$

and the arrival rates are $\lambda_1 = 0.5, \lambda_2 = 1$ and $\lambda_3 = 1.25$. Figure 3(a) shows that the optimal policy is non-decreasing in $n$ for each $s$. However, it is clear from Figure 3(b) that the service rates are not monotone in $s$ when queue length is 4.

In Example 3.2, the phase process starting at the highest arrival intensity state, 3, transitions to the lowest arrival intensity state, 1. This cyclic variation causes the optimal service rate to
be higher in state 2 as compared to state 3 for some congestion levels and thereby renders an optimal policy that is not monotone in $s$. These examples beg the question, is there a reasonable assumption under which the optimal policy is monotone in $s$? *Stochastic monotonicity* of the phase transition process is one such assumption.

### 3.2.1 Stochastic Monotonicity for Continuous Time Markov Chains

Intuitively, stochastic monotonicity means that given the arrival process is in a high arrival intensity state, the future states it will encounter are in some sense worse (in terms of arrival intensity) than if the process is in a low arrival intensity state. This leads to the following definitions (see for example Keilson and Kester [16])

**Definition 3.2.** Given two probability vectors $p$ and $q$, a stochastic matrix $M$ and a homogeneous Markov chain $\{X(t), t \geq 0\}$ with probability transition function $P(t) = P_{ij}(t)$.

1. $p$ stochastically dominates $q$ ($p \geq_{st} q$) iff $\sum_{i=n}^{N} p_i \geq \sum_{i=n}^{N} q_i$, $n = 1, 2, \ldots$.

2. Letting $M^i$ denote the $i^{th}$ row of the matrix, $M$ is called stochastically monotone if $M^k \geq_{st} M^l$, whenever $k > l$.

3. $\{X(t), t \geq 0\}$ is said to be stochastically monotone if $P(t)$ if monotone.

Note that the transition structure shown in Example 3.1 is stochastically monotone while that for Example 3.2 is not (uniformize and it is trivial). The following provides alternative
Figure 3: Structure of Optimal Policy for Example 3.2

(b) Optimal Rate for $n = 4$
methods for specifying when a transition matrix is stochastically monotone (again refer to Keilson and Kester [16]).

**Proposition 3.3.** For monotone matrices the following are equivalent.

1. M is monotone.
2. \((T^{-1}MT)_{ij} \geq 0\) where T is a square matrix with 1’s on or below the diagonal.
3. \(pM \geq_s qM\) for all probability vectors \(p, q\) with \(p \geq_s q\).
4. \(Mv\) is non-decreasing for all non-decreasing vectors \(v\).

Since for a continuous-time Markov chain with generator matrix \(Q\), the transition matrix \(P(t) = \sum_{n=0}^{\infty} \frac{(Q)^n}{n!}\), stochastic monotonicity implies that the generator for the phase process satisfies the property that \((T^{-1}QT)_{ij} \geq 0, i \neq j\). Choosing a uniformizing constant \(\bar{\eta}\) yields \((T^{-1}(\bar{\eta}I + Q)T) \geq 0\) in all elements. Thus using Proposition 3.3 we have that \(Q := \bar{\eta}I + Q\) satisfies the property that \(Qv\) is non-decreasing for all non-decreasing vectors \(v\). This leads to the next result; the main result of this section.

**Theorem 3.4.** Suppose that the phase transition process is stochastically monotone.

1. For each \(n \in \mathbb{Z}^+\), \(y_{\alpha,k}(n, s)\) is non-decreasing function of \(s\).
2. There exists a discounted cost optimal policy, \(\mu_{\alpha}(n, s)\), that is non-decreasing in \(s\) for each \(n\).
3. There exists an average cost optimal policy \(\mu(n, s)\), that is non-decreasing in \(s\) for each \(n\).

**Proof.** We show the first result by induction. The second and third results follow in an analogous manner to Proposition 3.1. The statement holds trivially for \(k = 0\). Assume it holds for \(k\). Using the definitions \(\nu = \lambda_L + \bar{\eta} + \bar{u}\) and \(\bar{Q} = \bar{\eta}I + Q\) we have for \(s > 1\)

\[
(\alpha + \nu) (y_{\alpha,k+1}(n+1, s) - y_{\alpha,k+1}(n+1, s-1))
= \phi (y_{\alpha,k}(n, s)) - \phi (y_{\alpha,k}(n, s-1)) + \lambda_s (y_{\alpha,k}(n + 2, s) - y_{\alpha,k}(n + 2, s - 1))
+ (\lambda_s - \lambda_{s-1}) (y_{\alpha,k}(n + 2, s - 1) - y_{\alpha,k}(n + 1, s - 1))
+ (\lambda_L - \lambda_s) (y_{\alpha,k}(n + 1, s) - y_{\alpha,k}(n + 1, s - 1))
+ [\bar{u} (y_{\alpha,k}(n + 1, s) - y_{\alpha,k}(n + 1, s - 1))
- (\phi(y_{\alpha,k}(n + 1, s)) - \phi(y_{\alpha,k}(n + 1, s - 1)))]
+ \sum_{s'=1}^{L} \bar{Q}_{s,s'} y_{\alpha,k}(n + 1, s') - \sum_{s'=1}^{L} \bar{Q}_{s-1,s'} y_{\alpha,k}(n + 1, s')
\] (3.5)
Now as \( \phi(y) = \int_0^y \psi(x)dx \) and \( \psi(y) \leq \bar{u}, \)

\[
\phi(y_{\alpha,k}(n+1,s)) - \phi(y_{\alpha,k}(n+1,s-1)) = \int_{y_{\alpha,k}(n+1,s-1)}^{y_{\alpha,k}(n+1,s)} \psi(x)dx \\
\leq \bar{u}(y_{\alpha,k}(n+1,s) - y_{\alpha,k}(n+1,s-1)).
\]

So the term in (3.5) is non-negative as is the next to last term. Furthermore, since the inductive hypothesis and the assumption on \( Q \) implies the last term in (3.5) is non-negative. Thus, we conclude from the induction hypothesis that

\[
(\alpha + \nu)(y_{\alpha,k+1}(n+1,s) - y_{\alpha,k+1}(n+1,s-1)) \geq 0,
\]

as desired.

\section{Numerical Study}

This section provides two insights. First, we compare the optimal control policy with two natural heuristics. When the environment is changing, it seems a decision-maker that is not armed with the current research might take one of two courses. (S)he might choose to ignore the state change of the environment altogether, or she might treat each state change as permanent and react accordingly. In either case, the resulting control policies are heuristics when compared to the optimal control that takes into account both the phase and queue length processes. The first goal is then to compare the optimal policy with these heuristics. Second, as alluded to in Section 1, the current model can act as a heuristic itself when compared to a model with NHPP arrivals. We analyze when this is a reasonable approximation.

\subsection{Comparison with Heuristics}

The first heuristic that we consider uses the optimal control for an average cost problem where the arrival process is Poisson with the long-run mean arrival rate of the MMPP. When applied to the original model, this policy is a function of the queue length only. We call this heuristic the \textit{Average Rate Method} (ARM).

Since the state of the arrival process is known, the decision-maker may solve the stationary model with each potential arrival rate and change the service rate according to the current state of the arrival process. That is to say, a second heuristic is derived in the following way:

1. Compute the service rate control average cost optimal policy, \( \pi^h_s \), for a system with Poisson arrivals with rate \( \lambda_s \) for each intensity level \( s \in S \).

2. The heuristic for the Markov-Modulated queue is obtained by using \( \pi^h_s \) when the state of the process is \( (n, s) \).
This heuristic is referred to as the Phase Rate based Method (PRM). Note that the long-run average arrival rate used in computing the ARM policy is influenced by both the infinitesimal generator matrix and the arrival rates of the phase process while the PRM policy relies only on the arrival rates.

A numerical study comparing the performance of these heuristics for various test cases is provided in Examples 4.1 and 4.2. In all cases, the policies and average cost are computed using value iteration where the queue length is truncated at 50. We use the cost rate function $c(\mu) = e^\mu - 1$ and holding cost rate $h(n) = n$. Service rates are allowed to be chosen from $A = [0, 15]$. In each case the arrival rates change in accordance with the phase state $\{1, 2, \ldots, 8\}$ with the arrival rates as shown in Table 1.

<table>
<thead>
<tr>
<th>Case</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0.1</td>
<td>0.35</td>
<td>0.6</td>
<td>0.85</td>
<td>1.1</td>
<td>1.35</td>
<td>1.6</td>
<td>1.85</td>
</tr>
<tr>
<td>II</td>
<td>0.1</td>
<td>0.6</td>
<td>1.1</td>
<td>1.6</td>
<td>2.1</td>
<td>2.6</td>
<td>3.1</td>
<td>3.6</td>
</tr>
<tr>
<td>III</td>
<td>0.1</td>
<td>0.85</td>
<td>1.6</td>
<td>2.35</td>
<td>3.1</td>
<td>3.85</td>
<td>4.6</td>
<td>5.35</td>
</tr>
</tbody>
</table>

Table 1: Arrival Rate Parameters for Phase Transition Process in Examples 4.1 and 4.2

Example 4.1.

Suppose that phase process is a birth and death process on states $\{1, 2, \ldots, 8\}$ (recall Figure 1(a)). Fix $c > 0$. The transition rates for the phase process are $\eta_{i,i+1} = \eta_{i,i-1} = c$ for $2 \leq i \leq 7$, $\eta_{1,2} = c$ and $\eta_{8,7} = c$. Note that higher values of $c$ imply that the phase process transitions faster between the arrival phases. We referred to $c$ as the fluctuation rate scaling parameter. We perform the numerical analysis for three sets of arrival rates for the phase process provided in Table 1. For each set the parameter $c$ takes values 0.25, 0.50, 0.75 and 1.00 resulting in a total of 12 different scenarios for the arrival process. The results are provided in Table 2.

A few observations are in order. Ignoring the dynamic state information of the phase process (and using the ARM policy) is more costly when the fluctuation parameter is lower. This stands to reason since the phase process can be in a state for a long period of time, while the ARM policy assumes the arrival rate is the mean arrival rate. In Case III, when the arrival rate change is the most between phase states and with the slowest rate of changing states, the percent sub-optimality is above 9%. If we try to approximate the state changes with stationary processes (using PRM) we see that again, the percent sub-optimality is high (above 13%) but this time when the fluctuation parameter is highest.

Figures 4(a)-4(c) show the change in the average cost rate under the heuristic and optimal policies as a function of the parameter $c$ for the three arrival cases. Figure 4(d) shows a comparison of the heuristics and optimal policies for the arrival rates in Case III and $n = 2$ for various values of fluctuation rate parameter (the behavior is similar for other values of $n$).
<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Gain (% Sub-Optimal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrival Rates</td>
<td>c</td>
</tr>
<tr>
<td>Case I</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Case II</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
</tr>
<tr>
<td>Case III</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2: Average Cost Rates and Percentage Difference between Optimal and Heuristic Policies for Example 4.1.

It should be clear that the PRM policy outperforms the ARM policy in most cases except for high values of $c$ in Case III. Moreover, one should note that the performance of the ARM policy improves while that of PRM policy degrades in comparison to the optimal policy as the fluctuation rate parameter increases. Intuitively this seems reasonable since for low values of $c$ the phase process spends more time in each phase. Therefore the PRM policy, that in each phase applies the optimal policy for a stationary M/M/1 queue with arrival rate of that particular phase, performs better than the ARM policy. In fact, if $c = 0$, the PRM policy is optimal since the phase process is stationary with the arrival rate of initial phase.

At high values of $c$ (since the system sees more and more transitions), the arrival process behaves like a Poisson process with the average arrival rate of MMPP. Therefore the PRM policy gets penalized more in comparison to the ARM policy in this case. In Figure 4(d) we also see that the change in the optimal service rates as a function of phase state for each congestion level is lower for higher values of $c$.

**Example 4.2.**

In this example we study a cyclic phase process (cf. Figure 1(b)) on the states $\{1, 2, \ldots, 8\}$. The transition rates for the phase process are $\eta_{i,i+1} = c$ for $1 \leq i \leq 7$ and $\eta_{8,1} = c$. Similar to the previous example we perform the numerical analysis for 12 scenarios for the phase process; three different sets of arrival rates given in Table 1 and for each set of arrival rates, $c$ takes values 0.25, 0.50, 0.75 and 1.00.

Figures 5(a)-5(c) show the change in the average cost computed under the heuristics as well.
as the optimal policies as a function of parameter $c$ for the three arrival cases. It is interesting to observe that unlike Example 4.1, the ARM policy outperforms the PRM policy in almost all cases. As illustrated in Example 3.2, when the phase process has cyclic transitions, the optimal service rates may not be monotone in the phase process (for each fixed $n$). In fact, while the service rates for the PRM policy are monotone in the phase of the transition process for each congestion level, the service rates in the optimal policy may begin to decrease as the phase state increases. This can be more clearly observed in Figure 5(d) which shows a comparison of heuristic and optimal policies as a function of the phase state when $n = 2$ for various values of parameter $c$ and the arrival rates of Case III. Thus the ARM policy approximates the optimal policy better than the PRM policy which explains the observed performance difference.
Similar to the previous example, we observe that the performance of the ARM policy gets worse and that of PRM policy improves with a decrease in the fluctuation rate parameter $c$. Furthermore, the average cost percent differences provided in Table 3 show that the ARM policy performs extremely well for all three arrival rate cases (less than 5% from optimal for all cases). The PRM policy performs well for Case I but the degradation in its performance is quite significant (almost 35% from optimal for Case III with $c = 1$) when the difference in arrival rates is medium or high (Cases II and III).

These examples show that one can not rely on a particular heuristic method to perform well in all scenarios. In particular, we find that within the gamut of simple heuristic methods considered here, the transition structure and the transition rates of the phase process plays an important role in the selection of an appropriate approximation method.

### 4.2 Approximation of system with non-homogeneous Poisson arrivals

When the arrival process follows known rate changes, a non-homogeneous Poisson process is a reasonable modeling tool. In the classic work of Green and Kolesar [12] or Massey and Whitt [18] the analysis of queues with non-stationary arrivals is considered. From the standpoint of control, Yoon and Lewis [25] consider the case of admission and pricing control. One thing is certain from Yoon and Lewis's work, control of non-stationary processes can be computationally intensive. This is due the fact that to solve each instance the numerical approach requires the time to be discretized.

In this section we explore the possibility of computing an approximate average cost optimal
policy for a single server queue with non-homogeneous Poisson arrivals using the optimal policies for a system with a “suitable” Markov-modulated Poisson arrival process. Apart from the arrival process, other details are same as the setting described in Section 2. For the purpose of this study, we consider a NHPP with the following (periodic) rate function,

\[ \lambda(t) = \begin{cases} 
0.1 & \text{for } 0 \leq t < T/5 \\
2.0 & \text{for } T/5 \leq t < 2T/5 \\
4.0 & \text{for } 2T/5 \leq t < 3T/5 \\
2.0 & \text{for } 3T/5 \leq t < 4T/5, \\
0.1 & \text{for } 4T/5 \leq t < T, 
\end{cases} \]
where $T$ is the time period of the rate function. One can think of this rate function as a quantized version of a triangular waveform with time period $T$. Since the optimization criterion considered in this study is over an infinite time horizon, and the rate function for NHPP is a periodic function of time, the principle of optimality implies that only the time elapsed in the current period and the number of jobs in the system need to be included in the state space [25].

To compute the optimal policy numerically, the time period is divided into $n$ equally spaced segments of length $\Delta t = T/n$. Denote the state space for this discretized process as $X = \{(n, z) \mid n \in \mathbb{Z}^+, s \in \{0, \Delta t, \ldots, T - \Delta t\}\}$. Under this setting, the decision epochs are the time points $0, \Delta t, 2\Delta t, \ldots, T - \Delta t$. Let $\nu$ be a uniformizing rate of the process. An event (arrival, departure or dummy transition) occurs at a decision epoch with probability $1 - e^{-\nu \Delta t}$ and with probability $e^{-\nu \Delta t}$ no event occurs. The standard theory of Markov decision processes yields the following average cost optimality inequality (ACOI):

$$w(n, z) \geq \min_{x \in A} \left\{ (-g + h(n) + 1_{\{n>0\}} c(x) ) \Delta t + (1 - e^{-\nu \Delta t}) \left( \frac{\lambda(z)}{\nu} w(n + 1, z + \Delta t) ight. \\
+ 1_{\{n>0\}} \frac{x}{\nu} w(n - 1, z + \Delta t) + \left(1 - \frac{\lambda(z)}{\nu} \right) 1_{\{n>0\}} \frac{x}{\nu} w(n, z + \Delta t) \\
+ e^{-\nu \Delta t} w(n, z + \Delta t) \right\} \text{ for } n \in \mathbb{Z}^+, z = 0, \Delta t, \ldots, T - 2\Delta t,$$

and

$$w(n, T - \Delta t) \geq \min_{x \in A} \left\{ (-g + h(n) + 1_{\{n>0\}} c(x) ) \Delta t + (1 - e^{-\nu \Delta t}) \left( \frac{\lambda(T - \Delta t)}{\nu} w(n + 1, 0) ight. \\
+ 1_{\{n>0\}} \frac{x}{\nu} w(n - 1, 0) + \left(1 - \frac{\lambda(z)}{\nu} \right) 1_{\{n>0\}} \frac{x}{\nu} w(n, 0) \\
+ e^{-\nu \Delta t} w(n, 0) \right\} \text{ for } n \in \mathbb{Z}^+,$$

where $1_E$ is the indicator function of the event $E$. When the solution, $(w, g)$ to the ACOI exists, $w$ is called the relative value function and $g^*(x) = g$ is the optimal long-run expected average cost for any initial state $x$.

For computing the approximate policy, we used a Markov-modulated arrival process with a cyclic transition structure for the phase process. The associated generator matrix is

$$Q = \frac{5}{T} \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{bmatrix},$$

and the arrival rates are $\lambda_1 = 0.1, \lambda_2 = 2.0, \lambda_3 = 4.0, \lambda_4 = 2.0$ and $\lambda_5 = 0.1$. Note that the mean sojourn times for the phase process in any phase is $T/5$ units which is also the time spent by the NHPP at each intensity level. To keep the size of numerical problem manageable, we
truncated the queue state to 51. The discretization interval $\Delta t$, is 0.05 units. We used a cost rate function $c(\mu) = e^\mu - 1$ and holding cost function $h(n) = n$. Service rates can be selected from a set $A = [0, 10]$.

![Service Rate Graphs](image)

Figure 6: NHPP Policy vs Approximate MMPP policy for various congestion levels (T=4)

<table>
<thead>
<tr>
<th>Time Period (T)</th>
<th>Gain(% Sub-Optimal)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Optimal</td>
</tr>
<tr>
<td>4</td>
<td>8.5667</td>
</tr>
<tr>
<td>5</td>
<td>8.7750</td>
</tr>
<tr>
<td>6</td>
<td>8.7467</td>
</tr>
<tr>
<td>7</td>
<td>8.8225</td>
</tr>
</tbody>
</table>

Table 4: Average Cost Rates and Percentage Difference between Optimal and Approximate NHPP Policy

Figures 6 and 7 show a comparison of optimal policy with approximate MMPP policy. Table 4 gives the average cost percent difference between the performance of approximate and optimal policy for various test scenarios. This data shows that the approximate policies perform extremely well in all cases (less than 1% sub-optimal). Of course, we can not conclude from this limited analysis that MMPP approximate policies will perform well for NHPP in more general settings. Our motivation in presenting this analysis is to stimulate further research in this direction.
Figure 7: NHPP Policy vs Approximate MMPP policy for various congestion levels (T=7)

5 Conclusions

In this paper we investigate the problem of service rate control for a single server queue with non-stationary arrivals. We propose a framework based on the Markov modulated Poisson processes; a popular model amongst practitioners that is also relatively easy to analyze. Assuming that the goal is to minimize a combination of effort cost and holding cost incurred per unit time, we study this problem under both the discounted and average cost optimality criterion. In either case, we characterize the structure of an optimal service rate as being monotone in the queue length for each arrival rate but perhaps not monotone in the arrival rates for each queue length. In particular, we show that the manner in which the process switches between the arrival rates plays an important role in determining the structure of the optimal policy. We further prove that monotonicity in the arrival rates is recovered when the transition matrix governing the MMPP is stochastically monotone. In this case, a monotone switching curve is the optimal control.

There are several ways to extend our work. Our numerical study confirms that in some cases simple heuristics may perform well in the face of changing arrival rates. However, it also points out that careful selection based on the parameters of the system is required, and in many cases applying the proposed model is essential (as opposed to the heuristics). The second part of our numerical work points to the fact that our model can be used as a heuristic itself. We show that we can provide a policy for a system with non-homogenous Poisson arrivals using the
optimal policy for an MMPP/M/1 queue. Our results indicate that this may be a promising direction for future research.

Another problem of interest is that of the control of MMPP/M/1 queue with a finite buffer but with an explicit constraint on the job loss rate. Under this setting, while the technical conditions required for stability are not needed, handling the explicit constraint poses a significant challenge. We note that results provided by Ata [1] for the stationary arrival case may be useful.

Finally, we would like to point out that although we assume that the cost of effort function, \( c(\mu) \), is strictly convex, continuously differentiable and non-decreasing, our proofs (with minor modification) and results hold for more general cost of effort functions. Using the analysis presented by George and Harrison [11], it can be easily shown that the structural results for an optimal policy continue to hold when \( c(\mu) \) is assumed to be non-decreasing and continuous.

References


This appendix is dedicated to providing proofs of Propositions 2.1 and 2.2. The results of Dai [7] and Dai and Meyn [8] are utilized to show the stability of a stochastic model by establishing the stability of its fluid limit approximation. For the purpose of this analysis, consider the continuous time Markov process, $X^\pi(t) = \{(Q^\pi(t), S^\pi(t)), t \geq 0\}$, induced by an admissible stationary policy $\pi \in \Pi$ where $\{Q^\pi(t), t \geq 0\}$ and $\{S^\pi(t), t \geq 0\}$ represent the queue length and phase transition process for arrivals, respectively.

The proof approach for Proposition 2.1 follows closely that of Kaufman and Lewis [15] and is done in several steps. Let $\hat{\pi}$ be a policy that selects a constant rate $\hat{\mu} \in (\sum_{s=1}^{L} p_s \lambda_s, \bar{u}]$ whenever the queue is non-empty. Let $\hat{X}(t) = \{(\hat{Q}(t), \hat{S}(t)), t \geq 0\}$ be the Markov process induced by the policy $\hat{\pi}$ on state space $X = \{(n,s) | n \in \mathbb{Z}^+, s \in S\}$. Since the policy is fixed for the remainder of this section, in the interest of brevity we suppress dependence on $\hat{\pi}$. The norm of a state, $x = (n,s) \in X$, is defined to be $|x| := n + s$. For an initial state $X(0) = x$, we define the scaled queue length process

$$\tilde{Q}^x(t) := \frac{1}{|x|} Q^x(|x| t).$$

We will use a similar notation to denote scaled versions of other stochastic processes.

For each $s \in S$, let $\{\xi_s(k), k \geq 1\}$ be a sequence of i.i.d exponential random variables with mean $1/\lambda_s$. The sequence $\{\xi_s(k), k \geq 1\}$ represents the set of job inter-arrival times when the arrival process is in phase $s$. Also, let $\{\eta(k), k \geq 1\}$ be a sequence of i.i.d exponential random variables with mean 1 representing the set of job completion times. Based on these sequences, we define the following cumulative processes

$$E_s(t) = \max\{k \geq 0 \mid \xi_s(1) + \xi_s(2), \ldots, \xi_s(k) \leq t\} \quad \text{for } s \in S,$$

$$D(t) = \max\{k \geq 0 \mid \eta(1) + \eta(2), \ldots, \eta(k) \leq t\}.$$

Let $Y^x_s(t)$ be the cumulative amount of time the arrival process spends in phase $s$ until time $t$ when the initial state is $x$. Similarly, let $T^x_s(t)$ be the cumulative amount of time for which there is at least one customer in the queue, $I^x(t)$ be the cumulative amount of time when the
queue is empty and $W^x(t)$ be the total work done by the server until time $t$. We can now write
the following system of equations for the stochastic process induced by policy $\hat{\pi}$ when starting
from initial state $x$,

$$Q^x(t) = Q^x(0) + \sum_{s=1}^{L} E_s(Y^x_s(t)) - D(W^x(t)), \quad (6.1)$$

$$Q^x(t) \geq 0, \quad (6.2)$$

$$\sum_{s=1}^{L} Y^x_s(t) = t, \quad (6.3)$$

$$W^x(t) = \hat{\mu} T^x(t), \quad (6.4)$$

$$T^x(t) + I^x(t) = t, \quad (6.5)$$

$$\int_{0}^{\infty} Q^x(t)dI^x(t) = 0, \quad (6.6)$$

$$Y^x_s(t), T^x(t), I^x(t), W^x(t) \text{ start from zero and are non-decreasing in } t. \quad (6.7)$$

A few comments are in order. First, note that (6.6) imposes the constraint that the server is
idle only when the system is empty. For a subsequence $\{x_n, n \geq 1\}$ such that $|x_n| \to \infty$, any
limit point, $\bar{Q}(t)$, of the sequence $\{\bar{Q}^x_n, n \geq 1\}$ is called a fluid limit. It will be shown that
every fluid limit satisfies a set of equations known as the fluid model. A fluid model is called
stable if there exists a $t_0 > 0$ such that $\bar{Q}(t) = 0$ for all $t \geq t_0$ and for all fluid limits. We next
present the fluid model and convergence results for the scaled processes.

**Proposition 6.1.** Let $\{x_j \mid x_j \in X, j \geq 1\}$ be a sequence of initial states with $|x_j| \to \infty$. Then
with probability 1, there exists a subsequence, $\{x_{j_k}, k \geq 1\}$, such that

$$\left(\bar{Q}^{x_{j_k}}(0), \bar{S}^{x_{j_k}}(0)\right) \to (\bar{Q}(0), 0), \quad (6.8)$$

$$\left(\bar{Q}^{x_{j_k}}(t), \bar{T}^{x_{j_k}}(t)\right) \to (\bar{Q}(t), \bar{T}(t)) \text{ uniformly on compact sets (u.o.c.)}, \quad (6.9)$$

where $(\bar{Q}(t), \bar{T}(t))$ satisfy the following equations,

$$\bar{Q}(t) = \bar{Q}(0) + \sum_{s=1}^{L} p_s \lambda_s t - \bar{W}(t), \quad (6.10)$$

$$\bar{Q}(t) \geq 0, \quad (6.11)$$

$$\bar{W}(t) = \hat{\mu} \bar{T}(t), \quad (6.12)$$

$$\bar{T}(t) + \bar{I}(t) = t, \quad (6.13)$$

$$\int_{0}^{\infty} \bar{Q}(t)d\bar{I}(t) = 0, \quad (6.14)$$

$$\bar{T}(t), \bar{I}(t), \bar{W}(t) \text{ start from 0 and are non-decreasing in } t. \quad (6.15)$$

26
Proof.} Since \( Q^{x_j}(0) \leq 1, S^{x_j}(0) \leq 1 \) and \( 1 \leq S^{x_j}(0) \leq L \) for all \( j \in \mathbb{N} \), there exists a subsequence, \( \{x_{j_k}, k \geq 1\} \) such that \( (Q^{x_{j_k}}(0), S^{x_{j_k}}(0)) \rightarrow (Q(0), 0) \). For any fixed sample path \( \omega \) and \( 0 \leq s \leq t \), we have \( 0 \leq \bar{T}^{x_j}(t) = \bar{T}^{x_j}(s) \leq t-s \). Thus, the function \( \bar{T}^{x_j}(t) \) is uniformly Lipschitz of order 1. Since \( 0 \leq \bar{T}^{x_j}(t) \leq t \), it is also uniformly bounded for each \( j \geq 1 \). Therefore, the sequence \( \{\bar{T}^{x_j}(t), j \geq 1\} \) is equicontinuous. By Arzela-Ascoli theorem, any subsequence of \( \{\bar{T}^{x_j}(t), j \geq 1\} \) has a u.o.c. convergent subsequence.

Since the phase transition process is ergodic, for each \( s \in \{1, \ldots, L\} \) we have with probability 1 \( \lim_{t \to \infty} Y_s^x(t)/t = p_s \). Furthermore, from the strong law of large numbers for renewal processes, the following hold almost surely

\[
\lim_{t \to \infty} \frac{E_s(t)}{t} = \lambda_s \quad s \in S, \\
\lim_{t \to \infty} \frac{D(t)}{t} = 1.
\]

The above results can be used in a manner similar to Lemma 4.2 of [7], yield (with probability 1)

\[
\bar{Y}_s(t) = \lim_{k \to \infty} \frac{1}{|x_{j_k}|} Y^{x_{j_k}}(|x_{j_k}|, t) = p_s t \quad u.o.c., \text{ for } s \in S, 
\]

\[
\bar{E}_s(t) = \lim_{k \to \infty} \frac{1}{|x_{j_k}|} E(|x_{j_k}|, t) = \lambda_s t \quad u.o.c., \text{ for } s \in S
\]

\[
\bar{D}(t) = \lim_{k \to \infty} \frac{1}{|x_{j_k}|} D(|x_{j_k}|, t) = t \quad u.o.c.
\]

The equality in (6.10) follows from (6.1) and (6.16)-(6.18). Similarly, (6.12)-(6.15) follow directly from (6.4)-(6.7), respectively.

\textbf{Proof of Proposition 2.1:} We start by showing that the fluid model provided in (6.10)-(6.15) is stable. First note that \( \bar{T}(t), \bar{I}(t) \) and \( \bar{Y}_s(t) \) are Lipschitz continuous and therefore absolutely continuous and differentiable almost everywhere. Taking the derivative with respect to \( t \) in (6.10) and (6.12) yields

\[
\dot{Q}(t) = \sum_{s=1}^{L} p_s \lambda_s - \mu \bar{T}(t).
\]

Further, due to the non-idling constraint (6.14), \( \dot{I}(t) = 0 \) whenever \( \bar{Q}(t) > 0 \). Thus, from (6.13), \( \dot{I}(t) = 1 \) whenever \( \bar{Q}(t) > 0 \). So for \( \dot{Q}(t) > 0 \), we have

\[
\dot{Q}(t) = \sum_{s=1}^{L} p_s \lambda_s - \mu.
\]

Our choice of the stationary policy enabled by the stability condition (2.5), implies that \( \dot{Q}(t) < 0 \) whenever \( \bar{Q}(t) > 0 \). Thus from Lemma 5.2 of [7] we have that the fluid limit process for queue length is non-increasing and there exists a \( t_0 \geq 0 \) such that \( \bar{Q}(t) = 0 \) for all \( t \geq t_0 \). That is, the
fluid model is stable. The results of Theorem 4.2 of [7] imply that the Markov process induced by the stationary policy \( \hat{\pi} \), is positive recurrent and stationary distribution exists. Furthermore, since the embedded discrete time Markov chain for the process is irreducible, this process is ergodic. This implies that the long-run average cost under \( \hat{\pi} \), say \( g^{\hat{\pi}} \), is independent of the initial state.

To show that \( g^{\hat{\pi}} \) is finite, we use the results of Theorem 4.1(i) of [8]. Since the fluid model is stable and conditions A1) and A2) of Theorem 4.1 hold, it follows that \( \limsup_{t \to \infty} \frac{1}{t} \int_{0}^{t} \mathbb{E}_x[Q(u)]du < \infty \) for each initial condition \( x \). The linear holding costs then imply that the long-run average holding cost rate is finite. Moreover, the direct contribution to long-run cost rate due to serving at \( \hat{\mu} \) whenever the queue is not empty is at most \( c(\hat{\mu}) < \infty \). It therefore follows \( g^{\hat{\pi}} \) is finite.

It remains to consider the necessity of (2.5). Consider the Markov process induced by a policy that uses the highest available service rate whenever the queue is not empty. As shown by Yechiali[24] using the detailed balance equations for the steady state distribution a non-trivial invariant measure exists only if \( \bar{u} > \sum_{s=1}^{L} p_s \lambda_s \). The result follows and the proof is complete.

The remainder of this section is dedicated to proving Proposition 2.2 holds. We show that the optimal value and relative value functions satisfying the ACOI, (2.4), exist and can be obtained via limits from the discounted expected cost value functions. Thus, the structural results proved for the discounted cost case continue to hold for the average cost case. In proving these results, we verify the following set of assumptions (SEN) (included for completeness) provided by Sennott [21] hold.

- **SEN1**: There exist \( \delta > 0 \) and \( \epsilon > 0 \) such that for every state and action, there is a probability of at least \( \epsilon \) that the transition time will be greater that \( \delta \).

- **SEN2**: There exists \( B \) such that \( \tau(i, \mu) \leq B \) for every state \( i \) and control \( \mu \), where \( \tau(i, \mu) \) is the expected transition time out of state \( i \) when control \( \mu \) is chosen.

- **SEN3**: \( v_{\alpha}(i) < \infty \) for all states \( i \) and \( \alpha > 0 \).

- **SEN4**: There exists \( \alpha_0 > 0 \) and nonnegative numbers \( M_i \) such that \( w_\alpha(i) \leq M_i \) for every state \( i \) and \( 0 < \alpha < \alpha_0 \) where \( w_\alpha(i) = v_{\alpha}(i) - v_{\alpha}(0) \) for distinguished state \( 0 \). For every state \( i \), there exists an action \( \mu_i \) such that \( \sum_j P_{ij}(\mu_i)M_j < \infty \).

- **SEN5**: There exists \( \alpha_0 > 0 \) and a non-negative number \( N \) such that \( -N \leq w_\alpha(i) \) for every \( i \) and \( 0 \leq \alpha \leq \alpha_0 \).

- **SEN6**: For each state \( i \), the expected single stage discounted cost \( f_\alpha(i, \mu) \) is a lower semi-continuous (lsc) function on the product space \([0, \infty) \times A\). Note that \( f_0(i, \mu) \) is the un-discounted single stage cost.
• **SEN7**: For all states \(i\) and \(j\), the function \(L_{ij}(\alpha, \mu) = P_{ij}(\mu) \int_{t=0}^{\infty} e^{-\alpha t} e^{-\nu t} dt\) is a lsc function on the product space \([0, \infty) \times A\).

• **SEN8**: Assume that \(\alpha_n\) is a sequence of discount factors converging to 0 with the property that \(\pi_{\alpha_n}\), the associated sequence of \(\alpha\)-discount optimal policies, converge to a stationary policy \(\pi\). Then for each state \(i\), \(\lim \inf \tau(i, \pi_{\alpha_n}) \leq \tau(i, \pi)\).

**Proof of Proposition 2.2**: We begin by verifying the SEN assumptions. Since the uniformizing rate is strictly positive and finite, assumptions **SEN1** and **SEN2** hold. It was shown in Proposition 2.1 that under the stability condition (2.5), there exists a stationary policy that induces an ergodic Markov process with finite long-run average expected cost. Thus, the hypotheses of Lemma 2 of [21] hold; validating **SEN3** and **SEN4** assumptions. Let \(\bar{s} = \arg \min_{s \in \mathcal{S}} \{v_{\alpha}(0, s)\}\) and define the distinguished state as \(O = (0, \bar{s})\). It follows from Proposition 3.1 that for any \(\alpha > 0\), \(v_{\alpha}(0, s) \leq v_{\alpha}(n, s)\) for all \((n, s) \in \mathbb{X}\). Therefore, \(w_{\alpha}(n, s) \geq 0\) and **SEN5** is satisfied. **SEN6** holds since for each \((n, s) \in \mathbb{X}\), \(f_{\alpha}((n, s), \mu) = (c(\mu) + h(n))/(\alpha + \nu)\) is a continuous function on \([0, \infty) \times A\).

There is no decision to be made when \(n = 0\). Fix \(n \geq 1\) and \(s \in \{1, 2, \ldots, L\}\) and note,

\[
L_{(n,s),(n',s')}^{\alpha,\mu} = \begin{cases} \lambda_s s' \alpha + \nu & \text{if } n' = n+1, s' = s, \\ \mu(\alpha + \nu) s' & \text{if } n' = n-1, s' = s, \\ Q_{ss'} ^{\alpha + \nu} & \text{if } n' = n, s' \in \mathcal{S}, \\ 0 & \text{otherwise.} \end{cases}
\] (6.19)

So the function \(L_{x,x'}^{\alpha,\mu}\) is jointly continuous in \(\alpha\) and \(\mu\) for each \(x, x' \in \mathbb{X}\), therefore **SEN7** holds. Finally, since for any policy \(\pi\), \(\tau(i, \pi) = 1/\nu\), **SEN8** holds.

**SEN1,SEN3,SEN6** and **SEN7** are required in Theorem 11 of [21] to prove the first result. Since the holding costs are assumed to be linear, the hypotheses of Proposition 4 of [21] are satisfied. The last two results follow from Theorem 12 (and its proof) of [21].