Stabilizing a Queue Subject to Action-Dependent Server Performance

Michael Lin, Richard J. La and Nuno C. Martins

Abstract—We consider a discrete-time system comprising an unbounded queue that logs tasks, a non-preemptive server, and a stationary non-work-conserving scheduler. Time is uniformly divided into intervals we call epochs, within which new tasks arrive according to a Bernoulli process. At the start of each epoch, the server is either busy working on a task or is available. In the latter case, the scheduler either assigns a new task to the server or allows a rest epoch. In addition to the aforementioned availability state, we assume that the server has an additional action-dependent state that takes values in a finite set of positive integers. The action-dependent state, which evolves according to a given Markov decision process whose input is the action chosen by the scheduler, is non-increasing during rest periods, and is non-decreasing otherwise. The action-dependent state determines the probability that a task can be completed within an epoch. The scheduler has access to the queue size and the entire state of the server. In this article, we show that stability, whenever viable, can be achieved by a simple scheduler whose decisions are based on the availability state, a threshold applied to the action-dependent state, and a flag that indicates when the queue is empty. The supremum of the long-term service rates achievable by the system, subject to the requirement that the queue remains stable, can be determined by a finite search.

I. INTRODUCTION

Recent advances in information and communication networks and sensor technologies brought a wide range of new applications to reality. These include wireless sensor networks in which sensors equipped with communication modules are powered by renewable energy, such as solar energy and geothermal energy, as well as emerging technologies, such as unmanned aerial vehicles (UAVs). In many cases, the performance of the component systems as well as that of the overall system can be analyzed with the help of a suitable queueing model.

Unfortunately, most existing queueing models are inadequate because, in many of these new applications, the performance of the servers (e.g., a human supervisor monitoring and controlling multiple UAVs) is not time-invariant and typically depends on the history of past workload or the types of performed tasks. As a result, little is known about the performance and stability of such systems and, hence, new methodologies and theory are needed.

In this study, we first propose a new framework for studying the stability of systems in which the efficiency of servers is time-varying and is dependent on their past utilization. Using the proposed framework, we then study the problem of designing a task scheduling policy with simple structure, which is optimal in that it keeps the task queue stable whenever doing so is possible using some policy.

To this end, we consider a queueing system comprising the following three components:

- A first-in first-out unbounded queue registers a new task whenever it arrives and removes it as soon as work on it is completed. It has an internal state, its queue size, which indicates the number of uncompleted tasks in the queue.
- The server performs the work required by each task assigned to it. It has an internal state with two components. The first is the availability state, which indicates whether the server is available to start a new task or is busy. We assume that the server is non-preemptive, which in our context means that the server gets busy when it starts work on a new task, and it becomes available again only after the task is completed. The second component of the state is termed action-dependent and takes values in a finite set of positive integers, which represent parameters that affect the performance of the server. More specifically, we assume that the action-dependent state determines the probability that, within a given time-period, the server can complete a task. Hence, a decrease in performance causes an increase in the expected time needed to service a task. Such an action-dependent state could, for instance, represent the battery charge level of an energy harvesting module that powers the server or the status of arousal or fatigue of a human operator that assists the server or supervises the work.
- The scheduler has access to the queue size and the entire state of the server. When the server is available and the queue is not empty, the scheduler decides whether to assign a new task or to allow for a rest period. Our formulation admits non-work-conserving policies whereby the scheduler may choose to assign rest periods even when the queue is not empty. This allows the server to rest as a way to steer the action-dependent state towards a range that can deliver better long-term performance.

We adopt a stochastic discrete-time framework in which time is uniformly partitioned into epochs, within which new tasks arrive according to a Bernoulli process. The probability of arrival per epoch is termed arrival rate\(^1\). We constrain our analysis to stationary schedulers characterized by policies that are invariant under epoch shifts. We discuss our assumptions and provide a detailed description of our framework in Section II-B.1.

The authors are with the Department of Electrical & Computer Engineering and the Institute for Systems Research, the University of Maryland, College Park, MD 20742. Email: {mlin1025, hyongla, nmartins}@umd.edu.

This work is supported in part by AFOSR Grant FA95501510367 and NSF Grant ECCS 1446785.

\(^1\)Notice that, unlike the nomenclature we adopt here, arrival rate is commonly used in the context of Poisson arrival processes. This distinction is explained in detail in Section II-B.1.
A. Main Problems

The following are the main challenges addressed in this article:

Main Problems:

P-i) An arrival rate is qualified as stabilizable when there is a scheduler that stabilizes the queue. Given a server, we seek to compute the supremum of the set of stabilizable arrival rates.

P-ii) We seek to propose schedulers that have a simple structure and are guaranteed to stabilize the queue for any stabilizable arrival rate.

Notice that, as alluded to above in the scheduler description, we allow non-work-conserving policies. This means that, in addressing Problem P-i), we must allow policies that are a function not only of the queue size, but also of the action-dependent and availability states of the server. The design process for good policies is complicated by the fact that they are a function of these states with intricate dependence, illustrating the importance of addressing Problem P-ii).

B. Preview of Results

The following are our main results and their correspondence with Problems P-i) and P-ii).

R-i) In Theorem 3.1, we show that the supremum mentioned in P-i) can be achieved by some function whose domain is the finite set of action-dependent states. The fact that such a quantity can be determined by a finite search is not evident because the queue size is unbounded.

R-ii) As we state in Theorem 5.2, given a server, there is a threshold policy that stabilizes the queue for any stabilizable arrival rate. The threshold policy assigns a new task only when the server is available, the queue is not empty, and the action-dependent state is less than a threshold chosen to be the value (found by a finite search) at which the maximum referred to in R-i) is attained. This is our answer to Problem P-ii).

From this discussion we conclude that, to the extent that the stability of the system is concerned, we can focus solely on threshold policies outlined in R-ii). We emphasize the fact that, as we discuss in Remark 2 of Section III.2, Theorem 5.2 is valid even when the performance of the server is not monotonic with respect to the action-dependent state.

C. Related Literature

There exists a large volume of literature on queueing systems with time-varying service rate. In particular, with the rapid growth of wireless networks, there has been much interest in understanding and designing efficient scheduling policies with time-varying channel conditions that affect the probability of successful transmissions [2], [3], [6].

There are different formulations researchers considered for designing sound scheduling policies. For example, some adopted an optimization framework in which the objective is to maximize the aggregate utility of flows/users (e.g., [1], [2], [9], [11]). This approach allows the designers to carry out the trade-off between the aggregate throughout and fairness among flows/users.

Another line of research, which is more relevant to our study, focused on designing throughput optimal schedulers that can stabilize the system for any arrival rate (vector) that lies in the stability region (e.g., [3], [14], [19]). However, there is a major difference between these studies and our study.

In wireless networks, channel conditions and probability of successful transmission/decoding vary independently of the scheduling decisions chosen by the scheduler. In other words, the scheduler (together with the physical layer schemes) attempts to cope with or take advantage of time-varying channel conditions, which are beyond the control of resource managers and are not affected by scheduling decisions. In our study, on the other hand, the probability of successfully completing a task within an epoch depends on the past history of scheduling decisions. Consequently, the current scheduling decision affects the future efficiency of the server.

Another area that is closely related to our study is task scheduling for human operators/servers. The performance and management of human operators and servers (e.g., bank tellers, toll collectors, doctors, nurses, emergency dispatchers) has been the subject of many studies in the past, e.g., [5], [7], [18], [20]. Recently, with rapid advances in information and sensor technologies, human supervisory control, which requires processing a large amount of information in a short period, potentially causing information overload, became an active research area [12], [17].

As human supervisors play a critical role in the systems (e.g., supervisory control and data acquisition (SCADA)), there is a resurging interest in understanding and modeling the performance of humans under varying settings. Although this is still an active research area, it is well documented that the performance of humans depends on many factors, including arousal and perceived workload [7], [4], [10], [18]. For example, the well-known Yerkes-Dodson law suggests that moderate levels of arousal are beneficial, leading to the inverted-U model [20].

In closely related studies, Savla and Frazzoli [15], [16] investigated the problem of designing a task release control policy. They assumed periodic task arrivals and modeled the dynamics of server utilization, which determines the service time of the server, using a differential equation; the server utilization increases when the server is busy and decreases when it is idle. They showed that, when all tasks bring identical workload, a policy that allows a new task to be released to the server only when its utilization is below a suitably chosen threshold, is maximally stabilizing [16, Theorems III.1 and III.2]. In addition, somewhat surprisingly, they proved that when task workloads are modeled using independent and identically distributed (i.i.d) random variables, the maximum achievable throughput increases compared to the homogeneous workload cases.

Paper Structure: A stochastic discrete-time model is de-
in Section III. In it we also introduce notation, key concepts and propose a Markov Decision Process (MDP) framework that is amenable to performance analysis and optimization. Our main results are discussed in Section III while Section IV describes an auxiliary MDP that is central to our analysis. The proofs of our results are presented in Section V and Section VI provides concluding remarks.

II. STOCHASTIC DISCRETE-TIME FRAMEWORK

In the following subsection, we describe a discrete-time framework that follows from assumptions on when the states of the queue and the server are updated and how actions are decided. In doing so, we will also introduce the notation used to represent these discrete-time processes. A probabilistic description that leads to a tractable MDP formulation is deferred to Section II-B.

A. State Updates and Scheduling Decisions: Timing and Notation

We consider an infinite horizon problem in which the physical (continuous) time set is \( \mathbb{R}_+ \), which we partition uniformly into half-open intervals of positive duration \( \Delta \) as follows:

\[
\mathbb{R}_+ = \bigcup_{k=0}^{\infty} \left[ k\Delta, (k+1)\Delta \right)
\]

Each interval is called an epoch, and epoch \( k \) refers to \([k\Delta, (k+1)\Delta)\). Our formulation and results are valid regardless of the epoch duration \( \Delta \). We reserve \( t \) to denote continuous time, and \( k \) is the discrete-time index we use to represent epochs.

Each epoch is subdivided into three half-open subintervals denoted by stages \( S_1, S_2 \) and \( S_3 \) (see Fig. 1). As we explain below, stages \( S_1 \) and \( S_2 \) are allocated for basic operations of record keeping, updates and scheduling decisions. Although, in practice, the duration of these stages is a negligible fraction of \( \Delta \), we discuss them here in detail to clarify the causality relationships among states and actions. We also introduce notation used to describe certain key discrete-time processes that are indexed with respect to epoch number.

1) Stage \( S_1 \): The following updates take place during stage \( S_1 \) of epoch \( k + 1 \):

We assume that at most one new task arrives during each epoch. In addition, the server can work on at most one task at any given time and, within each epoch, the scheduler assigns at most one task to the server. However, these assumptions are not critical to our main findings and can be relaxed to allow more general arrival distributions and batch processing.

The queue size at time \( t = k\Delta \) is denoted by \( Q_k \), and is updated as follows:

- If, during epoch \( k \), a new task arrives and no task is completed then \( Q_{k+1} = Q_k + 1 \).
- If, during epoch \( k \), either (i) a new task arrives and a task is completed or (ii) no new task arrives and no task is completed, then \( Q_{k+1} = Q_k \).
- If, during epoch \( k \), no new task arrives and a task is completed then \( Q_{k+1} = Q_k - 1 \).

The availability state of the server at time \( t = k\Delta \) is denoted by \( W_k \) and takes values in

\[
\mathbb{W} \overset{\text{def}}{=} \{A, B\}.
\]

We use \( W_k = B \) to indicate that the server is busy working on a task at time \( t = k\Delta \). If it is available at time \( t = k\Delta \), then \( W_k = A \). The update mechanism for \( W_k \) is as follows:

- If \( W_k = A \), then \( W_{k+1} = A \) when either no new task was assigned during epoch \( k \), or a new task was assigned and completed during epoch \( k \). If \( W_k = A \) and a new task is assigned during epoch \( k \) which is not completed until \( t = (k+1)\Delta \), then \( W_{k+1} = B \).
- If \( W_k = B \) and the server completes the task by time \( t = (k+1)\Delta \), then \( W_{k+1} = A \). Otherwise, \( W_{k+1} = B \).

We use \( S_k \) to denote the action-dependent state at time \( t = k\Delta \), and we assume that it takes values in

\[
\mathbb{S} \overset{\text{def}}{=} \{1, \ldots, n_x\}.
\]

The action-dependent state is non-decreasing while the server is working and is non-increasing when it is idle. In Section II-B, we describe an MDP that specifies probabilistically how \( S_k \) transitions to \( S_{k+1} \), conditioned on whether the server worked or rested during epoch \( k \).

Without loss of generality, we assume that \( Q_k, W_k \) and \( S_k \) are initialized as follows:

\[
Q_0 = 0, \quad W_0 = A, \quad S_0 = 1
\]

The overall state of the server is represented compactly by \( Y_k \), which takes values in \( \mathbb{Y} \), defined as follows:

\[
Y_k \overset{\text{def}}{=} (S_k, W_k), \quad \mathbb{Y} \overset{\text{def}}{=} \mathbb{S} \times \mathbb{W}
\]

In a like manner, we define the overall state for the MDP taking values in \( \mathbb{X} \) as follows:

\[
X_k \overset{\text{def}}{=} (Y_k, Q_k), \quad \mathbb{X} \overset{\text{def}}{=} \mathbb{S} \times \left((\mathbb{W} \times \mathbb{N}) \setminus (\mathbb{B}, 0)\right)
\]

From the definition of \( \mathbb{X} \), it follows that when the queue is empty, there is no task for the server to work on and, hence, it cannot be busy.
2) **Stage S2:** It is during stage S2 of epoch \( k \) that the scheduler issues a decision based on \( X_k \); let \( \mathcal{A} \) be the set of possible actions that the scheduler can request from the server, where \( \mathcal{R} \) and \( \mathcal{W} \) represent ‘rest’ and ‘work’, respectively. The assumption that the server is non-preemptive and the fact that no new tasks can be assigned when the queue is empty, lead to the following set of available actions for each possible state \( x = (s, w, q) \) in \( X \):

\[
\mathcal{A}_x = \begin{cases} 
\mathcal{R} & \text{if } q = 0, \text{ (impose ‘rest’ when queue is empty)} \\
\mathcal{W} & \text{if } q > 0 \text{ and } w = B, \text{ (non-preemptive server)} \\
\mathcal{A} & \text{otherwise.} 
\end{cases}
\]

We denote the action chosen by the adopted scheduling policy at epoch \( k \) by \( A_k \), which takes values in \( \mathcal{A}_{X_k} \). As we discuss in Section II-C, we focus on the design of stationary policies that determine \( A_k \) as a function of \( X_k \).

3) **Stage S3:** A task can arrive at any time during each epoch, but we assume that work on a new task can be assigned to the server only at the beginning of stage S3. More specifically, the scheduler acts as follows:

- If \( W_k = \mathcal{A} \) and \( A_k = \mathcal{W} \), then the server starts working on a new task at the head of the queue when stage S3 of epoch \( k \) begins.
- When \( W_k = \mathcal{A} \), the scheduler can also select \( A_k = \mathcal{R} \) to signal that no work will be performed by the server during the remainder of epoch. Once this ‘rest’ decision is made, a new task can be assigned no earlier than the beginning of stage S3 of epoch \( k + 1 \). Since the scheduler is non-work-conserving, it may decide to assign such ‘rest’ periods as a way to possibly reduce \( S_{k+1} \) and to improve future performance.
- If \( W_k = \mathcal{B} \), the server was still working on a task at time \( t = k\Delta \). In this case, because the server is non-preemptive, the scheduler picks \( A_k = \mathcal{W} \) to indicate that work on the current task is ongoing and must continue until it is completed and no new task can be assigned during epoch \( k \).

**B. State Updates and Scheduling Decisions: Probabilistic Model**

Based on the formulation outlined in Section II-A, we proceed to describe a discrete-time MDP that models how the states of the server and queue evolve over time for any given scheduling policy.

1) **Arrival Process:** We assume that tasks arrive according to a Bernoulli process \( \{B_k \mid k \in \mathbb{N}\} \). The probability of arrival for each epoch ( \( P(B_k = 1) \) ) is called the arrival rate and is denoted by \( \lambda \), which is assumed to belong to \((0, 1)\). Although we assume Bernoulli arrivals to simplify our analysis and discussion, more general arrival distributions (e.g., Poisson distributions) can be handled only with minor changes as it will be clear.

Notice that, as we discuss in Remark 1 below, our nomenclature for \( \lambda \) should not be confused with the standard definition of arrival rate for Poisson arrivals. Since our results are valid irrespective of \( \Delta \), including when it is arbitrarily small, the remark also gives a sound justification for our adoption of the Bernoulli arrival model by viewing it as a discrete-time approximation of the widely used Poisson arrival model.

**Remark 1:** It is a well-known fact that, as \( \Delta \) tends to zero, a Poisson process in continuous time \( t \), with arrival rate \( \lambda \), is arbitrarily well approximated by \( B_{\lfloor t/\Delta \rfloor} \) with \( \lambda = \Delta \lambda \).

2) **Action-Dependent Server Performance:** In our formulation, the efficiency or performance of the server during an epoch is modeled with the help of a *service rate function* \( \mu : S \to (0, 1) \). More specifically, if the server works on a task during epoch \( k \), the probability that it completes the task by the end of the epoch is \( \mu(S_k) \). This holds irrespective of whether the task is newly assigned or inherited as ongoing work from a previous epoch. Thus, the service rate function \( \mu \) quantifies the effect of the action-dependent state on the performance of the server. The results presented throughout this article are valid for any choice of \( \mu \) with codomain \((0, 1)\).

3) **Dynamics of the Action-Dependent State:** We assume that \( S_{k+1} \) is equal to either \( S_k \) or \( S_k + 1 \) when \( A_k \) is \( \mathcal{W} \) and \( S_{k+1} \) is either \( S_k \) or \( S_k - 1 \) if \( A_k \) is \( \mathcal{R} \). This is modeled by the following transition probabilities specified for every \( s \) and \( s' \) in \( S \).

\[
P_{S_{k+1} | S_k, A_k}(s' | s, \mathcal{W}) = \begin{cases} 
\rho_{s,s+1} & \text{if } s < n_x \text{ and } s' = s + 1,
1 - \rho_{s,s+1} & \text{if } s < n_x \text{ and } s' = s,
1 & \text{if } s = n_x \text{ and } s' = n_x,
0 & \text{otherwise},
\end{cases}
\]

\[
P_{S_{k+1} | S_k, A_k}(s' | s, \mathcal{R}) = \begin{cases} 
\rho_{s,s-1} & \text{if } s > 1 \text{ and } s' = s - 1,
1 - \rho_{s,s-1} & \text{if } s > 1 \text{ and } s' = s,
1 & \text{if } s = 1 \text{ and } s' = 1,
0 & \text{otherwise}.
\end{cases}
\]

where the parameters \( \rho_{s,s'} \), which take values in \((0, 1)\), model the likelihood that the action-dependent state will transition to a greater or lesser value, depending on whether the action is \( \mathcal{W} \) or \( \mathcal{R} \), respectively.

4) **Transition probabilities for \( X_k \):** We consider that \( S_{k+1} \) is independent of \( (W_{k+1}, Q_{k+1}) \) when conditioned on \((X_k, A_k)\). Under this assumption, the transition probabilities for \( X_k \) can be written as follows:

\[
P_{X_{k+1} | X_k, A_k}(x' | x, a) = P_{S_{k+1} | S_k, A_k}(s' | s, a) \times P_{W_{k+1} | Q_{k+1}, X_k, A_k}(w', q' | x, a)
\]

\[
P_{X_{k+1} | X_k, A_k}(x' | s, a) = \begin{cases} 
\rho_{s,s+1} & \text{if } s < n_x \text{ and } s' = s + 1,
1 - \rho_{s,s+1} & \text{if } s < n_x \text{ and } s' = s,
1 & \text{if } s = n_x \text{ and } s' = n_x,
0 & \text{otherwise},
\end{cases}
\]

\[
P_{X_{k+1} | X_k, A_k}(x' | x, a) = \begin{cases} 
\rho_{s,s-1} & \text{if } s > 1 \text{ and } s' = s - 1,
1 - \rho_{s,s-1} & \text{if } s > 1 \text{ and } s' = s,
1 & \text{if } s = 1 \text{ and } s' = 1,
0 & \text{otherwise}.
\end{cases}
\]

This assumption is introduced to simplify the exposition. However, more general scenarios in which the probability of task completion within an epoch depends on the total service received by the task prior to epoch \( k \) can be handled by extending the state space and explicitly modeling the total service received by the task in service.
for every $x, x' \in \mathcal{X}$ and $a \in \mathcal{A}_x$. 

We assume that, within each epoch $k$, the events that (a) there is a new task arrival during the epoch and (b) a task being serviced during the epoch is completed by the end of the epoch are independent when conditioned on $X_k$ and $\{A_k = \mathcal{W}\}$. Hence, the transition probability $P_{W_{k+1}, Q_{k+1}}|X_k, A_k$ in (3) is given by the following:

$$
P_{W_{k+1}, Q_{k+1}}|X_k, A_k(x', q') \mid x, \mathcal{W} = \begin{cases} 
\mu(s_k) & \text{if } w' = A \text{ and } q' = q, \\
\mu(s_k) (1 - \lambda) & \text{if } w' = A \text{ and } q' = q - 1, \\
(1 - \mu(s_k)) \lambda & \text{if } w' = B \text{ and } q' = q + 1, \\
(1 - \mu(s_k))(1 - \lambda) & \text{if } w' = B \text{ and } q' = q, \\
0 & \text{otherwise,}
\end{cases}$$

$$
P_{W_{k+1}, Q_{k+1}}|X_k, A_k(x', q') \mid x, \mathcal{W} = \begin{cases} 
\lambda & \text{if } w' = A \text{ and } q' = q, \\
1 - \lambda & \text{if } w' = A \text{ and } q' = q, \\
0 & \text{otherwise.}
\end{cases}
$$

**Definition 1:** (MDP $X$) The MDP with input $A_k$ and state $X_k$, which at this point is completely defined, is denoted by $X$. 

Table I summarizes the notation for MDP $X$.

| $\mathcal{S}$ | set of action-dependent states $\{1, \ldots, n_s\}$ |
| $\mathcal{W} \equiv \{A, B\}$ | server availability ($A = \text{available}, B = \text{busy}$) |
| $W_k$ | server availability at epoch $k$ (takes values in $\mathcal{W}$) |
| $Y_k \equiv (S_k, W_k)$ | server state at epoch $k$ (takes values in $Y$) |
| $Q_k$ | queue size at epoch $k$ (takes values in $\mathcal{Y}$) |
| $X_k \equiv (Y_k, Q_k)$ | system state at epoch $k$ (takes values in $X$) |
| $A_k \equiv \{\mathcal{A}, \mathcal{W}\}$ | possible actions (\mathcal{A} = \text{rest}, \mathcal{W} = \text{work}) |
| $M_k$ | MDP whose state is $X_k$ at epoch $k \in \mathbb{N}$ |
| $\mathcal{A}_x$ | set of actions available at a given state $x \in \mathcal{X}$ |
| $\mathcal{A}_k$ | action chosen at epoch $k$. |
| PMF | probability mass function |

### C. Evolution of the system state under a stationary policy

We start by defining the class of policies that we consider throughout the paper.

**Definition 2:** A stationary randomized policy is specified by a mapping $\theta : X \rightarrow [0, 1]$ that determines the probability that the server is assigned to work on a task or rest, as a function of the system state, according to

$$
P_{A_k \mid X_k, \ldots, X_0}(\mathcal{W} \mid x_k, \ldots, x_0) = \theta(x_k) \quad \text{and} \quad P_{A_k \mid X_k, \ldots, X_0}(\mathcal{A} \mid x_k, \ldots, x_0) = 1 - \theta(x_k).
$$

**Definition 3:** The set of admissible stationary randomized policies satisfying (1) is denoted by $\Theta_R$.

**Convention:** We adopt the convention that, unless stated otherwise, a positive arrival rate $\lambda$ is pre-selected and fixed throughout the paper. Although the statistical properties of $X$ and associated quantities subject to a given policy depend on $\lambda$, we simplify our notation by not labeling them with $\lambda$.

From (3) - (10), we conclude that $X$ subject to a policy $\theta$ in $\Theta_R$ evolves according to a time-homogeneous Markov chain (MC), which we denote by $X^\theta = \{X_k^\theta \mid k \in \mathbb{N}\}$. Also, provided that it is clear from the context, we refer to $X^\theta$ as the system.

The following is the notion of system stability we adopt in our study.

**Definition 4 (System stability):** For a given policy $\theta$ in $\Theta_R$, the system $X^\theta$ is stable if it satisfies the following properties:

**i.** There exists at least one recurrent communicating class.

**ii.** All recurrent communicating classes are positive recurrent.

**iii.** The number of transient states is finite.

We find it convenient to define $\Theta_R(\lambda)$ to be the set of randomized policies in $\Theta_R$, which stabilize the system for the fixed arrival rate $\lambda$.

Before we proceed, let us point out a useful fact under any stabilizing policy $\theta$ in $\Theta_R(\lambda)$.

**Lemma 1:** A stable system $X^\theta$ has a unique positive recurrent communicating class, which is aperiodic. Therefore, there is a unique stationary probability mass function (PMF) for $X^\theta$.

**Proof:** Please see Appendix II for a proof.

**Definition 5:** Given a fixed arrival rate $\lambda > 0$ and a stabilizing policy $\theta$ in $\Theta_R(\lambda)$, we denote the unique stationary PMF and positive recurrent communicating class of $X^\theta$ by $\pi^\theta = (\pi^\theta(x); x \in \mathcal{X})$ and $C_{\pi^\theta}$, respectively.

### III. MAIN RESULT

Our main results are Theorems 3.1 and 3.2 where we state that a soon-to-be defined quantity $\lambda^*$, which can be computed efficiently, is the least upper bound of all arrival rates for which there exists a stabilizing policy in $\Theta_R$ (see Definition 4). The theorems also assert that, for any arrival rate $\lambda$ less than $\lambda^*$, there is a stabilizing deterministic threshold policy in $\Theta_R$ with the following structure:

$$
\theta^\tau(s, w, q) \overset{\text{def}}{=} \begin{cases} 
\phi^\tau(s, w) & \text{if } q > 0, \\
0 & \text{otherwise}
\end{cases}
$$

where $\tau$ lies in $\mathcal{S} \cup \{n_s + 1\}$, and $\phi^\tau$ is a threshold policy that acts as follows:

$$
\phi^\tau(s, w) \overset{\text{def}}{=} \begin{cases} 
0 & \text{if } s \geq \tau \text{ and } w = A, \\
1 & \text{otherwise}.
\end{cases}
$$

Notice that, when the server is available and the queue is not empty, $\theta^\tau$ assigns a new task only if $s$ is less than the threshold $\tau$ and lets the server rest otherwise.

In Section IV, we introduce an auxiliary MDP with finite state space $\mathcal{Y}$, which can be viewed as the server state in
X when the queue size $Q_k$ is always positive. First, using the fact that $Y$ is finite, we demonstrate that, for every $\tau$ in $\mathbb{S} \cup \{n_4 + 1\}$, the auxiliary MDP subject to $\phi_\tau$ has a unique stationary PMF, which we denote by $\pi^\tau$. Then, we show that, for any stable system $X^\theta$ under some policy $\theta$ in $\Theta_R$, we can find a threshold policy $\phi_\tau$ for the auxiliary MDP, which achieves the same long-term departure rate of completed tasks as in $X^\theta$. As a result, the maximum long-term departure rate of completed tasks in the auxiliary MDP among all threshold policies $\phi_\tau$ with $\tau$ in $\mathbb{S} \cup \{n_4 + 1\}$ serves as an upper bound on the arrival rate $\lambda$ for which we can hope to find a stabilizing policy $\theta$ in $\Theta_R$.

Making use of this observation, we define the following important quantity:

$$\lambda^* \overset{\text{def}}{=} \max_{\tau \in \mathbb{S} \cup \{n_4 + 1\}} \left( \sum_{(s, w) \in Y} \pi^\tau(s, w) \phi_\tau(s, w) \mu(s) \right)$$  \hfill (7)

From the definition of the stationary PMF $\pi^\tau$, $\lambda^*$ can be interpreted as the maximum long-term departure rate of completed tasks under any threshold policy of the form in (5), assuming that the queue is always non-empty.

The following are the main results of this paper.

**Theorem 3.1:** (Necessity) If, for a given arrival rate $\lambda$, there exists a stabilizing policy in $\Theta_R$, then $\lambda \leq \lambda^*$.

A proof of Theorem 3.1 is provided in Section V-A.

**Theorem 3.2:** (Sufficiency) Let $\tau^*$ be a maximizer of (7). If the arrival rate $\lambda$ is strictly less than $\lambda^*$, then $\theta_{\tau^*}$ stabilizes the system.

Please see Section V-B for a proof of Theorem 3.2.

**Remark 2:** The following important observations are direct consequences of (7) and Theorems 3.1 and 3.2:

- The computation of $\lambda^*$ in (7) along with a maximizing threshold $\tau^*$ relies on a finite search that can be carried out efficiently.
- The theorems are valid for any choice of service rate function $\mu$ that takes values in $(0, 1)$. In particular, $\mu$ could be multi-modal, increasing or decreasing.
- The search that yields $\lambda^*$ and an associated $\tau^*$ does not require knowledge of $\lambda$.

We point out two key differences between our study and [15], [16]. The model employed by Savla and Frazzoli assumes that the service time function is convex, which is analogous to our service rate function being unimodal. In addition, a threshold policy is proved to be maximally stabilizing only for identical task workload. In our study, however, we do not impose any assumption on the service rate function, and the workloads of tasks are modeled using i.i.d. random variables.\(^3\)

### IV. An Auxiliary MDP \(\overline{Y}\)

In this section, we describe an auxiliary MDP whose state takes values in $\overline{Y}$ and is obtained from $X$ by artificially removing the queue-length component. We denote this auxiliary MDP by $\overline{Y}$ and its state at epoch $k$ by $\overline{Y}_k = (\overline{s}_k, \overline{w}_k)$ in order to emphasize that it takes values in $\overline{Y}$. The action chosen at epoch $k$ is denoted by $\overline{A}_k$. We use the overline to denote the auxiliary MDP and any other variables associated with it, in order to distinguish them from those of the server state in $X$.

As it will be clear, we can view $\overline{Y}$ as the server state of the original MDP $X$ for which infinitely many tasks are waiting in the queue at the beginning, i.e., $Q_0 = \infty$. As a result, there is always a task waiting for service when the server becomes available.

The reason for introducing $\overline{Y}$ is the following: (i) $\overline{Y}$ is finite and, hence, $\overline{Y}$ is easier to analyze than $X$, and (ii) we can establish a relation between $X$ and $\overline{Y}$, which allows us to prove the main results in the previous section by studying $\overline{Y}$ instead of $X$. This simplifies the proofs of the theorems in the previous section.

- **Admissible action sets:** As the queue size is no longer a component of the state of $\overline{Y}$, we eliminate the dependence of admissible action sets on $q$, which was explicitly specified in (1) for MDP $X$, while still ensuring that the server is non-preemptive. More specifically, the set of admissible actions at each element $\overline{Y} = (\overline{s}, \overline{w})$ of $\overline{Y}$ is given by

\[
\overline{A}_k = \begin{cases} \{\overline{w}\} & \text{if } \overline{w} = \overline{B}, \quad \text{(non-preemptive server)} \\ \overline{A} & \text{if } \overline{w} = \overline{A}. \end{cases}
\]

Consequently, for any given realization of the current state $\overline{Y}_k = (\overline{s}_k, \overline{w}_k)$, $\overline{A}_k$ is required to take values in $\overline{A}_k$.

- **Transition probabilities:** We define the transition probabilities that specify $\overline{Y}$, as follows:

\[
P_{\overline{Y}_{k+1}|\overline{Y}_k, \overline{A}_k}(\overline{Y}' \mid \overline{Y}, \overline{A}) \overset{\text{def}}{=} P_{\overline{s}_{k+1}|\overline{s}_k, \overline{A}_k}(\overline{s}' \mid \overline{s}, \overline{A}) \times P_{\overline{w}_{k+1}|\overline{w}_k, \overline{A}_k}(\overline{w}' \mid \overline{w}, \overline{A})
\]

where $\overline{Y}$ and $\overline{Y}'$ are in $\overline{Y}$, and $\overline{A}$ is in $\overline{A}_k$. Subject to these action constraints, the right-hand terms of (9) are defined, in connection with $X$, as follows:

\[
P_{\overline{s}_{k+1}|\overline{s}_k, \overline{A}_k}(\overline{s}' \mid \overline{s}, \overline{A}) \overset{\text{def}}{=} P_{S_{k+1}|S_k, A_k}(s' \mid s, A)
\]

\[
P_{\overline{w}_{k+1}|\overline{w}_k, \overline{A}_k}(\overline{w}' \mid \overline{w}, \overline{A}) \overset{\text{def}}{=} \begin{cases} \mu(\overline{s}) & \text{if } \overline{w}' = \overline{A} \\ 1 - \mu(\overline{s}) & \text{if } \overline{w}' = \overline{B} \end{cases}
\]

\[
P_{\overline{w}_{k+1}|\overline{w}_k, \overline{A}_k}(\overline{w}' \mid \overline{w}, \overline{A}) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } \overline{w}' = \overline{A} \\ 0 & \text{if } \overline{w}' = \overline{B} \end{cases}
\]

\(^{3}\)To be more precise, our assumptions correspond to the case with exponentially distributed workloads. However, as mentioned earlier, this assumption can be relaxed to allow more general workload distributions.
A relation between the transition probabilities of $X$ and $Y$: From the definition above and [4], we can deduce the following equality: for all $q \geq 1$,

$$P_{\bar{Y}, k+1} = \sum_{q' = 0}^{\infty} P_{W, k+1, Q, k+1}(\bar{Y} | \bar{Y}, q'),$$

(12)

which holds for any $\bar{Y}$ in $\mathcal{Y}$ and $\bar{Y}$ in $\mathcal{Y}$. Notice that the right-hand side (RHS) of (12) does not change when we vary $q$ across the positive integers. From this, in conjunction with [3], [9] and [10], we have, for all $q \geq 1$,

$$P_{\bar{Y}, k+1} = \sum_{q' = 0}^{\infty} P_{X, k+1, A, k}(\bar{Y} | \bar{Y}, q'),$$

(13)

The equality in (13) indicates that $P_{\bar{Y}, k+1} = \bar{Y}$ also characterizes the transition probabilities of the server state $\bar{Y}_k = (S_k, W_k)$ in $X$ when the current queue size is positive. This is consistent with our earlier viewpoint that $\bar{Y}$ can be considered the server state in $X$ initialized with infinite queue length at the beginning. We will explore this relationship in Section 5.1 where we use $\bar{Y}$ to prove Theorems 3.1 and 3.2.

A. Stationary policies and stationary PMFs of $Y$

Analogously to the MDP $X$, we only consider stationary randomized policies for $\bar{Y}$, which are defined below.

Definition 6 (Stationary randomized policies for $\bar{Y}$): We restrict our attention to stationary randomized policies acting on $\bar{Y}$, which are specified by a mapping $\phi : \mathcal{Y} \to [0, 1]$, as follows:

$$P_{\bar{Y}_k, \pi_k, \varphi_k} = \phi(\bar{Y}_k),$$

$$P_{\bar{Y}_k, \pi_k, \varphi_k} = 1 - \phi(\bar{Y}_k),$$

for every $k \in \mathbb{N}$ and $\bar{Y}_k, \pi_k, \varphi_k$ in $\mathcal{Y}$. The set of all stationary randomized policies for $\bar{Y}$ which honor [3] is defined to be $\Phi_R$.

a) Recurrent communicating classes of $\bar{Y}$: The MDP $\bar{Y}$ subject to a policy $\phi \in \Phi_R$ is a finite-state time-homogeneous MC and is denoted by $\bar{Y}^{\phi} = \{Y_k^\phi, k \in \mathbb{N}\}$. Because $\mathcal{Y}$ is finite, for any policy $\phi \in \Phi_R$, $\bar{Y}^{\phi}$ has a positive recurrent communicating class and a stationary distribution [8]. In fact, there are at most two positive recurrent communicating classes as explained below.

Define a mapping $\bar{T} : \Phi_R \to \mathbb{S} \cup \{0\}$, where

$$\bar{T}(\phi) \overset{\Delta}{=} \max(\bar{s} \in \mathbb{S} | \phi(\bar{s}, A) = 1), \quad \phi \in \Phi_R.$$  

We assume that $\bar{T}(\phi) = 0$ if the set on the RHS is empty.

Case 1. $\bar{T}(\phi, A) > 0$: First, from the definition of $\bar{T}(\phi)$, clearly all states $(\bar{s}, \bar{w})$ with $\bar{s} \geq \bar{T}(\phi)$ communicate with each other, but none of these states communicates with any other state $(\bar{s}', \bar{w}')$ with $\bar{s}' < \bar{T}(\phi)$ because $\phi(\bar{T}(\phi), A) = \phi(\bar{T}(\phi), B) = 1$. Second, because $\phi(1, A) > 0$ by assumption, all states $(\bar{s}', \bar{w}')$ with $\bar{s}' < \bar{T}(\phi)$ communicate with states $(\bar{s}, \bar{w})$ with $\bar{s} \geq \bar{T}(\phi)$. Together with the first observation, this implies that these states $(\bar{s}', \bar{w}')$ with $\bar{s}' < \bar{T}(\phi)$ are transient. Therefore, there is only one positive recurrent communicating class given by

$$\mathcal{Y}^{\phi} \overset{\Delta}{=} \{(\bar{s}, \bar{w}) \in \mathcal{Y} | \bar{s} \geq \bar{T}(\phi)\}.$$  

Case 2. $\bar{T}(\phi, A) = 0$: In this case, it is clear that $(1, A)$ is an absorbing state and forms a positive recurrent communicating class by itself. Hence, if $T(\phi) = 0$, as all other states communicate with $(1, A)$, the only positive recurrent communicating class is $\{1, A\}$ and all other states are transient. On the other hand, if $T(\phi) > 1$, for the same reason explained in the first case, $\mathcal{Y}^{\phi}$ gives rise to a second positive recurrent communicating class, and all other states $(\bar{s}', \bar{w}')$ with $\bar{s}' < \bar{T}(\phi)$, except for $(1, A)$, are transient.

In our study, we often limit our discussion to randomized policies $\phi \in \Phi_R$ with $\phi(1, A) > 0$. For this reason, for notational convenience, we define the set of randomized policies satisfying this condition by $\Phi_R^+$. The reason for this will be explained in the subsequent section.

The following proposition is an immediate consequence of the above observation.

Corollary 1: For any policy $\phi \in \Phi_R^+$, $\mathcal{Y}^{\phi}$ has a unique stationary PMF, which we denote by $\bar{\pi}^{\phi} = (\bar{\pi}^{\phi}(\bar{y}); \bar{y} \in \mathcal{Y})$.

b) Existence of a policy for $Y$ with an identical steady-state distribution of server state in $X^{\theta}$: One of key facts which we will make use of in our analysis is that, for any stabilizing policy $\theta \in \Theta_R(\lambda)$, we can find a policy $\phi \in \Phi_R^+$ which achieves the same steady-state distribution of server state. To this end, we first define, for each $\bar{y}$ in $\mathcal{Y}$,

$$\mathcal{L}_{\mathcal{Y}} \overset{\Delta}{=} \{q \in \mathbb{N} | (\bar{y}, q) \in \mathcal{X}\}.$$

Definition 7 (Policy projection map $Q$): We define a mapping $Q : \Theta_R(\lambda) \to \Phi_R$, where

$$Q(\theta) \overset{\Delta}{=} \phi^\theta, \quad \theta \in \Theta_R(\lambda),$$

with

$$\phi^\theta(\bar{y}) = \sum_{q \in \mathcal{L}_{\mathcal{Y}}} \theta(\bar{y}, q) \pi^\theta(\bar{y}, q), \quad \bar{y} \in \mathcal{Y}. \quad (15)$$

We first present a lemma that proves useful in our analysis.

Lemma 2: For every stabilizing policy $\theta \in \Theta_R(\lambda)$, we have $\phi^\theta(1, A) > 0$.

Proof: Please see Appendix III for a proof.

An obvious implication of the lemma is that $Q(\theta)$ belongs to $\Phi_R^+$ for every $\theta \in \Theta_R(\lambda)$, and there exists a unique stationary PMF for $Y^{\theta}(\mathcal{Y})$, namely $\pi^{\theta}(\mathcal{Y})$.

The following lemma shows that the steady-state distribution of the server state in $X$ under policy $\theta \in \Theta_R(\lambda)$ is identical to that of $\mathcal{Y}$ under policy $Q(\theta)$.

Lemma 3: Suppose that $\theta \in \Theta_R(\lambda)$. Then, we have

$$\pi^{\theta}(\mathcal{Y}) = \sum_{q \in \mathcal{L}_{\mathcal{Y}}} \pi^\theta(\bar{y}, q), \quad \bar{y} \in \mathcal{Y}. \quad (16)$$

Proof: A proof is provided in Appendix III.
V. PROOFS OF MAIN RESULTS

In this section, we begin with a comment on the long-term average departure rate of completed tasks when the system is stable. Then, we provide the proofs of Theorems 3.1 and 3.2 in Sections V-A and V-B respectively.

Remark 3: Recall from our discussion in Section II that, under a stabilizing policy \( \theta \) in \( \Theta_R(\lambda) \), there exists a unique stationary PMF \( \pi^\theta \). Consequently, the average number of completed tasks per epoch converges almost surely as \( k \) goes to infinity. In other words,

\[
\lim_{k \to \infty} \sum_{\tau=0}^{k-1} \mathbb{I}\{ \text{a task is completed at epoch } \tau \text{ in } X^\theta \} = \sum_{x \in X} \mu(s)\pi^\theta(x)\theta(x) \overset{\text{def}}{=} \nu^\theta \text{ with probability 1,}
\]

where \( s, w \) and \( q \) are the coordinates of \( x = (s, w, q) \). We call \( \nu^\theta \) the long-term service rate of \( \theta \) (for the given arrival rate \( \lambda > 0 \)). Moreover, because \( \theta \) is assumed to be a stabilizing policy, we have \( \nu^\theta = \lambda \).

A. A proof of Theorem 3.1

In order to prove Theorem 3.1, we make use of a similar notion of long-term service rate of \( Y^\phi \), which can be viewed in most cases as the average number of completed tasks per epoch. (Step 1) We first establish that, for every stabilizing policy \( \theta \), we can find a related policy \( \phi \) in \( \Phi_R \) whose long-term service rate equals that of \( \theta \) or, equivalently, the arrival rate \( \lambda \). (Step 2) We prove that \( \lambda^* = \lambda^{**} \) in (7) equals the maximum long-term service rate achievable by any policy in \( \Phi_R \). Together, they tell us \( \lambda \leq \lambda^* \).

- Long-term service rate of \( Y^\phi \): The long-term service rate associated with \( Y^\phi \) under policy \( \phi \) in \( \Phi_R \) is defined as follows. First, for each \( \phi \) in \( \Phi_R \), let \( \Pi(\phi) \) be the set of stationary PMFs of \( Y^\phi \). Clearly, by Corollary 1, for any \( \phi \) in \( \Phi_R \), there exists a unique stationary PMF and \( \Pi(\phi) \) is a singleton. The long-term service rate of \( \phi \) in \( \Phi_R \) is defined to be

\[
\nu^\phi \overset{\text{def}}{=} \sup_{\phi \in \Phi_R} \left( \sum_{y \in Y} \mu(s)\pi^\phi(y)\phi(y) \right). \tag{17}
\]

Recall that \( \bar{Y} \) is the pair \((\bar{\sigma}, \bar{w})\) taking values in \( \bar{Y} \).

Step 1: The following lemma illustrates that the long-term service rate achieved by \( Q(\theta) \) in \( \Phi_R^+ \) equals that of \( \theta \).

Lemma 4: Suppose that \( \theta \) is a stabilizing policy in \( \Theta_R(\lambda) \). Then, \( \nu^Q(\theta) = \nu^\theta = \lambda \).

Proof: First, note

\[
\nu^Q(\theta) \overset{(a)}{=} \sum_{y \in Y} \mu(s)\pi^Q(\theta)(y)\phi(y) \overset{(b)}{=} \sum_{y \in Y} \mu(s)\left( \sum_{q \in L} \pi^\theta(y,q)\theta(y,q) \right) \overset{(c)}{=} \sum_{x \in X} \mu(s)\pi^\theta(x)\theta(x) \overset{(d)}{=} \nu^\theta,
\]

where \( (b) \) follows from Lemma 3 and \( (c) \) results from rearranging the summations in terms of \( x = (s, w, q) \). Finally \( (a) \) and \( (d) \) hold by definition. The lemma follows from Remark 3 that \( \nu^\theta \) is equal to \( \lambda \).

Since \( Q(\theta) \) belongs to \( \Phi_R^+ \) as explained earlier, Lemma 4 implies

\[
\lambda = \nu^Q(\theta) \leq \max_{\phi \in \Phi_R^+} \nu^\phi \overset{\text{def}}{=} \lambda^{**}. \tag{18}
\]

Step 2: We shall prove that \( \lambda^* = \lambda^{**} \) in two steps. First, we establish that there is a stationary deterministic policy that achieves \( \lambda^{**} \). Then, we show that, for any stationary deterministic policy, we can find a deterministic threshold policy that achieves the same long-term service rate, thereby completing the proof of Theorem 3.1.

Let us define \( \Phi_D \) to be a subset of \( \Phi_R \), which consists only of stationary deterministic policies for \( Y \). In other words, if \( \phi \in \Phi_D \), then \( \phi(y) \in \{0, 1\} \) for all \( y \in Y \). Theorem 9.1.8 in [13, p. 451] tells us that (i) the state space is finite and (ii) the set of available actions is finite for every state, there exists a deterministic stationary optimal policy. Thus,

\[
\lambda^{**} = \max_{\phi \in \Phi_R^+} \nu^\phi = \max_{\phi \in \Phi_D} \nu^\phi. \tag{19}
\]

While the equality in (19) simplifies the computation of the maximum long-term service rate achievable by some \( \phi \) in \( \Phi_R \), it requires a search over a set of \( 2^n \) deterministic policies in the worst case. Thus, when \( n_k \) is large, it can be computationally expensive. As we show shortly, it turns out that the maximum long-term service rate on the RHS of (19) can always be achieved by a deterministic threshold policy of the form in (6).

Definition 8: Recall from (6) that, for a given \( \tau \in \mathbb{S} \cup \{n_k + 1\} \), \( \phi_{\tau} \) is the following deterministic threshold policy for \( Y \):

\[
\phi_{\tau}(y) = \begin{cases} 0 & \text{if } \exists \tau \geq \tau_0 \text{ and } w = A \\ 1 & \text{otherwise} \end{cases}
\]

The following lemma shows that, for each deterministic policy \( \phi \) satisfying \( \phi(1, A) = 1 \), there is a deterministic threshold policy with the same long-term service rate.

Lemma 5: Suppose that \( \phi \) is a policy in \( \Phi_D \) with \( \phi(1, A) = 1 \). Then, \( \nu^\phi = \nu^{\phi_{\tau}} \), where \( \tau' = T(\phi) + 1 \).

Proof: We begin with the following facts that will be utilized in the proof.

F1. The postulation that the server is non-preemptive, which we formally impose in (5), means that after the server initiates work on a task, it will be allowed to rest only after the task is completed. This implies that any policy \( \phi \) in \( \Phi_D \) satisfies \( \phi(\Sigma, B) = 1 \) for all \( \Sigma \) in \( \mathbb{S} \).

F2. From (2) and (10), we know that \( S_{k+1} \) is never less than \( S_k \) while the server is working.

From F1 and F2 stated above, we conclude that the following holds for any \( \phi \) in \( \mathbb{S} \):

\[
\phi(\sigma, A) = 1 \implies P(\Sigma_{k+1}^\phi \geq \sigma \mid \Sigma_k^\phi = \sigma) = 1 \tag{20}
\]

Here, we recall that \( \Sigma_k = (\Sigma_k^\phi, \bar{W}_k) \) represents the state of \( Y^\phi \) at epoch \( k \).
The implication in (20) leads us to the following important observation: suppose that a deterministic policy \( \phi \) in \( \Phi_D \) satisfies \( \phi(\sigma, A) = 1 \) for some \( \sigma \) greater than 1. Then, all states \((\pi, \pi')\) with \( \pi < \sigma \) are transient and, therefore,

\[
\pi^\phi(\pi, \pi') = 0 \text{ if } \pi < \sigma. \tag{21}
\]

The reason for this is that (i) because \( \phi(1, A) = 1 \), all states \((\pi, \pi')\) with \( \pi < \sigma \) communicate with every state \((\pi', \pi'')\) with \( \pi' > \sigma \), and (ii) none of the states \((\pi, \pi')\) with \( \pi' > \sigma \) communicates with any state \((\pi, \pi')\) with \( \pi < \sigma \) since \( \phi(\sigma, A) = \phi(\sigma, B) = 1 \).

**F3.** The above observation means that, given a deterministic policy \( \phi \) in \( \Phi_D \), every state \((\pi, \pi')\) with \( \pi < T(\phi) \) is transient and \( \pi^\phi(\pi, \pi') = 0 \).

**F4.** Moreover, the remaining states \((\pi, \pi')\) in \( Y^\phi \) with \( \pi < T(\phi) \) communicate with each other and their period is one (because it is possible to transition from \((T(\phi), A)\) to itself. Since \( Y^\phi \) is finite, it forms an aperiodic, positive recurrent communicating class of \( Y^\phi \).

We will complete the proof of Lemma 5 with the help of following lemma.

**Lemma 6:** Suppose that \( \phi \) and \( \tilde{\phi} \) are two deterministic policies in \( \Phi_D \) satisfying \( \phi(1, A) = \phi(1, A) = 1 \). Then,

\[
\tilde{T}(\phi) = \tilde{T}(\phi) \implies \pi^\phi = \pi^\tilde{\phi} \tag{22}
\]

**Proof:** If \( \tilde{T}(\phi) = \tilde{T}(\phi) \), **F3** states that, for any state \((\pi, \pi')\) with \( \pi < T(\phi) \), we have \( \pi^\phi(\pi, \pi') = \pi^\tilde{\phi}(\pi, \pi') = 0 \). Furthermore, **F4** tells us that the positive recurrent communicating classes are identical, i.e., \( Y^\phi = Y^\tilde{\phi} \). From **F1** and the definition of mapping \( T \), we conclude that, for all states \((\pi, \pi')\) in \( Y^\phi \), \( \phi(\pi, \pi') = \phi(\pi, \pi') \). This in turn means that, for all \((\pi, \pi')\) in \( Y^\phi \), we have \( \pi^\phi(\pi, \pi') = \pi^\tilde{\phi}(\pi, \pi') \). \( \square \)

Let us continue with the proof of Lemma 5. Select \( \tilde{\phi} = \phi_{\tau -} \) with \( \tau = T(\phi) + 1 \). Then, Lemma 6 tells us that \( \pi^\phi = \pi^\tilde{\phi} \).

From the definition of \( \tau^\phi \) in (17), Lemma 5 is now a direct consequence of this observation and **F4**. \( \square \)

The proof of Theorem 3.1 will be completed with the help of the following intermediate result. It tells us that we can focus only on the deterministic policies \( \phi \) with \( \phi(1, A) = 1 \).

**Lemma 7:** There exists a deterministic policy \( \phi^* \) with \( \phi^*(1, A) = 1 \), whose long-term service rate equals \( \lambda^* \).

**Proof:** A proof can be found in Appendix [V]. \( \square \)

Proceeding with the proof of the theorem, Lemma 7 tells us that some deterministic policy \( \phi^* \) with \( \phi^*(1, A) = 1 \) achieves the long-term service rate equal to \( \lambda^* \). This, together with Lemma 5 proves that there exists a deterministic threshold policy that achieves \( \lambda^{**} \) and, as a result, we must have \( \lambda^* = \lambda^{**} \).

**B. A proof of Theorem 3.2**

We prove the theorem by contradiction: suppose that the theorem is false and there exists an arrival rate \( \lambda_1 < \lambda^* \) for which the system is not stable under the policy \( \theta_{\tau^{-}} \). We demonstrate that this leads to a contradiction.

For notational convenience, we denote the unique stationary distribution of \( Y^\phi_{\tau^{-}} \) by \( \pi^\tau^{-} \). In addition, for each \( x_0 \in X \), we define two sequences of distributions \( \{ \xi_k^{x_0}; k \in \mathbb{N} \} \) and \( \{ \phi_k^{x_0}; k \in \mathbb{N} \} \), where \( \xi_k^{x_0} \) and \( \phi_k^{x_0} \) are the distribution of the server state \( Y_k \) and the system \( X_k \), respectively, at epoch \( k \in \mathbb{N} \) under the policy \( \theta_{\tau^{-}} \), conditional on \( \{X_0 = x_0\} \).

We make use of the following lemma to complete the proof of the theorem.

**Lemma 8:** Suppose that the system is not stable under the policy \( \theta_{\tau^{-}} \). Then, for every \( \varepsilon > 0 \) and initial state \( x_0 \in X \), there exists finite \( T'(\varepsilon, x_0) \) such that, for all \( k \geq T'(\varepsilon, x_0) \), we have \( \| \xi_k^{x_0} - \pi^{\tau^{-}} \|_1 < \varepsilon \).

**Proof:** A proof is provided in Appendix [V]. \( \square \)

Proceeding with the proof of the theorem, for every \( k \in \mathbb{N} \),

\[
\mathbb{P}_{\theta_{\tau^{-}}}[\text{complete a task at epoch } k \mid X_0 = x_0] = \sum_{x \in X} \theta_{\tau^{-}}(x) \mu(s) \Phi_k^{x_0}(x) \tag{23}
\]

\[
= \sum_{y \in Y} \phi_{\tau^{-}}(y) \mu(s) \left( \xi_k^{x_0}(y) - \phi_k^{x_0}(y, 0) \right).
\]

First, since the system is assumed to be not stable, Lemma 8 tells us \( \lim_{k \to \infty} \xi_k^{x_0} = \pi^{\tau^{-}} \) for all \( x_0 \in X \). Second, one can argue that the MDP \( X^{\theta_{\tau^{-}}} \) is irreducible because the state \((1, A, 0)\) communicates with all other states, and vice versa. In addition, \((1, A, 0)\) is aperiodic because the probability of transitioning from \((1, A, 0)\) to itself is positive. Therefore, all states are either null recurrent or transient if the system is not stable. Since \( |Y| = 2 \) is finite, this means \( \mathbb{P}_{\theta_{\tau^{-}}}[Q_k = 0 \mid X_0 = x_0] = \sum_{y \in Y} \phi_k^{x_0}(y, 0) \) converges to zero as \( k \) goes to \( \infty \). Thus, for all \( x_0 \in X \), the probability in (23) converges to

\[
\sum_{y \in Y} \phi_{\tau^{-}}(y) \mu(s) \pi^{\tau^{-}}(y) = \lambda^* \text{ as } k \to \infty. \tag{24}
\]

Making use of the reverse Fatou's lemma [2] and the convergence in (24), given any initial distribution of \( X_0 \), we obtain

\[
\mathbb{E} \left[ \limsup_{k \to \infty} \frac{1}{k} \sum_{\tau = 0}^{k-1} 1 \{ \text{complete a task at epoch } \tau \} \right]
\]

\[
\geq \limsup_{k \to \infty} \mathbb{E} \left[ \frac{1}{k} \sum_{\tau = 0}^{k-1} 1 \{ \text{complete a task at epoch } \tau \} \right]
\]

\[
= \limsup_{k \to \infty} \frac{1}{k} \sum_{\tau = 0}^{k-1} \mathbb{E} 1 \{ \text{complete a task at epoch } \tau \}
\]

\[
= \lambda^*.
\]

This implies that, for \( \delta \equiv 0.5(\lambda^* - \lambda_1) > 0 \), we must have

\[
\mathbb{P} \left[ \limsup_{k \to \infty} \frac{1}{k} \sum_{\tau = 0}^{k-1} 1 \{ \text{complete a task at epoch } \tau \} > \lambda_1 + \delta \right] > 0. \tag{25}
\]
On the other hand, for all $k \in \mathbb{N}$, the number of completed tasks up to epoch $k$ cannot exceed the sum of the initial queue size $Q_0$ and the number of arrivals up to epoch $k$. Thus,

$$
\limsup_{k \to \infty} \frac{1}{k} \sum_{\tau=0}^{k-1} 1 \{ \text{complete a task at epoch } \tau \}
\leq \limsup_{k \to \infty} \left( \frac{1}{k} \sum_{\tau=0}^{k-1} 1 \{ \text{a task arrives at epoch } \tau \} + \frac{Q_0}{k} \right) = \lambda_1 \text{ almost surely (by the strong law of large numbers).}
$$

Clearly, this contradicts the earlier inequality in (25).

VI. CONCLUSION

We investigated the problem of designing a task scheduling policy when the efficiency of the server is allowed to depend on the past utilization, which is modeled using an internal state of the server. First, we proposed a new framework for studying the stability of the queue length of the system. Second, making use of the new framework, we characterized the set of task arrival rates for which there exists a stabilizing stationary scheduling policy. Moreover, finding this set can be done by solving a simple optimization problem over a finite set. Finally, we identified an optimal threshold policy that stabilizes the system whenever the task arrival rate lies in the interior of the aforementioned set for which there is a stabilizing policy.

APPENDIX I

A PROOF OF PROPERTY P1

We will prove the claim by contradiction. The decomposition theorem of MCs tells us that $X$ can be partitioned into a set consisting of transient states and a collection of irreducible, closed recurrent communicating classes $\{C^1, C^2, \ldots\}$ [8]. Since $X^\theta$ is assumed stable, all recurrent communicating classes $C_m$, $m = 1, 2, \ldots$, are positive recurrent. Suppose that the claim is false and there is more than one positive recurrent communicating class. We demonstrate that this leads to a contradiction.

First, we show that $C_m$, $m = 1, 2, \ldots$, include a state $(s_m, B, q_m)$ for some $s_m \in S$ and $q_m > 0$. If this is not true, every state in $C_m$ is of the form $(s, A, q)$ and $\theta(s, A, q) = 0$ because $C_m$ is closed [8]. But, this implies that, starting with any state in $C_m$, the scheduler will never assign a task to the server and, consequently, all states in $C_m$ must be transient, which contradicts that $C_m$ is positive recurrent. For the same reason, each $C_m$ must include a state $(s_m, A, q_m)$ with $\theta(s_m) > 0$, which implies that $C_m$ is aperiodic.

Second, if some state $(s, B, q)$ is in $C_m$, $m = 1, 2, \ldots$, then so are all the states $(s', w, q')$ for all $s' \geq s$, $w$ in $W$, and $q' \geq q$. The fact that $(s, B, q)$ communicates with $(s', w, q')$, $s' \geq s$ and $w$ in $W$, which means that these states belong to $C_m$ as well, is obvious. In essence, it is evident that $(s', B, q)$ communicates with $(s, B, q')$ for all $q' \geq q$. In order to see why $(s', A, q')$, $s' \geq s$ and $q' \geq q$, also lie in $C_m$, consider the following two cases: if $\theta(s', A, q) = 0$, then clearly $(s', A, q)$ communicates with $(s', A, q + 1)$. On the other hand, if $\theta(s', A, q) > 0$, $(s', A, q)$ communicates with $(s', B, q + 1)$, which in turn communicates with $(s', A, q + 1)$. The claim now follows by induction.

Note that, the above two observations together imply that there exists finite $q^* \equiv \max\{q_1, q_2\}$ such that all states $(n_x, w, q)$, $w$ in $W$, and $q \geq q^*$, belong to both $C^1$ and $C^2$. This, however, contradicts the earlier assumption that $C^1$ and $C^2$ are disjoint recurrent communicating classes.

APPENDIX II

A PROOF OF LEMMA 2

First, note that every state in $C_\theta$ communicates with some state of the form $(1, A, q)$, which must lie in $C_\theta$ as well; this can be seen from the fact that, after each epoch the server works (resp. rests), the queue size (resp. the action-dependent state) decreases by one with positive probability. As a result, $X^\theta$, starting at $x = (s, w, q)$ in $C_\theta$, can reach a state $(1, A, q')$ for some $q'$ in $\mathbb{N}$, after at most $(s - 1) + q$ epochs with positive probability.

Since $(1, A, q')$ is in $C_\theta$ (because $C_\theta$ is closed [?]), $\bar{q}$ also belong to $C_\theta$. In this turn means that there exists $q^*$ such that (a) $(1, A, q^*)$ is in $C_\theta$, hence $\pi^*(1, A, q^*) > 0$, and (b) $\theta(1, A, q^*) > 0$. Otherwise, the states $(1, A, \bar{q})$, $\bar{q} \geq q'$, must be transient and cannot belong to $C_\theta$, leading to a contradiction.

APPENDIX III

A PROOF OF LEMMA 3

For notational convenience, let $\phi = Q(\theta)$. Taking advantage of the fact that there is a unique stationary PMF of $Y^\phi$, it suffices to show that the distribution given in (16) satisfies the definition of stationary PMF:

$$
\pi^\phi(y) = \sum_{y' \in Y} \pi^\phi(y') P^\phi_{y', y} \text{ for all } y \in Y,
$$

where $P^\phi$ denotes the one-step transition matrix of $Y^\phi$.

• Right-hand side of (26): Using the policy $\phi$ in place,

$$
\sum_{y' \in Y} \pi^\phi(y') P^\phi_{y', y} = \sum_{y' \in Y} \pi^\phi(y')(\phi(y')P^\phi_{y', y} + (1 - \phi(y'))P^\phi_{y', y})
= \sum_{y' \in Y} \pi^\phi(y')(\phi(y')(P^\phi_{y', y}) - (P^\phi_{y', y}) + P^\phi_{y', y}),
$$

where $P^\phi$ (resp. $P^w$) denotes the one-step transition matrix of $Y$ under a policy that always rests (resp. works on a new task) when available.

Substituting (15) for $\phi(y')$ in (27) and using the given
expression \( \pi^\theta(y') = \sum_{q \in \mathcal{F}} \pi^\theta(y', q) \) in (16), we obtain
\[
(27) = \sum_{y' \in Y, q' \in \mathcal{F}} \theta(y', q') \pi^\theta(y', q')(P_{y' y}^w - \circledast_{y' y}) + \sum_{y' \in Y, q' \in \mathcal{F}} \pi^\theta(y', q') \circledast_{y' y}
\]
\[
= \sum_{x' \in X} \theta(x') \pi^\theta(x') (P_{y' y}^w - \circledast_{y' y}) + \sum_{x' \in X} \pi^\theta(x') \circledast_{y' y}.
\]
(28)

- **Left-hand side of (26):** Using (16), we get
\[
\pi^\theta(y) = \sum_{q \in \mathcal{F}} \pi^\theta(y, q).
\]
(29)

For notational ease, we denote \((y, q)\) on the RHS of (29) simply by \(x\). Since \(\pi^\theta\) is the unique stationary PMF of \(X^\theta\), we have
\[
\pi^\theta(y, q) = \sum_{x \in X} \pi^\theta(x') P_{x', x}^\theta
\]
\[
= \sum_{x' \in X} \pi^\theta(x') \theta(x') (P_{x' x}^w - P_{x' x}^\theta) + \sum_{x' \in X} \pi^\theta(x') P_{x' x}^\theta,
\]
where \(P^\theta\) is the one-step transition matrix of \(X^\theta\), and \(P^\circledast\) (resp. \(P^w\)) is the one-step transition matrix under a policy that always rests (resp. assigns a new task) when the server is available and at least one task is waiting for service.

Substituting (30) in (29) and rearranging the summations, we obtain
\[
\pi^\theta(y) = \sum_{x' \in X} \pi^\theta(x') \theta(x') \sum_{q \in \mathcal{F}} (P_{x' x}^w - P_{x' x}^\theta) + \sum_{x' \in X} \pi^\theta(x') P_{x' x}^\theta.
\]
(31)

By comparing (28) and (31), in order to prove (26), it suffices to show
\[
P_{y' y}^w = \sum_{q \in \mathcal{F}} P_{y' y, q}^\theta, a \in A.
\]

Note that
\[
\sum_{q \in \mathcal{F}} P_{y' y, q}^\theta = P_{Y_{k+1} | (Y_k, Q_k, A_k)}(y' | (y', q'), a).
\]
(32)

Clearly, conditional on \(\{(Y_k, A_k) = (y', a)\}\), \(Y_{k+1}\) does not depend on the queue size at epoch \(k\). As a result, the RHS of (32) does not depend on \(q'\) and is equal to \(P_{Y_{k+1} | Y_k, A_k} (y' | y', a) = P_{y' y}^\theta\).

**APPENDIX IV**

**PROOF OF LEMMA 7**

In order to prove the lemma, it suffices to show the following: suppose that \(\phi'\) is a deterministic policy with \(\phi'(1, A) = 0\) and achieves \(\lambda^{**}\). Then, we can find another deterministic policy \(\phi^*\) with \(\phi^*(1, A) = 1\) which achieves \(\lambda^{**}\).

Suppose that \(\phi'\) satisfies \(\phi'(1, A) = 0\) and \(\pi^{\phi'} = \lambda^{**}\). First, we can show that \(\mathcal{T}(\phi') > 1\) by contradiction: assume that \(\mathcal{T}(\phi') = 0\). Then, from the discussion in Section IV-A, we know that \(\{(1, A)\}\) is the unique positive recurrent communicating class of \(Y^{\phi'}\) and, hence, \(\nu^{\phi'} = 0\), thereby contradicting the earlier assumption that \(\nu^{\phi'} = \lambda^{**} > 0\).

Since \(\mathcal{T}(\phi') > 1\), we know that there are two positive recurrent communicating classes \(-\{(1, A)\}\) and \(Y^{\phi'}\) and all other states are transient. But, the first positive recurrent communicating class does not contribute to the long-term service rate and, as a result, the long-term service rate \(\nu^{\phi'} = \lambda^{**}\) is achieved by a stationary PMF only over \(Y^{\phi'}\), which we denote by \(\pi^{\phi'}\).

Consider a new deterministic policy \(\phi^+\) with
\[
\phi^+(s, w) = \begin{cases} 1 & \text{if } 1 \leq \tau \leq \mathcal{T}(\phi'), \\ \phi^+(s, w) \text{ otherwise.} \end{cases}
\]

By construction, clearly \(\mathcal{T}(\phi') = \mathcal{T}(\phi^+)\) and, hence, \(Y^{\phi'} = Y^{\phi^+}\). Because \(\phi^+(s, w) = \phi^+(s, w)\) if \(\tau \geq \mathcal{T}(\phi')\), it follows that \(\pi^{\phi^+} = \pi^{\phi'}\) and, as a result, \(\nu^{\phi^+} = \lambda^{**}\).

**APPENDIX V**

**PROOF OF LEMMA 8**

Denote the one-step transition matrix of \(Y^{\phi^+}\) by \(P^\star\). We prove Lemma 8 with the help of the following two lemmas.

**Lemma 9:** For every \(\varepsilon > 0\), there exists finite \(T_1(\varepsilon)\) such that, for all \(k \geq T_1(\varepsilon)\) and any distribution \(\beta\) over \(Y\), we have
\[
\|\beta (P^\star)^k - \pi^\gamma\|_1 < 0.5 \varepsilon.
\]

**Proof:** Let \(\beta (P^\star)^k \text{ def } \lim_{k \to \infty} (P^\star)^k\), whose rows are equal to \(\pi^\gamma\). Then, given any distribution \(\beta\) over \(Y\), we have \(\beta (P^\star)^k = \pi^\gamma\). Using this equality, for all sufficiently large \(k\), we have
\[
\|\beta (P^\star)^k - \pi^\gamma\|_1 = \|\beta (P^\star)^k - \pi^\gamma\|_1 \leq \|\beta\|_1 \| (P^\star)^k - \pi^\gamma\|_1 = \| (P^\star)^k - \pi^\gamma\|_1 < 0.5 \varepsilon.
\]

The last inequality follows from the definition of \(P^\star\). 

**Lemma 10:** Suppose that the system is not stable. Then, for all \(\varepsilon > 0\), positive integer \(N\), and initial state \(x_0 \in X\), there exists finite \(T_2(\varepsilon, N, x_0)\) such that, for all \(k \geq N + T_2(\varepsilon, N, x_0)\), we have
\[
\|\xi^k_{x_0} - \xi^k_{x_0} (1, A)^N\|_1 < 0.5 \varepsilon.
\]

**Proof:** Let \(P^\theta\) be the one-step transition matrix of \(Y\) under policy \(\phi_1\) that always chooses \(\theta\) when the server is available. We denote the row of \(P^\theta\) (resp. \(P^\circledast\)) corresponding to the server state \(y = (s, u)\) by \(P^\theta_y\) (resp. \(P^\circledast_y\)).

By conditioning on \(X_{k-N}\) and using the equality
\( \xi_k^\infty(y) = \sum_{q \in Y} \psi_k^\infty(y, q) \), we can rewrite \( \xi_{k-N}^\infty \) as

\[
\xi_{k-N+1}^\infty = \sum_{s \in S} \left[ \frac{\xi_{k-N}^\infty(s, A)}{P^*(s, A)} + \sum_{w \in W} \left( \sum_{q \in Y} \psi_{k-N}^\infty(s, w, q) \right) P^*(s, w) \right] = \sum_{s \in S} \left[ \frac{\xi_{k-N}^\infty(s, A)}{P^*(s, A)} + \frac{\xi_{k-N}^\infty(s, B)}{P^*(s, B)} \right]
\]

Define \( \gamma_{\ell-1}^\infty = \sum_{s \in S} \left[ \psi_{\ell-1}^\infty(s, A, 0) \left( P^*(s, A) - P^*(s, B) \right) \right], \ell \in \mathbb{N} \). Applying (33) iteratively, we obtain

\[
\xi_{k+1}^\infty = \xi_{k-N}^\infty(\Phi^*)^N + \sum_{\ell=1}^{N} \gamma_{k-N}^\infty(\Phi^*)^{\ell-1}.
\]

Subtracting the first term on the RHS of (34) from both sides and taking the norm,

\[
\left\| \xi_{k-N}^\infty - \xi_{k-N}^\infty(\Phi^*)^N \right\|_1 \leq \sum_{\ell=1}^{N} \left\| \gamma_{k-N}^\infty(\Phi^*)^{\ell-1} \right\|_1 \leq \sum_{\ell=1}^{N} \left\| \gamma_{k-N}^\infty \right\|_1 \left\| (\Phi^*)^{\ell-1} \right\|_\infty = \sum_{\ell=1}^{N} \left\| \gamma_{k-N}^\infty \right\|_1.
\]

Substituting the expression for \( \gamma_{\ell-1}^\infty \) and using the inequality

\[
\left\| P_{y}^\Phi - P_{y}^\Psi \right\|_1 \leq 2 \forall y \in Y,
\]

we get

\[
\sum_{\ell=1}^{N} \left\| \gamma_{k-N}^\infty \right\|_1 \leq 2 \sum_{\ell=1}^{N} \left( \sum_{s \in S} \psi_{\ell-1}^\infty(s, A, 0) \right).
\]

Recall that if \( X_t^\beta \) is not stable, all states are either null recurrent or transient. This implies that, for all \( x_0 \in X \),

\[
\lim_{t \to \infty} \psi_{t-N}^\infty(s, A, 0) = 0 \forall s \in \mathbb{S}. \text{ Therefore, for all } \varepsilon > 0, \text{ and fixed } N, \text{ and an initial state } x_0, \text{ there exists finite } T_2(\varepsilon, N, x_0) \text{ such that, for all } k \geq T_2(\varepsilon, N, x_0), \text{ we have } \psi_{k-N}^\infty(s, A, 0) < \frac{\varepsilon}{4N^3} \text{. Consequently, the RHS of (35) is smaller than 0.5}\varepsilon \text{ for all } k \geq T_2(\varepsilon, N, x_0) + N.\]

We now proceed with the proof of Lemma 8

\[
\left\| \xi_{k-N}^\infty - \pi^*_{\Phi^\infty} \right\|_1 \leq \left\| \xi_{k-N}^\infty - \xi_{k-N}^\infty(\Phi^*)^N \right\|_1 + \left\| \xi_{k-N}^\infty(\Phi^*)^N - \pi^*_{\Phi^\infty} \right\|_1 \leq 0.5\varepsilon.
\]

Since Lemma 9 holds for any distribution \( \beta \), if we choose \( \beta = \xi_{k-N}^\infty \), for all \( N \geq T_1(\varepsilon) \), the second term in (36) is smaller than 0.5\varepsilon. In addition, Lemma 10 tells us that we can find \( T_2(\varepsilon, N, x_0) \) such that, for all \( k \geq N + T_2(\varepsilon, N, x_0) \), the first term in (36) is upper bounded by 0.5\varepsilon. Thus, it is clear that Lemma 8 holds with \( T(\varepsilon, x_0) = T_1(\varepsilon) + T_2(\varepsilon, T_1(\varepsilon), x_0) \).