Authors and Affiliations

Rob Shone [a,*] (r.shone@lancaster.ac.uk)

Kevin Glazebrook [a] (k.glazebrook@lancaster.ac.uk)

Konstantinos G. Zografos [a] (k.zografos@lancaster.ac.uk)

[a] Centre for Transport and Logistics (CENTRAL),

Department of Management Science,

Lancaster University Management School,

Bailrigg,

Lancaster,

United Kingdom.

LA1 4YX

* Corresponding author

Resource allocation in congested queueing systems with time-varying demand: An application to airport operations

Rob Shone, Kevin Glazebrook, Konstantinos G. Zografos

Lancaster University Management School, Bailrigg, Lancaster, UK

Abstract

Motivated by the need to develop time-efficient methods for minimizing operational delays at severely congested airports, we consider a problem involving the distribution of a common resource between two sources of time-varying demand. We formulate this as a dynamic program in which the objective is based on second moments of stochastic queue lengths and show that, for sufficiently high volumes of demand, optimal values can be well-approximated by quadratic functions of the system state. We identify conditions which enable the strong performance of myopic policies and develop approaches to the design of heuristic policies by means of approximate dynamic programming (ADP) methods. Numerical experiments suggest that our ADP-based heuristics, which require very little computational effort, are able to improve substantially upon the performances of more naive decision-making policies, particularly if exogenous system parameters vary considerably as functions of time.

Keywords: Queueing, Optimization, Approximate dynamic programming, Airport operations, Aviation

1 Introduction

In the modern world there exists a need to develop fast, responsive and easily adaptable methods for computing optimal (or near-optimal) solutions to problems in which resources must be allocated dynamically in order to satisfy time-varying demands from multiple sources. A particularly pertinent example of this can be found in the real-time deployment of scarce resources at busy airports. With increasing numbers of airports operating at the very limits of their capacity in order to maximize runway throughput rates in the face of ever-intensifying demand from airlines and other stakeholders, efficient usage of runway capacity is essential if significant operational delays are to be avoided.

In this paper, we consider the optimal dynamic control of service facilities which can be modelled as dual-class queueing systems with random, nonstationary patterns of demand and adjustable service rates. Various authors have formulated problems in which service rates may be adjusted dynamically in response to a stationary pattern of demand [1, 3, 13, 18, 34, 50], and used the techniques of stochastic dynamic programming to characterize optimal policies. In real-world applications, however, there is usually a need

to consider time-varying demand. Classical applications which have received a lot of attention in the literature include telephone call centres and hospital emergency departments [11, 21, 51]. In these settings, service rates are typically controlled via the setting of hourly staffing levels.

Although the results in this paper are intended to be of general interest, our primary motivation is the importance of reducing operational (queueing) delays at busy airports. The question of how to avoid excessive flight delays during peak congested periods has become increasingly topical in recent years due to relentless air traffic growth, which has placed the resources of many of the world's busiest airports under seemingly unsustainable pressure [16, 20]. In 2016, more than 15 million minutes of delays were experienced by air passengers due to inefficiencies in the air traffic management system [28].

Measures aimed at reducing operational delays on a particular day of operations at an airport can be separated into two distinct categories:

- **Strategic interventions:** These are made in advance of the day of operations in question, and involve designing schedules and setting 'declared capacity' constraints in such a way that airline requests can be satisfied without causing large queues of traffic to form during congested periods [54]. For example, "demand smoothing" [43] may help to achieve this.

- **Tactical interventions:** These are made on the day of operations itself in response to real-time observed queues of traffic and other important factors, which may include weather and wind conditions and information about future expected arrival times.

In this paper we assume that the properties of the random, nonstationary processes by which customers arrive in the system have been predetermined and are not subject to any real-time control. In an airport context, this means we take the schedule of flights as a 'given' and consider the optimization of tactical interventions only. The adjustment of time-varying service rates for different customer classes is within the realm of tactical control, and in an airport context this should be interpreted as the dynamic allocation of resources between two sources of demand (arrivals and departures). We can also consider other applications involving the dynamic allocation of service capacity, including make-to-stock production systems [26, 37] in which product inventories are updated based on anticipated consumer demand, maintenance problems with different classes of machines subject to breakdown [10] and resource allocation in wireless networks [15].

Tactical control of airport operations constitutes a relatively new area of research which began with Gilbo [19], who considered a problem involving the allocation of service rates between arrivals and departures based on linear capacity constraints. Gilbo's work, as well as that of other authors [7, 12], assumed deterministic queueing dynamics. Jacquillat and Odoni [30] introduced a dynamic optimization framework based on stochastic queue dynamics which represents a considerable advancement on earlier work. Their model incorporates the dynamic control of runway configurations and arrival/departure service rates and includes stochastic models for the evolution of weather and wind conditions.

This paper differs from previous research by focusing on the optimization of tactical decision-making during a congested period of time in which servers at the service facility are likely to be continuously busy. We suppose that, during the period of interest, the probability of any server becoming idle is negligible, and refer to this as the "Servers Always Busy" (SAB) assumption. This assumption, which becomes asymptotically correct as demand rates increase towards infinity, facilitates an elegant mathematical analysis which

can be used as the basis for a heuristic approach involving the use of approximate dynamic programming techniques to compute parametric value function approximations.

At a severely congested airport, there may indeed be lengthy periods of time during a typical day during which usage of the runways is constantly in demand from aircraft waiting to perform take-offs and landings. This is especially likely to be the case if weather conditions are poor, or if wind conditions prevent certain runway configurations from being used [4, 30, 43]. Naturally, the "busy" times of day are those at which capacity utilization decisions will have the greatest impact. In particular, lengthy queues and operational delays may result from suboptimal decision-making. Thus, we conjecture that optimizing tactical decision-making during a busy portion of the day may indeed go a long way towards alleviating the worst effects of congestion and optimizing overall performance.

An important measure of daily performance at an airport is the proportion of flight delays which exceed a certain length of time. In the airline industry, it is common practice to refer to departures or arrivals which take place within 15 minutes of their scheduled times as "on-time" [40, 44, 53]. For this reason, it is important that our objective function takes into account not only the average queueing delays over a period, but also their variability. In this paper we consider an objective function based on weighted second moments of queue lengths (measured in terms of service phases). Objectives based on second moments have previously been considered in [30, 31] and are appropriate to use in problems where control must be exercised over the tails of queue length distributions.

The key contributions of this paper are as follows:

- We formulate an optimization problem based on the dynamic, tactical control of service rates for two distinct customer classes in response to nonstationary demand rates and capacity constraints. We then discuss a natural method for approximating the queueing dynamics of the system during a busy period and, based on this idea, introduce a "surrogate problem" which offers greater mathematical tractability than our original optimization problem.

- For the surrogate problem, we derive *myopic* decision-making policies (which optimize immediate outcomes) and also prove certain structural properties of *optimal* policies. We argue that, in certain plausible cases for the system parameters, the myopic and optimal policies obtained using our analyses of the surrogate problem should also perform well in the original problem.

- We develop various heuristic policies for the original optimization problem based on properties of the surrogate problem. One such heuristic applies 'approximate' policy improvement to a myopic policy, while two others are based on quadratic approximations for the value function. We test these heuristics against more 'naive' decision-making policies using simulation experiments.

Our work builds upon earlier research by introducing a new methodological approach to address problems involving optimal resource allocation in highly congested queueing systems. Our modelling assumptions are similar to those made in the existing airport literature, but our solution methods (including the development of value function approximations based on simplified queueing dynamics) are innovative and provide an alternative to rigorous dynamic programming methods which tend to be slow and reliant upon computational 'fixes', e.g. discretizations of the state and action spaces. We anticipate several exciting possibilities for future research, including the potential for our busy-period analyses to be incorporated within hybridized optimization methods for large-scale transportation problems.

In Section 2 we formulate the problem of optimizing tactical decision-making during a pre-specified period of operations as a finite-horizon Markov decision process in which we aim to minimize a cost function which takes into account system congestion levels and operational costs associated with redeployment of resources. The "servers always busy" (SAB) approximation is discussed in Section 3. In Section 4 we introduce our 'surrogate' problem, discuss myopic policies and develop quadratic approximations for the value function. Section 5 describes our heuristic methods, and Section 6 provides the results of our numerical experiments. In Section 7 we summarize the contributions of this paper.

2 Model formulation

2.1 The demand and service processes

There exists a well-established convention for airports to be modelled as queueing systems in which arrival and departure queues evolve independently according to $M(t)/E_k(t)/1$ dynamics [30, 32, 33, 35, 42, 43]. The use of nonhomogeneous Poisson processes to model 'arrivals' to the runway (i.e. take-offs and landings) allows demand rates to be calibrated according to the schedule of operations for the day and other relevant operational factors. For service times, the $E_k(t)$ (time-dependent Erlang) assumption offers the advantage of mathematical tractability and also allows many types of $M(t)/G(t)/1$ queues to be well-approximated due to the "richness" of the Erlang family of distributions [24]. The treatment of arrivals and departures as independent queueing processes is arguably an oversimplification, since aircraft arriving at an airport will enter the queue for departures some time later; however, empirical studies have suggested that this assumption is adequate for practical purposes [45].

In this paper we consider a queueing system in which two classes of customer arrive via separate, independent nonstationary random processes. In keeping with the airport paradigm, we will refer to the customer classes as $\mathcal{A}$ (for arrivals) and $\mathcal{D}$ (for departures). We will consider two possible cases for the random processes governing customer arrival times:

- **Case 1:** Arrivals for customers of classes $\mathcal{A}$ and $\mathcal{D}$ follow separate, independent nonhomogeneous Poisson processes with demand rate functions $\lambda(t)$ and $\eta(t)$ respectively ($t \geq 0$).

- **Case 2:** An individual customer i (of either class) is scheduled to arrive at time ω_i. However, their actual arrival time is $\omega_i - \mathbb{E}[\xi_i] + \xi_i$, where the ξ_i are mutually independent, exponentially-distributed random variables and $\mathbb{E}[\xi_i]$ takes a uniform value of either $\bar{\xi}_\mathcal{A} > 0$ or $\bar{\xi}_\mathcal{D} > 0$ depending on whether customer i is of class $\mathcal{A}$ or class $\mathcal{D}$.

As mentioned above, models with Poisson arrivals have been relied upon extensively in the previous aviation literature and empirical studies have justified their use [14, 29, 52]. In recent years, however, there has been a trend towards considering models in which inter-arrival time variances are smaller than one would observe under a Poisson model [27, 36]. A key factor behind this trend has been the development of projects such as NextGen in the USA [38] and SESAR in Europe [17], which are intended to make flight arrival times more predictable through the use of four-dimensional trajectory-based operations. We therefore include Case 2, which bears similarities to pre-scheduled random arrival (PSRA) models considered in recent literature [9, 23, 25, 36], as an alternative to the Poisson model.

We assume that service times for all customers are Erlang-distributed and consist of $k \geq 1$ i.i.d service phases; these phases are exponentially distributed with intensity rates $k\mu(t)$ and $k\nu(t)$ for classes $\mathcal{A}$ and $\mathcal{D}$ respectively, where $t \geq 0$ is the continuous time variable.

The service rates $\mu(t)$ and $\nu(t)$ are within the realm of tactical control in our model and are modelled as piecewise constant functions, as in [19, 30, 31]. We consider a period of operations $[0, T]$ and let $0 = t_0 < t_1 < ... < t_S = T$ be a partition of this interval into S equal-length time intervals. For each $s \in \{1, 2, ..., S\}$, the interval $[t_{s-1}, t_s)$ will be referred to as time interval s. At the beginning of interval $s \geq 1$, a decision-maker chooses a pair of service rates (μ_s, ν_s) for the respective customer classes and these remain constant until the next interval, i.e. $\mu(t) = \mu_s$ and $\nu(t) = \nu_s$ for $t \in [t_{s-1}, t_s)$. For each $s \in \{1, 2, ..., S\}$, there is a set of linear constraints which governs the service rates μ_s, ν_s that may be chosen. We will use J_s to denote the number of linear constraints in interval s and write these constraints as

$$\gamma_{js}\mu_s + \nu_s \leq \theta_{js} \quad j \in \{1, 2, ..., J_s\},$$

where the parameters γ_{js} and θ_{js} are positive constants. We will also assume that $\gamma_{1s} < \gamma_{2s} < ... < \gamma_{J_s,s}$ and that the intersection points between pairs of constraints are located in the positive quadrant of the μ_s-ν_s plane, in order to avoid redundancy. For example, in Figure 1, the decision-maker may choose any point in the region R_s enclosed by the lines L_{1s}, L_{2s}, L_{3s} and the coordinate axes, and this corresponds to a pair of service rates (μ_s, ν_s) to be used during interval s.
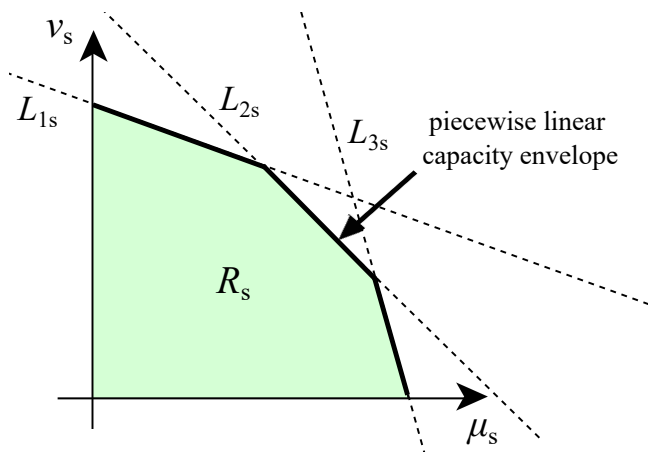

Figure 1: The piecewise linear capacity envelope for selection of μ_s and ν_s at the beginning of time interval s. For each $j \in \{1, 2, 3\}$, the line L_{js} has equation $\gamma_{js}\mu_s + \nu_s \leq \theta_{js}$.

In fact, it can be assumed that the decision-maker always chooses a point on the upper boundary of R_s, since the objectives that we consider in this paper imply that one should always choose the largest possible value of ν_s for a given value of μ_s (this can be shown using a simple sample path argument, which we have omitted here). It follows that the decision-maker only needs to choose one value (μ_s) at each time interval, with ν_s then being given via the dominant linear constraint.

The assumption of linear constraints is well-suited to the airport applications that we are considering, since it is standard practice in the literature to represent the maximum number of departures that can be processed in any time interval as a function of the number of arrivals by means of a *piecewise linear capacity envelope* [19]. This capacity envelope can be constructed using an intricate procedure described in [46]

which depends on aircraft separation requirements, approach speeds, runway occupancy times and other data. Recently, Jacquillat and Odoni [30] recommended interpreting each point on the piecewise linear envelope as a pair of *average* service rates for arrivals and departures (taking into account the uncertainty involved in actual take-off and landing times) and referred to the envelope as an *operational throughput envelope* (OTE). This interpretation makes more sense in the context of our paper, since we are treating each μ_s (or ν_s) as an *expected* number of service completions per unit time for class $\mathcal{A}$ (or $\mathcal{D}$) during interval s, rather than an upper bound. In reality, the exact shape of an OTE is interval-dependent and cannot be predicted with certainty for future time intervals, since it depends on weather conditions and also the mix and sequencing of different types of aircraft entering the queues for the runways. However, airports invariably have sophisticated meteorological equipment available for making weather forecasts and the mixture of air (or ground) traffic at an arbitrary point in time is heavily dependent on the predetermined schedule for arrivals (or departures), so it is reasonable to suppose that the OTEs for time intervals in the not-too-distant future can be predicted with a high degree of confidence.

The system evolves as a pair of independent, first-come-first-served queues of type $M(t)/E_k(t)/1$ or $PSRA/E_k(t)/1$, depending on the nature of the arrivals processes. We assume that the service rates selected at the beginning of any time interval are chosen with knowledge of the latest observed numbers of customers waiting in both queues, the random distribution of future arrival times and the service rate constraints for future intervals. In the PSRA case, the history of previous customer arrivals affects future random arrival times and we assume that this historical information is also known.

2.2 The objective function

We assume that the objective of the decision-maker is to find an admissible dynamic policy π for choosing the service rates $\mu_1, \mu_2, ..., \mu_S$ which minimizes

$$\sum_{s=1}^{S} \mathbb{E}^{\pi}\left[\alpha X_s^2 + Y_s^2 | x_0, y_0\right] + \beta \sum_{s=2}^{S} [k\tau(\mu_s - \mu_{s-1})]^2, \tag{1}$$

where X_s and Y_s are the respective numbers of Erlang service *phases* for customers of classes $\mathcal{A}$ and $\mathcal{D}$ remaining to be processed at the end of interval s, x_0 and y_0 are the initial values for these phase counts, k is the number of Erlang service phases per customer (as stated in Section 2.1), $\tau = t_1 - t_0$ is the length of each time interval, and $\alpha > 0$ and $\beta \geq 0$ are constants. The quantity $[k\tau(\mu_s - \mu_{s-1})]^2$ (which may be interpreted as the squared difference in the expected number of service phase completions between intervals s and $s-1$ if the server remains constantly in use) represents the *switching cost* of changing the service rate (for class $\mathcal{A}$) from μ_{s-1} to μ_s. Thus, the decision-maker aims to minimize the sum of two terms: a weighted, aggregated measure of the congestion in the system over all time intervals, and a penalty which increases with the frequency and severity of changes made to the service rates.

Our motivation for including the second term in (1) is that, in many applications, there may be some cost associated with altering the service rates due to the disruption caused by having to re-allocate physical and/or personnel resources. At airports, for example, changes in the service rates for arrivals and departures may require re-deployment of ground resources (e.g. apron stands, buses, marshallers) and in some cases, complicated operations such as runway configuration changes might be required. In inventory problems, frequent changes to the production rates for different products may result in inefficiencies due

to set-up costs, etc. However, in some applications the inclusion of a switching cost may be unnecessary and throughout this paper we make special reference to the case $\beta = 0$.

Weighted second moments of queue lengths have previously been used in [30] as a performance measure, although in our case the formulation is slightly different because X_s and Y_s represent the numbers of service *phases* (not customers) present after s time intervals. Throughout this paper, we refer to the number of phases present as the *phase count*. The factor α allows for the possibility of class $\mathcal{A}$ customers being more expensive to keep waiting than class $\mathcal{D}$ customers, or vice versa.

The use of weighted second moments in our objective function can be justified in various different ways. Due to the nature of the Erlang service process, if $n \geq 1$ customers are present in one of the queues (including the customer being served), then the phase count r satisfies $(n-1)k + 1 \leq r \leq nk$, where k is the Erlang parameter. As explained in [30], the total of the expected delays for $n = \lceil r/k \rceil$ customers of the same class present in the system (where $\lceil \cdot \rceil$ denotes the ceiling function) is indeed of order n^2 and therefore also of order r^2. Furthermore, taking into account the second moments of X_s and Y_s assists us in finding policies which will restrict the tails of their distributions, so that the system performs better with respect to measures such as the proportion of "on-time" service completions (as discussed in the introduction). In real-world applications, of course, practitioners are likely to be interested in meaningful performance measures rather than convoluted objective functions, and therefore it is desirable to show that a policy which performs well with respect to the objective 1 also achieves strong results with respect to other criteria. Later in this paper, we will examine the performance of our heuristics with respect to an alternative performance measure (expected waiting times) using numerical experiments.

2.3   MDP formulation

The problem that we are considering can be modelled as a Markov decision process (MDP) [39, 41] with a finite time horizon in which we aim to minimize expected total cost. Below, we describe the state space, action space, transition probabilities and single-stage cost function for the MDP.

- The **state space** of the MDP, which we will denote by $\mathcal{X}$, consists of vectors of the form

$$\mathbf{x} = \begin{cases} (x, y, \bar{\mu}) \in \mathbb{Z}_{\geq 0}^2 \times \mathbb{R}_{\geq 0}, & [\textbf{Poisson case}], \\ (x, y, z, w, \bar{\mu}) \in \mathbb{Z}_{\geq 0}^4 \times \mathbb{R}_{\geq 0}, & [\textbf{PSRA case}], \end{cases} \tag{2}$$

where x and y are the latest observed phase counts for the respective customer classes and $\bar{\mu}$ is the service rate for class $\mathcal{A}$ used in the previous time interval (this can be given a dummy value for the first interval). In the PSRA case, the extra state variables z and w are the numbers of class $\mathcal{A}$ and $\mathcal{D}$ customers respectively whose arrivals are 'imminent', in the sense that they have not yet arrived despite their earliest possible arrival times having passed (in other words, their exponential 'clocks' have already started running). Note that the memoryless property of the ξ_i variables implies that we do not need to know specifically *which* of the pre-scheduled arrivals have not yet occurred; we only require the number of imminent arrivals for each customer type. The inclusion of $\bar{\mu}$ as a state variable implies that the state space is uncountable, unless we are considering the special case $\beta = 0$ in which case $\bar{\mu}$ may be omitted.

- The **action space** at interval s is one-dimensional and consists of the continuous range of service

rates μ_s which satisfy the feasibility constraints described in Section 2.1 (based on the relevant OTE). Specifically, given a set of constraints of the form $\gamma_{js}\mu_s + \nu_s \leq \theta_{js}$ in interval s, we require $0 \leq \mu_s \leq \min_j\{\theta_{js}/\gamma_{js}\}$. As explained in Section 2.1, we can then assume that the complementary service rate ν_s is given by $\theta_{js} - \gamma_{js}\mu_s$.

- **Transition probabilities** for our MDP are governed by the underlying nonstationary queueing dynamics. For example, in the case of Poisson arrivals, these probabilities are based on the Chapman-Kolmogorov equations for $M(t)/E_k(t)/1$ queues (see, for example, [22]). We have included a full description of these equations in Appendix A.1.

- The **one-stage cost function** is related to the objective function (1) and can be defined as

$$f_s(\mathbf{x}, \mu_s) := \mathbb{E}[\alpha X_s^2 + Y_s^2 | \mathbf{x}, \mu_s] + \beta(k\tau)^2(\mu_s - \bar{\mu})^2. \tag{3}$$

- The **Bellman optimality equation** can be written, for $\mathbf{x} \in \mathcal{X}$, as

$$V_s(\mathbf{x}) = \min_{\mu_s \in \Theta_s} \left\{ f_s(\mathbf{x}, \mu_s) + \mathbb{E}[V_{s+1}(\mathbf{X}_s)|\mathbf{x}, \mu_s] \right\} \qquad (s \in \{1, 2, ..., S\}), \tag{4}$$

where Θ_s is the feasible set of service rates for interval s, $\mathbf{X}_s$ is the random state at the end of interval s and we set $V_{S+1}(\mathbf{x}) := 0$ for all states $\mathbf{x}$ since our objective function (1) does not include costs incurred beyond the end of slot S. The functions $V_1, V_2, ..., V_S$ are referred to as *value functions*. We also use $V(\mathbf{x}) := V_1(\mathbf{x})$ to denote the optimal expected total cost starting from state $\mathbf{x}$.

The inclusion of *phase counts* in our cost and objective functions makes sense from an optimization perspective, since we should be able to obtain stronger-performing policies by considering the fine-grain state of the system. However, we note that the phase counts are a mathematical construct and in physical applications would not be observable by a decision-maker. Instead, decisions would be taken in response to the numbers of *customers* in the queues. In this paper, we take the approach of finding an optimal *phase-dependent* policy (allowing decisions to depend on the exact phase counts), and then aggregating this to obtain a *customer-dependent* policy which chooses the same decision at any two states $(x, y, z, w, \bar{\mu})$ and $(\hat{x}, y, z, w, \bar{\mu})$ such that $x, \hat{x} \in \{(n-1)k+1, ..., nk\}$ for some $n \in \mathbb{N}$. Various methods of aggregation are possible, but in this paper we adopt a simple averaging method which works as follows: given that n and m customers are observed in the $\mathcal{A}$ and $\mathcal{D}$ queues respectively and $\bar{\mu}$ is the service rate for the previous time interval, let the service rate chosen for class $\mathcal{A}$ be the mean of the service rates specified by the phase-dependent policy at the k^2 states $(x, y, z, w, \bar{\mu})$ satisfying both $(n-1)k \leq x \leq nk$ and $(m-1)k \leq y \leq mk$. Our simulation results in Section 6 incorporate this aggregation method.

Unfortunately, the MDP formulated above cannot be solved using conventional dynamic programming (DP) methods such as value iteration or policy iteration. The reasons for this include the continuous state and action spaces, the large number of possible random transitions between consecutive time intervals and the fact that the transition probabilities themselves can only be obtained accurately via repeated recourse to slow numerical methods. Indeed, any computationally feasible DP algorithm would require an excessive amount of time to obtain a mere approximation to an optimal solution. The rest of this paper focuses on the development of heuristic policies which should be effective under busy conditions.

3 The "Servers Always Busy" (SAB) approximation

As discussed in the introduction, we aim to make progress by restricting attention to an interval of time $[0, T]$ during which there is a high probability of both queues remaining non-empty. Initially, we will consider the first of the two cases for arrival times described in Section 2, in which customers arrive via nonhomogeneous Poisson processes. For class $\mathcal{A}$, we use $\lambda(t)$, $\mu(t)$ and $X(t)$ to denote (respectively) the arrival rate, the service rate and the number of phases remaining to be completed (i.e. the phase count) at time t, where $t \in [0, T]$ is a continuous time variable. For class $\mathcal{D}$, the corresponding functions are denoted by $\eta(t)$, $\nu(t)$ and $Y(t)$. Let us also define two random variables $\hat{X}(t)$ and $\hat{Y}(t)$ by

$$\hat{X}(t) = x_0 + kA(t) - B(t), \quad \hat{Y}(t) = y_0 + kD(t) - C(t) \qquad \forall t \in [0, T], \tag{5}$$

where $x_0 = X(0)$ and $y_0 = Y(0)$ are the initial phase counts and $A(t)$, $B(t)$, $D(t)$ and $C(t)$ are mutually independent Poisson random variables:

$$A(t) \sim \mathrm{Po}\left(\int_0^t \lambda(u)du\right), \quad B(t) \sim \mathrm{Po}\left(k\int_0^t \mu(u)du\right),$$

$$D(t) \sim \mathrm{Po}\left(\int_0^t \eta(u)du\right), \quad C(t) \sim \mathrm{Po}\left(k\int_0^t \nu(u)du\right). \tag{6}$$

Note that $B(t)$ and $C(t)$ are phased-based measures (service phase completion counts), but $A(t)$ and $D(t)$ are customer-based measures (customer arrival counts). If the servers are always busy during $[0, T]$ then the customer arrival and service phase completion processes for the individual queues are indeed independent Poisson processes and we have $\hat{X}(t) \equiv X(t)$ and $\hat{Y}(t) \equiv Y(t)$, i.e. $\hat{X}(t)$ and $\hat{Y}(t)$ mirror the physical state of the system. If the servers are *almost* always busy during $[0, T]$ then we can expect $\hat{X}(t)$ and $\hat{Y}(t)$ to provide good approximations for $X(t)$ and $Y(t)$. In such cases, we can approximate the probability distributions at $t \in [0, T]$ by exploiting their relations to the random variables in (6). If t is sufficiently large, we can make this process more practical computationally by claiming that, by the central limit theorem, $\hat{X}(t)$ and $\hat{Y}(t)$ are approximately normally distributed with means and variances given as follows:

$$\hat{X}(t) \sim \mathrm{N}\left(x_0 + \int_0^t k(\lambda(u) - \mu(u))du, \int_0^t (k^2\lambda(u) + k\mu(u))du\right),$$

$$\hat{Y}(t) \sim \mathrm{N}\left(y_0 + \int_0^t k(\eta(u) - \nu(u))du, \int_0^t (k^2\eta(u) + k\nu(u))du\right). \tag{7}$$

On the other hand, suppose we are considering the second case (PSRA) described in Section 2, whereby customer arrival times are randomly displaced from their pre-scheduled times. For simplicity we will describe the situation only for the first queue (with state $X(t)$); the situation for the $Y(t)$ queue is obviously analogous. Let ω_i denote the scheduled arrival time of the ith customer, so that $\omega_i - \bar{\xi}_{\mathcal{A}}$ is their earliest possible time of arrival. We consider a random variable

$$\hat{X}(t) = x_0 + kA(t) - B(t) \tag{8}$$

where $x_0 = X(0)$ and $B(t) \sim \text{Po}\left(k \int_0^t \mu(u)du\right)$ (as before), but this time

$$A(t) = H_1(t) + H_2(t) + ... + H_m(t) \tag{9}$$

with $m = \max\{i : \omega_i - \bar{\xi}_\mathcal{A} < t\}$. Here, $H_i(t)$ is a Bernoulli random variable equal to 1 if the ith customer arrives before time t and 0 otherwise. Hence,

$$\mathbb{E}[A(t)] = \sum_{i=1}^{m} p_i(t), \quad \text{Var}[A(t)] = \sum_{i=1}^{m} p_i(t)(1 - p_i(t)),$$

where

$$p_i(t) = 1 - \exp\left(-\frac{t - \omega_i + \bar{\xi}_\mathcal{A}}{\bar{\xi}_\mathcal{A}}\right).$$

If approximations are required for state probabilities, then (as in the Poisson case) we can make use of the central limit theorem for sufficiently large t and write

$$\mathbb{P}(\hat{X}(t) \leq n) \approx \Phi\left(\frac{n + 0.5 - \mathbb{E}[\hat{X}(t)]}{\sqrt{\text{Var}[\hat{X}(t)]}}\right) \tag{10}$$

for $n \in \mathbb{N}$, with $\mathbb{E}[\hat{X}(t)]$ and $\text{Var}[\hat{X}(t)]$ derived from (8) (this holds since the variables $H_i(t)$ are independent of each other, albeit not identically distributed). However, the ADP methods that we develop later in this paper do not actually require approximations for state probabilities.

Taking into account the importance of being able to apply our methods to real-world scenarios, we tested the quality of the SAB approximation using data obtained from John F. Kennedy International Airport in New York. Our experiment included time-dependent Poisson demand rates for arrivals and departures which were constructed using a particular flight schedule, and two different sets of capacity constraints (OTEs) based on alternative weather scenarios. Based on these capacity constraints, a naive 'greedy' heuristic was used to select the time-varying service rates and the quality of the SAB approximation was then tested by comparing its estimates for the objective function (1) with 'exact' values derived using Chapman-Kolmogorov equations. A full description of our experimental procedure and results can be found in Appendix A.2, and we also summarize the results here:

- We found that, throughout a period of time beginning at roughly 4:30PM and ending at roughly 10:00PM on the day of operations that we considered, the arrivals queue and the departures queue both had idleness probabilities smaller than 0.05 in the 'good weather' case. In the 'poor weather' case, these probabilities were consistently smaller than 0.03.

- The error incurred by using the SAB approximation to estimate the objective function (1) over a period of 22 consecutive 15-minute time intervals, beginning at 4:30PM and ending at 10:00PM, was only 0.26% in the good weather case, and 0.02% in the poor weather case.

- In the poor weather case, it was possible to consider a longer time period beginning at 3:15PM and ending at 11:15PM without incurring a SAB approximation error greater than 0.5%.

Based on these findings, we propose to develop heuristic optimization strategies which offer the poten-

tial to improve the quality of decision-making at JFK Airport during a schedule-dependent 'busy' period of the day (e.g. 4:30PM to 10:00PM) by making use of the SAB approximation.

4   A surrogate optimization problem

As discussed earlier, the MDP formulated in Section 2.3 cannot be solved using conventional dynamic programming (DP) methods. We therefore aim to develop heuristic policies which are based on the simplified queueing dynamics discussed in Section 3.

In this section we introduce a surrogate optimization problem which closely resembles the problem formulated in Section 2, especially in cases where customer demand is consistently high in relation to service capacity. As discussed in the introduction, we intend to utilize this surrogate problem by (i) characterizing myopic policies and investigating the conditions under which these perform well; (ii) investigating the nature of optimal policies by finding quadratic approximations for the value functions in terms of the state variables; (iii) developing heuristic methods for the *original* (not the surrogate) optimization problem based on the properties of these myopic policies and quadratic approximating functions.

Please note that throughout this section we will make repeated use of the variables x, y, z, w and $\bar{\mu}$ (as in (2)) to represent a generic observed state at the beginning of an arbitrary time interval s, without including an explicit s-dependence on these variables.

The surrogate problem differs from the problem formulated in Section 2 in the following ways:

1. For each $s \in \{1, 2, ..., S\}$, we remove the constraint $\mu_s \in \Theta_s$ and allow the service rate μ_s (for class $\mathcal{A}$) to take any positive or negative real value, with the corresponding service rate ν_s (for class $\mathcal{D}$) then being given by the dominant linear constraint. Specifically, given a set of J_s linear constraints $\{\gamma_{js}\mu_s + \nu_s \leq \theta_{js}\}_{j=1}^{J_s}$, the feasible region for (μ_s, ν_s) is now

$$\{(\mu_s, \nu_s) \in \mathbb{R}^2 : \nu_{js} \leq \theta_{js} - \gamma_{js}\mu_s \ \ \forall j \in \{1, 2, ..., J_s\}\}$$

   and is therefore an infinite region given by the intersection of closed half-planes.

2. Following the selection of $(\mu_s, \nu_s) \in \mathbb{R}^2$ in time interval s, the random phase counts X_s and Y_s at the end of interval s are given by

$$X_s = x + kA_s - \text{sgn}(\mu_s)B_s, \quad Y_s = y + kD_s - \text{sgn}(\nu_s)C_s, \tag{11}$$

   where $\text{sgn}(\cdot)$ denotes the signum function, x and y are the phase counts at the beginning of interval s, and A_s, B_s, C_s and D_s are mutually independent, discrete random variables. In particular, A_s and D_s (which represent the numbers of customer arrivals during interval s) obey the same laws as in the original problem and these depend on which case we are considering for the demand processes (either Poisson or PSRA). We define B_s and C_s as follows:

$$B_s \sim \text{Po}(k\tau|\mu_s|), \ \ C_s \sim \text{Po}(k\tau|\nu_s|).$$

   Note that, as a result of (11), a negative service rate causes the phase count to increase by 1 rather than decreasing by 1 at the end of a service phase completion.

11

We will denote the one-step cost function in the surrogate problem by

$$g_s(\mathbf{x}, \mu_s) := \mathbb{E}[\alpha X_s^2 + Y_s^2 | \mathbf{x}, \mu_s] + \beta(k\tau)^2(\mu_s - \bar{\mu})^2, \tag{12}$$

so that it essentially has the same definition as $f_s(\mathbf{x}, \mu_s)$ (in the original problem) except that the expectation is with respect to the dynamics of the surrogate problem. The objective in the surrogate problem is to find a dynamic policy π for choosing $\mu_1, \mu_2, ..., \mu_S \in \mathbb{R}$ which minimizes

$$U(\mathbf{x}) := \mathbb{E}^\pi \left[\sum_{s=1}^{S} g_s(\mathbf{x}_s, \mu_s) \,\middle|\, \mathbf{x}_0 = \mathbf{x} \right]. \tag{13}$$

Throughout Sections 4 and 5 we will assume that the time interval duration τ is equal to one. This is without loss of generality, since the demand rates and service rates can always be re-scaled appropriately. We should also note that the objective function (13) implies that, as in the original problem, the service rates (μ_s, ν_s) should always be chosen on the upper boundary of the feasible region. Therefore the notation in (12) makes sense, since the complementary service rate ν_s is given as a deterministic function of μ_s via the dominant linear constraint.

By assuming that the random variables in (11) are independent, we implicitly relax the constraints on the phase counts X_s and Y_s by allowing them to take negative values. Effectively, we are considering a mathematical abstraction of our original problem in which servers remain busy at all times and can cause phase counts to either increase or decrease (depending on the signs of the service rates). Naturally, this abstraction makes it easier to perform analyses since we are approximating the queueing dynamics of the original system using independent random variables. At the same time, our analyses should provide useful insights into the nature of optimal policies during busy periods in which phase counts are likely to remain positive. In addition, given that (μ_s, ν_s) must be located on the upper boundary of the feasible region, a negative value for either μ_s or ν_s implies that $|\mu_s - \nu_s|$ is relatively large, i.e. it implies imbalanced service rates. We can argue that imbalanced service rates are unlikely to be chosen by the decision-maker under certain plausible cases for the parameters of our problem, as described below:

(i) The nature of the objective function (1) implies that, under an optimal policy, the decision-maker should try to maintain an appropriate balance between the phase counts X_s and Y_s. Hence, imbalanced service rates will usually not be advantageous unless there is a significant imbalance between demand rates and/or service rate constraints for classes $\mathcal{A}$ and $\mathcal{D}$.

(ii) The incorporation of non-zero switching costs (i.e. setting $\beta > 0$) will tend to deter the decision-maker from making drastic changes to the service rates from one time interval to the next, thereby reducing the likelihood of a negative service rate being chosen.

(iii) If the variances of A_s and D_s are relatively small (which may happen if we are considering pre-scheduled arrival times, for example) then the variances of X_s and Y_s will also be small, further reducing the likelihood of the decision-maker needing to select imbalanced service rates.

To summarize the above, each of the following characteristics of our original problem should tend to increase the likelihood of optimal policies and optimal costs being similar for the original and surrogate problems: (i) consistently high demand-to-capacity ratios over all time intervals; (ii) similarity between

classes $\mathcal{A}$ and $\mathcal{D}$ with respect to time-varying demand rates and service rate constraints; (iii) non-zero switching costs; (iv) small variances for customer arrival times. <mark>Of course, these conditions may not necessarily be satisfied in real-world applications, but if they *are* satisfied then this should imply a stronger performance for the heuristics that we derive in Section 5. Numerical experiments can be used to illustrate the effects of varying demand rates, service rate constraints etc. (see Section 6).</mark>

In searching for an optimal policy which minimizes (13), a useful starting point is to find a *myopic* policy which operates by minimizing the expected single-step cost $g_s(\mathbf{x}, \mu_s)$ at each time interval s. In Section 4.1 we provide characterizations for myopic policies, and in Section 4.2 we examine the nature of optimal policies for our surrogate problem.

4.1   Myopic policies

Our assumptions regarding the set of linear constraints $\{\gamma_{js}\mu_s + \nu_s \leq \theta_{js}\}_{j=1}^{J_s}$ are the same as in the original version of the problem. Specifically, we assume that $\gamma_{1s} < \gamma_{2s} < ... < \gamma_{J_s,s}$ and that the intersection point between any pair of consecutive constraints $\gamma_{js}\mu_s + \nu_s = \theta_{js}$ and $\gamma_{j+1,s}\mu_s + \nu_s = \theta_{j+1,s}$, which we will denote by σ_{js}, lies within $[0, \theta_{J_s,s}/\gamma_{J_s,s}]$ (otherwise there would be at least one redundant constraint in the original problem). Using this notation, we can state that

$$\nu_s = \begin{cases} \theta_{1s} - \gamma_{1s}\mu_s, & \text{if } \mu_s \leq \sigma_{1s}, \\ \theta_{2s} - \gamma_{2s}\mu_s, & \text{if } \sigma_{1s} \leq \mu_s \leq \sigma_{2s}, \\ \vdots & \vdots \\ \theta_{J_s-1,s} - \gamma_{J_s-1,s}\mu_s, & \text{if } \sigma_{J_s-2,s} \leq \mu_s \leq \sigma_{J_s-1,s}, \\ \theta_{J_s,s} - \gamma_{J_s,s}\mu_s, & \text{if } \mu_s \geq \sigma_{J_s-1,s}. \end{cases} \tag{14}$$

We are seeking to find a myopic decision which minimizes the function $g_s(\mathbf{x}, \mu_s)$ defined in (12). For ease of notation, let the intervals $(-\infty, \sigma_{1s}), (\sigma_{1s}, \sigma_{2s}), ..., (\sigma_{J_s-2,s}, \sigma_{J_s-1,s}), (\sigma_{J_s-1,s}, \infty)$ be denoted by $\Omega_{1s}, \Omega_{2s}, ..., \Omega_{J_s,s}$ respectively. On each of the intervals Ω_{js}, we can check to see if $g_s(\mathbf{x}, \mu_s)$ attains a local minimum. First, evaluating the means and variances of X_s and Y_s, we have

$$\mathbb{E}[X_s|\mathbf{x}, \mu_s] = x + k(\mathbb{E}[A_s|\mathbf{x}] - \mu_s), \quad \mathbb{E}[Y_s|\mathbf{x}, \mu_s] = y + k(\mathbb{E}[D_s|\mathbf{x}] - (\theta_s - \gamma_s\mu_s)), \tag{15}$$

$$\text{Var}[X_s|\mathbf{x}, \mu_s] = k(k\text{Var}[A_s|\mathbf{x}] + |\mu_s|), \quad \text{Var}[Y_s|\mathbf{x}, \mu_s] = k(k\text{Var}[D_s|\mathbf{x}] + |\theta_s - \gamma_s\mu_s|). \tag{16}$$

Hence, evaluating the second moment of X_s, we have

$$\mathbb{E}[X_s^2|\mathbf{x}, \mu_s] = (\text{terms independent of } \mu_s) + k^2\mu_s^2 - k\mu_s\left(2x + 2k\mathbb{E}[A_s|\mathbf{x}]\right) + k|\mu_s|, \tag{17}$$

where $\mathbb{E}[A_s|\mathbf{x}]$ depends on the nature of the demand process for class $\mathcal{A}$; for example, in the case of Poisson arrivals with demand rate $\lambda(t)$, A_s is independent of $\mathbf{x}$ and we have $\mathbb{E}[A_s] = \int_{t_{s-1}}^{t_s} \lambda(u)du$. Similarly, given an interval on which $\nu_s = \theta_{js} - \gamma_{js}\mu_s$, we have

$$\mathbb{E}[Y_s^2|\mathbf{x}, \mu_s] = (\text{terms independent of } \mu_s) + k^2(\theta_{js} - \gamma_{js}\mu_s)^2 - k(\theta_{js} - \gamma_{js}\mu_s)\left(2y + 2k\mathbb{E}[D_s|\mathbf{x}]\right)$$
$$+ k|\theta_{js} - \gamma_{js}\mu_s|. \tag{18}$$

It follows from (17) and (18) that $g_s(\mathbf{x}, \mu_s)$ is a continuous, piecewise quadratic function of μ_s. The boundaries between the quadratic pieces include the vertices σ_{js} (for $j = 1, 2, ..., J_s$) and also the end-points of the interval $[0, \theta_{J_s,s}/\gamma_{J_s,s}]$ at which the service rates change sign. On each quadratic piece, we can check to see if a local minimum is attained. The expressions for these local minima depend on the dominant linear constraints in the regions in which they are found, and also on the signs of μ_s and ν_s in these regions. For example, if a local minimum is attained at some point $\mu_s \in (0, \theta_{J_s,s}/\gamma_{J_s,s})$ in which the jth constraint is dominant, then the expression for this local minimum is

$$\frac{2\alpha x - 2\gamma_{js}y + 2k(\alpha\mathbb{E}[A_s|\mathbf{x}] - \gamma_{js}\mathbb{E}[D_s|\mathbf{x}] + \gamma_{js}\theta_{js} + \beta\bar{\mu}) + \gamma_{js} - \alpha}{2k(\alpha + \gamma_{js}^2 + \beta)}. \tag{19}$$

Other possible cases are discussed in Appendix A.3. To simplify the search for a global minimizer of $g_s(\mathbf{x}, \mu_s)$ we provide the following result, which is proved in Appendix A.3.

Proposition 4.1. *The function $g_s(\mathbf{x}, \mu_s)$ is continuous and unimodal with respect to $\mu_s \in \mathbb{R}$. Hence, for each state $\mathbf{x} \in \mathcal{X}$, there exists $\mu_s^{\text{MYP}}(\mathbf{x}) \in \mathbb{R}$ such that $g_s(\mathbf{x}, \mu_s)$ is strictly decreasing for $\mu_s < \mu_s^{\text{MYP}}(\mathbf{x})$, and strictly increasing for $\mu_s > \mu_s^{\text{MYP}}(\mathbf{x})$.*

From Proposition 4.1 it follows that the service rate $\mu_s^{\text{MYP}}(\mathbf{x})$ which minimizes $g_s(\mathbf{x}, \mu_s)$ can be found using a simple procedure which requires a negligible amount of computation time. For each $j \in \{1, 2, ..., J_s\}$, let $\mu_{js}^{\text{MYP}}(\mathbf{x})$ denote the unique local minimum of $g_s(\mathbf{x}, \mu_s)$ if we assume that the relationship $\gamma_{js}\mu_s + \nu_s = \theta_{js}$ applies for all $\mu_s \in \mathbb{R}$ (it is proved in Appendix A.3 that this value $\mu_{js}^{\text{MYP}}(\mathbf{x})$ exists, although in certain cases it may be a non-differentiable point). The resulting value $\mu_{js}^{\text{MYP}}(\mathbf{x})$ may or may not be contained in Ω_{js}. If $\mu_{js}^{\text{MYP}}(\mathbf{x}) \in \Omega_{js}$ for some $j \in \{1, 2, ..., J_s\}$, then Proposition 4.1 implies that $\mu_{js}^{\text{MYP}}(\mathbf{x})$ is also the global optimizer $\mu_s^{\text{MYP}}(\mathbf{x})$. On the other hand, if $\mu_{js}^{\text{MYP}}(\mathbf{x}) \notin \Omega_{js}$ for all $j \in \{1, 2, ..., J_s\}$, then the global minimum is one of the vertices $\{\sigma_{1s}, ..., \sigma_{J_s,s}\}$, which we will denote by σ_s^*.

With this notation, we can conclude that the myopic service rate is given by

$$\mu_s^{\text{MYP}}(\mathbf{x}) = \begin{cases} \mu_{js}^{\text{MYP}}(\mathbf{x}), & \text{if } \exists \, j \in \{1, 2, ..., J_s\} \text{ such that } \mu_{js}^{\text{MYP}}(\mathbf{x}) \in \Omega_{js}, \\ \sigma_s^*, & \text{otherwise.} \end{cases} \tag{20}$$

4.2 Optimal policies

Although the surrogate problem that we are considering enables myopic service rates to be calculated quite easily, finding a policy π which minimizes the objective function (13) remains a challenging problem. Our approach in this section is to look for a simple approximation to $U(\mathbf{x})$ (specifically, a parametric function of the system state $\mathbf{x}$) and use this approximation to develop a heuristic approach.

In Section 4.2.1 we present a general result which holds for both types of arrivals distributions and arbitrary $\beta \geq 0$, provided that there is a single linear constraint in each time interval. Section 4.2.2 concerns a special case where $\beta = 0$ and a fixed constraint $\gamma\mu_s + \nu_s = \theta$ applies for all time intervals. The case of multiple linear constraints is discussed briefly in Section 4.2.3.

14

4.2.1 A single, time-varying linear constraint

Let us consider the case where a single linear constraint of the form $\gamma_s \mu_s + \nu_s \leq \theta_s$ applies at time interval $s \in \{1, 2, ..., S\}$. The Bellman equation for the surrogate problem can be written

$$U_s(\mathbf{x}) = \min_{\mu_s \in \mathbb{R}} \left\{ g_s(\mathbf{x}, \mu_s) + \mathbb{E}[U_{s+1}(\mathbf{X}_s)|\mathbf{x}, \mu_s] \right\} \qquad (s \in \{1, 2, ..., S\}), \tag{21}$$

where $U_1, ..., U_S$ are finite-stage value functions and $U_{S+1}(\mathbf{x}) := 0$ for all $\mathbf{x} \in \mathcal{X}$. We aim to find the minimizing value of μ_s in (21), assuming that the relationships $\nu_s = \theta_s - \gamma_s \mu_s$ hold for intervals s.

Let us begin by considering $U_S(\mathbf{x})$, the optimal return under state $\mathbf{x}$ with only one interval remaining until the end of the horizon. Naturally, a myopic choice of service rate is optimal at this interval. Although we are considering a single linear constraint, the cost function $g_S(\mathbf{x}, \mu_S)$ still exhibits piecewise quadratic behavior due to the nature of the variances in (16) as the service rates change sign (this is explained further in Appendix A.3). Ideally, we would like to express $U_S(\mathbf{x})$ as a simple quadratic function of the state variables, but the piecewise behavior of $g_S(\mathbf{x}, \mu_S)$ prevents this. However, we can derive a *lower bound* for $U_S(\mathbf{x})$ by omitting the modulus signs in (16) and writing

$$\begin{aligned} g_S(\mathbf{x}, \mu_S) \geq \alpha \big[& k^2 \mu_S^2 - k\mu_S \left(2x + 2k\mathbb{E}[A_S|\mathbf{x}] - 1 \right) \\ & + k^2(\theta_S - \gamma_S \mu_S)^2 - k(\theta_S - \gamma_S \mu_S)\left(2y + 2k\mathbb{E}[D_S|\mathbf{x}] - 1 \right) \\ & + \beta k^2 (\mu_S - \bar{\mu})^2 + (\text{terms independent of } \mu_S). \end{aligned} \tag{22}$$

Equality holds in (22) if and only if μ_S and $\nu_S = \theta_S - \gamma_S \mu_S$ are both non-negative, which (as we have argued earlier) is likely to be the case for many states $\mathbf{x} \in \mathcal{X}$ under plausible modelling assumptions. Let us consider the random variables A_S and D_S. In the case of Poisson arrivals, these are actually independent of the system state $\mathbf{x}$. In the case of pre-scheduled arrivals, A_S and D_S depend on $\mathbf{x}$, but it is useful to note that $\mathbb{E}[A_S|\mathbf{x}]$ and $\mathbb{E}[D_S|\mathbf{x}]$ are linear in their dependencies on the state variables z and w (which represent the numbers of 'imminent arrivals') respectively. Indeed, this linearity property holds for arbitrary $s \in \{1, 2, ..., S\}$, since we can write

$$\mathbb{E}[A_s|\mathbf{x}] = zp_0 + \sum_{j=1}^{S-s} L_{s+j} p_j, \quad \text{Var}[A_s|\mathbf{x}] = zp_0(1 - p_0) + \sum_{j=1}^{S-s} L_{s+j} p_j(1 - p_j), \tag{23}$$

$$\mathbb{E}[D_s|\mathbf{x}] = wq_0 + \sum_{j=1}^{S-s} M_{s+j} q_j, \quad \text{Var}[D_s|\mathbf{x}] = wq_0(1 - q_0) + \sum_{j=1}^{S-s} M_{s+j} q_j(1 - q_j), \tag{24}$$

where L_j (resp. M_j) is the number of customers of class $\mathcal{A}$ (resp. $\mathcal{D}$) scheduled to arrive at the beginning of interval j, and p_j (resp. q_j) is the probability that a customer of class $\mathcal{A}$ (resp. $\mathcal{D}$) scheduled to arrive at the beginning of interval $s + j$ arrives during interval s. (Note: this notation makes sense when $j = 0$ since, due to the memoryless property, we can treat all customers who have not yet arrived by interval s as if they were scheduled to arrive at interval s.) For simplicity we have assumed in (23)-(24) that all customers scheduled to arrive during interval j have the same scheduled arrival time (this might be at the beginning of interval j, for example), but these expressions can easily be modified if this is not the case.

Minimizing the right-hand side of (22) over $\mu_S \in \mathbb{R}$ yields a local minimum at

$$\mu_S = \frac{2\alpha x - 2\gamma_S y + 2k(\alpha\mathbb{E}[A_S|\mathbf{x}] - \gamma_S\mathbb{E}[D_S|\mathbf{x}] + \gamma_S\theta_S + \beta\bar{\mu}) + \gamma_S - \alpha}{2k(\alpha + \gamma_S^2 + \beta)}, \tag{25}$$

and with this value of μ_S, the right-hand side of (22) can be expressed as

$$U_S^{\mathrm{LB}}(\mathbf{x}) := \mathbf{x}^{\mathrm{T}}\mathbf{P}_S\mathbf{x} + \mathbf{q}_S^{\mathrm{T}}\mathbf{x} + r_S, \tag{26}$$

where $\mathbf{P}_S$ is a matrix of order $n \times n$ (where n is the number of state variables), $\mathbf{q}_S \in \mathbb{R}^n$ and $r_S \in \mathbb{R}$. Importantly, $\mathbf{P}_S$, $\mathbf{q}_S$ and r_S do not have any dependence on $\mathbf{x}$ or μ_s. We can infer from (26) that $U_S^{\mathrm{LB}}(\mathbf{x})$ is a quadratic function of the state variables. Note that the number of state variables n can be 2, 3, 4 or 5, depending on the type of arrivals distribution and the value of β. For example, in the case of Poisson arrivals with $\beta = 0$, we have $\mathbf{x} = (x, y) \in \mathbb{Z}^2$ and it can be shown that

$$\mathbf{P}_S = \frac{\alpha}{\alpha + \gamma_S^2}\begin{bmatrix} \gamma_S^2 & \gamma_S \\ \gamma_S & 1 \end{bmatrix}, \quad \mathbf{q}_S = \frac{1}{\alpha + \gamma_S^2}\begin{bmatrix} \alpha(\alpha - \gamma_S) + 2k\alpha\gamma_S(\gamma_S\mathbb{E}[A_S|\mathbf{x}] + \mathbb{E}[D_S|\mathbf{x}] - \theta_S) \\ \gamma_S(\gamma_S - \alpha) + 2k\alpha(\gamma_S\mathbb{E}[A_S|\mathbf{x}] + \mathbb{E}[D_S|\mathbf{x}] - \theta_S) \end{bmatrix}.$$

(We have omitted the expression for r_S because it is too long.) Evidently, $U_S^{\mathrm{LB}}(\mathbf{x})$ is a lower bound for $U_S(\mathbf{x})$ since the latter is obtained by minimizing $g_S(\mathbf{x}, \mu_S)$, whereas the former is obtained by minimizing an expression which is bounded above by $g_S(\mathbf{x}, \mu_S)$. Let us proceed by obtaining lower bounds for $U_s(\mathbf{x})$ for each $s \in \{1, 2, ..., S - 1\}$. We propose the following lemma.

Lemma 4.2. *Suppose that, for a particular interval $s \in \{1, 2, ..., S - 1\}$, there exists a function $U_{s+1}^{\mathrm{LB}} : \mathcal{X} \to \mathbb{R}$ such that*

$$U_{s+1}^{\mathrm{LB}}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}_{s+1}\mathbf{x} + \mathbf{q}_{s+1}^{\mathrm{T}}\mathbf{x} + r_{s+1} \leq U_{s+1}(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X}, \tag{27}$$

where $\mathbf{P}_{s+1}$ is a real-valued matrix of order $n \times n$, $\mathbf{q}_{s+1} \in \mathbb{R}^n$, $r_{s+1} \in \mathbb{R}$ and n is the number of state variables. Let $U_s^{\mathrm{LB}} : \mathcal{X} \to \mathbb{R}$ be defined by

$$U_s^{\mathrm{LB}}(\mathbf{x}) = \min_{\mu_s \in \mathbb{R}} \left\{ g_s(\mathbf{x}, \mu_s) + \psi_s(\mu_s) + \mathbb{E}[U_{s+1}^{\mathrm{LB}}(\mathbf{X}_s)|\mathbf{x}, \mu_s] \right\}, \tag{28}$$

where

$$\psi_s(\mu_s) = \begin{cases} 2k(\alpha + \tilde{A}_{s+1})\mu_s, & \text{if } \mu_s < 0, \\ 2k(1 + \tilde{B}_{s+1})(\theta_s - \gamma_s\mu_s), & \text{if } \mu_s > \theta_s/\gamma_s, \\ 0, & \text{otherwise,} \end{cases} \tag{29}$$

and $\tilde{A}_{s+1}$ and $\tilde{B}_{s+1}$ are the coefficients of x^2 and y^2 respectively in the expression for $U_{s+1}^{\mathrm{LB}}(\mathbf{x})$. Then there exists a real-valued $n \times n$ matrix $\mathbf{P}_s$, a vector $\mathbf{q}_s \in \mathbb{R}^n$ and a constant $r_s \in \mathbb{R}$ such that

$$U_s^{\mathrm{LB}}(\mathbf{x}) = \mathbf{x}^{\mathrm{T}}\mathbf{P}_s\mathbf{x} + \mathbf{q}_s^{\mathrm{T}}\mathbf{x} + r_s \leq U_s(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X}. \tag{30}$$

Proof of the lemma can be found in Appendix A.4. We note that $\psi_s(\mu_s)$ is a piecewise linear function which 'compensates' for the fact that $\mathrm{Var}[X_s|\mathbf{x}, \mu_s]$ and $\mathrm{Var}[Y_s|\mathbf{x}, \mu_s]$ are not linear functions of μ_s. If

either μ_s or $\nu_s = \theta_s - \gamma_s \mu_s$ is negative then $\psi_s(\mu_s)$ takes an artificial non-zero value in order to ensure that a quadratic representation of $U_s^{\text{LB}}(\mathbf{x})$ is still possible.

By using the result of Lemma 4.2 iteratively, we can find suitable values $\mathbf{P}_s$, $\mathbf{q}_s$ and r_s satisfying (30) for each $s \in \{1, 2, ..., S\}$. We can then write

$$U_s^{\text{LB}}(\mathbf{x}) \leq U_s(\mathbf{x})$$

for any state $\mathbf{x} \in \mathcal{X}$ and $s \in \{1, 2, ..., S\}$. Generally it is not worthwhile to write expressions for $\mathbf{P}_s$, $\mathbf{q}_s$ and r_s in terms of the system parameters, since these expressions are too convoluted. However, an interesting special case occurs when $\beta = 0$ and the OTE is the same for each time interval, so that $\gamma_s = \gamma$ and $\theta_s = \theta$ for each $s \in \{1, 2, ..., S\}$. This is the subject of the next subsection.

4.2.2 A single, fixed linear constraint with $\beta = 0$

The following lemma states what may be proved in this special case.

Lemma 4.3. *Suppose $\beta = 0$ and we have a single linear constraint $\gamma \mu_s + \nu_s \leq \theta$ for each $s \in \{1, 2, ..., S\}$ (where $\gamma, \theta > 0$). Then, for a particular slot $s \in \{1, 2, ..., S\}$:*

(a) In the case of Poisson arrivals, we have

$$U_s^{\text{LB}}(\mathbf{x}) = \tilde{A}_s(\gamma x + y)^2 + \tilde{B}_s(\gamma x + y) + \tilde{C}_s(\alpha x - \gamma y) + \tilde{D}_s \qquad \forall \mathbf{x} \in \mathcal{X}, \tag{31}$$

where $\tilde{A}_s$, $\tilde{B}_s$, $\tilde{C}_s$ and $\tilde{D}_s$ are constants which depend on the system parameters. In particular, the values of $\tilde{A}_s$, $\tilde{B}_s$ and $\tilde{C}_s$ are given by

$$\tilde{A}_s = \frac{(S-s+1)\alpha}{\alpha + \gamma^2}, \qquad \tilde{B}_s = \frac{2k\alpha \sum_{j=s}^{S}(S-j+1)(\gamma \mathbb{E}[A_j] + \mathbb{E}[D_j] - \theta)}{\alpha + \gamma^2},$$

$$\tilde{C}_s = \frac{\alpha^2 - \gamma^3 + (S-s+1)\alpha\gamma(\gamma-1)}{(\alpha + \gamma^2)^2}. \tag{32}$$

(b) In the case of pre-scheduled arrivals, we have

$$U_s^{\text{LB}}(\mathbf{x}) = \tilde{A}_s(\gamma x + y)^2 + \tilde{C}_s(\alpha x - \gamma y) + (\tilde{F}_s + \tilde{G}_s z + \tilde{H}_s w)(\gamma x + y)$$
$$+ \tilde{J}_s z^2 + \tilde{K}_s w^2 + \tilde{L}_s zw + \tilde{M}_s z + \tilde{P}_s w + \tilde{Z}_s \qquad \forall \mathbf{x} \in \mathcal{X}, \tag{33}$$

where $\tilde{A}_s$, $\tilde{C}_s$, $\tilde{F}_s$, $\tilde{G}_s$, $\tilde{H}_s$, $\tilde{J}_s$, $\tilde{K}_s$, $\tilde{L}_s$, $\tilde{M}_s$, $\tilde{P}_s$ and $\tilde{Z}_s$ are constants which depend on the system parameters. In particular, $\tilde{A}_s$ and $\tilde{C}_s$ have the same values as in (32).

Furthermore, the service rate $\mu_s^{\text{OPT}}(\mathbf{x})$ which attains the minimum in (28) is

$$\mu_s^{\text{OPT}}(\mathbf{x}) = \begin{cases} \mu_s^{\text{MYP}}(\mathbf{x}) + \delta, & \text{if } s < S, \\ \mu_s^{\text{MYP}}(\mathbf{x}), & \text{if } s = S, \end{cases} \tag{34}$$

where

$$\delta = \frac{\gamma(1-\gamma)\tilde{A}_{s+1} + (\alpha + \gamma^2)\tilde{C}_{s+1}}{2k(\alpha + \gamma^2)} = \frac{\alpha^2 - \gamma^3}{2k(\alpha + \gamma^2)^2} \tag{35}$$

and $\mu_s^{\mathrm{MYP}}(\mathbf{x})$ is the myopic service rate:

$$\mu_s^{\mathrm{MYP}}(\mathbf{x}) = \frac{2\alpha x - 2\gamma y + \gamma - \alpha + 2k(\alpha \mathbb{E}[A_s|\mathbf{x}] - \gamma \mathbb{E}[D_s|\mathbf{x}] + \gamma\theta)}{2k(\alpha + \gamma^2)}. \tag{36}$$

The proof of Lemma 4.3 can be found in Appendix A.6. If we consider a problem in which the cost function $g_s(\mathbf{x}_s, \mu_s)$ is replaced by $g_s(\mathbf{x}_s, \mu_s) + \psi_s(\mu_s)$, then (35) provides us with a simple condition for the optimality of a myopic policy. We state this below as a corollary.

Corollary 4.4. *Consider a decision problem in which we aim to minimize*

$$U^{\mathrm{LB}}(\mathbf{x}) = \mathbb{E}^\pi \left[\sum_{s=1}^S (g_s(\mathbf{x}_s, \mu_s) + \psi_s(\mu_s)) \,\middle|\, \mathbf{x}_0 = \mathbf{x} \right], \tag{37}$$

where π is a policy for choosing the service rates $\mu_1, \mu_2, ..., \mu_S \in \mathbb{R}$. We assume a single constraint $\gamma\mu_s + \nu_s \leq \theta$ applies for all $s \in \{1, 2, ..., S\}$ and $\beta = 0$. Then a myopic policy is optimal if and only if $\alpha^2 = \gamma^3$.

Proof. The value functions for the problem described in the corollary are $U_s^{\mathrm{LB}}(\mathbf{x})$, defined recursively via (28). By Lemma 4.3, we see that an optimal policy (with respect to $U_s^{\mathrm{LB}}(\mathbf{x})$) differs from a myopic policy by adding a constant term $(\alpha^2 - \gamma^3)/(2k(\alpha + \gamma^2)^2)$ to the myopic service rate, and hence the myopic and optimal policies coincide if and only if $\alpha^2 = \gamma^3$. $\square$

It is important to note that the relationship $\alpha^2 = \gamma^3$ holds if $\alpha = \gamma = 1$, in which case the two classes of customer are equivalent with respect to both congestion charges and resource consumption rates. This might apply to many types of application in which there is no obvious reason to treat one class of customer differently from another. The fixed service rate adjustment δ increases with α but decreases with γ, which is logical given the physical interpretations of these parameters.

Although Corollary 4.4 is related to $U^{\mathrm{LB}}(\mathbf{x})$ rather than our original value function $U(\mathbf{x})$, the differences $U(\mathbf{x}) - U^{\mathrm{LB}}(\mathbf{x})$ should be relatively small if (i) the optimal policy for the surrogate problem rarely chooses negative service rates, and (ii) phase counts for the surrogate problem rarely become negative. At the beginning of Section 4 we suggested conditions under which these effects should be observed.

In Appendix A.5 we have provided a numerical example in which customers of classes $\mathcal{A}$ and $\mathcal{D}$ arrive according to Poisson processes with complementary sinusoidal demand rates. Our example considers the case $\alpha = \gamma = 1$ and uses Lemma 4.3 to compute the exact values of $U_s^{\mathrm{LB}}(\mathbf{x})$ for $s \in \{1, 2, ..., S\}$. Estimated upper bounds for $U(\mathbf{x})$ are then obtained by simulating the performance of a myopic policy in the surrogate problem. Since the optimal value $U(\mathbf{x})$ is (by definition) bounded above by the value $U^{\mathrm{MYP}}(\mathbf{x})$ obtained under a myopic policy, the estimates of $U^{\mathrm{MYP}}(\mathbf{x}) - U^{\mathrm{LB}}(\mathbf{x})$ obtained via simulation can be interpreted as estimated upper bounds for $U(\mathbf{x}) - U^{\mathrm{LB}}(\mathbf{x})$. Our results indicate that

- The difference $U^{\mathrm{MYP}}(\mathbf{x}) - U^{\mathrm{LB}}(\mathbf{x})$ tends to increase as demand rates become more nonstationary, and also increases as the demand-to-capacity ratio increases.

- However, even in cases where total demand consistently exceeds capacity by as much as 100%, $U^{\mathrm{MYP}}(\mathbf{x}) - U^{\mathrm{LB}}(\mathbf{x})$ is generally smaller than 0.25%.

We also repeated our experiments with different values of α and γ, with similar results (please refer to Appendix A.5 for full details). The second of the bulleted points above is especially relevant in the context of this paper, since in problems where demand greatly exceeds capacity, we can expect the dynamics of our surrogate problem to closely resemble those of the original problem. We therefore have some evidence to suggest that, in the case of a single, fixed linear constraint with $\beta = 0$, myopic policies are very close to optimal in the surrogate problem - and this should also apply to the original problem if demand rates are sufficiently high. Our numerical results in Section 6 will demonstrate the effects of time-varying service rate constraints and positive β values on the performances of myopic policies.

We will provide one further result here which concerns a 'deterministic fluid flow analogue' of our original optimization problem, in which queueing dynamics are deterministic and the workloads in the two queues are represented as continuous variables. Although this fluid flow model represents a significant departure from our original stochastic optimization problem, it does yield a condition for equivalence between the fluid flow analogues of $U(\mathbf{x})$ and $V(\mathbf{x})$ which may offer a useful practical guide for decision-making in the stochastic problem. Consider a variant of the problem formulated in Section 2 in which the state space is $\mathbb{R}^2_{\geq 0}$ and the phase counts at time $t \in [t_{s-1}, t_s)$ during interval $s \in \{1, 2, ..., S\}$ are

$$X(t) = x + k \int_{t_{s-1}}^{t} (\lambda(u) - I(X(u) > 0)\mu_s)\, du, \quad Y(t) = y + k \int_{t_{s-1}}^{t} (\eta(u) - I(Y(u) > 0)\nu_s)\, du,$$

where (x, y) is the state at the beginning of interval s, $I(\cdot)$ denotes the indicator function, $\lambda(t)$ and $\eta(t)$ are continuous flow rates and μ_s and ν_s are the chosen service rates for interval s, which are both non-negative and satisfy $\gamma\mu_s + \nu_s = \theta$. The phase counts at the end of interval s are $X_s = X(t_s)$ and $Y_s = Y(t_s)$. The value function for this problem will be denoted by $V^{\mathrm{FF}}(\mathbf{x})$, where

$$V^{\mathrm{FF}}(\mathbf{x}) = \min_{\pi} \sum_{s=1}^{S} \left[\alpha X_s^2 + Y_s^2 | \mathbf{x}, \pi \right]$$

and π is a decision-making policy which, in this deterministic setting, is simply a vector of service rates. In the 'surrogate' version of our fluid flow problem, we do not enforce non-negativity for either the service rates or the phase counts and we replace the definitions of $X(t)$ and $Y(t)$ above by

$$X(t) = x + k \int_{t_{s-1}}^{t} (\lambda(u) - \mu_s)du, \quad Y(t) = y + k \int_{t_{s-1}}^{t} (\eta(u) - \nu_s)du,$$

with $U^{\mathrm{FF}}(\mathbf{x})$ used instead of $V^{\mathrm{FF}}(\mathbf{x})$ to represent the corresponding value function.

Proposition 4.5. *In the deterministic fluid flow problem:*

(a) The finite-stage value functions $U_1^{\mathrm{FF}}, ..., U_S^{\mathrm{FF}}$ satisfying the Bellman equations

$$U_s^{\mathrm{FF}}(\mathbf{x}) = \min_{\mu_s \in \mathbb{R}} \left[\alpha X_s^2 + Y_s^2 + U_{s+1}^{\mathrm{FF}}(\mathbf{X}_s) | \mathbf{x}, \mu_s \right] \tag{38}$$

19

(with $U_{S+1}^{\mathrm{FF}}(\mathbf{x}) := 0$ for all $\mathbf{x} \in \mathcal{X}$) are given by

$$U_s^{\mathrm{FF}}(\mathbf{x}) = \tilde{A}_s(\gamma x + y)^2 + \tilde{B}_s(\gamma x + y) + \tilde{D}_s \quad \forall \mathbf{x} \in \mathcal{X}, \ s \in \{1, 2, ..., S\},$$

where $\tilde{A}_s$ and $\tilde{B}_s$ have the same definitions as in (32) and $\tilde{D}_s$ is constant. For each interval $s \in \{1, 2, ..., S\}$, the myopic and optimal service rates for the surrogate problem are given by

$$\mu_s^{\mathrm{OPT}}(\mathbf{x}) = \mu_s^{\mathrm{MYP}}(\mathbf{x}) = \frac{\alpha x - \gamma y + k \int_{t_{s-1}}^{t_s} (\alpha \lambda(u) - \gamma \eta(u)) du + k \gamma \theta}{k(\alpha + \gamma^2)}. \tag{39}$$

(b) Suppose the system is initialized in a state $\mathbf{x} = (x_0, y_0)$ which satisfies

$$0 \le \frac{\alpha x_0 - \gamma y_0 + \alpha \int_{t_0}^{t_1} \lambda(u) du - \gamma \int_{t_0}^{t_1} \eta(u) du + \gamma \theta}{\alpha + \gamma^2} \le \frac{\theta}{\gamma}, \tag{40}$$

and also suppose that

$$0 \le \frac{\alpha \int_{t_{s-1}}^{t_s} \lambda(u) du - \gamma \int_{t_{s-1}}^{t_s} \eta(u) du + \gamma \theta}{\alpha + \gamma^2} \le \frac{\theta}{\gamma} \tag{41}$$

for each $s \in \{2, ..., S\}$. Let $n_0 = \min\{x_0, y_0\}$. Then

$$\lim_{n_0 \to \infty} |U^{\mathrm{FF}}(\mathbf{x}) - V^{\mathrm{FF}}(\mathbf{x})| = 0$$

and the service rates in (39) are optimal in both versions of the fluid flow problem.

Proposition 4.5, which is proved in Appendix A.7, provides us with conditions involving the demand rates and capacity envelope parameters which ensure that the simple myopic policy represented by (39) is also an optimal policy for both the surrogate problem and the original problem. Although the result concerns a modified version of the problem with deterministic fluid flow dynamics, the conditions (40) and (41) (or similar conditions adapted to the case of pre-scheduled arrivals) might also imply a strong performance for the aforementioned policy in the original stochastic problem, especially if the variances for end-of-interval phase counts X_s and Y_s are small.

4.2.3 Multiple linear constraints

We now discuss the case of multiple linear constraints. As before, let J_s be the number of constraints in interval s, so that the jth constraint (for $j \in \{1, 2, ..., J_s\}$) has the form $\gamma_{js}\mu_s + \nu_s \le \theta_{js}$. Also, let $(n_s, n_{s+1}, ..., n_S)$ be an $(S - s + 1)$-tuple, such that $n_i \in \{1, 2, ..., J_i\}$ for $i \in \{s, s+1, ..., S\}$. The tuple $(n_s, n_{s+1}, ..., n_S)$ represents a sequence of individual constraints 'chosen' from the constraint sets for intervals $s, s + 1, ..., S$. Given a particular tuple $(n_s, n_{s+1}, ..., n_S)$ we can obtain a quadratic function $U_{(n_s, n_{s+1}, ..., n_S)}^{\mathrm{LB}}(\mathbf{x})$ such that

$$U_{(n_s, n_{s+1}, ..., n_S)}^{\mathrm{LB}}(\mathbf{x}) \le U_s(\mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X}$$

by assuming that the relationship $\nu_i = \theta_{n_i, i} - \gamma_{n_i, i}\mu_i$ holds for all $\mu_i \in \mathbb{R}$ in interval $i \in \{s, s+1, ..., S\}$ and then constructing the quadratic function $U_{(n_s, n_{s+1}, ..., n_S)}^{\mathrm{LB}}(\mathbf{x})$ in exactly the same way as described in Section 4.2.2, replacing single-constraint parameters such as γ_s and θ_s by γ_{js} and θ_{js} as appropriate. Conceptually,

for each of the intervals $i \in \{s, s+1, ..., S\}$, we are drawing a straight line over the top of the OTE for interval i and deriving a lower bound by supposing that the decision-maker is allowed to choose any point on this line. However, the lower bound $U_{(n_s,n_{s+1},...,n_S)}^{\mathrm{LB}}(\mathbf{x})$ obtained in this way is unlikely to be very tight if there are many constraints, since the local minima found on the unbounded straight lines will rarely satisfy the constraints for the actual surrogate problem. A tighter lower bound is given by

$$U_s^{\mathrm{MLB}}(\mathbf{x}) := \max_{(n_s,n_{s+1},...,n_S)} U_{(n_s,n_{s+1},...,n_S)}^{\mathrm{LB}}(\mathbf{x}),$$

but this is a piecewise quadratic function of $\mathbf{x}$ which cannot be computed easily due to the large number of possible tuples $(n_s, n_{s+1}, ..., n_S)$ that one must consider as $S - s$ increases.

A computationally feasible algorithm would need to be based on a simple representation of $U_s^{\mathrm{LB}}(\mathbf{x})$, similar to the quadratic functions that we were able to obtain in the case of single linear constraints. In Section 5 we introduce a number of heuristic policies, two of which are based on the idea of approximating the value functions $V_s(\mathbf{x})$ using quadratic functions of the state variables.

5 Design of heuristics

In this section we introduce 4 different heuristics. All of these are dynamic decision-making policies which choose a service rate μ_s for class $\mathcal{A}$ in response to an observed system state $\mathbf{x} \in \mathcal{X}$ at time interval s (with the service rate for class $\mathcal{D}$, ν_s, being given by the OTE for interval s). We present the 4 heuristics in increasing order of the amount of computational effort required.

All of the heuristics that we develop in this section are based on optimal (or near-optimal) solutions to the *surrogate* problem described in Section 4, in which negative service rates and phase counts are permitted. However, in Section 6 we test the heuristics by applying them to the original problem formulated in Section 2 (which does not permit negative service rates or phase counts).

1. Myopic policy (MYP)

Under this policy, we choose the service rate $\mu_s^{\mathrm{MYP}}(\mathbf{x})$ which minimizes the expected single-stage cost at the end of interval s in the surrogate problem. As discussed in Section 4.1, $\mu_s^{\mathrm{MYP}}(\mathbf{x})$ can be found by searching for a local minimum of the function $g_s(\mathbf{x}, \mu_s)$ in each of the intervals Ω_{js} (for $j = 1, 2, ..., J_s$), and choosing an appropriate boundary point if no local minimum is found.

When implementing this heuristic in the original problem, we will need to enforce non-negativity by choosing $\mu_s = 0$ if the myopic service rate $\mu_s^{\mathrm{MYP}}(\mathbf{x})$ defined in (20) is negative, or $\mu_s = \theta_{J_s,s}/\gamma_{J_s,s}$ if $\mu_s^{\mathrm{MYP}}(\mathbf{x})$ is greater than $\theta_{J_s,s}/\gamma_{J_s,s}$ (in which case we need to avoid a negative ν_s value). Due to the unimodality of $g_s(\mathbf{x}, \mu_s)$ proved by Proposition 4.1, it follows that this 'truncated' version of the myopic policy does indeed minimize $g_s(\mathbf{x}, \mu_s)$ if the extra constraints $\mu_s, \nu_s \geq 0$ are enforced.

2. Approximate Policy Improvement (API)

This heuristic offers a computationally inexpensive method of improving upon a myopic policy. The principle is that, at time interval $s \in \{1, 2, ..., S-1\}$ and state $\mathbf{x} \in \mathcal{X}$, we choose the service rate μ_s which

minimizes the *estimated* two-step return

$$g_s(\mathbf{x}, \mu_s) + g_{s+1}\big(\mathbb{E}[\mathbf{X}_s|\mathbf{x}, \mu_s],\ \mu_{s+1}^{\mathrm{MYP}}(\mathbb{E}[\mathbf{X}_s|\mathbf{x}, \mu_s])\big), \tag{42}$$

where $\mathbb{E}[\mathbf{X}_s|\mathbf{x}, \mu_s]$ is the expected state at the end of interval s given that we choose service rate μ_s. Effectively, we are relying upon a forecast of the next state in order to predict the service rate $\mu_{s+1}^{\mathrm{MYP}}(\mathbb{E}[\mathbf{X}_s|\mathbf{x}, \mu_s])$ chosen by the myopic policy at the beginning of interval $s + 1$, rather than taking into account the full distribution of $\mu_{s+1}^{\mathrm{MYP}}(\mathbf{X}_s)$ as a function of the random phase counts X_s and Y_s (which would require a prohibitive amount of computational effort). This is why we describe the heuristic as *approximate* policy improvement. For the final time interval S, a myopic service rate can be chosen. We also use the same truncation method as in the myopic policy, so that $\mu_s, \nu_s \in [0, \theta_{J_s,s}/\gamma_{J_s,s}]$.

In principle, we could extend this idea and obtain an approximate m-step improvement policy by calculating the best service rate for interval s under the assumption that myopic service rates are chosen for intervals $s+1, s+2, ..., s+m$. This method would also rely upon predictions of the myopic service rates for future time intervals based on expected values of the phase counts. However, our experiments have shown that any benefit obtained by implementing m-step improvement as opposed to one-step improvement is usually very slight, and in some cases it actually performs worse - perhaps because the forecasts for future phase counts become less reliable as the 'look-ahead distance' increases.

3. Coefficient Averaging (CAV)

This heuristic is motivated by the discussion in Section 4.2.3 and is intended as a fairly naive but inexpensive method of acquiring quadratic approximations for the value functions in a problem with multiple linear constraints. It uses backwards induction, starting from the final time interval S. For each of the linear constraints $\gamma_{jS}\mu_S + \nu_S \leq \theta_{jS}$ effective at interval S, we construct the function

$$U_{jS}^{\mathrm{LB}}(\mathbf{x}) := \min_{\mu_S \in \mathbb{R}} \left\{ g_{jS}(\mathbf{x}, \mu_S) + \psi_{jS}(\mu_S) \right\},$$

where $g_{jS}(\mathbf{x}, \mu_S)$ is similar to the usual single-step expected cost $g_S(\mathbf{x}, \mu_S)$ except that it assumes the relationship $\nu_S = \theta_{jS} - \gamma_{jS}\mu_S$ holds for all $\mu_S \in \mathbb{R}$, and $\psi_{jS}(\mu_S)$ is similar to the function $\psi_S(\mu_S)$ defined in (29) except that γ_S and θ_S are replaced by γ_{jS} and θ_{jS} respectively. The analyses in Section 4.2.1 imply that, for each $j \in \{1, 2, ..., J_S\}$, $U_{jS}^{\mathrm{LB}}(\mathbf{x})$ is a quadratic function of $\mathbf{x}$.

We then create a new function $U_S^{\mathrm{CAV}}(\mathbf{x})$ by considering each of the variables x^2, y^2, xy etc. individually and, for each variable, averaging the relevant coefficients which appear in the functions $U_{jS}^{\mathrm{LB}}(\mathbf{x})$. For example, if we are considering Poisson arrivals with $\beta = 0$, then (as shown in Section 4.2.1) the functions $U_{jS}^{\mathrm{LB}}(\mathbf{x})$ are linear combinations of the variables

$$x^2,\ y^2,\ xy,\ x,\ y.$$

We therefore form a new function $U_S^{\mathrm{CAV}}(\mathbf{x})$ which is also a linear combination of the above variables. The coefficient of x^2 in $U_S^{\mathrm{CAV}}(\mathbf{x})$ is the arithmetic mean of the coefficients of x^2 in the functions $U_{1S}^{\mathrm{LB}}(\mathbf{x}), U_{2S}^{\mathrm{LB}}(\mathbf{x}), ..., U_{J_S,S}^{\mathrm{LB}}(\mathbf{x})$, and similarly for the coefficients of y^2, xy, etc. Then, in order to obtain the

function $U_{S-1}^{CAV}(\mathbf{x})$, we start by obtaining the functions $U_{j,S-1}^{LB}(\mathbf{x})$ defined by

$$U_{j,S-1}^{LB}(\mathbf{x}) = \min_{\mu_{S-1} \in \mathbb{R}} \left\{ g_{j,S-1}(\mathbf{x}, \mu_{S-1}) + \psi_{j,S-1}(\mu_{S-1}) + \mathbb{E}[U_S^{CAV}(\mathbf{X}_{S-1})|\mathbf{x}, \mu_{S-1}] \right\}.$$

As demonstrated by the analyses in Section 4.2.1, the functions $U_{j,S-1}^{LB}(\mathbf{x})$ are linear combinations of the same variables as the functions $U_{jS}^{LB}(\mathbf{x})$. We can therefore obtain a function $U_{S-1}^{CAV}(\mathbf{x})$ using the same coefficient-averaging procedure as that used for $U_S^{CAV}(\mathbf{x})$, and we repeat these steps to obtain functions $U_{S-2}^{CAV}(\mathbf{x}), U_{S-3}^{CAV}(\mathbf{x}), ..., U_1^{CAV}(\mathbf{x})$. The CAV heuristic chooses the service rate $\mu_s^{CAV}(\mathbf{x}) \in [0, \theta_{J_s,s}/\gamma_{J_s,s}]$ which attains the minimum in the expression for $U_s^{CAV}(\mathbf{x})$.

4. Least-squares Fitting (LSF)

This heuristic, like the previous one, is based on backwards induction and the use of a parametric function of the system state to approximate the value functions $V_s(\mathbf{x})$. In this case we assume, *a priori*, that $V_s(\mathbf{x})$ can be well-approximated by a quadratic function of the state variables and use least-squares fitting (after 'sampling' the values at a selection of states) to find suitable values of the coefficients that appear in this quadratic function. The steps are as follows:

1. Set $s = S$ and $\hat{V}_{S+1}(\mathbf{x}) = 0$ for all states $\mathbf{x} \in \mathcal{X}$.

2. Select a certain number of system states in $\mathcal{X}$ either randomly or systematically (choosing a large number of states should yield a stronger performance at the expense of greater computation time). Let $M_s \subset \mathcal{X}$ denote the set of states chosen.

3. For each $\mathbf{x} \in M_s$, calculate the value

$$W_s(\mathbf{x}) := \min_{\mu_s \in \Theta_s} \left\{ g_s(\mathbf{x}, \mu_s) + \mathbb{E}[\hat{V}_{s+1}(\mathbf{X}_s)|\mathbf{x}, \mu_s] \right\}. \tag{43}$$

4. Use least-squares regression to minimize

$$\sum_{\mathbf{x} \in M_s} \left(W_s(\mathbf{x}) - Q_s(\mathbf{x}) \right)^2 \tag{44}$$

where $Q_s(\mathbf{x})$ is quadratic in the state variables and the minimization is carried out with respect to the coefficients in $Q_s(\mathbf{x})$. Let $\hat{V}_s(\mathbf{x})$ denote the function $Q_s(\mathbf{x})$ which minimizes (44).

5. Reduce s by 1. If $s \geq 1$ return to step 2; otherwise, stop.

The LSF heuristic chooses the service rate $\mu_s^{LSF}(\mathbf{x}) \in \Theta_s$ which attains the minimum in (43).

6   Numerical experiments

In this section we evaluate the performances of the heuristic policies proposed in Section 5 using results gathered from numerical experiments with randomly-generated system parameters. Noting that we cannot feasibly compute optimal solutions to the optimization problem formulated in Section 2, we instead take the approach of comparing the performances of our SAB-based heuristics with that of a more naive

'baseline' decision-making policy which might be employed by a decision-maker in practice. In Section 6.1 we describe our baseline heuristic. The majority of our numerical results are presented in Section 6.2, and in Section 6.3 we investigate the effects of our heuristics on customer waiting times.

6.1 A naive baseline heuristic

The naive heuristic that we introduce here, which we will refer to as a "myopic demand ratio" (MDR) heuristic, is based on a measure of the demand placed on the server by customers of class $\mathcal{A}$ as a proportion of the total demand imposed by both customer classes over a single time interval. The heuristic works as follows: at the beginning of interval $s \in \{1, 2, ..., S\}$, the decision-maker observes the phase counts x and y for classes $\mathcal{A}$ and $\mathcal{D}$ respectively and also calculates the expected numbers of customers arriving during interval s, $\mathbb{E}[A_s|\mathbf{x}]$ and $\mathbb{E}[D_s|\mathbf{x}]$. Taking into account that customers of class $\mathcal{A}$ are more expensive to hold in the queue by a factor α, the decision-maker calculates the 'effective demand ratio' as

$$\frac{\alpha(x + k\mathbb{E}[A_s|\mathbf{x}])}{y + k\mathbb{E}[D_s|\mathbf{x}]}. \tag{45}$$

They then select the unique point (μ_s, ν_s) on the OTE for interval s which results in the ratio μ_s/ν_s being equal to the effective demand ratio in (45).

The principle of making decisions based on short-term expected queue lengths has previously been adopted by Jacquillat and Odoni [31], who proposed a heuristic for air traffic control based on setting the service rate for arriving aircraft to be as close as possible to the 'effective arrival demand'.

It should be noted that the naive MDR heuristic - unlike the MYP, API, CAV and LSF heuristics developed in Section 5 (which we will refer to from this point on as the 'SAB heuristics') - is not based on any of the analyses in Section 4, and therefore does not rely upon a "servers always busy" (SAB) assumption. This suggests that any superiority that the SAB heuristics might have over the MDR heuristic might diminish as the SAB assumption becomes less reliable, i.e. as the system becomes less busy. Later in this section we will provide evidence to show that this is indeed the case.

6.2 Results from simulation experiments

We have created 28,800 test scenarios and used Monte Carlo simulation to test the performances of the 5 heuristics within each scenario. Details of the methods we have used for generating the parameters can be found in Appendix A.8. It should be noted that our heuristics are designed for use during congested periods, and therefore we have generated the random demand rates, initial phase counts and service rate constraints in such a way that server idleness is unlikely to occur.

For each scenario, we simulated the random evolution of the system under the 4 SAB heuristics and the MDR heuristic, carrying out 500 replications for each policy and ensuring that the same random number stream was used for all 5 policies. In all experiments, the random evolutions of the phase counts were simulated accurately but the decisions made under the various heuristic policies were made *without* reference to the exact phase counts; instead, these were made in response to the numbers of *customers* in the two queues and the aggregation method described in Section 2 was used to translate decisions based on phase counts to decisions based on observed numbers of customers.

Table 1 shows pairwise comparisons between the 5 heuristics across all 28,800 scenarios. Each cell (i, j)

in the table has 3 rows. The first row shows (in bold font) a 95% confidence interval for the percentage increase (or decrease) in the mean total cost when using heuristic i as opposed to heuristic j, taking into account the results from all scenarios. The second row shows an analogous confidence interval for the 50% of scenarios in which customers arrive via nonhomogeneous Poisson processes, and the third row shows the results for the remaining 50% of scenarios in which customers arrive via pre-scheduled random arrival (PSRA) distributions. For clarity, '$-$' denotes a reduced cost, i.e. an improvement.

Policy	MDR	MYP	API	CAV	LSF
MDR		**+3.47±0.12**	**+8.47±0.17**	**+11.91±0.17**	**+16.11±0.21**
Poisson	x	+2.89±0.18	+8.48±0.25	+13.82±0.24	+15.65±0.29
PSRA		+4.05±0.17	+8.47±0.24	+10.00±0.24	+16.56±0.31
MYP	**−2.50±0.10**		**+4.83±0.10**	**+8.35±0.13**	**+12.25±0.15**
Poisson	−1.88±0.15	x	+5.47±0.15	+10.97±0.20	+12.53±0.21
PSRA	−3.12±0.13		+4.21±0.13	+5.72±0.15	+11.98±0.22
API	**−6.65±0.10**	**−4.19±0.06**		**+3.39±0.08**	**+7.03±0.10**
Poisson	−6.60±0.14	−4.72±0.10	x	+5.25±0.13	+6.67±0.13
PSRA	−6.70±0.14	−3.67±0.08		+1.53±0.10	+7.40±0.14
CAV	**−9.43±0.11**	**−6.94±0.09**	**−2.85±0.08**		**+3.68±0.08**
Poisson	−10.95±0.15	−8.98±0.14	−4.43±0.12	x	+1.44±0.07
PSRA	−7.91±0.15	−4.90±0.10	−1.21±0.08		+5.90±0.13
LSF	**−12.31±0.12**	**−9.98±0.10**	**−6.07±0.08**	**−3.20±0.06**	
Poisson	−12.03±0.17	−10.22±0.14	−5.79±0.11	−1.31±0.06	x
PSRA	−12.58±0.17	−9.73±0.14	−6.34±0.11	−5.10±0.10	

Table 1: Pairwise comparisons of the 5 heuristic policies across all 28,800 scenarios

The results indicate that the LSF heuristic performs best, and is able to improve upon the myopic policy by about 10% on average. The CAV heuristic is the second strongest, but (unlike the LSF) appears to become significantly weaker if pre-scheduled arrivals are considered rather than Poisson arrivals. In the PSRA case, the LSF and CAV heuristics both have to estimate a greater number of coefficients in their quadratic approximating functions, and the LSF heuristic seems able to adapt better to this than the more naive CAV heuristic. The API heuristic consistently performs better than the myopic policy, but is usually not able to match the more sophisticated CAV and LSF heuristics. The naive MDR heuristic is not able to perform as well as any of the SAB heuristics. Overall, as expected, the performances of the 4 heuristics are ranked in an order which corresponds with their computational requirements.

We now present further analyses of our numerical results, focusing only on the relative strengths of the 4 SAB heuristics (the naive MDR heuristic will be discussed again later in this section, when different levels of server 'busyness' are considered). We can gain greater insight into the relative strengths of the SAB heuristics by breaking down the results according to the values of individual system parameters. Figure 2 shows comparisons between the API, CAV and LSF policies for various different values of β (recall that β controls the weight of the 'switching costs' for changing service rates). The policies are compared with respect to the improvements they achieve against the myopic policy.

When $\beta = 0$, the API heuristic is relatively strong. This is because, under an optimal policy, service rates may change radically from one interval to the next and API explicitly takes into account the piecewise linear nature of the OTE and the fact that different regions of the OTE correspond to constraints with
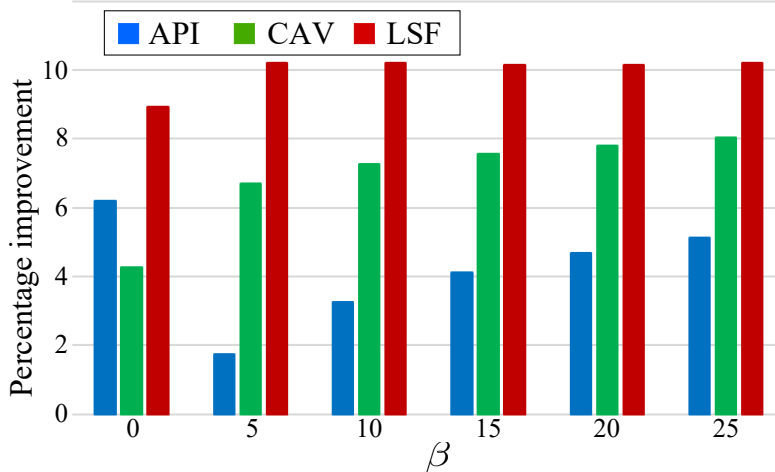
Figure 2: The average percentage improvements achieved by the API, CAV and LSF heuristics versus the MYP heuristic for various different values of β

different gradients, whereas the quadratic value function approximations relied upon by the CAV and LSF heuristics would be most suitable if the OTE was an infinite straight line. As β becomes larger, each OTE effectively becomes narrower as we only need to consider a small interval of possible μ_s values as realistic candidates for optimality. We can see that for relatively small, positive values of β, API is much worse than CAV and LSF, but as β becomes larger API seems to improve. This may be because, as β becomes large, the API policy's prediction of the next system state becomes almost unimportant with respect to predicting the next service rate that will be chosen ($\mu_{s+1}^{\mathrm{MYP}}(\mathbf{X}_s)$), so the API policy is effectively optimizing a two-step cost without any major uncertainty about the next service rate. Meanwhile, the CAV and LSF heuristics still maintain some superiority over API, but lack the ability to prioritize any particular segment of the OTE when approximating the shapes of the value functions.

Further analysis of our results suggests that the API, CAV and LSF heuristics all tend to increase their advantages over the myopic policy as (i) the parameter δ_{OTE} (which represents the variation in the OTEs from interval to interval - see Appendix A.8) increases; (ii) the number of time intervals S increases; (iii) the lengths of the time intervals increase. Tables 2, 3 and 4 show the percentage improvements (in the form of 95% confidence intervals) gained by the API, CAV and LSF policies against the myopic policy for various different values of these parameters. Increasing the amount of variability in the time-dependent customer arrival rates also appears to harm the performance of the myopic policy (the parameter(s) which control the demand rate variability are explained in Appendix A.8). We also tried varying the parameters k and α, but did not observe any significant ensuing effects on the relative strengths of the various policies.

We make two additional observations here regarding the results in Tables 2–4:

- The values shown in the last row of Table 2 (for $\delta_{\mathrm{OTE}} = 0.5$) are quite large. This suggests that a myopic policy may be very weak if the OTEs vary greatly from interval to interval.

- By making comparisons across the columns of Table 3 we can see that the API policy's relative strength against the CAV and LSF policies seems to diminish as S increases. This may be because the API policy only looks ahead by one time interval − which makes it better-suited for problems with short time horizons.

26

OTE parameter	API	CAV	LSF
$\delta_{\text{OTE}} = 0.1$	1.61 ± 0.07	2.17 ± 0.07	3.05 ± 0.07
$\delta_{\text{OTE}} = 0.2$	2.28 ± 0.07	3.92 ± 0.10	5.63 ± 0.09
$\delta_{\text{OTE}} = 0.3$	3.63 ± 0.10	6.85 ± 0.15	9.64 ± 0.14
$\delta_{\text{OTE}} = 0.4$	5.58 ± 0.15	9.61 ± 0.21	13.94 ± 0.20
$\delta_{\text{OTE}} = 0.5$	7.80 ± 0.20	12.02 ± 0.28	17.46 ± 0.27

Table 2: Percentage improvements for the API, CAV and LSF heuristics versus the MYP heuristic for various different values of δ_{OTE}, the OTE variability parameter

No. of time intervals	API	CAV	LSF
$S = 12$	3.85 ± 0.14	4.50 ± 0.14	5.77 ± 0.15
$S = 18$	4.42 ± 0.16	6.30 ± 0.18	8.38 ± 0.18
$S = 24$	4.14 ± 0.14	7.09 ± 0.20	10.12 ± 0.20
$S = 30$	4.30 ± 0.14	7.96 ± 0.22	12.04 ± 0.24
$S = 36$	4.27 ± 0.13	8.91 ± 0.23	13.68 ± 0.25

Table 3: Percentage improvements for the API, CAV and LSF heuristics versus the MYP heuristic for various different values of S, the number of time intervals

Our experiments were implemented in Python on a 2.5 Ghz quad-core processor with 8 GB of RAM, and we observed that none of the 4 heuristics required any significant amount of computational effort in comparison to dynamic programming algorithms. The MYP and API heuristics use closed-form (albeit somewhat messy) formulas for their decision-making and are able to return decisions for all S time intervals in less than one second. The CAV and LSF heuristics differ from the more simple MYP and API heuristics by requiring the pre-computation of a set of quadratic coefficients for each time interval. However, for the parameter values that we considered, we still found that the total amount of time required by the CAV heuristic to pre-compute quadratic coefficients and then return decisions for all time intervals was less than one second. The LSF heuristic requires slightly more time due to the need to 'sample' a number of states at each time interval as part of its least-squares fitting technique, as explained in Section 5. As described in Appendix A.8, we used a constant sample size $|M_s|$=500 in all of our experiments. With this value of $|M_s|$ we found that the total amount of computation time required by the LSF heuristic was between one and ten seconds, depending mainly on the number of time intervals, the number of linear constraints for the service rates and also the nature of the arrivals process (Poisson or PSRA).

Next, we decided to test the sensitivity of our results to the 'busyness' of the system by sampling 4800 of the 28,800 previously-tested scenarios and, for each of these 4800 scenarios, repeating our experiments several times with re-scaled demand rates. For these experiments, we restricted attention to scenarios with $\beta = 0$, in order to simplify the aggregate cost function to $\sum_{s=1}^{S}(\alpha X_s^2 + Y_s^2)$, without the extra component related to switching costs. Our intention was to demonstrate that, with $\beta = 0$, reducing the demand rates causes the difference in performance between the SAB heuristics and the more naive MDR heuristic to become smaller. With $\beta > 0$ the picture becomes less clear because, in that case, the switching cost component of the objective function becomes relatively more important as demand rates decrease (due to the expected phase counts becoming smaller), and this in turn may actually cause the SAB heuristics to increase their advantage over the MDR heuristic since they are able to deal with switching costs in a more

Time interval length	API	CAV	LSF
$5 \leq \tau < 10$	2.59 ± 0.07	5.30 ± 0.14	7.87 ± 0.16
$10 \leq \tau < 15$	3.99 ± 0.11	6.93 ± 0.17	10.06 ± 0.19
$15 \leq \tau < 20$	4.85 ± 0.14	7.66 ± 0.19	10.90 ± 0.20
$20 \leq \tau < 25$	5.33 ± 0.16	7.86 ± 0.21	11.07 ± 0.22

Table 4: Percentage improvements for the API, CAV and LSF heuristics versus the MYP heuristic for various different ranges of τ, the length of each time interval

sophisticated way. We re-scaled the demand rates as follows:

- For scenarios with Poisson arrivals, the Poisson demand rate functions $\lambda(t)$ and $\eta(t)$ were replaced by $p\lambda(t)$ and $p\eta(t)$ respectively, where $0 < p < 1$.

- For PSRA scenarios, the numbers of customers scheduled to arrive in the various time intervals were all multipled by p before being rounded to the nearest integer.

We tried each of the values $p = 0.75$, $p = 0.67$, $p = 0.5$ and $p = 0.25$ and, in each case, re-ran our numerical experiments for the 4800 sampled scenarios. Table 5 summarizes our results, with the results for $p = 1$ (from earlier experiments) included for comparison purposes. For each value of p, the proportion of time that the system spends idle under the MDR heuristic (estimated via the simulation experiments) is shown in bold. As predicted, the SAB heuristics tend to lose some of their superiority over the MDR heuristic as the system becomes less busy. The CAV and LSF heuristics become useless in the $p = 0.5$ and $p = 0.25$ cases due to the large amount of server idle time (which renders the quadratic value function approximations completely inaccurate). This highlights the importance of making sure the system is likely to be sufficiently busy before employing these quadratic-based heuristics. However, the myopic and API heuristics still retain some superiority over MDR even in the $p = 0.5$ and $p = 0.25$ cases.

	vs. MYP	vs. API	vs. CAV	vs. LSF
$p = 1$ (**<0.1% idle time**)	12.63 ± 0.28	17.60 ± 0.37	16.06 ± 0.36	19.93 ± 0.40
$p = 0.75$ (**3.8% idle time**)	12.16 ± 0.26	17.68 ± 0.35	12.18 ± 0.72	10.22 ± 1.12
$p = 0.67$ (**10.8% idle time**)	10.63 ± 0.23	14.84 ± 0.30	2.80 ± 1.18	8.74 ± 1.96
$p = 0.5$ (**34.0% idle time**)	7.02 ± 0.17	7.89 ± 0.23	-35.81 ± 3.09	-80.97 ± 4.87
$p = 0.25$ (**66.4% idle time**)	4.74 ± 0.13	1.70 ± 0.26	-80.79 ± 6.32	-155.56 ± 8.97

Table 5: Percentage cost increases resulting from using the MDR heuristic as opposed to each of the 4 SAB heuristics for various different values of p, with proportions of idle time shown under each p value

6.3 Analysis of customer waiting times

Our experiments in this section so far have been designed to compare the performances of the heuristic policies with respect to the objective function (1). A practitioner might also wish to know the effects of the heuristics on measures of performance which are more directly of interest, such as average waiting times and maximum waiting times. With this in mind, we carried out another set of experiments involving 4800 randomly-generated scenarios. In these experiments we used the parameter values $\alpha = 1$ and $\beta = 0$

in order to make the objective function (1) as suitable as possible for minimizing the average customer waiting time across all time intervals, for classes $\mathcal{A}$ and $\mathcal{D}$ combined. All other parameters were generated as described in Appendix A.8.

Table 6 shows pairwise comparisons between the 4 SAB heuristics based on the results from all 4800 scenarios. The figures in the table should be interpreted the same way as those in Table 1. That is, the figure in cell (i, j) of the table is a 95% confidence interval for the percentage difference in the average customer waiting time (across all time intervals) obtained by using heuristic i as opposed to heuristic j, with the Poisson and PSRA cases also shown separately.

Policy	MYP	API	CAV	LSF
MYP		**+3.52±0.12**	**+5.17±0.16**	**+7.18±0.21**
Poisson	x	+4.08±0.19	+5.71±0.26	+8.02±0.32
PSRA		+2.96±0.15	+4.63±0.20	+6.33±0.26
API	**−3.25±0.10**		**+1.57±0.08**	**+3.46±0.11**
Poisson	−3.74±0.16	x	+1.64±0.13	+3.71±0.17
PSRA	−2.76±0.13		+1.54±0.13	+3.21±0.13
CAV	**−4.65±0.14**	**−1.46±0.10**		**+1.88±0.09**
Poisson	−5.07±0.22	−1.41±0.13	x	+2.17±0.14
PSRA	−4.23±0.17	−1.52±0.10		+1.60±0.13
LSF	**−6.29±0.17**	**−3.21±0.10**	**−1.75±0.09**	
Poisson	−6.96±0.26	−3.41±0.17	−2.01±0.13	x
PSRA	−5.63±0.21	−3.01±0.12	−1.48±0.12	

Table 6: Pairwise comparisons of the 4 SAB heuristics with respect to average customer waiting times

Comparing the results in Table 6 with those in Table 1, we see that when average waiting times are used as a performance measure (with $\alpha = 1$ and $\beta = 0$) as opposed to the usual objective function (1), the ranking order of the 4 SAB heuristics remains the same, although the percentage improvements against the myopic policy are somewhat smaller - which is perhaps unsurprising, given that our objective function is not directly based on waiting times.

Figure 3 illustrates the differences in tail behavior between the 4 SAB heuristics. For each individual scenario and each value of p in $\{5, 10, ..., 90, 95\}$ we have calculated the pth percentile of the customer waiting time distribution under the myopic policy and then calculated the proportions of waiting times which equal or exceed this (myopic) percentile under each of the 4 SAB heuristics. The aggregated results are plotted in the figure. Our method of construction obviously implies that the purple myopic curve should be almost linear (the nonlinearity being due to ties, since there are zeros in the data), but the curves for the other 3 heuristics are slightly more convex in shape and tend to remain below the myopic curve, which reflects the fact that the proportions of customers exceeding the critical waiting time thresholds (as given by the myopic policy) are usually smaller under the API, CAV and LSF policies.

7 Conclusions and further work

In this paper we have considered a problem involving dynamic allocation of resources between two sources of time-varying demand, with an intended application to the tactical control of scarce resources
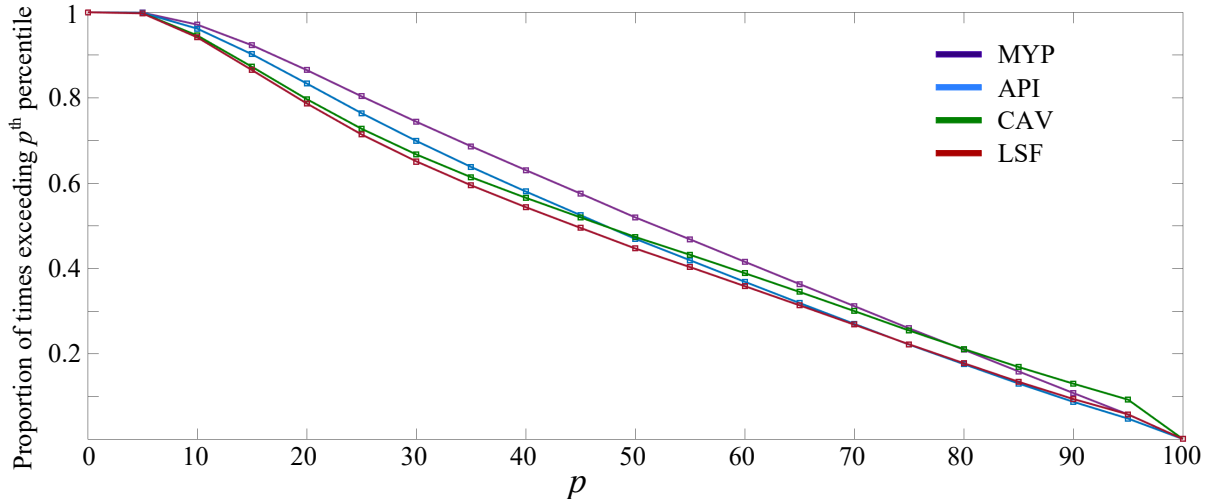
Figure 3: Proportions of customer waiting times exceeding the myopic policy's p^{th} percentile

at busy airports. We have developed and tested a number of heuristic approaches, all of which are based on an approximation of the queueing dynamics (the SAB approximation) during busy times of day. These heuristics are much faster to implement than conventional optimization procedures based on dynamic programming. Indeed, the MYP, API and CAV essentially require no computation time at all, while the LSF heuristic requires only a few seconds. Although our numerical experiments have assumed that the parameters or functions governing the customer arrival rates (e.g. $\lambda(t)$ and $\eta(t)$, in the case of Poisson arrivals) are predetermined and not subject to change, the low computational requirements of our heuristics imply that they could easily be re-run in order to prescribe new decisions at short notice if the arrival rates (or other system parameters) are subject to stochastic variations during on-line implementation.

A key finding of our work is that, under quite plausible modelling assumptions (e.g. similarity between the customer classes with respect to demand rates and service rate constraints), myopic policies based on SAB queue dynamics appear to perform strongly. This is worthy of note since, in general stochastic optimization problems, one usually cannot rely on a myopic policy to be anywhere near optimal. Furthermore, our numerical experiments have identified circumstances in which one can improve significantly on the performance of a myopic policy through the use of more sophisticated heuristics. These circumstances include long durations for time intervals and highly time-dependent service rate constraints.

Comparisons between our SAB-based heuristics suggest that:

- The LSF heuristic performs strongest overall, but is also the most computationally expensive.

- The CAV heuristic is strong in many cases but also appears to be more sensitive to the value of β and the type of arrivals distribution.

- The API heuristic offers a computationally inexpensive method of improving upon a myopic policy and appears to perform best in the $\beta = 0$ case.

As Table 5 shows, however, the ranking order of the heuristics can change dramatically if demand rates are not high enough for the SAB assumption to be justified.

From a methodological perspective, the approximate DP methods that we have developed in this paper - in particular, the least-squares fitting (LSF) heuristic - bear certain similarities to reinforcement learning (RL) methods which have been applied to many different types of dynamic control problems in the last few decades (see [6, 47]). Our LSF heuristic operates by randomly sampling a set of states at each time interval and fitting a quadratic to the value function estimates obtained at these states. In this respect it is similar to RL algorithms which use value function approximation and/or neural networks. We believe that our approach is particularly efficient for this particular problem because, as a result of our analyses in Section 4, we are able to use analytical expressions for the second moments in our surrogate problem to efficiently obtain value function estimates at the various time intervals via backwards induction, without needing to simulate the random transitions of the process and explore potentially sub-optimal action choices (as a conventional RL algorithm might do). In future work we intend to investigate whether the use of simulation-based RL algorithms might enable us to obtain strong parametric value function approximations under a broader range of conditions, including non-busy periods.

There are several interesting possibilities for extensions of our work. In this paper we have assumed, *a priori*, that the period of operation is one in which the servers remain almost constantly in use. At an airport, this might correspond to a period of several hours during part of the day (e.g. a late-afternoon peak period in which many aircraft are expected to arrive). The question of how our approach may be combined with a method for decision-making during non-busy times of day is left open for further research, as is the related issue of how optimizing decisions during busy periods might affect the performance of the system during any subsequent, non-busy periods. As we have asserted elsewhere, however, it is intuitively clear that if a decision-maker's main interest is in mitigating the worst effects of congestion, then optimizing decisions during the busiest times of day should help greatly.

In the context of airport operations, our work continues the long-standing convention of modelling queues of aircraft as $M(t)/E_k(t)/1$ queues, but we have also adapted our heuristic methods to the case of pre-scheduled random arrivals, which has received considerable recent attention in the literature. With airport applications in mind, there are various other possible generalizations of our model. We have not incorporated a stochastic model of weather evolution, and have instead relied upon the implicit assumption that reliable weather forecasts are available which enable the OTEs for all time intervals (which, in practice, would be affected by weather variations, as well as the types of aircraft scheduled to arrive at different times of day and any scheduled runway configuration changes) to be predicted with a high degree of confidence. Also, although our incorporation of 'switching penalties' for changing service rates is intended to model - to some extent - the costs involved in dynamically re-deploying airport resources, our model would have to be extended in order to facilitate more complicated operations (e.g. runway configuration changes) which might broaden the range of options available to a decision-maker.

airport resources (OR-MASTER)".

References

[1] K.M. Adusumilli and J.J. Hasenbein. Dynamic admission and service rate control of a queue. *Queueing Systems*, 66(2):131–154, 2010.

[2] M. Arns, P. Buchholz, and A. Panchenko. On the Numerical Analysis of Inhomogeneous Continuous-Time Markov Chains. *INFORMS Journal on Computing*, 22(3):416–432, 2010.

[3] B. Ata and S. Shneorson. Dynamic Control of an M/M/1 Service System with Adjustable Arrival and Service Rates. *Management Science*, 52(11):1778–1791, 2006.

[4] C. Barnhart, P. Belobaba, and A.R. Odoni. Applications of Operations Research in the Air Transport Industry. *Transportation Science*, 37(4):368–391, 2003.

[5] C. Barnhart, D. Fearing, A. Odoni, and V. Vaze. Demand and capacity management in air transportation. *EURO Journal on Transportation and Logistics*, 1:135–155, 2012.

[6] D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[7] D. Bertsimas, M. Frankovich, and A.R. Odoni. Optimal selection of airport runway configurations. *Operations Research*, 59(6):1407–1419, 2011.

[8] J.H. Bookbinder. Multiple Queues Of Aircraft Under Time-Dependent Conditions. *INFOR: Information Systems and Operational Research*, 24:280–288, 1986.

[9] M.V. Caccavale, A. Iovanella, C. Lancia, G. Lulli, and B. Scoppola. A model of inbound air traffic: The application to Heathrow airport. *Journal of Air Transport Management*, 34:116–122, 2014.

[10] T.B. Crabill. Optimal Control of a Maintenance System with Variable Service Rates. *Operations Research*, 22:736–745, 1974.

[11] M. Defraeye and I. Van Nieuwenhuyse. Staffing and scheduling under nonstationary demand for service: A literature review. *Omega*, 58:4–25, 2016.

[12] P. Dell'Olmo and G. Lulli. A dynamic programming approach for the airport capacity allocation problem. *IMA Journal of Management Mathematics*, 14(3):235–249, 2003.

[13] Y. Dimitrakopoulos and A.N. Burnetas. Customer equilibrium and optimal strategies in an M/M/1 queue with dynamic service control. *European Journal of Operational Research*, 252(2):477–486, 2016.

[14] W.J. Dunlay and R. Horonjeff. Stochastic properties of enroute air traffic - An empirical investigation. *Journal of Aircraft*, 13(5):376–381, 1976.

[15] A. Eryilmaz and R. Srikant. Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. *IEEE/ACM Transactions on Networking*, 15(6):1333–1344, 2007.

[16] Eurocontrol. *CODA digest: Delays to air transport in Europe (Annual 2013)*. Report prepared by Eurocontrol's Central Office for Delay Analysis (CODA), Brussels, Belgium, 2014.

[17] European Commission. SESAR 2020: developing the next generation of European Air Traffic Management. Accessed May 21, 2018, 2018.

[18] J.M. George and J.M. Harrison. Dynamic Control of a Queue with Adjustable Service Rate. *Operations Research*, 49(5):720–731, 2001.

[19] E. Gilbo. Airport capacity: Representation, estimation, optimization. *IEEE Trans. Control Systems Tech.*, 1(3):144–154, 1993.

[20] Steer Davies Gleave. *Impact Assessment of Revisions to Regulation 95/93*. Study prepared for the European Commission (DG MOVE), London, UK, 2011.

[21] L.V. Green, P.J. Kolesar, and W. Whitt. Coping with time-varying demand when setting staffing requirements for a service system. *Production and Operations Management*, 16(1):13–29, 2007.

[22] D. Gross and C. Harris. *Fundamentals of Queueing Theory*. John Wiley & Sons, New York, 1998.

[23] G. Guadagni, C. Lancia, S. Ndreca, and B. Scoppola. Queues with exponentially delayed arrivals. *ArXiv e-prints 1302.1999*, 2013.

[24] S. Gupta. *Transient Analysis of $D(t)/M(t)/1$ Queuing System with Applications to Computing Airport Delays*. Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, 2010.

[25] C. Gwiggner and S. Nagaoka. Data and queueing analysis of a Japanese air-traffic flow. *European Journal of Operational Research*, 235(1):265–275, 2014.

[26] A.Y. Ha. Inventory Rationing in a Make-to-Stock Production System with Several Demand Classes and Lost Sales. *Management Science*, 43(8):1093–1103, 1997.

[27] M. Hansen, T. Nikoleris, D.J. Lovell, K. Vlachou, and A.R. Odoni. Use of Queueing Models to Estimate Delay Savings from 4D Trajectory Precision. *Proc. 8th USA/Europe Air Traffic Management R&D Seminar, Napa, CA*, 2009.

[28] International Air Transport Association. Fact sheet: Single European Sky (SES). Accessed February 28, 2018, 2017.

[29] A. Jacquillat. *A queuing model of airport congestion and policy implications at JFK and EWR*. Unpublished Master's thesis, Massachusetts Institute of Technology, Cambridge, MA, 2012.

[30] A. Jacquillat and A.R. Odoni. An Integrated Scheduling and Operations Approach to Airport Congestion Mitigation. *Operations Research*, 63(6):1390–1410, 2015.

[31] A. Jacquillat, A.R. Odoni, and M.D. Webster. Dynamic Control of Runway Configurations and of Arrival and Service Departure Rates at JFK Airport Under Stochastic Queue Conditions. *Transportation Science*, 51(1):155–176, 2017.

[32] P. Kivestu. *Alternative methods of investigating the time-dependent $M/G/K$ queue*. S.M. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, 1976.

[33] B.O. Koopman. Air-Terminal Queues under Time-Dependent Conditions. *Operations Research*, 20(6):1089–1114, 1972.

[34] N. Lee and V.G. Kulkarni. Optimal arrival rate and service rate control of multi-server queues. *Queueing Systems*, 76(1):37–50, 2014.

[35] K. Malone. *Dynamic queuing systems: Behavior and approximations for individual queues and networks*. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1995.

[36] T. Nikoleris and M. Hansen. Queueing models for trajectory-based aircraft operations. *Transportation Science*, 46(4):501–511, 2012.

[37] A.P. Perez and P. Zipkin. Dynamic Scheduling Rules for a Multiproduct Make-To-Stock Queue. *Operations Research*, 45(6):919–930, 1997.

[38] Joint Planning and Development Office. Concept of Operations for the Next Generation Air Transportation System. 2010.

[39] W.B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley & Sons, New Jersey, 2007.

[40] J.T. Prince and D.H. Simon. Multimarket Contact and Service Quality: Evidence from On-Time Performance in the U.S. Airline Industry. *Academy of Management Journal*, 52(2):336–354, 2009.

[41] M.L. Puterman. *Markov Decision Processes - Discrete Stochastic Dynamic Programming*. Wiley & Sons, New York, 1994.

[42] N. Pyrgiotis, K.M. Malone, and A.R. Odoni. Modelling delay propagation within an airport network. *Transportation Research Part C*, 27:60–75, 2013.

[43] N. Pyrgiotis and A.R. Odoni. On the Impact of Scheduling Limits: A Case Study at Newark Liberty International Airport. *Transportation Science*, 50(1):150–165, 2016.

[44] L. Sherry, D. Wang, and G. Donohue. Air travel consumer protection: Metric for passenger on-time performance. *Transportation Research Record*, pages 22–27, 2007.

[45] I. Simaiakis. *Analysis, modeling and control of the airport departure process*. Unpublished doctoral thesis, Massachusetts Institute of Technology, Cambridge, MA, 2012.

[46] M.A. Stamatopoulos, K.G. Zografos, and A.R. Odoni. A decision support system for airport strategic planning. *Transportation Research Part C*, 12:91–117, 2004.

[47] R. Sutton and A. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, 1998.

[48] P. Swaroop, B. Zou, M.O. Ball, and M. Hansen. Do more US airports need slot controls? A welfare based approach to determine slot levels. *Transportation Research Part B*, 46(9):1239–1259, 2012.

[49] N.M. van Dijk. Uniformization for nonhomogeneous Markov chains. *Operations Research Letters*, 12:283–291, 1992.

[50] R.R. Weber and S. Stidham. Optimal Control of Service Rates in Networks of Queues. *Advances in Applied Probability*, 19(1):202–218, 1987.

[51] W. Whitt. What you should know about queueing models to set staffing requirements in service systems. *Naval Research Logistics*, 54(5):476–484, 2007.

[52] T.R. Willemain, H. Fan, and H Ma. Statistical analysis of intervals between projected airport arrivals. 2004.

[53] J. Yimga. Airline on-time performance and its effects on consumer choice behavior. *Research in Transportation Economics (In Press)*, 2017.

[54] K.G. Zografos, M.A. Madas, and K.N. Androutsopoulos. Increasing airport capacity utilisation through optimum slot scheduling: review of current developments and identification of future needs. *Journal of Scheduling*, 20(1):3–24, 2017.