Multi-Class Advance Patient Scheduling

by

Mohamad Sadegh Shirani Faradonbeh

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Management Sciences

Waterloo, Ontario, Canada, 2021

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The problem of advance scheduling of service appointments for patients arriving to a healthcare facility received a lot of attentions in the literature of operations management. Broadly speaking, the main goal is on the efficient assignment of patients entering the system to the next operating days, advance in time and in a dynamic manner. In particular, the problem of multi-class advance patient scheduling, that aims to incorporate differences in the priority levels of patient classes, is of interest in many situations. In this setting, one needs to address important challenges to efficiently utilize the limited and costly resources of the underlying healthcare facilities. Furthermore, a reliable scheduling policy needs to reserve sufficient capacity for high-priority patients, in order to avoid long waiting-times for urgent cases in the future. Accordingly, at every time instant, the policy needs to consider all outstanding appointments, as well as uncertainties in the future demand.

This work presents the first theoretically tractable framework for design and analysis of efficient advance scheduling policies in a multi-class setting. First, we provide a realistic formulation of the problem that reflects both the transient as well as the long-term behavior of scheduling policies. Then, we study optimal policies that efficiently schedule patients of different classes and characterize the resulting coarse-grained fluid dynamics, as well as the finer dynamics of diffusion approximation. In fact, the former yields to a simple policy that schedules all patients on the day of their arrival, and also sets the stage for the analysis of the latter stochastic dynamical model. Then, we proceed towards considering diffusion processes based on which the study of scheduling policies becomes a Brownian control problem. Finally, by leveraging a dynamic programming approach, we characterize the optimal policy and validate it through numerical implementations.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Hossein Abouee Mehrizi, for his valuable advice and continuous support during my master studies.

I also would like to thank my lovely wife Narges, for her tremendous supports, effective encouragements, and endless companionship in the recent few years.

Table of Contents

Chapter 1

Introduction

Demand uncertainties together with resource constraints have created a huge backlogged workload in the healthcare systems. For example, due to Covid-19 pandemic, Ontario has a backlog of nearly 16 million medical procedures[1]. Magnetic Resonance Imaging (MRI) with 477,301 cases is the procedure with the most backlogs[2]. The goal of this study is to develop implementable, yet reliable and efficient algorithms for multi-class patient scheduling. The process of patient scheduling assigns available resources in a healthcare system to patients, over an operating time horizon (Truong 2015). The fundamental goal in autonomous systems for appointment scheduling is matching supply and demand, considering the resulting operating costs as well as the trade-offs in patient wait time for different classes. Technically, the former captures the efficiency in utilizing limited resources, while the latter signifies the importance of timely access to the corresponding services.

Motivated by data collected from more than 70 hospitals in the province of Ontario, Canada, we study the problem of assigning appointment times to patients of multiple

[1] 16 million health-care procedures backlogged in Ontario: OMA

[2] Almost 16 million medical procedures built up in Ontario pandemic backlog

classes over a time horizon, subject to unpredictable daily fluctuations in demand. Patients are classified either by type or by priority, and their service requests arrive randomly to the underlying healthcare facility. Clearly, significant parameters for comparing the patients include the coefficients that reflect the waiting cost per unit time and service times that measure the time-length needed for providing and accomplishing different services. Therefore, the scheduler may set an appointment time for patients right after their request arrive, or alternatively make them wait according to their class and the state of the system at the time of their arrival.

We assume that at every time, the scheduling window consists of H upcoming days starting from the current day, and the number of days-to-schedule, H, is a priori fixed. So, a scheduling procedure needs to cope with

1. waiting time of a patient before receiving an appointment time,
2. waiting time of a patient after receiving an appointment until the appointment date,
3. and the utility cost a hospital needs to incur for a scheduled daily workload.

The broad objective is to find appointment scheduling policies for minimizing the waiting and utility costs in a (long) time horizon. We formulate the problem by assuming that service times are deterministic for all patients, and demands consist of stochastic processes with statistical properties that vary among different classes of patients. To have a realistic setting, we assume that the utility cost is a convex function of the scheduled workload. Indeed, this convexity reflects the well-known fact that the cost of scheduling each patient for a given day increases with the currently scheduled workload for that day.

Solving the problem using conventional methods leads to computational intractabilities caused by the large number of decision variables involved. More specifically, since the state of the system at any time depends on the arrival, scheduling, and appointment time of each patient, the number of possibilities grows exponentially with the total number of days, rendering the computational complexity of the problem NP-hard. Therefore, we establish an equivalent asymptotic problem at both fluid and diffusion scales. Considering the optimization problem in the asymptotic regime reveals the fundamental complicating factors (Armony et al. 2019). In the scaling scheme, we let the arrival rates tend to infinity while increasing the service capacity in proportion to the arrivals. We show that in the fluid regime, the daily variations of the demand process vanish, i.e., it remains the same demand over different days. As a result, the problem becomes a dynamic scheduling problem with deterministic arrivals. The consequence is that a simple scheduling policy that serves all patients on the arrival day performs optimally.

Moreover, analysis of the fluid regime provides the basis for studying the problem at the finer level of modeling the arrivals by diffusion processes. Technically, we show that the behavior of the fluctuations around the fluid arrival-rate constitute a Brownian motion. Therefore, we derive the corresponding Brownian Control Problem (**BCP**); in lights of its technical framework, we find optimal scheduling policies and interpret their prescriptions.

To this end, we take the following steps. First, we formulate the scheduling problem using a diffusion process. To characterize the optimal scheduling policy, we restrict the admissible set of policies to those that meet an appropriate differentiability condition. This lets us state the multi-class patient scheduling problem as a BCP. To deal with the obtained BCP, we employ the Dynamic Programming Principle (DPP) to tackle the problem. Specifically, the DPP splits the original problem into optimal control problems over two time intervals. Then, we derive the Hamilton-Jacobi-Bellman equation (HJB). The HJB is the infinitesimal representation of the DPP and describes the local behavior of the optimal control problem around the current state of the system. The result of the

HJB is interpreted as an optimal threshold policy for the scheduling problem where the threshold is given for each specific class of patients and for every day in the scheduling window. We determine the asymptotic optimal scheduling policy based on these thresholds and numerically validate the performance for the multi-class advance patient scheduling problem under study.

1.1 Relevant Literature

The existing literature of appointment scheduling is rich in general. Along this direction, Cayirli and Veral (2003) focus on appointment scheduling in outpatient services and present important considerations for formulating and modeling the problem. Gupta and Denton (2008) identify challenges in designing appointment systems and discuss potential directions for further investigations. In the work of Hall et al. (2012), applications of queueing models and stochastic processes for improving scheduling in the healthcare systems are demonstrated. Moreover, Ahmadi-Javid et al. (2017) review the papers that focus on designing and planning outpatient appointment systems. Further details together with comprehensive discussions about the different approaches can be found in the aforementioned references.

For studying the waiting times, there are two main categories in the existing literature. The first stream focuses on the concept of *direct wait*, which is the gap between the scheduled appointment time and the actual service time. This approach is widely used by the community and focuses on minimizing the direct-wait-time using techniques such as sequencing and restricting the appointments to a given interval (Denton and Gupta 2003, Hassin and Mendel 2008, Zacharias and Pinedo 2014, Kuiper et al. 2017).

There is also another line of works that adopt an asymptotic approach for modeling the arrival process. Araman and Glynn (2012) show that under non-punctual arrivals, the limit of scheduled traffic is a fractional Brownian motion. Under punctual arrivals assumption, Armony et al. (2019) investigate the problem of scheduling within a certain time interval for a finite population of customers. The problem is addressed in the fluid and diffusion scales in the presence of no-shows and the asymptotically optimal schedule to minimize the sum of customers waiting time and server overtime costs is characterized.

On the other hand, the *indirect wait* points out the interval between the patient request and the scheduled appointment time. As stated in the work of Gupta and Denton (2008), modeling and control of indirect waiting is an open problem. However, dynamic day-to-day scheduling of patients to appointment days is the main approach to control indirect waiting (McManus et al. (2003)). Two main paradigms of dynamic scheduling are *allocation scheduling* and *advance scheduling*. In allocation scheduling, patients wait to be notified until the appointment day while in advance scheduling, patients receive the appointment time upon request (Truong (2015)).

There is a vast literature focusing on models for allocation scheduling. Gerchak et al. (1996) employ a stochastic dynamic programming model to characterize an optimal policy for allocating capacity to regular and emergency surgeries. Min and Yih (2010) provide a structural analysis of a multi-priority scheduling problem to understand the properties of the optimal policy. Ayvaz and Huh (2010) study a system of multi-class customers in the sense of reactions to the delays in service, and propose a simple threshold heuristic policy. Huh et al. (2013) focus on allocating multiple resources to two classes of elective and emergency jobs. Several computationally-efficient policies are developed and their performances are examined by numerical experiments. Min and Yih (2014) propose a scheduling procedure for resource allocation to multiple classes of patients with time-dependent priority.

The alternative approach of considering the problem as a advance scheduling one received significant attentions as well. For example, Patrick et al. (2008) use the framework of Markov Decision Processes (MDP) for scheduling patients of different priorities. To tackle the "curse of dimensionality", they utilize approximate dynamic programming and develop heuristic methods for finding optimal control policies. Similarly, Gocgun and Ghate (2012) leverage an approximate dynamic programming method together with a Lagrangian relaxation. Efficiency of the proposed approximation is corroborated by performing extensive numerical experiments. Taking into account no-shows and cancellations, Liu et al. (2010) use data of an actual clinic to provide efficient dynamic policies for patient scheduling. Feldman et al. (2014) also develop heuristics to solve the scheduling problem with no-shows, considering the patient preferences regarding the appointment times.

More recently, Truong (2015) considers a model of advance scheduling with two urgent and regular demand classes. The patients of urgent class need to be served on the arrival day. However, the regular patients may receive an appointment time in the future. She proposes an algorithm to efficiently compute the exact optimal policy. Specifically, she shows that in the case of scheduling one class in advance, the cost of optimal advance scheduling is identical to the cost of optimal allocation scheduling. Parizi and Ghate (2016) provide a Markov decision process model for a class of advance scheduling problems with no-shows, cancellations, and overbooking. Wang and Truong (2018) develop an online algorithm to solve the problem of multi-priority patient scheduling with cancellations. Sauré et al. (2020) focus on a multi-class multi-priority patient scheduling problem and consider both deterministic and stochastic service times. They perform a comprehensive numerical investigation based on approximate dynamic programming to evaluate the performance of the proposed method(s).

1.2 Preliminaries and Notations

For $k, n \in \mathbb{N}$, by $C^k(\mathbb{R}^n)$, we denote the set of real-valued functions on $\mathbb{R}^n$ that their first k derivatives exist and are continuous. Further, let $\mathcal{D}^K[0,\infty)$ be the space of K-dimensional real-valued functions on $[0,\infty)$ that are Right-Continuous and the Left Limit exists (RCLL). We assume that $\mathcal{D}^K[0,\infty)$ is equipped with the Skorokhod J_1 topology[3].

For a sequence of stochastic processes $\{X_k\}$ and a stochastic process X, we denote the convergence in distribution by $X_k \Rightarrow X$. It is equivalent to the following: letting $\{P_k\}$ be the distributions of $\{X_k\}$ and P be the distributions of X, then $\int_{\mathbb{R}^K} \varphi \mathrm{d}P_k \to \int_{\mathbb{R}^K} \varphi \mathrm{d}P$ as $k \to \infty$, for all bounded and continuous functions φ in $\mathcal{D}^K[0,\infty)$[4]. By $[K]$ we refer to the set of all integers from 1 to K; $[K] := \{1, 2, \ldots, K\}$.

Next, for $t \in \mathbb{R}$, let t^- be the left-limit of the identity function at the point t. Define $\mathbb{R}_+^K := \left\{x \in \mathbb{R}^K : x \geq 0\right\}$, and denote the norm of the vector $x \in \mathbb{R}^K$ by $|x| = \sqrt{\sum_{k=1}^K x_k^2}$. For two real numbers a and b, let $a \wedge b = \min(a,b)$, as well as $a \vee b = \max(a,b)$ and $a^+ = \max\{a, 0\}$. We denote by β an arbitrary positive constant. For a function $\varphi \in C^2\left(\mathbb{R}^{N+H}\right)$ with the arguments $(x,y) \in \mathbb{R}^{N+H}$, let $\partial_{x_n}\varphi$ be the partial derivative with respect to the n-th argument x_n, while $\partial_{y_h}\varphi$ is the partial derivative with respect to the $(N+h)$-th argument y_h. We show the second order partial derivative of φ with respect to the n-th argument by $\partial_{x_n}^2\varphi$. Finally, $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra of the real line, which is the smallest σ-algebra that contains all open sets of $\mathbb{R}$.

[3] more details can be found in Section 3.3 of Whitt (2002)

[4] for details, see page 99 of Chen and Yao (2001)

Chapter 2

Problem Formulation

Consider a service system providing service to N classes of patients during the time window $[0, T]$, for some $T \in \mathbb{N}$. After arriving to the system, each patient waits to get an appointment time at which he will return to receive the corresponding service. It is assumed that in each period, patients can be scheduled to the next H periods, including the current period. Therefore, at any time t, we have a *waiting queue* vector $W(t) \in \mathbb{R}^N$ to capture the states of N queues, each of which corresponding to one of the patient classes. That is, $W_n(t)$ is the number of patients of class n waiting to receive their appointment time slot. Further, there is a scheduling table, denoted by $S(t) \in \mathbb{R}^{N \times H}$, with N rows (i.e., the number of classes) and H columns (i.e., the booking window), that reflects the scheduling decisions made up to time t.

Since all the scheduled patients will be eventually served, we do not need to distinguish the classes of patients after they receive their appointment times. That is, we only need to keep track of the capacities occupied by the scheduled patients over the next H days, stored in the vector $(Q_1, \ldots, Q_H)$. Note that, the booking system is dynamic in the sense that at the end of every day, the vector Q need to shift one step to the left, i.e., $Q_h \leftarrow Q_{h+1}$, for all $h < H$. Therefore, today's record (Q_1) disposes, and the records of the *new* last day (i.e., Q_H,) falls to zero because a new day becomes available to be used for scheduling patients on.

Let a_n^i denote the i-th inter-arrival time for class n patients, which has the mean value of $1/\lambda_n$ and the standard deviation of v_n^a. We assume that the inter-arrival times are i.i.d. within each class and independent between different classes. Let T_n^i denote the arrival time of the i-th patient from class n. So, we have $T_n^0 = 0,\ T_n^i = \sum_{j=1}^{i} a_n^j$, for all $n \in [N]$. Moreover, let $A_n(t)$ represent the number of class n patients who arrived by time t. Then, $A_n(t)$ satisfies

$$A_n(t) = \max\{i \geq 0 : T_n^i \leq t\}. \tag{2.1}$$

Suppose that p_n^i denotes the time that the i-th patient from class n receives his appointment date $\alpha_n^i \in [H]$. We define the *control signal* $S_{n,h}(t)$ as the number of class n patients scheduled up to time t to be served in h days after receiving their appointment time slot. Formally, for all $n \in [N]$ and $h \in [H]$, we define

$$S_{n,h}(t) = \#\{i \geq 0 : \alpha_n^i = h,\ p_n^i \leq t\}, \quad n \in [N],\ h \in [H]. \tag{2.2}$$

where $\#$ gives the number of elements in the set. Therefore, for $i \leq t_1 < t_2 < i+1$, the quantity $S_{n,h}(t_2) - S_{n,h}(t_1)$ indicates the number of class n patients who received an appointment time in the interval $(t_1, t_2]$ to be served on day $i+h$. Define $Q(i^-)$ as the number of patients scheduled to be served on day i (i.e., at some time t satisfying $i-1 \leq t < i$). Then, we have,

$$Q\left(i^-\right) = \sum_{n=1}^{N} \sum_{h=1}^{H} m_n \Big(S_{n,h}\left(i^- - h + 1\right) - S_{n,h}\left(i^- - h\right) \Big), \quad i = 1, 2, \ldots, T, \tag{2.3}$$

where $m := (m_1, \ldots, m_N)$ is the vector containing the average service time of different classes. Note that patients who receive appointment time h in time interval $[i-1, i)$ will be served in time interval $[i+h-2, i+h-1)$ that is days $i+h-1$. Then, aggregating the assignments of the last H days for different classes gives (2.3). Note also that, for all n and h, $S_{n,h} : \mathbb{R} \mapsto \mathbb{R}_+$ is a non-decreasing, RCLL function. Let $S_{n,h}$ vanish on $\mathbb{R}_-$. For the sake of brevity, hereafter, we eliminate the superscript of i in (2.3) and call $Q(i)$ as the *scheduled workload* of day i.

Based on (2.1) and (2.2), the dynamics of the waiting queues satisfy

$$W_n(t) = A_n(t) - \sum_{h=1}^{H} S_{n,h}(t), \quad t > 0, \ \forall n \in [N], \tag{2.4}$$

that is simply subtracting the scheduled patients from the total arrival. Note that we assume the system is initially empty so that the number of patients at time $t = 0$ is zero.

Now, we define the optimization criteria which is desired to be minimized by the scheduling process. Let $C_{0,n} \in \mathbb{R}_+$ be the waiting cost of a patient of class n for one unit of time before receiving an appointment. Denote also by $C_{n,h} \in \mathbb{R}_+$ the waiting cost of a class n patient who receives an appointment time for h days later. We use $u : \mathbb{R}_+ \mapsto \mathbb{R}_+$ to show the utility cost function that reflects the cost of resources and staffs for one day. The objective is to minimize the following total cost during the time horizon $[0, T]$:

$$\mathbb{E}\left[\sum_{n=1}^{N} \int_0^T C_{0,n} W_n^+(t) \mathrm{d}t + \sum_{n=1}^{N} \sum_{h=1}^{H} C_{n,h} S_{n,h}(T) + \sum_{i=1}^{T} u(Q(i))\right], \tag{2.5}$$

where $S(t) := \left[S_{n,h}(t)\right]_{n \in [N], h \in [H]}$ for all $0 \le t \le T$.

Assumption 2.0.1 *For all $n \in [N]$, we have $0 = C_{n,1} \le C_{n,2} \le \ldots \le C_{n,H}$.*

According to Assumption 2.0.1, the scheduler may schedule patients on the same day of their arrival at no cost. Generally, Assumption 2.0.1 expresses that $C_{n,h}$ is a non-decreasing function of h. This intuitively indicates that that, the system tends to give the patients the earliest affordable appointment time.

Assumption 2.0.2 *For all $n \in [N]$, we have $C_{0,n} > C_{n,H}$.*

Assumption 2.0.2 states that for all patients, the cost of waiting to receive an appointment time is more than the cost of waiting to be served after receiving the appointment. Thus, scheduling some patients in advance decreases the total expected cost of the system.

Assumption 2.0.3 *We have $u(x) := U\left((x-\kappa)^+\right)^2$, where U is a given constant and κ is the nominal daily capacity of the system.*

In general, the utility cost function u is required to be convex for a realistic setting. That is, the marginal cost of scheduling more patients for a given day is increasing as the scheduled workload of that day grows. Specifically, when we schedule above the nominal daily capacity (which is denoted by κ), scheduling each additional patient results a higher cost. For that reason, the decision-maker tends to schedule some patients on the next days with lower occupied capacity. Henceforth, we let $\kappa = \sum_{n=1}^{N} \lambda_n m_n$, which is the required daily capacity if the system experiences deterministic arrival rate of λ_n for class n, $n \in [N]$. Under this setting, u can also be interpreted as the overtime utility cost, in the sense that the system tolerates the utility cost only in case of scheduling above the nominal capacity.

Now, based on (2.2)-(2.5), we define the Multi-class Advance Scheduling Problem

(**MASP**) as follows,

$$
\begin{aligned}
&\inf_{S} \mathbb{E}\left[\sum_{n=1}^{N}\int_{0}^{T} C_{0,n} W_{n}^{+}(t)\mathrm{d}t + \sum_{n=1}^{N}\sum_{h=1}^{H} C_{n,h} S_{n,h}\left(T\right) + \sum_{i=1}^{T} U\left(\left(Q(i) - \sum_{n=1}^{N}\lambda_n m_n\right)^{+}\right)^{2}\right] \\
&\quad s.t. \\
&\quad W_n(t) = A_n\left(t\right) - \sum_{h=1}^{H} S_{n,h}\left(t\right), \quad t > 0,\ \forall n \in [N], \\
&\quad Q(i) = \sum_{n=1}^{N}\sum_{h=1}^{H} m_n\Big(S_{n,h}\left(i-h+1\right) - S_{n,h}\left(i-h\right)\Big), \quad i \in \mathbb{N}.
\end{aligned}
\tag{MASP}
$$

Taking into account Assumptions 2.0.1-2.0.3, the above definition of the cost function captures the fact that the scheduler aims to give the patients the earliest appointment to not exceed the daily scheduled workload Q too much from the nominal capacity $\sum_{n=1}^{N}\lambda_n m_n$.

Due to difficulties arising to solve this problem, we next introduce an asymptotic model that is tractable and helps us to characterize the optimal policy for the MASP.

Chapter 3

Asymptotic Analysis of MASP

In this section, we approximate the MASP by a Brownian Control Problem (**BCP**). To this end, we define a sequence of closely related systems such that in the k-th system we speed up the clock by k. Another interpretation of this scaling scheme is to scale up time by k. That is, each day in the k-th system is equal to k days in the original system. Under this interpretation, the scheduling window will be considered as $h = k, 2k, \ldots, Hk$, with respect to the initial system. A similar discussion is given in (Armony et al. 2019).

Speeding up the system leads to an increase in the system demand for appointment times; thus, we expect to have more appointment slots to be set for each day. More precisely, we assume that the n-th class arrival rate in the k-th system is $\lambda_n^k := k\lambda_n$. So, we rewrite the MASP for the k-th system by adding a superscript k to all quantities:

$$\inf_{S^k} \mathbb{E}\left[\sum_{n=1}^{N}\left(\int_0^T C_{0,n}^k W_n^{k,+}(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h}^k S_{n,h}^k(T)\right) + \sum_{i=1}^{T} U^k\left(\left(\left(Q^k(i) - \sum_{n=1}^{N}\lambda_n^k m_n^k\right)^+\right)^2\right]\right] \tag{MASP^k-1}$$

$$W_n^k(t) = A_n^k(t) - \sum_{h=1}^{H} S_{n,h}^k(t), \quad t > 0,\ n \in [N], \tag{MASP^k-2}$$

$$Q^k(i) = \sum_{n=1}^{N}\sum_{h=1}^{H} m_n^k\left(S_{n,h}^k(i-h+1) - S_{n,h}^k(i-h)\right), \quad i \in \mathbb{N}. \tag{MASP^k-3}$$

Note that, in the scaled system, the average service time vector remains the same, i.e., $m^k = (m_1^k, \ldots, m_N^k) = m$, while the nominal daily capacity increases with the rate k.

To study the behaviour of MASP^k as $k \to \infty$, consider a probability space $(\Omega, \mathbb{F}, \mathbb{P})$ with the N-dimensional standard Brownian Motion $B = (B_1, \ldots, B_N)$, and the natural filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$. We denote by $\mathbb{E}$ expectation with respect to $\mathbb{P}$.

We first investigate the fluid limit model of (MASP^k-1)-(MASP^k-3) as $k \to \infty$. The result of the fluid analysis provides the basis to define the diffusion approximation of the MASP and develop the corresponding BCP.

3.1 Fluid Model for MASP

In the fluid regime, the focus is on the first-order deterministic approximation of the MASP. We start by defining the fluid-scaled cumulative arrival process of the k-th system as:

$$\bar{A}_n^k(t) = \frac{A_n^k(t)}{k}. \tag{3.1}$$

By the Functional Law of Large Numbers for renewal processes (see e.g., Theorem 5.10 in Chen and Yao (2001)), as $k \to \infty$, (2.1) gives:

$$\bar{A}_n^k(t) \Rightarrow \lambda_n t. \tag{3.2}$$

As a result, for large values of k, it approximately holds that

$$A_n^k(t) \approx k\lambda_n t. \tag{3.3}$$

Denoting by $\overline{W}_n^k$ the fluid-scaled waiting queue of the n-th class patients, by (MASP^k-2) we have,

$$\overline{W}_n^k(t) = \frac{W_n^k(t)}{k} = \frac{A_n^k(t) - \sum_{h=1}^{H} S_{n,h}^k(t)}{k} = \bar{A}_n^k(t) - \sum_{h=1}^{H} \bar{S}_{n,h}^k(t),$$

where

$$\bar{S}_{n,h}^k(t) := \frac{S_{n,h}^k(t)}{k}, \quad n \in [N], h \in [H]. \tag{3.4}$$

Note that $\bar{S}_{n,h}^k(t)$ shows the fluid-scaled control matrix. Taking limit as $k \to \infty$ and using (3.2), we get,

$$\overline{W}_n(t) = \lambda_n t - \sum_{h=1}^{H} \bar{S}_{n,h}(t), \quad n \in [N], \tag{3.5}$$

where $\overline{W}_n(t)$ and $\bar{S}_{n,h}(t)$ are the limits of $\overline{W}_n^k(t)$ and $\bar{S}_{n,h}^k(t)$, respectively. The following lemma states the monotonicity of the control process with respect to time.

Lemma 3.1.1 *For all $n \in [N]$ and $h \in [H]$, the function $\bar{S}_{n,h}(\cdot)$ is non-decreasing.*

Proof. Appendix A.1.

Further, by (MASP^k-3), we can write the fluid-scaled scheduled workload of day i as,

$$\bar{Q}^k(i) = \frac{Q^k(i)}{k} = \sum_{n=1}^{N}\sum_{h=1}^{H} m_n \left(\bar{S}_{n,h}^k(i-h+1) - \bar{S}_{n,h}^k(i-h)\right).$$

Let $\bar{Q}(i)$ denote the limit of $\bar{Q}^k(i)$ as $k \to \infty$. So, we have

$$\bar{Q}(i) = \sum_{n=1}^{N}\sum_{h=1}^{H} m_n \left(\bar{S}_{n,h}(i-h+1) - \bar{S}_{n,h}(i-h)\right), \quad i = 1, 2, \ldots T. \tag{3.6}$$

We next derive the cost function of the fluid model. For large values of k, (MASP^k-1) yields to,

$$\begin{aligned}
&\sum_{n=1}^{N}\left(\int_0^T C_{0,n}^k W_n^{k,+}(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h}^k S_{n,h}^k\left(T\right)\right) + \sum_{i=1}^{T} U^k\left(\left(Q^k(i) - \sum_{n=1}^{N} k\lambda_n m_n\right)^+\right)^2 \\
&\approx \sum_{n=1}^{N}\left(\int_0^T C_{0,n}^k k\overline{W}_n^{k,+}(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h}^k k\bar{S}_{n,h}^k\left(T\right)\right) + \sum_{i=1}^{T} U^k\left(\left(k\bar{Q}^k(i) - \sum_{n=1}^{N} k\lambda_n m_n\right)^+\right)^2
\end{aligned} \tag{3.7}$$

To derive a meaningful cost function that admits finite Real values for the limit problem, we assume that the cost parameters vary with k as below:

$$C_{0,n}^k := \frac{C_{0,n}}{k}, \quad C_n^k := \frac{C_n}{k}, \quad U^k := \frac{U}{k^2}. \tag{3.8}$$

We can now provide the Fluid Multi-class Advance Patient Scheduling Problem (**FMASP**), based on relations (3.5)-(3.8) and Lemma 3.1.1, as follows:

$$\overline{V}_T = \min_{\bar{S}} \mathbb{E}\left[\sum_{n=1}^{N}\left(\int_0^T C_{0,n}\overline{W}_n^+(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h}\bar{S}_{n,h}\left(T\right)\right) + \sum_{i=1}^{T} U\left(\left(\bar{Q}(i) - \sum_{n=1}^{N}\lambda_n m_n\right)^+\right)^2\right] \tag{FMASP-1}$$

s.t.

$$\overline{W}_n(t) = \lambda_n t - \sum_{h=1}^{H} \bar{S}_{n,h}\left(t\right), \quad t > 0,\ n \in [N], \tag{FMASP-2}$$

$$\bar{Q}(i) = \sum_{n=1}^{N}\sum_{h=1}^{H} m_n\Big(\bar{S}_{n,h}\left(i-h+1\right) - \bar{S}_{n,h}\left(i-h\right)\Big), \quad i \in [T]. \tag{FMASP-3}$$

Note that in the fluid model, the system experiences a deterministic time-independent arrival rate λ_n for the n-th class as indicated in the constraint (FMASP-2). The following theorem characterizes the optimal scheduling policy for the fluid model.

Theorem 3.1.1 *Consider (FMASP-1)-(FMASP-3). Then, for all $T > 0$, we have $\overline{V}_T = 0$, and* *the optimal scheduling policy of can be characterized as*

$$\bar{S}^*_{n,h}(t) = \begin{cases} \lambda_n t, & h = 1 \\ 0, & h = 2, \ldots, H. \end{cases} \tag{3.9}$$

Proof. Appendix A.2.

As Theorem 3.1.1 shows, when there is no uncertainty in the demand, the optimal scheduling policy is to serve all the arrivals on the same day. In that case, no patient waits either before or after receiving an appointment time, and thus, all the waiting costs are zero. On the other hand, by assumption 2.0.3, scheduling some patients for the next days increases both the system utility cost and the patients waiting cost.

3.2 Diffusion Model for MASP

In this section, the goal is to take into account the uncertainty in the arrival process. Similar to (3.1), define the following scaled and centered version of the arrival process related to the k-th system:

$$\widehat{A}^k_n(t) = \frac{A^k_n(t) - k\lambda_n t}{\sqrt{k}}. \tag{3.10}$$

In view of Theorem 3.1.1, our focus is to derive deviations from optimal policy of the fluid model, which is serving all arrivals on the same day. To define the scaled cumulative scheduling control process of the k-th system, we are interested in the policies that satisfy the following relations as $k \to \infty$:

$$\frac{S^k_{n,h}(t)}{k} \approx \bar{S}^*_{n,h}(t)t, \quad \forall n \in [N], h \in [H], \tag{3.11}$$

where $\bar{S}^*$ is given in (3.9). That is, even in the diffusion model, to meet the average arrival rate and conserve the stability of the system for large values of k, the scheduling rate of the n-th class of patients should be around $k\lambda_n m_n$. We consequently define the following scaled and centered scheduling control processes:

$$\widehat{S}^k_{n,1}(t) = \frac{S^k_{n,1}(t) - k\lambda_n t}{\sqrt{k}}, \quad \forall n \in [N], \tag{3.12}$$

$$\widehat{S}^k_{n,h}(t) = \frac{S^k_{n,h}(t)}{\sqrt{k}}, \ h = 2, \dots, H, \quad \forall n \in [N]. \tag{3.13}$$

Applying Functional Central Limit Theorem (FCLT) for renewal processes (see e.g. Corollary 2.1 in Whitt (2016)), as $k \to \infty$, $\widehat{A}^k_n$ converges to a Brownian motion with zero drift and diffusion coefficient $\sigma_n := \lambda_n^{3/2} v^a_n$. That means, for large values of k, using (3.10), we have:

$$A^k_n(t) \approx k\lambda_n t + \sqrt{k}\sigma_n B_n(t), \quad n \in [N]. \tag{3.14}$$

Then, we define the scaled waiting queue as:

$$\begin{aligned} \widehat{W}^k_n(t) = \frac{W^k_n(t)}{\sqrt{k}} &= \frac{A^k_n(t) - \sum_{h=1}^H S^k_{n,h}(t)}{\sqrt{k}} \\ &= \frac{A^k_n(t) - kt\lambda_n - \sum_{h=1}^H S^k_{n,h}(t) + kt\lambda_n}{\sqrt{k}} \\ &= \widehat{A}^k_n(t) - \sum_{h=1}^H \widehat{S}^k_{n,h}(t). \end{aligned} \tag{3.15}$$

Let X_n and $\mathcal{S}_{n,h}$ denote the limits of $\widehat{W}^k_n$ and $\widehat{S}^k_{n,h}$ as $k \to \infty$, respectively. So, letting $k \to \infty$ in equation (3.15), we get:

$$X_n(t) = -\sum_{h=1}^H \mathcal{S}_{n,h}(t) + \sigma_n B_n(t), \quad n \in [N]. \tag{3.16}$$

Next, considering (2.3), (3.12), and (3.13), we define the following scaled and centered scheduled workload process as:

$$\begin{aligned} \widehat{Q}^k(i) &= \frac{Q^k(i) - \sum_{n=1}^N k\lambda_n m_n}{\sqrt{k}} \\ &= \frac{\sum_{n=1}^N k\lambda_n m_n + \sum_{n=1}^N \sum_{h=1}^H \sqrt{k} m_n \left(\widehat{S}^k_{n,h}(i-h+1) - \widehat{S}^k_{n,h}(i-h)\right) - \sum_{n=1}^N k\lambda_n m_n}{\sqrt{k}} \\ &= \sum_{n=1}^N \sum_{h=1}^H m_n \left(\widehat{S}^k_{n,h}(i-h+1) - \widehat{S}^k_{n,h}(i-h)\right). \end{aligned}$$

Denoting by $\mathcal{Q}(i)$ the limit of $\widehat{Q}^k(i)$ as $k \to \infty$, we get:

$$\mathcal{Q}(i) = \sum_{n=1}^{N} \sum_{h=1}^{H} m_n \left(\mathcal{S}_{n,h}\left(i-h+1\right) - \mathcal{S}_{n,h}\left(i-h\right) \right). \tag{3.17}$$

Note that (3.17) is in accordance with the definition of Q in (2.3). Specifically, both $\mathcal{Q}(i)$ and $Q(i^-)$ capture the total workload scheduled for day i in the associated system.

To derive the cost function (MASP^k-1) as $k \to \infty$, for large values of k, we can write,

$$\begin{aligned}\widehat{\phi}^k(S) &:= \sum_{n=1}^{N} \left(\int_0^T C^k_{0,n} W^{k+}_n(t)\mathrm{d}t + \sum_{h=1}^{H} C^k_{n,h} S^k_{n,h}\left(T\right) \right) + \sum_{i=1}^{T} U^k \left(\left(Q^k(i) - \sum_{n=1}^{N} \lambda^k_n m^k_n \right)^+ \right)^2 \\ &\approx \sum_{n=1}^{N} \left(\int_0^T C^k_{0,n} \sqrt{k} \widehat{W}^{k+}_n(t)\mathrm{d}t + \sum_{h=1}^{H} C^k_{n,h} \sqrt{k} \widehat{S}^k_{n,h}\left(T\right) \right) + \sum_{i=1}^{T} U^k \left(\sqrt{k} \left(\widehat{Q}^k(i) \right)^+ \right)^2.\end{aligned} \tag{3.18}$$

Similar to (3.8), to obtain the cost function of the diffusion problem, we define the following scaling scheme for the cost parameters:

$$C^k_{0,n} := \frac{C_{0,n}}{\sqrt{k}}, \quad C^k_n := \frac{C_n}{\sqrt{k}}, \quad U^k := \frac{U}{k}. \tag{3.19}$$

Let ϕ denote the limit of $\widehat{\phi}^k(S)$ as $k \to \infty$. Then, substituting (3.19) in (3.18) and letting $k \to \infty$, we get:

$$\phi(\mathcal{S}) = \sum_{n=1}^{N} \left(\int_0^T C_{0,n} X^+_n(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h} \mathcal{S}_{n,h}\left(T\right) \right) + \sum_{i=1}^{T} U \left(\mathcal{Q}(i) \right)^2. \tag{3.20}$$

Considering (3.16), (3.17), and (3.20), we have:

$$\inf_{\mathcal{S}} \mathbb{E} \left[\sum_{n=1}^{N} \left(\int_0^T C_{0,n} X^+_n(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h} \mathcal{S}_{n,h}\left(T\right) \right) + \sum_{i=1}^{T} U \left(\mathcal{Q}(i) \right)^2 \right] \tag{3.21a}$$

s.t.

$$X_n(t) = -\sum_{h=1}^{H} \mathcal{S}_{n,h}\left(t\right) + \sigma_n B_n(t), \quad n \in [N], \tag{3.21b}$$

$$\mathcal{Q}(i) = \sum_{n=1}^{N} \sum_{h=1}^{H} m_n \left(\mathcal{S}_{n,h}\left(i-h+1\right) - \mathcal{S}_{n,h}\left(i-h\right) \right), \quad i \in [T]. \tag{3.21c}$$

The mathematical formulation given in (3.21a)-(3.21c) represents a multi-class advance scheduling problem with Brownian noise. The evolution of the waiting queue X with respect to the control process and the Brownian motion is given by (3.21b), and (3.21c) determines the daily scheduled workload. Moreover, the cost function defined in (3.21a) reflects the scheduler inclinations; that is, assigning the arriving patients to the earliest possible appointment time, whereas keeping the scheduled capacity balanced over the booking horizon.

We next restate the problem defined in (3.21a)-(3.21c) as a Stochastic Control Problem, see e.g. Fleming and Soner (2006), Pham (2009), Touzi (2012). To this end, we assume that for all $n \in [N]$ and $h \in [H]$, $\mathcal{S}_{n,h}$ is continuous and differentiable with respect to time, and denote its derivative by $s_{n,h}$. Further, based on (3.13), Lemma 3.1.1 implies that $\mathcal{S}_{n,h}(\cdot)$ is a non-decreasing function of time for all $n \in [N]$ and $h = 2, 3, \ldots, H$. We assume that for all $n \in [N]$, $\mathcal{S}_{n,1}(\cdot)$ is non-decreasing as well. Regarding (3.21b), any decline in $\mathcal{S}_{n,1}(\cdot)$ over time translates into rescheduling some patients in the associated advance scheduling problem. We summarize the properties of the new control matrix s below.

Definition 3.2.1 *The set of* ***Admissible*** *policies $\mathcal{A}_\xi$ is the class of functions*

$$s(t, \omega) = [s_{n,h}(t, \omega)]_{n \in [N], h \in [H]},$$

such that for all $n \in [N]$ and $h \in [H]$, the function $s_{n,h} : \mathbb{R} \times \Omega \mapsto \mathbb{R}_+$ satisfies the following conditions:

(a) $(t, \omega) \mapsto s_{n,h}(t, \omega)$ is a $\mathcal{B}(\mathbb{R}_+) \times \mathcal{F}$-measurable function.

(b) $s_{n,h}(t, \omega)$ is $\mathcal{F}_t$-adapted.

(c) If $t < 0$, then $s_{n,h}(t) = 0$ for all $n \in [N]$, $h \in [H]$.

(d) For all $n \in [N]$, $h \in [H]$, and $t \geq 0$, we have $0 \leq s_{n,h}(t) \leq \xi$.

The last step before stating the Brownian Multi-class Advance Scheduling Problem is rewriting the utility cost so that it will be compatible with the Markov problem setting.

Lemma 3.2.1 *Let $\mathcal{Q}(i)$ be the scheduled daily capacity for day i as defined in (3.21c). Then, for any $\xi > 0$, it holds that*

$$\sum_{i=1}^{T} \mathcal{Q}(i)^2 = \sum_{n=1}^{N} \sum_{h=1}^{H} \int_0^{T-h+1} 2m_n Y_h(t) s_{n,h}(t)\,\mathrm{d}t, \tag{3.22}$$

where,,

$$Y_h(t) = \sum_{n=1}^{N} m_n \left(\int_{\lfloor t \rfloor}^{t} s_{n,h}(r)\,\mathrm{d}r + \sum_{l=h+1}^{H} \int_{i-l}^{i-l+1} s_{n,l}(r)\,\mathrm{d}r \right).$$

Proof. Appendix A.3.

In words, $Y_h(t)$ is the scheduled workload for day $\lfloor t \rfloor + h$ up to time t that satisfies[1],

$$\big(Y_1(i), Y_2(i), \ldots, Y_{H-1}(i), Y_H(i)\big) = \big(Y_2(i^-), Y_3(i^-), \ldots, Y_H(i^-), 0\big), \quad i > 0, i \in [T]. \tag{3.23}$$

For any $i \in [T]$, the quantity $\mathcal{Q}(i)$ includes the control process $s(t)$ such that $t \in [i-H, i)$, see (3.21c). That means, at any time $t \in [i-H, i)$, the scheduler decision $s(t)$ depends on future realizations of the Brownian motion. This is in contradiction to the definition of admissible policies, see Definition 3.2.1. However, Lemma 3.2.1 enables us to use the equivalent marginal cost rate.

Now, we can state the Brownian Multi-class Advance Scheduling Problem **(BMASP)** in an infinite time horizon. Toward this end, we introduce the discount rate $\gamma > 0$. Then, according to

[1]Day i corresponds to the time interval $i - 1 \leq t < i$.

(3.21a)-(3.21c), Definition 3.2.1, and Lemma 3.2.1, for any $\xi > 0$, we have,

$$\inf_{s\in\mathcal{A}_\xi} \mathbb{E}\left[\sum_{n=1}^{N}\int_0^\infty e^{-\gamma t}\left(C_{0,n}X_n^+(t)+\sum_{h=1}^{H}\left(C_{n,h}+2Um_nY_h(t)\right)s_{n,h}(t)\right)\mathrm{d}t\right] \tag{BMASP-1}$$

s.t.

$$\mathrm{d}X_n(t) = -\sum_{h=1}^{H} s_{n,h}(t)\,\mathrm{d}t + \sigma_n \mathrm{d}B_n(t), \quad n\in[N], \tag{BMASP-2}$$

$$\mathrm{d}Y_h(t) = \sum_{n=1}^{N} m_n s_{n,h}(t)\,\mathrm{d}t, \quad h\in[H], \tag{BMASP-3}$$

$$\left(Y_1(i), Y_2(i), \ldots, Y_{H-1}(i), Y_H(i)\right) = \left(Y_2\left(i^-\right), Y_3\left(i^-\right), \ldots, Y_H\left(i^-\right), 0\right), \quad i\in\mathbb{N}. \tag{BMASP-4}$$

To interpret the formulation BMASP given above, note that (BMASP-1) captures the optimization criteria of an appointment system over an infinite time horizon. Relation (BMASP-2) determines the evolution of the waiting queue, where the drift term reflects the scheduler decision, and the diffusion term captures the arrivals randomness as a Brownoian motion. Finally, (BMASP-3) gives the dynamics of the scheduled workload process and (BMASP-4) reflects the rolling scheduling horizon effect which opens new daily capacity to the organizer.

Taking into account the dynamics of the waiting queue, which is given in (BMASP-2), the first term of the cost function encourages the system to schedule more patients over the booking horizon. The second term of the cost function, in view of the system dynamics (BMASP-2) and (BMASP-3), indicates the scheduling trade-off. That is, the utility cost term U prevents the system to schedule many patients for each day. However, the term $C_{n,h}$ makes the later appointments more costly for the system, in lights of Assumption 2.0.1.

Recall that the fluid model discussed in Section 3.1 does not capture the value of the advance scheduling due to failure in incorporating the uncertainty in the arrival process. However, the BMASP captures the trade-off between scheduling patients in the current period versus scheduling them in a later period considering the uncertainty in the future arrivals.

Chapter 4

The BMASP Solution

This section develops a dynamic programming approach to study the BMASP defined in (BMASP-1)-(BMASP-4). We restrict our attention to the Markovian control policies. That is, at any time $t \geq 0$, the scheduling policy depends on the history of the system only through the current state of the system (the state of the system at time t). Let fix $\xi > 0$ throughout this section. We have the following definition.

Definition 4.0.1 *A measurable function $\pi : \mathbb{R}^N \times \mathbb{R}_+^H \mapsto \mathbb{R}_+^{N \times H}$ is a stationary Markovian control if for all $(x, y) \in \mathbb{R}^N \times \mathbb{R}_+^H$, $\pi(x, y)$ is an admissible policy, i.e., $\pi(x, y) \in \mathcal{A}_\xi$, that $\mathcal{A}_\xi$ is defined in Definition 3.2.1.*

We denote by Π_ξ the set of all stationary Markovian control policies π. At any time $t \geq 0$, we can write

$$s(t) = \pi\left(X(t), Y(t)\right),$$

where $X = (X_1, \ldots, X_N)$ and $Y = (Y_1, \ldots, Y_H)$ are states of the system defined in (BMASP-2) and (BMASP-3). In fact, $\pi(x, y)$ is the scheduler decision whenever the system is in state (x, y). Note that for any fixed $\mathrm{s} \in [0, \xi]^{N \times H}$, the constant control $\pi\left(X(t), Y(t)\right) = \mathrm{s}$ is within Π_ξ.

Next, we examine existence and uniqueness results for the solutions of the differential equations describing the dynamics of the system in (BMASP-2)-(BMASP-4).

Lemma 4.0.1 *Fix an arbitrary policy $\pi \in \Pi_\xi$ and choose $r > 0$ arbitrarily. Suppose that the state of the system at time r is given by a $\mathcal{F}_r$-measurable random variable $(X(r), Y(r))$ valued in $\mathbb{R}^N \times \mathbb{R}_+^H$, such that*

$$\mathbb{E}\left[|X(r)|^2 + |Y(r)|^2\right] = \mathbb{E}\left[\sum_{n=1}^{N} X_n(r)^2 + \sum_{h=1}^{H} Y_h(r)^2\right] < \infty. \tag{4.1}$$

Then, there exists a unique[1] *strong solution to the set of following equations that starts from $(X(r), Y(r))$ at time r,*

$$\begin{cases} \mathrm{d}X_n(t) = -\sum_{h=1}^{H} s_{n,h}(t)\,\mathrm{d}t + \sigma_n \mathrm{d}B_n(t), & r < t \le \lfloor r \rfloor + 1,\ n \in [N], \\ \mathrm{d}Y_h(t) = \sum_{n=1}^{N} m_n s_{n,h}(t)\,\mathrm{d}t, & r < t \le \lfloor r \rfloor + 1,\ h \in [H], \end{cases} \tag{4.2}$$

where $s(t) = \pi(X(t), Y(t)), r \le t \le \lfloor r \rfloor + 1$. Furthermore, we have,

$$\mathbb{E}\left[\sup_{r \le t \le \lfloor r \rfloor + 1} \left\{|X(t)|^2 + |Y(t)|^2\right\}\right] \le \beta e^{\beta}\left(1 + \mathbb{E}\left[|X(r)|^2 + |Y(r)|^2\right]\right). \tag{4.3}$$

Proof. Appendix A.4.

By Lemma 4.0.1, starting at any time in a day, there is a stochastic process that reflects the system dynamics within that day. Further, this process depends only on the initial state of the system and the control policy. Lemma 4.0.1 establishes the existence and uniqueness of the stochastic process describing the system dynamics for every admissible control policy within the whole time horizon, as stated in the following theorem.

[1]This is the path-wise uniqueness which is equivalent to indistinguishability. The processes $\{\chi_t^1\}_{t\ge0}$ and $\{\chi_t^2\}_{t\ge0}$ are called indistinguishable if $\mathbb{P}\{\chi_t^1 = \chi_t^2, \forall t \ge 0\} = 1$.

Theorem 4.0.1 *For any policy $\pi \in \Pi_\xi$ and any initial state $(x, y) \in \mathbb{R}^N \times \mathbb{R}_+^H$, there exists a unique strong solution to (4.4).*

$$
\begin{cases}
X(0) = x, \; Y(0) = y, & \\
\mathrm{d}X_n(t) = -\sum_{h=1}^{H} s_{n,h}(t)\,\mathrm{d}t + \sigma_n \mathrm{d}B_n(t), & t > 0, \; n \in [N], \\
\mathrm{d}Y_h(t) = \sum_{n=1}^{N} m_n s_{n,h}(t)\,\mathrm{d}t, & t > 0, \; h \in [H], \\
\big(Y_1(i), Y_2(i), \ldots, Y_{H-1}(i), Y_H(i)\big) = \left(Y_2(i^-), Y_3(i^-), \ldots, Y_H(i^-), 0\right), & i \in \mathbb{N}.
\end{cases}
\tag{4.4}
$$

where $s(t) = \pi(X(t), Y(t)), t \geq 0$.

Proof. Appendix A.5.

Now, for any policy $\pi \in \Pi_\xi$ and any $(x, y) \in \mathbb{R}^N \times \mathbb{R}_+^H$, we denote by $J(\pi; x, q)$ the *cost functional* of the BMASP. That means,

$$
J(\pi; x, y) = \mathbb{E}_{x,y}^{\pi}\left[\sum_{n=1}^{N} \int_0^\infty e^{-\gamma t}\left(C_{0,n} X_n^+(t) + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n Y_h(t)\right) \pi\left(X(t), Y(t)\right)\right) \mathrm{d}t\right],
\tag{4.5}
$$

where $\mathbb{E}_{x,y}^{\pi}[\cdot]$ denotes the expectation with respect to the probability distribution on the path space of (X, Y) which is the solution of (4.4) corresponds to the control policy π and the initial state (x, y). When there is no ambiguity about the control policy and the initial state, we simply use $\mathbb{E}$. We also define the *value function* V as follows,

$$
V(x, y) = \inf_{\pi \in \Pi_\xi} J(\pi; x, y).
\tag{4.6}
$$

Given an initial state (x, y), the quantity of the value function is the minimum operating cost of the system. The cost function $J(\pi; x, y)$ is always non-negative; that means, for all $\pi \in \Pi_\xi$ and $(x, y) \in \mathbb{R}^N \times \mathbb{R}_+^H$, we have $J(\pi; x, y) \geq 0$. Therefore, V is well defined as an extended positive real number, i.e., $V \in (\mathbb{R}_+ \cup \infty)$. Further, π^* is called an *optimal control policy* for the BMASP, if $V(x, y) = J(\pi^*; x, y)$. Now, the main goal is to characterize the optimal policy π^*. To this end, we state the following Dynamic Programming Principle (**DPP**).

Theorem 4.0.2 (Dynamic programming principle) *Assume that the value function V, defined in (4.6), is continuous. Let X and Y satisfy (BMASP-2)-(BMASP-4) and $(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H$ be fixed. Then, for any stopping time τ valued in $[0,\infty]$, we have*

$$\begin{aligned} V(x,y) = \inf_{\pi \in \Pi_\xi} \mathbb{E}\Bigg[\sum_{n=1}^{N} \int_0^\tau e^{-\gamma t} \left(C_{0,n} X_n^+(t) + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n Y_h(t)\right) s_{n,h}(t) \right) \mathrm{d}t \\ + e^{-\gamma\tau} V(X(\tau), Y(\tau)) \Bigg], \end{aligned} \tag{DPP}$$

where $s(t) = \pi(X(t), Y(t)), t \geq 0$.

Proof. Appendix A.6.

According to Theorem 4.0.2, we can split the BMASP into two parts with respect to the time horizon. Specifically, we solve a control problem that starts from time τ, given the system state $(X(\tau), Y(\tau))$. Then, considering the result of optimization for after τ, we find the optimal policy over the time interval $[0,\tau]$. Our goal is to characterize the optimal policy by finding the infinitesimal version of the DPP as $\tau \to 0$. We have Lemma 4.0.2 that is a straightforward application of Itô Formula (see e.g. Theorem 4.2.1 of Oksendal (2013)).

Lemma 4.0.2 *Fix an admissible control input $s \in \mathcal{A}_\xi$. For any function $v \in C^2\left(\mathbb{R}^{N\times H}\right)$ and any $i \leq t_1 < t_2 < i+1$, for some $i \in (\mathbb{N} \cup 0)$, we have,*

$$\begin{aligned} &\mathbb{E}\left[e^{-\gamma t_2} v\left(X(t_2), Y(t_2)\right) - e^{-\gamma t_1} v\left(X(t_1), Y(t_1)\right)\right] \\ =& \mathbb{E}\left[\int_{t_1}^{t_2} e^{-\gamma t} \left(\sum_{n=1}^{N} \sum_{h=1}^{H} \left(m_n \partial_{q_h} v\left(X(t), Y(t)\right) - \partial_{x_n} v\left(X(t), Y(t)\right)\right) s_{n,h}(t) \right) \mathrm{d}t \right] \\ &+ \mathbb{E}\left[\int_{t_1}^{t_2} e^{-\gamma t} \left(\frac{1}{2} \sum_{n=1}^{N} \sigma_n^2 \partial_{x_n}^2 v\left(X(t), Y(t)\right) - \gamma v\left(X(t), Y(t)\right) \right) \mathrm{d}t \right]. \end{aligned} \tag{4.7}$$

Proof. Appendix A.7.

Next, we define a notation that will be used in the following discussions. For any function $\varphi \in C^2\left(\mathbb{R}^{N+H}\right)$, define,

$$\delta^{\varphi}_{n,h}(x,y) = \partial_{x_n}\varphi(x,y) - C_{n,h} - 2Um_n y_h - m_n \partial_{y_h}\varphi(x,y). \tag{4.8}$$

Now, we state Hamilton-Jacobi-Bellman equation (**HJB**)[2] that describes the local behaviour of the value function defined in (4.6).

Theorem 4.0.3 (Hamilton-Jacobi-Bellman) *Suppose that $V \in C^2\left(\mathbb{R}^{N+H}\right)$, where V is the value function defined in (4.6). Then, V is a solution of the HJB equation, i.e., for all $(x,y) \in \mathbb{R}^N \times \mathbb{R}^H_+$, we have*

$$\sum_{n=1}^{N}\left(C_{0,n}x_n^+ + \frac{1}{2}\partial^2_{x_n}V(x,y)\sigma_n^2 + \inf\left\{\sum_{h=1}^{H} -\delta^V_{n,h}(x,y)\mathrm{s}_{n,h}\right\}\right) = \gamma V(x,y), \tag{HJB}$$

that the infimum is taken over all constant inputs $\mathrm{s} \in [0,\xi]^{N\times H}$.

Proof. Appendix A.8.

Theorem 4.0.3 demonstrates that the value function of the BMASP satisfies the partial differential equation given in (HJB). We show that there is a measurable function $\pi^\star = [\pi^\star_{n,h}]_{n\in[N],h\in[H]}$ such that for any $(x,y) \in \mathbb{R}^N \times \mathbb{R}^H_+$, the n-th row of $\pi^\star$ (i.e., $\pi^\star_n = [\pi^\star_{n,1},\ldots,\pi^\star_{n,H}]$) is a solution to the following optimization problem:

$$\min_{\mathrm{s}_n\in[0.\xi]^H}\left\{\sum_{h=1}^{H} -\delta^V_{n,h}(x,y)\mathrm{s}_{n,h}\right\}, \quad (x,y)\in\mathbb{R}^N\times\mathbb{R}^H_+,\ n\in[N]. \tag{4.9}$$

Since (4.9) defines a linear optimization problem with respect to the control s, we can find $\pi^\star(x,y)$ for any given state (x,y).

Theorem 4.0.4 *Suppose that $V \in C^2\left(\mathbb{R}^{N+H}\right)$, and define*

$$\pi^\star_{n,h}(x,y) = \begin{cases} \xi, & \delta^V_{n,h}(x,y) \geq 0, \\ 0, & \delta^V_{n,h}(x,y) < 0. \end{cases} \tag{4.10}$$

Then, $\pi^\star$ is a Borel measurable solution to the optimization problem given in (4.9).

[2] It is also called the dynamic programming equation.

Proof. Appendix A.9.

Recall that $s_{n,h}(t)$ is the rate of assigning class n patients to day $\lfloor t \rfloor + h$ at time t (see (2.2)). According to (4.10), for all $n \in [N]$ and $h \in [H]$, if at time t the coefficient of $\mathrm{s}_{n,h}$ is non-positive, $\pi^\star$ schedules the class n patients to day $\lfloor t \rfloor + h$ with the highest possible rate ξ. Otherwise, $\pi^\star$ does not assign any patient to day $\lfloor t \rfloor + h$ at the current time. We state the following verification theorem that justifies $\pi^\star$ as an optimal Markovian policy for the BMASP.

Theorem 4.0.5 (Verification theorem) *Let v be a function in $C^2(\mathbb{R}^{N+H})$ that satisfies a quadratic growth condition as follows*

$$\exists \beta > 0, \quad |v(x,y)| \leq \beta \left(1 + |x|^2 + |y|^2\right), \quad \forall (x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H . \tag{4.11}$$

Suppose that,

$$\sum_{n=1}^{N} \left(C_{0,n} x_n^+ + \frac{1}{2} \partial_{x_n}^2 v(x,y) \sigma_n^2 + \inf \left\{ \sum_{h=1}^{H} -\delta_{n,h}^v(x,y) \mathrm{s}_{n,h} \right\} \right) = \gamma v(x,y), \quad (x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H, \tag{4.12}$$

$$v\left(X(i), Y(i)\right) - v\left(X(i), Y\left(i^-\right)\right) = \sum_{h=1}^{H} \partial_{y_h} v\left(X(i), Y\left(i^-\right)\right) \Delta Y_h(i), \quad i \in \mathbb{N}, \tag{4.13}$$

$$\lim_{T \to \infty} e^{-\gamma T} \mathbb{E}\left[v(X(T), Y(T))\right] = 0. \tag{4.14}$$

where the infimum in (4.12) is taken over all constant inputs $\mathrm{s} \in [0, \xi]^{N \times H}$ and $\Delta Y_h(i) = Y_h(i) - Y_h(i^-)$. Then, $v(x,y) \leq V(x,y)$ for all $(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H$, where V is the value function defined in (4.6).

Further, fix the initial state (x_0, y_0) and let (X^*, Y^*) denote the solution of (BMASP-2)-(BMASP-4) corresponding to the Markovian policy $\pi^\star$ defined in (4.10) and the initial state (x_0, y_0). If

$$\lim_{T \to \infty} e^{-\gamma T} \mathbb{E}\left[v(X^*(T), Y^*(T))\right] = 0, \tag{4.15}$$

then $v(x_0, y_0) = V(x_0, y_0)$ and $\pi^\star$ is an optimal Markovian policy in Π_ξ.

Proof. Appendix A.10

Considering Theorem 4.0.5, we interpret $\pi^\star$ given in (4.10) as a threshold scheduling policy for the MASP and validate its performance numerically.

Chapter 5

Numerical Implementations

Taking into account Theorem 4.0.5, the optimal policy of the asymptotic multi-class advance scheduling problem, given in (BMASP-1)-(BMASP-4), is $\pi^\star$ defined in (4.10). This means that it is optimal to assign patients of class n to day h according to the sign of $\delta^V_{n,h}(x,y)$, given in (4.8), where (x,y) is the current state of the system. Now, we translate this result into the context of MASP. Recalling the definition of $\delta^V_{n,h}(x,y)$ given in (4.8), we have for any $n \in [N]$ and $h \in [H]$,

$$-\delta^V_{n,h}(x,y) = C_{n,h} + 2Um_n y_h + m_n \partial_{y_h} V(x,y) - \partial_{x_n} V(x,y). \tag{5.1}$$

Fixing n and h, we interpret $-\delta^V_{n,h}(x,y)$ as the marginal cost of scheduling one patient of class n on day h in the booking window, when the system is at state (x,y). The first term, $C_{n,h}$, is the waiting cost of a class n patient that receives an appointment for h day from now. The second term, $2Um_n y_h$, is the cost of assigning workload of m_n to day h with occupied capacity of y_h. Eventually, $m_n \partial_{y_h} V(x,y) - \partial_{x_n} V(x,y)$ is the marginal cost of moving one patient with the workload of m_n from x_n to y_h. Considering this interpretation of $-\delta^V_{n,h}(x,y)$ corresponding to the BMASP, we define function f to reflect the extra available capacity of $\sum_{n=1}^N \lambda_n m_n$ in MASP.

For all $n \in [N]$ and $h \in [H]$, let

$$f_{n,h}(x,y) = C_{n,h} + 2Um_n \left(q_h - \sum_{n=1}^{N} \lambda_n m_n \right)^+ + m_n \partial_{y_h} V(x,y) - \partial_{x_n} V(x,y). \quad (5.2)$$

Given the current state of the system (x, y) and values of $\partial_{x_n} V(x,y)$ and $\partial_{y_h} V(x,y)$ for all $n \in [N]$ and $h \in [H]$, our proposed scheduling policy is as follows. For each $n \in [N]$, find the set $\rho_n = \left\{ h \in [H] : f_{n,h}(x,y) \leq 0 \right\}$. If ρ_n is empty, let patients of class n wait. Otherwise, suppose that h_n^* is the smallest element of ρ_n. Assign a patient of class n to day h_n^*. As long as ρ_n is not empty, repeat this procedure to schedule patients of class n who are waiting to get their appoitnment date.

To investigate the performance of the proposed policy, we simulate a system with four classes of patients. We assume that the arrival of each class follows a Poisson process with rates $\lambda = (10, 20, 30, 40)$. The vector of waiting cost before receiving an appointment time is $C_0 = (1000, 50, 20, 1)$. We also let

1. For class 1: $C_{1,1} = 0, C_{1,h} = C_{1,h-1} + 10, \quad h > 1,$

2. For class 2: $C_{1,1} = 0, C_{1,h} = C_{1,h-1} + 2, \quad h > 1,$

3. For class 3: $C_{1,1} = 0, C_{1,h} = C_{1,h-1} + 1, \quad h > 1,$

4. For class 4: $C_{1,1} = 0, C_{1,h} = C_{1,h-1} + 0.1, \quad h > 1.$

The average workload vector is set to be $m = (0.12, 0.1, 0.08, 0.05)$. Therefore, the nominal daily capacity of the system is $\sum_{n=1}^{N} \lambda_n m_n = 7.6$ and scheduling beyond this result in an extra cost according to the cost function in MASP.

Considering a booking window of 30 days, we simulate the system over a time horizon of 3000 days considering several values for the utility cost U and different initial system states. We compare the performance of the proposed scheduling policy against the fluid optimal policy which is clarified in Theorem 3.1.1. That is, assigning all the patients to the day of their arrival.

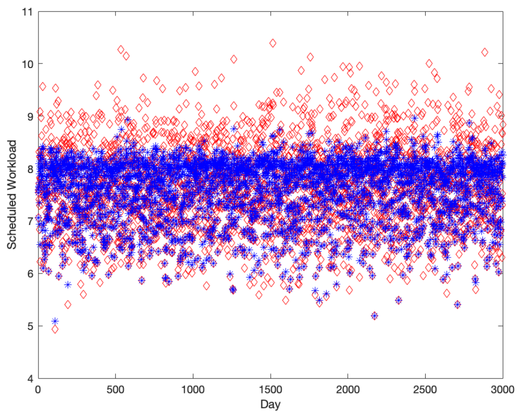
In order to numerically calculate f given in (5.2), we need the derivatives of the value function $\partial_{y_h} V$ and $\partial_{x_n} V$. We consider a linear approximation for these derivatives and find the coefficients using a brute-force search. To be more specific, we assume that $\partial_{y_h} V = \beta_1 U y_h$ and $\partial_{x_n} V = \beta_2 C_{0,n} x_n$ and seek for (β_1, β_2) that give the minimum operating cost for the simulated system.

5.1 Simulation Results

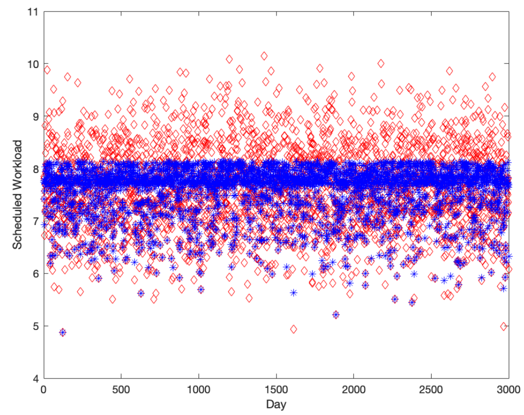
In this section, we study the performance of our scheduling policy by simulating an appointment system. Specifically, we show that employing the proposed policy results in the high utilization of the resources while provides patients with timely access to the required service. Figure 5.1 illustrates four appointment systems with the same system parameters except the utility cost coefficient. Blue stars in figures show the scheduled workload based on the proposed policy and the red diamonds show the scheduled daily workload resulted from assigning the patients to the day of their arrival. According to Figure 5.1, the proposed policy pushes the system to keep the daily balance within a smaller neighborhood around the nominal capacity as the utility cost increases.

The efficiency of the resource utilization also matters when the system needs to take into account a drop in the supply. Although we assumed that the system is empty at the beginning, it is not the case in healthcare facilities. In Figure 5.2, we examine the appointment system while 80% of the nominal capacity of the first 30 days is initially occupied. Our proposed policy assists the system to respond properly to this issue. In Figure 5.2b, the policy response to this pre-occupation issue is more smooth concerning the high utility cost. That is, the system resolves the extra demand issue over a longer time horizon and uses less overtime capacity. However, in view of the affordable utility cost in Figure 5.2a, the policy allows the system to utilize overtime capacity significantly.

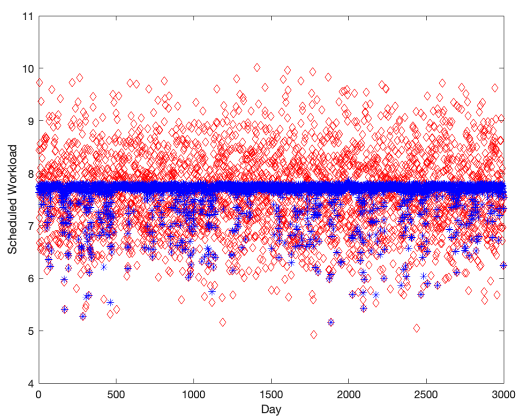
We also demonstrate the structure of the scheduling policy by simulating the system over 100



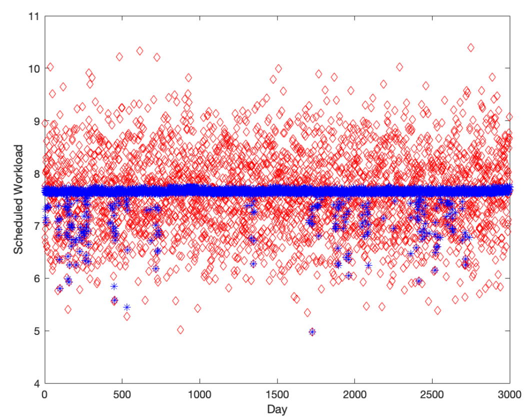
(a) U = 50



(b) U = 200

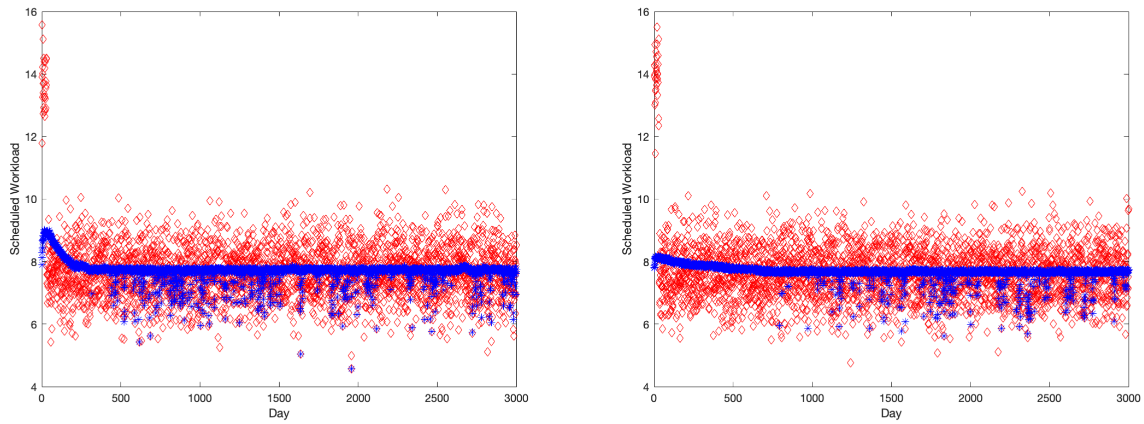


(c) U = 1000



(d) U = 10000

Figure 5.1: Results of implementing the proposed policy for different values of utility cost coefficient.



(a) U = 1000

(b) U = 4000

Figure 5.2: Results for an appointment system that 80% of the nominal capacity of the first 30 days is occupied at the beginning.

different sample paths of the arrival process. We assume that the system is initially fully occupied. That means, the nominal capacity of the first 30 days is unavailable to the scheduler. Figure 5.3 shows the average number of the scheduled patients over the booking window for different classes. According to Figure 5.3, more than 45 percent of patients received an appointment time for the arrival day, and more than 25 percent of patients were scheduled to be served within one day after getting the appointment time. The high cost of waiting before receiving an appointment time forces the system to schedule all patients just after their arrival time.

To compare the performance of our proposed policy with the case of scheduling all patients on the day of their arrival, we define the following cost ratio.

$$\text{cost ratio} = \frac{\text{System operating cost under the proposed policy}}{\text{System operating cost under the fluid policy}}. \tag{5.3}$$

According to (5.3), the lower cost ratio demonstrates a higher performance of the proposed policy comparing to the policy obtained based on the fluid analysis. We provided the result of simulating an appointment system for 200 different sample paths of arrivals in Figure 4. Specifically, Figure

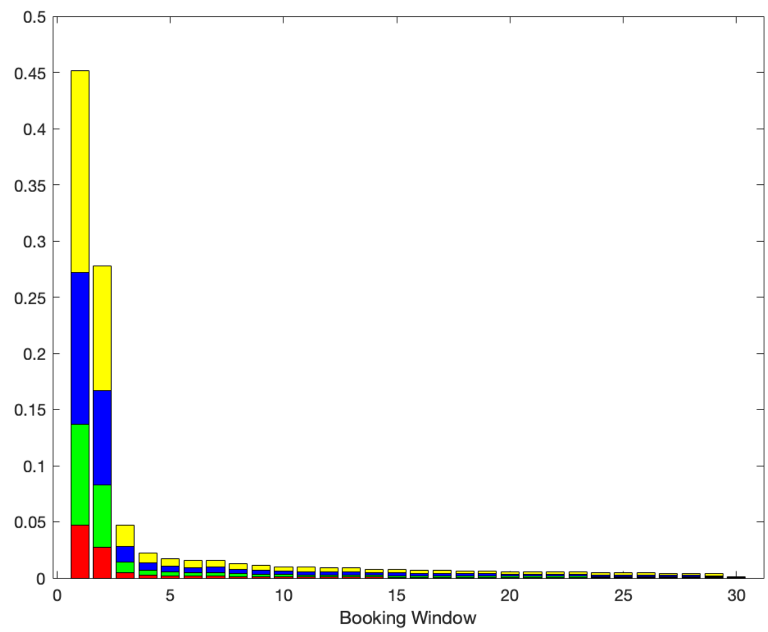
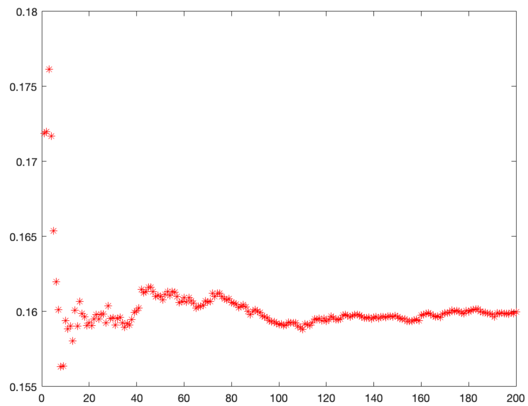
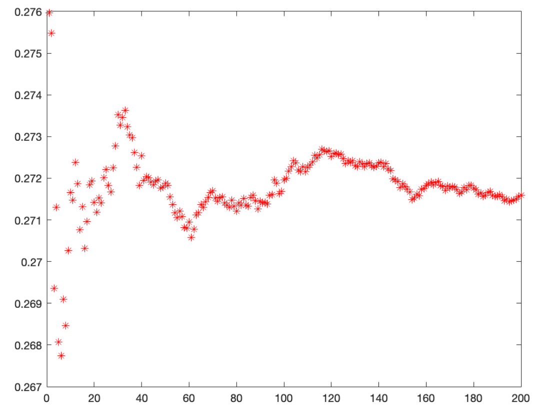


Figure 5.3: The average of the proposed policy over 100 different sample paths of the arrival process. The nominal capacity of the first 30 days is fully occupied. (red: class 1, green: class 2, blue: class 3, yellow: class 5)



(a) Initially empty system



(b) Initially fully occupied system

Figure 5.4: Comparing the average cost ratio (5.3) for 200 simulations

4 shows the average cost ratio of an initially empty system is around 0.16. In other words, employing our proposed policy decreases the system operating cost by as much as 84%. On the other hand, Figure 4 gives the result for an initially fully occupied system. The given policy still improves the system performance by 73%.

Chapter 6

Conclusion and Future Work

In this work, the first analytical results are provided for studying the problem of scheduling multiple classes of patients advance in time. Toward this aim, a realistic formulation of the problem is presented, which reflects the utility and waiting costs trade-off. Next, we formulate a theoretically tractable stochastic control problem that captures the asymptotic behavior of the system. We obtained an optimal policy relying on a proposed optimal solution to the corresponding fluid-regime problem.

Then, the presented results are interpreted in the original setting to design an efficient advance scheduling policy. At the last step, numerical implementations demonstrate the efficacy of the proposed policy to manage the daily fluctuations in demand while providing timely access to the desired service for patients. Moreover, the importance of taking into account the randomness of the arrival process is signified by comparing the system operating cost under optimal policies, in both fluid and diffusion regimes.

In future studies, we aim to incorporate the stochasticity of service times in the advance scheduling problem. Further extensions may consider the no-shows effect on the optimal policy, that is, to adapt the scheduling policy with the setting that some patients skip their appointment times. Investigating the consequences of the above extensions, as well as combinations of them, on the structure of optimal policies, will constitute an interesting direction for future work.

References

Amir Ahmadi-Javid, Zahra Jalali, and Kenneth J Klassen. Outpatient appointment systems in healthcare: A review of optimization studies. *European Journal of Operational Research*, 258(1):3–34, 2017.

Victor F Araman and Peter W Glynn. Fractional brownian motion with $h < 1/2$ as a limit of scheduled traffic. *Journal of Applied Probability*, 49(3):710–718, 2012.

Mor Armony, Rami Atar, and Harsha Honnappa. Asymptotically optimal appointment schedules. *Mathematics of Operations Research*, 44(4):1345–1380, 2019.

Nur Ayvaz and Woonghee Tim Huh. Allocation of hospital capacity to multiple types of patients. *Journal of Revenue and Pricing Management*, 9(5):386–398, 2010.

Dimitir P Bertsekas and Steven Shreve. *Stochastic optimal control: the discrete-time case*. 2004.

Tugba Cayirli and Emre Veral. Outpatient scheduling in health care: a review of literature. *Production and Operations Management*, 12(4):519–549, 2003.

Hong Chen and David D Yao. *Fundamentals of queueing networks: Performance, asymptotics, and optimization*, volume 4. Springer, 2001.

Brian Denton and Diwakar Gupta. A sequential bounding approach for optimal appointment scheduling. *IIE transactions*, 35(11):1003–1016, 2003.

Jacob Feldman, Nan Liu, Huseyin Topaloglu, and Serhan Ziya. Appointment scheduling under patient preference and no-show behavior. *Operations Research*, 62(4):794–811, 2014.

Wendell H Fleming and Halil Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

Yigal Gerchak, Diwakar Gupta, and Mordechai Henig. Reservation planning for elective surgery under uncertain demand for emergency surgery. *Management Science*, 42(3):321–334, 1996.

Yasin Gocgun and Archis Ghate. Lagrangian relaxation and constraint generation for allocation and advanced scheduling. *Computers & Operations Research*, 39(10):2323–2336, 2012.

Diwakar Gupta and Brian Denton. Appointment scheduling in health care: Challenges and opportunities. *IIE transactions*, 40(9):800–819, 2008.

Shlomo Halfin and Ward Whitt. Heavy-traffic limits for queues with many exponential servers. *Operations research*, 29(3):567–588, 1981.

Randolph W Hall et al. *Handbook of healthcare system scheduling*. Springer, 2012.

J Michael Harrison. Brownian models of queueing networks with heterogeneous customer populations. In *Stochastic differential systems, stochastic control theory and applications*, pages 147–186. Springer, 1988.

Refael Hassin and Sharon Mendel. Scheduling arrivals to queues: A single-server model with no-shows. *Management science*, 54(3):565–572, 2008.

Woonghee Tim Huh, Nan Liu, and Van-Anh Truong. Multiresource allocation scheduling in dynamic environments. *Manufacturing & Service Operations Management*, 15(2):280–291, 2013.

Alex Kuiper, Michel Mandjes, and Jeroen de Mast. Optimal stationary appointment schedules. *Operations Research Letters*, 45(6):549–555, 2017.

Nan Liu, Serhan Ziya, and Vidyadhar G Kulkarni. Dynamic scheduling of outpatient appointments under patient no-shows and cancellations. *Manufacturing & Service Operations Management*, 12(2):347–364, 2010.

Michael L McManus, Michael C Long, Abbot Cooper, James Mandell, Donald M Berwick, Marcello Pagano, and Eugene Litvak. Variability in surgical caseload and access to intensive care services. *The Journal of the American Society of Anesthesiologists*, 98(6):1491–1496, 2003.

Daiki Min and Yuehwern Yih. An elective surgery scheduling problem considering patient priority. *Computers & Operations Research*, 37(6):1091–1099, 2010.

Daiki Min and Yuehwern Yih. Managing a patient waiting list with time-dependent priority and adverse events. *RAIRO-Operations Research*, 48(1):53–74, 2014.

Bernt Oksendal. *Stochastic differential equations: an introduction with applications*. Springer Science & Business Media, 2013.

Mahshid Salemi Parizi and Archis Ghate. Multi-class, multi-resource advance scheduling with no-shows, cancellations and overbooking. *Computers & Operations Research*, 67:90–101, 2016.

Jonathan Patrick, Martin L Puterman, and Maurice Queyranne. Dynamic multipriority patient scheduling for a diagnostic resource. *Operations research*, 56(6):1507–1525, 2008.

Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.

P. Protter. *Stochastic Integration and Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2013. ISBN 9783662100615. URL https://books.google.ca/books?id=2a37CAAAQBAJ.

Martin I Reiman. Open queueing networks in heavy traffic. *Mathematics of operations research*, 9(3):441–458, 1984.

Walter Rudin. *Principles of mathematical analysis*, volume 3. McGraw-hill New York, 1964.

Antoine Sauré, Mehmet A Begen, and Jonathan Patrick. Dynamic multi-priority, multi-class patient scheduling with stochastic service times. *European Journal of Operational Research*, 280(1):254–265, 2020.

Nizar Touzi. *Optimal stochastic control, stochastic target problems, and backward SDE*, volume 29. Springer Science & Business Media, 2012.

Van-Anh Truong. Optimal advance scheduling. *Management Science*, 61:150211105729003, 02 2015. doi: 10.1287/mnsc.2014.2067.

Xinshang Wang and Van-Anh Truong. Multi-priority online scheduling with cancellations. *Operations Research*, 66(1):104–122, 2018.

Ward Whitt. *Stochastic-process limits: an introduction to stochastic-process limits and their application to queues*. Springer Science & Business Media, 2002.

Ward Whitt. Efficiency-driven heavy-traffic approximations for many-server queues with abandonments. *Management Science*, 50(10):1449–1461, 2004. ISSN 00251909, 15265501. URL http://www.jstor.org/stable/30046186.

Ward Whitt. A review of basic fclt's. 2016.

Christos Zacharias and Michael Pinedo. Appointment scheduling with no-shows and overbooking. *Production and Operations Management*, 23(5):788–801, 2014.

Appendix A

Proofs of Technical Results

A.1 Proof of Lemma 3.1.1

First, fix n and h. Then, according to (3.4), it is clear that $\bar{S}^k_{n,h}$ is non-decreasing. Thus, if $0 \leq t_1 \leq t_2 \leq T$, then

$$\bar{S}^k_{n,h}\left(t_1\right) \leq \bar{S}^k_{n,h}\left(t_2\right). \tag{A.1}$$

Taking limit by letting $k \to \infty$, the left-hand side of (A.1) gives $\bar{S}_{n,h}\left(t_1\right) \leq \bar{S}^k_{n,h}\left(t_2\right)$. Now, by taking a similar limit from the right-hand-side, we get the desired result. $\square$

A.2 Proof of Theorem 3.1.1

Fix $T > 0$ and set

$$\bar{J}(\bar{S}) = \mathbb{E}\left[\sum_{n=1}^{N}\left(\int_0^T C_{0,n}\overline{W}_n^+(t)\mathrm{d}t + \sum_{h=1}^{H} C_{n,h}\bar{S}_{n,h}\left(T\right)\right) + \sum_{i=1}^{T} U\left(\left(\bar{Q}(i) - \sum_{n=1}^{N}\lambda_n m_n\right)^+\right)^2\right]. \tag{A.2}$$

All terms of $\bar{J}(\bar{S})$ are non-negative for any arbitrary control process $\bar{S}$ that is non-negative and non-decreasing. It implies that $\bar{J}(\cdot) \geq 0$. As a result, we have $\overline{V}_T \geq 0$. Now, it is straightforward to check that under the control process $\bar{S}^*$ given by (3.9), for all $t \in [0, T]$ it holds that $\overline{W}_n(t) = 0$, and for all $i \in [T]$ we have $\bar{Q}(i) = \sum_{n=1}^{N} \lambda_n m_n$. Based on Assumption 2.0.1, we get $C_{n,1} = 0$ for all $n \in [N]$. Therefore, we have $\bar{J}(\bar{S}^*) = 0$. That is, $\bar{S}^*$ is the optimal policy for (FMASP-1)-(FMASP-3), under which it holds that $\overline{V}_T = 0$. $\square$

A.3 Proof of Lemma 3.2.1

Fix $i \in [T]$ and denote by $\mathcal{Q}_i(t)$ the scheduled capacity up to time t for the i-th day. For t with $\lfloor t \rfloor < i \leq \lfloor t \rfloor + H \leq T$, according to (3.17), it holds that,

$$\mathcal{Q}_i(t) = \sum_{n=1}^{N} m_n \left(\int_{\lfloor t \rfloor}^{t} s_{n,i-\lfloor t \rfloor}(r)\, \mathrm{d}r + \sum_{h=i-\lfloor t \rfloor+1}^{i \wedge H} \int_{i-h}^{i-h+1} s_{n,h}(r)\, \mathrm{d}r \right), \tag{A.3}$$

Then, by the definition, it holds that $\mathcal{Q}(i)^2 = \mathcal{Q}_i(i)^2$. We can write,

$$\mathrm{d}\mathcal{Q}_i(t)^2 = 2\mathcal{Q}_i(t)\mathrm{d}\mathcal{Q}_i(t). \tag{A.4}$$

Now, note that the scheduler may assign the patients to day i in time interval $[(i-H) \vee 0, i)$. Therefore, taking integral from both sides of (A.4) gives

$$\mathcal{Q}_i(i)^2 = \int_{(i-H)\vee 0}^{i} 2\mathcal{Q}_i(t)\mathrm{d}\mathcal{Q}_i(t) = \sum_{h=1}^{i \wedge H} \int_{i-h}^{i-h+1} 2\mathcal{Q}_i(t)\mathrm{d}\mathcal{Q}_i(t).$$

Based on definition 3.2.1 part (e), for any $t < 0$, $s_{n,h}(t) = 0$, $n \in [N], h \in [H]$. Then, considering (A.3), it gives,

$$\mathcal{Q}(i)^2 = \sum_{h=1}^{H} \int_{i-h}^{i-h+1} 2\mathcal{Q}_i(t)\mathrm{d}\mathcal{Q}_i(t) = \sum_{h=1}^{H} \sum_{n=1}^{N} 2m_n \int_{i-h}^{i-h+1} \mathcal{Q}_i(t) s_{n,i-\lfloor t \rfloor}(t)\, \mathrm{d}t.$$

Now if $i-h \leq t < i-h+1$, then $\lfloor t \rfloor = i-h$. Hence, using the definition of Y, we get

$$\mathcal{Q}(i)^2 = \sum_{h=1}^{H} \sum_{n=1}^{N} 2m_n \int_{i-h}^{i-h+1} Y_h(t) s_{n,h}(t) \, \mathrm{d}t.$$

Now, we can write

$$\begin{aligned}
\sum_{i=1}^{T} \mathcal{Q}(i)^2 &= \sum_{i=1}^{T} \sum_{h=1}^{H} \sum_{n=1}^{N} 2m_n \int_{i-h}^{i-h+1} Y_h(t) s_{n,h}(t) \, \mathrm{d}t \\
&= \sum_{i=1}^{T} \sum_{h=1}^{H} \sum_{n=1}^{N} 2m_n \int_{i-h}^{i-h+1} Y_h(t) s_{n,h}(t) \, \mathrm{d}t \\
&= \sum_{n=1}^{N} \sum_{h=1}^{H} \sum_{i=1}^{T} 2m_n \int_{i-h}^{i-h+1} Y_h(t) s_{n,h}(t) \, \mathrm{d}t \\
&= \sum_{n=1}^{N} \sum_{h=1}^{H} 2m_n \int_{0}^{T-h+1} Y_h(t) s_{n,h}(t) \, \mathrm{d}t.
\end{aligned}$$

□

A.4 Proof of Lemma 4.0.1

First, note that the drift and the diffusion coefficients of X defined in (BMASP-2) satisfy a uniform Lipschitz condition. That is, there exists a constant $\beta \geq 0$, such that for all $x_1, x_2 \in \mathbb{R}^N$ and for any constant control $\mathrm{s} \in \mathbb{R}_+^{N \times H}$, we have

$$\left| \sum_{h=1}^{H} \mathrm{s}_{n,h} - \sum_{h=1}^{H} \mathrm{s}_{n,h} \right| + |\sigma_n - \sigma_n| = 0 \leq \beta \, |x_1 - x_2| \,, \quad n \in [N].$$

Further, by Definition 3.2.1, for any $n \in [N]$, we can write,

$$\mathbb{E}\left[\int_0^T \left(\sum_{h=1}^{H} s_{n,h}(t) \right)^2 + \left(\sum_{h=1}^{H} m_n s_{n,h}(t) \right)^2 \mathrm{d}t \right] < \infty, \quad \forall T > 0.$$

Now, applying Theorem 3.1 of Touzi (2012) completes the proof. □

A.5 Proof of Theorem 4.0.1

Starting from (x, y), according to Lemma 4.0.1, there exists a process $\{X(t), Y(t)\}_{0\leq t\leq 1}$ that satisfies the relations in (4.2) for $r = 0$ and $X(0) = x, Y(0) = y$. Also, (4.3) implies that $\mathbb{E}\left[|X(1)|^2 + |Y(1)|^2\right] < \infty$. Now, define $x_1 = X(1)$ ans $y_1 = (Y_2(1), Y_3(1), \ldots, Y_H(1), 0)$. It holds that $\mathbb{E}\left[|x_1|^2 + |y_1|^2\right] < \infty$. Reusing Lemma 4.0.1 gives that there exists a process $\{X(t), Y(t)\}_{1\leq t\leq 2}$ that satisfies the desired conditions in (4.2). By repeating the procedure we obtain the process $\{X(t), Y(t)\}_{i\leq t\leq i+1}$ starting from (x_i, y_i) that satisfies (4.2) for any $i \in \mathbb{N}$. With a abuse of notations, define $\{X(t), Y(t)\}_{t>0} := \left\{\{X(t), Y(t)\}_{i<t\leq i+1}\right\}_{i\in\{0\cup\mathbb{N}\}}$ and $X(0) := x, Y(0) = y$. It is straightforward to check that $\{X(t), Y(t)\}_{t>0}$ is a solution to (4.4).

To see the uniqueness, let $\{\mathcal{X}(t), \mathcal{Y}(t)\}_{t\geq 0}$ be another solution to (4.4). Therefore, $\{\mathcal{X}(t), \mathcal{Y}(t)\}_{r\leq t<r+1}$ is a solution to (4.2) for $r \in \{0 \cup \mathbb{N}\}$. But, recall that the solution of (4.2) is unique. That means $\{X(t), Y(t)\}_{t\geq 0}$ and $\{\mathcal{X}(t), \mathcal{Y}(t)\}_{t\geq 0}$ are indistinguishable. $\square$

A.6 Proof of Theorem 4.0.2

Step 1. Using the tower property of expectations, by (4.5) we have,

$$\begin{aligned}
&J(\pi; x, y) \\
&= \mathbb{E}^{\pi}_{x,y}\left[\sum_{n=1}^{N}\int_0^{\infty} e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)s_{n,h}(t)\right)\mathrm{d}t\right] \\
&= \mathbb{E}^{\pi}_{x,y}\left[\sum_{n=1}^{N}\int_0^{\tau} e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)s_{n,h}(t)\right)\mathrm{d}t\right] \\
&+ \mathbb{E}^{\pi}_{x,y}\left[e^{-\gamma\tau}\mathbb{E}^{\pi}_{X(\tau),Y(\tau)}\left[\sum_{n=1}^{N}\int_{\tau}^{\infty} e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)s_{n,h}(t)\right)\mathrm{d}t\right]\right] \\
&= \mathbb{E}^{\pi}_{x,y}\left[\sum_{n=1}^{N}\int_0^{\tau} e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)s_{n,h}(t)\right)\mathrm{d}t + e^{-\gamma\tau}J(\pi; X(\tau), Y(\tau))\right]
\end{aligned} \tag{A.5}$$

In the last equality, we used the path-wise uniqueness of the solutions of (BMASP-2)- (BMASP-4) that implies

$$X^{(x,y)}(r) = X^{\left(X^{(x,y)}(\tau),Y^{(x,y)}(\tau)\right)}(r) \quad r \geq \tau,$$
$$Y^{(x,y)}(r) = Y^{\left(X^{(x,y)}(\tau),Y^{(x,y)}(\tau)\right)}(r), \quad r \geq \tau,$$

where $\left(X^{(x,y)}(r), Y^{(x,y)}(r)\right)$ is the solution to (4.4) starting from (x, y).

Note that by the definition of value function (4.6), we know that,

$$V(\cdot) \leq J(\pi; \cdot), \quad \forall \pi \in \Pi_\xi. \tag{A.6}$$

Now, (A.5) and (A.6) imply that,

$$J(\pi; x, y) \geq \inf_{\pi \in \Pi_\xi} \mathbb{E}^{\pi}_{x,y} \left[\sum_{n=1}^{N} \int_0^{\tau} e^{-\gamma t} \left(C_{0,n} X_n^+(t) + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n Y_h(t)\right) s_{n,h}(t) \right) \mathrm{d}t \right.$$
$$\left. + e^{-\gamma\tau} V(X(\tau), Y(\tau)) \right].$$

Takeing infimum over $\pi \in \Pi_\xi$ in the LHS, we get,

$$\begin{aligned} V(x, y) \geq \inf_{\pi \in \Pi_\xi} \mathbb{E}^{\pi}_{x,y} \Bigg[&\sum_{n=1}^{N} \int_0^{\tau} e^{-\gamma t} \left(C_{0,n} X_n^+(t) + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n Y_h(t)\right) s_{n,h}(t) \right) \mathrm{d}t \\ &+ e^{-\gamma\tau} V(X(\tau), Y(\tau)) \Bigg]. \end{aligned} \tag{A.7}$$

Step 2. To prove the reverse inequality, fix an arbitrary $s \in \mathcal{A}_\xi$ and τ. For any $\omega \in \Omega$ and $\varepsilon > 0$, let $s^{\omega,\varepsilon}$ be an ε-optimal control process for $V(X^\omega(\tau(\omega)), Y^\omega(\tau(\omega)))$. That means,

$$V(X^\omega(\tau(\omega)), Y^\omega(\tau(\omega))) + \varepsilon \geq J\left(\pi^\varepsilon; X^\omega(\tau(\omega)), Y^\omega(\tau(\omega))\right), \tag{A.8}$$

where $\pi^\varepsilon\left(X^\omega(t), Y^\omega(t)\right) = s^{\omega,\varepsilon}(t), t \geq 0$. We define $\widehat{\pi}$ as follows[1],

$$\widehat{\pi}\left(X^\omega(t), Y^\omega(t)\right) := \begin{cases} s^\omega(t), & t < \tau(\omega) \\ s^{\omega,\varepsilon}(t), & t \geq \tau(\omega) \end{cases}.$$

[1]The measurability of $\widehat{\pi}$ is the result of Measurable Selection Theorem (see e.g. Chapter 7 of Bertsekas and Shreve (2004)). For a similar approach, see Theorem 3.3.1 of Pham (2009).

Again by the tower property of expectation and similar to (A.5), we have,

$$J(\widehat{\pi};x,y) = \mathbb{E}_{x,y}^{\widehat{\pi}^{\omega}}\Bigg[\sum_{n=1}^{N}\int_0^{\tau} e^{-\gamma t}\left(C_{0,n}X_n^{\omega,+}(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h^{\omega}(t)\right)s_{n,h}^{\omega}(r)\right)\mathrm{d}t \\ + e^{-\gamma\tau}J\left(\pi^{\varepsilon}; X^{\omega}(\tau(\omega)), Y^{\omega}(\tau(\omega))\right)\Bigg]$$

Then, for any $\pi \in \Pi_{\xi}$, (A.6) and (A.8) imply,

$$V(x,y) \leq J(\pi;x,y) \leq \mathbb{E}_{x,y}^{\pi}\Bigg[\sum_{n=1}^{N}\int_0^{\tau} e^{-\gamma t}\left(C_{0,n}X_n^{\omega,+}(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h^{\omega}(t)\right)s_{n,h}^{\omega}(r)\right)\mathrm{d}t \\ + e^{-\gamma\tau}V\left(X^{\omega}(\tau(\omega)), Y^{\omega}(\tau(\omega))\right) + e^{-\gamma\tau}\varepsilon\Bigg].$$

where $s^{\omega}(t) = \pi\left(X^{\omega}(t), Y^{\omega}(t)\right), t \geq 0$. The arbitrariness of s and $\varepsilon > 0$ give,

$$V(x,y) \leq \mathbb{E}_{x,y}^{\pi}\Bigg[\sum_{n=1}^{N}\int_0^{\tau} e^{-\gamma t}\left(C_{0,n}X_n^{\omega,+}(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h^{\omega}(t)\right)s_{n,h}^{\omega}(r)\right)\mathrm{d}t \\ + e^{-\gamma\tau}V\left(X^{\omega}(\tau(\omega)), Y^{\omega}(\tau(\omega))\right)\Bigg]. \tag{A.9}$$

Combining (A.7) and (A.9) completes the proof. □

A.7 Proof of Lemma 4.0.2

Denoting $v\left(X(t), Y(t)\right)$ simply by v and using Itô Formula (see e.g. Theorem 4.2.1 of Oksendal (2013)), we have,

$$\begin{aligned}\mathrm{d}\left(e^{-\gamma t}v\right) &= e^{-\gamma t}\left(\sum_{n=1}^{N}\partial_{x_n}v\mathrm{d}X_n + \sum_{h=1}^{H}\partial_{q_h}v\mathrm{d}Y_h + \frac{1}{2}\sum_{n=1}^{N}\partial_{x_n}^2 v\sigma_n^2\mathrm{d}t - \gamma v\mathrm{d}t\right)\\ &= e^{-\gamma t}\left(\sum_{n=1}^{N}\sum_{h=1}^{H}\left(\partial_{q_h}vm_n - \partial_{x_n}v\right)s_{n,h}\left(t\right)\mathrm{d}t + \frac{1}{2}\sum_{n=1}^{N}\partial_{x_n}^2 v\sigma_n^2\mathrm{d}t + \sum_{n=1}^{N}\partial_{x_n}v\sigma_n\mathrm{d}B_n(t) - \gamma v\mathrm{d}t\right).\end{aligned}$$

Taking integration from both sides gives,

$$
\begin{aligned}
& e^{-\gamma t_2} v\left(X(t_2), Y(t_2)\right) - e^{-\gamma t_1} v\left(X(t_1), Y(t_1)\right) \\
& = \int_{t_1}^{t_2} \mathrm{d}\left(e^{-\gamma t} v\left(X(t), Y(t)\right)\right) \\
& = \int_{t_1}^{t_2} e^{-\gamma t}\left(\sum_{n=1}^{N}\sum_{h=1}^{H}\left(\partial_{y_h} v\left(X(t), Y(t)\right) m_n - \partial_{x_n} v\left(X(t), Y(t)\right)\right) s_{n,h}(t)\right) \mathrm{d}t \\
& + \int_{t_1}^{t_2} e^{-\gamma t}\left(\frac{1}{2}\sum_{n=1}^{N}\sigma_n^2 \partial_{x_n}^2 v\left(X(t), Y(t)\right) - \gamma v\left(X(t), Y(t)\right)\right) \mathrm{d}t \\
& + \int_{t_1}^{t_2} e^{-\gamma t}\sum_{n=1}^{N}\sigma_n \partial_{x_n} v\left(X(t), Y(t)\right) \mathrm{d}B_n(t).
\end{aligned}
$$

Taking expectation of both sides, the zero-mean property of Itô integrals gives,

$$
\begin{aligned}
& \mathbb{E}\left[e^{-\gamma t_2} v\left(X(t_2), Y(t_2)\right) - e^{-\gamma t_1} v\left(X(t_1), Y(t_1)\right)\right] \\
= & \mathbb{E}\left[\int_{t_1}^{t_2} e^{-\gamma t}\left(\sum_{n=1}^{N}\sum_{h=1}^{H}\left(m_n \partial_{q_h} v\left(X(t), Y(t)\right) - \partial_{x_n} v\left(X(t), Y(t)\right)\right) s_{n,h}(t)\right) \mathrm{d}t\right] \\
+ & \mathbb{E}\left[\int_{t_1}^{t_2} e^{-\gamma t}\left(\frac{1}{2}\sum_{n=1}^{N}\sigma_n^2 \partial_{x_n}^2 v\left(X(t), Y(t)\right) - \gamma v\left(X(t), Y(t)\right)\right) \mathrm{d}t\right].
\end{aligned}
$$

That is the desired result and the proof is complete. □

A.8 Proof of Theorem 4.0.3

Step 1. Fix $(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H$ as the initial state of the system. For an arbitrary $\mathrm{s} \in [0,\xi]^{N\times H}$, consider the constant control $s(r) = \mathrm{s}$. Based on (A.9) and (4.7), for $0 < \tau < 1$ we have,

$$\begin{aligned}
0 \le\, & \mathbb{E}\left[\sum_{n=1}^{N}\int_0^\tau e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)\mathrm{s}_{n,h}\right)\mathrm{d}t\right] \\
& + \mathbb{E}\left[e^{-\gamma\tau}V\left(X(\tau),Y(\tau)\right) - V(x,y)\right] \\
& = \mathbb{E}\left[\sum_{n=1}^{N}\int_0^\tau e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)\mathrm{s}_{n,h}\right)\mathrm{d}t\right] \\
& + \mathbb{E}\left[\int_0^\tau e^{-\gamma t}\left(\sum_{n=1}^{N}\sum_{h=1}^{H}\left(\partial_{q_h}V\left(X(t),Y(t)\right)m_n - \partial_{x_n}V\left(X(t),Y(t)\right)\right)\mathrm{s}_{n,h}\right)\mathrm{d}t\right] \\
& + \mathbb{E}\left[\int_0^\tau e^{-\gamma t}\left(\frac{1}{2}\sum_{n=1}^{N}\partial_{x_n}^2V\left(X(t),Y(t)\right)\sigma_n^2 - \gamma V\left(X(t),Y(t)\right)\right)\mathrm{d}t\right]
\end{aligned} \tag{A.10}$$

We simply denote $V\left(X(t),Y(t)\right)$ by V where there is no ambiguity. Divide both sides by τ. Then, take the limit as $\tau \to 0$ and pass it through the expectation using Dominated Convergence Theorem (see e.g. Theorem 11.32 in Rudin (1964)), relying on the fact that $V \in C^2\left(\mathbb{R}^{N\times H}\right)$. We have,

$$\begin{aligned}
0 \le\, & \mathbb{E}\left[\lim_{\tau\to 0}\frac{1}{\tau}\sum_{n=1}^{N}\int_0^\tau e^{-\gamma t}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_nY_h(t)\right)\mathrm{s}_{n,h}\right)\mathrm{d}t\right] \\
& + \mathbb{E}\left[\lim_{\tau\to 0}\frac{1}{\tau}\int_0^\tau e^{-\gamma t}\left(\sum_{n=1}^{N}\sum_{h=1}^{H}\left(\partial_{y_h}Vm_n - \partial_{x_n}V\right)\mathrm{s}_{n,h} + \frac{1}{2}\sum_{n=1}^{N}\partial_{x_n}^2V\sigma_n^2 - \gamma V\right)\mathrm{d}t\right].
\end{aligned}$$

A straightforward application of Mean Value Theorem (see e.g. Theorem 5.10 in Rudin (1964)) yields,

$$0 \le \sum_{n=1}^{N}\left(C_{0,n}x_n^+ + \sum_{h=1}^{H}\left(C_{n,h} + 2Um_ny_h + \partial_{y_h}Vm_n - \partial_{x_n}V\right)\mathrm{s}_{n,h} + \frac{1}{2}\partial_{x_n}^2V\sigma_n^2\right) - \gamma V.$$

Since s is arbitrary, we have,

$$0 \leq \sum_{n=1}^{N} \left(C_{0,n} x_n^+ + \frac{1}{2}\partial_{x_n}^2 V \sigma_n^2 + \inf_{\mathrm{s}} \left\{ \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n y_h + \partial_{y_h} V m_n - \partial_{x_n} V \right) \mathrm{s}_{n,h} \right\} \right) - \gamma V. \tag{A.11}$$

Step 2. Now, suppose that s^* is the optimal control process of the BMASP; that means,

$$V(x,y) = \mathbb{E}\left[\sum_{n=1}^{N} \int_0^\infty e^{-\gamma t} \left(C_{0,n} X_n^{*,+}(t) + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n Y_h^*(t) \right) s_{n,h}^*(t) \right) \mathrm{d}t \right].$$

where X^* and Y^* are respectively the solutions of (BMASP-2) and (BMASP-3) associated to control process s^*. Using the DPP, for $\delta > 0$ we can write,

$$V(x,y) = \mathbb{E}\left[\sum_{n=1}^{N} \int_0^\delta e^{-\gamma t} \left(C_{0,n} X_n^{*,+}(t) + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n Y_h^*(t) \right) s_{n,h}^*(t) \right) \mathrm{d}t + V(X^*(\delta), Y^*(\delta)) \right].$$

Similar to step 1, it gives,

$$0 = \sum_{n=1}^{N} \left(C_{0,n} x_n^+ + \sum_{h=1}^{H} \left(C_{n,h} + 2U m_n y_h + \partial_{y_h} V m_n - \partial_{x_n} V \right) s_{n,h}^*(t) + \frac{1}{2}\partial_{x_n}^2 V \sigma_n^2 \right) - \gamma V. \tag{A.12}$$

Combining (A.11) and (A.12) in the light of (4.8) gives the HJB. □

A.9 Proof of Theorem 4.0.4

Fix $n \in [N]$ and $h \in [H]$. Since $V \in C^2(\mathbb{R}^{N+H})$, we know that $\delta_{n,h}^V \in C^1(\mathbb{R}^{N+H})$. Fix an arbitrary real number a. We want to show that $\{(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H : \pi_{n,h}^\star(x,y) > a\}$ is a Borel measurable set. Then, by Definition 11.13 of Rudin (1964), $\pi_{n,h}^\star$ is a measurable function. If $a < 0$ or $a \geq \xi$, the measurability of $\{(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H : \pi_{n,h}^\star(x,y) > a\}$ is obvious. In case that $0 \leq a < \xi$, then

$$\{(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H : \pi_{n,h}^\star(x,y) > a\} = \{(x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H : \delta_{n,h}^V \geq 0\}.$$

But, the RHS is measurable since $\delta_{n,h}^V$ is continuous and measurable, see e.g. Example 11.14 and Theorem 11.15 in Rudin (1964). It is straightforward to check that $\pi^\star$ minimizes the linear problem given in (4.9). □

A.10 Proof of Theorem 4.0.5

Fix $T > 0$. For $l \in \mathbb{N}$, define $\theta_l = \inf \left\{ t \geq 0, \int_0^t \left| \sum_{n=1}^N \sigma_n \partial_{x_n} v\left(X(r), Y(r)\right) \right|^2 \right\} \mathrm{d}r$. That means θ_l is a stopping time. For a similar treatment see e.g. page 50 of Pham (2009).

Considering the discontinuities of Y defined in (BMASP-4), we apply Itô's formula for semi-martingales, see Theorem II-33 of Protter (2013), to $e^{-\gamma t} v(X(t), Y(t))$, we get,

$$
\begin{aligned}
& e^{-\gamma(\theta_l \wedge T)} v(X(\theta_l \wedge T), Y(\theta_l \wedge T)) - v(x,y) = \\
& \int_0^{\theta_l \wedge T} \sum_{h=1}^H e^{-\gamma t} \partial_{y_h} v\left(X(t), Y\left(t^-\right)\right) \mathrm{d}Y_h(t) \\
& + \int_0^{\theta_l \wedge T} \sum_{n=1}^N e^{-\gamma t} \partial_{x_n} v\left(X(t), Y\left(t^-\right)\right) \mathrm{d}X_n(t) \\
& + \int_0^{\theta_l \wedge T} \left(\frac{1}{2} \sum_{n=1}^N e^{-\gamma t} \sigma_n^2 \partial_{x_n}^2 v\left(X(t), Y\left(t^-\right)\right) - e^{-\gamma t} \gamma v\left(X(t), Y\left(t^-\right)\right) \right) \mathrm{d}t \\
& + \sum_{0 \leq t < (\theta_l \wedge T)} e^{-\gamma t} \left[v\left(X(t), Y(t)\right) - v\left(X(t), Y\left(t^-\right)\right) \right] \\
& - \sum_{0 \leq t < (\theta_l \wedge T)} e^{-\gamma t} \left[\sum_{h=1}^H \partial_{y_h} v\left(X(t), Y\left(t^-\right)\right) \Delta Y_h(t) \right].
\end{aligned}
$$

Recall the dynamics of the state processes (BMASP-2) and (BMASP-3), by the zero-mean property of the stopped integrals and (4.13) we get

$$
\begin{aligned}
v(x,y) = {} & \mathbb{E}\left[e^{-\gamma(\theta_l \wedge T)} v\left(X(\theta_l \wedge T), Y(\theta_l \wedge T)\right) \right] \\
& + \mathbb{E}\left[\sum_{n=1}^N \sum_{h=1}^H \int_0^{\theta_l \wedge T} e^{-\gamma t} \left(\partial_{x_n} v\left(X(t), Y\left(t^-\right)\right) - m_n \partial_{y_h} v\left(X(t), Y\left(t^-\right)\right) \right) s_{n,h}(t) \, \mathrm{d}t \right] \\
& + \mathbb{E}\left[\int_0^{\theta_l \wedge T} e^{-\gamma t} \left(\gamma v\left(X(t), Y\left(t^-\right)\right) - \frac{1}{2} \sum_{n=1}^N \sigma_n^2 \partial_{x_n}^2 v\left(X(t), Q\left(t^-\right)\right) \right) \mathrm{d}t \right].
\end{aligned}
\tag{A.13}
$$

By (4.12) we can write,

$$
\sum_{n=1}^N \left(C_{0,n} x_n^+ + \frac{1}{2} \partial_{x_n}^2 v(x,y) \sigma_n^2 - \sum_{h=1}^H \delta_{n,h}^v(x,y) \mathrm{s}_{n,h} \right) \geq \gamma v(x,y), \quad (x,y) \in \mathbb{R}^N \times \mathbb{R}_+^H, \ \mathrm{s} \in [0,\xi]^{N \times H}.
$$

It implies,

$$\sum_{n=1}^{N}\left(C_{0,n}x_n^+ + \sum_{h=1}^{H}\left(C_{n,h}+2Um_n y_h\right)\mathrm{s}_{n,h}\right) \geq$$
$$\gamma v(x,y) + \sum_{n=1}^{N}\left(\sum_{h=1}^{H}\left(\partial_{x_n}v(x,y) - m_n\partial_{y_h}v(x,y)\right)\mathrm{s}_{n,h} - \frac{1}{2}\partial^2_{x_n}v(x,y)\sigma_n^2\right).$$

Plugging it in (A.13), we get

$$\begin{aligned} v(x,y) \leq\ & \mathbb{E}\left[e^{-\gamma(\theta_l\wedge T)}v\left(X\left(\theta_l\wedge T\right), Y\left(\theta_l\wedge T\right)\right)\right] \\ & + \mathbb{E}\left[\int_0^{\theta_l\wedge T} e^{-\gamma t}\left(\sum_{n=1}^{N}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h}+2Um_nY_h(t)\right)s_{n,h}\left(t\right)\right)\right)\mathrm{d}t\right]. \end{aligned}$$

We may take the limit as $l \to \infty$ and using (4.11) and the fact that $\mathrm{s} \in [0,\xi]^{N\times H}$, we apply Dominated Convergence Theorem. It gives,

$$\begin{aligned} v(x,y) \leq\ & \mathbb{E}\left[e^{-\gamma T}v\left(X(T), Y(T)\right)\right] \\ & + \mathbb{E}\left[\int_0^{T} e^{-\gamma t}\left(\sum_{n=1}^{N}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h}+2Um_nY_h(t)\right)s_{n,h}\left(t\right)\right)\right)\mathrm{d}t\right]. \end{aligned}$$

Sending T to infinity, according to (4.14) we get

$$v(x,y) \leq \mathbb{E}\left[\int_0^{\infty} e^{-\gamma t}\left(\sum_{n=1}^{N}\left(C_{0,n}X_n^+(t) + \sum_{h=1}^{H}\left(C_{n,h}+2Um_nY_h(t)\right)s_{n,h}\left(t\right)\right)\right)\mathrm{d}t\right].$$

Considering (4.5) and (4.6) imply $v(x,y) \leq V(x,y)$ for all $(x,y) \in \mathbb{R}^N \times \mathbb{R}^H_+$.

Now, by (4.10) and (4.12) we get for all $(x,y) \in \mathbb{R}^N \times \mathbb{R}^H_+$,

$$\sum_{n=1}^{N}\left(C_{0,n}x_n^+ + \frac{1}{2}\partial^2_{x_n}v(x,y)\sigma_n^2 - \sum_{h=1}^{H}\delta^v_{n,h}(x,y)\xi\mathbf{1}_{\delta^v_{n,h}(x,y)\geq 0}\right) = \gamma v(x,y).$$

It gives,

$$\sum_{n=1}^{N}\left(C_{0,n}x_n^+ + \sum_{h=1}^{H}\left(C_{n,h}+2Um_n y_h\right)\xi\mathbf{1}_{\delta^v_{n,h}(x,y)}\right) =$$
$$\gamma v(x,y) + \sum_{n=1}^{N}\left(\sum_{h=1}^{H}\left(\partial_{x_n}v(x,y) - m_n\partial_{y_h}v(x,y)\right)\xi\mathbf{1}_{\delta^v_{n,h}(x,y)} - \frac{1}{2}\partial^2_{x_n}v(x,y)\sigma_n^2\right).$$

By (A.13), we have,

$$
\begin{aligned}
v(x_0, y_0) &= \mathbb{E}\left[e^{-\gamma(\theta_l \wedge T)} v\left(X^*(\theta_l \wedge T), Y^*(\theta_l \wedge T)\right)\right] \\
&\quad + \mathbb{E}\left[\int_0^{\theta_l \wedge T} e^{-\gamma t}\left(\sum_{n=1}^{N}\left(C_{0,n} X_n^{*,+}(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2U m_n Y_h^*(t)\right) \xi \mathbf{1}_{\delta_{n,h}^v(x,y)}\right)\right) \mathrm{d}t\right].
\end{aligned}
$$

Similar to above, by taking limit and according to Dominated Convergence Theorem as well as (4.15), we can write

$$
v(x_0, y_0) = J(\pi^\star; x_0, y_0) = \mathbb{E}\left[\int_0^{\infty} e^{-\gamma t}\left(\sum_{n=1}^{N}\left(C_{0,n} X_n^{*,+}(t) + \sum_{h=1}^{H}\left(C_{n,h} + 2U m_n Y_h^*(t)\right) s_{n,h}(t)\right)\right) \mathrm{d}t\right].
$$

Therefore, $v(x_0, y_0) = V(x_0, y_0)$. $\square$