OPTIMAL DISCHARGE CONTROL IN CAPACITY-CONSTRAINED SYSTEMS: A DYNAMIC PROGRAMMING APPROACH

by

Xin Yuan Zhang

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
Graduate Department of Mechanical and Industrial Engineering
University of Toronto

© Copyright 2016 by Xin Yuan Zhang
Abstract

Optimal Discharge Control in Capacity-Constrained Systems: A Dynamic Programming Approach

Xin Yuan Zhang
Master of Applied Science
Graduate Department of Mechanical and Industrial Engineering
University of Toronto
2016

In this thesis, we study a stochastic discharge control problem with patient readmissions in a capacity-constrained hospital unit. The problem is formulated as a multi-stage dynamic program where the objective is to determine the optimal discharge policy so as to minimize the expected total number of patients in the system until it is emptied. In the one-bed unit case, we fully characterize the optimal policy as a two-threshold policy (TTP). In the multi-bed unit case, the problem is more challenging since the associated dynamic programming equations are not fully decomposable bed-by-bed. To this end, we propose a so-called Marginal Policy Iteration (MPI) algorithm, that decomposes the joint optimization problem into easier one-bed marginal optimization problems and establish convergence of this algorithm to the optimal policy.
Acknowledgements

First, I would like to thank my supervisor, Dr. Michael Jong Kim for his guidance and support throughout my master study at U of T. I would also thank Dr. Hossein Abouee-Mehrizi and Ya-Tang Chuang with whom co-working on this project has been a privilege. I must also thank my the members of my examination committee, Dr. Dragan Banjevic and Dr. Jonathan Yu-Meng Li for their advice and help on the thesis.
## Contents

1 Introduction \( \quad \) \hfill 1
   1.1 Motivation \hfill 1
   1.2 Related Literature \hfill 2
   1.3 Contributions \hfill 4
   1.4 Road-map \hfill 4

2 Problem Formulation \hfill 5
   2.1 Model and Preliminaries \hfill 5
      2.1.1 States and Actions \hfill 5
      2.1.2 Discharge Policies and Transition Rules \hfill 7
   2.2 The Optimality Criteria \hfill 10
      2.2.1 The Value Function \hfill 10
      2.2.2 The Dynamic Programming Equations \hfill 11

3 Discharge Control in the One-Bed Setting \hfill 13
   3.1 Model \hfill 13
      3.1.1 The Myopic Policy \hfill 15
   3.2 DP Analysis \hfill 15
      3.2.1 Preliminary Results \hfill 15
      3.2.2 The Optimal Threshold Policies \hfill 20
   3.3 Numerical Example \hfill 23
3.3.1 OTP Problem ................................................. 23
3.3.2 TTP Problem ................................................. 25
3.3.3 Comparison Between Discharge Policies ....... 26
3.4 Summary ............................................................ 28

4 Discharge Control in the Multiple Bed Setting 29
4.1 Model .............................................................. 29
4.2 DP Analysis ....................................................... 30
  4.2.1 Preliminary Results ........................................ 30
  4.2.2 The Optimal Discharge Policies .................... 33
4.3 Marginal Policy Iteration (MPI) Algorithm ........ 37
  4.3.1 The MPI Algorithm ...................................... 37
  4.3.2 Optimality of the Limit Policy ....................... 40
4.4 Numerical Example ............................................. 42
4.5 Summary ............................................................ 44

5 Conclusion .......................................................... 45
5.1 Future Work ..................................................... 45

Bibliography .......................................................... 47
Chapter 1

Introduction

1.1 Motivation

In a capacity-constrained hospital unit such as the intensive care unit (ICU), clinicians may need to decide to discharge patients currently in bed upon the admission of new patients with more urgent conditions. In practice, patients in bed can be discharged even if they are not in a sufficiently healthy state, and this may result in the possibility of readmissions. A recent study in Canada estimated that 9% of patients discharged from acute care returned to the hospital within 30 days of their initial discharge [10]. In fact, unplanned returns to hospitals are a contributing factor to overcrowding and, as a result, long patient wait times. Moreover, urgent returns to hospitals are very costly to the healthcare system. For example, inpatient readmission within 30 days of discharge to hospitals costs more than $1.8 billion a year in Canada (excluding physician fees for the services) which accounts for more than 11% of the total inpatient costs [10]. In the U.S., approximately 20% patients that are discharged from a hospital are readmitted within 30 days. The cost of preventable readmissions to hospitals in the U.S. is about $25 billion annually. However, despite the significant impact of readmissions arising from discharge decisions, most existing works do not include this feature.

Motivated by these concerns, we aim to provide a mathematical model that explicitly describes the dynamics of patients being readmitted to the hospital, and from which
clean structural properties of the optimal discharge policy can be established.

1.2 Related Literature

Most relevant works to discharge decisions with readmissions can be found in the health-care operations/management literature. In Chan et al. [6], they examine the effects of ICU discharge strategies under uncertainty on the patient mortality and readmission. Specifically, they consider a hospital with several classes of patients and different expected length of stays. They model the system as a finite-horizon discrete-time dynamic program in which the length of stay is memoryless. They assume that if a patient is discharged prior to completing treatment due to the arrival of a more acute patient a cost is incurred; thus, they do not explicitly model patients being readmitted to hospital. They analyze a family of simple index policies and show that an index policy is optimal in certain regimes. In addition, Hosseinifard et al. [16] study ICU discharging decisions when a new critically ill patient arrives to the ICU and there is no bed available. Assuming that the length of stay is not memoryless, they formulate the problem as a stochastic program and evaluate various scenarios for discharging a patient from the ICU using an optimization-simulation method. KC and Terwiesch [17] empirically investigate the relationship between the occupancy level of an ICU and the length of stay. They observe that when the occupancy level is high, the chance that a patient is discharged earlier is higher. Also, they find that as the length of stay decreases, the rate of readmission increases. Other relevant contributions on readmission analysis in hospital units are the works of Armony et al [1], Helm et al [14] and Zhang et al [28].

Congestion in systems with rejoining customers also arise in queuing models. The most relevant work is by Chan et al [7], who consider a single class queuing network with two stations where customers who leave the first station may join the second station and return to the first station after a random amount of time. Using a deterministic fluid model, they analyze the system dynamics and the equilibrium of the system without
investigating the control of queues. In addition, Yom-Tov and Mandelbaum [27] study a single server Markovian system with time-varying arrival rate where customers may return to the system. They analyze the system using both fluid and diffusion processes, and propose a time-varying square-root staffing policy for such a system. In fact, dynamic control of queues are generally challenging even without the rejoining aspects. We refer the reader to the works of Ata [2], George and Harrison [13] and Van Mieghem [26], which only consider simple controls such as admission decisions or varying the service/arrival rate of the system.

Other streams of research that are relevant to the discharge control problem are perishable inventory management and maintenance optimization. Most recent contributions include the works of Chen et al [8], [9], Li et al [24], Li and Yu [23], Kim [19] and Kim and Makis [22]. Although the type of controls and objectives that those works consider are different than the ones in our work, they all attempt to solve control problems in capacity constrained systems in which the underlying state process follows a Markov degradation process.

As we formulate the control problem as a Markov decision process (MDP), our analysis uses standard results from discrete-time stochastic control. Some classical techniques can be found in [4], [5] and [15]. In the multi-bed setting, our system consists of weakly coupled one-bed systems so that the control problem cannot be fully decomposable into independent subproblems. Bertsekas [2] studies dynamic programming problems involving coupled subsystems. An approximate DP approach is developed to approximate the cost-to-go in such problems. However, the weak coupling they consider is only the constraints on the action sets. Other methods that attempt to approximately solve MDPs with coupled subproblems only apply to some special cases of weak coupling ( [11], [12], [25], [20], [21])
1.3 Contributions

The main contribution of this thesis is providing a stylized model that captures the essential dynamics of discharge control with readmissions and within which the analysis of the optimal polices still remains tractable. In addition, we are able to demonstrate that the optimal discharge policy in the one-bed setting is a two threshold policy that has practical interpretation: we only discharge patients in the sufficiently good or bad health condition states to avoid system congestion. Furthermore, we show that the problem in the multi-bed setting is a joint optimization problem that cannot be decomposed to independent one-bed problems. We connect the problem to MDP analysis with weakly coupled subsystems and propose a new policy iteration approach to solve the corresponding dynamic programming equations. The result leads to practical solutions to discharge decisions with readmissions that can be validated numerically.

1.4 Road-map

The remainder of the thesis is organized as follows. We introduce in Chapter 2 the basic model formulation and other preliminaries that are required for our analysis. In particular, we first formulate the problem so that the objective is clearly demonstrated with the associated dynamic programming equations. Chapter 3 includes the main results for the one-bed setting. The optimal discharge policy is fully characterized, and numerical examples are provided. In Chapter 4, we consider the multi-bed system, which cannot be fully decomposed to isolated one-bed systems. The problem becomes challenging both analytically and computationally. We introduce a new policy iteration algorithm that decomposes the computations into tractable subproblems related to the one-bed setting. We then show the optimality of the limiting policy generated by this algorithm. The results are presented with numerical examples. Concluding remarks are provided in Chapter 5.
Chapter 2

Problem Formulation

2.1 Model and Preliminaries

Consider the scenario where patients arrive to a hospital unit (e.g. intensive care unit) with $K$ beds at periods $t = 1, 2, \ldots$. We denote the number of new arrivals in period $t$ by $A_t$, and model $(A_t : t \geq 1)$ as a sequence of independent and identically distributed random variables taking values in $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$. We assume that at most one new patient arrives at the unit in period $t$, that is, $A_t \in \{0, 1\}$ with arrival probability $\lambda$.

As patients arrive, they wait in a queue until a bed in the unit becomes available for use. Whenever the bed becomes available, the patient at the front of the queue moves “into the unit”, occupies the bed, and treatment of that patient begins.

2.1.1 States and Actions

We model the health state of patients as follows. Patients waiting in the queue can be in one of $L$ states $\mathcal{X} = \{1, \ldots, L\}$, where state $L$ represents the worst possible health condition and state 1 represents a relatively good condition. Similarly, patients occupying a bed in the unit can be in one of $L + 1$ state $\mathcal{X}_0 = \{0\} \cup \mathcal{X}$ where state 0 represents a “perfect health” condition so no treatment is required. At period $t$, we denote the state of a patient occupying the $k$-th bed by $X_{ik} \in \mathcal{X}_0$, and as the patient receives treatment,
his health state evolves according to a time-homogeneous Markov chains \((X^k_t : t \geq 0)\). We denote the health state of all the patients in the unit by a random vector \(X_t = (X^k_t)_{k \in \{1, \ldots, K\}} \in \mathcal{X}_0^K\).

We assume population homogeneity and that the health dynamics of different patients are independent of one another. Accordingly, for all \(k\), \(X^k_t\) has the identical transition probability matrix \(P = (p_{x,y})\) and initial distribution \(\beta(x) = \mathbb{P}(X^k_0 = x), x \in \mathcal{X}\), that is, we assume that patients do not arrive to the unit in a “perfect health” condition.

**Assumption 2.1.1.** We assume the following properties for the transition matrix \(P\):

1. the treatment does not worsen a patient’s health, that is, \(p_{x,y} = 0\) for all \(y > x\),
2. patients recover in a “smooth fashion”, that is, \(p_{x,y} \geq p_{x+1,y}\) for all \(y \leq x\),
3. we have boundary conditions \(p_{0,0} = 1\) and \(p_{L,0} > 0\).

The above properties are modest assumptions on the evolution of \((X^k_t)\) which models the recovery process of a patient under treatment.

When a patient occupying the \(k\)-th bed reaches a “perfect health” condition at period \(t\), that is, \(X^k_t = 0\), the patient immediately leaves the unit and the bed is vacated. The next patient at the front of the queue enters the unit (and therefore occupies the \(k\)-th bed) with distribution \(\beta\), and does so instantaneously. Thus, having \(X^k_t = 0\) is equivalent to having the \(k\)-bed unoccupied at time \(t\) and no one is in the queue, otherwise it should immediately replaced by a new patient.

Suppose the hospital unit (decision maker) has the ability to discharge a patient from the unit before they reach a “perfect health” condition. Once the patient has been discharged, they vacate the bed, and the patient waiting at the front of the queue enters the unit. For simplicity, we assume that discharging a patient takes exactly one period to complete, so that the health state of the new patient in the unit at the beginning of the next period has distribution \(\beta\).

If the decision maker discharges a patient before they reach health state 0, they do so at the risk of possible readmission. In particular, for all \(k\), if the patient occupying the
Chapter 2. Problem Formulation

The $k$-th bed in health state $X_t^k \in \mathcal{X}$ is discharged at an arbitrary period $t$, at the beginning of the next period $t+1$, that patient may return back to the bed, in the worst state $L$, with probability $q(X_t^k) \in [0, 1]$.

**Remark 2.1.2.** Returning patients generally incur additional loads to the system. Instead of having an artificial cost, we assume the readmission penalty is imposed through transitioning to a worse system state. In our stylized model, we let the readmitted patient to be in the worst state $L$. Although this assumption seems to be restrictive in practical settings, it leads to simpler analysis and clear structure results for our problem. One can always weaken it in an extended model.

Similar to the recovery process, we require some regularity on the readmission probabilities.

**Assumption 2.1.3.** We assume that mapping $x \mapsto q(x)$ is increasing and concave in $x \in \mathcal{X}$, with boundary condition $q(0) = 0$.

At beginning of each period $t$, let $N_t$ represent the current system size, that is the total number of patients that are waiting in the queue or occupying a bed in the unit. Then, the system state at any period $t$ is given by the pair $(N_t, X_t)$. We note that a system state of $(N_t, X_t) = (0, 0)$ denotes an empty system. Let $K^*$ denote the number of occupied beds in which a patient with state $X_t \in \mathcal{X}$ receive treatment, that is, given $N_t \in \mathbb{N}$, $K^*(N_t) = \min\{N_t, K\}$. Based the assumption that a patient immediately leaves the unit when he reaches state 0, we have $X_t^k = 0$ if and only if $K^*(N_t) < K$.

### 2.1.2 Discharge Policies and Transition Rules

A *discharge policy* $\pi = (\pi^k)_{k \in \{1, ..., K\}}$ is a $K$-dimensional vector function whose components are bed-centralized/marginal discharge policies $\pi^k$. For any $k \in \{1, ..., K\}$, the marginal policy $\pi^k$ is a mapping $\pi^k : N_0 \times X_0^K \rightarrow \{0, 1\}$, where the patient in the $k$-th bed is discharged at period $t$ if $\pi^k(N_t, X_t) = 1$, and not discharged if $\pi^k(N_t, X_t) = 0$. Recall that $X_t^k = 0$ if and only if the $k$-th bed becomes empty and the state transition
Chapter 2. Problem Formulation

no longer depends on the discharge action. For simplicity, we assume that we always take a discharge action in this situation. Thus, we consider $\Pi$ the class of all admissible discharge policies $\pi = (\pi^k)$ such that for any $k$, $\pi^k(N_t, X_t) = 1$ if $X^k_t = 0$.

For short, we will write $x^{k-} = (x^1, \ldots, x^{k-1}, x^{k+1}, \ldots, x^K)$, so that any state $x \in \mathcal{X}^K$ can be written as $x = (x^k; x^{k-})$ for some $k$. Thus, we write the discharge policy as $\pi^k(x^k; x^{k-})$ to emphasize it is applied to the $k$-th bed with centralized state $x^k$, considering other beds’ state $x^{k-}$ as fixed parameters.

**Remark 2.1.4.** The beds we consider in this problem are indistinguishable. In other words, given any policy $\pi$, the $k$-dependence does not affect the transition dynamics so that the state vector $X_t$ is invariant under a permutation of $k$. Therefore, a marginal discharge policy $\pi^k$ actually applies to the $x^k$-valued bed rather than the $k$-th bed.

Under any admissible discharge policy $\pi \in \Pi$, if we let $D_t = \pi(N_t, X_t) \in \{0, 1\}^K$ denote the discharge decision at period $t$, then the health state $X_t$ and the queue length $N_t$ are governed by the following dynamics.

For transition in $N_t$, we have, for all $t \geq 0$ and $N_t \in \mathbb{N}_0$,

$$N_{t+1} = N_t + A_{t+1} - \sum_{k=1}^{K^*(N_t)} D_t^k S^k_{t+1} - \sum_{k=1}^{K^*(N_t)} (1 - D_t^k) W^k_{t+1},$$

(2.1.1)

where $S^k_{t+1} = 1$ if the patient in bed $k$ does not return back to the queue given he/she was discharged, that is, $D^k_t = 1$, and $S^k_{t+1} = 0$ otherwise, and $W^k_{t+1} = 1$ if the patient is in health state $X^k_t \in \mathcal{X}$ and fully recovers by the next period to a “good-as-new” state given he/she was not discharged, that is, $D^k_t = 0$, and $W^k_{t+1} = 0$ otherwise. We therefore have $\mathbb{P}(S^k_{t+1} = 1) = 1 - q(X^k_t)$, and $\mathbb{P}(W^k_{t+1} = 1) = p_{X^k_t, 0}$.

The transition in $X_t$ is slightly more complicated. In the boundary cases, when $N_t < 2K$, we can have empty beds in the next period $t + 1$, that is, $N_{t+1} < K$. If the patient
on the $k$-th bed leaves in period $t$, we let $O_{t+1}^k = 1$ if a new patient will be assigned to that bed and $O_{t+1}^k = 0$ otherwise. We also note that any order of assignment will result in the same joint process $(N_s, X_s)_{s \geq t}$. Thus, we simply assume that patients are assigned to beds in the order of index number $k$.

Therefore, for any $k$, if $O_{t+1}^k = 1$ such that $X_{t+1}^k = x \in \mathcal{X}$ we have the transition in $X_t^k$ given by

$$
P^\pi(X_{t+1}^k = x \mid N_t, X_t^k, D_t^k) = \begin{cases} 
\beta(x) & \text{if } X_t^k = 0 \\
D_t^k \left[ (1 - q(X_t^k)) \beta(x) + \mathbf{1}_{\{x=L\}} q(X_t^k) \right] \\
+ (1 - D_t^k)(p_{X_t^k,x} + p_{X_t^k,0}) \beta(x) & \text{if } X_t^k \in \mathcal{X}.
\end{cases}
$$

(2.1.2)

On the other hand, if $O_{t+1}^k = 0$ for some bed $k$, then we have

$$
P^\pi(X_{t+1}^k = 0 \mid N_t, X_t^k, D_t^k) = \begin{cases} 
1 & \text{if } X_t^k = 0, \\
D_t^k(1 - q(X_t^k)) + (1 - D_t^k)p_{X_t^k,0} & \text{if } X_t^k \in \mathcal{X},
\end{cases}
$$

(2.1.3)

and for all $x \in \mathcal{X}$

$$
P(X_{t+1}^k = x \mid N_t, X_t^k, D_t^k) = D_t^k \mathbf{1}_{\{x=L\}} q(X_t^k) + (1 - D_t^k)p_{X_t^k,x}
$$

(2.1.4)

When there are sufficient patients to be assigned in the next period $t+1$, that is, $N_t > 2K$ such that $N_{t+1} \geq K$, we have $O_{t+1}^k = 1$ for all $k$.

**Remark 2.1.5.** The assignment actions $O_{t+1}^k$ are not true decision variables. They are random variables that only depend on $N_{t+1}$, whose transition is fully characterized in 2.1.1. They only affect the transition dynamics in the boundary case when the number of patients in the system is sufficiently low. For this reason, we also observe that the discharge decision will depend on both $X_t$ and $N_t$. 
2.2 The Optimality Criteria

Given a policy \( \pi \in \Pi \), we denote by \( T_0(\pi) \) the stopping time when the system reaches the empty state, that is,

\[
T_0(\pi) = \inf\{ t \geq 0 : (N_t, X_t) = (0, 0) \}. \tag{2.2.1}
\]

**Assumption 2.2.1.** In order to be able to vacate the system, we further assume the system size can be eventually lowered to zero using any admissible discharge policy, that is, \( T_0(\pi) < \infty \) a.s..

Therefore, for any discharge policy \( \pi \in \Pi \), the system state process \((N_t, X_t)_{t \geq 0}\) is a renewal process with the cycle length characterized by the stopping time \( T_0(\pi) \).

### 2.2.1 The Value Function

The objective of the decision maker is to find a discharge policy \( \pi \in \Pi \), if it exists, that minimizes the expected total system size until it is emptied, that is, over the renewal cycle. In particular, for all \( n \in \mathbb{N}_0 \) and \( x = [x^k] \in \mathcal{X}^K \), we consider the following stochastic control problem with the value function given by

\[
V(n, x) = \inf_{\pi \in \Pi} \mathbb{E}^\pi \left[ \sum_{t=0}^{T_0} N_t \bigg| N_0 = n, X_0 = x \right], \tag{2.2.2}
\]

subject to the dynamics of \((N_t, X_t)^\pi\) defined in (2.1.1) - (2.1.4).

**Remark 2.2.2.** The minimization problem defined in (2.2.2) is a well-proposed objective as long as there are non-trivial discharge policies such that the effective leaving rate \( \mu_\pi \) of the system is larger than the arrival rate \( \lambda \). While it is difficult to determine the explicit form of \( \mu_\pi \), one can always consider a variant of our model in which the above objective is well-defined. For example, if our system has a finite buffer size \( B \), that is, \( \lambda = 0 \) when \( N_t = B \), the optimal value \( V_B \) is bounded above by \( B \cdot \mathbb{E}^\pi_B[T_0] \). For the boundedness of \( T_0 \) in \( L^1 \), we assume \( \lambda < 1 \) and \( \max\{p_{x,0}, 1 - q(x)\} > 0 \) for all \( x \in \mathcal{X} \). Then, under any
admissible policy \(\pi\), there exists a sufficiently large \(T^{\pi}\) such that \(\mathbb{P}_B(T_0 \leq T^{\pi}) = p_0 > 0\). It follows that
\[
E_B^\pi[T_0] \leq \sum_{i=1}^{\infty} i \cdot T^{\pi} \cdot p_0^i (1 - p_0)^{i-1} = \frac{T^{\pi}}{p_0},
\]
which implies the optimal value \(V_B \leq B \cdot \frac{T^{\pi}}{p_0}\) so that the optimization problem is well-posed with a finite buffer system. One can easily verify that this differs from the original problem only in the boundary values, so the same structural results must still hold.

### 2.2.2 The Dynamic Programming Equations

Using standard dynamic programming arguments, it can be shown that the value function \(V(n, x)\) defined in (2.2.2) uniquely satisfies the dynamic programming equations given by
\[
V(0, 0) = 0 \quad (2.2.3)
\]
and for all \(n \in \mathbb{N}, x \in \mathcal{X}^K\),
\[
V(n, x) = n + \min_{d \in \{0, 1\}^K} \{E[V(N_1, X_1) | N_0 = n, X_0 = x, D_0 = d]\} \quad (2.2.4)
\]
Furthermore, the value function can be obtained as the limit
\[
V(n, x) = \lim_{s \to \infty} V_s(n, x), \quad (2.2.5)
\]
where \(V_s(n, x)\) is the unique solution of the \(s\)-stage dynamic programming equations
\[
V_s(n, x) = n + \min_{d \in \{0, 1\}^K} \{E[V_{\sigma-1}(N_1, X_1) | N_0 = n, X_0 = x, D_0 = d]\}, \quad \sigma = 1, \ldots, s,
\]
\[
V_0(n, x) = 0. \quad (2.2.6)
\]
The stochastic optimization problem defined in (2.2.2) requires solving the dynamic programming equations defined in (2.2.4), which is computationally overwhelming due to its high dimensional state space and complicated transition dynamics. Furthermore, as the discharge decision of each bed are coupled due to the shared queue, the multi-bed optimization problem cannot be decomposed into \( K \) independent one-bed control problems. This makes the structure of the value function and discharge policy difficult to analyze.

To overcome these difficulties, in the next section, we first study the discharge control problem when \( K = 1 \), that is, optimization in a one-bed unit. We fully characterize the structure of the optimal discharge policy through an analysis of the dynamic programming equations in this case. With the established theoretical results and techniques, we then investigate the multi-bed problem in Chapter 4.
Chapter 3

Discharge Control in the One-Bed Setting

3.1 Model

We now study the discharge control problem in a one-bed hospital unit. Recall that at most one new patient arrives to the unit in period $t$, that is, the number of new arrivals $A_t$ is a binary random variable taking values in $\{0, 1\}$ with arrival probability $\Pr(A_t = 1) = \lambda$.

As patients arrive, they wait in a queue until the bed in the unit becomes available for use. The health state of the patient on the bed evolves according to a Markov chain $(X_t, t \geq 0)$ with parameters $(\mathcal{X}, \mathbf{P}, \beta)$ defined in the previous section. Similar to the general case, discharge dynamics with readmission probability $q(x)$ and policy $\pi$ are defined. Accordingly, for the remainder of this section, we will use the same notations as the previous section and simply suppress the dependence on $k$. Note that all the general assumptions still apply to the one-bed case.

For all $n \in \mathbb{N}_0$ and $x \in \mathcal{X}$, we define the same minimization problem with value
function given by

$$V(n, x) = \inf_{\pi \in \Pi} \mathbb{E}^\pi \left[ \sum_{t=0}^{T_0} N_t \mid N_0 = n, X_0 = x \right], \quad (3.1.1)$$

where $T_0(\pi) \triangleq \inf\{t \geq 0 : (N_t, X_t) = (0, 0)\}$.

In the one-bed case, we have the corresponding dynamic programming equations given by

$$V(0, 0) = 0, \quad (3.1.2)$$

and for all $n \in \mathbb{N}$ and $x \in \mathcal{X}$,

$$V(n, x) = n + \min_{d \in \{0, 1\}} \{ \mathbb{E}[V(N_1, X_1) \mid N_0 = n, X_0 = x, D_0 = d] \}$$

$$= n + \min \left\{ q(x) \mathbb{E}_{\lambda}[V(n + A_1, L)] + (1 - q(x)) \mathbb{E}_{\lambda}[V(n - 1 + A_1, \beta)], \right. \left. \mathbb{E}_{\lambda}[V(n + A_1, \tilde{p}_x)] + p_{x, 0} \mathbb{E}_{\lambda}[V(n - 1 + A_1, \beta)] \right\},$$

where $V(n, \beta) = \mathbb{E}_{\beta}[V(n, x)]$ and $V(n, \tilde{p}_x) = \sum_{y=1}^{x} p_{x,y} V(n, y)$. In addition, the value function can be obtained as the limit of the $s$-stage value functions

$$V(n, x) = \lim_{s \to \infty} V_s(n, x),$$

where $V_s(n, x)$ satisfies the following dynamic programming equations

$$V_\sigma(n, x) = n + \min_{d=0, 1} \{ \mathbb{E}[V_{\sigma-1}(N_1, X_1) \mid N_0 = n, X_0 = x, D_0 = d] \}, \quad \sigma = 1, \ldots, s,$$

$$V_0(n, x) = 0. \quad (3.1.4)$$

Before we continue, we make a few additional assumptions. For any real-valued function $g : \mathcal{X}_0 \to \mathbb{R}$ we define its linear interpolation $\bar{g} : [0, N] \to \mathbb{R}$ in the obvious way: for each $x \in \mathcal{X}_0$, $\bar{g}(x) = g(x)$, and for each $z \in (x, x + 1)$, $x \in \mathcal{X}$, $\bar{g}(y)$ is linearly interpolated between $g(x)$ and $g(x + 1)$. A function $g$ defined over $\mathcal{X}_0$ is said to be concave (convex) if its associated linear interpolation $\bar{g}$ is concave (convex).
3.1.1 The Myopic Policy

A myopic discharge policy can be easily derived using the 2-stage dynamic programming equations. For a two-stage problem, that is, we stop the decision process after two periods, it follows from the equations defined in (3.1.4) that for any $n \in \mathbb{N}$ and $x \in \mathcal{X}$

$$V_2(n, x) = \lambda + 2n - \max\{1 - q(x), p_{x,0}\}. \quad (3.1.5)$$

Therefore, in a two-stage problem, the optimal control is fully characterized by the functions $q(x)$ and $p_{x,0}$. This motivates us to construct a discharge policy for the general problem: Given any $n \in \mathbb{N}$ and $x \in \mathcal{X}$, the myopic policy (MYP) is to discharge a patient if

$$H(x) = 1 - p_{x,0} - q(x) \geq 0. \quad (3.1.6)$$

In other words, we make our discharge decisions by comparing only the probabilities of not leaving the unit within one period. Although the MYP is an easily implementable discharge policy, it obviously ignores the potential benefits from the treatment over multi-periods, which is one of the most important features that makes our problem non-trivial. We therefore want to investigate the trade-off between discharging and non-discharging over multi-periods using the dynamic programming equations.

3.2 DP Analysis

3.2.1 Preliminary Results

We begin our analysis with the following structural result for $V(n, x)$ as a function of $n$.

**Proposition 3.2.1.** For any fixed $x \in \mathcal{X}$, $V(n, x)$ is nondecreasing in $n$ for all $n \in \mathbb{N}_0$.

**Proof.** We consider the $s$-stage dynamic programming equations and then proceed by way of mathematical induction. The induction hypothesis is trivial for $s = 1$, and is easy
to check for $s = 2$; indeed, for any $x \in \mathcal{X}$ and $n \in \mathbb{N}$,

$$V_2(n, x) = \lambda + 2n - \max \{1 - q(x), p_{x,0}\},$$

$V_2(n, x)$ is nondecreasing in $n$. In other words, for all $n \in \mathbb{N}_0$ and $x \in \mathcal{X}$

$$\Delta V_2(n, x) = V_2(n + 1, x) - V_2(n, x) \geq 0.$$

Assume now that the induction hypothesis holds for some $s \geq 0$, that is, $V_s(n, x)$ is nondecreasing in $n$ for all $n$.

It follows from equation (3.1.4) that for any $n \in \mathbb{N}$,

$$V_{s+1}(n, x) = n + \min_{d=0,1} \{\mathbb{E}[V_s(N_1, X_1) \mid n, x, d]\}, \tag{3.2.1}$$

where $\mathbb{E}[V_s(N_1, X_1) \mid n, x, d]$ is nondecreasing in $n$ by the induction assumption. The minimum of nondecreasing functions is still nondecreasing. Combining this with the boundary value $V_{s+1}(0, 0) = 0$, it follows that $V_{s+1}(n, x)$ is a nondecreasing function in $n$ for any fixed $x \in \mathcal{X}$. By induction, the nondecreasing property holds for any $s \geq 1$. Therefore, $V(n, x) = \lim_{s \to \infty} V_s(n, x)$ is nondecreasing in $n$ for any fixed $x$, which completes the proof. \hfill $\square$

Now, we consider the property of $V(n, x)$ as a function of $x$. For any function $g : \mathcal{X}_0 \to \mathbb{R}$, define the mean operator $P(g)(x) = \sum_{y=0}^{x} g(y)p_{x,y}$. We show that $P(g)(x)$ has the following property.

**Lemma 3.2.2.** Suppose Assumption 2.1.1.2 is satisfied, that is, $p_{x,y}$ is nonincreasing in $x$ for $y \leq x$. For any nondecreasing function $g : \mathcal{X}_0 \to \mathbb{R}$, the mean operator $P(g)(x) = \sum_{y=0}^{x} p_{x,y}g(y)$ is also nondecreasing in $x$. 


Proof. We define the first order difference function to be

\[
\Delta Pg(x) = Pg(x + 1) - Pg(x) = \sum_{y=0}^{x} g(y)(p_{x+1,y} - p_{x,y}) + g(x + 1)p_{x+1,x+1}.
\]

Considering

\[
p_{x+1,x+1} = \sum_{y=0}^{x} p_{x,y} - \sum_{y=0}^{x} p_{x+1,y} = \sum_{y=0}^{x} (p_{x,y} - p_{x+1,y}),
\]

it follows from the assumption that

\[
\Delta Pg(x) = \sum_{y=0}^{x} (g(x + 1) - g(y))(p_{x,y} - p_{x+1,y}) \geq 0
\]

which completes the proof.

Lemma 3.2.2 immediately gives the monotonicity of \(V(n,x)\) in \(x\).

**Proposition 3.2.3.** Suppose Assumption 2.1.1 and 2.1.3 are satisfied. For any fixed \(n \in \mathbb{N}\), \(V(n,x)\) is nondecreasing in \(x\) for all \(x \in \mathcal{X}_0\).

**Proof.** We use mathematical induction. The induction hypothesis trivially holds for \(s = 1\) and \(s = 2\); indeed, for any \(x \in \mathcal{X}\), \(n \in \mathbb{N}\),

\[
V_2(n, x) = \lambda + 2n + \min \{q(x), k - p_{x,0}\},
\]

which nondecreasing for all \(x \in \mathcal{X}\).

Assume now that the induction hypothesis holds for some \(s \geq 1\), that is, \(V_s(n,x)\) is nondecreasing in \(x\) for any fixed \(n\). We only need to show that \(V_{s+1}(n,x) = n + \min_{d=0,1} \{E[V_s(N_1,X_1) | n, x, d]\}\) is nondecreasing in \(x\).
We first note that for \( d = 1 \),

\[
\mathbb{E}[V_s(N_1, X_1) \mid n, x, 1] = \mathbb{E}_\lambda \left[ q(x) \left( V_s(n + A_1, L) - V_s(n - 1 + A_1, \beta_0) \right) + V_s(n - 1 + A_1, \beta_0) \right],
\]

where

\[
V_s(n, \beta_0) = \begin{cases} 
V_s(n, \beta) & \text{if } n \in \mathbb{N} \\
0 & \text{if } n = 0
\end{cases}
\]

Since \( q(x) \) is nondecreasing in \( x \) and \( V_s(n + 1, L) - V_s(n, \beta_0) \geq 0 \) for all \( n \), it follows that \( \mathbb{E}[V_s(N_1, X_1) \mid n, x, 1] \) non-decreasing in \( x \).

Also, for \( d = 0 \), we have

\[
\mathbb{E}[V_s(N_1, X_1) \mid n, x, 0] = \mathbb{E}_\lambda \left[ \sum_{y=1}^{x} p_{x,y}(V_s(n + A_1, y)) + p_{x,0}V_s(n - 1 + A_1, \beta_0) \right] = \mathbb{E}_\lambda \left[ \sum_{y=0}^{x} p_{x,y}(g_s(y; n + A_1)) \right] \tag{3.2.3}
\]

where the function \( g_s : \mathcal{X}_0 \to \mathbb{R} \) is given by

\[
g_s(x; n) = \begin{cases} 
V_s(n, y) & \text{if } x \in \mathcal{X}, \\
0 & \text{if } x = 0
\end{cases}
\]

which is a nondecreasing function by the induction hypothesis. This therefore gives the nondecreasing property of \( \mathbb{E}[V_s(N_1, X_1) \mid n, x, 0] \) in \( x \) by Lemma 3.2.2.

Therefore, \( V_s(n, x) \) is nondecreasing in \( x \) for all \( s \geq 1 \) by induction, so its limit \( V(n, x) \) is also nondecreasing in \( x \).

In general, we cannot establish additional structure properties such as convexity of the truncated mean operator \( \mathbf{P}(g)(x) \) as a function of \( x \). However, by assuming further conditions on \( g(x) \) and \( p_{x,y} \), we can show that the function \( \mathbf{P}(g)(x) \) is convex. We now
consider the following *sufficient* condition for the convexity of \( P(g)(x) \).

**Assumption 3.2.4.** For all \( y \leq x \), \( p_{x,y} \) is concave in \( x \), that is, \( p_{x+2,y} - 2p_{x+1,y} + p_{x,y} \geq 0 \).

Assumption (3.2.4) is a modest condition for the desired convexity. Intuitively, this assumption states that patients receiving treatment recover “smoothly” over their health state.

**Lemma 3.2.5.** Suppose Assumption 2.1.1.2 and Assumption 3.2.4 are satisfied. For any nondecreasing function \( g : X_0 \to \mathbb{R} \), the mean operator \( P(g)(x) = \sum_{y=0}^{x} p_{x,y} g(y) \) is nondecreasing convex in \( x \).

**Proof.** The monotonicity proof is already given. Let us now consider the convexity property. We define the second order difference function to be

\[
\Delta^2 P g(x) = \Delta P g(x+1) - \Delta P g(x).
\]

Then, considering

\[
\Delta P g(x) = \sum_{y=0}^{x} (g(x+1) - g(y))(p_{x,y} - p_{x+1,y}),
\]

we have

\[
\Delta^2 P g(x) = \sum_{y=0}^{x+1} (g(x+2) - g(y))(p_{x+1,y} - p_{x+2,y})
\]

\[
- \sum_{y=0}^{x} (g(x+1) - g(y))(p_{x,y} - p_{x+1,y})
\]

\[
\geq (g(x+2) - g(x+1))(p_{x+1,x+1} - p_{x+2,x+1})
\]

\[
+ \sum_{y=0}^{x} (g(x+2) - g(y))(2p_{x+1,y} - p_{x+2,y} - p_{xy})
\]

\[
\geq 0.
\]

The last inequality follows from the concavity of \( p_{x,y} \). \( \square \)
3.2.2 The Optimal Threshold Policies

Given the above structural results, we are now in a position to fully characterize the structure of the optimal discharge policy of our control problem in the one-bed setting, which is the main result of this chapter.

**Theorem 3.2.6.** Suppose Assumptions 2.1.3-3.2.4 are satisfied. Then, the optimal discharge policy is characterized by a two-threshold policy (TTP). In particular, for any $N_t = n \in \mathbb{N}$ there exist thresholds $\tau_n, n \in \mathcal{X}$ such that at period $t$:

1. If $X_t \leq \tau_n$ or $X_t \geq \tau_n$, it is optimal to discharge the patient in the unit;

2. otherwise, it is optimal to not discharge the patient.

*Proof.* We first note that the dynamic programming equations (3.1.3) can be rewritten as for all $n \in \mathbb{N}$ and $x \in \mathcal{X}$,

$$V(n, x) = n + \min_{d \in \{0, 1\}} \{E[V(N_1, X_1) | N_0 = n, X_0 = x, D_0 = d]\}$$

$$= n + E[V(N_1, X_1) | n, x, 1]$$

$$+ \min \left\{0, \begin{array}{c}
E[V(N_1, X_1) | n, x, 0] - E[V(N_1, X_1) | n, x, 1]
\end{array} \right\}.$$

It follows that the optimal decision is to discharge if and only if

$$G_n(x) = E[V(N_1, X_1) | n, x, 0] - E[V(N_1, X_1) | n, x, 1] \geq 0$$

(3.2.5)

Therefore, to characterize the structure of the optimal discharge policy amounts to an understanding of the level-crossing properties of the function $G_n(x)$ defined in (3.2.5).

Then, for all $n \in \mathbb{N}$ and $x \in \mathcal{X}$,

$$E[V(N_1, X_1) | n, x, 1] = q(x) \left[ E_\lambda \left[ V(n + A_1, L) - V(n - 1 + A_1, \beta_0) \right] \right]$$

$$+ E_\lambda \left[ V(n - 1 + A_1, \beta_0) \right].$$
and

\[
\mathbb{E}[V(N_1, X_1) | n, x, 0] = \mathbb{E}_\lambda \left[ \sum_{y=1}^{x} p_{x,y} (V(n + A_1, y) - V(n - 1 + A_1, \beta_0)) \right.

\left. + V(n - 1 + A_1, \beta_0) \right],
\]

which implies

\[
G_n(x) = \mathbb{E}_\lambda \left[ \sum_{y=1}^{x} p_{x,y} (V(n + A_1, y) - V(n - 1 + A_1, \beta_0)) \right.

\left. - q(x) [V(n + A_1, L) - V(n - 1 + A_1, \beta_0)] \right].
\]

Since \(\mathbb{E}[V(N_1, X_1) | n, x, 1]\) is concave in \(x\) and \(\mathbb{E}[V(N_1, X_1) | n, x, 0]\) is convex in \(x\) by Lemma 3.2.5, \(G_n(x)\) is therefore a convex function in \(x\) for all \(x \in \mathcal{X}\) and \(n \in \mathbb{N}\).

In particular, for any fixed \(n \in \mathbb{N}\), the function \(G_n(x)\) changes signs at most twice over \(\mathcal{X}\). Thus, the 0-superlevel-set of \(G_n(x)\) can be expressed as

\[
\{ x \in \mathcal{X} : G_n(x) \geq 0 \} = \{ x \in \mathcal{X} : x \leq \tau_n, x \geq \tau_n \},
\]

for some \(\tau_n, \tau_n \in \mathcal{X}\), which completes the proof. \(\square\)

For any given state \((n, x) \in \mathbb{N} \times \mathcal{X}\), the exact thresholds values can be computed through characterizing the 0-superlevel-set of the difference function \(G_n(x)\), that is given by,

\[
G_n(x) = \mathbb{E}_\lambda \left[ \sum_{y=1}^{x} p_{x,y} (V(n + A_1, y) - V(n - 1 + A_1, \beta_0)) \right.

\left. - q(x) [V(n + A_1, L) - V(n - 1 + A_1, \beta_0)] \right].
\] (3.2.6)

In particular, given any \((n, x) \in \mathbb{N} \times \mathcal{X}\), it is optimal to discharge a patient if and only if \(G_n(x) \geq 0\).

By the convexity of \(G_n(x)\), we first note that it is optimal to discharge patients in the
good health condition region. The intuition behind this result is that we only discharge someone when the risk of readmission is sufficiently low. On the other hand, discharging a patient in a sufficiently bad state seems less intuitive. One way to interpret this result is that patients in a bad state cannot be efficiently cured under the treatment given at the current hospital unit: As such, a reasonable action to take is to move those patients to different care units or hospitals, when the risk of having future readmissions is relatively low. Clearly, the particular values of the optimal thresholds depend on the model parameters. We have the following sufficient condition on the function $q(x)$ so that a one-threshold policy (OTP) is optimal.

**Corollary 3.2.7.** Suppose $q(L) = 1$. Then the optimal discharge policy is characterized by a one-threshold policy (OTP). In other words, for any $n \in \mathbb{N}$, there exists a unique threshold $\tau_n \in \mathcal{X}$ so that it is optimal to discharge when $X_t \leq \tau$ and not discharge otherwise.

**Proof.** Consider the function $\Delta_n : \mathcal{X} \to \mathbb{R}$ defined by

$$\Delta_n(x) = \sum_{y=1}^{x} p_{x,y}(V(n,y) - V(n - 1, \beta_0)) - q(x) [V(n,L) - V(n - 1, \beta_0)]$$

so that $G_n(x) = \mathbb{E}_\lambda [\Delta_{n,A_t}(x)]$.

If $q(L) = 1$, then for any $n \in \mathbb{N}$, we have

$$\Delta_n(L) = \sum_{y=1}^{L} p_{L,y}(V(n,y) - V(n - 1, \beta_0)) - [V(n,L) - V(n - 1, \beta_0)] \leq 0,$$

where the inequality is due to the monotonicity of $V(n,x)$.

Therefore, $G_n(x)$ is convex on $\mathcal{X}$ and $G_n(L) \leq 0$. This implies $G_n(x)$ changes sign at most once on $\mathcal{X}$, which leads to the desired structure result for the optimal policy. $\square$

We also obtain a generic constraint over the values of the thresholds, and it only depends on the functions $q(x)$ and $p_{x,0}$. In some cases, this constraint allows us to estimate
the values of $\tau_n$ and $\tau_n \in \mathcal{X}$ without solving for $V(n,x)$ in the dynamic programming equations.

**Proposition 3.2.8.** Suppose the function $H(x) = 1 - p_{x,0} - q(x)$ has more than one zero such that its first zero and last zero are respectively $h_1$ and $h_2$. Then, for all $n$, the threshold $\tau_n$ is bounded above by $h_1$ and $\tau_n$ is bounded below by $h_2$.

**Proof.** By the monotonicity of $V$ in $x$, we have for any $n \in \mathbb{N}$

$$
\Delta_n(x) = \sum_{y=1}^{x} p_{x,y} (V(n,y) - V(n-1,\beta)) - q(x) [V(n,L) - V(n-1,\beta)] \\
\leq \sum_{y=1}^{x} p_{x,y} (V(n,L) - V(n-1,\beta)) - q(x) [V(n,L) - V(n-1,\beta)] \\
= (V(n,L) - V(n-1,\beta))(1 - p_{x,0} - q(x)),
$$

which implies that

$$
G_n(x) \leq \mathbb{E}_\lambda [(V(n + A_1,L) - V(n - 1 + A_1,\beta))] (1 - p_{x,0} - q(x)) \\
= \mathbb{E}_\lambda [(V(n + A_1,L) - V(n - 1 + A_1,\beta))] H(x).
$$

Therefore, given the first and last zero of $H(x)$ to be $h_1$ and $h_2$, we have $\tau_n \leq h_1$ and $\tau_n \geq h_2$. 

In the next section, we will illustrate the structural results using a numerical example.

### 3.3 Numerical Example

#### 3.3.1 OTP Problem

We first show an example of the discharge control problem in which a one-threshold policy (OTP) is optimal.

Consider a $K = 1$ discharge control problem with health condition state space $\mathcal{X}_0 = \{0,1,2,...,9\}$ and the treatment process is governed by the following lower triangular
transition matrix

\[
P = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.58 & 0.42 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.55 & 0.40 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.52 & 0.38 & 0.05 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.47 & 0.34 & 0.04 & 0.05 & 0.1 & 0 & 0 & 0 & 0 & 0 \\
0.42 & 0.30 & 0.04 & 0.05 & 0.08 & 0.11 & 0 & 0 & 0 & 0 \\
0.36 & 0.26 & 0.03 & 0.05 & 0.06 & 0.10 & 0.14 & 0 & 0 & 0 \\
0.28 & 0.21 & 0.02 & 0.04 & 0.05 & 0.08 & 0.10 & 0.22 & 0 & 0 \\
0.20 & 0.14 & 0.015 & 0.024 & 0.036 & 0.054 & 0.086 & 0.145 & 0.30 & 0 \\
0.10 & 0.07 & 0.008 & 0.013 & 0.018 & 0.028 & 0.045 & 0.070 & 0.168 & 0.48
\end{pmatrix},
\]

with a new arrival distribution

\[
\beta = \begin{pmatrix}
0 & 0.10 & 0.17 & 0.03 & 0.18 & 0.16 & 0.02 & 0.16 & 0.02 & 0.16
\end{pmatrix}
\]

We assume there is at most one arrival per period with probability \( \lambda = 0.1 \) and the readmission probability is given by

\[
Q = \begin{pmatrix}
0 & 0 & 0.14 & 0.28 & 0.41 & 0.53 & 0.65 & 0.77 & 0.89 & 1.00
\end{pmatrix}.
\]

We note that the transition probabilities \( p_{x,y} \) and \( q(x) \) satisfy the monotonicity and concavity assumptions mentioned in (3.1)-(3.4). In particular, \( q(L) = 1 \) so that an OTP is optimal according to Corollary 3.2.7.

We implemented the standard value iteration procedure in Matlab to compute the optimal values with the parameters defined above. As a result, the values of the difference function \( G_n(x) \) is graphed in Figure 3.3.2 for \( n = 1 \).

As plotted in Figure 3.3.2, \( G_1(x) \) is a convex function that crosses zero once around \( x = 2 \). Thus, at every decision time \( t \), given state \((1, x)\), Theorem 3.2.6 suggests that we
Figure 3.3.1: The graph of the difference function $G_n(x)$ is plotted for $n = 1$. The threshold value is marked by a filled circle at $x = 2$.

discharge the patient whenever $G_1(x) \geq 0$, that is, when $x \leq 2$. As described in Corollary 3.2.7, the optimal discharge policy is a one-threshold policy (OTP) when $q(L) = 1$. This simply implies that the benefits from discharging dominate over non-discharging when patients are in health states 1 and 2. On the other hand, for $x \geq 2$, since the risk of having readmissions is sufficiently high, we should not discharge a patient.

3.3.2 TTP Problem

We next illustrate a two-threshold policy (TTP) that can be obtained by modifying some parameters of the OTP problem defined above. In particular, if we have $q(x)$ to be smaller for large values of $x$, $G_1(x)$ in the OTP problem may lead to a TTP because of its convexity, which is depicted in the following example.

We consider the same problem setting as in the OTP problem defined above, except the readmission distribution is given by

$$Q = \begin{pmatrix} 0 & 0 & 0.12 & 0.24 & 0.35 & 0.46 & 0.50 & 0.54 & 0.58 & 0.62 \end{pmatrix}.$$ 

We compute the values of $G_n(x)$ through value iterations and the resulting $G_1(x)$ is
plotted in Figure 3.3.2 Thus, at every decision time $t$, given state $(1, x)$, Theorem 3.2.6

suggests that we discharge the patient whenever $G_1(x) \geq 0$, that is, when $x \leq 2$ or $x = 9$.

### 3.3.3 Comparison Between Discharge Policies

We conclude our numerical analysis with a performance comparison between the optimal discharge policy and the myopic policy in the OTP setting. Recall that the MYP policy suggests us to discharge a patient whenever $H(x) = 1 - p_{x,0} - q(x) \geq 0$. For the same set of parameters given above, $H(x)$ is plotted in Figure 3.3.3.

As plotted in Figure 3.3.3, $H(x)$ crosses zero once around $x = 6$. Thus, at every decision time $t$, given state $(1, x)$, the myopic policy suggests that we discharge the patient whenever $H(x) \geq 0$, that is, when $x \leq 6$. Therefore, the myopic policy (MYP) is a one-threshold policy as well.

Both OTP and MYP suggest that we discharge the patients only when they are relatively healthy, that is, when $x$ is small. However, the threshold values are very different for each discharge policy. This implies applying the myopic policy may lead significant
larger system sizes compared to the optimal OTP. To illustrate this, we simulated the total expected system size using MYP and OTP with the same parameter set as in the OTP problem. We ran 1000 simulations to compute the expected system size under each discharge policy. The computation was coded in Matlab on a Intel i7 CPU at 3.40 GHZ with 16.0GB RAM.

As a result of the simulation, a performance comparison of the two discharge policies MYP and OTP is provided in Table 3.3.1.

<table>
<thead>
<tr>
<th>((N_0, X_0))</th>
<th>(V_{OTP})</th>
<th>(V_{MYP})</th>
<th>(V_{MYP} - V_{OTP})</th>
<th>Percentage Gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2,6)</td>
<td>8.7</td>
<td>10.8</td>
<td>2.1</td>
<td>19.44 %</td>
</tr>
<tr>
<td>(5,6)</td>
<td>40.6</td>
<td>49.9</td>
<td>9.3</td>
<td>18.64 %</td>
</tr>
<tr>
<td>(5,1)</td>
<td>40.2</td>
<td>47.6</td>
<td>7.4</td>
<td>15.88 %</td>
</tr>
<tr>
<td>(10,6)</td>
<td>147.0</td>
<td>172.5</td>
<td>25.5</td>
<td>14.78%</td>
</tr>
<tr>
<td>(15,6)</td>
<td>312.8</td>
<td>360.2</td>
<td>47.4</td>
<td>13.17%</td>
</tr>
<tr>
<td>(15,1)</td>
<td>310.4</td>
<td>359.6</td>
<td>49.2</td>
<td>13.68 %</td>
</tr>
</tbody>
</table>

Table 3.3.1: A performance comparison between the OTP and MYP is summarized for different starting system state \((N_0, X_0)\).

Table 3.3.1 shows that the OTP performs better than the MYP for any starting
state \((N_0, X_0)\). In particular, using the OTP gives a 19.44% total system size reduced over the MYP when there are less patients in the system. This shows the importance of determining the optimal policy for the multi-stage dynamic programming rather than simply considering a two stage approximate solution to the actual problem. On the other hand, we also observe that the relative benefits of using OTP becomes less substantial when the starting system size is large. This motivates the potential of developing a heuristic policy from the MYP, under which the resulting expected system size should be bounded between \(V_{OTP}\) and \(V_{MYP}\).

### 3.4 Summary

In this chapter, we considered the discharge control problem in a one-bed unit setting. We first established some preliminary structural results on the mean operator and the value function through analyzing the dynamic programming equations. The main result of the analysis is that a two threshold policy (TTP) is optimal for the discharge control in the one-bed setting. We also presented illustrative numerical experiments comparing the optimal policy to a myopic policy (MYP).

In the next chapter, we investigate the problem in the general multi-bed setting, that is, when \(K \geq 2\). Due to its special problem structure, the analysis will diverge from the one-bed problem and lead to qualitatively different results.
Chapter 4

Discharge Control in the Multiple Bed Setting

In the multi-bed case, that is $K \geq 2$, we face a shared “resource” problem that is not present in the one-bed case: all the beds allocate waiting patients from the same line. This leads to a more complicated boundary setting in contrast to the $K$-isolated beds situation in which each beds has its own queue.

4.1 Model

For $K \in \mathbb{N}$, recall that the value function is given by, for all $n \in \mathbb{N}_0$ and $x = (x^k) \in \mathcal{X}^K_0$,

$$V(n, x) = \inf_{\pi \in \Pi} \mathbb{E}^\pi \left[ \sum_{t=0}^{T_0} N_t \mid N_0 = n, X_0 = x \right], \quad (4.1.1)$$

that uniquely satisfies the dynamic programming equations given by

$$V(0, 0) = 0,$$

$$V(n, x) = n + \min_{d \in \{0, 1\}^K} \{ \mathbb{E}[V(N_1, X_1) \mid n, x, d] \}, \quad n \in \mathbb{N}, x \in \mathcal{X}. \quad (4.1.2)$$

In addition, the value function can be still obtained as the limit of the $s$-stage value
functions, for all \( x \in \mathcal{X}_0^K \), \( n \in \mathbb{N}_0 \)

\[
V(n, x) = \lim_{s \to \infty} V_s(n, x),
\]

(4.1.2)

where \( V_s(n, x) \) satisfies the following dynamic programming equations,

\[
V_s(n, x) = n + \min_{d \in \{0, 1\}^K} \{ \mathbb{E}[V_{s-1}(N_1, X_1) \mid N_0 = n, X_0 = (x), D_0 = d] \}, \quad \sigma = 1, \ldots, s,
\]

\[
V_0(n, x) = 0.
\]

(4.1.3)

### 4.2 DP Analysis

#### 4.2.1 Preliminary Results

We begin by verifying that some of the structure properties that were obtained for the \( K = 1 \) bed case hold in the general case.

**Proposition 4.2.1.** For any fixed \( x \in \mathcal{X}^K \), \( V(n, x) \) is nondecreasing in \( n \) for all \( n \in \mathbb{N}_0 \).

**Proof.** We consider the \( s \)-stage dynamic programming equations and then proceed by way of mathematical induction.

The induction hypothesis is trivial for \( s = 1 \); indeed, for any \( x \in \mathcal{X}^K \), \( V_1(n, x) = n \) is linearly increasing in \( n \). Assume now that the induction hypothesis holds for some \( s \geq 1 \), that is, \( V_s(n, x) \) is nondecreasing in \( n \) for all \( n \).

It follows from equation (4.1.3) that for any \( n \in \mathbb{N} \),

\[
V_{s+1}(n, x) = n + \min_{d \in \{0, 1\}^K} \{ \mathbb{E}[V_s(N_1, X_1) \mid n, x, d] \}
\]

(4.2.1)

where \( \mathbb{E}[V_s(N_1, X_1) \mid n, x, d] \) are positive weighted sums of \( V_s(n, x) \), therefore nondecreasing in \( n \). The minimum of functions is still nondecreasing. Combining this with the boundary value \( V_{s+1}(0, 0) = 0 \), it follows that \( V_{s+1}(n, x) \) is a nondecreasing function in \( n \) for any fixed \( x \in \mathcal{X}^K \). By induction, the property holds for any \( s \geq 1 \). Therefore, \( V(n, x) = \lim_{s \to \infty} V_s(n, x) \) is nondecreasing in \( n \) for any fixed \( x \), which completes the
As in the $K = 1$ case, given any $x = (x^k, x^{k-})$, we want to study the property of $V(n, x^k, x^{k-})$ as a function of $x^k$ for fixed $n$ and $x^k \in \mathcal{X}^{K-1}$. Let us consider first the mean operator $P(g)(x^k; x^{k-}) = \sum_{y=1}^{x^k} p_{x^k,y} g(y; x^{k-})$.

Lemma 4.2.2. Suppose Assumptions 2.1.1.2 is satisfied, that is, $p_{x,y}$ is nonincreasing for $y \leq x$, and for any $k \in \{1, ..., K\}$, $g(x^k; x^{k-})$ is nondecreasing in $x^k$ for any fixed $x^{k-} \in \mathcal{X}_0^{K-1}$, the mean operator $P(g)(x^k; x^{k-}) = \sum_{x^k_0=0}^{x^k} p_{x^k_0,y} g(x^k_1; x^{k-})$ is also nondecreasing in $x^k$ on $\mathcal{X}_0$ for all fixed $x^{k-} \in \mathcal{X}_0^{K-1}$.

Proof. For any fixed values of $x^{k-} \in \mathcal{X}_0^{K-1}$, we note that $g(x; x^{k-})$ is simply a nondecreasing univariate function of $x^k$. Then the result trivially follows from the univariate case.

Lemma 4.2.2 gives the monotonicity of $V(n, x)$ in $x^k$ for any fixed $n$ and $x^{k-}$.

Proposition 4.2.3. Suppose Assumption 2.1.1 and 2.1.3 are satisfied. Given any $k$ and fixed $n \in \mathcal{N}$ and $x^{k-} \in X^{K-1}$, $V(n, x^k, x^{k-})$ is nondecreasing in $x^k$ on $\mathcal{X}_0$.

Proof. For fixed $n$ and $x^{k-}$, $V(n, x^k, x^{k-})$ is simply a univariate function of $x^k$ and we can proceed by mathematical induction. As in the $K = 1$ case, the induction hypothesis trivially holds when $s = 1$. Suppose we assume the monotonicity holds for some $s \geq 1$. Then we consider the $(s+1)$-stage dynamic programming equations, in particular, when $n = 1$, we have

$$V_{s+1}(1, x, 0) = 1 + \min_{d \in \{0, 1\}} \{ \mathbb{E}[V_s(N_1, X_1) \mid 1, x, d] \}.$$
The expected cost-to-go terms can be rewritten as

\[
\min_{d \in \{0, 1\}} \{ \mathbb{E}[V_s(N_1, X_1) \mid 1, x, d] \}
\]

\[
= \min \left\{ \mathbb{E}_{\lambda} \left[ q(x) \left( V_s(1 + A_1, L, \beta_0) - V_s(A_1, \beta_0, \beta_0) \right) + V_s(A_1, \beta_0, \beta_0) \right], \right. \\
\left. \mathbb{E}_{\lambda} \left[ \sum_{x=1}^{x} p_{x,x_1} \left( V_s(1 + A_1, x_1, \beta_0) - V_s(A_1, \beta_0, \beta_0) \right) + V_s(A_1, \beta_0, \beta_0) \right] \right\},
\]

where

\[
V_s(n, x^1, \beta_0) = \begin{cases} 
\sum_{x^K=1}^{L} \beta(x^K) \left( \cdots \left( \sum_{x^2=1}^{L} \beta(x^2)V_s(n, x^1, x^2, \ldots, x^K) \right) \cdots \right) & \text{if } n \geq K \\
\sum_{x^1=1}^{L} \beta(x^1) \left( \cdots \left( \sum_{x^2=1}^{L} \beta(x^2)V_s(n, x^1, x^2, \ldots, x^1, 0) \right) \cdots \right) & \text{if } n < K.
\end{cases}
\]

By the induction hypothesis and the monotonicity of the mean operator, \(V_{s+1}(1, x, 0)\) is nondecreasing in \(x\). Similarly, for \(n \geq 2\), the dynamic programming equation is given by

\[
V_{s+1}(n, x) = n + \min_{d \in \{0, 1\}^K} \left\{ \mathbb{E}[V_s(N_1, X_1) \mid N_0 = n, X_0 = (x), D_0 = (d^k, d^{k^-})] \right\}.
\]

where \(d^{k^-} = (d^1, d^2, \ldots, d^{k-1}, d^{k+1}, \ldots, d^K)\) denotes the discharge actions applied beds other than the \(k\)-th bed. Given any discharge control \(d^{k^-} \in \{0, 1\}^K\), for \(d^k = 1\), we have the expected cost-to-go given by

\[
\mathbb{E}[V_s(N_1, X_1) \mid n, x, d = (1, d^{k^-})] = \mathbb{E}_{\lambda} E_{d^{k^-}} \left[ q(x^k) \left( V_s(n + A_1 - Q_1^{d^{k^-}}, L, \beta_0) - V_s(n - 1 + A_1 - Q_1^{d^{k^-}}, \beta_0, \beta_0) \right) \\
+ V_s(n - 1 + A_1 - Q_1^{d^{k^-}}, \beta_0, \beta_0) \right],
\]

where

\[
Q_1^{d^{k^-}} = \sum_{j \neq k} d^j S_j + \sum_{j \neq k} (1 - d^j)W_j
\]

denotes the number of patients on the \(k^-\) beds leaving the unit under the discharge policy.
Similarly, for \( d^k = 0 \), we have the expected cost-to-go given by

\[
\mathbb{E}[V_s(n_1, X_1) | n, x, y, d = (0, d^k)] = \mathbb{E}_{\lambda}E_{d^k} \left[ \sum_{x^k_1=1}^{x^k} p_{x^k, x^k_1} \left( V_s(n + A_1 - Q^{d^k-1}_1, x^k_1, \beta_0) - V_s(n - 1 + A_1 - Q^{d^k-1}_1, \beta_0, \beta_0) \right) + V_s(n - 1 + A_1 - Q^{d^k-1}_1, \beta_0, \beta_0) \right]
\]

which are both functions nondecreasing in \( x^k \) for fixed \( n \) and \( x^{k-} \). Thus, the nondecreasing property holds for any \( s \geq 1 \) and the limit \( V(n, x) \), which completes the proof.

**Lemma 4.2.4.** Suppose Assumptions 2.1.1-3.2.4 are satisfied, and that for any \( k \), \( g(x^k; x^k-) \) is convex in \( x^k \) for any fixed \( x^k- \). Then the mean operator \( P(g)(x^k; x^k-) = \sum_{x^k_1=0}^{x^k} p_{x^k, x^k_1} g(x^k_1; x^k-) \) is convex in \( x^k \).

**Proof.** For any fixed values of \( x^k- \in \mathcal{X}_0^{K-1} \), we note that \( g(x; x^k-) \) is simply a nondecreasing univariate function of \( x^k \). Then everything trivially follows from the univariate case.

**4.2.2 The Optimal Discharge Policies**

Similar to the one-bed setting, for any fixed discharge control \( d^{k-} \in \{0, 1\}^{K-1} \), we can define a difference function \( G_{d^{k-}}(x^k; x^k-) : \mathbb{N} \times \mathcal{X} \rightarrow \mathbb{R} \) given by

\[
G_{d^{k-}}(x^k; x^k-) = \mathbb{E} \left[ V(N_1, X^k_1, X^k_1) | d = (0, d^{k-}) \right] - \mathbb{E} \left[ V(N_1, X^k_1, X^k_1) | d = (1, d^{k-}) \right]
\]

(4.2.2)

which represents the marginal difference between non-discharging and discharging in the \( k \)-th bed while fixing the discharge actions for other beds.

**Proposition 4.2.5.** For any \( k \) and fixed \( d^{k-} \in \{0, 1\}^{K-1} \), \( x^k- \in \mathcal{X}^{K-1} \) and \( n \in \mathbb{N} \), \( G_{d^{k-}}(x^k; x^k-) \) is convex in \( x^k \) on \( \mathcal{X} \).
Proof. For \( n = 1 \), it immediately follows from the univariate result. For \( n \geq 2 \), for any fixed \( d^k \in \{0, 1\}^{K-1} \), the expected cost-to-go for \( d^k = 1 \) is given by

\[
\mathbb{E}[V(N_1, X_1) \mid n, x^k, x^k, d = (1, d^k)] = \mathbb{E}_A E_{d^k} \left[ q(x^k)(V(n + A_1 - Q_1^{d^k}, L, X_1^k) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X_1^k)) \\
+ V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X_1^k) \right],
\]

which is a concave function in \( x^k \) by the properties of \( q(x^k) \).

Similarly, for \( d^k = 0 \), we have

\[
\mathbb{E}[V(N_1, X_1) \mid n, x^k, x^k, d = (0, d^k)] = \mathbb{E}_A E_{d^k} \left[ \sum_{x_1^k = 1}^{x^k} p_{x^k, x_1^k} \left( V(n + A_1 - Q_1^{d^k}, x^k, X_1^k) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X_1^k) \right) \\
+ V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X_1^k) \right],
\]

which is a convex function in \( x^k \) by Lemma (4.2.4).

Then, given any \( d^k \in \{0, 1\}^{K-1} \), it follows that for fixed \( x^k \in X^{K-1} \), \( n \in \mathbb{N} \),

\[
G_n^{d^k} (x^k; x^k) \\
= \mathbb{E} \left[ V(X_1^k, X_1^k, N_1) \mid d = (0, d^k) \right] - \mathbb{E} \left[ V(X_1^k, X_1^k, N_1) \mid d = (1, d^k) \right] \\
= \mathbb{E}_A E_{d^k} \left[ \sum_{x_1^k = 1}^{x^k} p_{x^k, x_1^k} \left( V(n + A_1 - Q_1^{d^k}, x^k, X_1^k) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X_1^k) \right) \\
- q(x^k)(V(n + A_1 - Q_1^{d^k}, L, \beta_0) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X_1^k)) \right]
\]

is a convex function in \( x^k \) on \( X \).

In the multi-bed setting, knowing the properties of \( G_n^{d^k} (x^k; x^k) \) for fixed \( d^k \) does not immediately give the optimal discharge policy in the \( k \)-th bed. In fact, it follows from the dynamic programming equations that the \( k \)-th component of the optimal discharge
policy $\pi^* = (\pi^{*k})$ is uniquely characterized by, for any $k \in \{1, ..., K\}$,

$$
\pi^{*k}(x^k, n, x^{k-}) = \begin{cases} 
0 & \text{if } G^k_n(x^k, x^{k-}) < 0, \\
1 & \text{if } G^k_n(x^k, x^{k-}) \geq 0,
\end{cases}
$$

and the marginal difference function $G^k_n$ is given by

$$
G^k_n(x^k, x^{k-}) = \sum_{d^{k-} \in \{0,1\}^{K-1}} \mathbb{I}_{\{\pi^{*k-(n,x)} = d^{k-}\}} \cdot G^{d^{k-}}_{d^{k-}}(x^k, x^{k-})
$$

where $\pi^{*k-} = (\pi^{*1}, ..., \pi^{*k+1}, \pi^{*k-1}, ..., \pi^{*K})$ denotes the other marginal components of $\pi^*$.

In order to characterize the optimal policy, we need to solve for $G^k_n$ in a system of $K$ symmetric equations as described in (4.2.4). As the solutions to those equations must be identical by symmetry, that is, $G^j_n \equiv G^k_n$ for any $j, k \in \{1, ..., K\}$, we can suppress the dependence of $G^k_n$ on $k$ and simply write it as $G_n$. This further implies that the optimal discharge policy must be symmetric in the sense that it is invariant under a permutation of its marginal components. In other words, $\pi^*$ must have identical components so that $\pi^{*j} = \pi^{*k}$ for any $j, k \in \{1, ..., K\}$.

As in the one-bed setting, a generic constraint on the thresholds, which only depends on the functions $q(x)$ and $p_{x,0}$, can also be obtained using $G_n$.

**Proposition 4.2.6.** For any $k$, fixed $x^{k-} \in X^{K-1}$ and $n \in \mathbb{N}$, there exists a constant $C^k_n \in \mathbb{R}^+$ such the marginal difference $G_n(x^k, x^{k-})$ as a function of $x^k$ in $x$ is bounded above by $H(x^k)C^k_n$.

**Proof.** It is a direct result from analyzing the difference functions with fixed policies as defined in (4.2.2).
Considering, for any $d^k \in \{0, 1\}^{K-1}$,

$$G_n^{d^k} (x^k, x^{k-}) = \mathbb{E}_\lambda E_{d^k} \left[ \sum_{x^k_1=1}^{x^k} p_{x^k_1 \cdot x^k} \left( V(n + A_1 - Q_1^{d^k}, x^k_1, X^{k-}) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X^{k-}) \right) \right. $$

$$- q(x^k)(V(n + A_1 - Q_1^{d^k}, L, \beta_0) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X^{k-})) \right] \leq \mathcal{H}(x^k) \mathbb{E}_\lambda E_{d^k} \left[ V(n + A_1 - Q_1^{d^k}, L, X^{k-}) - V(n - 1 + A_1 - Q_1^{d^k}, \beta_0, X^{k-}) \right],$$

we therefore have for some positive constant $C_n^k \in \mathbb{R}^+$,

$$G_n^k (x^k, x^{k-}) = \sum_{d^k \in \{0, 1\}^{K-1}} \mathbb{I}_{\{\pi^{k-} (n, x) = d^k\}} \cdot G_n^{d^k} (x^k, x^{k-}) \leq C_n^k \mathcal{H}(x^k),$$

for all $x^k \in \mathcal{X}$, which completes the proof. 

As the above results suggest, it is difficult to fully characterize the structure of the optimal discharge policy in the multi-bed setting. In particular, we cannot easily establish the convexity of $G_n(x^k, x^{k-})$ although each of its partition $G_n^{d^k} (x^k, x^{k-})$ is convex in $x^k$ for all fixed $x^{k-}$. The level-crossing properties of $G_n(x^k, x^{k-})$ in the centralized patient state variable $x^k$ depends on other patients’ state $x^{k-}$, which requires a deeper analysis of the system of equations defined in (4.2.4).

In addition, it is also computationally challenging to directly solve the dynamic programming equations unless it is a very small state space. The computation time grows exponentially as the number of beds $K$ increases. However, if we fix a policy for all beds except the $k$-th bed, the marginal difference function $G_n^{d^k} (x^k, x^{k-})$ is a much easier object to deal with both analytically and computationally. Therefore, we propose an algorithm that breaks the complicated joint optimization problem into easier subproblems. We show how under this approach, the limiting policy generated will still achieve joint optimality.
4.3 Marginal Policy Iteration (MPI) Algorithm

In a $K \geq 2$ setting, for any $k \in \{1, ..., K\}$, suppose we fix a $K - 1$ bed discharge policy $\pi^{k-} : \mathbb{N} \times \mathcal{X}_0^K \rightarrow \{0, 1\}^{K-1}$ for all the beds other than the $k$-th bed. Then, a natural control problem to consider is how one should sequentially discharge patients on the $k$th bed in order to minimize the expected system size if other beds use discharge policy $\pi^{k-}$. In other words, we can define a marginal optimization problem with the value function given by

\[
W^{\pi^{k-}}(x^k; n, x^{k-}) = \inf_{\pi_k} \mathbb{E}^\pi \left[ \sum_{t=0}^{T_0} N_t \mid x^k, n, x^{k-} \right],
\]

where $\pi = (\pi^k; \pi^{k-})$ is the discharge policy with fixed $K - 1$ marginal policies. Solving for (4.3.1) is a well-defined discrete time stochastic control problem, and it uniquely satisfies the following dynamic programming equation:

\[
W^{\pi^{k-}}(x^k; n, x^{k-}) = n + \min_{d^k \in \{0, 1\}} \left\{ \mathbb{E} \left[ W^{\pi^{k-}}(X^k_1; N_1, X^{k-}_1) \mid (d^k, \pi^{k-}) \right] \right\},
\]

which can be solved as in the $K = 1$ setting.

4.3.1 The MPI Algorithm

Motivated by the tractability of the marginal optimization problem, we define the marginal policy iteration algorithm (MPI) given in Algorithm 1 below.

The idea is to optimize the discharge control for one-bed while fixing the discharge policies for the other beds, and repeat this procedure for all beds in a round-robin manner one bed at a time. For simplicity, we initialize the iteration with a symmetric policy whose components are identical marginal discharge policies.
Algorithm 1 MPI

1. **Initiation** Set \( m = 1 \) and choose any symmetric policy \( \rho_1 = (\rho^k)_{k \in \{1, \ldots, K\}} \) with identical components \( \rho^k : \mathbb{N} \times \mathcal{X}^k \to \{0, 1\} \).

2. **Policy Improvement**

3. Let \( k = m \mod K \). Optimize the control for the \( k \)-th bed while applying \( \rho^{k-}_m \) to the \( k^- \) beds. In particular, we solve the marginal dynamic programming problem given by

\[
W^\rho_1(x^k; n, x^{k^-}) = n + \min_{d^m} \left\{ \mathbb{E} \left[ W^\rho_1(x^k_1; N_1, X^{k^-}_1) | (d^1, \rho^{k-}_m) \right] \right\} .
\]  

4. Update the marginal discharge policy \( \rho_{m+1} \) using \( G^{\rho_m}_n(x^k; x^{k-}) \):

\[
\rho^{k}_{m+1}(x^k; n, x^{k^-}) = \begin{cases} 
1 & \text{if } G^{\rho_m}_n(x^k; x^{k^-}) \geq 0 \\
0 & \text{if } G^{\rho_m}_n(x^k; x^{k^-}) < 0,
\end{cases}
\]

and let \( \rho^j_{m+1} = \rho^j_m \) for all \( j \neq k \).

5. **if** \( \rho_{m+1} \equiv \rho_m \) **then**

6. Stop and output a discharge policy \( \pi = (\pi^k) \) with \( \pi^k = \rho^k_m \) for all \( k \).

7. **else**

8. Set \( m = m + 1 \) and repeat policy improvement with a new fixed policy obtaining from \( \rho_m \).

We first note that by construction, there is always an improvement in the expected system size \( W^\rho_{m} \) after each iteration.

**Lemma 4.3.1.** Given any \( \rho_1 \), \( W^\rho_{m} \) is monotonically decreasing in \( m \), that is, for any \( n \in \mathbb{N}, x \in \mathcal{X}^K \), \( W^\rho_{m+1}(n, x) \leq W^\rho_{m}(n, x) \).

**Proof.** It follows from the definition of \( W^\rho_{m} \).

For any \( m \geq 1 \) such that \( k = m \mod K \) and \( n \in \mathbb{N}, x \in \mathcal{X}^K \), we have

\[
W^\rho_{m+1}(n, x) = \inf_{\pi^{k+1}} \mathbb{E}^{(\rho^k_{m+1}, \pi^{k+1})} \left[ \sum_{t=0}^{T_0} N_t | (n, x) \right]
\]

\[
\leq \mathbb{E}^{(\rho_m)} \left[ \sum_{t=0}^{T_0} N_t | (n, x) \right]
\]

\[
= W^\rho_{m}(n, x),
\]

which completes the proof.

Recall that the optimal discharge policy \( \pi^* = (\psi^*) \) that solves (2.2.2) should consist
Chapter 4. Discharge Control in the Multiple Bed Setting

Lemma 4.3.2. The optimal value function $V$ defined in (2.2.2) is a lower bound of the marginal value functions, that is, for any $\rho_1$, $V \leq W_{\rho_1}^m$ for all $m \geq 1$. In addition, $\pi^*$ is a fixed point of our policy improvement procedure, that is, for any $m$, $\rho_{m+1} = \pi^*$ if $\rho_m = \pi^*$.

Proof. The first result directly follows from the optimality of $\pi^*$, for any $n \in \mathbb{N}, x \in \mathcal{X}^K$, $\pi^*$ is a fixed point policy, WLOG, suppose there exists a $k$ such that $\rho_{m+1}^k \neq \psi^*$ given $\rho_m^k = \psi^*$. Then by the monotonicity of $W_{\rho_1}^m$ in $m$ and $W^\pi^* = V$, we have $W_{\rho_1}^{\pi^*} = W_{\pi^*}^m = V$, that is, the $\rho_{m+1} \neq \pi^*$ achieves the optimality. This contradicts that the optimal policy must be symmetric, which requires $\rho_{m+1}^k = \rho_m^k = \psi^*$ so that $\rho_{m+1} = \pi^*$.

Lemma 4.3.3. For any $\rho_1$, there exists a $W_{\rho_1}^\infty$ such that $W_{\rho_1}^m \to W_{\rho_1}^\infty$ as $m \to \infty$, and the limiting function $W_{\rho_1}^\infty$ is a symmetric function in $x$, that is, $W_{\rho_1}^\infty$ is invariant under a permutation of $x^k$, that is, for any $j,k$, $W_{\rho_1}^\infty(x^k;n,x^{k-}) = W_{\rho_1}^\infty(x^j;n,x^{j-}) = W_{\rho_1}^\infty(n,x)$.

Proof. By the two previous lemmas, for any starting policy $\rho_1$, $W_{\rho_1}^m$ is monotonically decreasing in $m$ and bounded below by the optimal value $V$ for all $m$. Therefore, the convergence result follows from the monotone convergence theorem. The symmetry property trivially follows from the fact that $W_{\rho_1}^m(x^k;n,x^{k-}) = W_{\rho_1}^{m+1}(x^j;n,x^{j-}) = W_{\rho_1}^\infty(n,x)$ for $j = k + 1 \mod K$ in the limit.
4.3.2 Optimality of the Limit Policy

**Theorem 4.3.4.** For any $\rho_1$, $W_{m}^{\rho_1} \to V$ as $m \to \infty$. In other words, for all $n \in \mathbb{N}, x \in \mathcal{X}^K$, $W_{\infty}^{\rho_1}(n, x) = V(n, x)$. Therefore, the limiting discharge policy generated by the MPI algorithm is joint optimal for the problem defined in (2.2.2).

**Proof.** By Lemma (4.3.3), for any $\rho_1$, there exists a limiting function $W_{\infty}^{\rho_1}$ that is symmetric in $x$. It follows from the construction of algorithm that the $W_{\infty}^{\rho_1}$ is uniquely determined by a system of governing equations given by, for all $n \in \mathbb{N}$,

\[
\begin{cases}
W_{\infty}^{\rho_1}(x^1; n, x^1) = n + \min_{d^2} \{ \mathbb{E} \left[ W_{\infty}^{\rho_1} \left( X_1^1; N_1, X_1^1 \right) | (d^1, \rho_{\infty}^1) \right] \} \\
\quad \quad \quad \quad \quad \quad \quad \quad \quad \ldots \\
W_{\infty}^{\rho_1}(x^K; n, x^K) = n + \min_{d^K} \{ \mathbb{E} \left[ W_{\infty}^{\rho_1} \left( X_1^K; N_1, X_1^K \right) | (d^K, \rho_{\infty}^K) \right] \},
\end{cases}
(4.3.4)
\]

where the fixed policy $\rho_{\infty}^k$ in each equation is the argument of the minimum operator in the other equation.

For any $\rho_1$, $W_{\infty}^{\rho_1}(n, x)$ is a symmetric function in $x$, so the expected cost-to-go that needs to be minimized is the same for all the equations defined in (4.3.4), that is,

\[
\mathbb{E} \left[ W_{\infty}^{\rho_1} \left( X_1^j; N_1, X_1^j \right) | (d^1, ..., d^K) \right] = \mathbb{E} \left[ W_{\infty}^{\rho_1} \left( X_1^k; N_1, X_1^k \right) | (d^1, ..., d^K) \right],
(4.3.5)
\]

for all $j, k \in \{1, ..., K\}$.

It follows that for all $x \in \mathcal{X}^K$ and $n \in \mathbb{N}$, solving for $W_{\infty}^{\rho_1}$ in the system of equations in (4.3.4) is equivalent to doing so in the following dynamic programming equation:

\[
W_{\infty}^{\rho_1}(n, x) = n + \min_{d^{(i)}} \left\{ \cdots \left\{ \min_{d^K} \mathbb{E} \left[ W_{\infty} \left( X_1^1, X_1^2 \right) | (d^1, ..., d^K) \right] \right\} \cdots \right\}. 
(4.3.6)
\]

However, as the order of minimization does not matter by symmetry, a function satisfies (4.3.6) if and only if it satisfies the dynamic programming equation given by

\[
W_{\infty}^{\rho_1}(n, x) = n + \min_{d} \left\{ \mathbb{E} \left[ W_{\infty} \left( N_1, X_1^1, X_1^2 \right) | (d^1, ..., d^K) \right] \right\},
(4.3.7)
\]
which is identical to the DP equations associated with the original problem $K$-bed problem defined in (2.2.2).

Then, by the unique fixed point property of the Bellman operator, $W_{\infty}^{\rho_1} \equiv V$ for any $\rho_1$. Thus, the MPI algorithm will eventually generate discharge policies that are jointly optimal, independent of the initial policy $\rho_1$.

Theorem 4.3.4 shows that the limiting policy obtained using the MPI algorithm achieves joint optimality. As a practical procedure, this marginal improvement technique only requires us to solve a sequence of one-bed optimization problems. We next discuss the use and performance of the MPI algorithm in a numerical example.
4.4 Numerical Example

For computational tractability, we consider a discharge control problem in which there are only two beds with at most one arrival in each period, that is, $K = 2$ and $A_t \in \{0, 1\}$ with probability $\lambda = 0.1$. The transition matrix $P$, readmission distribution $Q$ and arrival distribution $\beta$ are the same ones that were considered in the one-bed numerical example in Chapter 3. That is, the treatment process is governed by the following lower triangular transition matrix

$$
P = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.58 & 0.42 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.55 & 0.40 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.52 & 0.38 & 0.05 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.47 & 0.34 & 0.04 & 0.05 & 0.1 & 0 & 0 & 0 & 0 & 0 \\
0.42 & 0.30 & 0.04 & 0.05 & 0.08 & 0.11 & 0 & 0 & 0 & 0 \\
0.36 & 0.26 & 0.03 & 0.05 & 0.06 & 0.10 & 0.14 & 0 & 0 & 0 \\
0.28 & 0.21 & 0.02 & 0.04 & 0.05 & 0.08 & 0.10 & 0.22 & 0 & 0 \\
0.20 & 0.14 & 0.015 & 0.024 & 0.036 & 0.054 & 0.086 & 0.145 & 0.30 & 0 \\
0.10 & 0.07 & 0.008 & 0.013 & 0.018 & 0.028 & 0.045 & 0.070 & 0.168 & 0.48
\end{pmatrix},
$$

with a new arrival distribution

$$
\beta = \begin{pmatrix}
0 & 0.10 & 0.17 & 0.03 & 0.18 & 0.16 & 0.02 & 0.16 & 0.02 & 0.16
\end{pmatrix}.
$$

We assume there is at most one arrival per period with probability $\lambda = 0.1$ and the readmission probability is given by

$$
Q = \begin{pmatrix}
0 & 0 & 0.14 & 0.28 & 0.41 & 0.53 & 0.65 & 0.77 & 0.89 & 1.00
\end{pmatrix}.
$$

We implemented the MPI procedure in MATLAB to solve for the joint optimal discharge policy. At each iteration, similar to the one-bed problem, we used standard value
iteration methods to solve the dynamic programming equations. In our two bed setting, the computation converged within 6 iterations, which took 231 seconds to complete on a Intel i7 CPU at 3.40 GHZ with 16.0GB RAM. An example of the function $G_n(x^1; x^2)$ is plotted in Fig. 4.4.1.

Figure 4.4.1: The graph of the difference function $G_n(x^1; x^2)$ is plotted for $n = 2$. The threshold values are marked by filled circles.

At every decision time $t$, given state $(2, x^1; x^2)$, the MPI algorithm tells us to discharge a patient with state $x^1$ whenever $G_n(x^1; x^2) \geq 0$. We note that the threshold values depend on the health state of the other patient $x^2$. Note, this dependence does not appear in the myopic policy MYP or the one-bed one threshold policy OTP. We next compare the performance of these policies by simulating the total system size under different discharge policies. The myopic policy is simply the MYP described in section 3.1.1. The one-bed optimal policy is generated with arrival probability $\lambda_2$. To facilitate our discussion, we refer to the one-bed optimal policy as OTP and we denote the expected total system size using different policies as $V_{OTP}, V_{MYP}$ and $V_{MPI}$, respectively. The simulation results are provided in Table 4.4.1.

Table 4.4.1 shows that the MPI policy outperforms both the MYP and the OTP. Also,
the OTP is always significantly better than the MYP. At the same time, the benefits of using the MPI procedure comparing with the OTP becomes marginal when the starting system size is large. Combining this with its computational efficiency, the OTP is a good suboptimal discharge policy that can be suitably applied in some cases. The above numerical results conclude our discussion on the discharge control in a multi-bed hospital unit.

### 4.5 Summary

In this chapter, we investigated the multi-bed discharge control which requires solving a joint optimization problem due to the shared queue. We demonstrated that this problem is generally challenging as it cannot be fully decomposed to independent marginal problems. We proposed a policy improvement algorithm MPI that uses the values derived from a sequence of marginal optimization problems to approximate the joint problem. We proved that the MPI algorithm generates a limiting policy that is jointly optimal.
Chapter 5

Conclusion

In this thesis, we considered optimal discharge control with readmission in a hospital unit. The problem was modeled as a stochastic dynamic program and the objective was to determine the optimal discharge policy that minimizes the total expected system size until the system is empty. In the one-bed setting, we fully characterized the optimal policy as a two-threshold policy. In the general multi-bed setting, we proposed an efficient policy iteration algorithm, the MPI algorithm, and proved it converge to the jointly optimal policy.

5.1 Future Work

As our main motivation was to provide a basic stylized model to stochastic discharge control with readmission, there are many possible model extensions and modifications that can be considered for future research. In particular, modeling choices and parameter constraints assumed in Chapter 2 can be generalized to obtain a more realistic model. An example would be considering a more general state distribution of readmitted patients.

The results obtained in Chapter 3 and Chapter 4 certainly provide a initial description of the structure the optimal discharge policy. Still, developing efficient heuristic policies with good performance guaranteed may be an interesting future research topic. This is especially appealing in settings with high-dimensional system state spaces. Fur-
thermore, in the multi-bed setting, the marginal policy improvement (MPI) technique can certainly be applied to a more general class of joint optimization problems. One may further investigate the boundary of problems within which the technique yields an optimal solution.

From the numerical results provided in Chapter 4, we observe that applying a discharge policy that is optimal in the one-bed setting generally performs well in the multi-bed problem. This motivates us to further develop approximation methods for more general MDPs with nearly-decomposable structures that are patterned after our problem.
Bibliography


[10] Canada Institute for Health Information. “All-Cause Readmission to Acute Care and Return to the Emergency Department”. 2012. CIHI.


