

**DATA-DRIVEN RECONFIGURABLE SUPPLY CHAIN DESIGN AND  
INVENTORY CONTROL**

A Dissertation  
Presented to  
The Academic Faculty

By

Satya S. Malladi

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy in the  
School of Industrial and Systems Engineering

Georgia Institute of Technology

August 2018

Copyright © Satya S. Malladi 2018

# **DATA-DRIVEN RECONFIGURABLE SUPPLY CHAIN DESIGN AND INVENTORY CONTROL**

Approved by:

Professor Alan L. Erera, Advisor  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Professor Chelsea C. White III,  
Advisor  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Professor Alejandro Toriello  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Professor Enlu Zhou  
School of Industrial and Systems  
Engineering  
*Georgia Institute of Technology*

Professor Soumen Ghosh  
Scheller College of Business  
*Georgia Institute of Technology*

Professor David A. Goldberg  
School of Operations Research and  
Information Engineering  
*Cornell University*

Date Approved: May 3, 2018

Arise, awake, and stop not till the goal is reached.

*Katha Upanishad and Lectures of Swami Vivekananda*

To my parents, who inspire and motivate me constantly.

## ACKNOWLEDGEMENTS

My PhD would not have taken shape if it is not for the optimal learning environment provided by my amazing advisors. Both my advisors, Prof. Alan Erera and Prof. Chip White, have always been pillars of immense support to me and have over and over taken the risk of trusting my abilities and ideas. I am thankful to Prof. Erera for always giving me honest feedback and forthright advice on every topic to enable my growth. As his protégé, I have come to realize the quintessence of quality and practicality in academic research. I am grateful to Prof. White for his profound insight and foresight to ensure effective progress. Under his guidance, I have learned to appreciate the importance of the credibility that theoretical foundations render to practical research. I feel fortunate to have found my role models in my advisors, both of whom are very approachable.

I wish to acknowledge the support of the members of my dissertation committee to my academic pursuits. I thank Prof. Alejandro Toriello for his constant encouragement and positivity, both as a teacher as well as a committee member. I express my gratitude to Prof. David Goldberg for providing valuable advice on the impactful positioning of my research. I appreciate Prof. Enlu Zhou's wonderful mentorship to help build my academic career effectively. I thank Prof. Soumen Ghosh for enriching my business perspective of logistics research.

As a graduate student, I have learned extensively through numerous courses offered by Industrial and Systems Engineering (ISyE) and other departments. I find that classroom learning has been refreshingly beneficial to me and appreciate getting the opportunity to learn from the great teachers at Tech. I would like to mention the pleasant collegiality I have shared with Dr. Yanling Chang, Dr. Ilke Bakir, Dr. Mathias Klapp, Dr. Erin Garcia, Caglar Caglayan, Dr. Can Zhang, Dr. Amy Musselman, Dr. Weihong Hu, Dr. Luke Marshall, and Jan Vlachy. Whenever the going got tough, I have always leaned on a very understanding network of friends. My dearest friend, Samaneh Ebrahimi, has always been available to

share my joys and troubles. I cannot thank her enough for being such a great friend. I thank Yasaman Shahi for being the lively addition to my day on many days. I have enjoyed my time in the "Big Potatoes" ISyE gang along with Geet Lahoti, Dr. Chitta Ranjan, Samaneh and Yasaman. I will always cherish the numerous game nights, sunny hikes, and memorable conversations I have engaged in with fellow graduate student friends, namely, Janani, Preethi, Lakshmi, Usha, Balachandra, Swarnava, Prasanth, Trilochan, Vigna, Ravi, and Vaidhi. I have found that there is no better stress-buster than the camaraderie we have.

I am deeply thankful to my parents for ingraining in me the courage to pursue knowledge. My mother, Dr. Varalakshmi Malladi, has demonstrated to me that with determination and commitment to a goal, one can scale new heights everyday. My father, Mr. Swamy MBVLN, has instilled in me a deep respect for learning and a passion to excel. They have always picked me up in my lows and healed me with love. My late grandparents, Mr. Kameswara Rao Garimella and Mrs. Satyavathi Garimella, had always dreamed of a bright future for me. I am thankful to have grown up in the presence of their optimism and discipline. I thank my brother, Dr. Sriram Malladi, for not only providing a constant anchor of stability to me over the last five years, but also giving me clear-sighted advice on difficult career decisions. I acknowledge the unique motivation I have received from my relatives and in-laws, including my aunts, my parents-in-law, Harika, and Anupama. I specifically thank my aunt, Dr. Padmavathi Garimella, for always encouraging me and setting an example to emulate.

Finally, and most importantly, my partner-in-PhD-and-life, my husband Dr. Phanisri Pradeep Pratapa, has always rooted for me in the PhD game. I am thankful to him for his steadfast support towards my career ambitions.

Above all, I am deeply grateful to Goddess Mother Durga for bestowing upon me the courage, the capability, and the opportunity to pursue PhD in Industrial Engineering at Georgia Tech.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS</b> . . . . .	v
<b>LIST OF TABLES</b> . . . . .	xii
<b>LIST OF FIGURES</b> . . . . .	xiv
<b>Chapter 1: INTRODUCTION AND BACKGROUND</b> . . . . .	1
<b>Chapter 2: A DYNAMIC MOBILE PRODUCTION CAPACITY AND INVENTORY CONTROL PROBLEM</b> . . . . .	9
2.1 Introduction . . . . .	9
2.2 Problem Description . . . . .	11
2.2.1 Formulation . . . . .	13
2.3 Related Literature . . . . .	15
2.4 Bounds on the Optimal Expected Cost Function . . . . .	17
2.4.1 Perfect Information Relaxation: Lower Bound (PIR) . . . . .	17
2.4.2 Most Flexible System: Lower Bound (LB) . . . . .	18
2.4.3 Fixed Capacity System: Upper Bound (UB) . . . . .	19
2.5 Heuristics . . . . .	21
2.5.1 Myopic Policy (MP) . . . . .	22
2.5.2 Rollout of Fixed Future (RF) . . . . .	23

2.5.3	Lookahead with Approximate Fixed Future (LAF)	26
2.5.4	$L = 2$ Heuristics	28
2.6	Computational Study	30
2.6.1	Instance Set 1	31
2.6.2	Instance Set 2	37
2.6.3	Instance Set 3	39
2.7	Conclusion	43

**Chapter 3: INVENTORY CONTROL WITH MODULATED DEMAND AND  
A PARTIALLY OBSERVED MODULATION PROCESS . . . . . 44**

3.1	Introduction	44
3.1.1	Literature Review	45
3.2	Problem Description and Preliminary Results	48
3.2.1	Problem Definition	49
3.2.2	The POMDP Model and Preliminary Results	51
3.2.3	$L(x, y)$ Analysis	53
3.2.4	Definition of A1	57
3.3	When A1 holds	58
3.3.1	Main Result	58
3.3.2	A1 Analysis	58
3.3.3	Computing the Expected Cost Function, $v_n$	61
3.4	When A1 Does Not Hold	63
3.4.1	A Lower Bound, $v^L$	63
3.4.2	An Upper Bound, $v^U$	64



3.4.3	An Upper Bound on $v^U - v^L$ . . . . .	65
3.4.4	A Tighter Lower Bound, $v'$ . . . . .	67
3.5	Reorder Cost Case . . . . .	69
3.5.1	$K$ -convexity and Optimal $(s, S)$ Policies . . . . .	69
3.5.2	Bounds on $s_n$ and $S_n$ . . . . .	70
3.5.3	A Partition based on . . . . .	71
3.6	Conclusions . . . . .	72

**Chapter 4: DATA-DRIVEN CONTROL OF DISTRIBUTED RECONFIGURABLE  
PRODUCTION - INVENTORY SYSTEMS . . . . . 74**

4.1	Introduction . . . . .	74
4.1.1	Literature Review . . . . .	76
4.1.2	chapter Outline . . . . .	78
4.2	Problem Statement and Preliminary Results . . . . .	78
4.2.1	Problem Statement . . . . .	79
4.2.2	POMDP Model and General Results . . . . .	81
4.2.3	Complexity Analysis . . . . .	82
4.3	Bounds and Approximate Value Function Based on $L = 1$ Case . . . . .	83
4.4	Joint Control and Global-Local Control . . . . .	85
4.5	Solving the $L = 1$ Case . . . . .	88
4.5.1	A stationary approximation of $v_t^F$ . . . . .	88
4.5.2	A piecewise linear and convex approximation of $\hat{v}_t^F$ . . . . .	90
4.6	Heuristics . . . . .	90
4.6.1	Myopic Policy (MP) . . . . .	91

4.6.2	No Flexibility Policy (NF)	92
4.6.3	Joint Rollout of Stationary Future (JR)	92
4.6.4	Lookahead Strategy of Joint Controller (LAJ)	94
4.6.5	Global-Local Rollout of Stationary Future (GLR)	95
4.7	Computational Study and Results	96
4.7.1	Instance design	96
4.7.2	Results	97
4.7.3	Demand modeling: epoch-invariant vs. epoch-variant	99
4.7.4	Complete observability of modulation process	101
4.8	Conclusion	103
<b>Chapter 5: CONCLUSIONS AND FUTURE RESEARCH</b>		<b>104</b>
5.1	Summary of Results	104
5.2	Recommendations for Future Research	107
5.2.1	Resource Mobility and Sharing	107
5.2.2	Learning-based Demand Modeling	107
<b>Appendix A: Chapter 2</b>		<b>110</b>
A1	Proof of Proposition 1	110
A2	Proof of Proposition 2	111
A3	Algorithm for LIU	111
A4	Instance Sets for Computational Study	111
A4.1	Instance Set 1	112
A4.2	Instance Set 2	113

A4.3	Instance Set 3 . . . . .	113
A5	A Simple Strategy for Module Allocation . . . . .	114
A6	Additional Tables . . . . .	114
<b>Appendix B: Chapter 3 . . . . .</b>		<b>117</b>
B1	No Reorder Cost Case . . . . .	117
B2	Reorder Cost Case . . . . .	122
<b>Appendix C: Chapter 4 . . . . .</b>		<b>126</b>
C1	Analysis for the $L = 1$ Case . . . . .	126
C2	Proof of Proposition 1 . . . . .	134
C3	Lookahead with Global-Local Stationary Future (LAGL) . . . . .	136
C4	Results: Additional Tables . . . . .	136
C5	DM Assumes Epoch-invariant System . . . . .	138
C6	Complete Observability of Modulation Process . . . . .	139
<b>REFERENCES . . . . .</b>		<b>139</b>

## LIST OF TABLES

2.1	Instance Set 1 - Usefulness of mobile modularity . . . . .	33
2.2	Instance Set 1 - Variation of bounds w.r.t. OPT across $T$ , $Y$ , $b$ , $h$ , and $\bar{K}^M$ .	34
2.3	Instance Set 3 - Average computation time in seconds for $T = 15$ instances across $L$ . . . . .	40
2.4	Instance Set 3 - Performance of heuristics w.r.t. $UB_{\min}$ across $L$ and $Y$ . . .	43
4.1	Average performance of heuristics w.r.t. NF across $L$ for $\theta = 0.2$ . . . . .	98
4.2	Average computation times in seconds across $L$ . . . . .	98
4.3	Value of mobility (% savings over NF) using JR with $\theta = 0.2$ across $K^S$ and $K^M$ for $L = 10$ . . . . .	99
4.4	Value of non-stationarity w.r.t. NF across $L$ , when the DM assumes steady state, epoch-invariant demand distributions in a Markov-modulated system with partial observability . . . . .	100
4.5	Comparison of heuristic performance w.r.t. NF for $T = 30$ when the DM assumes epoch-invariant and epoch-variant demand distributions in a Markov-modulated system with partial observability . . . . .	101
4.6	Comparison of heuristic performance w.r.t. NF for $T = 30$ and $\hat{P} =$ $\{\{0.95, 0.05, 0\}, \{0.05, 0.9, 0.05\}, \{0, 0.05, 0.95\}\}$ when the DM assumes epoch-invariant and epoch-variant demand distributions in a Markov-modulated system with partial observability . . . . .	102
4.7	Additional value of mobility (% savings over NF) due to complete observ- ability using JR with $\theta = 0.2$ across $K^S$ and $K^M$ for $L = 10$ . . . . .	103

A.1	Instance Set 1 - Variation of bounds w.r.t. OPT for HH instances across $T$ , $Y$ , $b$ , $h$ , and $\bar{K}^M$ . . . . .	114
A.2	Instance Set 1 - Value addition of mobile modularity w.r.t. OPT across $Y$ , $T$ , $\bar{K}^M$ , $h$ , and $b$ . . . . .	115
A.3	Instance Set 1 - Performance of heuristics w.r.t. OPT on HH instances w.r.t. $b$ , $h$ , and $\bar{K}^M$ . . . . .	115
A.4	Instance Set 1 - Performance of heuristics w.r.t. OPT on all instances w.r.t. $b$ , $h$ , and $\bar{K}^M$ . . . . .	116
C.1	Average performance of GLR w.r.t. NF across $L$ and $\theta$ . . . . .	136
C.2	Average performance of LAJ w.r.t. NF across $L$ and $\theta$ . . . . .	137
C.3	Average performance of JR w.r.t. NF across $L$ and $\theta$ . . . . .	137
C.4	Average value of mobility (% savings over NF) using JR with $\theta = 0.2$ across $K^S$ and $K^M$ . . . . .	137
C.5	Average performance of heuristics w.r.t. NF across $L$ , when the DM assumes epoch-invariant and epoch-variant demand distributions in a Markov-modulated system with partial observability . . . . .	138
C.6	Value of mobility (% savings over NF) under complete observability using JR with $\theta = 0.2$ across $K^S$ and $K^M$ for $L = 10$ . . . . .	139

## LIST OF FIGURES

1.1	Mobile modular production-inventory system . . . . .	4
1.2	Demand-influencing, partially observed modulation state . . . . .	5
1.3	Mobile modular production-inventory system with transshipment . . . . .	6
2.1	Sequence of events within a period . . . . .	12
2.2	Module movement tendency of the optimal policy . . . . .	32
2.3	Effect of variability of demand on movement tendency . . . . .	33
2.4	Instance Set 1 - Performance of heuristics w.r.t. OPT across $T$ and $Y$ on HH instances and all instances . . . . .	35
2.5	Instance Set 2 - Module movement tendency across $G$ . . . . .	37
2.6	Instance Set 2 - Overall performance w.r.t. $UB_{\min}$ across $Y$ , $G$ , $\bar{K}^M$ , and $b$ .	38
2.7	Instance Set 2 - Performance w.r.t. $UB_{\min}$ across $T$ . . . . .	39
2.8	Instance Set 3 - Average computational effort relative to MP across $L$ . . . .	40
2.9	Instance Set 3 - Performance of heuristics w.r.t. $UB_{\min}$ , when initialized to $u_{\min}$ , across $L$ , $T$ , $\bar{K}^M$ , and $b$ . . . . .	41
2.10	Instance Set 3 - Performance of heuristics when initialized to $u_{\min}$ and $u_{\text{simple}}$	42
3.1	Example of $\mathcal{P}_1$ with $N = 3$ and $M = 7$ . . . . .	56
3.2	$\hat{\mathbf{x}}^{d_5, z}$ and $\{\lambda(d_5, z, \mathbf{x}) : \mathbf{x} \in X\}$ . . . . .	61
3.3	$v_2^U(\mathbf{x}, s)$ and $v_2^L(\mathbf{x})$ at $s = 13$ . . . . .	66

3.4	Example of partition of $X$ , with $N = 3$ and $M = 7$ . . . . .	72
4.1	Centralized and decentralized decision-making . . . . .	87
4.2	Underlying Markov chain of the modulation process in the instance set . . .	96

## SUMMARY

In this dissertation, we examine resource mobility in a supply chain that attempts to satisfy geographically distributed demand through resource sharing, where the resources can be inventory and manufacturing capacity. Our objective is to examine how resource mobility, coupled with data-driven analytics, can result in supply chains that without customer service level reduction blend the advantages of distributed production-inventory systems (e.g., fast fulfillment) and centralized systems (e.g., economies of scale, less total buffer inventory, and reduced capital expenditures). Transportability of production capacity allows a judicious overall investment due to risk pooling, in addition to better responsiveness to demands. Given this novel form of manufacturing flexibility, namely, mobility of production capacity, the value addition perceived operationally will guide the strategic decisions on investment in necessary infrastructure. Towards smart decision-making, it is also of importance to understand the power of effective information management through data-driven demand learning, capturing various fluctuations of demand, such as seasonal patterns, migratory tendencies, trends, etc.

In the first part of this dissertation, we introduce the problem of planning the logistics of a multi-location production-inventory system with transportable production capacity and propose near-optimal heuristic methodologies to effectively manage its response to uncertainty. A computational study of problems with stationary demand distributions, which should benefit least from mobile capacity, demonstrates the effectiveness of the suboptimal policies. For problems with twenty locations, the best heuristic solution cost provides 13% savings over a system with an optimal fixed capacity allocation. Greater savings result when the number of locations increases.

Next, we present an analysis of a single location inventory control problem with a partially observed, demand-influencing Markov-modulation process. We present an easy-to-compute newsvendor-styled criterion to determine the optimal myopic base stock policy



as a function of the belief of what the current modulation state is. We prove that this criterion linearly partitions the belief space into regions with unique optimal myopic base stock levels. We then prove that this policy is optimal for finite and infinite horizon problems when an inventory position attainability assumption holds.

Finally, we consider the problem of planning the logistics of a multi-location production-inventory system with the options of transportability of production capacity and transshipment of inventory, while facing demands influenced by a partially observed, Markov-modulation process. We propose efficient heuristic methods that result in cost savings as high as 26% on ten location instances over a system with no flexibility. Additionally, the computational performance of our heuristic approaches scales significantly well with the problem size. The results reinforce the value addition due to the production capacity transportability, independent of the option of transshipment flexibility.

# CHAPTER 1

## INTRODUCTION AND BACKGROUND

Logistics management is the set of planning and execution processes that target efficient and effective flow and storage of goods, services, and information between the point of origin and the point of consumption in a supply chain. Logistics systems should *respond to information* about uncertainty, in *anticipation* of new information. The quality of information and the responsiveness of the system determine its efficiency and effectiveness. Broadly, logistics efficiency and effectiveness may be achieved by two sets of approaches *information management and response management*. Improving the accuracy of demand estimation through better-informed demand models with greater predictive power over uncertainty, for example through data-driven learning approaches, falls under information management. Response management, on the other hand, includes approaches that enhance *supply flexibility* and design *optimized response* to information. Today's supply chains are grappling with many disruptive forces that shape information on the demand side, such as digitization and impatience of customers, increased competition in the market, and rapid technological adoption [1]. On the supply side also, disruptive technologies have been on the rise. Responses to the transforming demand patterns have evolved to be real-time, highly customer-centric, and on-demand. In particular, the mobility of production capacity is the latest breakthrough in the manufacturing industry. Major players in the pharmaceutical industry are developing manufacturing innovations such as transportable production facilities for synthesis of pharmaceutical outputs. Bayer demonstrated the production of fertilizer intermediates in a twenty-foot-equivalent-sized container [2]. Pfizer is in collaboration with Glaxo Smithkline to commercialize portable, miniature, and modular manufacturing technology for drug production [3]. Novartis developed a refrigerator-sized, on-demand pharmacy unit for fast tablet production [4]. E-commerce giant Amazon won a

patent for the logistics of mobile, additive manufacturing based fabrication and fulfillment [5, 6]. These developments signal a future where transportability of production facilities is commonplace.

The aspect of mobility of production resources has multiple advantages for logistics operations. First, production capacity can be shared across multiple locations that satisfy demand. Response is relatively faster due to on-demand resource flexibility. Second, some of the design-redundancy in the decentralized, no-flexibility system can be relaxed and fewer production resources may be sufficient for addressing shifting patterns of demands across locations. Third, due to the ease of modifying the stock of production resources on-demand over time, even poor initial production capacity allocation configurations would not hurt the system much; the logistics system will be quite resilient. Fourth, the mobility of production capacity and inventory enables customizability through manufacturing postponement in addition to rapid response to demand. The key to realizing these advantages lies in designing good solution methods that jointly manage production resource-sharing across locations and resource use at each location. The core challenge in designing an optimized response when mobility is incorporated into the logistics system is that managing dynamic production capacity allocation across multiple locations requires centralized decision-making that is inherently coupled with location-wise production capacity use. Location-wise production capacity use is determined by location-wise inventory management. The problem is further complicated by the ability to transship inventory in addition to transporting production capacity.

In this dissertation, we specifically concern ourselves with the transportability of production capacity in *multi-location production-inventory systems*. A multi-location production - inventory system is one where production and inventory control decisions are managed at multiple locations simultaneously. We specifically examine systems where production capacity can be relocated over time between locations, and also where inventory can be moved between locations when needed. With the goal of boosting logistics efficiency

in the transforming landscape of customer expectations and manufacturing innovations, we investigate the following problem settings in the domain of data-driven reconfigurable supply chain design and inventory control and provide solution methods.

- Chapter 2: We consider a multi-location production - inventory system with supply flexibility stemming from the mobility of production capacity between locations. To develop an understanding of the value of mobile production capacity in this setting, we build a Markov decision process (MDP) model for the problem of determining production and demand fulfillment decisions over a finite decision horizon. We develop heuristics by extending ideas from approximate dynamic programming to the problems of this type facing stationary and location-independent stochastic demand, test the heuristics, and obtain insights through an extensive computational study.
- Chapter 3: In order to better manage information, we develop a novel and generalized representation of demand uncertainty. We model demands modulated by an underlying partially observed Markov-modulation process. We study a single location inventory control system with a demand-influencing modulation process, of which the decision-maker is partially informed. We model the problem of determining an optimal replenishment policy over an infinite horizon as a partially observed MDP (POMDP). We present a newsvendor-styled criterion to determine an optimal myopic base stock policy and prove that when an attainability assumption holds, this myopic policy is optimal over the infinite horizon. We demonstrate the linear partition of the belief space induced by the proposed newsvendor criterion numerically.
- Chapter 4: We revisit the multi-location production-inventory system in Chapter 2, now with supply flexibility arising from the mobility of both production capacity and inventory between locations. We represent demand uncertainty using the partially observed Markov modulation process introduced in Chapter 3. This chapter synthesizes the frameworks of Chapters 2 and 3 with the goals of analyzing the useful-

ness of mobility of transportable capacity when inventory also can be transshipped and enriching the problem to capture demand volatility and temporal variation of uncertainty. We model the problem as a POMDP to determine two sets of decisions: a) inventory and production capacity relocation and b) replenishment at all locations. We propose approximate dynamic programming based heuristics based on two approaches, namely, joint control and global-local control. The former makes both sets of decisions jointly while the latter splits the control into making the global movement decisions in a centralized fashion first followed by the local replenishment decisions in a decentralized fashion.

The specific major contributions of this dissertation are now summarized.

1. The primary contributions of the investigation of the multi-location production - inventory system with production capacity transportability (see Figure 1.1) in Chapter 2 are:

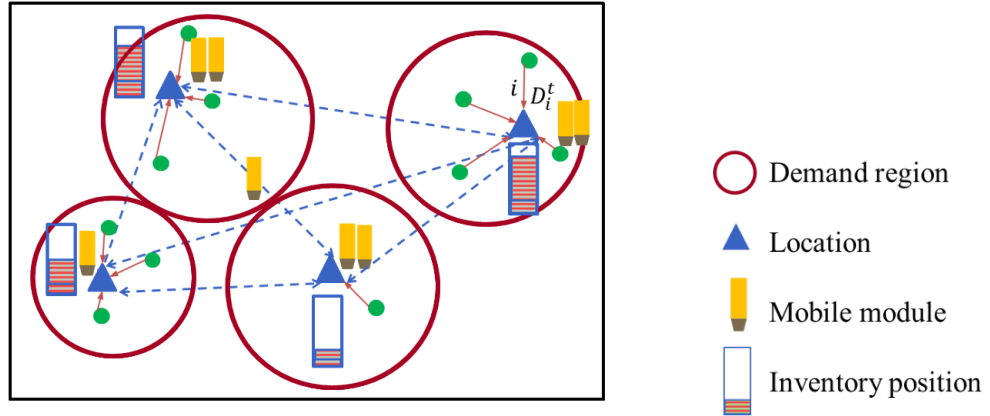


Figure 1.1: Mobile modular production-inventory system

- a formulation of the problem as an MDP model;
- the development of computationally efficient value function approximation-based heuristics that scale well with the number of locations; and
- a computational study that demonstrates the value of mobility of production capacity.

2. The primary contributions of the analysis of a single location inventory control problem with incomplete information about the underlying demand-influencing modulation process (see Figure 1.2) in Chapter 3 are:

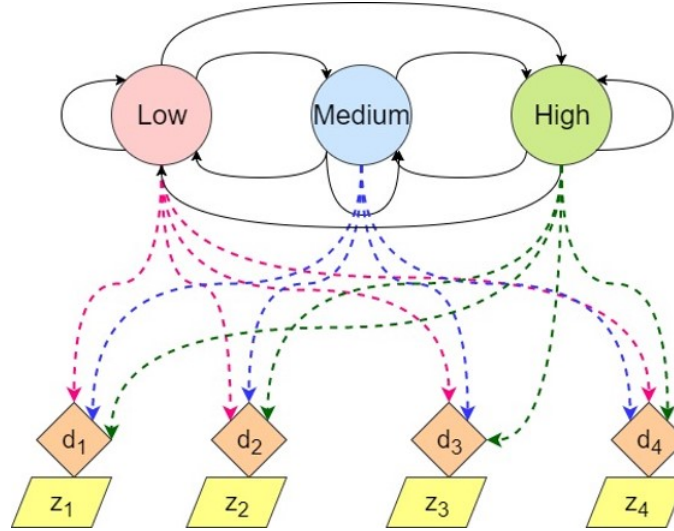


Figure 1.2: Demand-influencing, partially observed modulation state

- a formulation of the problem as a POMDP;
  - the development of an easy-to-compute newsvendor-styled criterion to determine an optimal myopic base stock policy as a function of the belief of the current modulation state; and
  - the proposition of the optimality of the myopic optimal base stock policy for finite and infinite horizon problems when an inventory position attainability assumption holds.
3. The primary contributions of the investigation of the multi-location production-inventory system with production capacity transportability, inventory transshipment, and a demand-influencing modulation process that is not completely observed (see Figure 1.3) in Chapter 4 are:

- a formulation of the problem as a POMDP;

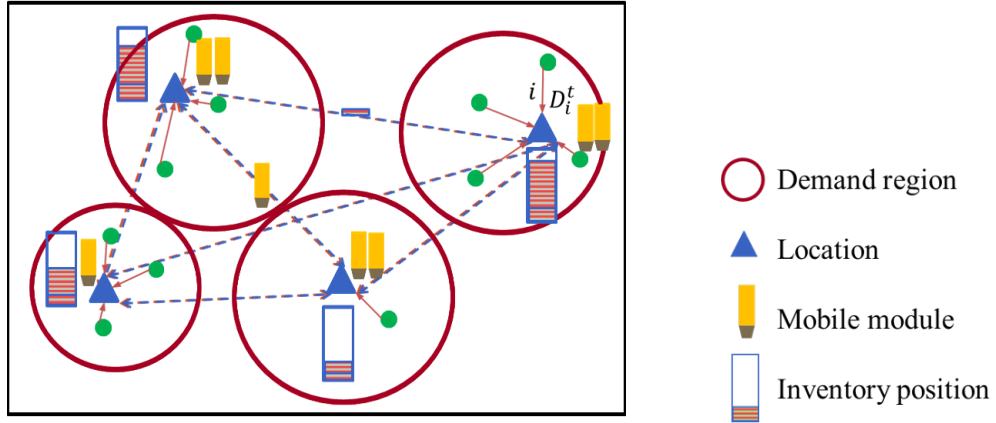


Figure 1.3: Mobile modular production-inventory system with transshipment

- the development of computationally efficient heuristics based on joint and global-local approaches, drawing the best features from centralized control and decentralized control; and
- a computational study that a) substantiates the value of mobility of multiple resources, namely production capacity and inventory, operated jointly as well as independently, b) corroborates the value of modeling a partially observed Markov modulation process, and c) appraises the value of complete observability of the modulation process.

Thus, we design an optimized response to uncertainty in a complex multi-location production-inventory system by managing known information in a data-driven learning environment. We now present our insights derived from the analysis performed in each of the three chapters.

1. Our takeaways from Chapter 2 are presented below.

- A value addition of about 13% is achieved computationally over the best no-flexibility system, using near-optimal heuristics when the number of locations is twenty, even under stationary demands.
- It is computationally inferred that the value addition increases with increase in

the number of locations, increase in the horizon length, and decrease in movement cost in the problem.

- Decomposition by locations results in a value function approximation that forms the backbone of a well-performing rollout heuristic.

2. Our insights from Chapter 3 are listed here.

- The proposed optimal myopic policy is a base stock policy with a newsvendor-styled criterion, allowing quick computation.
- The belief space is linearly partitioned into regions with unique optimal myopic base stock levels by the proposed optimality criterion.
- The linear partition thus obtained is independent of the demand outcome values.
- The proposed optimal policy is computationally significantly superior to existing optimal solution search methods.

3. We present our takeaways from Chapter 4 here.

- It is computationally ascertained that both transportability of product capacity and inventory transshipment capability lead to significant value addition over a system with no mobility when acting independently. The value addition observed when both forms of flexibility are active is higher than that with any one of the two forms of flexibility.
- Around 26% savings are observed over the no-flexibility system for the considered instances with ten locations, when both forms of flexibility are available.
- When the number of locations is ten, if transshipment is the only available flexibility, the independent savings over a no-flexibility system amount to 18% and if only transportability of production capacity at the same unit cost is allowed, independent savings of about 22% emerge.



- It is determined that greater savings are created as the number of locations and length of the horizon increase and as movement costs decrease.
- A significant additional value addition (6%) is observed by modeling demands in the proposed dynamic data-driven learning fashion instead of an aggregate static demand distribution.
- The value of the modeling dynamically updated demand distributions is significantly higher computationally for modulation transition dynamics with higher staying probabilities than leaving probabilities for all modulation states.
- Complete observability of the modulation process improves the savings over a no-flexibility system by a significant amount of 6 – 26% computationally on instances with ten locations.

## **CHAPTER 2**

### **A DYNAMIC MOBILE PRODUCTION CAPACITY AND INVENTORY CONTROL PROBLEM**

#### **2.1 Introduction**

Mobile manufacturing capacity is an emerging innovation in the chemical process industry and in the additive manufacturing industry. E-commerce giant Amazon has recently filed a patent for mobile additive manufacturing in a make-to-order setting [5]. Bayer, a pharmaceutical and agricultural chemicals company, has developed a containerized production unit that operates an intensified continuous batch process for fertilizer production, and has shown that mobile production units require lower setup costs than fixed facilities [2]. Pfizer is conducting a large-scale collaborative research project on miniaturized modular production technology for oral drugs [3]. Novartis has developed a refrigerator-sized on-demand pharmacy that can produce common drugs [4]. Supply chain systems that rely on mobile and modular production capacity may have many benefits, including:

- capacity sharing (via production module movement) may allow a smaller total capacity investment;
- perishable products and time-sensitive demand may be better served by producing locally;
- demand variation over time may be better accommodated by relocating capacity inexpensively; and
- new markets for products can be tested with recoverable, transportable production modules.

To develop a better understanding of supply chain systems that rely on mobile and modular production capacity, this chapter explores a dynamic production-inventory planning

problem in this emerging context. Suppose that product demands arise in a number of locations, and that each location can host one or more transportable production units referred to hereafter as *modules*. At any given time, a fleet of production modules is available and is deployed across the system. The module counts at each location at any given time will be referred to as the capacity configuration, and this configuration can be altered by moving modules between locations. The mobile modular production and inventory problem (MMPIP), introduced in this chapter, seeks to determine an optimal policy for managing capacity configurations, production, and inventory over a finite planning horizon to serve uncertain demands. In this problem setting, in addition to the typical trade-off between inventory and shortage costs, there exists a trade-off between the cost of relocating modules and the typical inventory costs.

We define an initial base problem in this study, and for simplicity we assume that demand in one location cannot be satisfied by production in another location; inventory transshipment between locations is also not modeled. This modeling choice was made to focus squarely on capacity movement. There also may be some systems where production needs to be local for a variety of reasons. We also focus this initial work on a typical make-to-stock inventory system, and note that in one of the important applications in chemical manufacturing that holding inventory is typical and that the time required for production is significantly larger than the time taken for demand fulfillment. Finally, we will study these systems in cases where location-wise demands are stationary and independent. Such settings should provide the least value for mobile production capacity, but are still useful to demonstrate the primary modeling ideas.

Successful operation of a production system with transportable, modular capacity also depends on other important considerations that will not be specifically addressed in this chapter. For example, effective inbound logistics systems for the inputs to production must be available in each location, and must be rapidly scalable with allocated production capacity. Similarly, outbound distribution systems must also accommodate potentially changing

production rates. Efficient movement and rapid setup and breakdown of the mobile production modules is also necessary for such systems to be effective. These considerations will be considered out of the scope of this study.

We organize this chapter into the following sections. Section 2.2 provides a formal problem definition and a formulation using a Markov decision process (MDP), and Section 2.3 places the research in the context of related literature. Section 2.4 presents bounds on the optimal cost function. In Section 2.5, we propose various heuristics for finding sub-optimal model solutions. Section 2.6 provides a numerical study of computational experiments using the heuristics on three different sets of instances. We summarize our findings and conclude the chapter in Section 2.7.

## 2.2 Problem Description

Consider a production-inventory system facing uncertain demands at a set of  $L$  locations, operated over a finite time horizon. Each location can produce product and store inventory of product to meet its demand over time, with unmet demand backlogged. At any time, a fleet of homogeneous production modules is available and distributed across the system, with some units installed in place and the others moving between locations. Production capacity at each location is limited by the number of production modules presently installed there. The objective is to determine module movement, production, and inventory decisions over time to minimize total system costs.

To make these decisions, consider a decision model with a planning horizon  $T$  discretized into  $T - 1$  consecutive, equal-duration decision periods,  $\{1, \dots, T - 1\}$ . Each location  $i$  must satisfy or backlog demand  $D_i(t)$  during decision epoch  $t$ . Let  $Y(t)$  be the total number of production modules available in the fleet at time  $t$ , and suppose that each module can produce a maximum of  $G$  units of product per time period. Finally, let  $\alpha_{ij}$  be the time that a production module is unable to produce when moved from location  $i$  to  $j$ , measured in fractional periods. Then, in each period two types of primary decisions are

made: module movement decisions to relocate capacity between locations, and production decisions to use installed modules at each location. When combined with observed demand, production decisions imply a change in the inventory position at each location each period.

Our models assume the following sequence of events in each period, as depicted in Figure 2.1. First, module movement decisions are determined and executed, yielding a new capacity configuration. Next, given the current capacity configuration, production decisions are determined and executed, yielding a post-replenishment inventory state across all locations. Finally, uncertain demands are observed and the inventory state is updated after filling or backlogging demands. Costs are incurred for module movement, production, inventory holding, and demand backlogging. Module movement costs are modeled as separable functions of number of modules relocated between specific location pairs, while production and inventory costs are modeled as typical in the supply chain planning literature.

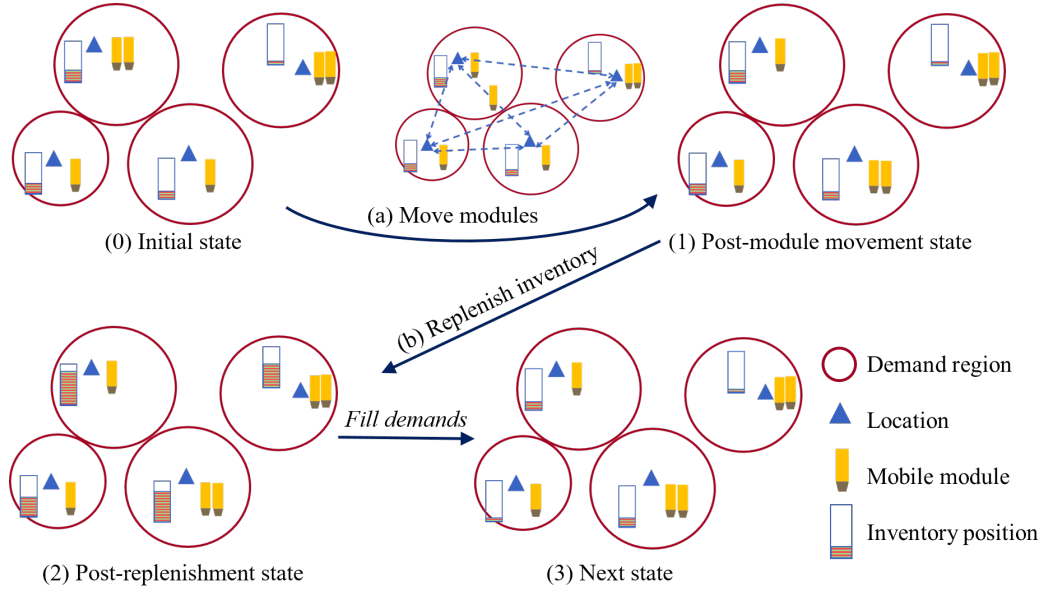


Figure 2.1: Sequence of events within a period

In this chapter, we will define a simple problem in this setting that we denote as the *mo-*

*bile modular production and inventory problem* (MMPIP). The MMPIP specifically models systems of the type introduced above with the following additional assumptions:

- Location demands are modeled as discrete random variables, and are independent and stationary across locations and time periods;
- Total available capacity is constant throughout the planning horizon, such that  $Y(t) = Y$  for all  $t$ ;
- Module movement costs are time-invariant, linear, and separable in the number of modules moved between location pairs;
- Module movements require no movement or setup time such that  $\alpha_{ij} = 0$  for all pairs of locations  $i$  and  $j$ , and thus modules moved during period  $t$  are immediately available for production at a new location in that period;
- Production costs are linear in the number of units produced, and per period inventory holding costs and backordering costs are linear in the number of units; and
- Production decisions executed during period  $t$  create new items available for immediate use in period  $t$ .

In the next subsection, we present a detailed formulation for the MMPIP.

### 2.2.1 Formulation

We formulate the MMPIP as a Markov decision process (MDP) for a finite horizon with  $T$  periods. Decisions are made in the first  $T - 1$  epochs,  $\mathcal{T} = \{1, \dots, T - 1\}$ , and epoch  $T$  models a horizon end state. At every decision epoch  $t \in \mathcal{T}$ , the multi-dimensional state variable is composed of  $2L$  components: the number of modules,  $u_i(t)$ , before modules are moved and the inventory position,  $s_i(t)$ , before production at each location  $i$ . The multi-dimensional action variable at  $t$  is composed of  $L^2$  components: the number of modules to move,  $\Delta_{ij}^M(t)$ , from location  $i$  to location  $j \neq i$  and the production quantity,  $q_i(t)$ , at each location  $i$ . Let  $D_i(t)$  be the discrete demand random variable for location  $i$  in period

$t$ , where stationarity implies that the probability mass function for  $D_i(t)$  is given by  $P_i$  for all time periods  $t$ .

Module movement decisions outbound from location  $i$  in period  $t$  are limited by the current number of modules location there,  $u_i(t)$ . The production decision at each location  $i$  in period  $t$  is limited by the capacity provided by the post-movement module state,  $u_i(t+1)$ . Note that this coupling of production decisions to module movement decisions defines the extension that this model proposes to single-location inventory control models.

Costs are incurred for actions in this system as follows. Module movement costs are linear in the number of modules moved between locations  $i$  and  $j$ , and given by  $K_{ij}^M \Delta_{ij}^M(t)$  in period  $t$ . Production and inventory costs are separable by location  $i$ , and follow the usual form found in base stock models. Production costs are linear, and given by  $c_i q_i(t)$  in period  $t$ . Holding costs accrue for each unit of positive inventory position at the end of period  $t$ ,  $h_i(s_i(t) + q_i(t) - D_i(t))^+$  where  $(y)^+ \equiv \max(y, 0)$ . Similarly, backorder costs accrue for each unit of negative inventory position, and are given by  $b_i(D_i(t) - s_i(t) - q_i(t))^+$ . We assume no costs associated with any state in the final horizon period  $T$ .

If we let  $\xi(t) = (\{u_i(t)\}, \{s_i(t)\})$  represent the complete state variable tuple, we can formulate the MDP as follows:

$$V_t(\xi(t)) = \min_{\substack{\Delta_{ij}^M(t): \\ \sum_j \Delta_{ij}^M(t) \leq u_i(t)}} \min_{\substack{\forall i q_i(t): \\ q_i(t) \leq Gu_i(t+1)}} \mathbb{E}_D \left[ \sum_i \left\{ \sum_j K_{ij}^M \Delta_{ij}^M(t) + c_i q_i(t) \right. \right. \\ \left. \left. + h_i(s_i(t) + q_i(t) - D_i(t))^+ + b_i(D_i(t) - s_i(t) - q_i(t))^+ \right. \right. \\ \left. \left. + V_{t+1}(\xi(t+1)) \right\} \right], \quad \forall \xi(t), \quad \forall t \in \mathcal{T} \quad (2.1)$$

where

$$\begin{aligned} \xi(t) &= (s_1(t), s_2(t), \dots, s_L(t), u_1(t), u_2(t), \dots, u_L(t)), \quad \forall t \in \mathcal{T} \\ s_i(t+1) &= s_i(t) + q_i(t) - D_i(t), \quad \forall i \in \{1, \dots, L\}, \quad \forall t \in \mathcal{T} \\ u_i(t+1) &= u_i(t) - \sum_j \Delta_{ij}^M(t) + \sum_k \Delta_{ki}^M(t), \quad \forall i \in \{1, \dots, L\}, \quad \forall t \in \mathcal{T} \\ \sum_i u_i(1) &= Y \\ V_T(\xi(T)) &= 0 \quad \forall \xi(T). \end{aligned} \quad (2.2)$$

where  $V_t(\xi(t))$  is the expected cost-to-go function of MMPIP from decision epoch  $t$  to the end of the horizon. Extending this formulation and the solution approaches presented in this chapter to include a per-period discount rate is straightforward.

The above MDP formulation leads to a state space whose cardinality is exponential in the number of locations,  $L$ . The space of possible movement decisions each period is similarly large, and thus finding exact optimal solutions to this model will not be possible except for the smallest instances. We will thus develop heuristic solution methods for identifying high-quality suboptimal designs.

### 2.3 Related Literature

The emergence of reconfigurable, mobile, decentralized/distributed manufacturing units has generated significant interest in the manufacturing/process industry in recent years [7, 8, 9, 10, 11, 12, 13, 14, 15, 16]. Reconfigurable or mobile modular production systems are characterized by *transformability*: scalability, adaptability (modularity, universality, compatibility), and mobility [7, 12]. [17] and [18] present mathematical models for make-to-stock and make-to-order scenarios for *hyperconnected* mobile production, which relies on multiple threads of innovation and is attained by the interconnectivity of all logistics systems.

The MMPIP models both dynamic capacity allocation and also joint capacity and inventory management across multiple locations. When demands are deterministic, dynamic modular capacity allocation and mobile capacity routing can be viewed as special cases of the dynamic facility location problem (DFLP). The DFLP is the problem of determining locations and opening schedules for multiple facilities (equivalently, units of capacity) over the planning horizon [19]. [12] model a special case of the MMPIP, a mobile modular production-inventory problem with deterministic demands, as a DFLP. A DFLP with modular capacities is presented in [19], which provides a good review of DFLP literature. Shifting a module in the MMPIP is equivalent to opening a facility at the new location and



closing one at the old location in the DFLP. [20] present a capacitated DFLP with multiple facilities in the same site for which [21] allow transfer of capacity between sites. The mobile facility routing problem with deterministic demands [22] is a DFLP with mobile facilities with duration-based capacity. A study comparing the performance of conventional production systems with mobile modular production systems can be found in [12], which assumes a binomial option pricing model of value of future cash flows for two possible “states of the nature”. The feature of managing mobile capacity and controlling inventory under uncertain demands over all the locations is missing in all the versions of DFLP that have appeared thus far to the best of our knowledge. In order to manage the mobile modular production-inventory system effectively, it is necessary to consider inventory and capacity management simultaneously. Hence, the MMPIP cannot be treated as a special case of the existing DFLP settings.

In the context of joint capacity and inventory decision-making, [23] study a make-to-stock production system with the ability to buy and sell capacity and prove that a target interval policy is optimal in two cases (with and without carryover of inventory). Simultaneously planning inventory actions and capacity change decisions at a single location is also studied in [24, 25]. [26] jointly plan the location and inventory of a single facility facing spatially distributed deterministic demand, in the context of managing a sea base, which is a collection of ships that serves as a military base at sea. [27] present results on the analysis of priority-based operating rules for a production inventory system in which two locations with low and high demand variabilities choose to pool their capacities.

We model the MMPIP using a Markov decision process (MDP). We assume that the set of production locations is given, which may be determined using the qualitative and quantitative approach suggested by [9]. The MMPIP for realistically-sized problems is computationally intractable. Hence, our focus is on finding good heuristics that rely on techniques that include: tractable bounds, value function approximations [28, Ch. 10], and rollout methods [29]. For a rollout heuristic to be effective, the computation of the guid-

ance mechanism must be tractable [30, 31, 32, 33]. We use the fixed future case as a guidance mechanism in the rollout heuristic we propose. We propose multiple well-performing strategies that rely on the proposed bounds, for solving the MMPIP. The lookahead with approximate fixed future performs particularly well even when the number of locations is large. In the next section, we present bounds on the optimal total expected cost function.

## 2.4 Bounds on the Optimal Expected Cost Function

In this section, we present two lower bounds for the optimal expected cost function, specified by (2.1) and (2.2). We denote them as the *perfect information relaxation* (PIR) and the *most flexible system* (LB) lower bounding functions respectively. Related to the most flexible system bound, we also develop an upper bound based on a fixed module configuration.

### 2.4.1 Perfect Information Relaxation: Lower Bound (PIR)

As is typical in dynamic programming, the structure of the recursion in (2.1) ensures that decisions made in time period  $t$  do not anticipate the outcomes of  $D_i(\tau)$  for  $\tau \geq t$ . A common approach for developing a lower bound for  $V_i(\xi(t))$  is to suppose that this is not the case, and to solve a deterministic planning problem for each possible demand trajectory defined by outcomes of  $(\{D_i(t)\}, \{D_i(t+1)\}, \dots, \{D_i(T-1)\})$ . A lower bound on the optimal expected cost is then given by the probability-weighted sum of the resultant optimal objective function for each demand trajectory. The most straightforward way to solve planning problems is to simply use value iteration for each trajectory to solve the deterministic variant of (2.1) and (2.2) with the expectation removed.

Since the number of demand trajectories may be very large for longer horizon problems with many production locations, this PIR bound can be approximated by instead computing an average over a Monte Carlo sample of trajectories. We note that bounds of this type (sometimes also called *a posteriori* bounds) are quite common for dynamic programming models, but we have found that they are weak for MMPIP problem instances. We note that

PIR bounds for problems in which capacity is fixed at locations in advance also tend to be weak. The weakness of PIR bounds for the MMPIP does lead to a natural conjecture: there may be significant value to move production modules in response to specific demand trajectories, and therefore solving MMPIP problems effectively may lead to significant value beyond systems with fixed installed capacity.

#### 2.4.2 Most Flexible System: Lower Bound (LB)

We now present a lower bound for the MMPIP that we have found to be much tighter in computational experiments. A mobile modular production system is highly flexible, since the capacity configuration can be altered during the time horizon. Module movement costs, however, mitigate the value of this flexibility. An approach to developing lower bounds is to assume that production capacity can be allocated to locations immediately in any period, and without module movement costs, in response to the current vector inventory state.

Since lower bounds of this type do not depend on the current capacity configuration, we specify a lower bounding function  $\tilde{V}_t^{\mathcal{L}}(\{s_i(t)\})$ , where

$$\tilde{V}_t^{\mathcal{L}}(\{s_i(t)\}) \leq V_t(\{s_i(t)\}, \{u_i(t)\}) \quad \forall \{u_i(t)\}.$$

We compute  $\tilde{V}_t^{\mathcal{L}}$  recursively using a simpler dynamic program, where we assume that the lowest-cost capacity configuration can be used for each possible inventory state, as follows:

$$\begin{aligned} \tilde{V}_t^{\mathcal{L}}(\{s_i(t)\}) &= \min_{\substack{\{u_i(t+1)\} \\ \sum_i u_i(t+1)=Y}} \min_{q_i(t) \leq G u_i(t+1) \forall i} \mathbb{E}_D \sum_{i \in \mathcal{I}} \left[ h_i(s_i(t) + q_i(t) - D_i(t))^+ \right. \\ &\quad \left. + b_i(D_i(t) - s_i(t) - q_i(t))^+ + \tilde{V}_{t+1}^{\mathcal{L}}(\{s_i(t) + q_i(t) - D_i(t)\}) \right], \\ &\quad \forall t \in \mathcal{T}, \\ \tilde{V}_T^{\mathcal{L}}(\{s_i(T)\}) &= 0 \quad \forall \{s_i(T)\}. \end{aligned}$$

Note that it remains computationally expensive to compute the bounding function  $\tilde{V}_t^{\mathcal{L}}$

in general. It is more difficult than solving  $L$  independent inventory management problems. Since the total number of modules defines total production capacity, it is necessary in most instances to decide which locations, given a current inventory state, should be prevented from selecting an optimal unconstrained production value by restricting their capacities. However, it requires significantly less computation than solving the MMPIP problem to optimality. Consider solving both problems by value iteration. The number of capacity configurations possible is  $\mathcal{O}((Y+L)^{L-1})$ , since the number of ways to allocate  $Y$  modules to  $L$  locations is given by  $\binom{Y+L-1}{L-1}$ . Thus, at each epoch, for a given vector inventory state, value iteration for MMPIP requires  $\mathcal{O}((Y+L)^{L-1})$  times more effort than solving for LB.

#### 2.4.3 Fixed Capacity System: Upper Bound (UB)

A conventional production system with stationary capacity at all locations can be viewed as a mobile modular production system where the capacity configuration is fixed for the planning horizon. Given a fixed capacity configuration, an upper bound on the optimal expected cost given an initial inventory state can be determined by solving  $L$  independent, constrained inventory management problems. More specifically, let  $\tilde{V}_t^F(\{s_i(t)\}, \{u_i\})$  be the optimal cost-to-go function of the multilocation fixed system with capacity configuration  $\{u_i\}$  fixed from  $t$  until the end of the planning horizon, given initial inventory position state  $\{s_i(t)\}$ . Let  $V_{i,t}^F(s, C)$  be the optimal cost-to-go function of the single location inventory control problem with production capacity  $C$  and initial inventory position  $s$  at location  $i$  at time epoch  $t$ . When there are  $u_i$  modules at  $i$ , then  $C = Gu_i$ . To determine this upper

bound, we use the following optimality equations:

$$\begin{aligned}
\tilde{V}_t^F(\{s_i(t)\}, \{u_i\}) &= \sum_{i=1}^L V_{i,t}^F(s_i(t), Gu_i) \\
&= \sum_{i=1}^L \min_{q_i(t) \leq Gu_i} \mathbb{E}_D \left[ h_i(s_i(t) + q_i(t) - D_i(t))^+ \right. \\
&\quad \left. + b_i(D_i(t) - s_i(t) - q_i(t))^+ \right. \\
&\quad \left. + V_{i,t}^F((s_i(t) + q_i(t) - D_i(t), Gu_i) \right] \quad \forall \{s_i(t)\}, \forall t \in \mathcal{T} \\
\tilde{V}_T^F(\{s_i(T)\}, \{u_i\}) &= 0 \quad \forall \{s_i(T)\}.
\end{aligned}$$

We note that the computation of  $\tilde{V}_t^F(\cdot, \{u_i\})$  can be decoupled across locations. Let  $\mathcal{Q}$  be the cardinality of the inventory position state space for each location. Given capacity configuration  $\{u_i\}$ , for each location at each epoch computing the cost function above requires  $\mathcal{O}(\mathcal{Q}Gu_i)$  steps, and thus the total computational effort at each epoch is  $\mathcal{O}(\mathcal{Q}GY)$ . Again, this effort is significantly smaller than the corresponding  $\mathcal{O}(\mathcal{Q}^L(Y + L)^{L-1}GY)$  effort required for computing the optimal value function for MMPIP at each epoch for a given capacity configuration.

If we use this upper bounding approach beginning at the initial time epoch 1, we can also compute the minimum ( $\text{UB}_{\min}$ ) and the maximum ( $\text{UB}_{\max}$ ) possible expected total cost of a fixed system, given an initial inventory state of zero at all locations. Define the capacity configuration that corresponds to  $\text{UB}_{\min}$  as  $u_{\min} = \arg \min_{\{u_i\}} \tilde{V}_1^F(\bar{0}, \{u_i\})$ . Determining  $\text{UB}_{\min}$  and  $u_{\min}$  requires comparing  $\mathcal{O}((Y + L)^{L-1})$  capacity configurations, or solving a multiple choice knapsack problem [34, 35] as presented below:

$$\begin{aligned}
&\min \sum_{i=1}^L \sum_{y=0}^Y V_{i,1}^F(0, Gy) z_{iy} \\
&\text{s.t. } \sum_{i=1}^L \sum_{y=0}^Y y z_{iy} = Y \\
&\quad \sum_{y=0}^Y z_{iy} = 1 \quad \forall i \in \{1, \dots, L\} \\
&\quad z_{iy} \in \{0, 1\} \quad \forall i \in \{1, \dots, L\}, y \in \{0, \dots, Y\}
\end{aligned}$$

The idea of the knapsack formulation is to add one item for each location, where the size of the item is its number of assigned modules and the knapsack size is  $Y$ . The cost of adding item  $iy$  to the knapsack is the expected cost-to-go at epoch 0 for location  $i$  with  $y$  modules. In the multiple choice problem, a set of constraints ensures that exactly one item is chosen from each class of items (location). We note that the computational effort to solve this integer program in practice is much smaller than the effort required to determine the cost-to-go values for all locations and capacity levels.

We will see later fixed system configurations beginning at some time period can play the role of a base heuristic within rollout approaches for determining good (but suboptimal) dynamic policies for the MMPIP problem. Additionally, fixed system bounds will be used as benchmarks to assess the value addition created by mobile modular production systems and to evaluate the performance of heuristics for the MMPIP.

## 2.5 Heuristics

Since the MMPIP is characterized by a state space whose size is exponential in the number of locations,  $L$ , and the length of the horizon,  $T$ , the model suffers from a curse of dimensionality. Hence, we seek suboptimal policies for the problem constructed using approximate dynamic programming techniques, such as rollout algorithms and decomposition-based approaches, that do not require complete characterization of the optimal expected cost function. Following the terminology presented in [29], a one-step rollout algorithm is a value function approximation approach in which the decision at the current epoch is determined by approximating the cost-to-go by the expected cost of implementing a base policy for states beginning in the next epoch. We propose a decomposition-based rollout algorithm, RF, that approximates the cost-to-go function by assuming that the capacity configuration does not change again after decisions made in current epoch, and that an optimal inventory control policy is used for future replenishment decisions given this fixed capacity. Since this rollout can still be computationally expensive, we also present an alternative one-

step lookahead policy, LAF, that approximates the expected cost of the optimal RF rollout. A one-step lookahead policy is a suboptimal policy obtained by minimizing the sum of the immediate cost for the current period and an approximation of the cost-to-go function for the remaining horizon [36, Chapter 6]. In addition to these two core heuristics, we also develop additional value function approximation policies that can be used for special case problem instances with  $L = 2$  locations.

### 2.5.1 Myopic Policy (MP)

In a myopic policy, we ignore the cost-to-go in the next epoch, and thus  $V_{t+1}(\{s_i(t+1)\}, \{u_i(t+1)\}) \approx 0$  for all system states. The myopic action in the current epoch can be found by solving an integer linear program (IP) that extends the classic formulation for the discrete demand distribution newsvendor model. Let the set of demand outcomes from stationary distribution  $P_i$  at every location  $i$  be  $\{d_i^k\}$ , where for notational convenience  $k$  indexes the outcomes in  $\mathcal{K}_i$ . Given the current state  $(\{s_i\}, \{u_i\})$  of module allocation and inventory positions, the IP formulation (2.3) is given by:

$$\begin{aligned}
\text{MMPIP-MP.} \quad \min \quad & \sum_{i=1}^L \left[ \sum_{j=1}^L K_{ij}^M \Delta_{ij}^M + \sum_{k=1}^M p_i^k (h_i r_i^k + b_i o_i^k) \right] \\
& r_i^k \geq s_i + q_i - d_i^k, \quad \forall k \in \mathcal{K}_i, \quad i \in \{1, \dots, L\} \\
& o_i^k \geq d_i^k - s_i - q_i, \quad \forall k \in \mathcal{K}_i, \quad i \in \{1, \dots, L\} \\
& 0 \leq q_i \leq G \left( u_i - \sum_j \Delta_{ij}^M + \sum_l \Delta_{li}^M \right) \quad \forall i \in \{1, \dots, L\} \\
& \sum_j \Delta_{ij}^M \leq u_i \quad \forall i \in \{1, \dots, L\} \\
& q_i, \Delta_{ij}^M \in \mathcal{Z}^+ \quad \forall i, j; \quad r_i^k, o_i^k \in \mathcal{Z}^+ \quad \forall k, \forall i.
\end{aligned} \tag{2.3}$$

The decision variables are the module movements  $\{\Delta_{ij}^M\}$ , the replenishment quantities  $\{q_i\}$ , and the positive and negative parts of post-decision, post-information inventory po-

sition  $\{r_i^k\}$  and  $\{o_i^k\}$  respectively. MMPIP-MP minimizes the immediate cost of module movement and inventory holding or backordering. The first two constraints characterize a newsvendor problem. The third constraint ensures that the post-module movement production capacity is not exceeded at each location. The last constraint prevents removing more modules than those available at any location.

The formulation for the single location discrete demand distribution newsvendor model is proved to be a linear program [37]. We now present analogous structural results on the integer program MMPIP-MP to enable increased computational efficiency.

**Theorem 1.** *For integer values of  $G$ ,  $d_i^k$ ,  $u_i$ , and  $s_i$  for all  $k \in \mathcal{K}_i$  and  $i \in \{1, \dots, L\}$ ,*

- (a) *the integrality constraints on  $\{o_i^k\}$ ,  $\{r_i^k\}$ , and  $\{q_i\}$  for all  $k \in \mathcal{K}_i$ ,  $i \in \{1, \dots, L\}$  in the integer program MMPIP-MP are redundant.*
- (b) *when capacity per module  $G = 1$ , the linear programming relaxation of the integer program MMPIP-MP has an integral optimal solution.*

Proof of Theorem 1 is provided in Section A1. Theorem 1(a) implies that it is sufficient to impose integrality constraints on the variables  $\{\Delta_{ij}^M\}$  only. Theorem 1(b) presents a condition, namely,  $G = 1$ , under which, all the integrality constraints are redundant and thus the MMPIP-MP can be solved by its linear programming relaxation. This implies fast compute times when  $G = 1$  even for relatively large values of  $L$  and  $T$ .

### 2.5.2 Rollout of Fixed Future (RF)

In this rollout heuristic, the base heuristic assumes that beginning in the next period modules will be fixed at their current locations until the end of the horizon. Thus, this approach approximates the flexible capacity production system with one in which flexibility is only available for the current period. We propose integer linear program (2.4) to determine the optimal one-step decisions for this rollout, given the current state  $(\{s_i\}, \{u_i\})$  of module



allocation and inventory positions. This problem is an extension of the integer program for the myopic single-period action selection problem.

In (2.4), the objective is to minimize movement cost, expected inventory cost at all locations, and the expected optimal fixed future cost. Note that since the capacity state is fixed beginning in the next period, it is possible to decompose the optimal cost-to-go function by location, and to only require the local production capacity and inventory state as inputs to the precomputed functions  $V_{i,t+1}^F(s, C)$ . Thus, in addition to the  $\{q_i\}$  and  $\{\{\Delta_{ij}^M\}\}$  decision variables used in the myopic integer program, binary variables  $z_i(\Delta^M, q)$  are specified that take value 1 if  $\Delta^M$  modules are transferred to location  $i$ , and then used in the current period to produce  $q$  items (note that  $q \leq G(u_i + \Delta^M)$  and that  $\Delta^M$  may be negative). The first two constraints again are used to compute single-period underage or overage units. The next three constraints ensure that only one set of module movement and production decisions is made for each location, and that the module movement variables

result in the selected capacity state for each location  $i$ . The formulation is provided here:

$$\begin{aligned}
\text{MMPIP-RF. } \min & \sum_{i=1}^L \left[ \sum_{j=1}^L K_{ij}^M \Delta_{ij}^M + \sum_{k=1}^M p_i^k \{h_i r_i^k + b_i o_i^k\} \right. \\
& + \sum_{i=1}^L \sum_{\Delta^M = -u_i}^{Y-u_i} \sum_{x=0}^{G(u_i + \Delta^M)} z_i(\Delta^M, q) \sum_k p_i^k V_{i,t+1}^F(s_i + q - d_i^k, G(u_i + \Delta^M)) \left. \right] \\
& r_i^k \geq s_i + q_i - d_i^k, \quad \forall k \in \mathcal{K}_i, i \in \{1, \dots, L\} \\
& o_i^k \geq d_i^k - s_i - q_i, \quad \forall k \in \mathcal{K}_i, i \in \{1, \dots, L\} \\
& \sum_{\Delta^M = -u_i}^{Y-u_i} \sum_{q=0}^{G(u_i + \Delta^M)} z_i(\Delta^M, q) = 1 \quad \forall i \in \{1, \dots, L\} \\
& - \sum_j \Delta_{ij}^M + \sum_l \Delta_{li}^M = \sum_{\Delta^M = -u_i}^{Y-u_i} \sum_{q=0}^{G(u_i + \Delta^M)} \Delta^M z_i(\Delta^M, q) \quad \forall i \in \{1, \dots, L\} \\
& q_i = \sum_{\Delta^M = -u_i}^{Y-u_i} \sum_{q=0}^{G(u_i + \Delta^M)} q z_i(\Delta^M, q) \quad \forall i \in \{1, \dots, L\} \\
& z_i(\Delta^M, q) \in \{0, 1\}, \quad \forall q \in \{0, \dots, G(u_i + \Delta^M)\}, \\
& \Delta^M \in \{-u_i, \dots, Y - u_i\}, \quad i \in \{1, \dots, L\} \\
& q_i, \Delta_{ij}^M \in \mathcal{Z}^+, \quad \forall i, j \in \{1, \dots, L\}; \quad r_i^k, o_i^k \in \mathcal{Z}^+ \quad \forall k \in \mathcal{K}_i, \forall i \in \{1, \dots, L\}.
\end{aligned} \tag{2.4}$$

It should be clear that the integer program (2.4) includes a large number of binary variables for larger values of  $L$ ,  $Y$ , and  $G$ ; the variable count grows at  $\mathcal{O}(GY^2L)$ . Furthermore, before the formulation can be used, it is necessary to compute the function lookup tables  $V_{i,t+1}^F(s, C)$  for each possible module change  $\Delta_i^M \in \{-u_i, \dots, Y - u_i\}$  and inventory state  $s$  at all locations using the approach described earlier. We also note that (2.4) is a potentially useful model when  $V_{i,t+1}^F(s, C)$  is an approximate value function, decoupled by location, developed using any alternative approach and not necessarily limited to the case where these functions represent the optimal cost-to-go of the single location fixed capacity problem.

### 2.5.3 Lookahead with Approximate Fixed Future (LAF)

The complete rollout heuristic RF can be computationally expensive since the integer program becomes difficult in practice for larger problems. We therefore now develop an approximation of RF that is more computationally tractable. To do so, we approximate the single location cost function by the average of two piecewise linear and convex functions. Doing so bypasses the computationally expensive cost function lookup table modeling required in the RF heuristic. In this method, at every epoch decisions are made under the assumption that the cost-to-go function is approximated by the following expression:

$$V_{t+1}(\{s_i(t+1)\}, \{u_i(t+1)\}) \approx (\tilde{V}_{t+1}^F(\{\bar{s}_i(t+1)\}, \{u_i(t)\}) + \tilde{V}_{t+1}^F(\{s_i(t)\}, \{u_i(t+1)\}))/2,$$

where  $\bar{s}_i(t+1) = s_i(t) + q_i(t) - \lceil E[D_i(t)] \rceil$  and  $\lceil a \rceil$  rounds  $a$  to the nearest integer.

We can model this cost approximation using the optimal cost-to-go function of the fixed system by leveraging the structural properties which establish that cost-to-go function of a capacitated single location inventory control problem is convex in inventory position for a fixed capacity level and convex in capacity level for a fixed inventory position [38]. This result implies that  $V_{i,t}^F(s_i, G_i)$  is piecewise linear (due to discrete inventory state space) and convex in  $s_i$  for a fixed  $G_i$ . Thus, it can be represented as  $\max\{\gamma_j^i s_i + \hat{\gamma}_j^i : (\gamma_j^i, \hat{\gamma}_j^i) \in \Gamma_t^i(u_i)\}, \forall i \in \{1, \dots, L\}$ . Similarly, for a fixed  $s_i$ , the function  $V_{i,t}^F(s_i, G_i)$  is piecewise linear and convex in  $u_i$  (since  $G_i = G u_i$ ) that can be expressed as  $\max\{\theta_j^i u_i + \hat{\theta}_j^i : (\theta_j^i, \hat{\theta}_j^i) \in \Theta_t^i(s_i)\}, \forall i \in \{1, \dots, L\}$ .

This approximation is again implemented by modifying the integer program (2.3), as

follows:

$$\begin{aligned}
\text{MMPIP-LAF. } \min \sum_{i=1}^L & \left[ \sum_{j=1}^L K_{ij}^M \Delta_{ij}^M + \sum_{k=1}^M p_i^k \{h_i r_i^k + b_i o_i^k\} + (\zeta_i + \eta_i)/2 \right] \\
\zeta_i & \geq \gamma_j^i (s_i + q_i - [E[D_i(t)]]) + \widehat{\gamma}_j^i, \forall (\gamma_j^i, \widehat{\gamma}_j^i) \in \Gamma_{t+1}^i(u_i), i \in \{1, \dots, L\} \\
\eta_i & \geq \theta_j^i y_i + \widehat{\theta}_j^i, \forall (\theta_j^i, \widehat{\theta}_j^i) \in \Theta_{t+1}^i(s_i), i \in \{1, \dots, L\} \\
r_i^k & \geq s_i + q_i - d_i^k, \forall k \in \mathcal{K}_i, i \in \{1, \dots, L\} \\
o_i^k & \geq d_i^k - s_i - q_i, \forall k \in \mathcal{K}_i, i \in \{1, \dots, L\} \\
0 & \leq q_i \leq G y_i \forall i \in \{1, \dots, L\} \\
u_i - \sum_j \Delta_{ij}^M + \sum_l \Delta_{li}^M & = y_i \forall i \in \{1, \dots, L\} \\
q_i, y_i, \Delta_{ij}^M & \in \mathcal{Z}^+, \forall i, j; r_i^k, o_i^k \in \mathcal{Z}^+ \forall k \in \mathcal{K}_i, \forall i \in \{1, \dots, L\}; \\
\eta_i, \zeta_i & \in \mathbb{R} \forall i \in \{1, \dots, L\}.
\end{aligned} \tag{2.5}$$

In addition to the decision variables described in the implementation of MP,  $\zeta_i$  and  $\eta_i$  represent the single location future costs at  $i$  expressed as a function of next period's inventory when capacity is held at the initial level of  $G u_i$ , and as a function of new capacity when inventory is held at the initial level of  $s_i$  respectively. The number of integer variables required to represent the future cost-to-go is significantly lower by  $\mathcal{O}(GY^2L)$  in MIP (2.5) when compared with IP (2.4). This reduces the computational effort required to solve this MIP dramatically. Furthermore, when module capacity  $G = 1$ , this MIP reduces to a linear program.

**Theorem 2.** *For integer values of  $G$ ,  $d_i^k$ ,  $u_i$ , and  $s_i$  for all  $k \in \mathcal{K}_i$  and  $i \in \{1, \dots, L\}$ , when capacity per module  $G = 1$ , the linear programming relaxation of the mixed integer program MMPIP-LAF has an optimal solution where decision variables  $q$ ,  $y$ ,  $w$ ,  $r$ , and  $o$  are integral.*

Proof of Theorem 2 is presented in Section A2.

#### 2.5.4 $L = 2$ Heuristics

The general suboptimal policies that we have proposed so far rely on approximating the future value function with a form that decomposes by location. Of course, if we assume that production modules cannot be moved after the current time epoch, then indeed the *optimal* value function decomposes by location. In this section, we explore suboptimal policies and solution heuristics that no longer have this feature to understand how much incremental value may be gained. For computational tractability, we focus on small problem instances with  $L = 2$  production locations.

##### *Lookahead with Fixed or Purchasable Most Flexible Future (LFP)*

The cost function computation of the lower bound LB presented in Section 2.4.2 is coupled across locations and has severe computational drawbacks for larger numbers  $L$  of locations. However, it is complementary to the fixed capacity upper bound and models the case where there is no cost to moving production modules; we call this case the most flexible future. We therefore investigate the potential of building a value function approximation that blends these upper and lower bounds within a suboptimal lookahead policy.

We denote the blending heuristic LFP, denoting a lookahead with a fixed future capacity or a “purchasable” most flexible future capacity. The LFP assumes that the decision-maker chooses the best action at the current decision epoch by assuming that she has a choice between (i) keeping capacity fixed for the remainder of the planning horizon and (ii) making a one-time payment now to allow unlimited future module movements. That is, the cost-to-go function in Eq. 2.1 is approximated as follows:

$$V_{t+1}(\{s_i(t+1)\}, \{u_i(t+1)\}) \approx \min \left\{ \tilde{V}_{t+1}^F(\{s_i(t+1)\}, \{u_i(t+1)\}), \right. \\ \left. \kappa^F(t) + \tilde{V}_{t+1}^{\mathcal{L}}(\{s_i(t+1)\}) \right\},$$

where  $\tilde{V}_{t+1}^F$  as usual can be decomposed into  $\sum_{i=1}^L V_{i,t+1}^F(s_i(t+1), u_i(t+1))$ .

To use this approach, it is necessary to define the mobility purchase cost  $\kappa^F(t)$ . We use the following definition, which works well in practice:

$$\kappa^F(t) = \frac{1}{2} \bar{K}^M (T - t) P(\{D_1 = d_1^{\max}\} \cup \{D_2 = d_2^{\max}\}), \quad \forall t \in \mathcal{T}$$

. The above expression  $q^F(t)$  is designed to be an approximation of the cost of preventing stock-outs by moving modules. First, let  $\bar{K}^M$  be the average (directed) cost of moving a module between the pair of locations. Next, we approximate the stock-out likelihood using the probability of the maximum demand occurring at either location. Finally, we assume that given a stock-out, that it will be addressed by a module movement with probability one half. It is certainly possible to use different approximations of the cost purchasing future flexibility, but this approach led to good results for test instances.

We note that if the mobility purchase cost is very large, *i.e.*,  $\kappa^F(t) = \infty \forall t \in \mathcal{T}$ , the LFP policy is equivalent to **RF**. On the other hand, if mobility purchase cost  $\kappa^F(t) = 0 \forall t \in \mathcal{T}$ , we denote the resulting suboptimal policy as *lookahead with most flexible future* (**RLB**). Under the RLB policy, decisions are made each period assuming that in the next period there will be no additional cost of module movements.

#### *Lookahead with Iteratively Updated Cost-to-go (LIU)*

We now present a policy, LIU, generated by a simulation-based optimization method. The method seeks an approximate characterization of the expected cost function in the form of a lookup table covering the complete discrete state space of MMPIP. Once again, since such a method becomes intractable for larger values of  $L$ , we develop and test this heuristic policy using systems with only  $L = 2$  production locations.

In each pass of the algorithm presented in Section A3 of the appendix, the initial estimate of the cost-to-go function is set to the cost-to-go function of the fixed system. Demand outcomes are generated for all periods from the given stationary demand distributions by

Monte-Carlo simulation. The state at  $t = 1$ ,  $\xi(1)$ , has zero inventory at all locations and the best fixed module configuration  $u^{\min}$ . At every epoch  $t$ , at the current state in the sample trajectory, a new iterate is created by approximating the cost-to-go of the current state. This is done by blending the current approximation of the current state cost-to-do with a new estimate created by finding the expected cost of an optimal single-period action when using the current cost-to-go approximation at epoch  $t + 1$ . This approximation approach is repeated and cost-to-go functions updated along  $N$  sample trajectories and then the algorithm is terminated. The heuristic policy induced by the final estimate of the cost-to-go function is referred to as LIU.

Let  $\alpha_n$  be the blending coefficient used in the  $n^{\text{th}}$  iteration of the approach (for trajectory  $n$ ). The condition  $\sum_{n=1}^{\infty} \alpha_n \rightarrow \infty$  prevents premature stalling of the algorithm [28, Ch. 11]. Additionally, the condition  $\sum_{n=1}^{\infty} \alpha_n^2 < \infty$  ensures fast convergence of the iterates by limiting their variance. We test three blending coefficients, namely, a constant (0.5),  $1/(n + 1)$ , and  $1/(\text{number of visits to a particular state})$ , where  $n$  is the trajectory counter. We note that although the constant coefficient violates the second condition, it performs better than the other two candidates. Thus, we present computational results only for the constant blending coefficient.

## 2.6 Computational Study

We work with three different datasets for our study. Instance Set 1 is a set of  $L = 2$  instances that are small enough to allow computation of the exact optimal solution by value iteration. Instance Set 2 is a set of larger  $L = 2$  instances (with respect to module fleet size  $Y$ , horizon  $T$ , and capacity per module  $G$ ) for which the heuristic policies are compared with respect to  $\text{UB}_{\min}$ , the upper bound computed at the best fixed system configuration. Instance Set 3 focuses on problems with  $L \geq 2$  locations, and studies the performance of MP, RF, and LAF with respect to  $\text{UB}_{\min}$ . The design of these three sets of instances is presented in Section A4 of the appendix. To evaluate heuristic performance on all the

instance sets, on all forward dynamic programming paths, the initial state  $\xi(0)$  is set to zero inventory at all locations and the best fixed capacity configuration  $u^{\min}$ . Without loss of generality, we let  $\bar{K}^M$  be the average module movement cost between any two locations. We also let  $b_i = b$  and  $h_i = h$  for all locations  $i$ .

### 2.6.1 Instance Set 1.

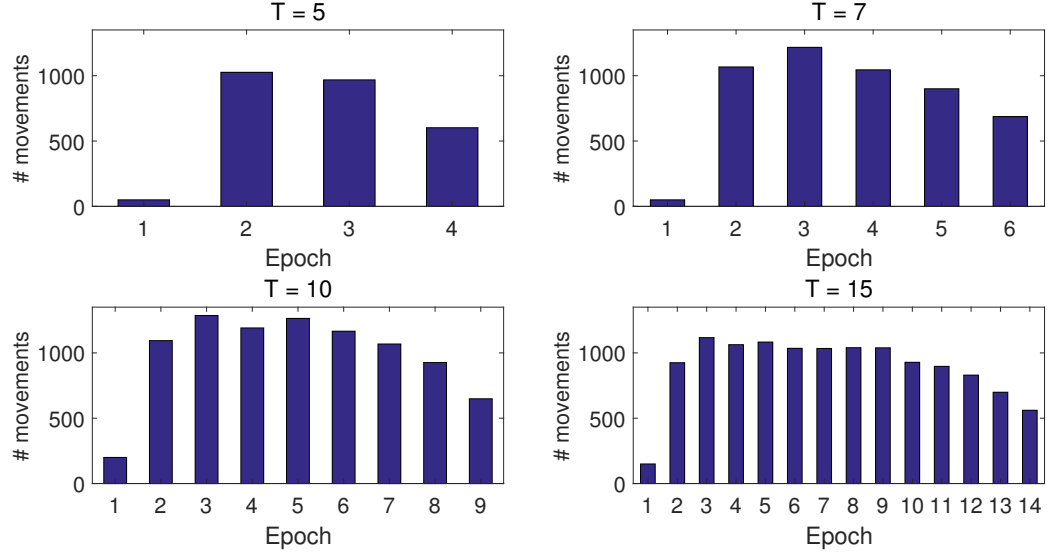
This set consists of 1280 instances with  $L = 2$ . We study the performance of the proposed bounds and heuristics with respect to the value of optimal total expected cost function (OPT).

We consider the subset of the instances for which the optimal policy prescribes module movement in at least one of 50 simulated trajectories. Of the 1280 instances, 595 instances exhibit movement in a simulation experiment with 50 sampling-based trajectories per instance. Figure 2.2a of histograms of optimal movement tendency shows that movements are observed for  $L = 2$  both in early as well as later periods over various lengths of the horizon. Figure 2.2b presents a bar plot of movement tendency for 50 simulated trajectories of one specific shifting instance from Instance Set 1 with the horizon length changed to  $T = 100$ . This movement tendency establishes that module movements is not caused by end effects and hints at effects of spatio-temporal dissimilarity and uncertainty.

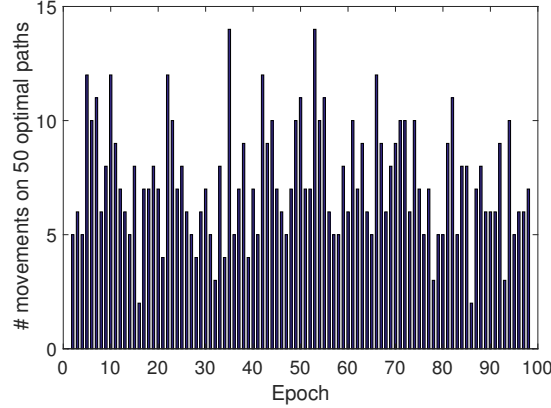
We now pursue patterns in module movement tendency, which is observed in about 50% of the instances of Instance Set 1. We label each instance as  $(AB, CD)$  where  $A, B, C, D \in \{L, H\}$  respectively indicate the level of expected demand (ED) at location 1, location 2, coefficient of variation (CV) at location 1 and location 2 respectively. For any location, the level of ED is high ( $H$ ) if it is greater than or equal to  $0.3Y$  and low ( $L$ ) otherwise. Likewise, CV is high( $H$ ) if it is greater than or equal to 0.5 and low ( $L$ ) otherwise. Based on such a classification of instances, we study the variation in movement tendency in Instance Set 1 across these classes.

In Figure 2.3, we observe that instances with high ED at both locations show the high-





(a) on Instance Set 1.



(b) for one instance with  $T = 100$ .

Figure 2.2: Module movement tendency of the optimal policy

est tendency to move modules (the top three classes with HH before comma). Instances with high CV at both locations and high ED at any one location come next. Instances with opposite ED classes and CV classes at the two locations show moderate movement behavior. The lowest movement tendency is observed when both locations have low ED or low CV. This classification establishes that the nature of uncertainty plays a major role in the movement inclination of an instance.

We use the above information to obtain the following infographic (Table 2.1) that demonstrates the usefulness of mobile modularity (based on the percentage of instances exhibiting movement in each subset, we classify as Very low: 0-10%, Low: 10-20%, Mod-

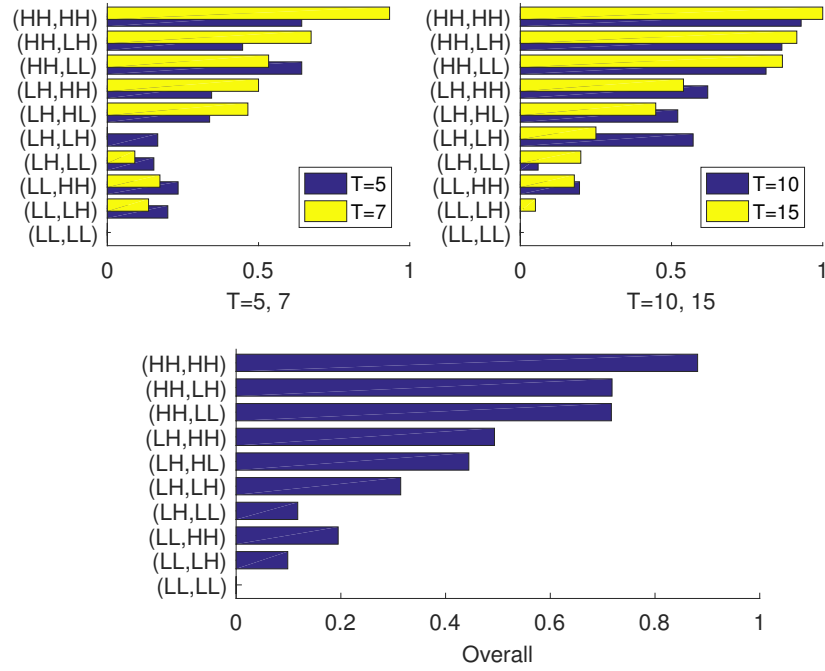


Figure 2.3: Effect of variability of demand on movement tendency

Table 2.1: Instance Set 1 - Usefulness of mobile modularity

	low ED at both	low, high EDs	high ED at both	Overall
high CV at both	Low	High	Very high	High
low, high CVs	Low	Moderate	Very high	Moderate
low CV at both	Very low	Low	Very high	Moderate
	Low	Moderate	High	

erate: 20-45%, High: 45-70%, and Very high:  $\geq 70\%$ ). Such a table can be used by firms to evaluate how useful mobile modularity can be to their specific demand behavior. The subset of Instance Set 1, which have high ( $> 0.3GY$ ) expected demand at both locations (and hence, are expected to show higher tendency of module movement than the other instances) are referred to as HH instances of Instance Set 1. Table 2.1 confirms the intuition that at high CVs and EDs, there is greater benefit of using mobile modularity. Another key takeaway is that the magnitude of ED of demand random variables has a stronger effect than the magnitude of their CV. Irrespective of how high the variability of demand is, at high average demands, there is immense utility of mobile modular systems.

In Tables 2.2, we study the variation in the quality of the bounds with the total number

Table 2.2: Instance Set 1 - Variation of bounds w.r.t. OPT across  $T$ ,  $Y$ ,  $b$ ,  $h$ , and  $\bar{K}^M$ 

(a) $T$					(b) $Y$				
$T$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>	$Y$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>
5	0.224	0.946	4.8	1.021	3	0.520	0.871	4.2	1.119
7	0.228	0.938	6.5	1.033	5	0.221	0.923	7.4	1.059
10	0.221	0.930	9.2	1.062	7	0.105	0.961	9.6	1.020
15	0.199	0.931	13.2	1.086	10	0.026	0.990	12.4	1.004
(c) $b$					(d) $h$				
$b$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>	$h$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>
1	0.218	0.938	6.4	1.036	0.5	0.234	0.927	10.5	1.060
2	0.218	0.934	10.5	1.065	1	0.201	0.945	6.4	1.041
(e) $\bar{K}^M$									
$\bar{K}^M$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>					
0.5	0.181	0.960	8.6	1.079					
5	0.255	0.912	8.2	1.021					
Overall	0.218	0.936	8.4	1.050					

of modules  $Y$ , length of the horizon  $T$ , backorder rate  $b$ , holding rate  $h$ , and movement cost  $\bar{K}^M$  on set 1. PIR is consistently bad for all  $T$  and grows worse with increase in  $Y$ . It is interesting to note that PIR, although bad, is significantly better on HH instances (Table A.1 in appendix). This signifies the role of uncertainty and information in situations that benefit from access to flexibility and confirms the intuition of using mobile modularity as a hedge against spatio-temporal uncertainty of demand. We note that the performance of LB and UB<sub>min</sub> deteriorates with increasing  $T$  but improves with increasing  $Y$ . This behavior is expected, as with increase in horizon length, period-wise difference between the bound and the optimal accumulate and hence, the bound will be farther from the optimal on longer horizons. When the number of modules is increased, a smaller fraction of modules will be moved, thus making the bounds tighter. The cost of the worst fixed configuration, UB<sub>max</sub>, although a very bad bound, gives us an insight into the value addition by mobile modularity under a) nonstationary demands, and b) poor forecast of stationary demand. As expected, UB<sub>max</sub> grows worse with increase in both  $T$  and  $Y$ . LB is tighter when  $Y$  is high (as

moving fewer modules is profitable) and moves more frequently than required increasing its optimality gap at higher  $\bar{K}^M$ ,  $b$ , and lower  $h$  (Table 2.2). When  $b$  is low and/or  $\bar{K}^M$  is high, there is not enough incentive in moving modules, making  $UB_{\min}$  very tight.

Table A.2 in the appendix provides the results on value addition due to mobile modularity and the fractions of costs spent by an optimally run mobile modular system on moving modules, backordering, and holding when compared to the best fixed system on Instance Set 1. A significant reduction (18% on average and 30% for  $T = 15$ ) in cost due to backordering is observed due mobile modularity (on HH instances).

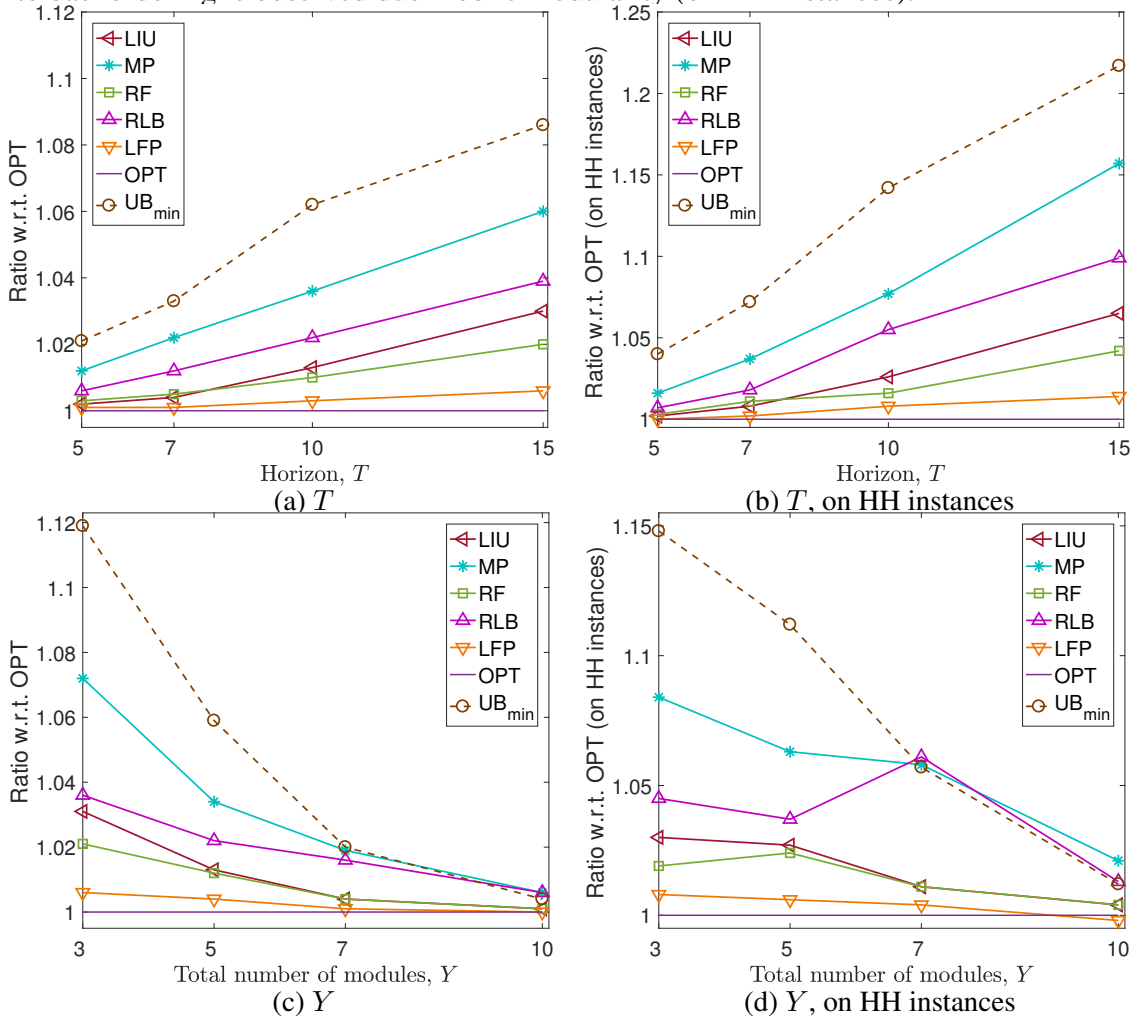


Figure 2.4: Instance Set 1 - Performance of heuristics w.r.t. OPT across  $T$  and  $Y$  on HH instances and all instances

Figure 2.4 presents the performance of heuristics over Instance Set 1 and HH instances

of Instance Set 1. In all the sub-figures of Figure 2.4, the ratios of MP, RF, RLB, and LFP with respect to the optimal solution are presented, compared with LIU (implemented with a constant blending coefficient of 0.5). The performance of MP improves with increase in  $Y$  and declines for larger horizon lengths. We note that MP leads to an overall average gap of about 9% (15% among shifting instances). RLB performs better than MP at all levels of  $Y$  and  $T$  with an overall average of 1.02 times the optimal MMPIP cost. RF outperforms RLB leading to an overall average of 1.01 times the optimal MMPIP cost. LFP’s performance is similar or better than that of LIU. For the  $L = 2$  problem, LFP leads to the lowest gap of 0.3% on average and it improves with increase in  $Y$ . LFP is computationally more intensive compared to MP, RLB or RF. MP is the most computationally efficient heuristic as it does not need to compute and store a lookup table over the entire state space while estimating the future cost. When only HH instances are considered, the performance of the proposed heuristics is still very good although the gaps are slightly higher. The lower gaps over the entire instance set can be attributed to the behavior of the fixed system’s cost, which guides the heuristics. We now discuss the movement tendency captured by these heuristics. All instances showing module movement in MP and RLB have movement tendency in the optimal policy as well. However, MP and RLB are more conservative than the optimal policy and exhibit movement in only 66% and 74% of the optimal policy’s shifting instances respectively. This result is intuitive as MP does not account for sustained future benefit while moving modules and RLB would not be eager to shift modules in the present due to its assumed zero cost future flexibility. RF and LFP show movement in most (more than 97%) of the instances that show movement by the optimal policy. All the heuristics work better at low  $b$  and high  $h$  (Table A.4). For high  $\bar{K}^M$ , LIU, RF, and LFP perform well but MP and RLB are closer to optimality when  $\bar{K}^M$  is low, as expected (Table A.4).

Focusing on the HH subset of Instance Set 1, we note that the most flexible system and the fixed system perform worse on HH instances than on the overall set, indicating the need

for mobile modularity. We observe that optimality gap of all the heuristics on HH instances is higher than the overall average, although, the magnitude of gap itself is still within 8%. Figure 2.4 show that LFP outperforms all the heuristics and is closely followed by RF.

### 2.6.2 Instance Set 2.

This set consists of 720 two location instances with relatively higher values of horizon  $T$ , module fleet size  $Y$ , more spread in backorder cost  $b$  and module movement cost  $K$ , and production capacity per module  $G$ . We compare the performance of the heuristics and the lower bound, most flexible system's optimal cost (LB), with respect to the upper bound, the optimal cost of the best fixed system ( $UB_{\min}$ ).

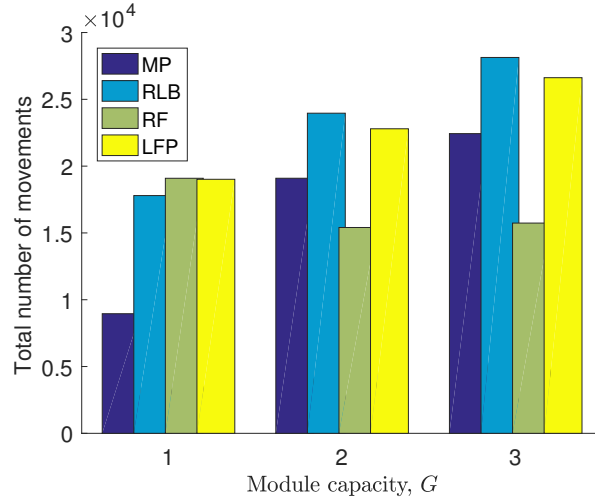


Figure 2.5: Instance Set 2 - Module movement tendency across  $G$

Figure 2.5 presents the module movement tendency at different module capacity levels. MP is the most conservative policy in moving modules as it aims to minimize the immediate costs only. RF restricts movements at higher capacity levels due to assumed fixed future that results in a greater impact due to removal at sending locations. As expected, RLB moves modules more frequently compared to MP and RF at  $G > 1$ . LFP exhibits a movement frequency, which falls between that of RF and RLB, as expected.

We observe that  $UB_{\max}$  increases with increase in  $T$ ,  $Y$ , and  $b$  but decreases with an increase in  $Y$ , as expected. It stands at 13.9 times  $UB_{\min}$  for  $G = 1$  and increases to 20.1

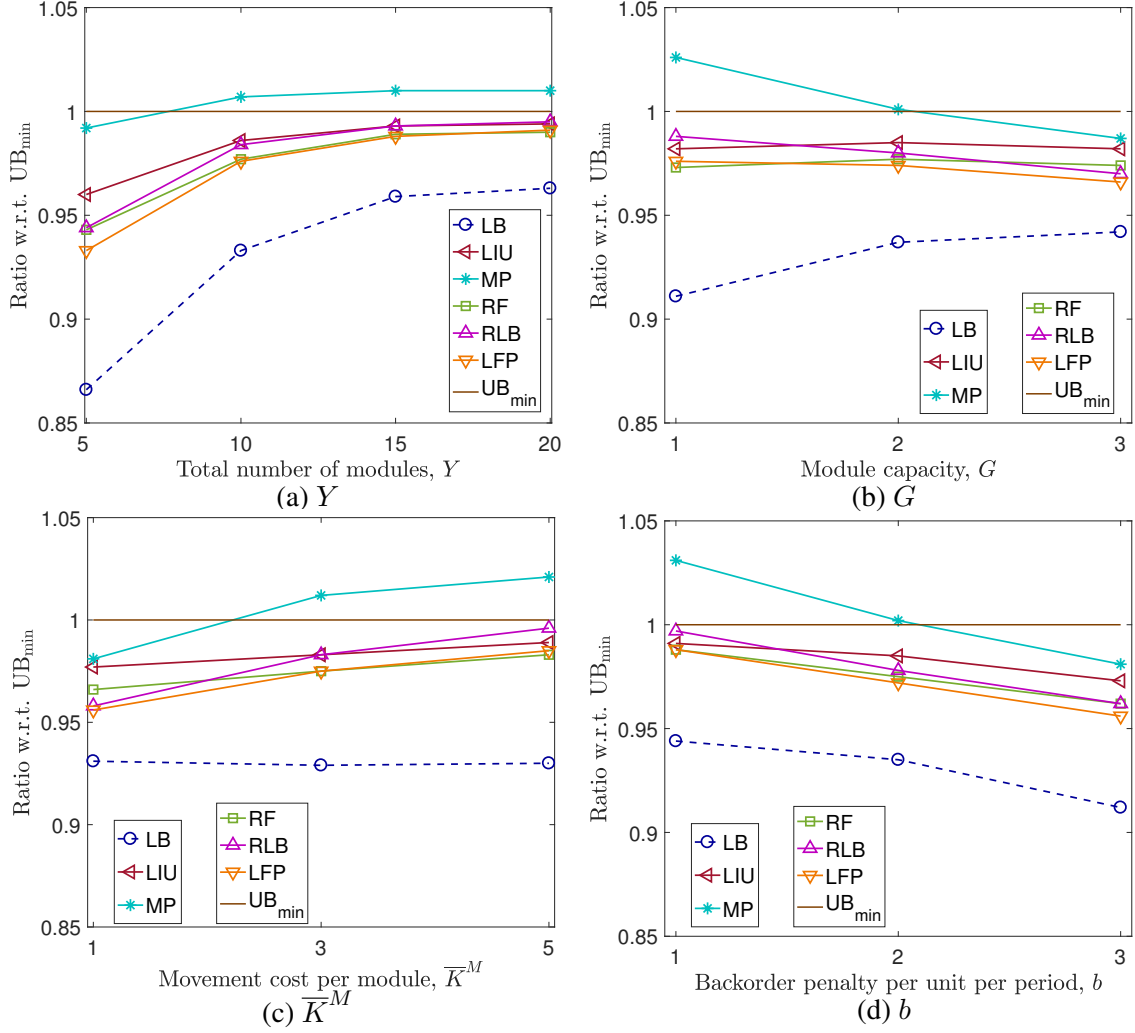


Figure 2.6: Instance Set 2 - Overall performance w.r.t.  $UB_{\min}$  across  $Y$ ,  $G$ ,  $\bar{K}^M$ , and  $b$

times  $UB_{\min}$  at  $G = 3$ . These magnitudes establish that for larger instance sizes, the risk of loss due to demand non-stationarity is higher as fixed systems do not have a recourse when demand distributions vary. We now focus our attention on Figs. 2.6 and 2.7a. The most flexible system lower bound LB is tighter for higher  $G$  and  $Y$ , and lower  $b$ ,  $\bar{K}^M$ , and  $T$ , which is expected as mobile modularity is not fully utilized in these situations. We see that RF and LFP perform much closer to optimality and significantly better than MP, LIU and RLB.

The best heuristic, LFP, delivers an average improvement of about 4% over  $UB_{\min}$  on Instance Set 2. However, focusing on the HH instances of Instance Set 2 (Fig. 2.7b),

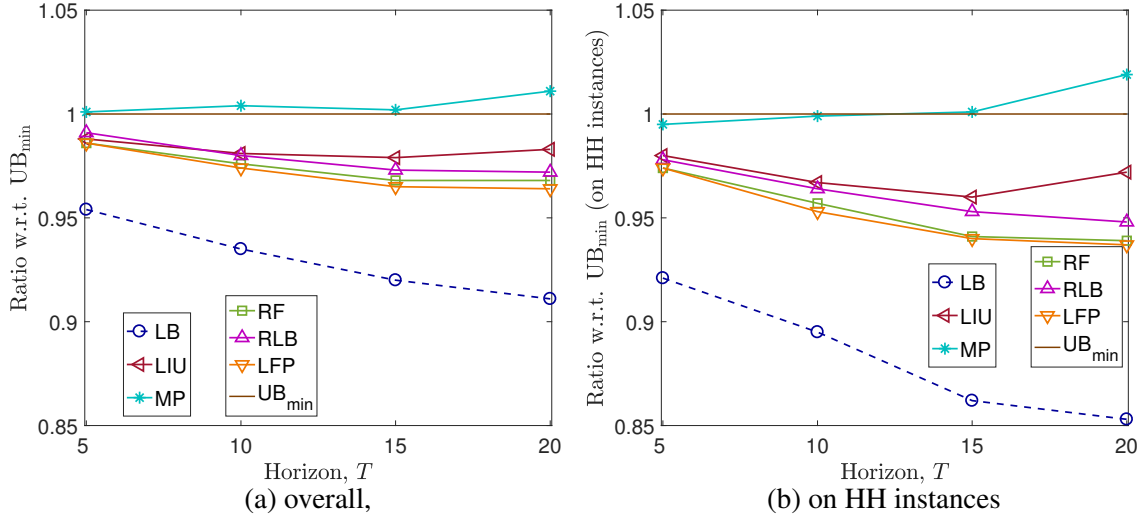


Figure 2.7: Instance Set 2 - Performance w.r.t.  $UB_{\min}$  across  $T$

we note a 5% average advantage over fixed systems for LFP, which is close to RF(4.8%), and higher than RLB (4%), LIU (3%), and MP (−0.4%). For certain configurations, LFP costs about 9% less than  $UB_{\min}$ . These results reveal promise in the resourcefulness of the proposed heuristics and have encouraged our pursuit of the solution of the problem for  $L > 2$ .

### 2.6.3 Instance Set 3.

This set consists of 540 instances generated by a procedure described in the appendix, with  $L \geq 2$ . We study the performance of three heuristics, *viz.*, the myopic policy(MP), the lookahead with approximate fixed future (LAF), and the rollout of fixed future (RF) on 50 simulated trajectories of every instance in this instance set. We conduct our study by setting the initial state of the system to have zero inventory positions at all locations and two different capacity configurations: a)  $u_{\text{simple}}$  and b)  $u_{\min}$ .  $u_{\text{simple}}$  is an allocation of the modules based on both the mean demand and variable component of demand at each location (see Appendix for details).  $u_{\min}$  is the best fixed system's capacity configuration.

The computations for the study on Set 3 are performed on a single server of an Intel Xeon Processor E5-2670 workstation. We first take stock of the times taken to compute the best upper bound  $UB_{\min}$  (one-time and offline) and the three heuristics MP, LAF, and



RF (Table 2.3) on  $T = 15$  instances. We note that LAF is significantly faster than RF and is comparable to MP in its order of magnitude. Figure 2.8 shows that, as  $L$  increases, the effort required to implement RF increases drastically and LAF is on average 35 times faster than RF for  $L = 20$ . Additionally, we note that using LAF takes less than twice the effort required for implementing the naive MP policy, even when  $L$  is high. Thus, LAF presents remarkable computational advantage over RF.

Table 2.3: Instance Set 3 - Average computation time in seconds for  $T = 15$  instances across  $L$

$L$	$UB_{\min}$ (per instance)	MP (per trajectory)	LAF	RF
2	0.08	0.05	0.08	0.21
3	0.23	0.07	0.13	0.33
5	0.87	0.10	0.21	0.74
10	7.30	0.23	0.48	3.6
20	96.0	0.77	1.30	33.7

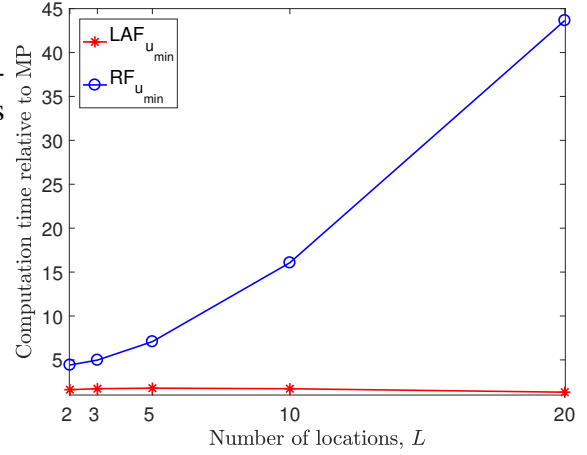


Figure 2.8: Instance Set 3 - Average computational effort relative to MP across  $L$

We observe that LAF and RF perform almost identically, suggesting that the approximation of fixed future proposed in LAF is efficient both in speed and quality. LAF and RF improve over  $UB_{\min}$  by 4-9% and the savings increase with  $L$  on average. Also, lesser savings are obtained for higher  $Y$  for any given  $L$ , as fewer modules would be moved between locations due to greater availability of capacity at each location. MP results in average costs that are sometimes higher than  $UB_{\min}$  also, indicating that naive approaches would not yield the advantage that can potentially be garnered by mobile modularity. Fig. 2.9a shows that the well-performing heuristics, RF and LAF, generally yield increasing benefit with increasing  $L$ .

Fig. 2.9c confirms the intuition that, with increasing  $\bar{K}^M$ , well-performing heuristics, such as LAF and RF, lead to lesser benefit over the fixed system. MP appears to be more sensitive to higher values of  $\bar{K}^M$ , as moves made for immediate cost benefits may induce

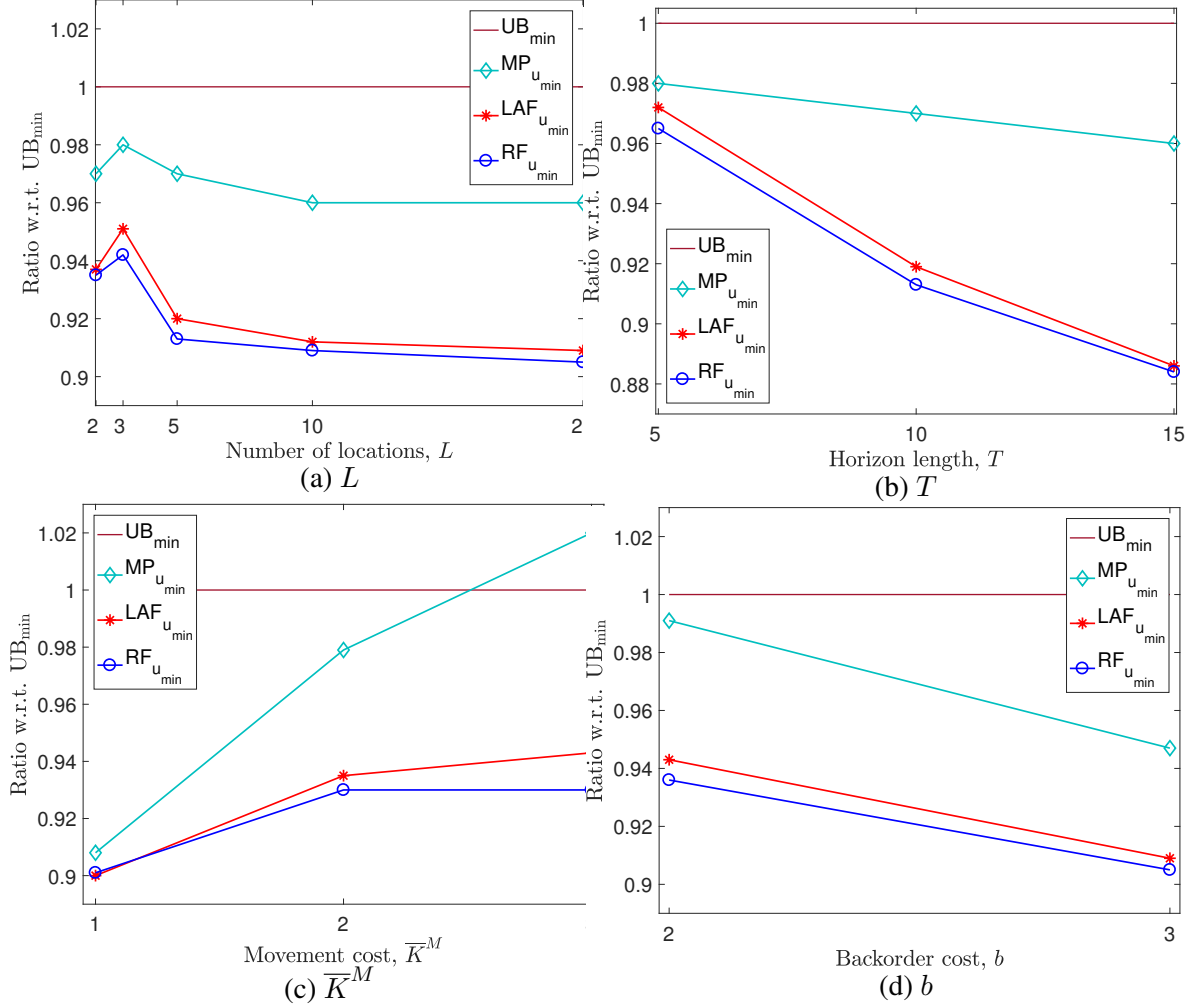


Figure 2.9: Instance Set 3 - Performance of heuristics w.r.t.  $UB_{\min}$ , when initialized to  $u_{\min}$ , across  $L$ ,  $T$ ,  $\bar{K}^M$ , and  $b$

more movements to the escape the bad state that was landed in. As the backorder rate per unit per period,  $b$ , increases, there is greater use of mobile modularity (Fig. 2.9d). In Fig. 2.9b, we note that higher value addition due to mobile modularity is attained on longer horizons due to accumulated benefit (about 12% for  $T = 15$ ). Motivated by this observation, we present at the performance of the heuristics across  $L$  for the longest possible horizon,  $T = 15$ , in Fig. 2.10b. We achieve an average value addition ranging between 9 to 16% on realistically sized instances for  $T = 15$ .

We also compare the resilience of the two systems of interest, namely, mobile modular systems and fixed systems, by studying the effect of configuring both systems to a simple

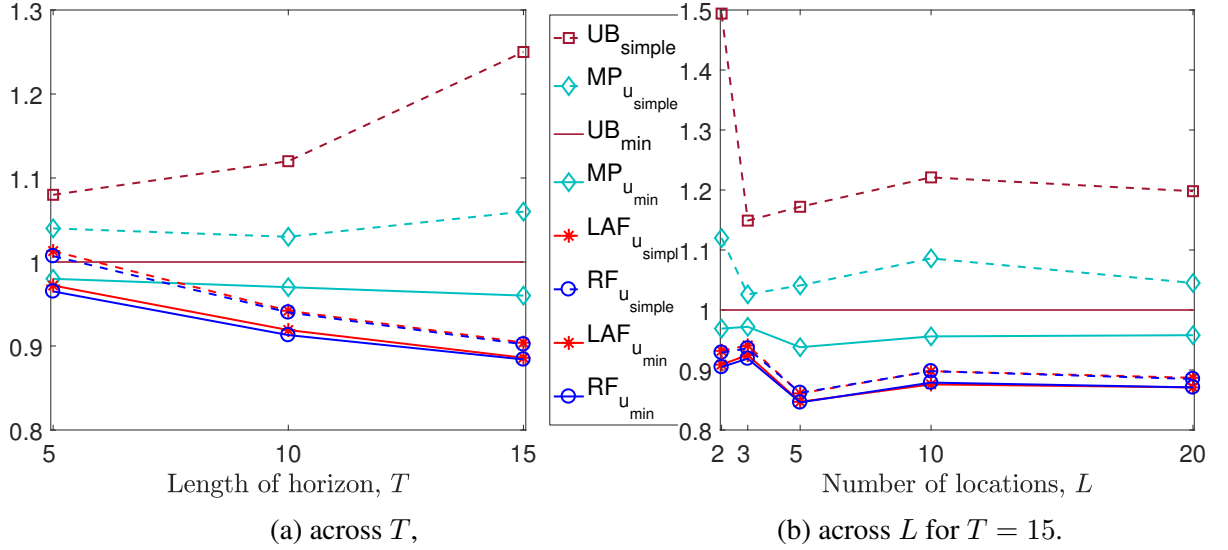


Figure 2.10: Instance Set 3 - Performance of heuristics when initialized to  $u_{\min}$  and  $u_{\text{simple}}$  initial capacity allocation based on location-wise expected demand and variability,  $u_{\text{simple}}$ , instead of  $u_{\min}$  (see Fig. 2.10). In Fig. 2.10a, we note that the gap between the implementations of RF with two different initial configurations reduces as the horizon grows longer. We may infer that the difference in heuristic performance is due to the amortization of a one time movement cost to switch from  $u_{\text{simple}}$  to the better capacity configuration  $u_{\min}$ . The cost of the fixed system configured to  $u_{\text{simple}}$  is about 25% higher than UB<sub>min</sub> when the length of the horizon,  $T$ , is 15 (Fig. 2.10b). However, setting the initial state to  $u_{\text{simple}}$  instead of  $u_{\min}$  leads to only about 2% decrement in the gaps of LAF and RF at  $T = 15$ . This observation establishes that the mobile modular system is indeed very resilient compared to fixed systems as it is able to retrieve most of the savings over fixed systems even with suboptimal initial configurations (about 25-30% savings over the fixed system configured to  $u_{\text{simple}}$ ). These observations indicate that LAF and RF perform robustly irrespective of the initial capacity configuration and mobile modular systems are very resilient relative to fixed systems.

Our computational study firmly establishes that mobile modular production systems respond to the uncertainty of demands in multi-location, multi-period production-inventory systems significantly more effectively (in terms of cost and resilience) than fixed systems.

Table 2.4: Instance Set 3 - Performance of heuristics w.r.t.  $UB_{\min}$  across  $L$  and  $Y$

L	Y	MP	LAF	RF
2	4	0.96	0.914	0.908
	5	0.97	0.960	0.961
3	6	0.96	0.913	0.897
	8	1.00	0.990	0.986
5	10	0.94	0.866	0.860
	12	1.00	0.974	0.965
10	19	0.93	0.851	0.848
	24	1.00	0.973	0.970
20	38	0.93	0.834	0.829
	48	1.00	0.985	0.981
Overall		0.97	0.93	0.92

## 2.7 Conclusion

We introduce the mobile modular production and inventory problem and present novel solution techniques that are suitable for multilocation instances also. The proposed heuristics for mobile modular production systems accomplish immense value addition over optimally run fixed systems, providing evidence of efficiency, practicality, and suitability of mobile modular production systems. Our heuristics perform very close to the optimal solution for  $L = 2$  and significantly better than the fixed system for  $L \geq 2$ . Through our computational study, we demonstrate the cost-effectiveness and resilience of mobile modular systems in comparison to fixed systems. Additionally, the heuristic LAF achieves remarkable computational efficiency. Future work may focus on better models of demand and the incorporation of inventory transshipment also in the model. Lead times of module movement and thus a down time in production for some of the modules may also be incorporated into the stochastic optimization framework.

## CHAPTER 3

### INVENTORY CONTROL WITH MODULATED DEMAND AND A PARTIALLY OBSERVED MODULATION PROCESS

#### 3.1 Introduction

We consider a periodic review, data driven inventory control problem over finite and infinite planning horizons with instantaneous replenishment. We assume that there are several interconnected processes: the completely observed inventory process that keeps track of the inventory level, the uncensored demand process, the action process that represents replenishment decisions, the underlying modulation process that affects demand, and the additional observation data (AOD) process that together with the demand process partially observes the modulation process. The inventory, demand, and action processes are common to inventory control problems. When completely observed by the demand and AOD processes, the modulation process models the case where demand is Markov-modulated. When the modulation process is only observed by the demand process and is assumed static, then the model conforms to the model considered by the Bayesian updating literature. The modulation process can represent an unknown static parameter or index of the demand process, the state of the world, etc., and can model dynamic exogenous factors, such as the weather, seasonal effects, and the underlying economy. The AOD process can model observations of the modulation process other than demand; e.g., macro-economic indicators.

In Section 3.2, we model this problem as a partially observed Markov decision process (POMDP) and present related preliminary results and a key assumption, A1, a generalization of the Veinott attainability assumption, assuming the reorder cost  $K = 0$ . We assume A1 holds in Section 3.3 and show there exists a myopic optimal base stock policy, the value

of the optimal base stock level is constant within regions of the belief space, and these regions can be described by a finite set of linear inequalities. A1 guarantees that the current base stock (i.e., order up to) level is always at least as great as the current inventory level. We then present conditions that imply A1 holds and present a procedure for computing the optimal cost function.

We assume A1 does not hold in Section 3.4, present lower and upper bounds on the optimal cost function based on the base stock policy that is optimal when A1 holds, and present an upper bound on the difference between these two bounds. Interestingly, we show that the upper bound on the optimal cost function is piecewise linear in the belief function for the finite horizon case but may not be continuous; hence, improved observation quality of the modulation process may not result in improved systems performance. We then present a tighter lower bound based on the assumption that A1 holds within  $\delta > 0$  and show that this tighter lower bound improves as  $\delta$  gets smaller.

We consider the  $K > 0$  case in Section 3.5 and assume throughout that A1 holds. We show that there exists an optimal  $(s, S)$  policy and determine upper and lower bounds on  $s$  and  $S$  for the finite and infinite horizon cases, where each bound and the values of  $s$  and  $S$  are dependent on the belief function of the modulation process. Each of these bounds and the values of  $s$  and  $S$  are shown to be constant within regions of the belief space described by a finite number of linear inequalities. An outline of an approach for determining an optimal  $(s, S)$  policy and the resultant expected cost function for the finite horizon case are presented in the appendix. Conclusions are presented in Section 3.6.

### 3.1.1 Literature Review

Inventory control has been studied extensively over six decades; see [39], [40], [41], [42], [43], [44], and [45] for detailed surveys. This survey is organized around various assumptions made in the literature regarding the modulation process. We also survey several non-parametric approaches. We first consider the  $K = 0$  case, followed by the  $K > 0$  case.

Assuming  $K = 0$ , the case where the modulation process is completely observed and static was first considered by [46], and [47, 48], various extensions of which are detailed in surveys by [39], [41], [42], and [43]. The case where the modulation process is completely observed and nonstationary was first considered by [49, 50], [51], and [52, 53]. A base stock policy dependent on the state of the modulation process (current demand distribution) was proved to be optimal in [49] and [50]. [51] developed computational approaches for determining the base stock level. [54] extended these results to the average cost criterion and to cyclic costs. [53] and [52] proved the existence of an optimal myopic base stock policy when the base stock level at the next decision epoch is guaranteed to exceed the current inventory position after satisfying demand (i.e., the attainability assumption) for independent and correlated nonstationary demands across time periods, respectively. [53] also provided sufficient conditions for this assumption. [55] studied an inventory system with the additional option of disposal of inventory at a cost and nonstationary demands in each period. [56] modeled explicit dependence of a generalized demand process on a modulation process with exogenous parameters and demand history and derived an upper bound on the optimal cost for scenarios such as Markov modulation and additive and multiplicative demand shocks. [57] modeled the modulation process as a completely observed underlying “state-of-the-world” in a continuous time framework similar to [51], with a Markov-modulated Poisson demand process. A “state-of-the-world” dependent base stock policy was proved to be optimal. When attainability of the next period’s base stock level is guaranteed, an optimal myopic policy was shown to exist. [58] extended the results of [57] to a discrete time system and obtained analogous results. [59] dealt with martingale-demand under a robust optimization framework.

[60], [61], and [62] analyzed the case where the modulation process is static, partially observed by the demand process, completely unobserved by the AOD process, and represents unknown parameters of a single stationary distribution. While [61] proved the optimality of a statistic-dependent base stock policy, [62] extended the results to determine

a Bayesian update on unknown parameters. [63] and [64] extended these results ([61]) to other distributions and compared this method with non-Bayesian mixture methods. [65] built on their work to prove the optimality of a myopic base stock policy for “parameter adaptive models” of demand, and [66] dealt with an unknown stationary distribution of demand partially observed by a scale parameter and a shape parameter. [67] presented a study of the Bayesian updating mechanism with and without nonstationarity and disposal.

Partial observability of demand outcomes results from limitations on the accuracy of inventory book-keeping (in [68] and [69]), and censoring (in [70, 71]). [70, 71] treated Markovian modulation of demand as a special case. Their problem formulation differs from our framework as they learn the unknown stationary demand distribution in a Bayesian fashion and the demand process (not the modulation process) is partially observed (censored). [72] presented an analysis of optimal policies for the Bayesian newsvendor problem with and without censoring.

For the case where the modulation process is partially observed by the demand process, completely unobserved by the AOD process, and dynamic, [73] proved the existence of an optimal state-dependent base stock policy for an uncapacitated inventory system. [74] proved the optimality of inflated state-dependent base stock policies for capacitated production systems under Markov-modulated demand and supply processes (extending [75]). [76] studied a completely unobserved Markov-modulated Poisson demand process in a continuous-review inventory system with reorder cost and lost sales (censoring).

[73] and [74], however, did not prove the existence of an optimal *myopic* state-dependent base stock policy, which we prove in this chapter, assuming A1 holds. Further, we show that the belief space can be partitioned into subsets by a finite set of linear inequalities and that the base stock level is constant within each of these subsets. Such regions have also been observed in the numerical example provided in [76]; however, no explanation is given for such behavior. The linear partition of the belief space we present provides an easily computed approach to determine an optimal base stock level for any given belief vector.



For the  $K > 0$  case, [77] and [78] proved that there exists an optimal  $(s, S)$  policy under finite and infinite horizons, respectively. [78] presented the first set of bounds on period-wise reorder points and base stock levels, which were later tightened by [79]. [80] extended [53, 52] to the  $K > 0$  case. More recently, [81] presented sufficiency conditions of divergence and  $K$ -convexity for the optimality of  $(s, S)$  policies under time-varying parameters and correlated demand variables modulated by an underlying “state-of-the-world” variable. Our results extended to the  $K > 0$  case lead to significantly reduced computational effort in determining the optimal policy compared to [76] when A1 holds.

More recently, nonparametric approaches for describing demand uncertainty have garnered interest. [82] presented a bootstrap procedure when lead time distribution is unknown. [83] and [84], [85] and [86] studied problems with partial information about the demand randomness, *viz.*, mean and variance, moments and shape, and censored data respectively. [87] obtained history-dependent base stock levels while simultaneously optimizing and learning the histogram of realized demand. For the distribution-free problem, [88] and [89] employed censoring using statistical estimators, [90] and [91] used machine learning techniques in conjunction with optimization, [92] and [93] studied the performance of the sample average approximation (SAA), and [94] applied a piecewise linear value function approximation. [95] estimated historical data for new products and presented an algorithm to perform price optimization. Future research may involve a blend of nonparametric approaches with Bayesian approaches, such as the approach presented in this chapter.

## 3.2 Problem Description and Preliminary Results

We describe the inventory control problem in Section 3.2.1. We then model the problem as a POMDP and present optimality equations and other standard results in Section 3.2.2. In Section 3.2.3 we present results associated with the single period expected cost function that will be useful in later sections and also present the condition A1.

### 3.2.1 Problem Definition

We consider an inventory control problem that involves the inventory process  $\{s(t), t = 0, 1, \dots\}$ , the modulation process  $\{\mu(t), t = 0, 1, \dots\}$ , the demand process  $\{d(t), t = 1, 2, \dots\}$ , the additional observation data (AOD) process  $\{z(t), t = 1, 2, \dots\}$ , and the action process  $\{a(t), t = 0, 1, \dots\}$ . These processes are linked by the state dynamics equation  $s(t+1) = f(y(t), d(t+1))$ , where  $y(t) = s(t) + a(t)$ , and the given conditional probability  $\Pr(d(t+1), z(t+1), \mu(t+1) \mid \mu(t))$ . We assume the single period cost accrued between decision epoch  $t$  and  $t+1$  is  $c(y(t), d(t+1))$ , where  $c(y, d)$  is convex in  $y$  and  $\lim_{|y| \rightarrow \infty} c(y, d) \rightarrow \infty$  for all  $d$ . We also assume that  $c(y, d)$  is piecewise linear in  $y$  for all  $d$  and that the facets describing  $c(y, d)$  intersect at integers. We will have particular interest in the case where  $f(y, d) = y - d$ , which assumes backlogging, and  $c(y, d) = p(d - y)^+ + h(y - d)^+$ , where  $p$  is the shortage penalty per period for each unit of stockout,  $h$  is the holding cost per period for each unit of excess inventory after demand realization, and  $(g)^+ = \max(g, 0)$ . Without loss of significant generality, this definition of single period cost does not include an ordering cost. It is straightforward to transform an inventory problem with a strictly positive ordering cost into an inventory problem with no ordering cost for a wide variety of cost and dynamic models of inventory position, e.g.,  $f(y, d) = y - d$  or  $f(y, d) = (y - d)^+$  and  $c(s, y, d) = c'(y - s) + p(d - y)^+ + h(y - d)^+$ , where in this case the single period cost accrued between decision epochs is dependent on  $s$  and  $c'$  is the cost per unit ordered.

We assume that the modulation, demand and AOD state spaces are all finite, the inventory process has a countable state space, and the action space is the set of non-negative integers. We assume the action at  $t$  can be selected based on  $s(t), d(t), d(t-1), \dots, z(t), z(t-1), \dots$ , and the prior probability mass vector  $\{\Pr(\mu(0) = \mu_i), \forall i\}$ . Thus, the inventory process is completely observed, demand is not censored, and the modulation process is partially observed by the demand and AOD processes. The problem is to determine a policy that minimizes the expected total discounted cost over the infinite horizon, where

we let  $\beta \in [0, 1)$  be the discount factor. It is assumed throughout that replenishment is instantaneous.

We remark that the inventory, demand, and action processes are all part of inventory control problems considered in the literature. As indicated in the literature review, the modulation process is also part of the structure of inventory control problems with Markov-modulated demand. The AOD process is intended to provide information about the modulation process, where appropriate, in addition to that provided by the demand process, such as macro-economic data. Throughout we assume demand realization is uncensored and completely revealed. This assumption is in contrast to the censored demand case where only sales data are available to the decision maker.

We note that the conditional probability  $\Pr(d(t+1), z(t+1), \mu(t+1) \mid \mu(t))$  is the product of two conditional probabilities:

1.  $\Pr(d(t+1), z(t+1) \mid \mu(t+1), \mu(t))$ , the demand and AOD probabilities, conditioned on the modulation process
2.  $\Pr(\mu(t+1) \mid \mu(t))$ , the state transition probabilities for the (Markov-modulated) modulation process.

The Baum-Welch algorithm is typically used to estimate parameters of a POMDP, *viz.*, observation and transition probabilities and initial belief state (see [96] for a review on POMDP training methods).

We remark that demand is i.i.d. under several assumptions including:

1. if  $z(t+1) = \mu(t+1)$  w.p.1  $\mu(t+1) = \mu(t)$  w.p.1 and  $\Pr(\mu(0) = \mu_i) = 1$  for some given  $i$ .
2. if  $z(t+1)$  is independent of  $\mu(t+1)$  and  $\mu(t)$  and  $\mu(t+1) = \mu(t)$  w.p.1.
3. if  $d(t+1)$  is independent of  $z(t+1)$ ,  $\mu(t+1)$ , and  $\mu(t)$ .

### 3.2.2 The POMDP Model and Preliminary Results

*Optimality equations.*

This problem can be recast as a partially observed Markov decision problem as follows. Results in [97] and [98] imply that  $(s(t), x(t))$  is a sufficient statistic, where  $N$  is the number of values the modulation process can take, the belief function  $\mathbf{x}(t) = \text{row}\{x_1(t), \dots, x_N(t)\}$ , is such that  $x_i(t) = \Pr(\mu(t) = \mu_i \mid d(t), \dots, d(1), z(t), \dots, z(1), x(0))$ , and  $\mathbf{x}(t) \in X = \{\mathbf{x} \in \mathbb{R}^N : \mathbf{x} \geq 0 \text{ and } \sum_{i=1}^N x_i = 1\}$ . For  $\mathbf{g} \in \mathbb{R}^N$ , let  $\mathbf{g}\mathbf{1} = \sum_{n=1}^N g_n$ . Thus, the inventory process is completely observed, the modulation process is partially observed through the demand and AOD processes, and the state of the modulation process is characterized by the belief function. Let

$$\begin{aligned} P_{ij}(d, z) &= \Pr(d(t+1) = d, z(t+1) = z, \mu(t+1) = j \mid \mu(t) = i), \\ \mathbf{P}(d, z) &= \{P_{ij}(d, z)\}, \\ \sigma(d, z, \mathbf{x}) &= \mathbf{x}\mathbf{P}(d, z)\mathbf{1} = \sum_{i=1}^N x_i \sum_j P_{ij}(d, z), \\ \boldsymbol{\lambda}(d, z, \mathbf{x}) &= \text{row}\{\lambda_1(d, z, \mathbf{x}), \dots, \lambda_N(d, z, \mathbf{x})\} = \mathbf{x}\mathbf{P}(d, z)/\sigma(d, z, \mathbf{x}), \\ \sigma(d, z, \mathbf{x}) &\neq 0, \\ L(\mathbf{x}, y) &= E[c(y, d)] = \sum_{d, z} \sigma(d, z, \mathbf{x})c(y, d). \end{aligned}$$

Define the operator  $H$  as

$$[Hv](\mathbf{x}, s) = \min_{y \geq s} \left\{ L(\mathbf{x}, y) + \beta \sum_{d, z} \sigma(d, z, \mathbf{x}) v(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(y, d)) \right\}. \quad (3.1)$$

Results in [99] guarantee that there exists a unique cost function  $v^*$  such that  $v^* = Hv^*$  and that this fixed point is the expected total discounted cost accrued by an optimal policy. We can restrict search for an optimal policy to  $t$ -invariant functions that select  $a(t)$  on the basis of  $(s(t), \mathbf{x}(t))$ , the function  $\psi$  such that  $\psi(s(t), \mathbf{x}(t)) = a(t)$  causing the minimum

in Equation 3.1 to be attained is an optimal policy, and  $\lim_{n \rightarrow \infty} \|v^* - v_n\| = 0$ , where the (finite horizon) cost function  $v_{n+1} = Hv_n$  for any given bounded function  $v_0$  and  $\|\cdot\|$  is the sup-norm. The function  $L(\mathbf{x}, y)$  is the expected single period cost, conditioned on belief  $\mathbf{x}$  and inventory level  $y$ . From the perspective of Bayes' Rule, note that  $\mathbf{x} = \mathbf{x}(t)$  can be thought of as the prior probability mass function of  $\mu(t)$ ,  $\sigma(d, z, \mathbf{x})$  is the probability that the demand and AOD processes will have realizations  $d = d(t+1)$  and  $z = z(t+1)$ , respectively, given  $\mathbf{x}$ , and  $\mathbf{x}(t+1) = \boldsymbol{\lambda}(d, z, \mathbf{x})$  is the posterior probability mass function of  $\mu(t)$ , given  $d, z$ , and  $\mathbf{x}$ .

#### *Piecewise linearity and concavity in $\mathbf{x}$ .*

Results in [97] guarantee that  $v_n(\mathbf{x}, s)$  is piecewise linear and concave in  $\mathbf{x}$  for each fixed  $s$  for all finite  $n$ , assuming  $v_0(\mathbf{x}, s)$  is also piecewise linear and concave in  $\mathbf{x}$  for each  $s$ . In the limit  $v^*(\mathbf{x}, s)$  may no longer be piecewise linear in  $\mathbf{x}$  for each  $s$ ; however, concavity will be preserved.

#### *Value of Information and Upper and Lower Bounds.*

Let  $q(d, z \mid i, j) = \Pr(d(t+1) = d, z(t+1) = z \mid \mu(t+1) = j, \mu(t) = i)$ , and assume  $\mathbf{Q} = \{q(d, z \mid i, j)\}$ , which we call the observation array. The observation array  $\mathbf{Q}$  is stochastic in the sense that  $q(d, z \mid i, j) \geq 0$  for all  $i, j, d, z$  and  $\sum_{d,z} q(d, z \mid i, j) = 1$  for all  $(i, j)$ . Following [100], the observation array  $\mathbf{Q}'$  is said to be at least as informative as the observation array  $\mathbf{Q}$  if there exists a stochastic array  $\mathbf{R} = \{r(d, z \mid d', z')\}$  such that  $\sum_{d',z'} q'(d', z' \mid i, j) r(d, z \mid d', z') = q(d, z \mid i, j)$  for all  $i, j, d, z$  (or equivalently,  $\mathbf{Q}'\mathbf{R} = \mathbf{Q}$ ). Consider two problems, the unprimed and primed problems, that are defined identically except the unprimed problem is associated with the observation array  $\mathbf{Q}$  and the primed problem is associated with the observation array  $\mathbf{Q}'$ . Let  $\{v_n\}$  and  $v^*$  be associated with the problem having observation array  $\mathbf{Q}$ , let  $\{v'_n\}$  and  $v^{*'}$  be associated with the problem having observation array  $\mathbf{Q}'$ , and assume there is a stochastic array  $\mathbf{R}$  such that  $\mathbf{Q}'\mathbf{R} =$

$\mathbf{Q}$ . Then, according to results in [100], for all  $(\mathbf{x}, s)$ ,  $v'_n(\mathbf{x}, s) \leq v_n(\mathbf{x}, s)$  for all  $n$  and  $v^{*'}(\mathbf{x}, s) \leq v^*(\mathbf{x}, s)$ . Thus, if the observation array  $\mathbf{Q}'$  is at least as informative as the observation array  $\mathbf{Q}$ , then the primed problem is guaranteed to perform as least as well as the unprimed system (i.e., the value of more accurate information about the modulation process is positive).

It is then straightforward to show (see [98]):

- (i) If the modulation process is only observed by the demand process (and hence the AOD process is not a function of  $\mu(t+1)$  and  $\mu(t)$  and hence provides no information regarding the state of the modulation state), then the resulting infinite and finite horizon cost functions are upper bounds on the cost functions of the general case.
- (ii) If  $\Pr(z(t+1) \mid \mu(t+1), \mu(t)) = 1$  if and only if  $z(t+1) = \mu(t+1)$  w.p.1 (the case where the modulation process is completely observed by the AOD process), then the resulting infinite and finite horizon cost functions are lower bounds on the cost functions of the general case.

For the case where  $\Pr(z(t+1) \mid \mu(t+1), \mu(t)) = \Pr(z(t+1) \mid \mu(t+1))$ , i.e., observation  $z(t+1)$  is independent of modulation  $\mu(t)$ , the matrix  $\{\Pr(z(t+1) \mid \mu(t+1))\}$  has rank 1 for case (i) and is the identity matrix for (ii).

### 3.2.3 $\mathbf{L}(\mathbf{x}, \mathbf{y})$ Analysis

We now examine  $L(\mathbf{x}, \mathbf{y})$  in more detail, where we assume throughout this section that  $f(y, d) = y - d$  and  $c(y, d) = p(d - y)^+ + h(y - d)^+$ . Let  $\{d_1, \dots, d_M\}$  be the set of all possible demand values, where  $d_m < d_{m+1}$ , for all  $m = 1, \dots, M - 1$ . Letting

$\sigma(d, \mathbf{x}) = \sum_z \sigma(d, z, \mathbf{x})$ , define for all  $m = 0, \dots, M$ ,

$$\begin{aligned} A_m(\mathbf{x}) &= h \sum_{k=1}^m \sigma(d_k, \mathbf{x}) - p \sum_{k=m+1}^M \sigma(d_k, \mathbf{x}), \\ B_m(\mathbf{x}) &= p \sum_{k=m+1}^M d_k \sigma(d_k, \mathbf{x}) - h \sum_{k=1}^m d_k \sigma(d_k, \mathbf{x}). \end{aligned}$$

Note,  $A_0(x) = -p$  and  $B_0(x) = p \sum_{k=1}^M d_k \sigma(d_k, \mathbf{x})$ . Proof of the next result, which provides structure that will prove useful, is straightforward.

**Lemma 1.** *For all  $\mathbf{x} \in X$ :*

$$(i) \quad L(\mathbf{x}, y) = \begin{cases} A_0(\mathbf{x})y + B_0(\mathbf{x}) = p \sum_{k=1}^M \sigma(d_k, \mathbf{x})(d_k - y), & y \leq d_1 \\ A_m(\mathbf{x})y + B_m(\mathbf{x}), & d_m \leq y \leq d_{m+1}, \quad m = 1, \dots, M-1 \\ A_M(\mathbf{x})y + B_M(\mathbf{x}) = h \sum_{k=1}^M \sigma(d_k, \mathbf{x})(y - d_k), & d_M \leq y. \end{cases}$$

(ii) *for all  $m = 1, \dots, M-1$ ,  $A_{m+1}(\mathbf{x}) = A_m(\mathbf{x}) + (h + p)\sigma(d_{m+1}, \mathbf{x})$ , and hence,*

$$A_{m+1}(\mathbf{x}) \geq A_m(\mathbf{x}).$$

(iii) *for all  $m = 1, \dots, M-1$ ,  $B_{m+1}(\mathbf{x}) = B_m(\mathbf{x}) - (p + h)d_{m+1}\sigma(d_{m+1}, \mathbf{x})$ , and hence,*

$$B_{m+1}(\mathbf{x}) \leq B_m(\mathbf{x}).$$

(iv) *for all  $m = 1, \dots, M$ ,  $A_{m-1}(\mathbf{x})d_m + B_{m-1}(\mathbf{x}) = A_m(\mathbf{x})d_m + B_m(\mathbf{x})$ .*

$$(v) \quad L(\mathbf{x}, y) = \max_{0 \leq m \leq M} \left[ A_m(\mathbf{x})y + B_m(\mathbf{x}) \right].$$

*Myopic Base Stock Policy: Linear Partition of Belief Space.*

Lemma 1 establishes that  $L(\mathbf{x}, y)$  is piecewise linear and convex in  $y$  for all  $\mathbf{x} \in X$ . Let  $s^*(\mathbf{x})$  be the smallest integer that minimizes  $L(\mathbf{x}, y)$  with respect to  $y$ . Note that it is sufficient to restrict  $s^*(\mathbf{x})$  to the set  $\{d_1, \dots, d_M\}$ . Hence,  $L(\mathbf{x}, s^*(\mathbf{x})) = \min_{1 \leq m \leq M} \left\{ A_m(\mathbf{x})d_m + B_m(\mathbf{x}) \right\}$ . Let  $\mathcal{P}_1$  be the partition of  $X$  composed of elements  $X_m = \{\mathbf{x} \in X : s^*(\mathbf{x}) =$

$d_m\}$ . Thus,  $\mathcal{P}_1 = \{X_m, m = 1, \dots, M\}$ , where  $X_m$  is non-null for all  $d_m$  such that  $\min\{s^*(\mathbf{x}) : \mathbf{x} \in X\} \leq d_m \leq \max\{s^*(\mathbf{x}) : \mathbf{x} \in X\}$ . We characterize  $X_m$  as follows.

**Lemma 2.** Let  $\mathbf{P}(d) = \sum_z \mathbf{P}(d, z)$ ,  $\forall d$ . For  $m = 1, \dots, M$ ,

$$X_m = \left\{ \mathbf{x} \in X : \mathbf{x} \sum_{k=1}^{m-1} \mathbf{P}(d_k) \underline{1} < p/(p+h) \leq \mathbf{x} \sum_{k=1}^m \mathbf{P}(d_k) \underline{1} \right\}. \quad (3.2)$$

Note that the criterion in Equation 3.2 can be re-written as:

$$\sum_{k=1}^{m-1} \sigma(d_k, \mathbf{x}) < p/(p+h) \leq \sum_{k=1}^m \sigma(d_k, \mathbf{x}),$$

where  $\sigma(d_k, \mathbf{x})$  is the probability of observing demand outcome  $d_k$  when the current belief is  $\mathbf{x}$ . This criterion is identical to the *newsvendor problem*'s criterion for determining the optimal base stock policy with the probability mass function of demand given by  $\sigma(d_k, x)$ ,  $\forall k$ .

Due to the linearity of  $\sigma(d, \mathbf{x})$  in  $\mathbf{x}$ , the above criterion results in a *linear partition* of the belief space. We note that the partition thus obtained is independent of the values of demand and AOD outcomes but depends only on the parameters,  $P_{ij}(d, z)$ ,  $p$ , and  $h$ . We remark that  $X_m$  for all  $m$  can be described by two inequalities linear in  $\mathbf{x}$ , which is true irrespective of the values  $N$  and  $M$  take, since  $L(\mathbf{x}, y)$  is piecewise linear in  $\mathbf{x}$  for fixed  $y$ .



**Example 1.** Let  $M = 7, N = 3, h = 1, p = 3, \mathbf{d} = [5, 10, 15, 20, 25, 30, 35]$ ,

$$\mathbf{P} = \begin{bmatrix} 0.0192 & 0.8744 & 0.1063 \\ 0.0437 & 0.4712 & 0.4851 \\ 0.4467 & 0.0313 & 0.522 \end{bmatrix}, \text{ and}$$

$$\mathbf{Q}^D = \begin{bmatrix} 0.0207 & 0.2321 & 0.0717 & 0.2054 & 0.1519 & 0.0346 & 0.2837 \\ 0.2697 & 0.208 & 0.2044 & 0.1942 & 0.0748 & 0.0427 & 0.0062 \\ 0.0283 & 0.0378 & 0.0429 & 0.0605 & 0.1335 & 0.3001 & 0.3969 \end{bmatrix}$$

where  $\mathbf{P} = \{P_{ij}\}$  and  $P_{ij} = \Pr(\mu(t+1) = j \mid \mu(t) = i)$ ,

$$\mathbf{Q}^D = \{q_{jd}^D\}, \quad q_{jd}^D = \Pr(d(t+1) = d \mid \mu(t+1) = j),$$

where  $q_{jd}$  is independent of  $i$ .

Note that  $P_{ij}(d) = P_{ij}q_{jd}^D$  is independent of  $z$ . The belief space is given by the triangle with vertices  $(1, 0, 0)$ ,  $(0, 1, 0)$ , and  $(0, 0, 1)$  (described by  $x_1 + x_2 + x_3 = 1$ ,  $x_1 \geq 0$ ,  $x_2 \geq 0$ , and  $x_3 \geq 0$ ), where modulation state  $n+1$  indicates a stronger economy than modulation state  $n$ , for all  $n$ . Figure 3.1 depicts the belief space,  $X$ , overlaid with the partition,  $\mathcal{P}_1$

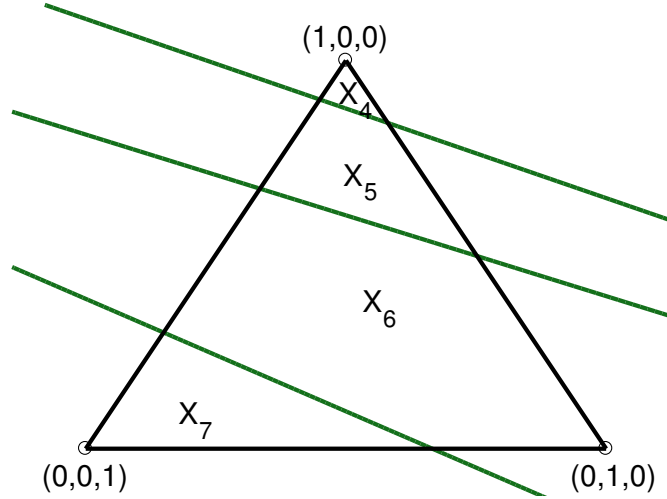


Figure 3.1: Example of  $\mathcal{P}_1$  with  $N = 3$  and  $M = 7$

(derived in Lemma 2).  $\mathcal{P}_1$  divides  $X$  into 4 regions of constant base stock level, viz.,  $X_4$

through  $X_7$ . For any belief vector in  $X_m$ , the optimal order-up-to level for the one period problem is  $d_m$ . Hence, the optimal myopic base stock levels are 20, 25, 30, and 35 in  $X_4$ ,  $X_5$ ,  $X_6$ , and  $X_7$  respectively.

If the AOD process is dependent on  $\mu(t+1)$  (e.g. current state of the economy), and has two outcomes  $\{z_1, z_2\}$  (e.g. real estate price levels) with  $\mathbf{R}^Z = \{r_{jz}^Z\}$ ,  $r_{jz}^Z = \Pr(z(t+1) = z \mid \mu(t+1) = j)$ , then  $P_{ij}(d, z) = P_{ij} q_{jd}^D r_{jz}^Z$ . The regions presented in Figure 3.1 do not change as Equation 3.2 does not depend on the values of the outcomes of either the AOD or demand processes. Note  $\lambda(d_1, e_3) = [0.28, 0.26, 0.46]$  when the AOD process is uninformative. Let  $\mathbf{R}^Z = [1, 0; 0.7, 0.3; 0, 1]$  and hence the AOD process is informative. Then,  $\lambda(d_1, z_1, e_3) = [0.61, 0.39, 0]$  and  $\lambda(d_1, z_2, e_3) = [0, 0.15, 0.85]$ . We note that the availability of additional observation data leads to substantially different updated belief functions.

### 3.2.4 Definition of A1

We now present a key assumption, A1.

**Assumption 1 (A1).**  $f(s^*(\mathbf{x}), d) \leq s^*(\boldsymbol{\lambda}(d, z, \mathbf{x}))$  for all  $\mathbf{x} \in X$ ,  $d \in \{d_1, \dots, d_M\}$ , and  $z$ .

Assuming  $f(s^*(\mathbf{x}), d)$  is the number of units of inventory after satisfying demand just after the current decision epoch,  $\boldsymbol{\lambda}(d, z, \mathbf{x})$  then becomes the belief function at the next decision epoch and  $s^*(\boldsymbol{\lambda}(d, z, \mathbf{x}))$  is the order-up-to level at the next decision epoch. A1 assumes that the amount of inventory after demand is satisfied never exceeds the order-up-to level at the next decision epoch. This assumption is always satisfied when demand is i.i.d. and is consistent with assumptions made in the inventory literature as early as [53, 52]. [53] provides a sufficiency condition for the optimality of myopic base stock levels, viz., non-decreasing base stock levels. [52] presents an attainability assumption that the remaining inventory in every period after placing an order and satisfying demand is less than the next period's base stock level under nonstationarity. A1 is a generalized attainability assumption

that we will show ensures the optimality of a myopic base stock policy for the general problem.

### 3.3 When A1 holds

#### 3.3.1 Main Result

We now present the main result of this section, which assumes throughout that A1 holds. Proof of the following result is provided in the appendix.

**Proposition 1.** *Assume A1 holds,  $L(\mathbf{x}, y)$  is piecewise linear in  $y$  for all  $\mathbf{x} \in X$ ,  $s^*(\mathbf{x})$  is the smallest integer that minimizes  $L(\mathbf{x}, y)$  with respect to  $y$ , and that  $f(y, d)$  is non-decreasing in  $y$  for each  $d$ . Then,  $v_n(\mathbf{x}, s) = v_n(\mathbf{x}, \max\{s^*(\mathbf{x}), s\})$  is non-decreasing and convex in  $s$  for all  $n$  and  $\mathbf{x}$ . Further, the myopic base stock policy that orders up to  $\max\{s^*(\mathbf{x}), s\}$  is an optimal policy.*

Thus, when A1 is satisfied and recalling that  $s^*(\mathbf{x})$  is determined using the inequalities presented in (3.2), ordering up to  $\max\{s, s^*(\mathbf{x})\}$  at every decision epoch is optimal for any finite horizon problem and the infinite horizon problem. This result ensures significantly less computational effort for computing the optimal policy compared to the procedure proposed in [76] for the special case where the modulation process is completely unobserved by the AOD process.

#### 3.3.2 A1 Analysis

Assuming A1, the above result ensures that there is an optimal policy that is a myopic base stock policy, where the order-up-to level is  $\max\{s^*(\mathbf{x}), s\}$ , given state  $(\mathbf{x}, s)$ . We note that  $s^*(\mathbf{x})$  is easily determined given the partition  $\mathcal{P}_1$ . We now examine conditions that imply A1. Proof of the following result is straightforward.

**Lemma 3.** *Assume A1 holds, apply the base stock policy “order up to  $\max\{s^*(\mathbf{x}), s\}$ ”, and assume  $s(t) \leq s^*(\mathbf{x}(t))$ . Then,  $s(\tau) \leq s^*(\mathbf{x}(\tau))$  for all  $\tau \geq t$ .*

Thus, once the inventory level falls at or below the base stock level, A1 guarantees that the inventory level will always fall at or below the base stock level at the next decision epoch.

We now present conditions that assure A1 holds. Let  $\mathcal{B}_0(\mathbf{x}) = \{\mathbf{x}\}$ ,  $\mathcal{B}_{n+1}(\mathbf{x}) = \{\boldsymbol{\lambda}(d, z, \mathbf{x}') : \mathbf{x}' \in \mathcal{B}_n(\mathbf{x}), d \in \{d_1, \dots, d_M\}, \forall z\}$ , and  $\mathcal{B}(\mathbf{x})$  be such that  $\mathcal{B}_n(\mathbf{x}) \subseteq \mathcal{B}(\mathbf{x})$  for all  $n$ .

**Assumption 2 (A2).** *There exists an  $m$  such that  $\mathcal{B}(\mathbf{x}) \subseteq X_m$ .*

Proof of the following result is straightforward.

**Lemma 4.** *Assume  $f(y, d) \leq y$  for all  $y$  and  $d$ . Then, A2 implies that the base stock level is stationary and hence A1 holds.*

We remark that if the modulation process is static and completely unobserved by both the demand and AOD processes and hence demand is i.i.d., then  $\mathbf{x}(t+1) = \mathbf{x}(t)$  for all  $t$ ,  $\mathcal{B}_n(\mathbf{x}) = \{\mathbf{x}\}$  for all  $n$ ,  $\mathcal{B}(\mathbf{x})$  can equal  $\{\mathbf{x}\}$ , and the base stock level is stationary. Let  $\mathbf{e}_m \in X$  be such that the  $m^{\text{th}}$  entry of  $\mathbf{e}_m$  is 1 and its remaining entries are zero. If  $\{\Pr(\mu(0) = \mu_i), \forall i\} = \mathbf{e}_m$  and the modulation process is static, then  $\mathbf{x}(t) = \mathbf{e}_m$  for all  $t$ ,  $\mathcal{B}_n(\mathbf{e}_m) = \{\mathbf{e}_m\}$  for all  $n$ ,  $\mathcal{B}(\mathbf{x})$  can equal  $\{\mathbf{e}_m\}$ , and once again the base stock level is stationary, irrespective of the demand and observation processes.

We now present a second set of conditions that imply A1 holds, following several preliminary results.

**Assumption 3 (A3).** *If  $i \leq i'$ , then  $\sum_{k=m}^M \sum_j P_{ij}(d_k) \leq \sum_{k=m}^M \sum_j P_{i'j}(d_k)$  for all  $m = 1, \dots, M$ .*

Define the binary operator for first order stochastic dominance,  $\preceq$ , as follows: for  $\mathbf{x}, \mathbf{x}' \in X$ ,  $\mathbf{x} \preceq \mathbf{x}' \iff \sum_{i=n}^N x_i \leq \sum_{i=n}^N x'_i \quad \forall n = 1, \dots, N$ .

**Lemma 5.** *Let A3 hold. Then,  $\mathbf{x} \preceq \mathbf{x}'$ , implies  $s^*(\mathbf{x}) \leq s^*(\mathbf{x}')$ .*

Let the modulation process represent the state of the economy, and assume the higher the state of the modulation process, the better the economy. Thus,  $e_1$  is the lowest performance level of the economy and  $e_N$  is the highest. Then, it is reasonable to assume that for higher performance levels of the economy, the probability of observing greater demand outcomes increases. Lemma 5 confirms the intuition that the optimal order quantities will be greater when the economy is performing better.

Let  $\mathbf{x}^{d,z} \in X$  be such that  $\mathbf{x}^{d,z} \preceq \boldsymbol{\lambda}(d, z, \mathbf{x}) \forall \mathbf{x} \in X$ . Existence is assured since  $e_1 \preceq \mathbf{x}$  for all  $\mathbf{x} \in X$ .

**Assumption 4 (A4).**  $f(s^*(e_N), d) \leq s^*(\mathbf{x}^{d,z})$  for all  $d$  and  $z$ .

**Lemma 6.** Assuming  $f(y, d)$  is non-decreasing in  $y$  for all  $d$ , A3 and A4 imply A1.

Ideally, we would want to select  $\mathbf{x}^{d,z}$  so that  $s^*(\mathbf{x}') \leq s^*(\mathbf{x}^{d,z})$  for all  $\mathbf{x}'$  such that  $\mathbf{x}' \preceq \boldsymbol{\lambda}(d, z, \mathbf{x}) \forall \mathbf{x} \in X$ , for all  $(d, z)$ , which would strengthen Lemma 6 as much as possible. We construct such an  $\mathbf{x}^{d,z}$  after the following preliminary result.

**Lemma 7.** The set  $\{\boldsymbol{\lambda}(d, z, \mathbf{x}) : \mathbf{x} \in X\} = \left\{ \sum_i \xi_i \boldsymbol{\lambda}(d, z, \mathbf{e}_i) : \xi_i \geq 0 \forall i, \sum_i \xi_i = 1 \right\}$ .

We remark that if  $\mathbf{x} \preceq \mathbf{x}'$  and  $\mathbf{x} \preceq \mathbf{x}''$ , then  $\mathbf{x} \preceq \alpha \mathbf{x}' + (1 - \alpha) \mathbf{x}''$  for all  $\alpha \in [0, 1]$ . Thus, if  $\mathbf{x}^{d,z}$  is such that  $\mathbf{x}^{d,z} \preceq \boldsymbol{\lambda}(d, z, \mathbf{e}_i)$  for all  $i$ , then  $\mathbf{x}^{d,z}$  is such that  $\mathbf{x}^{d,z} \preceq \mathbf{x}'$  for all  $\mathbf{x}' \in \{\boldsymbol{\lambda}(d, z, \mathbf{x}) : \mathbf{x} \in X\}$ .

We now construct  $\mathbf{x}^{d,z}$ . Let

$$\begin{aligned} \hat{x}_N^{d,z} &= \min \{ \lambda_N(d, z, \mathbf{e}_i), i = 1, \dots, N \} \\ \hat{x}_n^{d,z} &= \min \left\{ \sum_{k=n}^N \lambda_k(d, z, \mathbf{e}_i), i = 1, \dots, N \right\} - \sum_{k=n+1}^N \hat{x}_k^{d,z}, \\ &\quad \text{for } n = N - 1, \dots, 2 \\ \hat{x}_1^{d,z} &= 1 - \sum_{k=2}^N \hat{x}_k^{d,z}. \end{aligned}$$

By construction,  $\hat{\mathbf{x}}^{d,z} \preceq \boldsymbol{\lambda}(d, z, \mathbf{x}) \forall \mathbf{x} \in X$ . We now show that  $\hat{\mathbf{x}}^{d,z} \in X$  and that  $s^*(\mathbf{x}') \leq s^*(\hat{\mathbf{x}}^{d,z})$  for all  $\mathbf{x}' \in X$  such that  $\mathbf{x}' \preceq \boldsymbol{\lambda}(d, z, \mathbf{x}) \forall \mathbf{x} \in X$ .

**Lemma 8.** (i)  $\hat{x}^{d,z} \in X$ . (ii) For any  $x' \preceq \lambda(d, z, x) \forall x \in X$ ,  $s^*(x') \leq s^*(\hat{x}^{d,z})$ .

**Example 2.** Consider the problem in Example 1. As there is a stochastic ordering:  $e_1 \prec e_2 \prec e_3$ , A3 is satisfied. We remark that A4 can be verified by determining  $x^{d,z}$  using the process described above. Figure 3.2 shows  $\{\lambda(d_5, z, x) : x \in X\}$  and  $\hat{x}^{d_5,z}$ , where both  $\hat{x}^{d,z}$  and  $\lambda(d, z, x)$  are independent of  $z$  (e.g., where  $\Pr(z(t+1) \mid \mu(t+1), \mu(t))$  is independent of  $\mu(t+1)$  and  $\mu(t)$ ). Thus since Example 1 satisfies A1 by Lemma 6, ordering

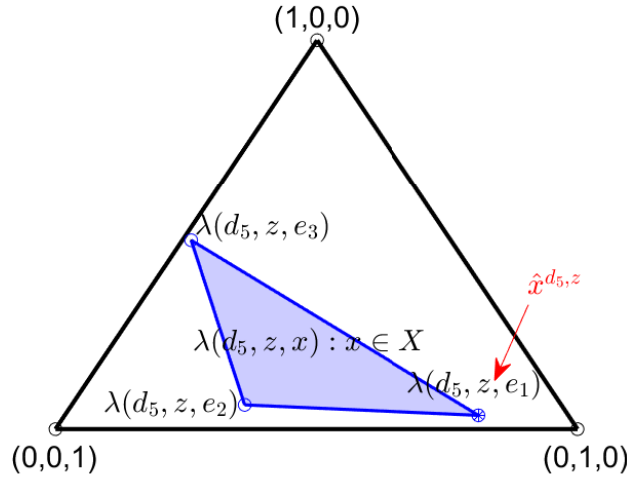


Figure 3.2:  $\hat{x}^{d_5,z}$  and  $\{\lambda(d_5, z, x) : x \in X\}$

up to the myopic base stock level at every decision epoch is optimal over finite and infinite horizons.

We remark that for the case where the modulation process is completely observed and  $f(y, d)$  is non-increasing in  $d$  for all  $y$ , A1 is equivalent to  $f(s^*(\mu(t)), d(t+1)) \leq s^*(\mu(t+1))$  for all  $d(t+1)$ , and hence  $f(s^*(\mu(t)), d_1) \leq s^*(\mu(t+1))$ . Note that this is equivalent to the attainability assumption presented by [52, 53] that guarantees the optimality of a myopic base stock policy for the completely observed nonstationary case.

### 3.3.3 Computing the Expected Cost Function, $v_n$

We now present a procedure for computing  $v_n(s, x)$ . We only consider the case where  $s = s^*(x)$  due to Proposition 1 and Lemma 3. For notational simplicity, we assume that

$\Pr(z(t+1) \mid \mu(t+1), \mu(t))$  is independent of  $\mu(t+1)$  and  $\mu(t)$ . Extension to the more general case is straightforward.

Assume  $v_0 = 0$ , let  $n = 1$ , and recall  $v_1(\mathbf{x}, s^*(\mathbf{x})) = L(\mathbf{x}, s^*(\mathbf{x}))$ . Note  $L(\mathbf{x}, y) = \mathbf{x}\bar{\gamma}_y$ , where  $\bar{\gamma}_y = \sum_{d,z} \mathbf{P}(d, z) \mathbf{1}_{c(y, d)}$ . Let  $\Gamma_1 = \{\bar{\gamma}_y\}$ , and note that if  $c(y, d) = p(d - y)^+ + h(y - d)^+$ , it is sufficient to consider only  $y \in \{d_1, \dots, d_M\}$ . Then,  $v_1(\mathbf{x}, s^*(\mathbf{x})) = \min \{\mathbf{x}\bar{\gamma} : \bar{\gamma} \in \Gamma_1\}$ . Assume there is a finite set  $\Gamma_n$  such that  $v_n(\mathbf{x}, s^*(\mathbf{x})) = \min \{\mathbf{x}\gamma : \gamma \in \Gamma_n\}$ . Then,

$$\begin{aligned}
v_{n+1}(\mathbf{x}, s^*(\mathbf{x})) &= L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{m=1}^M \sigma(d_m, \mathbf{x}) v_n(\boldsymbol{\lambda}(d_m, \mathbf{x}), f(s^*(\mathbf{x}), d_m)) \\
&= \min \{\mathbf{x}\bar{\gamma} : \bar{\gamma} \in \Gamma_1\} + \beta \sum_{m=1}^M \sigma(d_m, \mathbf{x}) v_n(\boldsymbol{\lambda}(d_m, \mathbf{x}), s^*(\boldsymbol{\lambda}(d_m, \mathbf{x}))) \\
&= \min \{\mathbf{x}\bar{\gamma} : \bar{\gamma} \in \Gamma_1\} + \beta \sum_{m=1}^M \sigma(d_m, \mathbf{x}) \min \{\boldsymbol{\lambda}(d_m, \mathbf{x})\gamma : \gamma \in \Gamma_n\} \\
&= \min_{\bar{\gamma}} \min_{\gamma_1} \dots \min_{\gamma_M} \left\{ \mathbf{x}\bar{\gamma} + \beta \sum_{m=1}^M \sigma(d_m, \mathbf{x}) \boldsymbol{\lambda}(d_m, \mathbf{x}) \gamma_m \right\} \\
&= \min_{\bar{\gamma}} \min_{\gamma_1} \dots \min_{\gamma_M} \left\{ \mathbf{x} \left[ \bar{\gamma} + \beta \sum_{m=1}^M \mathbf{P}(d_m) \gamma_m \right] \right\}
\end{aligned}$$

Thus,  $\Gamma_{n+1}$  is the set of all  $\gamma$  such that  $\gamma = \bar{\gamma} + \beta \sum_{m=1}^M \mathbf{P}(d_m) \gamma_m$ , where  $\bar{\gamma} \in \Gamma_1$  and  $\gamma_m \in \Gamma_n$  for all  $m = 1, \dots, M$ , and for all  $n$ ,  $v_n(\mathbf{x}, s^*(\mathbf{x}))$  is piecewise linear and concave in  $\mathbf{x}$ .

Let  $|\Gamma|$  be the cardinality of the set  $\Gamma$ . Then,  $|\Gamma_{n+1}| = |\Gamma_1| |\Gamma_n|^M$ , where  $|\Gamma_1| \leq M$ , and hence the cardinality of  $\Gamma_n$  can grow rapidly. Many of the vectors in the sets  $\Gamma_n$  are redundant and can be eliminated, reducing both computational and storage burdens. An exhaustive literature study of elimination procedures and other solution methods for solving POMDPs can be found in [101].

### 3.4 When A1 Does Not Hold

We now consider the case where A1 does not hold and examine the quality of the myopic policy “order up to  $\max\{s^*(\mathbf{x}), s\}$ ”, which from results in Section 3.3 we know is optimal if A1 does hold. We proceed by determining a lower bound on the optimal expected cost function and outlining a procedure for determining the expected cost of the “order up to  $\max\{s^*(\mathbf{x}), s\}$ ” policy. We note by example that this cost function may contain discontinuities and hence is not concave in  $\mathbf{x}$  and discuss the possible implications. Finally, we present a simple procedure for determining an upper bound on the difference between the expected cost function and the lower bound.

#### 3.4.1 A Lower Bound, $v^L$

We now present a lower bound on  $v_n(\mathbf{x}, s)$ . Let

$$\begin{aligned} [H^L v](\mathbf{x}, s) &= L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v(\boldsymbol{\lambda}(d, z, \mathbf{x}), s^*(\boldsymbol{\lambda}(d, z, \mathbf{x}))), \\ v_{n+1}^L &= H^L v_n^L, \quad v_0^L = 0, \end{aligned}$$

and  $v^L$  be the fixed point of  $H^L$ , which we note is independent of  $s$ .

**Proposition 2.** *For all  $\mathbf{x}$ ,  $s$ , and  $n$ ,  $v_n^L(\mathbf{x}) = v_n^L(\mathbf{x}, s^*(\mathbf{x})) \leq v_n(\mathbf{x}, s)$ .*

The proof follows from the fact that the controller always brings the inventory to  $s^*(\mathbf{x})$ , which is not feasible when the inventory is higher than  $s^*(\mathbf{x})$ . We remark that  $v_n^L(\mathbf{x})$  can be computed as was  $v_n(\mathbf{x}, s)$  for  $s \leq s^*(\mathbf{x})$ , in Section 3.3. A tighter lower bound, dependent on  $s$  for  $s > s^*(\mathbf{x})$ , would replace  $L(\mathbf{x}, s^*(\mathbf{x}))$  with  $L(\mathbf{x}, \max\{s^*(\mathbf{x}), s\})$  in the definition of the operator  $H^L$ . However, since such a definition of  $H^L$  would complicate later analysis, we have chosen not to use this tighter lower bound in the development of results in the sections to follow.



### 3.4.2 An Upper Bound, $v^U$

Let

$$\begin{aligned} [H^U v](\mathbf{x}, s) &= L(\mathbf{x}, \max\{s^*(\mathbf{x}), s\}) \\ &\quad + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(\max\{s^*(\mathbf{x}), s\}, d)), \end{aligned}$$

$v_{n+1}^U = H^U v_n^U$ ,  $v_0^U = 0$ , and let  $v^U$  be the fixed point of  $H^U$ . We remark that  $v^U$  is the expected cost to be accrued by the “order-up-to  $\max\{s^*(\mathbf{x}), s\}$ ” policy, which is feasible but may not be optimal when A1 is not satisfied, and hence represents an upper bound on the optimal cost function. It is straightforward to prove the following structural result.

**Proposition 3.** *For all  $n$  and  $\mathbf{x}$ ,  $v_n^U(\mathbf{x}, s) = v_n^U(\mathbf{x}, s^*(\mathbf{x}))$  for  $s \leq s^*(\mathbf{x})$ , and  $v_n^U(\mathbf{x}, s)$  is non-decreasing and convex in  $s$ .*

We show the following structural result in the appendix of this chapter.

**Lemma 9.** *For each  $n \geq 1$ , there is a partition  $\mathcal{P}_n$  of  $X$  that is defined by a finite set of linear inequalities such that on each element of this partition  $v_n^U$  is linear in  $\mathbf{x}$ . Further,  $\mathcal{P}_{n+1}$  is at least as fine as  $\mathcal{P}_n$  (i.e., if  $S \in \mathcal{P}_{n+1}$ , then there is an  $S' \in \mathcal{P}_n$  such that  $S \subseteq S'$ ).*

Thus,  $v_n(\mathbf{x}, s)$  is piecewise linear in  $\mathbf{x}$  for each  $s$ . Note that  $\mathcal{P}_1$  is identical to  $\mathcal{P}_1$  defined in Section 3.2.3. However, Example 3 shows that  $v_n^U(\mathbf{x}, s)$  may be discontinuous and hence not concave in  $\mathbf{x}$  for each  $s$ . Thus, according to [102], it may not be true that improved observation accuracy will improve the performance of the “order up to  $\max\{s^*(\mathbf{x}), s\}$ ” policy if A1 is not satisfied.

**Example 3.** *Assume  $f(y, d) = y - d$ ,  $c(y, d) = p(d - y)^+ + h(d - y)^+$ , and  $\beta = 0.9$ . Let*

$N = 2, M = 10, h = 1, p = 2, d = [0 \ 1 \ 2 \ 3 \ 4 \ 8 \ 12 \ 17 \ 18 \ 19],$

$$P = \begin{bmatrix} 0.4670 & 0.5330 \\ 0.4103 & 0.5897 \end{bmatrix}, \text{ and } Q = \begin{bmatrix} 0.1747 & 0.0115 \\ 0.01716 & 0.0278 \\ 0.1417 & 0.0537 \\ 0.1153 & 0.0611 \\ 0.1095 & 0.1012 \\ 0.0993 & 0.1176 \\ 0.0712 & 0.1215 \\ 0.0658 & 0.1612 \\ 0.0368 & 0.1667 \\ 0.0142 & 0.1777 \end{bmatrix}.$$

Then,  $\min_{\mathbf{x}} s^*(\mathbf{x}) = 12$  and  $\max_{\mathbf{x}} s^*(\mathbf{x}) = 17$ .

In Example 3, A1 does not hold. Figure 3.3 presents  $v_2^U(\mathbf{x}, s)$  and  $v_2^L(\mathbf{x}, s)$  for this example. We note the discontinuity in the expected cost function for two periods,  $v_2^U$ , obtained by implementing the myopic base stock policy when A1 does not hold.

We remark that although  $v_n^U(\mathbf{x}, s)$  is piecewise linear in  $\mathbf{x}$  for all  $s$  and  $n$ , in the limit as  $n$  approaches  $\infty$ , we may lose piecewise linearity. Thus, although implementing the policy “order up to  $\max\{s^*(\mathbf{x}), s\}$ ” is straightforward, determining  $v^U$ , or for that matter  $v_n^U$  for large  $n$ , is computationally demanding. For this reason, we seek an easily computable upper bound on  $v^U - v^L$  in the next section.

### 3.4.3 An Upper Bound on $v^U - v^L$

Let  $\Delta = \max_{(\mathbf{x}, s)} \left\{ L(\mathbf{x}, \max\{s^*(\mathbf{x}), s\}) - L(\mathbf{x}, s^*(\mathbf{x})) \right\}$ . We now present upper bounds on  $v_n^U(\mathbf{x}, s) - v_n^L(\mathbf{x})$  for all  $n$ . Let  $f(y, d) = y - d$  and  $c(y, d) = p(d - y)^+ + h(d - y)^+$ . Proof of Proposition 4 follows from a standard induction argument and the fact that the lower bound is independent of  $s$ .

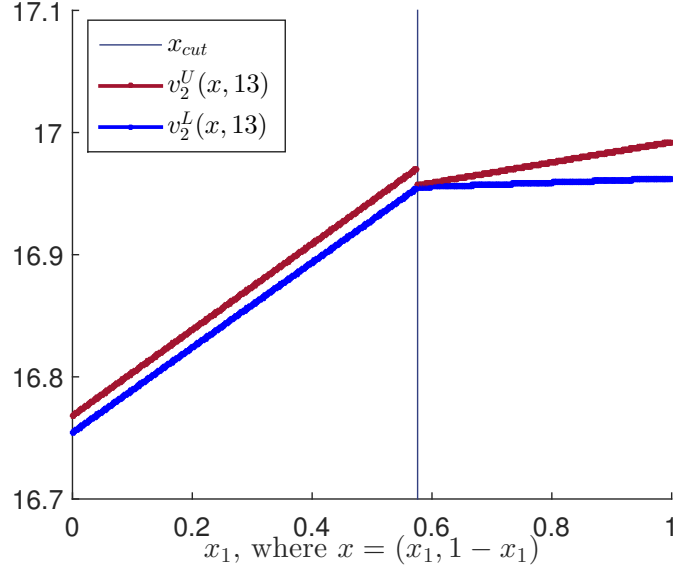


Figure 3.3:  $v_2^U(\mathbf{x}, s)$  and  $v_2^L(\mathbf{x})$  at  $s = 13$ .

**Proposition 4.** *For all  $\mathbf{x}$ ,  $s$ , and  $n$ ,*

$$v_n^U(\mathbf{x}, s) - v_n^L(\mathbf{x}) \leq (1 + \beta + \beta^2 + \dots + \beta^{n-1})\Delta$$

$$\text{and } v^U(\mathbf{x}, s) - v^L(\mathbf{x}) \leq \Delta/(1 - \beta).$$

We now determine  $\Delta$ . We only need to consider  $s$  such that  $s^*(\mathbf{x}) < s \leq \max_{\mathbf{x}} s^*(\mathbf{x}) - d_1$  and we only need to consider  $s^*(\mathbf{x})$  such that  $\min_{\mathbf{x}} s^*(\mathbf{x}) \leq s^*(\mathbf{x}) \leq \max_{\mathbf{x}} s^*(\mathbf{x}) - d_1$ . Hence,  $\Delta$  is finite. Let  $m$  and  $l$  be such that  $s^*(\mathbf{x}) = d_m$ ,  $s = d_l$  and  $d_m < d_l \leq \max_{\mathbf{x}} s^*(\mathbf{x}) - d_1$ . Then,

$$\begin{aligned} L(\mathbf{x}, \max\{s^*(\mathbf{x}), s\}) - L(\mathbf{x}, s^*(\mathbf{x})) &= L(\mathbf{x}, d_l) - L(\mathbf{x}, d_m) \\ &= A_l(\mathbf{x})d_l + B_l(\mathbf{x}) - A_m(\mathbf{x})d_m - B_m(\mathbf{x}), \end{aligned} \tag{3.3}$$

subject to the constraints  $\mathbf{x} \in X$  and

$$\sum_{k=1}^{m-1} \sigma(d_k, \mathbf{x}) < p/(p + h) \leq \sum_{k=1}^m \sigma(d_k, \mathbf{x}).$$

Note that all of the constraints on  $x$  are linear and that  $L(\mathbf{x}, d_l) - L(\mathbf{x}, d_m)$  is linear in  $\mathbf{x}$ .

Thus, maximizing (3.3) subject to the given constraints is a simple linear program. The number of LPs that require solution in order to completely determine  $\Delta$  is the number of  $(d_m, d_l)$  such that  $\min_{\mathbf{x}} s^*(\mathbf{x}) \leq d_m \leq \max_{\mathbf{x}} s^*(\mathbf{x}) - d_l$ , and  $d_m < d_l \leq \max_{\mathbf{x}} s^*(\mathbf{x}) - d_1$ .

**Example 4.** We now continue Example 3 and note:  $\Delta = 0.1398$ ,  $\max_{(\mathbf{x}, s)} (v_2^U(\mathbf{x}, s) - v_2^L(\mathbf{x})) = 1.3508$ , where the max occurs at  $\mathbf{x} = \mathbf{e}_1$ ,  $\Delta(1 + \beta) = 0.2656$  and hence,  $(v_2^U(\mathbf{x}, s) - v_2^L(\mathbf{x}))/v_2^L(\mathbf{x}) \leq \Delta(1 + \beta)/v_2^L(\mathbf{x}) = 0.2656/16.9624 = 0.0156$  at  $\mathbf{x} = \mathbf{e}_1$ . The term  $v_2^L(\mathbf{x})$  is computed by successively applying the operator  $H^L$  twice, using  $L(\mathbf{x}, s^*(\mathbf{x})) = \min_{1 \leq m \leq M} \{A_m(\mathbf{x})d_m + B_m(\mathbf{x})\}$  from Section 3.2.3. We obtain  $v_2^L(\mathbf{e}_1) = 16.9624$ . Note that the sub-optimal policy for  $n = 2$  is no more than 1.56% sub-optimal for  $\mathbf{x} = \mathbf{e}_1$ , indicating that this sub-optimal policy when A1 is violated may still be a high-quality heuristic.

#### 3.4.4 A Tighter Lower Bound, $v'$

We assume throughout this section that  $f(y, d) = y - d$  and  $c(y, d) = p(d - y)^+ + h(d - y)^+$ . When A1 does not hold, there will be a  $\delta$  such that

$$s^*(\mathbf{x}) - d \leq s^*(\lambda(d, z, \mathbf{x})) + \delta \quad (3.4)$$

for all  $d, z$  and  $\mathbf{x}$ . Assuming  $\delta$  satisfies Equation 3.4, we now consider a second problem, the primed problem. We show that the primed problem satisfies A1 and generates a tighter lower bound on the optimal cost function than the lower bound presented in Section 3.4.1.

Let  $P'_{ij}(d', z) = P_{ij}(d, z)$  for  $d' = d + \delta$ . Then,  $\sigma'(d', z, \mathbf{x}) = \sigma(d, z, \mathbf{x})$  and  $\lambda'(d', z, \mathbf{x}) = \lambda(d, z, \mathbf{x})$ . Let  $\sigma'(d', \mathbf{x}) = \sum_z \sigma'(d', z, \mathbf{x})$ . Define

$$\begin{aligned} A'_m(\mathbf{x}) &= h \sum_{k=1}^m \sigma'(d'_k, \mathbf{x}) - p \sum_{k=m+1}^M \sigma'(d'_k, \mathbf{x}), \\ B'_m(\mathbf{x}) &= p \sum_{k=m+1}^M d'_k \sigma'(d'_k, \mathbf{x}) - h \sum_{k=1}^m d'_k \sigma'(d'_k, \mathbf{x}), \end{aligned}$$

$$L'(\mathbf{x}, y') = \begin{cases} A'_0(\mathbf{x})y' + B'_0(\mathbf{x}), & y' \leq d'_1 \\ A'_m(\mathbf{x})y' + B'_m(\mathbf{x}), & d'_m \leq y' \leq d'_{m+1}, \quad m = 1, \dots, M-1 \\ A'_M(\mathbf{x})y' + B'_M(\mathbf{x}), & d'_M \leq y'. \end{cases}$$

Proof of the following result is straightforward.

**Lemma 10.** For  $d'_m = d_m + \delta$  and  $y' = y + \delta$ ,

(i)  $A'_m(\mathbf{x})d'_m + B'_m(\mathbf{x}) = A_m(\mathbf{x})d_m + B_m$  for all  $m$ .

(ii)  $L'(\mathbf{x}, y') = L(\mathbf{x}, y)$ .

We define  $s^{*'}(\mathbf{x})$  as follows:  $s^{*'}(\mathbf{x}) = d'_m$  if  $A'_m(\mathbf{x})d'_m + B'_m(\mathbf{x}) < A'_{m-1}(\mathbf{x})d'_{m-1} + B'_{m-1}(\mathbf{x})$  and  $A'_m(\mathbf{x})d'_m + B'_m(\mathbf{x}) \leq A'_{m+1}(\mathbf{x})d'_{m+1} + B'_{m+1}(\mathbf{x})$ . Lemma 10 then implies the following result.

**Lemma 11.** For all  $\mathbf{x}$  and  $m$  where  $d'_m = d_m + \delta$  and  $d' = d + \delta$ ,

(i)  $\{\mathbf{x} : s^*(\mathbf{x}) = d_m\} = \{\mathbf{x} : s^{*'}(\mathbf{x}) = d'_m\}$

(ii)  $s^{*'}(\mathbf{x}) - d' \leq s^{*'}(\lambda'(d', z, \mathbf{x}))$ .

Thus, the transformation has resulted in a problem that satisfies A1.

Define the operator  $H'$  as follows:

$$[H'v'](\mathbf{x}, s') = \min_{y' \geq s'} \left\{ L'(\mathbf{x}, y') + \beta \sum_{d', z} \sigma'(d', z, \mathbf{x}) v'(\boldsymbol{\lambda}'(d', z, \mathbf{x}), y' - d') \right\}.$$

We now present the main result of this section, where  $v'$  is the fixed point of  $H'$  and  $v'_{n+1} = H'v'_n$  for all  $n$ .

**Proposition 5.** Assume  $v_0^L(\mathbf{x}) = v'_0(\mathbf{x}, s) = v_0(\mathbf{x}, s) = 0$  for all  $\mathbf{x}$  and  $s$ . Then for all  $n$ ,  $v_n^L(\mathbf{x}) \leq v'_n(\mathbf{x}, s) \leq v_n(\mathbf{x}, s)$  and hence  $v^L(\mathbf{x}) \leq v'(\mathbf{x}, s) \leq v(\mathbf{x}, s)$  for all  $\mathbf{x}$  and  $s$ .

It is straightforward to show that  $v'$  is non-increasing in  $\delta$ ; hence, as  $\delta$  increases, the lower bound  $v'$  becomes weaker and can be shown to converge to  $v^L$ . Thus, there is incentive to choose  $\delta$  as small as possible to satisfy A1 in constructing a lower bound on  $v$ .

### 3.5 Reorder Cost Case

We now consider the case where there is a reorder cost  $K \geq 0$ , and assume throughout this section that A1 holds. The following results combine the ideas presented for the  $K = 0$  case with straightforward extensions of earlier results in the literature. Let the operator  $H^K$  be defined as:

$$[H^K v](\mathbf{x}, s) = \min_{y \geq s} \left\{ K\xi(y - s) + [Gv](\mathbf{x}, y) \right\},$$

where  $\xi(k) = 0$  if  $k = 0$  and  $\xi(k) = 1$  if  $k \neq 0$  and

$$[Gv](\mathbf{x}, y) = L(\mathbf{x}, y) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(y, d)).$$

We note that when  $K = 0$ ,  $H^K = H$ , as defined in Section 3.2.2.

We now assume that  $K > 0$ . Our objective is to present conditions under which  $(s, S)$  policies exist and how such policies can be computed.

#### 3.5.1 $K$ -convexity and Optimal $(s, S)$ Policies

We now present our first result following a key definition: the real-valued function  $g$  is  $K$ -convex if for any  $s \leq s'$ ,

$$g(\lambda s + (1 - \lambda)s') \leq \lambda g(s) + (1 - \lambda)(g(s') + K), \text{ for all } \lambda \in [0, 1].$$

Proof of the following result is a direct extension of results in [77] and elsewhere.

**Proposition 6.** Assume: (i)  $v(\mathbf{x}, s)$  is  $K$ -convex in  $s$  for all  $\mathbf{x}$ , (ii)  $f(y, d)$  is non-decreasing and convex in  $y$  for all  $d$ , and (iii)  $c(y, d)$  is convex in  $y$  and  $\lim_{|y| \rightarrow \infty} c(y, d) \rightarrow \infty$  for all  $d$ . Then,

1.  $[Gv](\mathbf{x}, y)$  is  $K$ -convex in  $y$  for all  $\mathbf{x}$ ,

2.  $[H^K v](\mathbf{x}, s)$  is  $K$ -convex in  $s$  for all  $\mathbf{x}$ , and

$$3. [H^K v](\mathbf{x}, s) = \begin{cases} K + [Gv](\mathbf{x}, S^*(\mathbf{x}, v)) & s \leq s^*(\mathbf{x}, v) \\ [Gv](\mathbf{x}, s) & \text{otherwise,} \end{cases}$$

where:  $S^*(\mathbf{x}, v)$  is the smallest integer minimizing  $[Gv](\mathbf{x}, y)$  with respect to  $y$ , and

$s^*(\mathbf{x}, v)$  is the smallest integer such that  $[Gv](\mathbf{x}, s^*(\mathbf{x}, v)) \leq K + [Gv](\mathbf{x}, S^*(\mathbf{x}, v))$ .

Thus, the fact that  $v(\mathbf{x}, s)$  is  $K$ -convex and non-decreasing in  $s$  for all  $\mathbf{x}$  leads to the existence of an optimal policy that is of  $(s, S)$  form: if the inventory drops below  $s$ , then order up to  $S$ ; otherwise, do not replenish.

### 3.5.2 Bounds on $s_n$ and $S_n$

Let  $v_0 = 0$ ,  $v_{n+1} = H^K v_n$  for all  $n \geq 0$ , and  $G_n(\mathbf{x}, y) = [Gv_n](\mathbf{x}, y)$ . Let  $S_n(\mathbf{x})$  be the smallest integer such that  $G_n(\mathbf{x}, S_n(\mathbf{x})) \leq G_n(\mathbf{x}, y)$  for all  $y$ , and let  $s_n(\mathbf{x})$  be the smallest integer such that  $G_n(\mathbf{x}, s_n(\mathbf{x})) \leq K + G_n(\mathbf{x}, S_n(\mathbf{x}))$ . Following [79], we now define four real-valued functions that represent bounds on the set  $\{(s_n(\mathbf{x}), S_n(\mathbf{x})) : n \geq 0\}$ . Let the values  $\underline{s}(\mathbf{x})$ ,  $\bar{s}(\mathbf{x})$ ,  $\underline{S}(\mathbf{x})$ , and  $\bar{S}(\mathbf{x})$  be the smallest integers such that:

$$L(\mathbf{x}, \underline{S}(\mathbf{x})) \leq L(\mathbf{x}, y) \quad \forall y \quad (3.5)$$

$$L(\mathbf{x}, \underline{s}(\mathbf{x})) \leq K + L(\mathbf{x}, \underline{S}(\mathbf{x})) \quad (3.6)$$

$$\beta K + L(\mathbf{x}, \underline{S}(\mathbf{x})) \leq L(\mathbf{x}, \bar{S}(\mathbf{x})), \quad \bar{S}(\mathbf{x}) \geq \underline{S}(\mathbf{x}) \quad (3.7)$$

$$L(\mathbf{x}, \bar{s}(\mathbf{x})) \leq L(\mathbf{x}, \underline{S}(\mathbf{x})) + (1 - \beta)K, \quad (3.8)$$

where, from earlier results,  $\underline{S}(\mathbf{x})$  can be restricted to the set  $\{d_1, \dots, d_M\}$  and where  $\underline{S}$  is identical to the functions  $s^*$  and  $S_0$ . We remark that the convexity of  $L(\mathbf{x}, y)$  in  $y$  for all  $\mathbf{x}$  insures that for each  $\mathbf{x}$ ,  $\underline{s}(\mathbf{x}) \leq \bar{s}(\mathbf{x}) \leq \underline{S}(\mathbf{x}) \leq \bar{S}(\mathbf{x})$ .

### 3.5.3 A Partition based on

Extending results in [79], we now show that for all  $\mathbf{x}$  and  $n$ ,  $\underline{s}(\mathbf{x}) \leq s_n(\mathbf{x}) \leq \bar{s}(\mathbf{x})$  and  $\underline{S}(\mathbf{x}) \leq S_n(\mathbf{x}) \leq \bar{S}(\mathbf{x})$  and that for the infinite horizon discounted case,  $\underline{s}(\mathbf{x}) \leq s^*(\mathbf{x}) \leq \bar{s}(\mathbf{x})$  and  $\underline{S}(\mathbf{x}) \leq S^*(\mathbf{x}) \leq \bar{S}(\mathbf{x})$ , where  $(s^*, S^*)$  represents an  $(s, S)$  belief-dependent optimal policy. Proof is presented in the appendix of this chapter.

**Proposition 7. (a)** *For the  $n$ -period problem, for all  $\mathbf{x}$ , there exists an optimal  $(s, S)$  policy  $(s_n(\mathbf{x}), S_n(\mathbf{x}))$ , where  $\underline{s}(\mathbf{x}) \leq s_n(\mathbf{x}) \leq \bar{s}(\mathbf{x}) \leq \underline{S}(\mathbf{x}) \leq S_n(\mathbf{x}) \leq \bar{S}(\mathbf{x})$ .*

**(b)** *For the infinite horizon problem, for all  $\mathbf{x}$  there is an epoch-invariant  $(s, S)$  policy  $(s^*(\mathbf{x}), S^*(\mathbf{x}))$ , where  $\underline{s}(\mathbf{x}) \leq s^*(\mathbf{x}) \leq \bar{s}(\mathbf{x}) \leq \underline{S}(\mathbf{x}) \leq S^*(\mathbf{x}) \leq \bar{S}(\mathbf{x})$ .*

Assume  $f(y, d) = y - d$  and  $c(y, d) = p(d - y)^+ + h(d - y)^+$  and recall from Lemma 2 that  $S^*(\mathbf{x}) = d_m$  if  $\mathbf{x}$  satisfies

$$\mathbf{x} \sum_{k=1}^{m-1} \mathbf{P}(d_k) \mathbf{1} < \frac{p}{p+h} \leq \mathbf{x} \sum_{k=1}^m \mathbf{P}(d_k) \mathbf{1}. \quad (3.9)$$

Given  $\underline{S}(\mathbf{x}) = d_m$ , let  $\underline{s}(\mathbf{x}) = d_i$ ,  $\bar{s}(\mathbf{x}) = d_j$ , and  $\bar{S}(\mathbf{x}) = d_n$  satisfy

$$A_i(\mathbf{x})d_i + B_i(\mathbf{x}) \leq K + A_m(\mathbf{x})d_m + B_m(\mathbf{x}) \quad (3.10)$$

$$A_j(\mathbf{x})d_j + B_j(\mathbf{x}) \leq (1 - \beta)K + A_m(\mathbf{x})d_m + B_m(\mathbf{x}) \quad (3.11)$$

$$\beta K + A_m(\mathbf{x})d_m + B_m(\mathbf{x}) \leq A_n(\mathbf{x})d_n + B_n(\mathbf{x}). \quad (3.12)$$

Let  $\bar{X}(\underline{s}, \bar{s}, \underline{S}, \bar{S})$  be the set of all  $\mathbf{x} \in X$  such that  $\underline{s} = d_i$ ,  $\bar{s} = d_j$ ,  $\underline{S} = d_m$ , and  $\bar{S} = d_n$  are the smallest integers satisfying the five linear inequalities in Eqs. 3.9 - 3.12. Note that the set of all  $\bar{X}(\underline{s}, \bar{s}, \underline{S}, \bar{S})$  such that  $\bar{X}(\underline{s}, \bar{s}, \underline{S}, \bar{S})$  is non-null is a partition of  $X$ .



**Example 5.** Consider Example 1 with reorder cost,  $K = 5$ . Each region in the triangle

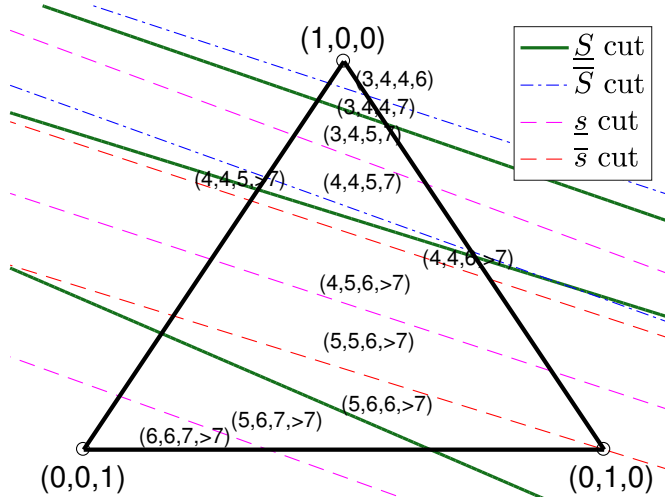


Figure 3.4: Example of partition of  $X$ , with  $N = 3$  and  $M = 7$

is described by  $(i, j, m, n)$  where,  $(\underline{s}, \bar{s}, \underline{S}, \bar{S}) = (d(i), d(j), d(m), d(n))$ . For example, the region labeled as  $(4, 4, 5, > 7)$  in Figure 3.4 has  $(\underline{s}, \bar{s}, \underline{S}, \bar{S}) = (20, 20, 25, 36)$ . This implies that  $s^*(\mathbf{x}) = d_4 = 20, \forall \mathbf{x} \in (4, 4, 5, > 7)$ . The search interval for  $S^*(\mathbf{x})$  is also significantly restricted to the demand outcomes between  $\underline{S}$  and  $\bar{S}$ , making the computation very easy. We remark that it is possible  $\bar{S} > d_M$ , as indicated (by  $> 7$ ) in Figure 3.4. The corresponding  $\bar{S}$  is 36 in  $X_5$  and  $X_6$  and it equals 38 in  $X_7$ .

A description of the determination of the sets  $\Gamma_n(s)$ , where  $v_n(\mathbf{x}, s) = \min\{\mathbf{x}\gamma : \gamma \in \Gamma_n(s)\}$ , can be found in the appendix of this chapter.

### 3.6 Conclusions

We have presented and analyzed an inventory control problem having a modulation process that affects demand and that is partially observed by the demand and AOD processes. The demand and AOD processes inform the decision maker (DM) about the state of the modulation process and hence inform the DM regarding future demand. We modeled the problem as a POMDP assuming that the DM knows the current number of items in the inventory and

the current belief function of the modulation process prior to making a replenishment decision. Current and past demand and AOD data are used to construct and update the belief function. This model was shown to generalize several of the Markov-modulated demand and the Bayesian updating models in the literature.

Assuming A1 holds and the reorder cost  $K = 0$ , we generalized results found throughout the literature that there exists an optimal policy that is a myopic base stock policy. We also developed a simple, easily implemented description of the optimal myopic base stock levels, as a function of the belief function. We determined conditions that imply A1 holds and an algorithm for computing the expected cost function.

When A1 is violated and  $K = 0$ , we examined the base stock policy that is optimal when A1 holds as a suboptimal policy and used the expected cost accrued by this suboptimal policy as an upper bound on the optimal expected cost function. We presented a lower bound on the optimal expected cost function and a bound on the difference between the upper and lower bounds. An example indicated that the bound on the difference between these two bounds can be quite small, indicating that even when A1 is violated, the optimal base stock policy for the case where A1 is not violated may be quite good. A thorough numerical investigation of the quality of this policy when A1 is violated is a topic for future research. We then presented a tighter lower bound that assumed A1 holds within a  $\delta > 0$  and showed that this tighter lower bound improves as  $\delta$  gets smaller.

When  $K > 0$  and A1 holds, we showed that there exists optimal  $(s, S)$  policies, dependent on the belief function, and determined upper and lower bounds on  $s$  and  $S$  for the finite and infinite horizon cases, where each bound is dependent on the belief function of the modulation process. We showed that each of these bounds and the values of  $s$  and  $S$  for the finite and infinite horizon cases are constant within regions of the belief space and that these regions can be described by a finite number of linear inequalities. We outlined an approach for determining an optimal  $(s, S)$  policy and the resultant expected cost function for the finite horizon case.

## CHAPTER 4

### DATA-DRIVEN CONTROL OF DISTRIBUTED RECONFIGURABLE PRODUCTION - INVENTORY SYSTEMS

#### 4.1 Introduction

We investigate a multi-period, multi-location production-inventory system under stochastic demand that allows backlogging, assumes instantaneous replenishment, and has the capability to relocate transportable production units and/or transship inventory between locations. Historically, transshipment has been a tool to reposition inventory in order to improve supply chain performance. We now add the capability of repositioning production capacity to further aid in improving the performance of a supply chain. Transportable production units, which we refer to as *modules*, have recently generated significant interest in manufacturing [6, 103, 3, 4]; we remark that manufacturing and/or storing the final, or near-final, product close to demand can enable fast fulfillment. The aim of this chapter is to help answer such questions as: (i) when, how much, and to where inventory and/or transportable production capacity should be relocated? (ii) what replenishment decisions should be made in coordination with this capability to relocate inventory and/or production capacity? Our intent is that this research will lead to the design of reconfigurable supply chains that share the advantages of centralized supply chain systems - having reduced buffer stock and reduced capital investments and expenditures, relative to distributed systems - while providing the fast fulfillment of distributed systems positioned in demand-dense geographical areas.

We develop and analyze a partially observed Markov decision process (POMDP) model for this system, propose efficient heuristics that determine replenishment decisions, when to transship and/or relocate production capacity, and hence determine the value of having

the capability to transship and relocate production capacity. The objective of the model is to minimize the expected total discounted cost criterion composed of backorder, holding, transshipment, and module relocation costs.

More specifically, we consider a distributed production - inventory system with  $L$  locations and  $Y$  transportable production modules. None, one, or more of the modules are located at each of the locations. At each decision epoch, we assume the (centralized) decision maker (DM) knows the current demand forecast, inventory level, and production capacity at each location. This production capacity is made up of fixed capacity and transportable capacity. The DM decides how the current inventory and transportable production capacity should be relocated. We assume these relocations occur instantaneously. Once the inventory and transportable production capacity have been relocated, the DM determines the replenishment decisions at each location based on current demand forecasts, the new inventory levels, and the new production capacities at the locations. Replenishment is instantaneous. Once replenishment is complete, demands at the locations are realized. Based on these realizations and possibly other data, the demand forecast is updated just before the next decision epoch.

Our model of data-driven demand forecasting assumes the existence of a stochastic process, the *modulation* process, that affects demand. The modulation process is governed by a Markov chain and is partially observed by the demand process and another process, the *additional observation data* process. The modulation process can model exogenous factors, such as current macro-economic conditions, the weather, and seasonal effects that can affect the demand process. Realizations of the observation process may provide additional data useful for understanding the current state of the modulation process, e.g., interest rates, unemployment rates, consumer price indices. We assume that the current belief function of the modulation process influences the current demand forecast.

#### 4.1.1 Literature Review

The problem considered in this chapter involves inventory transshipment, mobile capacity relocation, finite production capacity of each single location production facility, and a centralized controller determining transshipment, module relocation, and replenishment decisions.

Numerous innovative developments in manufacturing, such as containerized production for pharmaceutical manufacturing processes [103, 3, 4] and on-demand mobile production [6] necessitate the planning of logistics for flexible production and inventory systems that are characterized by resource mobility, interconnectivity, sharing, and decentralization [18]. [104] investigate the dynamic mobile production and inventory problem without the option of inventory transshipment under stationary and independent demands and have proposed heuristic approaches to solve the problem. A value addition of more than 10% over in-the-ground production systems was determined for a system of than twenty locations. [12] present a real options pricing based method of evaluating the value added by mobile containerized production systems. Other research that address the operational logistics of mobile facilities can be found in [22, 26].

Regarding multi-location inventory management with transshipment, [105, 106] considers the multi-location inventory control problem over a single period and multiple periods, respectively, under uncertain demands. It is proved that when the inventory addition and subtraction matrix has a Leontief structure, there exists a base stock policy that is optimal when attainable. In [107], a restricted Lagrangean dual -based lower bound and a dual relaxation based upper bound on the optimal cost of the multi-location problem are presented. The upper bound assumes the post ordering and shipment inventory position does not fall below the initial inventory position. [108] indicate that localized transshipment strategies are outperformed by centralized strategies. [109] propose heuristics for a problem that considers inventory held at a warehouse and allocated for distribution to various locations in a centralized fashion. [110] prove the optimality of order-up-to policies

at each location in a multi-location inventory control system with reactive transshipment for a long-run average cost criterion and present a heuristic for computation. The authors consider only replenishment decisions that result in non-negative inventory positions post replenishment at each location. [111] present a comparison of chain and group configurations of transshipment network design building on the ideas of manufacturing process flexibility [112] and restricted connectivity in a transshipment network [113]. [114] consider a multi-retailer, one warehouse framework that allows reactive transshipment either from the warehouse to the retailers and/or between retailers. The authors prove that it is optimal to adopt either retailer only, or warehouse only, or retailer first, or warehouse first protocols only, when considering transshipment. Various cost parameter thresholds based intervals are presented to indicate the system that is optimal in each regime.

We consider the data-driven online learning demand model presented in [115] and adopt it for the multi-location problem in this chapter. [115] analyze a single location, infinite capacity inventory control problem with demand and additional observation data influenced solely by a Markov *modulation* process. The modulation process is intended to model forces that may be partially observed, influence the demand process, but are not affected by actions taken by the DM (e.g., the macro-economy, air currents, tides). Demand realizations and other data (e.g., housing starts, consumer spending) represent observations of the modulation process. What is known to the DM about the modulation process is provided by the belief function, which is updated with new data using Bayes Rule. A base stock policy, having a base stock level dependent on the belief function, is proved to be optimal for the infinite horizon problem when an attainability assumption holds. The modulation process can be used to model the correlation between demands at different locations.

We consider approximate dynamic programming approaches that do not rely on maintaining the cost function's lookup table over the entire horizon to find good heuristic solutions to the multi-location mobile capacity and inventory control problem [116, 33, 30, 29, 28]. In particular, we are interested in rollout based heuristics which are known to perform

well on dynamic systems with stochasticity as suggested in [30] for solving the vehicle routing problem with stochastic demands. [29] provide a systematic classification-aimed analysis of rollout policies. Additionally, the literature suggests that centralized control is expected to perform better than decentralized control from a solution-quality perspective; however, there is an inherent tradeoff between solution quality and computational expense [117, 118, 119]. In the current chapter, as determining a centralized control policy is computationally intensive, we propose a decentralized control policy that performs comparably with the centralized control policy.

#### 4.1.2 chapter Outline

This chapter is organized as follows. Section 4.2 presents the problem definition, formulation as a partially observed Markov decision process and some preliminary results. In Section 4.3, we present bounds on the optimal cost function based on the single location optimal cost function and propose a bounds blending based estimate of the optimal cost function. We then propose two approaches, namely joint control and global control, in Section 4.4 for solving the problem. After presenting methods for tractable estimation of the single location cost function in Section 4.5, we propose our heuristics and benchmarks in Section 4.6. We then present a computational analysis on multilocation problems in Section 4.7 and finally, conclude the chapter in Section 4.8.

## **4.2 Problem Statement and Preliminary Results**

We now define the general  $L$  location,  $Y$  module problem statement in Section 4.2.1 and present the partially observed Markov decision process (POMDP) model of this problem and general results for the model in Section 4.2.2. Section 4.6 then presents promising heuristic approaches to the general problem; a decentralized control heuristic where there is a Global Controller (GC) that determines the inventory and module relation decisions, leaving local replenishment decisions up to the individual locations. This heuristic is based

on the solution of the  $L = 1, Y = 0$  problem, for which we provide initial results in Section 4.5.

#### 4.2.1 Problem Statement

Consider a distributed production-inventory system with  $L$  locations and  $Y$  portable manufacturing modules. At each decision epoch  $t$  we assume the decision-maker (DM) knows:

- $\mathbf{s}(t) = \{s_l(t), l = 1, \dots, L\}$ , where  $s_l(t)$  is the inventory level at location  $l$ ,
- $\mathbf{u}(t) = \{u_l(t), l = 1, \dots, L\}$ , where  $u_l(t) \in \{0, 1, \dots, Y_l'\}$  is the number of modules positioned at location  $l$ ,
- $\mathcal{I}(t) = \{\mathbf{d}(t), \dots, \mathbf{d}(1), \mathbf{z}(t), \dots, \mathbf{z}(1), \mathbf{x}(0)\}$ , where:
  - $d_l(t)$  is the demand realized during period  $(t - 1, t)$  that location  $l$  is required to fulfill (or back order) and  $\mathbf{d}(t) = \{d_l(t), l = 1, \dots, L\}$
  - $\mathbf{z}(t)$  represents data, in addition to the realization of demand, that might be of use to the DM,
  - $\mathbf{x}(0)$  is an a priori probability vector defined below.

We assume the *demand* process  $\{\mathbf{d}(t), t = 1, 2, \dots\}$  and *additional observation data* process  $\{\mathbf{z}(t), t = 1, 2, \dots\}$  are linked to the *modulation* process  $\{\mu(t), t = 0, 1, \dots\}$  through the given conditional probability  $P(\mathbf{d}(t + 1), \mathbf{z}(t + 1), \mu(t + 1) \mid \mu(t))$ , where  $\mathbf{x}(0) = \{x_i(0), \forall i\}$  for  $x_i(0) = P(\mu(0) = \mu_i)$ . A discussion of this general description of data-driven demand and learning and how it generalizes and extends the Markov-modulated demand and Bayesian updating literatures can be found in [115].

The chronology of events within period  $(t, t + 1)$  is as follows:

Step 1. Given  $\mathcal{I}(t)$ ,  $\mathbf{s}(t)$ , and  $\mathbf{u}(t)$ , the DM relocates inventory and modules to reach the post-movement state  $(\mathbf{s}'(t), \mathbf{u}'(t))$ , where we assume  $\sum_l s'_l(t) = \sum_l s_l(t)$  and  $\sum_l u'_l(t) = \sum_l u_l(t)$ . Necessarily,  $-(s_l(t))^+ \leq \Delta_l^S(t) \leq \sum_{k \neq l} (s_k(t))^+$ , where



$\Delta_l^S(t)$  is the non-negative amount of inventory relocated to location  $l$ . Thus,  $s'_l(t) = s_l(t) + \Delta_l^S(t)$  for all  $l$  and hence  $\mathbf{s}'(t) = \mathbf{s}(t) + \mathbf{\Delta}^S(t)$ , where  $\mathbf{\Delta}^S(t) = \{\Delta_l^S(t), l = 1, \dots, L\}$ . The decision variables are  $\mathbf{\Delta}^S(t)$  and  $\mathbf{u}'(t)$  for Step 1.

Step 2. Given  $\mathcal{I}(t)$ ,  $\mathbf{s}'(t)$ , and  $\mathbf{u}'(t)$ , the DM determines  $\mathbf{q}(t) = \{q_l(t), l = 1, \dots, L\}$ , where  $q_l(t)$  is the replenishment decision at location  $l$ . Necessarily,  $0 \leq q_l(t) \leq U_l + u'_l(t)G$ , where  $U_l$  is the fixed amount capacity at location  $l$  and  $G$  is the capacity of each module. Let  $y_l(t) = s'_l(t) + q_l(t)$ , the inventory level at location  $l$  after inventory and module relocation and replenishment but before demand realization, and assume  $\mathbf{y}(t) = \{y_l(t), l = 1, \dots, L\}$ . The decision variables are therefore  $\mathbf{q}(t)$ , or equivalently  $\mathbf{y}(t)$ , for Step 2, where necessarily,  $s'_l(t) \leq y_l(t) \leq s'_l(t) + U_l + u'_l(t)G$  for all  $l$ .

Step 3. The realizations of the random variables  $\mathbf{d}(t+1)$  and  $\mathbf{z}(t+1)$  become known and unfulfilled demands are backordered,  $\mathcal{I}(t+1) = \{\mathbf{d}(t+1), \mathbf{z}(t+1), \mathcal{I}(t)\}$ ,  $\mathbf{s}(t+1) = \mathbf{y}(t) - \mathbf{d}(t+1)$ , and  $\mathbf{u}(t+1) = \mathbf{u}'(t)$ .

Step 4.  $t = t + 1$ .

We assume that for location  $l$ ,  $c_l(y_l(t), d_l(t+1)) = b_l(d_l(t+1) - y_l(t+1))^+ + h_l(y_l(t+1) - d_l(t+1))^+ \geq 0$  is the single period cost accrued between  $t$  and  $t+1$ , where  $b_l$  and  $h_l$  are respectively the backorder and holding cost per unit per period and for all  $d_l$ ,  $c_l(y_l, d_l)$  is convex in  $y_l$  and  $\lim_{|y| \rightarrow \infty} c_l(y, d_l) = \infty$ .

We assume that the modulation and the observation state spaces are finite and that for each location, the demand state space is finite and the inventory state space  $\{\dots, -1, 0, 1, \dots\}$  is countable.

Let the single period  $(t, t+1)$  cost be:

$$\sum_l (K_l^{S+}(\Delta_l^S(t))^+ + K_l^{S-}(-\Delta_l^S(t))^+) + K^M \sum_l |u_l(t) - u'_l(t)|/2 + \sum_l c_l(y_l(t), d_l(t+1)),$$

where  $K_l^{S+}$  ( $K_l^{S-}$ ) is the cost of moving a unit of inventory to (from) location  $l$ , and  $K^M$  is the cost of moving a module from one location to another. A *feasible* policy determines  $(\mathbf{q}(t), \Delta^S(t), \mathbf{u}'(t))$  based on  $\mathcal{I}(t)$ ,  $\mathbf{s}(t)$ , and  $\mathbf{u}(t)$  for all  $t$ .

The problem criterion is the expected total discounted cost over the infinite horizon, where  $\beta \in [0, 1)$  is the discount factor. The problem is to determine a feasible policy that minimizes the criterion with respect to the set of all feasible policies.

#### 4.2.2 POMDP Model and General Results

This problem can be recast as a partially observed Markov decision process (POMDP) as follows. Results in [97] and [98] imply that  $(\mathbf{x}(t), \mathbf{s}(t), \mathbf{u}(t))$  is a sufficient statistic, where the *belief* function  $\mathbf{x}(t) = \{x_i(t), \forall i\}$  is such that  $x_i(t) = P(\mu(t) = \mu_i \mid \mathcal{I}(t))$  and  $\mathbf{x}(t) \in X = \{\mathbf{x} \geq 0 : \sum_i x_i = 1\}$ . Let

$$\begin{aligned} P_{ij}(\mathbf{d}, \mathbf{z}) &= P(\mathbf{d}(t+1) = \mathbf{d}, \mathbf{z}(t+1) = \mathbf{z}, \mu(t+1) = j \mid \mu(t) = i) \\ \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) &= \mathbf{x}P(\mathbf{d}, \mathbf{z})\mathbf{1} = \sum_i x_i \sum_j P_{ij}(\mathbf{d}, \mathbf{z}) \\ \lambda(\mathbf{d}, \mathbf{z}, \mathbf{x}) &= \{\lambda_j(\mathbf{d}, \mathbf{z}, \mathbf{x}), \forall j\} = \mathbf{x}P(\mathbf{d}, \mathbf{z})/\sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}), \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) \neq 0 \\ L(\mathbf{x}, \mathbf{y}) &= E[c(\mathbf{y}, \mathbf{d})] = \sum_{\mathbf{d}, \mathbf{z}} \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) c(\mathbf{y}, \mathbf{d}), \quad c(\mathbf{y}, \mathbf{d}) = \sum_l c_l(y_l, d_l). \end{aligned}$$

Thus, if  $\mathbf{x}$  is the prior belief function, then  $\lambda(\mathbf{d}, \mathbf{z}, \mathbf{x})$  is the posterior belief function, given realizations  $(\mathbf{d}, \mathbf{z})$ , and  $\sigma(\mathbf{d}, \mathbf{z}, \mathbf{x})$  is the probability that  $(\mathbf{d}, \mathbf{z})$  will be the demand and observation realizations, given prior  $\mathbf{x}$ . Define the operator  $H$  as follows:

$$[Hv](\mathbf{x}, \mathbf{s}, \mathbf{u}) = \min_{\Delta^S, \mathbf{u}', \mathbf{y}} \{G(\mathbf{x}, \mathbf{y}, v)\}, \quad \text{where} \quad (4.1)$$

$$\begin{aligned} G(\mathbf{x}, \mathbf{y}, v) &= \sum_l (K_l^{S+}(\Delta_l^S)^+ + K_l^{S-}(-\Delta_l^S)^+) + K^M \sum_l |u_l - u'_l|/2 + L(\mathbf{x}, \mathbf{y}) \\ &\quad + \beta \sum_{\mathbf{d}, \mathbf{z}} \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) v(\lambda(\mathbf{d}, \mathbf{z}, \mathbf{x}), \mathbf{y} - \mathbf{d}, \mathbf{u}') \end{aligned}$$

and where the minimization is with respect to

$$\begin{aligned}
& \sum_l u'_l = Y \\
& 0 \leq u'_l \leq Y'_l, \forall l \\
& \sum_l \Delta_l^S = 0, \\
& -(s_l)^+ \leq \Delta_l^S \leq \sum_{k \neq l} (s_k)^+, \forall l \\
& (s_l + \Delta_l^S) \leq y_l \leq (s_l + \Delta_l^S) + U_l + u'_l G, \forall l \\
& u'_l, \Delta_l^S, y_l \in \mathbb{Z}, \forall l
\end{aligned}$$

Results in [99] guarantee that there exists a unique  $v^*$  such that  $v^* = Hv^*$  and that this fixed point is the minimum expected total discounted cost over the infinite horizon. Further, a policy that causes the minimum in (4.1) to be attained is an optimal policy and is decision epoch invariant. For any given bounded function  $v_0$ , let  $\{v_n\}$  be such that  $v_{n+1} = Hv_n$ . Then,  $\lim_{n \rightarrow \infty} \|v^* - v_n\| = 0$ , where  $\|\cdot\|$  is the sup-norm.

Results in [97] guarantee that  $v_n(\mathbf{x}, \mathbf{s}, \mathbf{u})$  is piecewise linear and concave in  $\mathbf{x}$  for fixed  $(\mathbf{s}, \mathbf{u})$  for all  $n$ , assuming  $v_0(\mathbf{x}, \mathbf{s}, \mathbf{u})$  is also piecewise linear and concave in  $\mathbf{x}$  for fixed  $(\mathbf{s}, \mathbf{u})$ . In the limit,  $v^*(\mathbf{x}, \mathbf{s}, \mathbf{u})$  may no longer be piecewise linear in  $\mathbf{x}$  for fixed  $(\mathbf{s}, \mathbf{u})$ ; however, concavity will be preserved. Further, results in [115] extended to the finite capacity case indicate that improved observation quality of the modulation process through the demand and observation processes will result in a fixed point no greater than the fixed point associated with the original quality of observation.

#### 4.2.3 Complexity Analysis

We recall that the number of multiplications per successive approximations iteration of a completely observed MDP is  $|S|^2|A|$ , where  $S$  and  $A$  are the state and action spaces, respectively, and  $|W|$  is the cardinality of the set  $W$ ,  $W \in \{S, A\}$ . Assuming we approximate

the set of all inventory levels of location  $l$  with the finite set  $S_l$  and let  $A_l(s_l, u'_l) = \{y_l : s_l \leq y_l \leq s_l + U_l + u'_l G\}$ , then for the  $L$  location,  $Y$  module problem where demand is i.i.d. (e.g., the modulation process is completely observed and static), each successive approximation step requires  $\left(\prod_{l=1}^L |S_l|\right)^2 \prod_{l=1}^L (U_l + u'_l G + 1)$  multiplications. Let  $|S_l| = 50$ ,  $|A_l| = (U_l + u'_l G + 1) = 50$  for all  $l$ , and  $L = 10$ , assuming we only consider replenishment decisions. Then, this number of multiplications is in the order of  $10^{31}$  when the number of locations is ten, making use of successive approximations intractable. Therefore, we seek good sub-optimal approaches that significantly reduce this computational burden. Solving each of the  $L$  local replenishment problems for the i.i.d. case requires  $|S_l|^2 |A_l|$  multiplications per successive approximation iteration, and  $L$  of these are required. For  $L = 10$  and  $|S_l| = |A_l| = 50$ ,  $L|S_l|^2 |A_l|$  is in the order of  $10^5$ , which is a large but computationally manageable problem.

### 4.3 Bounds and Approximate Value Function Based on $L = 1$ Case

One sub-optimal approach, which we now consider, is to base heuristics on the most tractable problem, the single location inventory control problem, i.e., the  $L = 1, Y = 0$  case. Then, the operator  $H$  simplifies to  $H_l^F$  for location  $l$ , where

$$[H_l^F v_l^F](\mathbf{x}, s_l, u_l) = \min \{G_l^F(\mathbf{x}, y_l, v_l)\},$$

$$G_l^F(\mathbf{x}, y_l, v_l^F) = L_l^F(\mathbf{x}, y_l) + \beta \sum_{\mathbf{d}, \mathbf{z}} \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) v_l^F(\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x}), y_l - d_l)$$

and where the minimization is with respect to  $s_l \leq y_l \leq s_l + U_l + u_l G$ . In the appendix section C1, we will show that the fixed point of  $H_l^F$ ,  $v_l^F$ , is non-decreasing in capacity for fixed  $(\mathbf{x}, s_l)$ . This monotonicity result implies

$$\sum_l v_l^F(\mathbf{x}, s_l, Y_l') \leq v(\mathbf{x}, \mathbf{s}, \mathbf{u}).$$

Further, since the local controllers assume that there is no inventory transshipment or module relocation in the future,

$$v(\mathbf{x}, \mathbf{s}, \mathbf{u}) \leq \sum_l v_l^F(\mathbf{x}, s_l, u_l).$$

Thus, the solutions of the local replenishment problems provide upper and lower bounds on the cost function of the initial problem.

We now present a blending approach to approximate the optimal cost-to-go function of the POMDP presented in (4.1). Let  $\theta \in [0, 1]$ , be such that

$$v_l^{F,\theta}(\mathbf{x}, s_l, u_l) = (1 - \theta)v_l^F(\mathbf{x}, s_l, Y_l') + \theta v_l^F(\mathbf{x}, s_l, u_l) \text{ and } \tilde{v}^\theta(\mathbf{x}, \mathbf{s}, \mathbf{u}) = \sum_l v_l^{F,\theta}(\mathbf{x}, s_l, u_l).$$

Hence,  $\tilde{v}^\theta(\mathbf{x}, \mathbf{s}, \mathbf{u})$  is an approximation of  $v(\mathbf{x}, \mathbf{s}, \mathbf{u})$  that relies solely on the solution of the single location ( $L = 1, Y = 0$ ) problem. Then,

$$\begin{aligned} [H\tilde{v}^\theta](\mathbf{x}, \mathbf{s}, \mathbf{u}) = & \min_{\Delta^S, \mathbf{u}'} \left\{ \sum_l (K_l^{S+}(\Delta_l^S)^+ + K_l^{S-}(-\Delta_l^S)^+) + K^M \sum_l |u_l - u_l'|/2 \right. \\ & \left. + \min_{\mathbf{y}} \{L(\mathbf{x}, \mathbf{y}) + \beta \sum_{\mathbf{d}, \mathbf{z}} \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) \tilde{v}^\theta(\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x}), \mathbf{y} - \mathbf{d}, \mathbf{u}')\} \right\} \end{aligned} \quad (4.2)$$

In (4.2), the inner minimization is over all  $y_l$  such that  $s_l + \Delta_l^S \leq y_l \leq s_l + \Delta_l^S + U_l + u_l' G$ .

#### 4.4 Joint Control and Global-Local Control

A straightforward modification of (4.2) leads to the *joint control (JC) problem*:

$$\begin{aligned}
\text{JC: } \quad & \min_{\Delta^S, u', y} \sum_l \left\{ (K_l^{S+}(\Delta_l^S)^+ + K_l^{S-}(-\Delta_l^S)^+) + K^M \sum_l |u_l - u'_l|/2 \right. \\
& \left. + \sum_{d, z} \sigma(d, z, \mathbf{x}) \left[ c_l(y_l, d_l) + \beta v_l^{F, \theta}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y_l - d_l, u'_l) \right] \right\}, \\
& \text{subject to} \\
& \sum_l u'_l = Y \\
& \sum_l \Delta_l^S = 0, \\
& 0 \leq u'_l \leq Y'_l, \forall l \\
& -(s_l)^+ \leq \Delta_l^S \leq \sum_{k \neq l} (s_k)^+, \forall l \\
& (s_l + \Delta_l^S) \leq y_l \leq (s_l + \Delta_l^S) + U_l + u'_l G, \forall l \\
& u'_l, \Delta_l^S, y_l \in \mathbb{Z}, \forall l
\end{aligned} \tag{4.3}$$

We refer to this problem as the JC problem since the DM jointly determines inventory and production capacity relocations and replenishment decisions at all of the locations. We propose a distributed decision-making structure in which all the relocation decisions are made globally while replenishment decisions are made at the individual locations. In (4.2), consider the inner minimization and note the terms in the inner brackets are bounded below by

$$\begin{aligned}
& \sum_l \left[ (1 - \theta) \min_{y_l} \left\{ L_l(x, y_l) + \beta \sum_{d, z} \sigma(d, z, \mathbf{x}) v_l^F(\boldsymbol{\lambda}(d, z, \mathbf{x}), y_l - d_l, Y'_l) \right\} \right. \\
& \left. + \theta \min_{y_l} \left\{ L_l(\mathbf{x}, y_l) + \beta \sum_{d, z} \sigma(d, z, \mathbf{x}) v_l^F(\boldsymbol{\lambda}(d, z, \mathbf{x}), y_l - d_l, u'_l) \right\} \right],
\end{aligned} \tag{4.4}$$

where the first minimization is now relaxed to operate over all  $y_l$  such that  $s_l + \Delta_l^S \leq y_l \leq$

$s_l + \Delta_l^S + U_l + Y_l'G$  and the second minimization is over all  $y_l$  such that  $s_l + \Delta_l^S \leq y_l \leq s_l + \Delta_l^S + U_l + u_l'G$ . We note that the terms in (4.4) equal

$$\sum_l \left[ (1 - \theta)v_l^F(\mathbf{x}, s_l + \Delta_l^S, Y_l') + \theta v_l^F(\mathbf{x}, s_l + \Delta_l^S, u_l') \right],$$

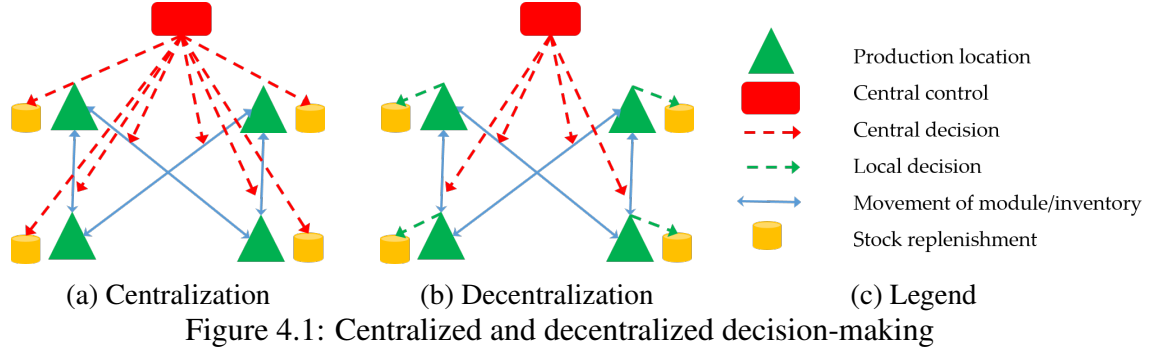
and hence, the fixed point of the operator  $\tilde{H}$ , evaluated at  $(\mathbf{x}, \mathbf{s}, \mathbf{u})$ , can be approximated by the *global control (GC) problem*:

$$\begin{aligned} \text{GC:} \quad & \min_{\Delta^S, \mathbf{u}'} \left\{ \sum_l (K_l^{S+}(\Delta_l^S)^+ + K_l^{S-}(-\Delta_l^S)^+) + K^M \sum_l |u_l - u_l'|/2 \right. \\ & \left. + \sum_l \left[ (1 - \theta)v_l^F(\mathbf{x}, s_l + \Delta_l^S, Y_l') + \theta v_l^F(\mathbf{x}, s_l + \Delta_l^S, u_l') \right] \right\}, \\ & \text{subject to} \\ & \sum_l u_l' = Y \\ & 0 \leq u_l' \leq Y_l', \forall l \\ & \sum_l \Delta_l^S = 0, \\ & -(s_l)^+ \leq \Delta_l^S \leq \sum_{k \neq l} (s_k)^+, \forall l \\ & u_l', \Delta_l^S \in \mathbb{Z}, \forall l \end{aligned} \tag{4.5}$$

We now present the following decentralized control (global-local) design.

Assume there is a global controller (GC) that selects  $(\Delta^S(t), \mathbf{u}'(t))$ , based on  $(\mathbf{x}(t), \mathbf{s}(t), \mathbf{u}(t))$ . The GC assumes that at each location  $l$  there is a local controller that:

1. Selects  $a_l(t)$  based on  $(\mathbf{x}(t), s_l'(t), u_l'(t))$  in order to minimize the infinite horizon discounted total cost criterion with single period  $(\Delta^S, \Delta^S + 1)$  cost  $c_l(s_l(\Delta^S), d_l(\Delta^S + 1))$  for all  $\Delta^S \geq t$ , given initial state  $s_l'(t)$ .
2. Assumes that there will be no inventory or module relocation in the future and hence has capacity  $U_l + u_l'(t)G$  over the infinite planning horizon.



Thus, the local controllers do not attempt to coordinate their replenishment decisions. All coordination is left up to the GC. We note, however, that the GC and the local controllers know the belief function and hence share the same information about the state of the modulation process.

These assumptions reduce the overall system coordination problem to determining  $(\Delta_l^S(t) u_l'(t))$  for all  $l$ , given  $(\mathbf{x}(t), \mathbf{s}(t), \mathbf{u}(t))$ , which eliminates determining the  $L$  location replenishment decisions at the system level, leaving these decisions up to the local controllers.

After determining the inventory transshipment and module movement quantities by solving the GC problem, a replenishment order must be placed for each location. We propose to replenish at each location up to the optimal myopic base stock level of the uncapacitated single location inventory control problem when demand is Markov-modulated with a partially observed modulation process [115], while satisfying the production capacity constraint. From [115], we recall that the optimal myopic base stock level  $s_l^*(\mathbf{x}) = d_m^l$ , the  $m$ -th demand outcome at location  $l$ , if and only if the following modified newsvendor criterion is satisfied:

$$\sum_{k=1}^{m-1} \sigma(d_k^l, \mathbf{x}) < b_l / (b_l + h_l) \leq \sum_{k=1}^m \sigma(d_k^l, \mathbf{x}),$$

where  $\sigma(d_k^l, \mathbf{x}) = \sum_{\mathbf{z}} \sigma(d_k^l, \mathbf{z}, \mathbf{x})$  is the probability of observing demand outcome  $d_k$  when the current belief is  $\mathbf{x}$ . We remark that the additional observation data vector,  $\mathbf{z}$ , can store the demand realization information from other locations as well. Hence, the local order up



to level at location  $l$  is given by

$$\hat{y}_l = \min \left\{ \max\{s_l^*(\mathbf{x}), s_l + \Delta_l^S\}, s_l + \Delta_l^S + U_l + u_l'G \right\}. \quad (4.6)$$

We have thus far investigated a decentralized approach for determining a good heuristic for the  $L$  location,  $Y$  module problem that involves a two-step approach for decision determination: (1) determine the solution to the finite capacity  $L = 1, Y = 0$  problem and (2) then use this solution to solve the GC problem. This decentralized approach has produced a computationally tractable approximation of  $\tilde{v}(\mathbf{x}, \mathbf{s}, \mathbf{u})$  in Section 4.2.2.

We also remark that the decentralized approach has variations that might prove of interest. For example, our current approach assumes each location does not coordinate replenishment decisions with other locations. An extension of the ‘totally local’ decision making assumption is for small numbers of the locations that are closely located geographically to coordinate their replenishment decisions (a ‘semi-global’ case). Further discussion can be found in [120].

## 4.5 Solving the $L = 1$ Case

To develop heuristics based on JC and GC approaches, we require the cost-to-go function of the fixed system (that is expressed as a sum of  $L$  location-wise cost-to-go functions) at every decision epoch. However, the cost-to-go function of the capacitated single location system is intractable due to curse of dimensionality. We propose the following approximations of it to allow developing decision support for large realistic systems.

### 4.5.1 A stationary approximation of $v_l^F$

We now present a tractable approximation of the value function of the  $L = 1$  problem having general modulated demand for all values of the belief function, noting that solution of the general  $L = 1$  problem is considerably more computationally complex than the

i.i.d. case mentioned above. Our intent is to fully define the GC problem presented in (4.5) using the approximation of  $v_l^F(\mathbf{x}, s_l, u_l)$  presented above and the approximation of the  $L = 1$  case that we now develop. Our approximation is based on an  $L = 1$  problem (i) having the tractability of its i.i.d. special case for a given fixed belief function that (ii) shares the structural properties of the optimal value function for the general modulated demand case and (iii) is defined over all values of the belief function. We proceed as follows. With respect to (i), assume that the modulation process has no dynamics and is completely unobserved, and hence  $\mathbf{x}(t+1) = \mathbf{x}(t)$  for all  $t$ . Given this assumption, the  $L = 1$  operator becomes

$$\begin{aligned} [\hat{H}_l^F \hat{v}_l^F](\mathbf{x}, s_l, u_l) = & \min_{s_l \leq y_l \leq s_l + U_l + G u_l} \left\{ \sum_{d_l} \sum_i x_i \Pr(d_l | i) [c_l(y_l, d_l) \right. \\ & \left. + \beta \hat{v}_l^F(\mathbf{x}, y_l - d_l, u_l)] \right\} \end{aligned} \quad (4.7)$$

which for given  $\mathbf{x}$  and  $u_l$ , requires essentially the same number of operations per successive approximations step as required in the i.i.d. case. With respect to (ii), results presented in the appendix insure that for the  $L = 1$  problem having general modulated demand, there exists an optimal policy that is a base stock policy, an optimal base stock level is non-increasing in capacity, and the optimal value function is non-increasing in capacity and convex in inventory level. Since the case where the modulation process has no dynamics and is completely unobserved is a special case of this problem, the fixed point of  $\hat{H}_l^F \hat{v}_l^F = \hat{v}_l^F$  inherits all of these characteristics. Requirement (iii) is satisfied by virtue of the definition of the approximation.

We now present a result that bounds the gap between  $\hat{v}_l^F$  and  $v_l^F$ .

**Proposition 1.** *We have  $v_l^F(\mathbf{x}, s_l, u_l) \geq \hat{v}_l^F(\mathbf{x}, s_l, u_l) - \rho/(1 - \beta)$  for all  $\mathbf{x}$ ,  $s_l$ , and  $u_l$ , where  $\rho = \sum_{d_l} k(d_l) c_l(\hat{y}_l, d_l)$  and  $k(d_l) = (\max_k \Pr(d_l | k) - \min_k \Pr(d_l | k))$ .*

The proof of this proposition is presented in the appendix section C2.

#### 4.5.2 A piecewise linear and convex approximation of $\widehat{v}_l^F$

We use the following approximation of  $\widehat{v}_l^F$ , drawing inspiration from the approximation of the cost-to-go function in the lookahead of fixed future (LAF) heuristic in [104].

$$\begin{aligned} \widehat{v}_l^F(\mathbf{x}(t+1), s_l(t+1), u_l(t+1)) &\approx (\widehat{v}_l^F(\mathbf{x}(t+1), \bar{s}_l(t+1), u_l(t)) \\ &\quad + \widehat{v}_l^F(\mathbf{x}(t+1), s_l(t), u_l(t+1))) / 2, \end{aligned}$$

where  $\bar{s}_l(t+1) = y_l(t) - \lceil E[D_l(t)] \rceil$  and  $\lceil a \rceil$  denotes the rounded  $a$ .

Since  $v_l^F(\mathbf{x}, s_l, u_l)$  is piecewise linear and convex in  $s_l$  when  $u_l$  is held constant and in  $u_l$  when  $s_l$  is held constant (from Proposition 3 and Proposition 6) and  $\widehat{v}_l^F(\mathbf{x}, s_l, u_l)$  inherits these properties as it is a stationary special case, the latter can be represented as  $\max\{\gamma_j^l s_l + \widehat{\gamma}_j^l : (\gamma_j^l, \widehat{\gamma}_j^l) \in \Gamma_t^l(u_l)\}$ ,  $\forall l \in \{1, \dots, L\}$  and as  $\max\{\theta_j^l u_l + \widehat{\theta}_j^l : (\theta_j^l, \widehat{\theta}_j^l) \in \Theta_t^l(s_l)\}$ ,  $\forall l \in \{1, \dots, L\}$ . The set  $\Gamma_t^l(u_l)$  ( $\Theta_t^l(s_l)$ ) is the set of coefficients describing the facets of the piecewise linear and convex function  $\widehat{v}_l^F(\mathbf{x}, s_l, u_l)$ , when  $u_l$  ( $s_l$ ) is held constant at time  $t$ .

The following expression is the proposed approximation:

$$\begin{aligned} &\widehat{v}_l^F(\mathbf{x}(t+1), s_l(t+1), u_l(t+1)) \\ &\approx \left( \max\{\gamma_j^l s_l + \widehat{\gamma}_j^l : (\gamma_j^l, \widehat{\gamma}_j^l) \in \Gamma_t^l(u_l(t))\} \right. \\ &\quad \left. + \max\{\theta_j^l u_l + \widehat{\theta}_j^l : (\theta_j^l, \widehat{\theta}_j^l) \in \Theta_t^l(s_l(t))\} \right) / 2. \end{aligned} \tag{4.8}$$

## 4.6 Heuristics

In this section, we take the next step toward the development of a computable heuristic by combining (4.1) and the bounds presented in Section 4.3. We now present all the heuristics we implement in our computational study. We begin with the naive solution method of determining dynamic decisions myopically. This is followed by a description of the policy

that does not consider inventory and module relocation. We then propose our heuristics resulting from the JC and the GC approaches.

#### 4.6.1 Myopic Policy (MP)

For the myopic policy (MP), the decision-maker optimizes over the one period cost to determine relocation and replenishment decisions. At every decision epoch with current state  $(\mathbf{x}, \mathbf{s}, \mathbf{u})$ , we solve the following integer program:

$$\begin{aligned}
\text{MP: } \quad & \min_{\Delta^{S, \mathbf{u}', \mathbf{y}}} \sum_l \left\{ (K_l^{S+} \Delta_l^{S+} + K_l^{S-} \Delta_l^{S-}) + K^M \sum_l |u_l - u'_l|/2 + \right. \\
& \left. + \beta \sum_n \sigma(d_l^n, \mathbf{x}) [h_l r_l^n + b_l o_l^n] \right\} \\
& \text{subject to } \sum_l u'_l = Y \\
& 0 \leq u'_l \leq Y'_l, \forall l \\
& \sum_l \Delta_l^{S+} = \sum_l \Delta_l^{S-}, \\
& 0 \leq \Delta_l^{S+} \leq \sum_{k \neq l} (s_k)^+, \quad 0 \leq \Delta_l^{S-} \leq -(s_l)^+, \forall l \\
& (s_l + \Delta_l^{S+} - \Delta_l^{S-}) \leq y_l \leq (s_l + \Delta_l^{S+} - \Delta_l^{S-}) + U_l + u'_l G, \forall l \\
& r_l^n \geq y_l - d_l^n, \quad o_l^n \geq d_l^n - y_l, \forall l, n \\
& r_l^n, o_l^n \in \mathbb{Z}^+, u'_l, \Delta_l^{S+}, \Delta_l^{S-}, y_l \in \mathbb{Z}, \quad \eta_l, \zeta_l \in \mathbb{R} \quad \forall l
\end{aligned} \tag{4.9}$$

MP accounts for transshipment quantities entering and leaving each location  $l$  as  $\Delta_l^{S+}$  and  $\Delta_l^{S-}$  respectively, the post module movement capacity count as  $u'_l$ , the post-replenishment inventory position as  $y_l$ , and the held and backlogged inventory quantities as  $r_l^n$ , and  $o_l^n$  for the  $n$ th demand scenario. Following the flow balance constraints for modules and inventory are the inventory accounting constraints.

The decision-making at multiple locations is dynamically coupled through the decision-maker's perception of the modulation state  $\mathbf{x}$ . As a natural consequence, the potential next

period's belief state will require consideration of all possible demand vectors over all the coupled locations, which is computationally intensive. However, we can circumvent this challenge by assuming that at each location, only the demand possibilities at that location, *i.e.*, local information, would affect the belief update. This is equivalent to assuming that the location-wise Markov-modulated demand random variables in this model are *conditionally mutually independent*, given the current belief of the modulation state. We note that this assumption will enable the implementation of JR and LAJ heuristics (presented below), although such an assumption is not required for executing the global-local control approach based GLR heuristic (also presented below).

#### 4.6.2 No Flexibility Policy (NF)

The No Flexibility policy (NF), our benchmark policy, does not permit inventory and module relocation, assumes that local replenishment is based on the policy presented in (4.6), and assumes that the fixed, static production capacities at the locations are selected in order to minimize the multi-location expected total cost with epoch-invariant steady state belief-based demand distributions. The steady state belief is given by the row matrix  $\pi = \pi \hat{P}$ , where  $\hat{P}$  is transition probability matrix of the modulation process. The corresponding static demand distribution at each location  $l$  is  $\pi O^l$  where  $O^l$  is the matrix  $\{\{O_{jk}^l\}\}$  such that  $O_{jk}^l$  is probability of observing demand outcome  $d_k$  at location  $l$  given that the modulation state is  $j$ .

#### 4.6.3 Joint Rollout of Stationary Future (JR)

The joint rollout of stationary future (JR) is based on the JC approach. In JR, at each decision epoch with current state  $(x, s, u)$ , the following integer program must be solved. In the integer program JR presented below, as the future cost term  $v_t^\theta$  is obtained from a lookup table and is a nonlinear expression, we adopt the following formulation that uses binary variables  $w(l, \Delta^S, \Delta^M, a)$  to choose the actions at the current epoch: transshipment

quantity  $\Delta^S$  entering location  $l$ , the number of modules  $u$  entering location  $l$ , and the production quantity  $a$  at location  $l$ . These binary variables enable suitable selection of  $v_l^\theta$  from lookup tables. Additionally, we assume that the local information assumption, which ensures  $\hat{\lambda}(d_l^n, \mathbf{x})$  sufficiently approximates  $\hat{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x})$  when picking the approximate, local cost-to-go function  $v_l^{F,\theta}$ ,  $\forall l$ , is satisfied.

In this approach, the number of binary variables required to solve the one period problem at every epoch grows linearly in  $L$  and quadratically in the total module strength  $Y$ . Hence, we present a lookahead approach in Section 4.6.4 to improve the computational efficiency of the joint controller's strategy using piecewise linear and convex approximation of  $v_l^{F,\theta}$  that reduces the number of binary variables used.

$$\begin{aligned} \text{JR:} \quad \min \quad & \sum_{l, \Delta^S, \Delta^M, q} w(l, \Delta^S, \Delta^M, q) \left\{ K_l^{S+} (\Delta^S)^+ + K_l^{S-} (-\Delta^S)^+ + K^M |\Delta^M|/2 \right. \\ & + \sum_n \sum_i x_i \sum_j P_{ij} O_{nj}^l \left[ h_l r_l^n + b_l o_l^n \right. \\ & \left. \left. + \beta \hat{v}_l^{F,\theta}(\hat{\lambda}(d_l^n, \mathbf{x}), s_l + \Delta^S + q - d_l^n, u_l + \Delta^M) \right] \right\}, \end{aligned}$$

subject to

$$\begin{aligned} & \sum_{l, \Delta^S, \Delta^M, q} w(l, \Delta^S, \Delta^M, q) \Delta^M = 0 \\ & \sum_{l, \Delta^S, \Delta^M, q} w(l, \Delta^S, \Delta^M, q) \Delta^S = 0, \\ & r_l^n \geq s_l + \sum_{\Delta^S, \Delta^M, q} w(l, \Delta^S, \Delta^M, q) (\Delta^S + q) - d_l^n, \forall l, n \\ & o_l^n \geq d_l^n - s_l - \sum_{\Delta^S, \Delta^M, q} w(l, \Delta^S, \Delta^M, q) (\Delta^S + q), \forall l, n \\ & r_l^n, o_l^n \in \mathbb{Z}^+, w(l, \Delta^S, \Delta^M, q) \in \{0, 1\}, \forall \Delta^S \in \{-(s_l)^+, \dots, \sum_{k \neq l} (s_k)^+\}, \\ & \Delta^M \in \{-u_l, \dots, Y_l' - u_l\}, q \in \{0, \dots, U_l + (u_l + \Delta^M)G\}, \\ & l \in \{1, \dots, L\}. \end{aligned} \tag{4.10}$$

#### 4.6.4 Lookahead Strategy of Joint Controller (LAJ)

The mixed integer program LAJ, presented below, makes use of the piecewise linear and convex approximation of the single location capacitated inventory control system's cost-to-go function presented in Section 4.5.2 in order to reduce the computational effort required to implement the JC approach. Using this functional approximation of the cost-to-go function reduces the number of integer variables by  $\mathcal{O}(GY^2LI)$  where G, Y, L, and I are respectively the capacity per module, the total number of production modules, the number of locations, and the available storage capacity at each location.

$$\begin{aligned}
\text{LAJ: } \quad & \min_{\Delta^S, u', y} \sum_l \left\{ (K_l^{S+} \Delta_l^{S+} + K_l^{S-} \Delta_l^{S-}) + K^M \sum_l |u_l - u'_l|/2 \right. \\
& \left. + \sum_n \sigma(d_l^n, \mathbf{x}) \left[ h_l r_l^n + b_l o_l^n + \beta(\zeta_l + \eta_l)/2 \right] \right\}, \\
& \text{subject to} \\
& \zeta_l \geq \gamma_j^l (y_l - [E[D_l(t)]]) + \hat{\gamma}_j^l \quad \forall (\gamma_j^l, \hat{\gamma}_j^l) \in \Gamma_{t+1}^l(u_l) \quad \forall l \\
& \eta_l \geq \theta_j^l u'_l + \hat{\theta}_j^l \quad \forall (\theta_j^l, \hat{\theta}_j^l) \in \Theta_{t+1}^l(s_l) \quad \forall l \\
& \sum_l u'_l = Y \\
& 0 \leq u'_l \leq Y'_l, \quad \forall l \\
& \sum_l \Delta_l^{S+} = \sum_l \Delta_l^{S-}, \\
& 0 \leq \Delta_l^{S+} \leq \sum_{k \neq l} (s_k)^+, \quad 0 \leq \Delta_l^{S-} \leq -(s_l)^+, \quad \forall l \\
& (s_l + \Delta_l^{S+} - \Delta_l^{S-}) \leq y_l \leq (s_l + \Delta_l^{S+} - \Delta_l^{S-}) + U_l + u'_l G, \quad \forall l \\
& r_l^n \geq y_l - d_l^n, \quad o_l^n \geq d_l^n - y_l, \quad \forall l, n \\
& r_l^n, o_l^n \in \mathbb{Z}^+, \quad u'_l, \Delta_l^{S+}, \Delta_l^{S-}, y_l \in \mathbb{Z}, \quad \eta_l, \zeta_l \in \mathbb{R} \quad \forall l
\end{aligned} \tag{4.11}$$

This heuristic utilizes significantly fewer integer variables compared to the integer program in (4.10). Additionally, we have the following result that shows LAJ can be solved as a linear program to obtain an optimal solution when module capacity equals 1. This result improves

the speed of implementing the JC approach drastically in such instances, in comparison with JR.

**Proposition 2.** *LAI can be solved exactly by relaxing all the integrality constraints when  $G = 1$ .*

The proof of this result follows from Proposition 2 [104].

#### 4.6.5 Global-Local Rollout of Stationary Future (GLR)

For the GLR heuristic, at every decision epoch with beginning state  $(\mathbf{x}, \mathbf{s}, \mathbf{u})$ , we first solve

1. the following integer program to determine the GC decisions, namely, the amount of inventory  $\Delta^S$  received at every location  $l$  and the number of production modules  $\Delta^M$  received at every location  $l$ :

$$\begin{aligned}
\text{GLR: } \quad & \min \sum_{l, \Delta^S, \Delta^M} w(l, \Delta^S, \Delta^M) \left\{ K_l^{S+} (\Delta^S)^+ + K_l^{S-} (-\Delta^S)^+ + K^M |\Delta^M|/2 \right. \\
& \left. + \widehat{v}_l^{F, \theta}(\mathbf{x}, s_l + \Delta^S, u_l + \Delta^M) \right\}, \\
& \text{subject to} \\
& \sum_{l, \Delta^S, u} w(l, \Delta^S, \Delta^M) \Delta^M = 0 \\
& \sum_{l, \Delta^S, \Delta^M} w(l, \Delta^S, \Delta^M) \Delta^S = 0, \\
& w(l, \Delta^S, \Delta^M) \in \{0, 1\}, \forall \Delta^S \in \{-(s_l)^+, \dots, \sum_{k \neq l} (s_k)^+\}, \\
& \Delta^M \in \{-u_l, \dots, Y_l' - u_l\}, l \in \{1, \dots, L\}.
\end{aligned} \tag{4.12}$$

2. We then determine the local controllers' replenishment decisions through the location-wise order-up-to-policy presented in (4.6), in which the quantity transshipped to any location  $l$  will be obtained using the solution of the above integer program GLR as

$$\Delta_l^S = \sum_{\Delta^S = -(s_l)^+}^{\sum_{k \neq l} (s_k)^+} \sum_{\Delta^M = -u_l}^{Y_l' - u_l} w(l, \Delta^S, \Delta^M) \Delta^S \text{ for all locations } l.$$



## 4.7 Computational Study and Results

We designed a set of instances that allows us to study the variation of heuristic quality and the value addition of mobility as a function of the number of locations  $L$ , the movement cost per unit of inventory between any pair of locations represented as  $K^S$ , and the movement cost per production module  $K^M$ .

### 4.7.1 Instance design

We generated a set of 100 instances in the following manner. The number of modulation states is  $N = 3$  for all instances. The transition matrix  $\hat{P}$  is set to  $\{\{0.7, 0.3, 0\}, \{0.5, 0.3, 0.2\}, \{0, 0.8, 0.2\}\}$ ; the underlying Markov chain is represented in Figure 4.2. There are

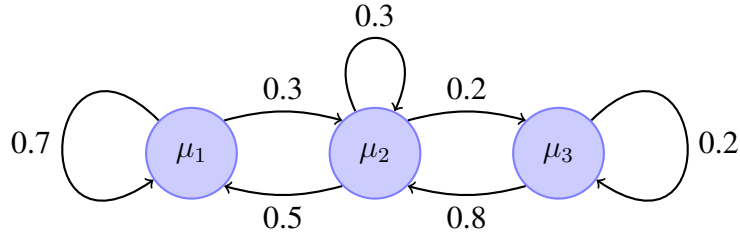


Figure 4.2: Underlying Markov chain of the modulation process in the instance set

three demand outcomes at every location, 0, 1, or 2. For each value of  $L \in \{2, 3, 5, 10\}$ , we generated discrete demand distributions at each location randomly for each modulation state such that the expected demands are either increasing with modulation state or decreasing with modulation state at each location. We ensured that exactly one of the three expected demands at each location lies in  $[0, 0.6)$ ,  $[0.6, 1.4)$ , and  $[1.4, 2]$ . The total number of modules  $Y$  and the module holding capacity at each location  $Y'_l, \forall l \in \{1, \dots, L\}$  are set to equal  $\lceil \frac{4}{3}L \rceil$ . We created a problem instance for each combination of the parameters fixed above and inventory movement cost per unit  $K^S \in \{0, 1.5, 2, 2.5, 10000\}$  for all  $l$ , module movement cost per production module  $K^M \in \{0, 1.5, 2, 2.5, 10000\}$ , holding cost per unit per period  $h_l = h = 1$  for all  $l$ , and backorder cost per unit per period  $b_l = b = 2$

for all  $l$ . Without loss of generality, we set the production cost  $c_l$  at all locations to zero in all the instances. Based on the analysis in [104], we note that capacity per module does not affect the results observed significantly, and hence we set the capacity per module  $G$  to 1 in all the instances.

#### 4.7.2 Results

We evaluated the heuristic policies, GLR, JR, LAJ, and the myopic policy MP on fifty sample trajectories of the instance set, obtained by Monte Carlo simulation, for five values of the blending coefficient  $\theta \in \{0, 0.2, 0.5, 0.8, 1\}$ . We compared their performance against the following benchmark policy, NF. For each instance, we computed the approximate cost-to-go value function of the  $L = 1, Y = 0$  problem with the various capacities and determined the minimum total fixed cost among all configurations. We then generated 50 sample demand trajectories at each epoch based on the current simulated modulation state. For each trajectory, the beginning state is the zero inventory position at all locations and the module configuration that minimizes the sum of the fixed expected total cost of the single location problems with the steady state belief-based distribution of demand as the epoch-invariant demand distribution at each location. We computed the upper bound  $\hat{v}_l(\mathbf{x}, s_l, u_l)$  for  $\mathbf{x} \in X', u \in \{0, \dots, Y\}, \forall s_l, \forall l$  in a one time offline pre-computation step. We approximated the belief space  $X = \{\mathbf{x} : \sum_{i=1}^N x_i = 1, x_i \geq 0\}$  with its non-empty, fixed, finite subset  $X' = \{\mathbf{x} : \sum_{i=1}^N x_i = 1, x_i \in \{0, 1/3, 2/3, 1\}\} \cup \pi$ , for  $\pi$  such that  $\pi = \pi \hat{P}$  when it exists [121].

We performed a forward dynamic programming pass or a forward rollout implementing the decision-making proposed by each method at each epoch. We obtained the average performance of each heuristic over the 50 simulated trajectories of each instance to analyze various resultant trends in comparison to NF.

We compared heuristic performance across values of the blending coefficient  $\theta$  in Tables C.1, C.2, and C.3 Table 4.1 (in Section C4) and found the best performance usually at

$\theta = 0.2$  for all the heuristics. Table 4.1 presents the comparison of the performance of all heuristics at  $\theta = 0.2$ . We note that the cost of the naive policy MP is very close to that of NF, sometimes exceeding it. This observation establishes the need for intelligent, dynamic heuristics that account for future costs. The proposed heuristics provide about a 20% – 25% reduction in cost compared to NF, in effect extracting 20% – 25% improvement in system performance from the two forms of mobile flexibility. As the number of locations  $L$  increases, we observe increasing value addition over NF generally. We note that heuristic quality is ordered as JR (best), LAJ, and GLR with the bandwidth of 5% variation.

Table 4.1: Average performance of heuristics w.r.t. NF across  $L$  for  $\theta = 0.2$

$L \setminus \theta = 0.2$	GLR	JR	LAJ	MP
2	0.903	0.890	0.903	1.126
3	0.705	0.661	0.690	0.945
5	0.833	0.797	0.796	0.986
10	0.726	0.673	0.695	0.997
Overall	0.792	0.755	0.771	1.013

JR and LAJ outperform GLR on average by 1% – 5% and 0% – 3% respectively. The strength of GLR is its unique usefulness while managing instances where different locations are coupled (or correlated) not only through the modulation process. JR and LAJ rely on the assumption that the demands at different locations are mutually independent, conditional on the belief state.

Table 4.2: Average computation times in seconds across  $L$

$L$	MP	GLR	JR	LAJ	$\hat{v}_l^F \forall l$
		per trajectory			per instance
2	0.06	0.09	0.84	0.16	0.60
3	0.09	0.24	3.12	0.25	1.75
5	0.16	0.62	20.7	0.43	7.45
10	0.36	3.55	258	0.93	65.8

Although JR gives the best performance, its computational demands, indicated by the average run time per trajectory in Table 4.2, makes it unattractive as a preferred heuristic for

online implementation. LAJ and GLR, on the other hand, offer significant computational advantages over JR and are amenable to online implementation. The time taken to compute the upper bound  $\hat{v}_l(\mathbf{x}, s_l, u_l)$  for  $\mathbf{x} \in X', u \in \{0, \dots, Y\}, \forall s_l, \forall l$  is a one time offline computation cost and it grows with  $L$  as expected (due to decentralization).

We now explore how the value of mobility of production capacity and inventory is affected by the movement cost structure. In Table 4.3, the module movement cost per module  $K^M$  and the transshipment cost per unit of inventory  $K^S$  are varied horizontally and vertically respectively when the number of locations  $L = 10$ . We observe decreasing

Table 4.3: Value of mobility (% savings over NF) using JR with  $\theta = 0.2$  across  $K^S$  and  $K^M$  for  $L = 10$

		Module movement cost $K^M$				
		0	1.5	2	2.5	10000
Transshipment cost $K^S$	0	47%	50%	48%	47%	46%
	1.5	41%	24.9%	24.6%	19.4%	18.6%
	2	50%	26.2%	15.5%	12.8%	12.6%
	2.5	45%	25.2%	10.7%	0.3%	0.1%
	10000	46%	22.8%	8.4%	-1%	-0.7%

savings generally when we go radially outwards from the top-left corner, indicating greater value of mobility when the cost of purchasing mobility is lower. Holding  $K^M$  constant for  $K^M \geq b$ , the savings decrease for increasing  $K^S$ . Holding  $K^S > 0$  constant, we observe higher savings at lower values of  $K^M$ . The value additions from the two forms of flexibility independently can be obtained from the last row and last column of Table 4.3. We infer that there is maximum benefit from these forms of flexibility when the average movement cost of a unit of capacity or a unit of inventory are less than or equal to the backorder cost per unit per period  $b = 2$ .

#### 4.7.3 Demand modeling: epoch-invariant vs. epoch-variant

We now consider how valuable modeling the inherent epoch-variability of demand distributions is. We accomplish this by repeating the experimental runs of heuristics, assuming

steady-state, epoch-invariant distributions of demand in a Markov-modulated system with partial observability. Table 4.4 considers the value lost in heuristic costs when epoch-invariability is assumed by the decision-maker. The values in Table 4.4 may be obtained as the difference between entries of Table C.5 presented in Section C5. We present the observations for  $\theta = 0.2$ , which is the best blending coefficient for all heuristics with epoch-variability-assuming decision-maker. We present  $\theta = 0.8$  results for GLR as it is the best blending coefficient for epoch-invariability - assuming decision-maker's GLR. We note that only GLR performs better under an epoch-invariability assumption. This finding is counter-intuitive. We tested the tractable heuristics GLR and LAJ on a longer horizon ( $T = 30$ )

Table 4.4: Value of non-stationarity w.r.t. NF across  $L$ , when the DM assumes steady state, epoch-invariant demand distributions in a Markov-modulated system with partial observability

$L$	GLR ( $\theta = 0.8$ )	GLR ( $\theta = 0.2$ )	JR ( $\theta = 0.2$ )	LAJ ( $\theta = 0.2$ )
2	-0.028	0.003	-0.013	0.011
3	-0.048	0.029	0.014	0.007
5	-0.046	0.001	0.006	0.010
10	-0.071	0.017	0.017	0.006
Overall	-0.048	0.013	0.006	0.009

with both epoch-invariant and epoch-variant assumptions of DM, assuming the modulation system is epoch-variant (with partially observed Markov modulation). We obtained the results presented in Table 4.5. These results indicate that all heuristics perform marginally better over very long horizons under a stationary assumption, when the steady state distribution of the corresponding Markov chain of the state-of-the-economy is supplied. These results are now convincing, given that there are significant probabilities of transitioning away from any given modulation state in the transition matrix  $\hat{P} = \{\{0.7, 0.3, 0\}, \{0.5, 0.3, 0.2\}, \{0, 0.8, 0.2\}\}$ . The opportunity to learn demands and infer a modulation state may be lost too quickly due to fast transition to a different modulation state. We understand that if the system is too volatile, a good steady state based epoch-invariant distribution works well to

extract comparable gain from flexibility.

Table 4.5: Comparison of heuristic performance w.r.t. NF for  $T = 30$  when the DM assumes epoch-invariant and epoch-variant demand distributions in a Markov-modulated system with partial observability

$L$	Non-stationary		Stationary	
	GLR ( $\theta = 0.2$ )	LAJ ( $\theta = 0$ )	GLR ( $\theta = 0.5$ )	LAJ ( $\theta = 0$ )
2	0.79	0.77	0.77	0.77
3	0.45	0.47	0.45	0.47
5	0.71	0.72	0.67	0.71
10	0.53	0.59	0.52	0.59
Overall	0.62	0.64	0.60	0.63

We now establish the value of modeling epoch-variability by comparing heuristic performance under the two demand assumptions of the DM in an epoch-variant system for a transition matrix  $\hat{P} = \{\{0.95, 0.05, 0\}, \{0.05, 0.9, 0.05\}, \{0, 0.05, 0.95\}\}$  that has very small probabilities of leaving the current modulation state (with all the other parameters of the set remaining the same as before). Once again, we present the results for the best  $\theta$ 's. In Table 4.6, even on very long horizons, even GLR extracts 6% more average savings under epoch-variant, belief-based demand modeling. These results confirm our intuition that when the modulation states are vastly different in their location-wise demand distributions and do not change too rapidly, our demand learning model produces significant value addition over assuming epoch-invariant demand models in a Markov-modulated system with partial observability.

#### 4.7.4 Complete observability of modulation process

We now pursue the value proposition of complete observability of modulation process. How useful is it to completely (or more accurately) observe the state-of-the-economy? We repeat the experiments on the instance set presented in Section 4.7.1 with the assumption that the modulation process is completely observed. This relaxes the constraint of the decision-maker being unaware of the current modulation state and hence results in a lower

Table 4.6: Comparison of heuristic performance w.r.t. NF for  $T = 30$  and  $\hat{P} = \{\{0.95, 0.05, 0\}, \{0.05, 0.9, 0.05\}, \{0, 0.05, 0.95\}\}$  when the DM assumes epoch-invariant and epoch-variant demand distributions in a Markov-modulated system with partial observability

$L$	Non-stationary		Stationary	
	GLR ( $\theta = 0.2$ )	LAJ ( $\theta = 0.2$ )	GLR ( $\theta = 0.2$ )	LAJ ( $\theta = 0$ )
2	0.43	0.44	0.48	0.49
3	0.42	0.42	0.48	0.48
5	0.29	0.30	0.36	0.43
10	0.29	0.29	0.35	0.51
Overall	0.36	0.36	0.42	0.48

cost than without complete observability. Table 4.7 presents the additional value of mobility under complete observability, or equivalently, the value of complete observability using JR on  $L = 10$  instances. These values are obtained as the difference of the values of Tables 4.3 and C.6. We note that end effects lead to some negative entries in the first column that are found to be positive when tested on longer horizons. Value of observability can significantly high (25.6% when  $K^S = K^M = 2.5$ ), particularly on cost structures that were indicating low value of mobility under partial observability. This finding signals that exploring techniques to improve observability of the state-of-the-economy would be greatly beneficial to the current logistics system. When costs are moderate ( $K^S = K^M = 1.5$ , there is about 8.5% improvement due to complete observability. We infer that under moderate movement costs, the value of mobility with and without complete observability is very high and under higher movement costs, with complete observability, the savings obtained are quite lucrative.

Thus, we conclude our computational analysis by emphasizing the significant impact of mobility, non-stationary demand modeling, and observability of the state-of-the-economy on the profitability of production-inventory systems and reiterating the computational advantage offered by the heuristics LAJ and GLR over JR.

Table 4.7: Additional value of mobility (% savings over NF) due to complete observability using JR with  $\theta = 0.2$  across  $K^S$  and  $K^M$  for  $L = 10$

		Module movement cost $K^M$				
		0	1.5	2	2.5	10000
Transshipment cost $K^S$	0	1%	1%	0%	1%	1%
	1.5	-3%	8.5%	8.5%	6.2%	5.8%
	2	-4%	6%	11.3%	13.9%	13.4%
	2.5	-1%	7.8%	15.3%	25.6%	22.2%
	10000	-3%	5.7%	11.3%	22%	1.3%

## 4.8 Conclusion

We present computationally efficient heuristics LAJ and GLR that improve in solution quality with the number of locations in the  $L$  location,  $Y$  module problem. We observe savings from mobility that are as high as 26% in some instances over systems with no flexibility. Also, we note that non-stationary modeling of demand in a non-stationary world allows about 6% more profits in some cases and complete observability of the state-of-the-economy increases value addition of mobility by 5 – 27% on the instances considered. Additionally, we infer that although joint control results in slightly lower costs, decentralized control heuristics perform significantly faster. Our results reinforce the value addition due to the production capacity portability, irrespective of the presence of transshipment flexibility.



## **CHAPTER 5**

### **CONCLUSIONS AND FUTURE RESEARCH**

#### **5.1 Summary of Results**

In this dissertation, we consider some emergent logistics systems and pursue solution methods to maximize their operational efficiency and ease. We adopt a multi-pronged approach of handling uncertainty - by managing response (on the supply side) and information (on the demand side). To deliver a better logistics response using the novel characteristic of mobility of production capacity, we also seek to manage revealed information effectively by proposing a data-driven learning model of demand that is affected by a Markovian state-of-the-world variable. Computational evidence in this research suggests that significant savings arise in a multi-location production - inventory system from resource mobility when the resource is either production capacity or inventory. Multi-resource mobility (for example, when both production capacity and inventory are transportable and shareable) also taps additional savings over and above single resource mobility. A data - informed non-stationary demand model is expected to have greater predictive power of anticipation over a stationary demand model. The remainder of this section summarizes the primary results of the research going into this dissertation.

In Chapter 2, we analyze a problem of dynamic logistics planning for a multi-location production - inventory system with transportable modular production capacity facing uncertain demands. In such systems, production modules provide units of capacity, and can be moved from one location to another to produce stock and satisfy demand. We formulate a dynamic programming model of a planning problem in this setting that considers production and inventory decisions. Given the size of the state and action spaces, we focus on developing suboptimal one period lookahead and rollout policies based on decomposition

by locations, and upper and lower bounds on the optimal cost function that help quantify the effectiveness of suboptimal policies as well as measure the value of transportable production capacity. In some cases, finding these suboptimal policies requires solving single-period problems to optimality. We propose mixed-integer linear programming models for these generalizations of newsvendor problems, and show under certain circumstances the feasible region polyhedra have only integer extreme points. A computational study of problems with stationary demand distributions, which should benefit least from mobile capacity, demonstrates the effectiveness of the suboptimal policies. For problems with 20 locations, the best heuristic solution cost provides 13% savings over a system with an optimal fixed capacity allocation. Greater savings result when the number of locations increases.

In Chapter 3, we consider a periodic review inventory control problem having an underlying modulation process that affects demand and that is partially observed by the uncensored demand process and a novel additional observation data (AOD) process. Letting  $K$  be the reorder cost, we present a condition, A1, which is a generalization of the Veinott attainability assumption, that guarantees the existence of an optimal myopic base stock policy if  $K = 0$  and the existence of an optimal  $(s, S)$  policy if  $K > 0$ , where both policies depend on the belief function of the modulation process. Assuming A1 holds, we show that (i) when  $K = 0$ , the value of the optimal base stock level is constant within regions of the belief space and that these regions can be described by a finite set of linear inequalities and (ii) when  $K > 0$ , the values of  $s$  and  $S$  and upper and lower bounds on these values are constant within regions of the belief space and that these regions can be described by a finite set of linear inequalities. Computational procedures for  $K \geq 0$  are outlined, and results for the  $K = 0$  case are presented when A1 does not hold. Special cases of this inventory control problem include problems considered in the Markov-modulated demand and Bayesian updating literatures.

In Chapter 4, we analyze the value of mobile production capacity in a supply chain with geographically distributed production facilities. For supply chains with fixed location

facilities, changing geographical demands has usually been met by transshipment, expanding or contracting production capacity at existing facilities, and either eliminating old or adding new facilities. The increasing interest in additive manufacturing, which can be mobile, and other forms of mobile production capacity (e.g., in the pharmaceutical industry) offers an opportunity for improved performance of next generation supply chain design and operations.

We model the  $L$  location,  $Y$  mobile production unit problem as a problem of sequential decision making under uncertainty to determine transshipment, mobile production capacity relocation, and replenishment decisions at each epoch. We include in this model a data-driven demand forecasting capability that assumes the existence of a partially observed stochastic process, the modulation process, that affects demand, is not affected by the actions of the decision maker, and reflects the reality that decision making environments are often affected by exogenous and partially observed forces (e.g., the macro economy, sea or air currents). We use a specially structured partially observed Markov decision process as our model, develop several heuristic procedures for determining policies, and compare these heuristics numerically, demonstrating that a decentralized decision making approach shows promise. We note that the value of mobility is 26% on some instances with 10 locations, the value of non-stationary modeling is significant - around 6% - on instances in which it is significantly more likely to stay in any modulation state than to transition away from it, and the value of complete observability of the modulation process is 5 – 27% on the instance set considered.

The two kinds of resources, namely production capacity and inventory, differ in some respects. Firstly, production capacity is a re-usable resource unlike inventory which serves a one time demand satisfaction purpose. Thus, relocating a unit of capacity appears to create much more impact at both the sending as well receiving locations than relocating a unit of inventory. Secondly, production capacity often requires an enabling setup for plug-and-produce at all the locations and thus an implicit setup time before it is ready to

use typically. We assume that this time is negligible relative to the length of the period throughout Chapters 2 and 4.

## **5.2 Recommendations for Future Research**

While the scope of research in supply chain logistics planning is very wide, some directions of future research in logistics planning under certainty that are direct extensions of the research in this dissertation are presented below.

### 5.2.1 Resource Mobility and Sharing

- Considering lead times of resource movement between locations.
- Considering setup times for resources that require a setup before usage.
- Allowing competition between locations instead of joint ownership.
- Exploring the impact of product perishability and demand impatience on the value addition from resource mobility
- Investigating human resource mobility and sharing in the context of skilled service providers (health care, specialized tasks, etc.)
- Considering design and operation questions in urban mobility, such as car sharing and ride sharing, from a resource mobility perspective.
- Analyzing the value proposition of other novel forms of flexibility in logistics systems such as mobile storage capacity (warehouses and smart lockers).

### 5.2.2 Learning-based Demand Modeling

- Modeling inventory control with lost sales under demand-influencing Markov-modulation.

- Modeling the problem of joint inventory control and pricing under demand-influencing Markov-modulation.
- Analyzing the resultant correlation between demands at different locations due to implicit coupling through the state-of-the-world.

# **Appendices**

## APPENDIX A

### CHAPTER 2

#### A1 Proof of Proposition 1

*Proof.* (a) As the newsvendor constraint matrices are TU already, their rows can be separated into partitions suitably to satisfy [122, Theorem 2.7, III.1]. Total unimodularity is retained in the presence of location-wise capacity constraints when  $G = 1$ . Hence, the constraint matrix of MMPIP-MP is TU when  $G = 1$ . As the right hand side is integral, the polyhedron of the IP presented here for the implementation of the myopic policy, MP, is integral.

(b) The objective function of each of the  $L$  newsvendor problems is given by

$$\sum_{k \in \mathcal{K}} p_i^k \{ h_i(s_i + q_i - d_i^k)^+ + b_i(d_i^k - s_i - q_i)^+ \}.$$

$h_i P(D_i \leq d_i^m) - b_i P(D_i > d_i^m)$  is the slope of the newsvendor objective function in  $d_i^m \leq q_i + s_i \leq d_i^{m+1}$ . As the function is convex, its slope is decreasing from left to right and increasing from right to left. The current problem involves re-allocation of capacity followed by replenishment. In the one period problem, a module movement from location  $i$  to location  $j$  occurs if

- i. the available capacity is greater than the unconstrained optimal order quantity at the sending location  $i$ , or
- ii. if the positive gain at the receiving location  $j$ ,  $h_j P(D_j \leq d_j^n) - b_j P(D_j > d_j^n)$  per unit is higher than the (absolute value of) loss at the sending location  $i$ ,  $h_i P(D_i \leq d_i^m) - b_i P(D_i > d_i^m)$ . The movement is feasible only if  $h_i P(D_i \leq d_i^m) - b_i P(D_i > d_i^m) + h_j P(D_j \leq d_j^n) - b_j P(D_j > d_j^n) > K_{ij}^M / G$ . The movement will be beneficial

until this inequality is reversed when either  $s_j + q_j$  or  $s_i + q_i$  reach a slope change point, *viz.*, a demand outcome (which is integral for this problem setup). Since the starting inventory ( $s_i$  or  $s_j$ ) is an integer and the order-up-to level at one of the locations is integral, it follows that the amount of effective potential inventory increment/decrement at the locations is integral and hence the other order-up-to level is also integral. This mechanism is in action for all module shifts induced by the cost structure and hence, irrespective of  $\{\{\Delta_{ij}^M\}\}$ 's integrality, for all  $i$ ,  $q_i$  will be take integer values even without integrality constraints.

□

## A2 Proof of Proposition 2

*Proof.* MMPIP-LAF is different from MMPIP-MP in the objective function that now contains  $\zeta_i$  and  $\eta_i$  terms additionally. The constraints now include the description of the  $\zeta_i$  and  $\eta_i$  terms as the maximum over piecewise linear facets of sets of convex curves. As these inventory and capacity curves have with integer slope transition points, the minimization of  $\sum_{i=1}^L (\zeta_i + \eta_i)/2$  ensures that the  $y_i$ 's and  $q_i$ 's are still integers in addition to the constraints of MMPIP-MP. Having integer values for  $y_i$  ensures that  $\Delta_{ij}^M$ 's are integers as  $y_i$ 's are placeholder variables. Integer values of  $q_i$ 's guarantees that  $r_i^k$  and  $o_i^k$  are integers as  $r_i^k + o_i^k = s_i + q_i - d_i^k$  and  $d_i^k$ 's are integers. Hence, MMPIP-LAF can be solved as a linear program.

□

## A3 Algorithm for LIU

Algorithm 1 is the value function approximation algorithm used to implement LIU.

## A4 Instance Sets for Computational Study

We generate three instance sets for our computational study.



---

**Algorithm 1:** Approximate Value Iteration over finite horizon: LIU

---

Step 1. Initialization:

- a. Initialize  $\bar{V}_t^0(\xi(t))$  for all states  $\xi(t)$ .
- b. Choose an initial state  $\xi(1)$ .
- c. Set  $n = 1$ .

Step 2. Choose a sample path  $d^n$ .

Step 3. For  $t = 1, \dots, T - 1$  do:

- a. Solve  $\hat{v}_t^n = \min_{a \in A(\xi^n(t))} \mathbb{E}_D \left( C_t(\xi^n(t), a, D) + \bar{V}_{t+1}^{n-1}(f(\xi^n(t), a, D)) \right)$  and let  $a_t^n$  be a minimizer.
- b. Blending:  $\bar{V}_t^n(\xi^n(t)) = \alpha_n \hat{v}_t^n + (1 - \alpha_n) \bar{V}_t^{n-1}(\xi^n(t))$ .
- c. Greedy trajectory following:  $\xi^n(t+1) = f(\xi^n(t), a_t^n, d^n(t))$ .

Step 4. Let  $n = n + 1$ . If  $n < N$ , go to step 2.

---

#### A4.1 Instance Set 1

$L = 2$ ,  $Y \in \{3, 5, 7, 10\}$ ,  $T \in \{5, 7, 10, 15\}$ ,  $G = 1$ ,  $c = 0$ ,  $b \in \{1, 2\}$ ,  $h \in \{0.5, 1\}$ ,  $\bar{K}^M \in \{0.5, 5\}$ , Initial allocation =  $\{u_{\min}\}$  = the best fixed system's capacity allocation

Demand distributions. For each of the two stations:

- Set expected total demand to  $\alpha GY$ ,  $\alpha = 0.75$ ;
- Randomly obtain fraction of expected demand at station 1:  $\beta \in [0.2, 0.8]$ ;  $\mu_1 = \beta\mu$ ;  
 $\mu_2 = (1 - \beta)\mu$ .  $\implies \mu_i^{\max} = \alpha\beta^{\max}Y = 0.6Y$ .
- Randomly generate three distinct samples each from  $\{0, \dots, \lceil \gamma\mu_1 \rceil\}$ ,  $\{0, \dots, \lceil \gamma\mu_2 \rceil\}$ ,  
 $\gamma = 1.1$ ;  $D_i^{\max} = \lceil \gamma\mu_i^{\max} \rceil = \lceil 0.66Y \rceil$  and sort them.
- Randomly generate three numbers from  $\{1, 2, 3, 4, 5, 6\}$  and obtain probabilities from them by dividing each by their sum.

For each combination of  $Y, T, b, h$ , and  $\bar{K}^M$ , obtain 10 different demand distributions. This completes instance generation. Thus, we generate a total of 1280 instances.

#### A4.2 Instance Set 2

$L = 2, Y \in \{5, 10, 15, 20\}, T \in \{5, 10, 15, 20\}, G \in \{1, 2, 3\}, c = 0, b \in \{1, 2, 3\}, h = 1,$   
 $\bar{K}^M \in \{1, 3, 5\}$ , Initial allocation =  $\{u_{\min}\}$  = the best fixed system's capacity allocation

Demand distributions. For each of the two stations:

- Set expected total demand to  $\alpha GY$ ,  $\alpha = 0.8$ ;
- Randomly obtain fraction of expected demand at station 1:  $\beta \in [0.2, 0.8]$ ;  $\mu_1 = \beta\mu$ ,  
 $\mu_2 = (1 - \beta)\mu. \implies \mu_i^{\max} = \alpha\beta^{\max}Y = 0.525GY$ .
- Randomly generate three distinct samples each from  $\{0, \dots, \lceil \gamma\mu_1 \rceil\}, \{0, \dots, \lceil \gamma\mu_2 \rceil\}$ ,  
 $\gamma = 1.5$ ;  $D_i^{\max} = \lceil \gamma\mu_i^{\max} \rceil = \lceil 0.7875GY \rceil$  and sort them.
- Randomly generate three numbers from  $\{1, 2, 3, 4, 5, 6\}$  and obtain probabilities from them by dividing each by their sum.

For each combination of  $Y, T, b, h$ , and  $\bar{K}^M$ , obtain 10 different demand distributions. This completes instance generation. Thus, we generate a total of 2160 instances.

#### A4.3 Instance Set 3

$L \in \{2, 3, 5, 10, 20\}, Y \in \{\lceil 1.9L \rceil, \lceil 2.4L \rceil\}, T \in \{5, 10, 15\}, G = 1, c = 0, b \in \{2, 3\},$   
 $h = 1, \bar{K}^M \in \{1, 2, 3\}$ , Initial allocation =  $\{u_{\min}\}$  = the best fixed system's capacity allocation

Demand distributions. For each location:

- Randomly generate three distinct samples each from  $\{0, 1, 2, 3\}$  and sort them.
- Randomly generate three numbers and obtain probabilities from them by dividing each by their sum.

For each combination of  $L, Y, T, b$ , and  $\bar{K}^M$ , obtain three different demand distributions. This completes instance generation. Thus, we generate a total of 540 instances.

## A5 A Simple Strategy for Module Allocation

Assign  $\{\lfloor \rho_i \rfloor Y\}$  to each location  $i$  where,

$$\rho_i = \frac{\sigma_i}{\sum_{j=1}^L \sigma_j} \times \beta + \frac{\mu_i}{\sum_{j=1}^L \mu_j} \times (1 - \beta),$$

$\beta = \sum_{j=1}^L \frac{z_\alpha \sigma_j}{\sum_{j=1}^L (z_\alpha \sigma_j + \mu_j)}$ ,  $z_\alpha = 1.64$  for a 95% service level.

If  $\sum_{i=1}^L \lfloor \rho_i Y \rfloor < Y$ , then allocate the remaining modules in the decreasing order of coefficient of variation,  $\sigma_i/\mu_i$ , of locations. The final module allocation thus obtained is represented by  $u_{\text{simple}}$ .

## A6 Additional Tables

Table A.1: Instance Set 1 - Variation of bounds w.r.t. OPT for HH instances across  $T$ ,  $Y$ ,  $b$ ,  $h$ , and  $\bar{K}^M$

(a) $T$					(b) $Y$				
$T$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>	$Y$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>
5	0.498	0.900	3.9	1.040	3	0.626	0.865	3.9	1.148
7	0.482	0.893	5.9	1.072	5	0.490	0.859	7.1	1.112
10	0.521	0.866	6.5	1.142	7	0.246	0.917	11.3	1.057
15	0.499	0.855	9.8	1.217	10	0.061	0.975	16.0	1.012
(c) $b$					(d) $h$				
$b$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>	$h$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>
1	0.500	0.883	5.0	1.086	0.5	0.525	0.861	7.7	1.132
2	0.500	0.874	8.1	1.148	1	0.472	0.898	5.3	1.101
(e) $\bar{K}^M$									
$\bar{K}^M$	PIR	LB	UB <sub>max</sub>	UB <sub>min</sub>					
0.5	0.437	0.928	6.8	1.180					
5	0.560	0.831	6.3	1.056					
Overall	0.500	0.879	6.6	1.117					

Table A.2: Instance Set 1 - Value addition of mobile modularity w.r.t. OPT across  $Y$ ,  $T$ ,  $\overline{K}^M$ ,  $h$ , and  $b$

	All			HH Instances		
$Y$	$\Delta K$	$-\Delta B$	$-\Delta H$	$\Delta K$	$-\Delta B$	$-\Delta H$
3	0.049	0.182	-0.011	0.055	0.223	-0.019
5	0.025	0.093	-0.010	0.044	0.170	-0.019
7	0.012	0.033	-0.001	0.026	0.097	-0.011
10	0.003	0.009	-0.002	0.008	0.033	-0.004
$T$	$\Delta K$	$-\Delta B$	$-\Delta H$	$\Delta K$	$-\Delta B$	$-\Delta H$
5	0.012	0.038	-0.005	0.025	0.075	-0.011
7	0.020	0.061	-0.004	0.035	0.123	-0.014
10	0.027	0.098	-0.010	0.057	0.221	-0.027
15	0.029	0.121	-0.004	0.065	0.304	-0.015
$\overline{K}^M$	$\Delta K$	$-\Delta B$	$-\Delta H$	$\Delta K$	$-\Delta B$	$-\Delta H$
0.5	0.026	0.112	-0.005	0.046	0.248	-0.021
5	0.019	0.047	-0.007	0.044	0.113	-0.012
$h$	$\Delta K$	$-\Delta B$	$-\Delta H$	$\Delta K$	$-\Delta B$	$-\Delta H$
0.5	0.025	0.093	-0.007	0.048	0.199	-0.017
1	0.019	0.065	-0.005	0.041	0.158	-0.017
$b$	$\Delta K$	$-\Delta B$	$-\Delta H$	$\Delta K$	$-\Delta B$	$-\Delta H$
1	0.020	0.061	-0.005	0.041	0.139	-0.012
2	0.024	0.097	-0.006	0.049	0.220	-0.021
Overall	0.022	0.079	-0.006	0.045	0.179	-0.017

Table A.3: Instance Set 1 - Performance of heuristics w.r.t. OPT on HH instances w.r.t.  $b$ ,  $h$ , and  $\overline{K}^M$

		LIU	MP	RF	RLB	LFP
$b$	1	1.020	1.066	1.014	1.032	1.005
	2	1.030	1.076	1.022	1.056	1.007
$h$	0.5	1.028	1.068	1.019	1.049	1.007
	1	1.022	1.075	1.017	1.039	1.005
$\overline{K}^M$	0.5	1.037	1.052	1.027	1.009	1.004
	5	1.013	1.090	1.009	1.078	1.008
Overall		1.025	1.071	1.018	1.044	1.006

Table A.4: Instance Set 1 - Performance of heuristics w.r.t. OPT on all instances w.r.t.  $b$ ,  $h$ , and  $\overline{K}^M$

		LIU	MP	RF	RLB	LFP
$b$	1	1.009	1.031	1.006	1.016	1.002
	2	1.016	1.034	1.013	1.023	1.003
$h$	0.5	1.014	1.032	1.010	1.022	1.003
	1	1.011	1.033	1.009	1.018	1.003
$\overline{K}^M$	0.5	1.020	1.023	1.015	1.005	1.003
	5	1.005	1.042	1.004	1.034	1.003
Overall		1.013	1.033	1.010	1.020	1.003

## APPENDIX B

### CHAPTER 3

#### B1 No Reorder Cost Case

**Proof of Lemma 2.** If  $s^*(\mathbf{x}) = d_m$ , then

$$\begin{aligned} A_{m-1}(\mathbf{x})d_{m-1} + B_{m-1}(\mathbf{x}) &> A_m(\mathbf{x})d_m + B_m(\mathbf{x}), \\ A_{m+1}(\mathbf{x})d_{m+1} + B_{m+1}(\mathbf{x}) &\geq A_m(\mathbf{x})d_m + B_m(\mathbf{x}), \end{aligned}$$

which leads to the result. □

**Proof of Proposition 1.** By induction. Letting  $v_0 = 0$ , we note that

$$v_1(s, \mathbf{x}) = \min_{y \geq s} L(\mathbf{x}, y) = L(\mathbf{x}, \max\{s^*(\mathbf{x}), s\})$$

for all  $\mathbf{x}$  and  $L(\mathbf{x}, \max\{s^*(\mathbf{x}), s\})$  is non-decreasing and convex in  $s$ . Thus, the result holds true for  $n = 1$  (and, trivially for  $n = 0$ ). Assume the result holds for  $n$ . Then, for  $s \leq s^*(\mathbf{x})$ ,

$$\begin{aligned} v_{n+1}(\mathbf{x}, s) &\leq L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(s^*(\mathbf{x}), d)) \\ &= L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), s^*(\boldsymbol{\lambda}(d, z, \mathbf{x}))) \\ &\quad \text{(using A1).} \end{aligned}$$

Also,

$$\begin{aligned}
v_{n+1}(\mathbf{x}, s) &\geq \min_{y \geq s} L(\mathbf{x}, y) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) \min_y v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(y, d)) \\
&= L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), s^*(\boldsymbol{\lambda}(d, z, \mathbf{x}))) \\
&= L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(s^*(\mathbf{x}), d)).
\end{aligned}$$

Thus, for  $s \leq s^*(\mathbf{x})$ ,

$$v_{n+1}(\mathbf{x}, s) = L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(s^*(\mathbf{x}), d)),$$

and  $v_{n+1}(\mathbf{x}, s) = v_{n+1}(\mathbf{x}, s^*(\mathbf{x}))$ . Assume  $s \geq s^*(\mathbf{x})$ . Note

$$v_{n+1}(\mathbf{x}, s) \leq L(\mathbf{x}, s) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(s, d)).$$

Also,

$$\begin{aligned}
v_{n+1}(\mathbf{x}, s) &\geq \min_{y \geq s} L(\mathbf{x}, y) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) \min_{y \geq s} v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(y, d)) \\
&= L(\mathbf{x}, s) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(s, d)),
\end{aligned}$$

and hence for  $s \geq s^*(\mathbf{x})$

$$v_{n+1}(\mathbf{x}, s) = L(\mathbf{x}, s) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(s, d))$$

and is non-decreasing and convex in  $s$ . □

**Proof of Lemma 5.** It is sufficient to show that if  $y \leq y'$  and  $\mathbf{x} \preceq \mathbf{x}'$ , then,

$$L(\mathbf{x}, y) - L(\mathbf{x}, y') \leq L(\mathbf{x}', y) - L(\mathbf{x}', y'),$$

which follows from the assumptions and [99, Lemma 4.7.2].  $\square$

**Proof of Lemma 7.** For any  $\mathbf{x} \in X$ , noting that  $\mathbf{x} = \sum_i x_i \mathbf{e}_i$ , it is straightforward to show that  $\lambda(d, z, \mathbf{x}) = \sum_i \xi_i \lambda(d, z, \mathbf{e}_i)$ , where  $\xi_i = x_i \sigma(d, z, \mathbf{e}_i) / \sum_j x_j \sigma(d, z, \mathbf{e}_j)$ .  $\square$

**Proof of Lemma 8.** We have the following:

(i) Clearly,  $0 \leq \hat{\mathbf{x}}_N^{d,z} \leq 1$  and  $\sum_{n=1}^N \hat{\mathbf{x}}_n^{d,z} = 1$ . It is sufficient to show  $0 \leq \hat{\mathbf{x}}_n^{d,z}, n = N-1, \dots, 1$ . Note

$$\begin{aligned} \sum_{k=n+1}^N \hat{\mathbf{x}}_k^{d,z} &= \min_{1 \leq i \leq N} \left\{ \sum_{k=n+1}^N \lambda_k(d, z, \mathbf{e}_i) \right\} \leq \sum_{k=n+1}^N \lambda_k(d, z, \mathbf{e}_i) \\ &\leq \sum_{k=n}^N \lambda_k(d, z, \mathbf{e}_i), \quad \forall i. \end{aligned}$$

Thus,  $\sum_{k=n+1}^N \hat{\mathbf{x}}_k^{d,z} \leq \min_{1 \leq i \leq N} \left\{ \sum_{k=n}^N \lambda_k(d, z, \mathbf{e}_i) \right\} = \sum_{k=n}^N \hat{\mathbf{x}}_k^{d,z}$ , and hence  $\hat{\mathbf{x}}_n^{d,z} \geq 0$ .

(ii) Let  $\mathbf{x}' \preceq \lambda(d, z, \mathbf{x}) \forall \mathbf{x} \in X$  and assume  $s^*(\hat{\mathbf{x}}^{d,z}) < s^*(\mathbf{x}')$ . Then by Lemma 5, there is an  $n \in \{1, \dots, N\}$  such that  $\sum_{k=n}^N x'_k > \sum_{k=n}^N \hat{\mathbf{x}}_k^{d,z}$ . However,  $\sum_{k=n}^N \hat{\mathbf{x}}_k^{d,z} = \min_{1 \leq i \leq N} \left\{ \sum_{k=n}^N \lambda_k(d, z, \mathbf{e}_i) \right\}$ , which leads to a contradiction of the assumption that  $\mathbf{x}' \preceq \lambda(d, z, \mathbf{x}) \forall \mathbf{x} \in X$ .  $\square$

**Proof of Lemma 9.** Assume  $f(y, d) = y - d$  and  $c(y, d) = p(d - y)^+ + h(d - y)^+$ , recall that elements of  $\mathcal{P}_1$  are sets of the form  $\{\mathbf{x} \in X : s^*(\mathbf{x}) = d_m\}$  for all  $d_m$  such that  $\min_{\mathbf{x}} s^*(\mathbf{x}) \leq d_m \leq \max_{\mathbf{x}} s^*(\mathbf{x})$ . Further recall that  $\{\mathbf{x} \in X : s^*(\mathbf{x}) = d_m\}$  is the set of all  $\mathbf{x}$  such that

$$\sum_{k=1}^{m-1} \sigma(d_k, \mathbf{x}) < p/(p+h) \leq \sum_{k=1}^m \sigma(d_k, \mathbf{x}),$$



or equivalently,

$$\mathbf{x} \sum_{k=1}^{m-1} \mathbf{P}(d_k) \underline{1} < p/(p+h) \leq \mathbf{x} \sum_{k=1}^m \mathbf{P}(d_k) \underline{1},$$

which represents two linear inequalities. Further, for  $\mathbf{x} \in \{\mathbf{x} \in X : s^*(\mathbf{x}) = d_m\}$ ,  $v_1^U(\mathbf{x}, s) = A_l(\mathbf{x})d_l + B_l(\mathbf{x})$  for  $l = \max\{s^*(\mathbf{x}), s\}$ , where we note

$$A_l(\mathbf{x})d_l + B_l(\mathbf{x}) = \mathbf{x} \left[ h \sum_{k=1}^l (d_l - d_k) \mathbf{P}(d_k) \underline{1} + p \sum_{k=l+1}^M (d_k - d_l) \mathbf{P}(d_k) \underline{1} \right],$$

where  $A_j(\mathbf{x})$  and  $B_j(\mathbf{x})$  are defined in Section 3.2.3. Thus, on each element of  $\mathcal{P}_1$ ,  $v_1^U$  is linear in  $\mathbf{x}$  for each  $s$  and each element of  $\mathcal{P}_1$  is described by a finite number of linear inequalities.

Let  $(\mathbf{x}, s)$  be such that  $d_l \leq \max\{s^*(\mathbf{x}), s\} \leq d_{l+1}$  for all  $\mathbf{x}$  in an element  $\{\mathbf{x} \in X : s^*(\mathbf{x}) = d_m\}$ . Further, let  $d_{l(d,z)} \leq \max\{s^*(\boldsymbol{\lambda}(d, z, \mathbf{x})), \max\{s^*(\mathbf{x}), s\} - d\} \leq d_{l(d,z)+1}$  for all  $\mathbf{x}$  in an element  $\{\mathbf{x} \in X : s^*(\boldsymbol{\lambda}(d, z, \mathbf{x})) = d_{m(d)}\}$ , which is the set of all  $\mathbf{x}$  such that:

$$\boldsymbol{\lambda}(d, z, \mathbf{x}) \sum_{k=1}^{m(d)-1} \mathbf{P}(d_k) \underline{1} < p/(p+h) \leq \boldsymbol{\lambda}(d, z, \mathbf{x}) \sum_{k=1}^{m(d)} \mathbf{P}(d_k) \underline{1},$$

or equivalently, for all  $\mathbf{x}$  such that  $\sigma(d, \mathbf{x}) \neq 0$ ,

$$\mathbf{x} \mathbf{P}(d, z) \sum_{k=1}^{m(d)-1} \mathbf{P}(d_k) \underline{1} < (p/(p+h)) \mathbf{x} \mathbf{P}(d, z) \underline{1} \leq \mathbf{x} \mathbf{P}(d, z) \sum_{k=1}^{m(d)} \mathbf{P}(d_k) \underline{1},$$

where we assume  $m$  and  $m(d)$  for all  $d$  have been chosen so that the finite set of linear

inequalities describes a non-null subset of  $X$ . We note that for such a subset,

$$\begin{aligned}
v_{n+1}^U(\mathbf{x}, s) &= A_l(\mathbf{x})d_l + B_l(\mathbf{x}) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) \times \\
&\times \left[ A_{l(d,z)}(\boldsymbol{\lambda}(d, z, \mathbf{x}))d_{l(d,z)} + B_{l(d,z)}(\boldsymbol{\lambda}(d, z, \mathbf{x})) \right] \\
&= \mathbf{x} \left[ h \sum_{k=1}^l (d_l - d_k) \mathbf{P}(d_k) \underline{1} + p \sum_{k=l+1}^M (d_k - d_l) \mathbf{P}(d_k) \underline{1} \right. \\
&\quad + \beta \sum_d \left[ h \sum_z \sum_{k=1}^{l(d,z)} (d_{l(d,z)} - d_k) \mathbf{P}(d, z) \mathbf{P}(d_k) \underline{1} \right. \\
&\quad \left. \left. + p \sum_z \sum_{k=l(d,z)+1}^N (d_k - d_{l(d,z)}) \mathbf{P}(d, z) \mathbf{P}(d_k) \underline{1} \right] \right].
\end{aligned}$$

The resulting partition  $\mathcal{P}_2$  is at least as fine as  $\mathcal{P}_1$  and each element in  $\mathcal{P}_2$  is described by a finite set of linear inequalities. We have shown that on each element in  $\mathcal{P}_2$ ,  $v_2^U(\mathbf{x}, s)$  is linear in  $\mathbf{x}$  for each  $s$ . A straightforward induction argument shows these characteristics hold for all  $n$ . We illustrate by example (through Example 3) how  $v_n^U(\mathbf{x}, s)$  may be discontinuous in  $\mathbf{x}$  for fixed  $s$ .  $\square$

**Proof of Proposition 5.** The result holds for  $n = 0$ ; assume the result holds for  $n$ . Then,

$$\begin{aligned}
v'_{n+1}(\mathbf{x}, s') &\geq \min_{y' \geq s'} \left\{ L'(\mathbf{x}, y') + \beta \sum_{d', z} \sigma'(d', z, \mathbf{x}) v_n^L(\boldsymbol{\lambda}'(d', z, \mathbf{x})) \right\} \\
&= \min_{y \geq s' - \delta} \left\{ L(\mathbf{x}, y) \right\} + \beta \sum_{d', z} \sigma'(d', z, \mathbf{x}) v_n^L(\boldsymbol{\lambda}'(d', z, \mathbf{x})) \\
&\geq L(\mathbf{x}, s^*(\mathbf{x})) + \beta \sum_{d', z} \sigma'(d', z, \mathbf{x}) v_n^L(\boldsymbol{\lambda}'(d', z, \mathbf{x})) \\
&= v_{n+1}^L(\mathbf{x}).
\end{aligned}$$

Further, note

$$\begin{aligned}
v'_{n+1}(x, s) &\leq \min_{y' \geq s} \left\{ L'(\mathbf{x}, y') + \beta \sum_{d', z} \sigma'(d', z, \mathbf{x}) v_n(\boldsymbol{\lambda}'(d', z, \mathbf{x}), y' - d') \right\} \\
&= \min_{y \geq s - \delta} \left\{ L(\mathbf{x}, y) + \beta \sum_{d, z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), y - d) \right\} \\
&\leq \min_{y \geq s} \left\{ L(\mathbf{x}, y) + \beta \sum_{d, z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), y - d) \right\} \\
&= v_{n+1}(\mathbf{x}, s).
\end{aligned}$$

The result follows by induction.  $\square$

## B2 Reorder Cost Case

**Proof of Proposition 6.** The proof of Proposition 6 is a direct extension of the results in [77].  $\square$

**Lemma 12.** For all  $\mathbf{x}$  and  $n$ :

- (i) if  $s \leq s'$ , then  $v_n(\mathbf{x}, s) \leq v_n(\mathbf{x}, s') + K$
- (ii) if  $y \leq y'$ , then  $G_n(\mathbf{x}, y') - G_n(\mathbf{x}, y) \geq L(\mathbf{x}, y') - L(\mathbf{x}, y) - \beta K$
- (iii) if  $s \leq s' \leq \underline{S}(\mathbf{x})$ , then  $v_n(\mathbf{x}, s) \geq v_n(\mathbf{x}, s')$
- (iv) if  $y \leq y' \leq \underline{S}(\mathbf{x})$ , then  $G_n(\mathbf{x}, y') - G_n(\mathbf{x}, y) \leq L(\mathbf{x}, y') - L(\mathbf{x}, y) \leq 0$ .

**Proof of Lemma 12.** (i) This result follows from the  $K$ -convexity of  $v_n(\mathbf{x}, s)$  in  $s$ , which is a direct implication of the second item of Proposition 6.

(ii) This result follows from the definition of  $G_n(\mathbf{x}, y)$ , the previous result (i), and the fact that  $f(y, d)$  is convex and non-decreasing.

(iii)  $G_n(\mathbf{x}, s_n(\mathbf{x})) \leq K + G_n(\mathbf{x}, S_n(\mathbf{x})) \leq K + G_n(\mathbf{x}, \underline{S}(\mathbf{x}))$  implies that  $s_n(\mathbf{x}) \leq S_n(\mathbf{x}) \leq \underline{S}(\mathbf{x})$  (This is an implication of the definitions of  $s_n(\mathbf{x})$  and  $S_n(\mathbf{x})$ , and

the fact that  $\underline{S}(\mathbf{x})$  minimizes  $L(\mathbf{x}, y)$  while  $S_n(\mathbf{x})$  minimizes the sum of  $L(\mathbf{x}, y)$  and a positive term.). It follows from the four cases of  $s \leq s' \leq \underline{S}(\mathbf{x})$  with respect to the value of  $s_n(\mathbf{x})$  that  $v_n(\mathbf{x}, s) \geq v_n(\mathbf{x}, s')$ .

(iv) This result follows from the definition of  $G_n(\mathbf{x}, y)$ , the non-decreasing nature of  $f(y, d)$  in  $y$  and (iii).

□

The proof of Proposition 7 requires four lemmas.

**Lemma 13.** *For all  $n$  and  $\mathbf{x}$ ,  $\underline{S}(\mathbf{x}) = S_0(\mathbf{x}) \leq S_n(\mathbf{x})$ .*

**Lemma 14.** *For all  $n$  and  $\mathbf{x}$ ,  $s_n(\mathbf{x})$  can be selected so that  $s_n(\mathbf{x}) \leq \bar{s}(\mathbf{x})$ .*

**Lemma 15.** *For all  $n$  and  $\mathbf{x}$ ,  $S_n(\mathbf{x})$  can be selected so that  $S_n(\mathbf{x}) \leq \bar{S}(\mathbf{x})$ .*

**Lemma 16.** *For all  $n$  and  $\mathbf{x}$ ,  $\underline{s}(\mathbf{x}) \leq s_n(\mathbf{x})$ .*

**Proof of Proposition 7.** The proof of these results follow from the proofs of Lemmas 2 - 5 in [79]. Proof of Proposition 7(a) follows from Lemmas 13 - 16, and Proposition 7(b) follows from (a) and Proposition 6.

□

### Determining $\Gamma_n(s)$

As was true for the  $K = 0$  case, when  $K > 0$ , there is a finite set of vectors  $\Gamma_n(s)$  such that  $v_n(\mathbf{x}, s) = \min\{\mathbf{x}\boldsymbol{\gamma} : \boldsymbol{\gamma} \in \Gamma_n(s)\}$  for all  $s$ . Note that  $\Gamma_0(s) = \{\underline{0}\}$  for all  $s$ , where  $\underline{0}$  is the column  $N$ -vector having zero in all entries. Given  $\{\Gamma_n(s) : \forall s\}$ , we now present an approach for determining  $\{\Gamma_{n+1}(s) : \forall s\}$ . Recalling Section 3.3.3, let  $\bar{\Gamma} = \{\bar{\gamma}_1, \dots, \bar{\gamma}_M\}$  be such that  $\min_y L(\mathbf{x}, y) = \min\{\mathbf{x}\boldsymbol{\gamma} : \boldsymbol{\gamma} \in \bar{\Gamma}\}$ . Note

$$G_n(\mathbf{x}, y) = L(\mathbf{x}, y) + \beta \sum_{d,z} \sigma(d, z, \mathbf{x}) v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(y, d)),$$

for  $y \in \{d_1, \dots, d_M\}$ . Then,

$$v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), f(y, d)) = \min\{\boldsymbol{\lambda}(d, z, \mathbf{x})\boldsymbol{\gamma} : \boldsymbol{\gamma} \in \Gamma_n(f(y, d))\}.$$

Let  $\Gamma'_n(y)$  be the set of all vectors of the form

$$\bar{\boldsymbol{\gamma}} + \beta \sum_{d,z} \mathbf{P}(d, z) \boldsymbol{\gamma}(d, z),$$

where  $\bar{\boldsymbol{\gamma}} \in \bar{\Gamma}$  and  $\boldsymbol{\gamma}(d, z) \in \Gamma_n(f(y, d))$ . Then,  $G_n(\mathbf{x}, y) = \min\{\mathbf{x}\boldsymbol{\gamma} : \boldsymbol{\gamma} \in \Gamma'_n(y)\}$  and

$$v_{n+1}(\mathbf{x}, s) = \begin{cases} K + G_n(\mathbf{x}, S_n(\mathbf{x})) & s \leq s_n(\mathbf{x}) \\ G_n(\mathbf{x}, s) & \text{otherwise,} \end{cases}$$

where  $S_n(\mathbf{x})$  and  $s_n(\mathbf{x})$  are the smallest integers such that

$$\begin{aligned} G_n(\mathbf{x}, S_n(\mathbf{x})) &\leq G_n(\mathbf{x}, y) \quad \forall y. \\ G_n(\mathbf{x}, s_n(\mathbf{x})) &\leq K + G_n(\mathbf{x}, S_n(\mathbf{x})). \end{aligned}$$

Let  $X_n(s', S')$  be the set of all  $\mathbf{x} \in X$  such that  $s_n(\mathbf{x}) = s'$  and  $S_n(\mathbf{x}) = S'$ . Thus, if  $\mathbf{x} \in X_n(s', S')$ , then  $s'$  and  $S'$  are the smallest integers such that

$$\begin{aligned} G_n(\mathbf{x}, S'(\mathbf{x})) &\leq G_n(\mathbf{x}, y) \quad \forall y. \\ G_n(\mathbf{x}, s'(\mathbf{x})) &\leq K + G_n(\mathbf{x}, S'(\mathbf{x})). \end{aligned}$$

Since  $G_n(\mathbf{x}, y)$  is piecewise linear and convex in  $\mathbf{x}$  for each  $y$ ,  $X_n(s', S')$  is described by a finite set of linear inequalities. We remark that  $\{X_n(s', S') : s' \leq S', \text{ and } X_n(s', S') \neq \emptyset\}$  is a partition of  $X$ . Further, we remark that if  $\overline{X}(\underline{s}, \bar{s}, \underline{S}, \bar{S}) \cap X_n(s', S') \neq \emptyset$ , then search for  $(s', S')$  can be restricted to  $\underline{s} \leq s' \leq \bar{s}$  and  $\underline{S} \leq S' \leq \bar{S}$ . Let  $\Gamma_{n+1}(s) = \{K\mathbf{1} + \boldsymbol{\gamma} :$

$\gamma \in \Gamma'_n(S')\}$  for all  $s \leq s'$ , and let  $\Gamma_{n+1}(s) = \Gamma'_n(s)$  for all  $s > s'$ . Thus,  $v_{n+1}(\boldsymbol{x}, s) = \min\{\boldsymbol{x}\gamma : \gamma \in \Gamma_{n+1}(s)\}$  for all  $s$ .

## APPENDIX C

### CHAPTER 4

#### C1 Analysis for the $L = 1$ Case

Assume  $v_0 = 0$ ,  $v_{n+1} = Hv_n$ , define  $G_n(\mathbf{x}, y) = G(\mathbf{x}, y, v_n)$  for all  $n$ , and let  $y_n^*(\mathbf{x}, C)$  be the smallest value that minimizes  $G_n(\mathbf{x}, y)$  with respect to  $y$ . We remark that

$$v_{n+1}(\mathbf{x}, s, C) = \begin{cases} G_n(\mathbf{x}, s) & \text{if } s \geq y_n^*(\mathbf{x}, C) \\ G_n(\mathbf{x}, s + C) & \text{if } s \leq y_n^*(\mathbf{x}, C) - C \\ G_n(\mathbf{x}, y_n^*(\mathbf{x}, C)) & \text{otherwise.} \end{cases}$$

We now present claims for structured results with respect to  $G_n$ ,  $v_n$ , and  $y_n^*$  based on results in [123] and [115].

**Proposition 3.** *For all  $n$ ,  $\mathbf{x}$ , and  $C$ ,*

- (i)  $G_n(\mathbf{x}, y)$  is convex in  $y$
- (ii)  $v_n(\mathbf{x}, s, C)$  is:
  - (a) convex in  $s$ ,
  - (b) non-decreasing for  $s \geq y_n^*(\mathbf{x}, C)$ ,
  - (c) non-increasing for  $s \leq y_n^*(\mathbf{x}, C) - C$ ,
  - (d) equal to  $v_n(\mathbf{x}, y_n^*(\mathbf{x}, C), C)$  otherwise
- (iii)  $v_{n+1}(\mathbf{x}, s, C) \geq v_n(\mathbf{x}, s, C)$  for all  $s$ .

**Proof of Proposition 3.** The convexity of  $G_0(\mathbf{x}, y)$  in  $y$  for all  $\mathbf{x}$  follows from the definitions and assumptions. Assume  $G_n(\mathbf{x}, y)$  is convex in  $y$  for all  $\mathbf{x}$ . It is then straightforward to show that item ii holds for  $n = n + 1$  and all  $(\mathbf{x}, C)$ . We remark that the

function  $g(y) = w(f(y))$  is convex and non-decreasing (non-increasing) if  $w$  is convex and non-decreasing (non-increasing) and if  $f$  is linear and non-decreasing. Hence,  $G_{n+1}(\mathbf{x}, y)$  is convex in  $y$  for all  $\mathbf{x}$ , and item i and item ii hold for all  $n$  by induction. Since  $v_1(\mathbf{x}, s, C) \geq v_0(\mathbf{x}, s, C)$ , a standard induction argument guarantees that item iii holds.  $\square$

Let  $v_n(\mathbf{x}, s) = v_n(\mathbf{x}, s, C)$ ,  $v'_n(\mathbf{x}, s) = v_n(\mathbf{x}, s, G)$ ,  $G_n(\mathbf{x}, y) = G(\mathbf{x}, y, v_n)$ , and  $G'_n(\mathbf{x}, y) = G(\mathbf{x}, y, v'_n)$ .

**Proposition 4.** Assume  $C \leq C'$ , and that  $y_n^*(\mathbf{x}, C) - d \leq y_n^*(\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x}), C)$  for all  $n$  and all  $(\mathbf{d}, \mathbf{z}, \mathbf{x})$ . Then for all  $n$ ,  $\mathbf{x}$ , and  $s$ ,

$$(i) \quad v'_n(\mathbf{x}, s, C) \leq v_n(\mathbf{x}, s, C)$$

$$(ii) \quad \text{If } y \leq y' \leq y_n^*(\mathbf{x}, C), \text{ then } G_n(\mathbf{x}, y') - G_n(\mathbf{x}, y) \leq G'_n(\mathbf{x}, y') - G'_n(\mathbf{x}, y)$$

$$(iii) \quad \text{If } s \leq s' \leq y_n^*(\mathbf{x}, C), \text{ then } v_{n+1}(\mathbf{x}, s', C) - v_{n+1}(\mathbf{x}, s, C) \leq v'_{n+1}(\mathbf{x}, s', C) - v'_{n+1}(\mathbf{x}, s, C).$$

$$(iv) \quad y_n^*(\mathbf{x}, C') \leq y_n^*(\mathbf{x}, C).$$

**Proof of Proposition 4.** Proof of item i is straightforward. Regarding item ii-item iv, note item ii holds for  $n = 0$ ; assume item ii holds for  $n$ . Then item iv also holds for  $n$ . We now outline the proof that item iii holds for  $n = n + 1$ . Recall

$$v_{n+1}(\mathbf{x}, s, C) = \begin{cases} G_n(\mathbf{x}, s + C) & \text{if } s \leq y_n - C \\ G_n(\mathbf{x}, s) & \text{if } s \geq y_n \\ G_n(\mathbf{x}, y_n) & \text{otherwise,} \end{cases}$$

where  $y_n = y_n^*(\mathbf{x}, C)$ , and

$$v'_{n+1}(\mathbf{x}, s, C) = \begin{cases} G'_n(\mathbf{x}, s + C') & \text{if } s \leq y'_n - C' \\ G'_n(\mathbf{x}, s) & \text{if } s \geq y'_n \\ G'_n(\mathbf{x}, y'_n) & \text{otherwise,} \end{cases}$$



where  $y'_n = y_n^*(\mathbf{x}, G)$ . Similar to the proof of 5 and the proof of [123, Theorem 3], there are two cases: (1)  $y_n - C \leq y'_n$ , (2)  $y'_n \leq y_n - C$ , which are more completely described as

$$y'_n - C' \leq y_n - C \leq y'_n \leq y_n,$$

$$y'_n - C' \leq y'_n \leq y_n - C \leq y_n,$$

respectively. For each case, there are 10 different sets of inequalities that the pair  $(s, s')$  can satisfy. Showing that item iii holds when  $n = n + 1$  for each of the 20 sets of inequalities is tedious but straightforward. We now show that for  $s \leq s'$ ,

$$v_{n+1}(\mathbf{x}, s', C) - v_{n+1}(\mathbf{x}, s, C) \leq v'_{n+1}(\mathbf{x}, s', C) - v'_{n+1}(\mathbf{x}, s, C)$$

implies that for  $y \leq y' \leq y_n$ ,  $G_{n+1}(\mathbf{x}, y') - G_{n+1}(\mathbf{x}, y) \leq G'_{n+1}(\mathbf{x}, y') - G'_{n+1}(\mathbf{x}, y)$ . Note

$$v_{n+1}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y' - d, C) - v_{n+1}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y - d, C) \leq$$

$$v'_{n+1}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y' - d, C) - v'_{n+1}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y - d, C)$$

for  $y - d \leq y' - d \leq y_n^*(\boldsymbol{\lambda}(d, z, \mathbf{x}), C)$ , which implies

$$G_{n+1}(\mathbf{x}, y') - G_{n+1}(\mathbf{x}, y) \leq G'_{n+1}(\mathbf{x}, y') - G'_{n+1}(\mathbf{x}, y)$$

for all  $y \leq y' \leq y_{n+1}^*(\mathbf{x}, C)$  assuming  $y_{n+1}^*(\mathbf{x}, C) - d \leq y_{n+1}^*(\boldsymbol{\lambda}(d, z, \mathbf{x}), C)$  for all  $(d, z, \mathbf{x})$ . A standard induction argument completes the proof.  $\square$

**Proposition 5.** Assume  $y_n^*(\mathbf{x}, C) - d_l \leq y_n^*(\boldsymbol{\lambda}(d, z, \mathbf{x}), C)$  for all  $n$  and all  $(d, z, \mathbf{x})$ . Then for all  $n$ ,  $s \leq s' \leq y_n^*(\mathbf{x}, C)$  implies:

$$(i) \quad v_n(\mathbf{x}, s', C) - v_n(\mathbf{x}, s, C) \geq v_{n+1}(\mathbf{x}, s', C) - v_{n+1}(\mathbf{x}, s, C),$$

$$(ii) \quad G_n(\mathbf{x}, s') - G_n(\mathbf{x}, s) \geq G_{n+1}(\mathbf{x}, s') - G_{n+1}(\mathbf{x}, s),$$

$$(iii) \quad y_n^*(\mathbf{x}, C) \leq y_{n+1}^*(\mathbf{x}, C).$$

**Proof of Proposition 5.** We note item i holds when  $n = 0$ . Assume item i holds for  $n = n - 1$ . Let  $y \leq y' \leq y_{n-1}^*(\mathbf{x}, C)$ , implying that  $y - d \leq y' - d \leq y_{n-1}^*(\mathbf{x}, C) - d \leq y_{n-1}^*(\boldsymbol{\lambda}(d, z, \mathbf{x}), C)$  for all  $(d, z, \mathbf{x})$ . Hence,

$$v_{n-1}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y' - d, C) - v_{n-1}(\boldsymbol{\lambda}(d, z, \mathbf{x}), y - d, C) \geq$$

$$v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), y' - d, C) - v_n(\boldsymbol{\lambda}(d, z, \mathbf{x}), y - d, C),$$

and thus item ii holds for  $n = n - 1$  for all  $y \leq y' \leq y_{n-1}^*(\mathbf{x}, C)$ . Letting  $y' = y_{n-1}^*(\mathbf{x}, C)$ , we observe

$$0 \geq G_{n-1}(\mathbf{x}, y_{n-1}^*(\mathbf{x}, C)) - G_{n-1}(\mathbf{x}, y) \geq G_n(\mathbf{x}, y_{n-1}^*(\mathbf{x}, C)) - G_n(\mathbf{x}, y);$$

hence, item iii holds for  $n = n - 1$ .

We now outline a proof that  $s \leq s' \leq y_n^*(\mathbf{x}, C)$  implies

$$v_n(\mathbf{x}, s') - v_n(\mathbf{x}, s) \geq v_{n+1}(\mathbf{x}, s') - v_{n+1}(\mathbf{x}, s). \quad (\text{C.1})$$

Following an argument in the proof of [123, Theorem 2], we consider two general cases: (1)  $y_n^*(\mathbf{x}, C) - C \leq y_{n-1}^*(\mathbf{x}, C)$  and (2)  $y_{n-1}^*(\mathbf{x}, C) \leq y_n^*(\mathbf{x}, C) - C$ . Letting the dependence on  $(\mathbf{x}, C)$  be implicit, cases (1) and (2) are more completely described as

$$y_{n-1}^* - C \leq y_n^* - C \leq y_{n-1}^* \leq y_n^*$$

$$y_{n-1}^* - C \leq y_{n-1}^* \leq y_n^* - C \leq y_n^*,$$

respectively. For each case, there are 10 different sets of inequalities that the pair  $(s, s')$  can satisfy. The values  $v_n(\mathbf{x}, s')$ ,  $v_n(\mathbf{x}, s)$ ,  $v_{n+1}(\mathbf{x}, s')$ , and  $v_{n+1}(\mathbf{x}, s)$  are well defined for each

of these inequalities in terms of  $G_{n-1}$  and  $G_n$ . Showing that (C.1) holds for each of these 20 different sets of inequalities is tedious but straightforward.

A standard induction argument completes the proof of the proposition.  $\square$

We now claim that  $v(\mathbf{x}, s, C)$  is convex in  $C$ .

**Proposition 6.** (i) *If  $y \in A(s, C)$  and  $y' \in A(s, C')$ , then  $\lambda y + (1 - \lambda)y' \in A(s, \lambda C + (1 - \lambda)C')$ .*

(ii) *If  $\xi \in A(s, \lambda C + (1 - \lambda)C')$ , then there is a  $y \in A(s, C)$  and a  $y' \in A(s, C')$  such that  $\xi = \lambda y + (1 - \lambda)y'$ .*

(iii) *For real-valued and continuous  $v$ ,*

$$\begin{aligned} & \min\{v(\xi) : \xi \in A(s, \lambda C + (1 - \lambda)C')\} \\ &= \min\{v(\lambda y + (1 - \lambda)y') : y \in A(s, C) \text{ and } y' \in A(s, C')\}. \end{aligned}$$

(iv) *For all  $(\mathbf{x}, s)$  and  $n$ ,  $v_n(\mathbf{x}, s, C)$  is convex in  $C$ .*

**Proof of Proposition 6.** (i)  $y \in A(s, C)$  and  $y' \in A(s, C')$  imply  $\lambda s \leq \lambda y \leq \lambda(s + C)$  and  $(1 - \lambda)s \leq (1 - \lambda)y' \leq (1 - \lambda)(s + C')$ ; summing terms implies the result.

(ii) Let  $X = (\lambda C + (1 - \lambda)C' + s)$  and  $\Delta^S = (X - \xi)/(X - s)$ . Note  $\Delta^S \in [0, 1]$  and  $\xi = \Delta^S s + (1 - \Delta^S)X$ . Let  $y = \Delta^S s + (1 - \Delta^S)(s + C)$  and  $y' = \Delta^S s + (1 - \Delta^S)(s + C')$ . Then,  $y \in A(s, C)$ ,  $y' \in A(s, C')$ , and  $\lambda y + (1 - \lambda)y' = \xi$ .

(iii) Proof by contradiction follows from items i and ii.

(iv) From item iii and the convexity of  $G_n(\mathbf{x}, y)$  in  $y$  for all  $n$  and  $y$  (by Proposition 3

item i), it follows that

$$\begin{aligned}
& v_n(\mathbf{x}, s, \lambda C + (1 - \lambda)C') \\
& \min\{G_n(\mathbf{x}, \lambda y + (1 - \lambda)y') : y \in A(s, C), y' \in A(s, C')\} \\
& \leq \min\{\lambda G_n(\mathbf{x}, y) + (1 - \lambda)G_n(\mathbf{x}, y') : y \in A(s, C), y' \in A(s, C')\} \\
& = \lambda v_n(\mathbf{x}, s, C) + (1 - \lambda)v_n(\mathbf{x}, s, C').
\end{aligned}$$

□

Clearly, the assumption that  $y_n^*(\mathbf{x}, C) - d_l \leq y_n^*(\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x}), C)$  for all  $n$  and all  $(\mathbf{d}, \mathbf{z}, \mathbf{x})$  is in general a challenge to verify *a priori*. Arguments in [123] suggest that as  $n$  gets large,  $y_n^*(\mathbf{x}, C)$  may converge in some sense to a function  $y_\infty^*(\mathbf{x}, C)$ . From [115],  $y_0^*(\mathbf{x}, C)$  is straightforward to determine. Let  $\hat{y}(\mathbf{x}, C) \geq y_\infty^*(\mathbf{x}, C) \geq y_n^*(\mathbf{x}, C)$  for all  $n$  and  $\mathbf{x}$ . Then  $\hat{y}(\mathbf{x}, C) - d_l \leq y_0^*(\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x}), C)$  for all  $(\mathbf{d}, \mathbf{z}, \mathbf{x})$  implies the above assumption holds. Determination of a function  $\hat{y}$  for the general case is a topic for future research. We present a special case where  $y_0^* = y_n^*$  for all  $n$  in appendix section.

We point out two key differences between the infinite capacity and the finite capacity cases when the reorder cost,  $K' = 0$ . First, when  $C$  is infinite, the smallest optimal base stock level  $y_n^*(\mathbf{x})$  is independent of the number of successive approximation steps, making it (relatively) easy to determine. Unfortunately, this result does not appear to hold when  $C$  is finite except for the situation considered in Proposition 7. This apparent fact has implementation implications for the controllers at the locations (determining the base stock levels for the capacitated case will in general be more difficult than for the infinite capacity case).

Second, Claims 4 and 6 state that  $v(\mathbf{x}, s, C)$  is non-decreasing and convex in  $C$ . We also know that  $v(\mathbf{x}, s, C)$  is convex in  $s$  (from Proposition 3, which is also true for the infinite capacity case) and concave and possibly piecewise linear in  $\mathbf{x}$  (from earlier cited results, which is also true for the infinite capacity case). We will find later that these structural

results will be computationally useful in determining solutions to the GC problem. The GC problem for determining  $(\Delta^S, \sigma, u')$ , given  $(x, s, u)$ , requires knowing  $v_l(x, s'_l, u'_l)$  for all  $l$ . We now consider approaches to compute or approximate  $v(x, s, C)$ .

We now present a special case where we claim that  $y_0^* = y_n^*$  for all  $n$ .

**Proposition 7.** *Assume that for all  $(d, z, x)$ ,  $y_0^*(\lambda(d, z, x), C) - C \leq y_0^*(x, C) - d \leq y_0^*(\lambda(d, z, x), C)$ . Then,  $y_n^*(x, C) = y_0^*(x, C)$  for all  $n$ .*

We remark that the left inequality in Proposition 7 essentially implies that although capacity may be finite, it is always sufficient to insure the inventory level after replenishment can be  $y_0^*(x, C)$ .

**Proof of Proposition 7.** By induction. Assume  $y_n^*(x, C) = y_0^*(x, C)$ . Note therefore,

$$v_{n+1}(x, s, C) = \begin{cases} G_n(x, s + C), & s \leq y_0^*(x, C) - C \\ G_n(x, s), & s \geq y_0^*(x, C) \\ G_n(x, y_0^*(x, C)) & \text{otherwise.} \end{cases}$$

Note

$$(i) \min_y G_{n+1}(x, y) \leq G_{n+1}(x, y_0^*(x, C))$$

$$(ii) \min_y G_{n+1}(x, y) \geq \min_y L(x, y) + \beta \sum_{d,z} \sigma(d, z, x) \min_y v_{n+1}(\lambda(d, z, x), y - d).$$

The minimum with respect to  $y$   $v_{n+1}(\lambda(d, z, x), y - d, C)$  is such that  $y_0^*(\lambda(d, z, x), C) - C \leq y - d \leq y_0^*(\lambda(d, z, x), C)$ . By assumption,  $y = y_0^*(x, C)$  satisfies these inequalities.

Thus,

$$\begin{aligned} \min_y G_{n+1}(x, y) &\geq L(x, y_0^*(x, y_0^*(x, C))) \\ &\quad + \beta \sum_{d,z} \sigma(d, z, x) v_{n+1}(\lambda(d, z, x), y_0^*(x, C) - d, C) \\ &= G_{n+1}(x, y_0^*(x, C)), \end{aligned}$$

and hence  $y_{n+1}^*(\mathbf{x}, C) = y_0^*(\mathbf{x}, C)$ . □

**Determination of  $\Gamma_n(s, C)$ :** Earlier results cited state that for each  $n$  and  $(s, C)$ , there is a finite set of vectors  $\Gamma_n(s, C)$  such that  $v_n(\mathbf{x}, s, C) = \min\{\mathbf{x}\gamma : \gamma \in \Gamma_n(s, C)\}$ .

We remark that determination of the  $\{\Gamma_n(s, C)\}$  is a numerical challenge for the standard POMDP (where there is only a single  $s$  and a single  $C$ ). Therefore, determining  $\{\Gamma_n(s, C)\}$  for all (or large enough)  $n$  and  $(s, C)$  may represent a formidable computational challenge.

The following result may be computationally exploitable. Assuming that

$$v_n(\mathbf{x}, s, C) = \min\{\mathbf{x}\gamma : \gamma \in \Gamma_n(s, C)\},$$

$$\begin{aligned} v_{n+1}(\mathbf{x}, s, C) &= \min_{y \in A(s, C)} \{L(\mathbf{x}, y) + \beta \sum_{\mathbf{d}, \mathbf{z}} \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) v_n(\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x}), y - d_l, C)\} \\ &= \min_{y \in A(s, C)} \{L(\mathbf{x}, y) + \beta \sum_{\mathbf{d}, \mathbf{z}} \sigma(\mathbf{d}, \mathbf{z}, \mathbf{x}) \min\{\boldsymbol{\lambda}(\mathbf{d}, \mathbf{z}, \mathbf{x})\gamma : \\ &\quad \gamma \in \Gamma_n(y - d_l, C)\}\} \\ &= \min_{y \in A(s, C)} \{\mathbf{x}\hat{\gamma}(y) + \beta \sum_{\mathbf{d}, \mathbf{z}} \min\{\mathbf{x}P(\mathbf{d}, \mathbf{z})\gamma : \gamma \in \Gamma_n(y - d_l, C)\}\} \end{aligned}$$

where  $\hat{\gamma}(y) = \sum_{\mathbf{d}, \mathbf{z}} P(\mathbf{d}, \mathbf{z}) \mathbf{1}_C(y, d_l)$ .

Let  $\hat{\Gamma}(s, C, y) = \{\hat{\gamma}(y) + \beta \sum_{\mathbf{d}, \mathbf{z}} P(\mathbf{d}, \mathbf{z}) \gamma_{\mathbf{d}, \mathbf{z}} : \forall \gamma_{\mathbf{d}, \mathbf{z}} \in \Gamma_n(y - d_l, C), \forall (\mathbf{d}, \mathbf{z})\}$ . Define  $\Gamma_{n+1}(s, C) = \cup_{y \in A(s, C)} \hat{\Gamma}_n(s, C, y)$ . Thus,  $v_{n+1}(\mathbf{x}, s, C) = \min\{\mathbf{x}\gamma : \gamma \in \Gamma_{n+1}(s, C)\}$ .

We therefore note that

$$\Gamma_{n+1}(s + 1, C) = \Gamma_{n+1}(s, C) \cup \hat{\Gamma}_n(s, C, s + C + 1) \sim \hat{\Gamma}_n(s, C, s).$$

**Finite-Memory Approximation:** As noted earlier, results in [69] and elsewhere imply that  $\{(\mathbf{x}(t), \mathbf{s}(t)), t = 0, 1, \dots\}$  is a sufficient statistic for the  $(L = 1, Y = 0)$  problem. We note that given  $\{\mathbf{d}(t), \dots, \mathbf{d}(t - \Delta^S + 1), \mathbf{z}(t), \dots, \mathbf{z}(t - \Delta^S + 1), \mathbf{x}(t - \Delta^S)\}$ , we can

determine  $\mathbf{x}(t)$ . Hence, there is a family of sufficient statistics for the  $L$  location,  $Y$  module problem. Further, under reasonable assumptions (see [124] and [69]), the larger  $\Delta^S$ , the less influential  $\mathbf{x}(t - \Delta^S)$  is for determining  $\mathbf{x}(t)$ , suggesting the possibility of exploring finite-memory policies for determining approximations to  $v(\mathbf{x}, s, C)$ .

**Finite Horizon Approximation:** For the case where the modulation process is completely observed (or where the above finite-memory approximation is assumed) and a single value iteration step is computationally intensive (due to, for example, a large state space), it may be useful to approximate the salvage value of a finite horizon MDP in order to approximate the infinite horizon MDP. Bounds on cost, the discount factor, the coefficient of ergodicity, and the length of the finite horizon can be used to bound the salvage value for a completely observed MDP that can be solved by specially structured POMDP solution techniques [125].

## C2 Proof of Proposition 1

*Proof.* Let  $v_0(\mathbf{x}, s, C) = \hat{v}_0(\mathbf{x}, s, C) = 0$ . Consider  $\mathbf{d} = (d_l, \mathbf{d}_{j \neq l})$ , where  $\mathbf{d}_{j \neq l}$  can be considered as additional observation data  $z$ . Let  $\sum_z \sigma(d_l, z, \mathbf{x}) = \sigma(d_l, \mathbf{x})$ .

$$\begin{aligned}
v_1(\mathbf{x}, s, C) &= \min_{s \leq y \leq s+C} \left\{ \sum_{d_l} \sigma(d_l, \mathbf{x}) [c(y, d_l)] \right\} \\
&= \min_{s \leq y \leq s+C} \left\{ \sum_{d_l} \sum_i x_i \sum_j \Pr(j | i) \Pr(d_l | j) [c(y, d_l)] \right\} \\
&= \min_{s \leq y \leq s+C} \left\{ \sum_d \sum_i x_i \left( \Pr(d_l | i) + \sum_j \Pr(j | i) \Pr(d_l | j) - \Pr(d_l | i) \right) [c(y, d_l)] \right\} \\
&\geq \min_{s \leq y \leq s+C} \left\{ \sum_{d_l} \sum_i x_i \left( \Pr(d_l | i) - \max_k \Pr(d_l | k) + \min_k \Pr(d_l | k) \right) [c(y, d_l)] \right\} \\
&\geq \min_{s \leq y \leq s+C} \left\{ \sum_{d_l} \sum_i x_i \Pr(d_l | i) c(y, d_l) \right. \\
&\quad \left. - \sum_{d_l} \left( \max_k \Pr(d_l | k) - \min_k \Pr(d_l | k) \right) c(y, d_l) \right\}
\end{aligned}$$

$$\begin{aligned}
&\geq \min_{s \leq y \leq s+C} \left\{ \sum_{d_l} \sum_i x_i \Pr(d_l \mid i) c(y, d_l) \right\} \\
&\quad + \min_{s \leq y \leq s+C} \left\{ - \sum_{d_l} \left( \max_k \Pr(d_l \mid k) - \min_k \Pr(d_l \mid k) \right) c(y, d_l) \right\} \\
&= \hat{v}_1(\mathbf{x}, s, C) + \min_{s \leq y \leq s+C} \left\{ - \sum_{d_l} k(d_l) c(y, d_l) \right\} \\
&= \hat{v}_1(\mathbf{x}, s, C) - \max_{s \leq y \leq s+C} \left\{ \sum_{d_l} k(d_l) c(y, d_l) \right\} \\
&= \hat{v}_1(\mathbf{x}, s, C) - \sum_{d_l} k(d_l) c(\hat{y}, d_l) = \hat{v}_1(\mathbf{x}, s, C) - u, \text{ where } u = \sum_{d_l} k(d_l) c(\hat{y}, d_l) \text{ and} \\
&\quad \hat{y} \in \{s, s+C\} \text{ due to convexity of } c(y, d_l) \forall y, d_l, \\
&\quad \text{where } k(d_l) = \left( \max_k \Pr(d_l \mid k) - \min_k \Pr(d_l \mid k) \right).
\end{aligned}$$

By induction and infinite summation,

$$v_n(\mathbf{x}, s, C) \geq \hat{v}_n(\mathbf{x}, s, C) - u(1 + \beta + \cdots + \beta^n); v(\mathbf{x}, s, C) \geq \hat{v}(\mathbf{x}, s, C) - u/(1 - \beta).$$

□



### C3 Lookahead with Global-Local Stationary Future (LAGL)

$$\begin{aligned}
\text{LAGL} \quad & \min_{\Delta^{S,u',y}} \sum_l \left\{ (K_l^{S+} \Delta_l^{S+} + K_l^{S-} \Delta_l^{S-}) + K^M \sum_l |u_l - u'_l|/2 + (\zeta_l + \eta_l)/2 \right\}, \\
& \text{subject to} \\
& \zeta_l \geq \gamma_j^l (s_l + \Delta_l^{S+} - \Delta_l^{S-}) + \widehat{\gamma}_j^l \quad \forall (\gamma_j^l, \widehat{\gamma}_j^l) \in \Gamma_{t+1}^l(u_l) \quad \forall l \\
& \eta_l \geq \theta_j^l u'_l + \widehat{\theta}_j^l \quad \forall (\theta_j^l, \widehat{\theta}_j^l) \in \Theta_{t+1}^l(s_l) \quad \forall l \\
& \sum_l u'_l = Y \\
& \sum_l \Delta_l^{S+} = \sum_l \Delta_l^{S-}, \\
& 0 \leq u'_l \leq Y'_l, \quad \forall l \\
& 0 \leq \Delta_l^{S+} \leq \sum_{k \neq l} (s_k)^+, \quad \forall l \\
& 0 \leq \Delta_l^{S-} \leq -(s_l)^+, \quad \forall l \\
& u'_l, \Delta_l^{S+}, \Delta_l^{S-} \in \mathbb{Z}, \quad \eta_l, \zeta_l \in \mathbb{R} \quad \forall l
\end{aligned} \tag{C.2}$$

**Proposition 8.** *LAGLR can be solved exactly by relaxing the integrality constraints.*

### C4 Results: Additional Tables

Table C.1: Average performance of GLR w.r.t. NF across  $L$  and  $\theta$

$L \backslash \theta$	0	0.2	0.5	0.8	1
2	1.66	0.90	0.91	0.91	0.92
3	1.09	0.70	0.71	0.73	0.74
5	1.44	0.83	0.83	0.84	0.85
10	1.74	0.73	0.74	0.77	0.78
Overall	1.48	0.79	0.80	0.81	0.82

Table C.2: Average performance of LAJ w.r.t. NF across  $L$  and  $\theta$ 

$L \backslash \theta$	0	0.2	0.5	0.8	1
2	0.89	0.90	0.94	0.97	0.99
3	0.73	0.69	0.75	0.79	0.81
5	0.84	0.80	0.84	0.89	0.91
10	0.76	0.70	0.74	0.81	0.84
Overall	0.80	0.77	0.82	0.86	0.89

Table C.3: Average performance of JR w.r.t. NF across  $L$  and  $\theta$ 

$L \backslash \theta$	0	0.2	0.5	0.8	1
2	0.89	0.89	0.90	0.92	0.93
3	0.71	0.66	0.67	0.70	0.72
5	0.85	0.80	0.80	0.81	0.83
10	0.74	0.67	0.69	0.73	0.75
JR - Overall	0.80	0.76	0.77	0.79	0.81

Table C.4: Average value of mobility (% savings over NF) using JR with  $\theta = 0.2$  across  $K^S$  and  $K^M$ 

		Module movement cost $K^M$				
		0	1.5	2	2.5	10000
Transshipment cost $K^S$	0	34%	34%	38%	36%	36%
	1.5	32%	23.4%	19.3%	19%	18.3%
	2	34%	23.3%	16.8%	20.7%	19.8%
	2.5	36%	20.9%	23.3%	18.2%	15.4%
	10000	34%	22.5%	18.8%	17.8%	0.3%

## C5 DM Assumes Epoch-invariant System

Table C.5: Average performance of heuristics w.r.t. NF across  $L$ , when the DM assumes epoch-invariant and epoch-variant demand distributions in a Markov-modulated system with partial observability

Steady-state, stationary demands				
$L$	GLR ( $\theta = 0.8$ )	GLR ( $\theta = 0.2$ )	JR( $\theta = 0.2$ )	LAJ( $\theta = 0.2$ )
2	0.883	0.906	0.876	0.914
3	0.679	0.734	0.674	0.697
5	0.797	0.834	0.804	0.806
10	0.696	0.743	0.691	0.701
Overall	0.764	0.804	0.761	0.780
Partially observed, Markov-modulated demands				
$L$	GLR ( $\theta = 0.8$ )	GLR ( $\theta = 0.2$ )	JR( $\theta = 0.2$ )	LAJ( $\theta = 0.2$ )
2	0.912	0.903	0.890	0.903
3	0.727	0.705	0.661	0.690
5	0.844	0.833	0.797	0.796
10	0.767	0.726	0.673	0.695
Overall	0.812	0.792	0.755	0.771

## C6 Complete Observability of Modulation Process

Table C.6: Value of mobility (% savings over NF) under complete observability using JR with  $\theta = 0.2$  across  $K^S$  and  $K^M$  for  $L = 10$

		Module movement cost $K^M$				
		0	1.5	2	2.5	10000
Transshipment cost $K^S$	0	48%	51%	48%	48%	47%
	1.5	38%	33.4%	33.1%	25.6%	24.4%
	2	46%	32.2%	26.8%	26.7%	26%
	2.5	44%	33%	26%	25.9%	22.3%
	10000	43%	28.5%	19.7%	21%	0.6%

## REFERENCES

- [1] Deloitte, “The digital transformation of customer services - Our point of view,” Tech. Rep., 2013, pp. 1–24.
- [2] F3, “Flexible, fast and future production processes,” Bayer Technology Services GMBH, Tech. Rep., 2014.
- [3] Pfizer, *Pfizer announces collaboration with GSK on next-generation design of portable, continuous, miniature and modular (PCMM) oral solid dose development and manufacturing units*, 2015.
- [4] MIT, *Pharmacy on demand*, 2016.
- [5] Amazon, *Amazon 3d printing patent*, 2015.
- [6] Geek Wire, *Amazon finally wins a patent for 3-D printing on demand, for pickup or delivery*, 2018.
- [7] S. Lier and M. Grünewald, “Net present value analysis of modular chemical production plants,” *Chemical Engineering & Technology*, vol. 34, no. 5, pp. 809–816, 2011.
- [8] S. Lier, D. Wörsdörfer, and J. Gesing, “Business models and product service systems for transformable, modular plants in the chemical process industry,” in *Product-Service Integration for Sustainable Solutions: Proceedings of the 5th CIRP International Conference on Industrial Product-Service Systems, Bochum, Germany, March 14th - 15th, 2013*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 227–238, ISBN: 978-3-642-30820-8.
- [9] U. Clausen, S. Langkau, I. Goedicke, and S. Lier, “Location and network planning for modular container plants in the process industry,” *Chemie Ingenieur Technik*, vol. 87, no. 10, pp. 1365–1372, 2015.
- [10] S. Lier, D. Wörsdörfer, and M. Grünewald, “Transformable production concepts: Flexible, mobile, decentralized, modular, fast,” *ChemBioEng Reviews*, vol. 3, no. 1, 2016.
- [11] D. Wörsdörfer and S. Lier, “Optimized modular production networks in the process industry,” in *Operations Research Proceedings 2015*, Springer International Publishing, 2017.

- [12] D. Wörsdörfer, S. Lier, and N. Crasselt, “Real options-based evaluation model for transformable plant designs in the process industry,” *Journal of Manufacturing Systems*, vol. 42, pp. 29–43, 2017.
- [13] S. Kessler and L. Bruell, “New production concepts for the process industry require modular logistics solutions,” *ChemBioEng Reviews*, vol. 2, no. 5, pp. 351–355, 2015.
- [14] Y. Koren, “Reconfigurable manufacturing systems,” in, ser. The Global Manufacturing Revolution: Product-Process-Business Integration and Reconfigurable Systems, John Wiley & Sons, Inc., 2010.
- [15] Z. M. Bi, S. Y. T. Lang, W. Shen, and L. Wang, “Reconfigurable manufacturing systems: the state of the art,” *International Journal of Production Research*, vol. 46, no. 4, pp. 967–992, 2008.
- [16] G. G. Rogers and L. Bottaci, “Modular production systems: a new manufacturing paradigm,” *Journal of Intelligent Manufacturing*, vol. 8, no. 2, pp. 147–156, 1997.
- [17] S. Marcotte, B. Montreuil, and L. C. Coelho, “Modeling of physical internet enabled hyperconnected modular production,” in *Proceedings of 2nd Physical Internet Conference, 2015*, 2015.
- [18] S. Marcotte and B. Montreuil, “Introducing the concept of hyperconnected mobile production,” *Progress in Material Handling Research*, 2016.
- [19] S. D. Jena, J.-F. Cordeau, and B. Gendron, “Dynamic facility location with generalized modular capacities,” *Transportation Science*, vol. 49, no. 3, pp. 489–499, 2015.
- [20] G. Ghiani, F. Guerriero, and R. Musmanno, “The capacitated plant location problem with multiple facilities in the same site,” *Computers & Operations Research*, vol. 29, no. 13, pp. 1903–1912, 2002.
- [21] M. T. Melo, S. Nickel, and F. da Gama, “Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning,” *Computers & Operations Research*, vol. 33, no. 1, pp. 181–208, 2005.
- [22] R. Halper and S. Raghavan, “The mobile facility routing problem,” *Transportation Science*, vol. 45, no. 3, pp. 413–434, 2011.
- [23] A. Angelus and E. L. Porteus, “Simultaneous capacity and production management of short-life-cycle, produce-to-stock goods under stochastic demand,” *Management Science*, vol. 48, no. 3, pp. 399–413, 2002.

- [24] J. R. Bradley and P. W. Glynn, “Managing capacity and inventory jointly in manufacturing systems,” *Management Science*, vol. 48, no. 2, pp. 273–288, 2002.
- [25] S. Rajagopalan and J. M. Swaminathan, “A coordinated production planning model with capacity expansion and inventory management,” *Management Science*, vol. 47, no. 1, pp. 1562–1580, 2001.
- [26] J. Qiu and T. C. Sharkey, “Integrated dynamic single-facility location and inventory planning problems,” *IIE Transactions (Institute of Industrial Engineers)*, vol. 45, no. 8, pp. 883–895, 2013.
- [27] A. V. Iyer and A. Jain, “Modeling the impact of merging capacity in production-inventory systems,” *Management Science*, vol. 50, no. 8, pp. 1082–1094, 2004.
- [28] W. B. Powell, *Approximate dynamic programming: Solving the curses of dimensionality (Wiley Series in Probability and Statistics)*, Second. Hoboken, New Jersey: John Wiley & Sons, Inc., 2007, pp. 1–638, ISBN: 9781118029176.
- [29] J. C. Goodson, B. W. Thomas, and J. W. Ohlmann, “A rollout algorithm framework for heuristic solutions to finite-horizon stochastic dynamic programs,” *European Journal of Operational Research*, vol. 258, no. 1, pp. 216–229, 2017.
- [30] N. Secomandi, “A rollout policy for the vehicle routing problem with stochastic demands,” *Operations Research*, vol. 49, no. 5, pp. 796–802, 2001.
- [31] D. P. Bertsekas, J. N. Tsitsiklis, and C. Wu, “Rollout algorithms for combinatorial optimization,” *Journal of Heuristics*, vol. 3, pp. 245–262, 1997.
- [32] G. Tesauro and G. R. Galperin, “On-line policy improvement using Monte Carlo search,” *Advances in Neural Information Processing Systems*, vol. 9, pp. 1068–1074, 1996.
- [33] I. O. Ryzhov, W. B. Powell, and P. I. Frazier, “The knowledge gradient algorithm for a general class of online learning problems,” *Operations Research*, vol. 60, no. 1, pp. 180–195, 2012.
- [34] T. Ibaraki, T. Hasegawa, K. Teranaka, and J. Iwase, “The multiple-choice knapsack problem,” *Journal of the Operations Research Society of Japan*, vol. 21, no. 1, pp. 59–95, 1978.
- [35] P. Sinha and A. A. Zoltners, “The multiple-choice knapsack problem,” *Operations Research*, vol. 27, no. 3, pp. 503–515, 1979.
- [36] D. P. Bertsekas, *Dynamic programming and optimal control*, 3rd. Athena Scientific, 2000, ISBN: 1886529094.

- [37] W. Chen, M. Dawande, and G. Janakiraman, “Integrality in stochastic inventory models,” *Production and Operations Management*, vol. 23, no. 9, pp. 1646–1663, 2014.
- [38] Y. Aviv and A. Federgruen, “Stochastic inventory models with limited production capacity and periodically varying parameters,” *Probability in the Engineering and Informational Sciences*, vol. 11, pp. 107–135, 1997.
- [39] S. C. Graves, A. H. G. Rinnooy Kan, and P. H. Zipkin, *Logistics of production and inventory*, ser. Handbooks in Operations Research and Management Science. Elsevier, 1993, vol. 4.
- [40] T.-M. Choi, *Handbook of newsvendor problems: Models, extensions and applications*, T.-M. Choi, Ed. New York: Springer, 2014, ISBN: 978-1-4614-3599-0 (Print) 978-1-4614-3600-3 (Online).
- [41] M. Khouja, “The single-period (news-vendor) problem: literature review and suggestions for future research,” *Omega*, vol. 27, no. 5, pp. 537–553, 1999.
- [42] N. C. Petruzzi and M. Dada, “Pricing and the newsvendor problem: A review with extensions,” *Operations Research*, vol. 47, no. 2, pp. 183–194, 1999.
- [43] Y. Qin, R. Wang, A. J. Vakharia, Y. Chen, and M. M. H. Seref, “The newsvendor problem: Review and directions for future research,” *European Journal of Operational Research*, vol. 213, no. 2, pp. 361–374, 2011.
- [44] R. Bellman, “Review,” *Management Science*, vol. 5, no. 1, pp. 139–141, 1958.
- [45] K. J. Arrow, H Scarf, and S Karlin, *Studies in the mathematical theory of inventory and production*. Stanford: Stanford Press, 1958.
- [46] K. J. Arrow, T. Harris, and J. Marschak, “Optimal inventory policy,” *Econometrica*, vol. 19, no. 3, pp. 252–272, 1951.
- [47] S. Karlin, “One-stage inventory models with uncertainty,” in *Studies in the mathematical theory of inventory and production*, K. J. Arrow, H Scarf, and S Karlin, Eds., Stanford: Stanford Press, 1958.
- [48] —, “Optimal inventory policy for the Arrow-Harris-Marschak dynamic model,” in *Studies in the mathematical theory of inventory and production*, K. J. Arrow, H Scarf, and S Karlin, Eds., Stanford: Stanford Press, 1958.
- [49] —, “Dynamic inventory policy with varying stochastic demands,” *Management Science*, vol. 6, no. 3, pp. 231–258, 1959.



- [50] ———, “Optimal policy for dynamic inventory process with stochastic demands subject to seasonal variations,” *Journal of the Society of Industrial and Applied Mathematics*, vol. 8, no. 4, pp. 611–629, 1959.
- [51] D Iglehart and S. Karlin, “Optimal policy for dynamic inventory process with non-stationary stochastic demands,” in *Studies in the Mathematical Theory of Inventory and Production*, K. J. Arrow, S. Karlin, and H Scarf, Eds., Stanford, CA: Stanford University Press, 1962, ch. 8.
- [52] A. F. Veinott Jr., “Optimal policy for a multi-product, dynamic, nonstationary inventory problem,” *Management Science*, vol. 12, no. 3, pp. 206–222, 1965.
- [53] A. F. Veinott Jr., “Optimal policy in a dynamic, single product, nonstationary inventory model with several demand classes,” *Operations Research*, vol. 13, no. 5, pp. 761–778, 1965.
- [54] P. Zipkin, “Critical number policies for inventory models with periodic data,” *Management Science*, vol. 35, no. 1, pp. 71–80, 1989.
- [55] T. E. Morton, “The nonstationary infinite horizon inventory problem,” *Management Science*, vol. 24, no. 14, pp. 1474–1482, 1978.
- [56] W. S. Lovejoy, “Stopped myopic policies in some inventory models with generalized demand processes,” *Management Science*, vol. 38, no. 5, pp. 688–707, 1992.
- [57] J.-S. Song and P. Zipkin, “Inventory control in a fluctuating demand environment,” *Operations Research*, vol. 41, no. 2, pp. 282–304, 1993.
- [58] S. P. Sethi and F. Cheng, “Optimality of  $(s, S)$  policies in inventory models with Markovian demand,” *Operations Research*, vol. 45, no. 6, pp. 931–939, 1997.
- [59] L. Xin and D. A. Goldberg, “Distributionally robust inventory control when demand is a martingale,” *ArXiv preprint arXiv:1511.09437*, 2015.
- [60] A Dvoretzky, J Keifer, and J Wolfowitz, “The inventory problem: II. Case of unknown distributions of demand,” *Econometrica*, vol. 20, no. 3, pp. 450–466, 1952.
- [61] H. Scarf, “Bayes solutions of the statistical inventory problem,” *Annals of Mathematical Statistics*, vol. 30, no. 2, pp. 490–508, 1959.
- [62] G. R. Murray Jr. and E. A. Silver, “A Bayesian analysis of the style goods inventory problem,” *Management Science*, vol. 12, no. 11, pp. 785–797, 1966.

- [63] K. S. Azoury and B. L. Miller, “A comparison of the optimal ordering levels of Bayesian and non-Bayesian inventory models,” *Management Science*, vol. 30, no. 8, pp. 993–1003, 1984.
- [64] K. S. Azoury, “Bayes solution to dynamic inventory models under unknown demand distribution,” *Management Science*, vol. 31, no. 9, pp. 1150–1160, 1985.
- [65] W. S. Lovejoy, “Myopic policies for some inventory models with uncertain demand distributions,” *Management Science*, vol. 36, no. 6, pp. 724–738, 1990.
- [66] M Lariviere and E Porteus, “Stalking information: Bayesian inventory management with unobserved lost sales,” *Management Science*, vol. 45, no. 3, pp. 346–363, 1999.
- [67] R. Kamath and T. P. M. Pakkala, “A Bayesian approach to a dynamic inventory model under an unknown demand distribution,” *Computers and Operations Research*, vol. 29, no. 2002, pp. 403–422, 2002.
- [68] A. Bensoussan, M. Cakanyildirim, and S. P. Sethi, “Partially observed inventory systems: The case of zero-balance walk,” *SIAM Journal on Control and Optimization*, vol. 46, no. 1, pp. 176–209, 2007.
- [69] O. L. Ortiz, A. L. Erera, and C. C. White III, “State observation accuracy and finite-memory policy performance,” *Operations Research Letters*, vol. 41, no. 5, pp. 477–481, 2013.
- [70] A. Bensoussan, M. Cakanyildirim, and S. P. Sethi, “A multiperiod newsvendor problem with partially observed demand,” *Mathematics of Operations Research*, vol. 32, no. 2, pp. 322–344, 2007.
- [71] A. Bensoussan, M. Cakanyildirim, J. A. Minjarez-Sosa, A. Royal, and S. P. Sethi, “Inventory problems with partially observed demands and lost sales,” *Journal of Optimization Theory and Applications*, vol. 136, no. 3, pp. 321–340, 2008.
- [72] X. Ding, M. L. Puterman, and A. Bisi, “The censored newsvendor and the optimal acquisition of information,” *Operations Research*, vol. 50, no. 3, pp. 517–527, 2002.
- [73] J. T. Treharne and C. R. Sox, “Adaptive inventory control for nonstationary demand and partial information,” *Management Science*, vol. 48, no. 5, pp. 607–624, 2002.
- [74] K. Arifoglu and S. Ozekici, “Optimal policies for inventory systems with finite capacity and partially observed Markov-modulated demand and supply processes,” *European Journal of Operations Research*, vol. 204, no. 3, pp. 421–438, 2010.

- [75] G Gallego and H Hu, “Optimal policies for production/inventory systems with finite capacity and Markov-modulated demand and supply processes,” *Annals of Operations Research*, vol. 126, no. 1, pp. 21–41, 2004.
- [76] E. Bayraktar and M. Ludkovski, “Inventory management with partially observed nonstationary demand,” *Annals of Operations Research*, vol. 176, no. 1, pp. 7–39, 2010.
- [77] H. Scarf, “The optimality of  $(S, s)$  policies in the dynamic inventory problem,” in *Mathematical Methods in the Social Sciences*, K Arrow, S Karlin, and P Suppes, Eds., 1960, ch. 13.
- [78] D. Iglehart, “Optimality of  $(s, S)$  policies in the infinite horizon dynamic inventory problem,” *Management Science*, vol. 9, no. 2, pp. 259–267, 1963.
- [79] A. F. Veinott Jr. and H. M. Wagner, “Computing optimal  $(s, S)$  inventory policies,” *Management Science*, vol. 11, no. 5, pp. 525–552, 1965.
- [80] A. F. Veinott Jr., “On the optimality of  $(s, S)$  inventory policies: New conditions and a new proof,” *SIAM Journal on Applied Mathematics*, vol. 14, no. 5, pp. 1067–1083, 1966.
- [81] L. G. Chen, L. W. Robinson, R. O. Roundy, and R. Q. Zhang, “Technical note - New sufficient conditions for  $(s, S)$  policies to be optimal in systems with multiple uncertainties,” *Operations Research*, vol. 63, no. 1, pp. 186–197, 2015.
- [82] J. H. Bookbinder and A. E. Lordahl, “Estimation of inventory re-order levels using the bootstrap statistical procedure,” *IIE Transactions*, vol. 21, no. 4, pp. 302–312, 1989.
- [83] G Gallego and I Moon, “The distribution free newsboy problem: Review and extensions,” *The Journal of the Operational Research Society*, vol. 44, no. 8, pp. 825–834, 1993.
- [84] H. Mamani, S. Nassiri, and M. R. Wagner, “Closed-form solutions for robust inventory management,” *Management Science*, 2016.
- [85] G. Perakis and G. Roels, “Regret in the newsvendor model with partial information,” *Operations Research*, vol. 56, no. 1, pp. 188–203, 2008.
- [86] O. Besbes and A. Muharremoglu, “On implications of demand censoring in the newsvendor problem,” *Management Science*, vol. 59, no. 6, pp. 1407–1424, 2013.

- [87] D. Klabjan, D. Simch-Levi, and M. Song, “Robust stochastic lot-sizing by means of histograms,” *Production and Operations Management*, vol. 22, no. 3, pp. 691–710, 2013.
- [88] W. T. Huh, R. Levi, P. Rusmevichientong, and J. B. Orlin, “Adaptive data-driven inventory control with censored demand based on Kaplan-Meier estimator,” *Operations Research*, vol. 59, no. 4, pp. 929–941, 2011.
- [89] B. GahYi, “The data-driven (s, S) policy: Why you can have confidence in censored demand data,” *Available at SSRN*, 2015.
- [90] D. Bertsimas and N. Kallus, “From predictive to prescriptive analytics,” *ArXiv preprint arXiv:1402.5481*, 2014.
- [91] B. GahYi and C. Rudin, “The big data newsvendor: Practical insights from machine learning,” *SSRN*, 2014.
- [92] R. Levi, G. Perakis, and J. Uichanco, “The data-driven newsvendor problem: New bounds and insights,” *Operations Research*, vol. 63, no. 6, pp. 1294–1306, 2015.
- [93] W. C. Cheung and D. Simchi-Levi, “Sampling-based approximation schemes for capacitated stochastic inventory control models,” *Available at SSRN*, 2015.
- [94] G. A. Godfrey and W. B. Powell, “An adaptive, distribution-free algorithm for the newsvendor problem with censored demands, with applications to inventory and distribution,” *Management Science*, vol. 47, no. 8, pp. 1101–1112, 2001.
- [95] K. J. Ferreira, B. H. A. Lee, and D. Simchi-Levi, “Analytics for an online retailer: Demand forecasting and price optimization,” *Manufacturing & Service Operations Management*, vol. 18, no. 1, pp. 69–88, 2016.
- [96] A. Atrash and J. Pineau, “A Bayesian method for learning POMDP observation parameters for robot interaction management systems,” in *In The POMDP Practitioners Workshop*, 2010.
- [97] R. D. Smallwood and E. J. Sondik, “The optimal control of partially observable Markov processes over a finite horizon,” *Operations Research*, vol. 21, no. 5, pp. 1071–1088, 1973.
- [98] E. J. Sondik, “The optimal control of partially observable Markov processes over the infinite horizon: Discounted costs,” *Operations Research*, vol. 26, no. 2, pp. 282–304, 1978.
- [99] M. L. Puterman, *Markov decision processes: Discrete stochastic dynamic programming*. 1994, p. 672, ISBN: 0471619779.

- [100] Y. Chang, A. L. Erera, and C. C. White III, "Value of information for a leader-follower partially observed Markov game," *Annals of Operations Research*, vol. 235, no. 1, pp. 129–153, 2015.
- [101] ———, "A leader-follower partially observed, multiobjective Markov game," *Annals of Operations Research*, vol. 235, no. 1, pp. 103–128, 2015.
- [102] C. C. White III and D. P. Harrington, "Application of Jensen's inequality to adaptive suboptimal design," *Journal of Optimization Theory and Applications*, vol. 32, no. 1, pp. 89–99, 1980.
- [103] Bayer Technology Services GMBH, "Flexible, fast and future production processes," 2014.
- [104] S. S. Malladi, A. L. Erera, and C. C. White III, "A dynamic mobile production capacity and inventory control problem," 2018.
- [105] U. S. Karmarkar, "Convex/Stochastic programming and multilocation inventory problems," *Naval Research Logistics*, vol. 26, no. 1, pp. 1–19, 1979.
- [106] ———, "The multiperiod multilocation inventory problem," *Operations Research*, vol. 29, no. 2, pp. 215–228, 1981.
- [107] ———, "The multilocation multiperiod inventory problem: Bounds and approximations," *Management Science*, vol. 33, no. 1, pp. 86–94, 1987.
- [108] N. Rudi, S. Kapur, and D. F. Pyke, "A two-location inventory model with transshipment and local decision making," *Management Science*, vol. 47, no. 12, pp. 1668–1680, 2001.
- [109] S. Axsäter, J. Marklund, and E. A. Silver, "Heuristic methods for centralized control of one-warehouse , N -retailer inventory systems," *Manufacturing & Service Operations Management*, vol. 4, no. 1, pp. 75–97, 2002.
- [110] Y. T. Herer, M. Tzur, and E. Yücesan, "The multilocation transshipment problem," *IIE Transactions (Institute of Industrial Engineers)*, vol. 38, no. 3, pp. 185–200, 2006.
- [111] R. W. Lien, S. M. R. Iravani, K. Smilowitz, and M. Tzur, "An efficient and robust design for transshipment networks," *Production and Operations Management*, vol. 20, no. 5, pp. 699–713, 2011.
- [112] W. C. Jordan and S. C. Graves, "Principles on the benefits of manufacturing process flexibility," *Management Science*, vol. 41, no. 4, pp. 577–594, 1995.

- [113] Y. T. Herer, M. Tzur, and E. Yücesan, “Transshipments: An emerging inventory recourse to achieve supply chain leagility,” *International Journal of Production Economics*, vol. 80, no. 3, pp. 201–212, 2002.
- [114] K. E. Wee and M. Dada, “Optimal policies for transshipping inventory in a retail network,” *Management Science*, vol. 51, no. 10, pp. 1519–1533, 2005.
- [115] S. S. Malladi, A. L. Erera, and C. C. White III, “Inventory control with modulated demand and a partially observed Modulation process,” Mar. 2018.
- [116] W. B. Powell, “Perspectives of approximate dynamic programming,” *Annals of Operations Research*, pp. 1–38, 2012.
- [117] P. Kouvelis and G. J. Gutierrez, “The newsvendor problem in a global market: Optimal centralized and decentralized control policies for a two-market stochastic inventory system,” *Management Science*, vol. 43, no. 5, pp. 571–585, 1997.
- [118] F. Bernstein and A. Federgruen, “Decentralized supply chains with competing retailers under demand uncertainty,” *Management Science*, vol. 51, no. 1, pp. 18–29, 2005.
- [119] F. Bernstein and G. A. DeCroix, “Inventory policies in a decentralized assembly system,” *Operations Research*, vol. 54, no. 2, pp. 324–336, 2006.
- [120] C. C. White III and K. Schlüssel, “Suboptimal design for large scale, multimodule systems,” *Operations Research*, vol. 29, no. 5, 1981.
- [121] W. S. Lovejoy, “Computationally feasible bounds for partially observed Markov decision processes,” vol. 39, no. 1, pp. 162–175, 1991.
- [122] G. L. Nemhauser and L. A. Wolsey, *Integer and combinatorial optimization*. New York, NY, USA: Wiley-Interscience, 1988, ISBN: 0-471-82819-X.
- [123] A Federgruen and P Zipkin, “An inventory model with limited production capacity and uncertain demands II: The discounted-cost criterion,” *Mathematics of Operations Research*, vol. 11, no. 2, pp. 208–215, 1986.
- [124] C. C. White III and W. T. Scherer, “Finite-memory suboptimal design for partially observed Markov decision processes,” *Operations Research*, vol. 42, no. 3, pp. 439–455, 1994.
- [125] C. C. White III and H. K. El-Deib, “Parameter imprecision in finite state, finite action dynamic programs,” *Operations Research*, vol. 34, no. 1, pp. 120–129, 1986.