Efficient Methods for Large-Scale Dynamic Optimization with Applications to Inventory Management Problems

Xujia Liu

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2023

Abstract

Efficient Methods for Large-Scale Dynamic Optimization with Applications to Inventory Management Problems

Xujia Liu

In this thesis, we study large-scale dynamic optimization problems in the context of inventory management. We analyze inventory problems with constraints coupling the items or facility locations in the inventory systems, and we propose efficient solutions that are asymptotically optimal or empirically near-optimal.

In Chapter 1, we analyze multi-item, single-location inventory systems with storage capacity limits which are formulated as both *unconditional* expected value constraints and *unconditional* probability constraints. We first show that problems with *unconditional* expected value constraints *only* can be solved to optimality through Lagrangian relaxation. Then, under an assumption on the correlation structure of the demands that is valid under most practical setting, we show that the original problem can be sandwiched between two other problems with expected value constraints only. One of these problems yields a feasible solution to the original problem that is asymptotically optimal as the number of items grows.

In Chapter 2, we consider the same problem but with *conditional* probability constraints, that impose limits on overflow frequency for *every* possible state in each period. We construct an efficient feasible solution in two steps. First, we solve an *unconditional* expected value constrained problem with reduced capacity. Second, in each period, given the state information, we solve a single-period convex optimization problem with a *conditional* expected value constraint. We further show that the heuristic is asymptotically optimal as number of items I grows. In addition, we design another efficient method for moderate values of

I, which works empirically well in an extensive numerical study. Moreover, we extract key managerial insights from the numerical study which are critical to decision making in real business problems.

In Chapter 3, we analyze single-item, multi-location systems on inventory networks that can be described by directed acyclic graphs (DAG). We propose an innovative reformulation of the problem so that Lagrangian relaxation can still be applied, which, instead of decomposing the problem by facility location, aggregates the state information, leading to a tractable lower bound approximation for the problem. The Lagrange multiplier, which provides information on the value function from the lower bound dynamic program, is used in designing a feasible heuristic. An extensive numerical study is conducted which suggests that both the lower bound approximation and upper bound heuristic perform very well.

Table of Contents

Acknowledgements

Five years of doctoral study witnesses my biggest growth in my academic career. I am deeply grateful for all the guidance and help received in this wonderful journey.

First, I would like to express my sincerest gratitude to my advisors, Professor Garud Iyengar and Professor Awi Federgruen, who have been providing enlightening guidance throughout my PhD study, without which I cannot reach at this point. In addition to the valuable advice on research, they have also offered the largest possible accommodation which allows me to complete the doctoral study while staying with my family during the hardest time of the pandemic.

I would also like to thank my co-author, Professor Daniel Guetta. He has been providing patient guidance not only on my research, but also on personal life. He himself is a great example to learn from, as a researcher, and, as a person.

I would like to extend my thanks to the rest of my research committee, Professor David Yao and Professor Martin Haugh.

I want to thank all the professors and lecturers in the department of IEOR, who have been providing useful and inspiring courses, and all the staff in the department, who have provided assistance to make the doctoral life easy and interesting.

Lastly and most importantly, I would like to thank my family, who have been providing enduring support throughout my five years of PhD study. Thank you, my parents, my wife, my twin brother, and my little daughter who was born in this period.

To my family.

Introduction

Most interesting multi-period decision problems can naturally be formulated as stochastic dynamic programming problems (DP). However, the state space of the DPs grows exponentially with the number of items describing the state of the system and the number of time periods in the planning horizon, both typically large numbers. Very often, the state dynamics decompose by item, but the constraints create complex state-dependent interdependencies, preventing the problems from decomposition.

Inventory management is an area where this kind of problems arise. For example, retail organizations such as Amazon, Walmart and Alibaba, manage complex inventory systems where a large number of different products with potentially correlated demand share common storage space or budgets with capacity limits. Each product has a stochastic demand, and associated costs including but not limited to holding, backlogging and ordering costs. The managers of the systems aim to minimize total costs over a planning horizon while satisfying customer demands and meeting complex operational requirements (constraints). The dimension of the state in this problem grows with the number of facilities in the inventory network, or number of products in the system, and thus can easily scale up to thousands or even millions for large organizations like Amazon. Linking constraints, e.g., storage or budget capacity limits shared by all products, or shipping flows connecting different facilities, prevent the dynamic program from decoupling by product or facility location.

Lagrangian relaxation is a natural decomposition approach to dynamic programming problems with weakly coupling constraints. By dualizing the coupling constraints, it removes the interdependencies between components of the state and decomposes a high dimensional DP into multiple low dimensional DPs.

However, the number of Lagrangian multipliers required to obtain strong duality grows with the dimension of the state, and is thus impractical per se. The problem becomes even more challenging when the linking constraints are non-convex, for example, capacity constraints of probability type limiting the overflow frequency of inventory level.

In this thesis, we propose efficient solutions to the aforementioned problems. Some of the solutions are asymptotically optimal as the dimension of state grows, and others are shown empirically near-optimal.

In Chapter 1, we analyze multi-item, single-location inventory systems with storage capacity limits. Both expected value type constraints limiting the expected overflow amount, and chance constraints limiting the overflow frequency, are imposed. The expectations and probabilities are evaluated given the state information at the beginning of the planning horizon, and thus, we refer to this type of constraints as *unconditional* constraints. In this chapter, we first show that problems with *unconditional* expected value constraints *only* can be solved to optimality through Lagrangian relaxation. The optimal policy to controls the inventory for each item independently using a so-called double base stock policy. However, the real complexity of the problem comes from the chance constraints which define a highly non-convex feasible region and link all products sharing the capacity limits, preventing the problem from decoupling by Lagrangian relaxation. We show that, when the correlation of the demands has a particular structure that is often seen in practice, the problem can be sandwiched between two other problems with expected value constraints only. The optimal policy for one of these problems yields a feasible solution to the original problem that is asymptotically optimal as the number of items grows.

In Chapter 2, we consider a problem similar to the one in Chapter 1, but with *conditional* type constraints. These constraints evaluate expectations or probabilities conditioning on the state information at the beginning of each period, and thus rendered the name *conditional*. This problem is considered much more complicated because the constraints are imposed for *every* possible state in each period, which are infinitely many for continuous state space. Taking the Lagrangian dual of any such dynamic problem with infinitely many constraints results in a dynamic program with infinitely many Lagrange multipliers. Solution of any such Lagrangian dual is, therefore, entirely prohibitive. Therefore, a fundamentally different approach is employed. In this chapter, we construct a feasible heuristic in two steps. First, we solve an *unconditional* expected value constrained problem with reduced capacity, which as has been shown in Chapter 1, can be efficiently solved to optimality. Second, in each period, given the state information, we solve a single-period convex optimization problem with a *conditional* expected value constraint, which gives a feasible solution to the original chance constrained problem. We further show, in an involved empirical process analysis, that the heuristic is asymptotically optimal as number of items I grows. In addition, we propose another efficient method designed for moderate values of I, which we show works empirically well in an extensive numerical study. Moreover, we extract key managerial insights from the numerical study which are critical to decision making in real business problems.

In Chapter 3, we extend the complexity of the problem in another dimension, the number of facility locations in the systems. We analyze single-item, multi-location systems on inventory networks that can be described by directed acyclic graphs (DAG). The problem is no longer a weakly coupling constrained problem since the shipping flows in the network connect different facilities in a complex way, In this chapter, we propose an innovative reformulation of the problem so that Lagrangian relaxation can still be applied, which, instead of decomposing the problem by facility location, aggregates the state information, leading to a tractable lower bound approximation for the problem. The Lagrange multiplier, which provides information on the value function from the lower bound dynamic program, is used in designing a feasible heuristic. We conduct an extensive numerical study consisting of 40,960 instances to assess the accuracy of the lower bound approximation and suboptimality of the upper bond heuristic. The average gap is 1.86% only with gaps below 10% in 99.88% of the instances. This suggests that both the lower bound approximation and upper bound heuristic perform very well.

Chapter 1: An Asymptotically Optimal Heuristic for Multi-Item Inventory Models with Unconditional Expected Value and Chance Inventory Constraints

1.1 Introduction and Summary

One of the most fundamental complications retailers face in managing their inventories is constraints on the total amount of inventory that can be stored across different items or stock keeping units. These constraints may arise from limited physical space, or from aggregate budget limits (resulting from cash flow or other financial considerations). For example, the inventory-to-sales ratio continues to be one of the important financial metrics monitored by financial analysts, and is therefore tightly controlled by CFOs. Because these constraints bound the *total* amount of inventory across *many* products, they introduce interdependencies among these products. A fundamental question is how these constraints impact both the structure and parameters of optimal or near-optimal procurement strategies.

This is one of the most basic stochastic multi-item inventory models, and was addressed as early as the sixties when inventory theory was at its infancy, in particular in seminal papers by Veinott Jr (1965) and Ignall and Veinott Jr (1969). However, these authors focus on settings where lead times are zero, i.e., orders are assumed to arrive instantaneously. (Under this assumption, the inventory constraints are simple deterministic and linear constraints of the inventories on hand.) To our knowledge, this restriction has been adopted by all published literature, recent contributions included.

Positive lead times, however, are ubiquitous and one of the main drivers of inventories to begin with. They render the problem especially challenging, since the aggregate inventory level at the beginning of a given period depends not only on the inventory positions and orders placed in the past, but also on intermediate stochastic demands that are realized over the lead time after the order. Depending on the specific criterion chosen, this relationship is, in addition, highly non-linear. Indeed, even the specification of the inventory constraints is far less than obvious.

The joint storage constraints could be specified in a way that precludes overflows under all possible demand scenarios, including the scenario where all demands during the procurement lead time are at their minimal possible levels, i.e., the smallest values of the support of the joint distribution, to the extent such smallest values exist. (For distributions without an essential infimum, such as the frequently used Normal distribution, this is, of course, not the case.) However this approach, even when feasible, is extremely conservative and leaves much of the capacity under-utilized.

A much more practical and flexible approach would control (a) the expected aggregate inventory in every period (expected value constraints), but also (b) the probability at which an overflow occurs in every period (chance constraints).

Living with a low odds overflow event is certainly acceptable under soft constraints imposed by budgetary or other financial considerations. It is also sensible in the case of physical inventory constraints, where occasional emergency measures can be taken to temporarily store extra inventory. Focusing on expected value constraints, as in (a), exclusively allows for frequent scenarios in which overflow occurs. Conversely, an exclusive control of the overflow probability, as under (b), fails to control the magnitude of the overflow.

Expected value constraints are, considerably easier, to deal with. Being separable in the decision variables and convex, they lend themselves easily to be solved via their Lagrangian dual. Under any vector of Lagrangian multipliers, the problem decomposes into I separate single-item, single-dimensional dynamic programs (for which, as an additional benefit, a simply structured policy is optimal), where I is the number of items. Moreover, there is strong duality, so that the policy optimizing the Lagrangian dual is optimal for the original problem as well. Finally, the Lagrangian dual, an unconstrained convex program in T variables, where T is the number of periods, can be optimized, to any desired precision, by efficient steepest ascent methods, like FISTA, see Beck and Teboulle (2009).

Chance constraints, controlling e.g. overflow probabilities or service levels, important as they are, are considerably harder to deal with, even in a single-item setting, see e.g. Jiang et al. (2019) and Wei et al.

(2021) that formulate service level constraints as chance constraints. See Section 1.2 for a review. It is probably for this reason that very few papers on inventory theory, even in single- or two-stage problem, consider such chance constraints.

Lagrangian duals of problems with such chance constraints, for example, exhibit duality gaps because they are not be convex in the decision variables. In multi-item settings, like ours, there is the added complication that Lagrangian relaxation fails to decompose the problem into tractable, e.g. single-item, dynamic programs. A tractable and computable exact optimal policy, therefore appears unachievable. Nevertheless, we are able to design a heuristic policy, of simple structure, which is asymptotically optimal as the number of items I grows to infinity. In fact the absolute optimality gap is $O(\sqrt{I})$, so that the relative optimality gap is $O(\frac{1}{\sqrt{I}})$. Moreover, this heuristic policy controls each item based on an item-specific control rule which determines orders or salvage batches based on the item's inventory position only. Moreover, the control rules have a very simple structure. This is similar to the policy which is *exactly* optimal, when only expected value constraints prevail.

An important proviso for establishing asymptotic optimality of a heuristic is that the constraints be stated as *unconditional* constraints given any specific starting state at the beginning of the first period. The problem becomes considerably harder when the constraints for a given period need to apply for every starting vector of inventory positions one lead time earlier. In Chapter 2, we will address this conditional version of the problem. We will establish a lower bound using an entirely different approach, as well as two feasible heuristic policies. One policy is asymptotically optimal under the *conditional* storage constraints. The second policy is shown, numerically, to exhibit extremely low optimality gaps even for moderate values of I over a large range of problem parameters; however, it cannot be shown to be asymptotically optimal.

Our contributions

We analyze a periodic review model with T periods. At the beginning of each period, the inventory position of each item can be adjusted by placing an order or by salvaging some of the inventory. Orders arrive and salvage batches deplete the inventory after a given lead time. Demands are given by general continuous distributions, with bounded support. In Section 1.7, we generalize our results to general multivariate Normal distributions, and similar families of distributions with unbounded support. In addition to variable order and salvaging costs, there are linear or convex holding and backlogging costs. In every period, we bound the expected value of the aggregate inventory level. In addition, in every period, the probability of the aggregate inventory level exceeding the prevailing inventory capacity must be smaller than a given tolerance.

When there are *only* expected value constraints, an exact optimal policy exists which is of a very simple structure and efficiently computable. Lagrangian relaxation is performed on the expected value constraints. Under any vector of Lagrange multipliers, the problem decomposes into I separate single-item dynamic programs. In each of these DPs, a so-called double base stock policy is optimal: in each period, when the starting inventory position is below (above) a first (second) base stock level, the inventory position is increased (decreased) to this level. There is strong duality, which means that the optimal value of the Lagrangian dual equals the optimal cost value, and the policy associated with the optimal solution of the Lagrangian dual, is optimal for the original problem as well. Finally, an ϵ-optimal solution of the Lagrangian dual can be computed with complexity $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$. Thus, remarkably, each item can be managed by itself, with the required coordination achieved by guiding their base stock levels through the common optimal Lagrange vector.

When chance constraints, controlling the overflow probabilities, are added, an exact optimal policy can no longer be found, and it is likely to be of a prohibitively complex structure, requiring each item to be governed on the basis of information about the entire I-dimensional system state. Instead, we derive a heuristic policy which is asymptotically optimal when the number of items I grows to infinity. We believe this to be the relevant asymptotic result because in many applications, the number of items I can go into the ten thousands to millions, if not more. The heuristic is not asymptotically optimal in T; however, the asymptotic behavior in T is unimportant in this context where the length of a reasonable planning horizon is naturally bounded. We also develop a lower bound that is asymptotically accurate as $I \to \infty$. (The absolute optimality gap is, in fact, $O(\sqrt{I})$ which implies that the relative gap is $O(\frac{1}{\sqrt{I}})$.) The heuristic is of the same simple structure as that in the expected value constrained problem, and can be efficiently computed with the same complexity of $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$. This asymptotic optimality result requires that the items can be partitioned into product lines, where each product line contains a bounded number of items, such that items within the same product line may be arbitrarily correlated, but are independent across product lines.

We achieve these results by showing that the problem, with chance constraints on the overflow probability in each period, can be sandwiched in between two other problems with expected value constraints only and slightly expanded (reduced) capacity levels.

The layout of this paper is as follows. Section 1.2 provides a literature review. The model and notation are introduced in Section 1.3. Section 1.4 shows how the problem with expected value constraints (only) can be solved effectively and to optimality by solving its Lagrangian dual. A policy of simple structure is optimal for this model. Section 1.5 shows how the problem with chance constraints added can be bounded from below and above by one in which all constraints are expected value constraints. Section 1.6 derives a heuristic policy for the original model, which is shown to be asymptotically optimal in the number of items under the above correlation structure assumption.

In practice as well as in inventory theory, demands are most commonly assumed to have a multivariate Normal distribution. (As one of many advantages, this family of distributions allows for any specification of the correlation structure via the variance-covariance matrix.) Therefore, in Section 1.7 we extend our results to general Normal demands as well as to a truncated Normal demand model in which demands are non-negative. Section 1.9 concludes this paper with generalizations to the infinite horizon problem with the long-run average cost criterion, and lost sales settings.

1.2 Literature Review

The stochastic, periodic review multi-item model with joint inventory constraints has a long history, starting with Veinott Jr (1965) and Ignall and Veinott Jr (1969). These papers consider a model with stationary demand distributions and parameters, and focus on conditions under which a myopic multi-item base stock policy is optimal. Most importantly, the authors assume zero lead times, in which case the joint inventory constraints translate into simple linear constraints bounding the sum of the inventory levels in each period. Ignall and Veinott Jr (1969) briefly discuss the case of positive lead times and observe that if the demand distributions are assumed to be bounded from below and one insists that the inventory constraints hold for every possible demand scenario, the constraints once again reduce to simple bounds on the sum of the inventory positions. As argued above, however, this approach is unrealistically conservative and leaves a good part of the inventory capacity unused. It also precludes the use of distributions which are not bounded from below, as with Normal distributions.

Later contributions, e.g., DeCroix and Arreola-Risa (1998), Beyer et al. (2001, 2002), Shi et al. (2016) and Chen and Li (2019) have all continued with this paradigm of stationary models with zero lead times.

Inventory models with chance constraints are, in fact, rare. This is surprising since the Type I service level constraint (which constrains the probability of running out of stock in any given period), dating back to the earliest days of the field, is precisely of this type. Exceptions include a few papers surveyed in Prékopa (2003), Section 7.3. This section deals with inventory models under chance constraints, published prior to 2003; the cited references all translate into two-stage stochastic programs. More recent contributions, dealing with single-item systems under Type-I service level constraints include Boyaci and Gallego (2001), Shang and Song (2006), Özer and Xiong (2008), Bijvank and Vis (2012) and Bijvank (2014). These papers focus on developing/analyzing heuristic policies of certain forms.

Most recently, Wei et al. (2021) show that, for this single-item model, under backlogging, a simple base stock policy is asymptotically optimal, when the service level goes to 100%. The base stock levels are computed by solving a linear program corresponding to a deterministic approximation of the problem. These authors also analyze the lost-sales version for the model, and in fact appear to be the first ones to theoretically study policy performance, in this model with Type-I service level constraints. They design a heuristic, which is, once again, shown to be asymptotically optimal when the service level goes to 100%. (Even for the unconstrained single-item lost-sales model, the structure of the optimal policy is unknown under general parameters.) See Section 1.9, for a treatment of this model.

Chance constraints in the form of Type I service level constraints were also addressed in the context of assemble-to-order systems. This approach was started by Baker (1985) with a more formal analysis for W-systems in Baker et al. (1986). (In a W-system, there are 2 final products, each with a product-specific component and one common component; thus there are 3 components in total.) Capacity allocation is only relevant for the common component. A specific simple allocation policy is assumed in Baker et al. (1986) while an optimal allocation policy is derived in Gerchak et al. (1988), when the service level targets for the two end products are the same, and Mirchandani and Mishra (2002) when they are general. These Type I constraints also appear in a few risk pooling models but all in single stage models, see Swaminathan and Srinivasan (1999), Alptekinoğlu et al. (2013) and most recently in Jiang et al. (2019). For a general treatment of mathematical programs under chance constraints, see Nemirovski and Shapiro (2006, 2007).

Our paper also contributes to the emerging literature on using Lagrangian relaxation techniques to approximately solve the dynamic programs that arise in the context of multi-item inventory and revenue management problems. The general technique was introduced by Adelman and Mersereau (2008) for a general class of dynamic programs, and subsequently applied in several contexts, including inventory theory. The authors also develop an LP based bound replacing the multi-dimensional value functions with separable approximations. (See Brown and Zhang (2021) for a recent theoretic comparison of the two types of approximations.) Examples of Lagrangian relaxation in dynamic programs are Kunnumkal and Topaloglu (2008, 2011) and Federgruen et al. (2018) dealing with versions of the single-item, one-warehouse multi-retailer problem.

However, none of these prior papers attempt to design an asymptotically optimal heuristic for the (very distinct) problems they addressed. Only a handful of papers have been able to show how this approach leads to asymptotically optimal heuristics. Marklund and Rosling (2012) consider a periodic review, one warehouse multiple retailers system in which the warehouse is replenished in cycles. For one given cycle with assumed known initial inventory at the warehouse – which cannot be replenished during the cycle – the problem is how to best procure the retailers over the course of the cycle to minimize the retailers' expected holding and backlogging plus shipment costs. A single unconditional expected value constraint ensures that the shipments from the warehouse in expectation do not exceed the given initial inventory. This problem was initiated by Jackson (1988).

The above result was extended in Nambiar et al. (2021) and Miao et al. (2022) to a system with multiple warehouses with a common cycle and known initial inventories at each warehouse. Marklund and Rosling (2012), Nambiar et al. (2021) and Miao et al. (2022) have in common that shipment times are assumed to be zero, and under these assumptions the authors derive asymptotically optimal heuristics.

Hu and Frazier (2017) and Zayas-Caban et al. (2019) consider a restless finite horizon bandit problem, i.e., one in which in every period multiple arms can be pulled. They consider randomized policies but impose unconditional constraints on the number of arms being pulled in any given period, being equal to a given number. Both papers apply Lagrangian relaxation on the coupling constraints and identify a policy which is asymptotically optimal in the number of bandits. Brown and Smith (2020) similarly consider dynamic selection policies, as arise e.g., in dynamic assortment problems, subject to constraints that for a given starting state the number of selected items does not exceed given bounds. Once again, based on Lagrangian relaxation, a heuristic is found that is asymptotically optimal in the number of items. Finally, Balseiro et al. (2021) establish the same, under the long-run average cost criterion, in a pricing model for networks in which, in every period, a resource can be moved between a pair of nodes if a demand arises there. Asymptotic optimality is in terms of the number of nodes.

1.3 Model and Notation

We consider a periodic review inventory control system with a planning horizon of T periods storing I distinct items with random demands. At the start of each period, one may place an order for any of the items or salvage some of their inventory. Orders take effect after a given lead time of ℓ periods. All stockouts are fully backlogged. (See Section 1.9 for a treatment of the lost-sales case.) A common storage facility or aggregate inventory budget allows for a given inventory capacity only. Without loss of generality, we assume that a unit of every item occupies one unit of storage capacity (the budget). In addition to variable order and salvaging costs, there are holding and backlogging costs, specified by linear or convex functions of the beginning-of-period inventory levels. Demands for the I items are specified by a general multivariate, possibly time dependent distribution with bounded support. As in almost all inventory models, demands are assumed to be independent across time, but not necessarily across products.

Our objective is to minimize expected total (discounted) costs over the full planning horizon, subject to joint inventory constraints.

We use the following notation:

- $c^i_{(t)}$: the per unit ordering cost for product i in period t. $c^i_{(t)} \in [\underline{c}, \bar{c}]$ for all i, t.
- $d^i_{(t)}$: the per unit salvaging cost for product i in period t. $d^i_{(t)} \in [\underline{d}, \bar{d}]$ for all i, t.
- $h^i_{(t)}$: the per unit holding cost for product i at the beginning of period t. $h^i_{(t)} \in [\underline{h}, \bar{h}]$ for all i, t. To

avoid an unbounded arbitrage opportunity, we assume, without loss of generality that for all i and t and $\tau \geq t$,

$$c^i_{(t)} + \delta h^i_{(t+1)} + \ldots + \delta^{\tau-t} h^i_{(\tau)} + \delta^{\tau-t} d^i_{(\tau)} > 0, \tag{1.1}$$

i.e., that it is impossible to make a profit by buying a product, holding it for any number of periods, and then salvaging it.

- $p^i_{(t)}$: the per unit backlogging cost for product i at the beginning of period t. $p^i_{(t)} \in [\underline{p}, \bar{p}]$ for all i, t.
- $\chi_{(t)}$: the inventory capacity at the beginning of period t. The capacity levels $\{\chi_{(t)}\}$ are, of course, assumed to grow with I, but not necessarily proportionally with I, exhibiting economies of scope. In fact the following assumption is sufficient:

  Assumption 1. *The capacity values $\{\chi_{(t)}\}$ grow (at least) at the rate of $\sqrt{I}$, i.e., $\chi_{(t)} = \Omega(\sqrt{I})$, for all $t = 1, \ldots, T$.*

- δ: the discount factor ($0 \leq \delta \leq 1$).
- $u^i_{(t)}$: the demand for product i in period t. We assume that the marginal demand for any item i and period t follows a general bounded continuous distribution with $L \leq u^i_{(t)} \leq U$, and without loss of generality, we assume $L = 0$.

  As far as the correlation structure among the items is concerned, we assume the following general structure.

  Assumption 2. *The items $\{1, \ldots, I\}$ can be clustered into disjoint product lines $\{G_1, \ldots, G_K\}$ such that any two items in different product lines have independent demands with $\max_k |G_k| \leq m$ for some constant m that does not grow with I.*

- $\dot{u}^i_{(t)}$: the demand for product i, during a lead time of ℓ periods which starts in period t, i.e., the demand during the interval $[t, t+\ell)$.
- $x^i_{(t)}$: the inventory position for product i at the beginning of period t *before* placement of an order.
- $\bar{x}^i_{(t)}$: the inventory position for product i, at the beginning of period t, but *after* placement of an order. (As we will shortly discuss, this will be the "decision variable" in our problem).

For any parameters or variables, we use bold characters to denote their vectors over all I items, e.g., $\boldsymbol{x}_{(t)} = [x^1_{(t)}, \ldots, x^I_{(t)}]$ and $\bar{\boldsymbol{x}}_{(t)} = [\bar{x}^1_{(t)}, \ldots, \bar{x}^I_{(t)}]$.

By a well-known accounting scheme, the total inventory on hand at the beginning of period $t+\ell$ is given by

$$\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+,$$

where $x^+ = \max\{x, 0\}$. For reasons explained in the Introduction, we consider expected value constraints of the type:

$$\mathbb{E}\left[\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \middle| \boldsymbol{x}_{(0)}\right] \leq \chi_{(t+\ell)} + R, \qquad t = 0, \ldots, T-\ell. \tag{1.2}$$

for some $R \geq 0$. Often $R = 0$; however, $R > 0$ may be chosen, to allow for limited overflows stored in secondary facilities (financed from secondary sources). In the latter case, we assume $R = o(I)$. Furthermore, we impose additional chance constraints of the following type:

$$\mathbb{P}\left(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{t+\ell} \middle| \boldsymbol{x}_{(0)}\right) \leq \beta, \qquad t = 0, \ldots, T-\ell \tag{1.3}$$

for a given (small) probability $0 < \beta < 1$ to control the odds of an overflow. We emphasize that both the expectations and the probabilities for all periods $t = 0, \ldots, T$ are assessed with respect to the initial state $\boldsymbol{x}_{(0)}$. In other words, they are unconditional expectations and probabilities, averaged out over all policy-dependent sequences $\{\bar{\boldsymbol{x}}_{(1)}, \ldots, \bar{\boldsymbol{x}}_{(T)}\}$. We emphasize that those constraints are weaker than, for example, the conditional probability constraints

$$\mathbb{P}\left(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{t+\ell} \middle| \boldsymbol{x}_{(t)}\right) \leq \beta, \qquad t = 0, \ldots, T-\ell. \tag{1.4}$$

Under (1.3), we may have intermediate states $\boldsymbol{x}_{(t)}$ for some $t = 0, \ldots, T$, where the capacity constraint is violated with very low probability, and others where it is violated with high probability, as long as the fraction of sample paths that result in a violation is smaller than β. Under the *stronger* chance constraints the *conditional* violation probability in every period is bounded regardless of the starting state $\boldsymbol{x}_{(t)}$, one lead time earlier.

The problem can be formulated as a dynamic program in which $\boldsymbol{x}_{(\tau)}$ is the state of the system in period τ. Let $V_{(\tau)}(\cdot)$ be the value function for period τ. Then

$$(DP) \quad V_{(\tau)}(\boldsymbol{x}_{(\tau)}) = \min_{\bar{\boldsymbol{x}}_{(t)} \geq \min\{0, \boldsymbol{x}_{(t)}\}} \left\{ \sum_{t=\tau}^{T-\ell} \delta^{t-\tau} \sum_{i=1}^{I} \left[c^i_{(t)}(\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)}(x^i_{(t)} - \bar{x}^i_{(t)})^+ + Q^i_{(t)}(\bar{x}^i_{(t)}) \right] \right\} \tag{1.5}$$

$$\text{s.t. constraints (1.2) (1.3),}$$

$$x^i_{(t+1)} = \bar{x}^i_{(t)} - u^i_{(t)}, \qquad \text{for } i = 1, \ldots, I \text{ and } t = 0, \ldots, T - \ell$$

where $Q^i_{(t)}(\bar{x}^i_{(t)}) = \delta^\ell \mathbb{E}\Big[h^i_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ + p^i_{(t+\ell)}(\dot{u}^i_{(t)} - \bar{x}^i_{(t)})^+ \Big]$ represents, the expected and discounted holding and backlogging costs that will be incurred at the beginning of period $t + \ell$. By accounting for the expectation of these costs in period t, we eliminate the need to keep track of intermediate ordering quantities that are in the pipeline.

$V_1(\boldsymbol{x}_{(1)})$ may be viewed as the value of a T-stage stochastic program, or the value of a dynamic program. Either way, the exact problem is intractable: the dynamic program, for example, has an I-dimensional state space. It is loosely coupled by constraints (1.2) and (1.3). Relaxing the inventory constraints via Lagrangian relaxation would not succeed in decoupling the dynamic program into single-item dynamic programs. Moreover, because the *chance* constraints are (highly) non-convex, a Lagrangian dual would only result in a lower bound, i.e., only weak duality prevails.

In the next section, we show that for a dynamic program with only unconditional expected value constraints, as in (1.2), Lagrangian relaxation succeeds in decoupling the problem and *strong* duality also prevails, i.e., the problem can be solved exactly through its Lagrangian dual. In Section 1.5, we show that, the dynamic program with both expected value constraints and chance constraints can be sandwiched between two dynamic programs with expected value constraints only.

Finally, Section 1.6 shows that the optimal strategy in the problem with the expected value constrains and reduced capacity levels, is feasible in the original problem (DP), and it provides a bound for its optimality gap. This policy is, additionally, of simple structure, and therefore very easy to implement. For the case where the demands have a correlation structure in accordance with Assumption 4, this policy is then shown to be asymptotically optimal as I grows to infinity.

1.4 The Expected Value Constrained Problem

In this section, we show that, a problem with unconditional expected value constraints only can be solved efficiently by computing the Lagrangian dual of the problem. For any perturbed capacity levels $\{\chi'_{(t)}\}_{t=0}^{T}$ with $\chi'_{(t)} \geq 0$ for all t, we use $W^*(\{\chi'_{(t)}\}_{t=0}^{T})$ to denote the optimal cost of problem with the unconditional expected value constraints (1.2) only, i.e.,

$$
\begin{aligned}
W^*(\{\chi'_{(t)}\}) = \min_{\bar{\boldsymbol{x}}_{(t)} \geq \min\{0, \boldsymbol{x}_{(t)}\}} \mathbb{E}\Big[&\sum_{t=0}^{T-\ell} \delta^{t-1} \sum_{i=1}^{I} \{ c^i_{(t)}(\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)}(x^i_{(t)} - \bar{x}^i_{(t)})^+ \\
&+ \delta^{\ell} h^i_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ + \delta^{\ell} p^i_{(t+\ell)}(\dot{u}^i_{(t)} - x^i_{(t)})^+ \} \Big] \\
\text{s.t. } \mathbb{E}&\left[\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \Big| \boldsymbol{x}_{(0)} \right] \leq \chi'_{(t+\ell)}, \quad t = 0, \ldots, T-\ell, \\
&x^i_{(t+1)} = \bar{x}^i_{(t)} - u^i_{(t)}, \quad \text{for } i = 1, \ldots, I \text{ and } t = 0, \ldots, T-\ell.
\end{aligned}
\tag{1.6}
$$

To apply Lagrangian relaxation, we assign a Lagrangian multiplier $\lambda_{(t)} \geq 0$ to the t-th constraint in (1.6) and the constraint is replaced by a penalty term in the objective. Thus, for a given vector $\boldsymbol{\lambda} \geq 0$, we get the relaxed program

$$
\begin{aligned}
W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}) = \min_{\bar{\boldsymbol{x}}_{(t)} \geq \min\{0, \boldsymbol{x}_{(t)}\}} \mathbb{E}\Big[&\sum_{t=0}^{T-\ell} \delta^{t-1} \sum_{i=1}^{I} \{ c^i_{(t)}(\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)}(x^i_{(t)} - \bar{x}^i_{(t)})^+ \\
&+ \delta^{\ell} h^i_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ + \lambda_{(t)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ + \delta^{\ell} p^i_{(t+\ell)}(\dot{u}^i_{(t)} - x^i_{(t)})^+ \} \Big] \\
&- \sum_{t=0}^{T-\ell} \delta^{t-1} \lambda_{(t)} \chi'_{(t+\ell)} \\
\text{s.t. } & x^i_{(t+1)} = \bar{x}^i_{(t)} - u^i_{(t)}, \quad \text{for } i = 1, \ldots, I \text{ and } t = 0, \ldots, T-\ell
\end{aligned}
\tag{1.7}
$$

For any given $\boldsymbol{\lambda} \geq 0$, the problem $W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ decomposes naturally into I separate inventory problems, with perturbed holding cost functions $\bar{h}^{\lambda,i}_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \equiv \delta^{\ell} h^i_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ + \lambda_{(t)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$.

Alternatively, we can compute each single-item dynamic program via one-dimensional recursively defined value functions:

$$W^{\lambda,i}_{(t)}(x^i_{(t)}) = \min_{\bar{x}^i_{(t)} \geq \min\{x^i_{(t)},0\}} \Big\{ c^i_{(t)}(\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)}(x^i_{(t)} - \bar{x}^i_{(t)})^+ + Q^i_{(t)}(\bar{x}^i_{(t)}) \\ + \lambda_{(t)}\mathbb{E}[(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+] - \lambda_{(t)}\chi'_{(t+\ell)} + \delta\mathbb{E}[W^{\lambda,i}_{(t+1)}(\bar{x}^i_{(t)} - u^i_{(t)})] \Big\}. \tag{1.8}$$

The value functions $W^{\lambda,i}_{(t)}(\cdot)$ are easily shown to be convex and differentiable everywhere, and twice differentiable almost everywhere. If $\bar{x}^i_{(t)} > 0$, then it can be easily seen to be the unique root of the function:

$$H(\bar{x}^i_{(t)}) := \begin{cases} c^i_{(t)} + {Q'}^i_{(t)}(\bar{x}^i_{(t)}) + \lambda_{(t)}\dot{F}^i_{(t)}(\bar{x}^i_{(t)}) + \delta\mathbb{E}[{W'}^{\lambda,i}_{(t+1)}(\bar{x}^i_{(t)} - u^i_{(t)})], & \text{if } \bar{x}^i_{(t)} > x^i_{(t)}, \\ -\, d^i_{(t)} + {Q'}^i_{(t)}(\bar{x}^i_{(t)}) + \lambda_{(t)}\dot{F}^i_{(t)}(\bar{x}^i_{(t)}) + \delta\mathbb{E}[{W'}^{\lambda,i}_{(t+1)}(\bar{x}^i_{(t)} - u^i_{(t)})], & \text{if } \bar{x}^i_{(t)} \leq x^i_{(t)}. \end{cases}$$

(1.9a)

(1.9b)

The following theorem characterizes the optimal policy for the dynamic program (1.8).

Theorem 1. *(a) Fix $\boldsymbol{\lambda}$. The optimal policy for the dynamic program* (1.8) *is a double base stock policy: there exists two parameters $S^{i,\boldsymbol{\lambda}}_{(t)}$ and $B^{i,\boldsymbol{\lambda}}_{(t)}$ for every period t, with $-\infty \leq S^{i,\boldsymbol{\lambda}}_{(t)} < B^{i,\boldsymbol{\lambda}}_{(t)} \leq +\infty$, and with $B^{i,\boldsymbol{\lambda}}_{(t)} \geq 0$, such that if the starting inventory position is below $S^{i,\boldsymbol{\lambda}}_{(t)}$, an order should be placed that brings it up to $S^{i,\boldsymbol{\lambda}}_{(t)}$, and if the starting inventory position is above $B^{i,\boldsymbol{\lambda}}_{(t)}$, the salvage option should be exercised to bring it down to $B^{i,\boldsymbol{\lambda}}_{(t)}$. Otherwise, it is optimal to leave the starting inventory position unchanged.*

In addition, the value functions $W^{\lambda,i}_{(t)}(x^i_{(t)})$ are convex.

(b) Let $S^{i,\boldsymbol{\lambda}}_{(t)}$ be defined in part (a). $S^{i,\boldsymbol{\lambda}}_{(t)}$ is decreasing in $\boldsymbol{\lambda}$ with respect to the regular partial order of $\mathbb{R}^T$. Specifically, if $\boldsymbol{\lambda}^1 \leq \boldsymbol{\lambda}^2$, then

$$S^{i,\boldsymbol{\lambda}^1}_{(t)} \geq S^{i,\boldsymbol{\lambda}^2}_{(t)}.$$

In particular, $S^{i,\boldsymbol{\lambda}}_{(t)} \leq S^{i,\mathbf{0}}_{(t)}$, for any $\boldsymbol{\lambda} \geq 0$.

(c) $S^{i,\mathbf{0}}_{(t)} \leq \bar{S} \equiv F^{-1}\left(\frac{\bar{p}-(1-\gamma)\underline{c}}{\bar{p}+\underline{h}}\right)$.

Proof. See Appendix A.1. □

Therefore, the optimal policy ρ_λ for any Lagrangian relaxed problem for any Lagrangian multiplier vector $\boldsymbol{\lambda}$ is a multi-item combination of double base stock policies.

The Lagrangian dual

$$\max_{\boldsymbol{\lambda}\geq 0} \; W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}) \tag{1.10}$$

is a concave program and can be solved via many numerical methods, such as a steepest ascent method, or FISTA (Beck and Teboulle, 2009). In Section 1.6, we provide a complete description of our algorithm using FISTA. The following lemma shows that the computation complexity of using FISTA to obtain an ϵ-optimal Lagrange multiplier of the Lagrangian dual is $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$.

Lemma 1. *(a)* $\lambda^*_{(t)} \leq \bar{p} + \bar{c}$ *for all* t*, where* $\bar{p}, \bar{c}$ *are defined in the notation section.*

(b) The complexity of using FISTA to obtain an ϵ*-optimal Lagrange multiplierof the Lagrangian dual is* $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$.

Proof. See Appendix B.5. □

Let $\boldsymbol{\lambda}^*$ denote the Lagrange multiplier vector which optimizes (1.10), and let $\rho^* = \rho_{\boldsymbol{\lambda}^*}$ denote the corresponding policy. The feasible region described by (1.6) is convex and admits at least one point $\bar{\boldsymbol{x}}_{(t)} = \mathbf{0}$. Thus, we have strong duality implying that

$$\boldsymbol{\lambda}^* \geq 0, \quad \mathbb{E}\left[\sum_{i=1}^{I}(\bar{x}^{\rho^*,i}_{(t)} - \dot{u}^i_{(t)})\Big|\boldsymbol{x}_{(0)}\right] \leq \chi'_{(t+\ell)}, \quad \boldsymbol{\lambda}^{*\top}\left(\mathbb{E}\left[\sum_{i=1}^{I}(\bar{x}^{\rho^*,i}_{(t)} - \dot{u}^i_{(t)})\Big|\boldsymbol{x}_{(0)}\right] - \chi'_{(t+\ell)}\right) = 0,$$

where $\bar{\boldsymbol{x}}^{\rho^*}$ denotes the decisions corresponding to the optimal policy ρ^*.

Theorem 2 (Solution of the expected value constrained problem)**.** *Consider the expected value constrained problem defining* $W(\{\chi'_{(t)}\})$*. This problem can be solved via its Lagrangian dual (relaxing* (1.6) *with Lagrangian penalties). For any given vector of Lagrangian multipliers* $\boldsymbol{\lambda} \geq 0$*, the optimal policy for the*

Lagrangian relaxed problem, is a double base stock policy. Moreover, the policy $\rho^* \equiv \rho_{\boldsymbol{\lambda}^*}$ which is optimal when $\boldsymbol{\lambda} = \boldsymbol{\lambda}^*$ is also feasible for the expected value constrained problem.

Remarkably, when there are only expected value constraints with respect to the aggregate inventory, like in (1.2) or (1.6), each item may be managed based on its own inventory position only; the required coordination between the item is achieved through the use of the common optimal Lagrange vector $\boldsymbol{\lambda}^*$.

1.5 Bounding the Problem with Mixed Constraints by Two Expected Value Constrained Dynamic Programs

In this section, we show that the original problem (1.5) can be sandwiched between two dynamic programs with (unconditional) expected value constraints only.

We first show that the chance constraints (1.3) can be restricted to expected value constraints corresponding with slightly reduced capacity levels $\{\chi'_{(t)}\}$. To establish this, we need the following lemma.

Lemma 2. *Consider any "product line-specific" strategy that prescribes decisions for each item i based on inventory information pertaining to its own product line only, i.e., the decision $\bar{x}^i_{(t)}$ is independent of $\bar{x}^{i'}_{(t)}$ for all $i \neq i'$ belonging to different product lines and all t by Assumption 4. Define $Y_{(t)} := \sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$ to denote the aggregate inventory level at the beginning of period $t + \ell$.*

(a)

$$\mathbb{P}\Big(\Big| Y_{(t)} - \mathbb{E}[Y_{(t)}] \Big| \geq s \Big) \leq 2e^{-\frac{s^2}{2\sigma^2}} \tag{1.11}$$

where $\sigma = \sqrt{K} m U (T - t + 1)$.

(b) Fix $t = 0, \ldots, T - \ell$. If for a given "product line-specific" strategy, $\mathbb{E}\big[\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+\big] \leq \chi_{(t+\ell)} - \gamma$, where

$$\gamma := \sigma \sqrt{-2 \ln(\frac{\beta}{2})}, \tag{1.12}$$

then $\mathbb{P}\big(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{(t+\ell)} \big) \leq \beta$.

Proof. (a) By Hoeffding (1963)'s inequality, see e.g. (2.11) in Wainwright (2019), it suffices to show that

$Y_{(t)}$ is a sub-Gaussian random variable with sub-Gaussian parameter σ.

For each product line G_k, define $Z_k := \sum_{i \in G_k} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$, then $Y_{(t)} = \sum_{k=1}^{K} Z_k$.

The total demand for product line i over the interval $[t, T]$ is bounded by $U(T-t+1)$. Consequently, it is suboptimal for any item i to have an inventory position $\bar{x}^i_{(t)} > U(T-t+1)$ as there is no arbitrage opportunity given (1.1). Thus, since every product line has at most m products, and since all $u^i_{(t)} \geq 0$, we have

$$0 \leq Z^k_{(t)} \leq \sum_{i \in G_k} \bar{x}^i_{(t)} \leq mU(T-t+1),$$

where the upper bound is independent of I. As a result, $Z^k_{(t)}$ is a sub-Gaussian random variable with sub-Gaussian parameter $\hat{\sigma} := mU(T-t+1)$.

For all $i \neq i'$ belonging to different product lines and all t, $\bar{x}^i_{(t)}$ is independent of $\bar{x}^{i'}_{(t)}$ by the "product line-specific" property of the strategy, and $\dot{u}^i_{(t)}$ is independent of $\dot{u}^{i'}_{(t)}$ by Assumption 4. Thus, Z_k is independent of $Z_{k'}$ for all $k \neq k'$. Therefore, $Y_{(t)}$ is sub-Gaussian with parameter $\sigma = \sqrt{K}\hat{\sigma} = \sqrt{K}mU(T-t+1)$.

(b) Thus, if for a given strategy, $\mathbb{E}[\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+] \leq \chi_{(t+\ell)} - \gamma$, the likelihood of an overflow beyond $\chi_{(t+\ell)}$ is bounded by $2e^{-\frac{\gamma^2}{2\sigma^2}}$. Equating $2e^{-\frac{\gamma^2}{2\sigma^2}} = \beta$ we get that by specifying $\gamma = \sigma\sqrt{-2\ln(\frac{\beta}{2})}$, if a solution satisfies $\mathbb{E}[\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+] \leq \chi_{(t+\ell)} - \gamma$, the overflow probability $\mathbb{P}(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{(t+\ell)}) \leq \beta$.

$\square$

Let C^* denote the optimal cost in the original problem (1.5), and $W^*(\{\chi'_{(t)}\})$ be as defined in Section 1.4. We now derive the following theorem.

Theorem 3. *(Bounding the chance constrained problem by expected value constrained problems)*

$$W^*(\{\chi_{(t)} + R\}_{t=0}^{T}) \leq C^* \leq W^*(\{\chi_{(t)} - \gamma\}_{t=0}^{T}),$$

where $R \geq 0$ is defined in (1.2) *and $\gamma \geq 0$ is defined in* (1.12).

Proof. The left inequality is simple. The two problems for which C^* and $W^*(\{\chi_{(t)} + R\}_{t=0}^T)$ are the optimal values are identical except for the fact that the former has additional chance constraints. Hence, $W^*(\{\chi_{(t)} + R\}_{t=0}^T) \leq C^*$.

To prove the right inequality, let $\rho^*(\gamma)$ denote the optimal policy to the problem defining $W^*(\{\chi_{(t)} - \gamma\}_{t=0}^T)$. By Assumption 1, $\chi_{(t)} - \gamma \geq 0$ for sufficiently large value of I. Thus, $W^*(\{\chi_{(t)} - \gamma\}_{t=0}^T)$ admits at least $\bar{\boldsymbol{x}}_{(t)} = \mathbf{0}$ as a feasible solution for each period t, and therefore, a feasible policy exists. Then, it suffices to show that $\rho^*(\gamma)$ is a feasible policy to the original problem (1.5). As shown in Section 1.4, $\rho^*(\gamma)$ employs double base stock policies for each item, thus it is a "product line-specific" policy. By Lemma 2, it satisfies the chance constraints (1.3). In addition, since $-\gamma \leq 0 \leq R$, it also satisfies the expected value constraints (1.2). Thus, $\rho^*(\gamma)$ is a feasible policy to the original problem (1.5). $\square$

Thus, we have sandwiched our problem with a feasible region that is complex and highly non-convex between two expected value constrained problems, which can be easily solved to optimality as shown in Section 1.4. In the next section, we show that the policy $\rho^*(\gamma)$ is an asymptotically optimal policy for the original problem with mixed constraints.

1.6 An Asymptotically Optimal Heuristic Policy for the Problem with Mixed Constraints

Consider the policy $\rho^*(\gamma)$ which is optimal for (the expected value constrained) problem $W^*(\{\chi_{(t)} - \gamma\}_{t=0}^T)$. As shown in the proof of Theorem 3, this policy is feasible for the original problem with mixed constraints. To provide a bound for its optimality gap we extend the bounds in Theorem 3. Let $\boldsymbol{\lambda}^*_{(\{\chi_{(t)} - \gamma\})}$ denote the optimal vector of Lagrangian multipliers in problem $W^*(\{\chi_{(t)} - \gamma\}_{t=0}^T)$.

Corollary 1.

$$W^*(\{\chi_{(t)} - \gamma\}_{t=0}^T) \geq C^* \geq W^*(\{\chi_{(t)} + R\}_{t=0}^T) \geq W^*(\{\chi_{(t)} - \gamma\}_{t=0}^T) - {\boldsymbol{\lambda}^*_{(\{\chi_{(t)} - \gamma\})}}^\top (R + \gamma)\boldsymbol{e}.$$

Proof. It is easily verified that $W^*(\{\chi'_{(t)}\})$ is a jointly convex function of the T-vector $\{\chi'_{(t)}\}_{t=0}^T$ and

$\boldsymbol{\lambda}^*_{(\{\chi_{(t)}\}-\gamma \boldsymbol{e})}$ is its gradient. The last bound follows. $\square$

While policy $\rho^*(\gamma)$ is always an effective heuristic, it can be shown that it is, in fact, asymptotically optimal under *independent* demands; or more generally, under Assumption 4.

Theorem 4. *(Asymptotic optimality)*

Under Assumption 4, the policy $\rho^*(\gamma)$ is asymptotically optimal for the original problem, and $W^*(\{\chi_{(t)}\} + R\boldsymbol{e})$ is an asymptotically accurate lower bound.

Proof. In view of corollary 1, the absolute optimality gap is bounded by $(R+\gamma)\|\boldsymbol{\lambda}^*_{(\{\chi_{(t)}\}-\gamma \boldsymbol{e})}\|_1 \leq T(R+\gamma)(\bar{p}+\bar{c})$, where the inequality follows from Lemma 1(a). R is sublinear in I by definition, and $\gamma = O(\sqrt{K}) = O(\sqrt{I})$ by Lemma 2. Asymptotic optimality of the policy $\rho^*(\gamma)$ follows from $C^* = \Omega(I)$. The latter follows from the fact that the problem with no capacity constraints, which can be decomposed into I problems, has an optimal cost $\Omega(I)$ that is a lower bound for C^*. $\square$

We finish this section with a formal description of the algorithm.

Algorithm

Step 1: Replace the capacity levels $\chi_{(t)}$ by $\chi_{(t)} := \chi_{(t)} - \gamma$ where γ is defined in Lemma 2.

Step 2: Select an arbitrary vector of Lagrange multipliers $\boldsymbol{\lambda}$ and solve $W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ in (1.7), as I separate single-dimensional dynamic programs, in each of which a double base stock policy is optimal.

Step 3: Apply FISTA, see Beck and Teboulle (2009) to update the vector $\boldsymbol{\lambda}$ until an ϵ-optimal vector $\boldsymbol{\lambda}$ is found for the (unconstrained) concave program:

$$W^*(\{\chi_{(t)}\} - \gamma \boldsymbol{e}) = \max_{\boldsymbol{\lambda} \geq \mathbf{0}} W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}).$$

Each time the vector $\boldsymbol{\lambda}$ is adapted, return to Step 2.

1.7 Normal Demands

1.7.1 A Normal Demand Model

In this section, we extend our results to the case where the demand vector $\boldsymbol{u}_{(t)}$, and hence the lead time vector $\dot{\boldsymbol{u}}_{(t)}$, follows a multivariate Normal distribution, i.e., $\boldsymbol{u}_{(t)} \sim \mathcal{N}(\boldsymbol{\mu}_{(t)}, \boldsymbol{\Sigma}_{(t)})$. Let $\sigma^i_{(t)}$ be the standard deviation of $u^i_{(t)}$. Without loss of generality, we assume $\sigma^i_{(t)} \leq \bar{\sigma}$ for some constant $\bar{\sigma}$ for all i, t. Since Normals have unbounded support, a different analysis technique is required.

Since any Normal distribution can take on negative values, $\mathbb{E}[\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+] = \Omega(I)$ even for $\bar{\boldsymbol{x}}_{(t)} = \mathbf{0}$. In order for the original expected value constraints (1.2) to admit a feasible solution as I grows, we need to make the following assumption with respect to the capacity levels, which is stronger than Assumption 1.

Assumption 3. *The capacity values $\{\chi_{(t)}\}$ grow (at least) linearly in I, i.e., $\chi_{(t)} = \Omega(I)$, for all $t = 1, \ldots, T$.*

The optimal policy ρ^*_γ to $W^*(\{\chi_{(t)} - \gamma \boldsymbol{e}\})$ is still an I-tuple of double base stock policies, as the proof of this structured result applies to general continuous distributions with bounded or unbounded support. Thus, to establish Theorem 3, Corrolary 1 and the asymptotic result in Theorem 4, it suffices to show that inequality (1.11) still holds under Normal demands for some sub-Gaussian parameter $\sigma = O(\sqrt{I})$.

We write $\boldsymbol{\Sigma}_{(t)} = \boldsymbol{S}_{(t)}\boldsymbol{P}_{(t)}\boldsymbol{S}_{(t)}$, where $\boldsymbol{S}_{(t)} \equiv \operatorname{diag}(\boldsymbol{\sigma}_{(t)})$ and $\boldsymbol{P}_{(t)}$ is the correlation matrix. Let $\boldsymbol{L}_{(t)}$ be the Cholesky decomposition of the correlation matrix (i.e., $\boldsymbol{P}_{(t)} = \boldsymbol{L}_{(t)}\boldsymbol{L}^\top_{(t)}$), and let $\boldsymbol{L}^i_{(t)}$ denote the i-th row of $\boldsymbol{L}_{(t)}$. We can then write $\boldsymbol{u}_{(t)} = \boldsymbol{\mu}_{(t)} + \boldsymbol{S}_{(t)}\boldsymbol{L}_{(t)}\boldsymbol{z}_{(t)}$, where $\boldsymbol{z}_{(t)}$ is a vector of independent standard Normals that generates the demand process. We will write $\boldsymbol{z}_{(0\to t)} \equiv \left(\boldsymbol{z}^\top_{(0)}, \ldots, \boldsymbol{z}^\top_{(t)}\right)^\top$.

By Theorem 2.26 in Wainwright (2019), it is sufficient to show that under policy ρ^*_γ, the functions $Y_{(t)}$ of the vector of *independent* standard Normals $\boldsymbol{z}_{(0\to t+\ell-1)}$ are σ-Lipschitz continuous for some $\sigma = O(\sqrt{I})$. We show this by establishing the following lemma.

Lemma 3. *Assume Assumptions 4 and 3 apply. Assume a given starting state $\boldsymbol{x}_{(0)}$. For two sequences $\{\boldsymbol{z}_{(t)}(\omega_1)\}_{t=0}^T$ and $\{\boldsymbol{z}_{(t)}(\omega_2)\}_{t=0}^T$, let $\{x^{*i}_{(t)}\}_{t=0}^T$ and $\{y^{*i}_{(t)}\}_{t=0}^T$ be the trajectories of starting inventory positions of item i under policy ρ^*_γ, respectively. Let $\{\bar{x}^{*i}_{(t)}\}_{t=0}^T$ and $\{\bar{y}^{*i}_{(t)}\}_{t=0}^T$ be the corresponding actions taken by ρ^*_γ, under $\{\boldsymbol{z}(\omega_1)\}_{t=0}^T$ and $\{\boldsymbol{z}(\omega_2)\}_{t=0}^T$, respectively.*

(a) For any i and t,

$$\big|x^{*i}_{(t)} - y^{*i}_{(t)}\big| \le \sum_{\tau=0}^{t-1} \big|\sigma^i_{(t)} \boldsymbol{L}^i_{(\tau)} \big(\boldsymbol{z}_{(\tau)}(\omega_1) - \boldsymbol{z}_{(\tau)}(\omega_2)\big)\big|.$$

(b) Let $\bar{\boldsymbol{L}}_{(t)} := \big|\boldsymbol{L}_{(t)}\big|$ be the element-wise absolute value of $\boldsymbol{L}_{(t)}$. For any t,

$$\begin{aligned}
&\Big|\sum_{i=1}^{I} \big(\bar{x}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_1)\big)^+ - \sum_{i=1}^{I} \big(\bar{y}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_2)\big)^+\Big| \\
\le\ & \Big(\sum_{\tau=0}^{t+\ell-1} \big\|\boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)}\big\|\Big) \big\|\boldsymbol{z}_{(0\to t+\ell-1)}(\omega_1) - \boldsymbol{z}_{(0\to t+\ell-1)}(\omega_2)\big\|,
\end{aligned}$$

i.e. $Y_{(t)}$ is Lipschitz continuous in $\boldsymbol{z}_{(0\to t+\ell-1)}$ with Lipschitz constant $\sum_{\tau=0}^{t+\ell-1} \big\|\boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)}\big\|$.

(c) Under Assumption 4,

$$\sum_{\tau=0}^{t+\ell-1} \big\|\boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)}\big\| \le \bar{\sigma} T m \sqrt{I} = O(\sqrt{I}), \tag{1.13}$$

where m is the constant defined in Assumption 4.

Proof. (a) We prove this by forward induction with respect to t. At $t = 0$, $\big|x^{*i}_{(0)} - y^{*i}_{(0)}\big| = 0$ by assumption, the inequality trivially holds.

Assume the inequality holds for a general value of t.

$$\begin{aligned}
\big|x^{*i}_{(t+1)} - y^{*i}_{(t+1)}\big| &= \big|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)} - u^i_{(t)}(\omega_1) + u^i_{(t)}(\omega_2)\big| \\
&\le \big|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\big| + \big|u^i_{(t)}(\omega_1) - u^i_{(t)}(\omega_2)\big| \\
&= \big|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\big| + \big|\sigma^i_{(t)} \boldsymbol{L}^i_{(t)} \big(\boldsymbol{z}_{(t)}(\omega_1) - \boldsymbol{z}_{(t)}(\omega_2)\big)\big| \\
&\le \big|x^{*i}_{(t)} - y^{*i}_{(t)}\big| + \big|\sigma^i_{(t)} \boldsymbol{L}^i_{(t)} \big(\boldsymbol{z}_{(t)}(\omega_1) - \boldsymbol{z}_{(t)}(\omega_2)\big)\big| \\
&\le \sum_{\tau=0}^{t} \big|\sigma^i_{(\tau)} \boldsymbol{L}^i_{(\tau)} \big(\boldsymbol{z}_{(\tau)}(\omega_1) - \boldsymbol{z}_{(\tau)}(\omega_2)\big)\big|
\end{aligned} \tag{1.14}$$

where the second inequality follows from the double base stock structure of ρ^*_γ, as is easily verified from considering the three possibilities (i) $x^{*i}_{(t)} \le S^{*i}_{(t)}$, (ii) $S^{*i}_{(t)} \le x^{*i}_{(t)} \le B^{*i}_{(t)}$ and (iii) $x^{*i}_{(t)} \ge B^{*i}_{(t)}$, in combination with the three analogous possibilities for $y^{*i}_{(t)}$. The last inequality follows from the induction assumption. This completes the induction proof.

(b) First, we show that

$$LHS := \left|\left(\bar{x}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_1)\right)^+ - \left(\bar{y}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_2)\right)^+\right| \leq \left|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\right| + \left|\dot{u}^i_{(t)}(\omega_1) - \dot{u}^i_{(t)}(\omega_2)\right| =: RHS.$$

Without loss of generality, assume $\bar{x}^{*i}_{(t)} > \bar{y}^{*i}_{(t)}$. Now, consider:

(i) If $\dot{u}^i_{(t)}(\omega_1) \geq \bar{x}^{*i}_{(t)}$ and $\dot{u}^i_{(t)}(\omega_2) \geq \bar{y}^{*i}_{(t)}$, $LHS = 0 \leq RHS.$

(ii) If $\dot{u}^i_{(t)}(\omega_1) \leq \bar{x}^{*i}_{(t)}$ and $\dot{u}^i_{(t)}(\omega_2) \geq \bar{y}^{*i}_{(t)}$, $LHS = \bar{x}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_1) \leq \bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)} \leq RHS.$

(iii) If $\dot{u}^i_{(t)}(\omega_1) \geq \bar{x}^{*i}_{(t)}$ and $\dot{u}^i_{(t)}(\omega_2) \leq \bar{y}^{*i}_{(t)}$, $LHS = \bar{y}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_2) \leq \bar{x}^{i}_{(t)} - \dot{u}^i_{(t)}(\omega_2) \leq \dot{u}^i_{(t)}(\omega_1) - \dot{u}^i_{(t)}(\omega_2) \leq RHS.$

(iv) If $\dot{u}^i_{(t)}(\omega_1) \leq \bar{x}^{*i}_{(t)}$ and $\dot{u}^i_{(t)}(\omega_2) \leq \bar{y}^{*i}_{(t)}$, $LHS = \bar{x}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_1) - \bar{y}^{*i}_{(t)} + \dot{u}^i_{(t)}(\omega_2) \leq \left|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\right| + \left|\dot{u}^i_{(t)}(\omega_2) - \dot{u}^i_{(t)}(\omega_1)\right| = RHS.$

Thus, we have $LHS \leq RHS.$ Then,

$$\begin{aligned}
\left|\left(\bar{x}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_1)\right)^+ - \left(\bar{y}^{*i}_{(t)} - \dot{u}^i_{(t)}(\omega_2)\right)^+\right| &\leq \left|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\right| + \sum_{\tau=t}^{\tau=t+\ell-1} \left|u^i_{(t)}(\omega_1) - u^i_{(t)}(\omega_2)\right| \\
&\leq \left|x^{*i}_{(t)} - y^{*i}_{(t)}\right| + \sum_{\tau=t}^{\tau=t+\ell-1} \left|u^i_{(t)}(\omega_1) - u^i_{(t)}(\omega_2)\right| \\
&\leq \sum_{\tau=0}^{t+\ell-1} \left|\sigma^i_{(\tau)} \boldsymbol{L}^i_{(\tau)}\left(\boldsymbol{z}_{(\tau)}(\omega_1) - \boldsymbol{z}_{(\tau)}(\omega_2)\right)\right|,
\end{aligned}$$

where the second inequality was shown in part (a) and the third inequality follows (1.14).

Thus,

$$
\begin{aligned}
&\Big|\sum_{i=1}^{I}\big(\bar{x}^{*i}_{(t)}-\dot{u}^{i}_{(t)}(\omega_1)\big)^{+}-\sum_{i=1}^{I}\big(\bar{y}^{*i}_{(t)}-\dot{u}^{i}_{(t)}(\omega_2)\big)^{+}\Big|\\
\le&\sum_{i=1}^{I}\Big|\big(\bar{x}^{*i}_{(t)}-\dot{u}^{i}_{(t)}(\omega_1)\big)^{+}-\big(\bar{y}^{*i}_{(t)}-\dot{u}^{i}_{(t)}(\omega_2)\big)^{+}\Big|\\
=&\sum_{\tau=0}^{t+\ell-1}\boldsymbol{e}^{\top}\big|\boldsymbol{s}_{(\tau)}\boldsymbol{L}_{(\tau)}\big(\boldsymbol{z}_{(\tau)}(\omega_1)-\boldsymbol{z}_{(\tau)}(\omega_2)\big)\big|\\
\le&\sum_{\tau=0}^{t+\ell-1}\boldsymbol{e}^{\top}\boldsymbol{s}_{(\tau)}\bar{\boldsymbol{L}}_{(\tau)}\big|\boldsymbol{z}_{(\tau)}(\omega_1)-\boldsymbol{z}_{(\tau)}(\omega_2)\big|\\
\le&\sum_{\tau=0}^{t+\ell-1}\|\boldsymbol{\sigma}_{(\tau)}^{\top}\bar{\boldsymbol{L}}_{(\tau)}\|\|\boldsymbol{z}_{(\tau)}(\omega_1)-\boldsymbol{z}_{(\tau)}(\omega_2)\|\\
\le&\Big(\sum_{\tau=0}^{t+\ell-1}\|\boldsymbol{\sigma}_{(\tau)}^{\top}\bar{\boldsymbol{L}}_{(\tau)}\|\Big)\|\boldsymbol{z}_{(0\to t+\ell-1)}(\omega_1)-\boldsymbol{z}_{(0\to t+\ell-1)}(\omega_2)\|,
\end{aligned}
$$

where the second to last inequality follows from the Cauchy Schwartz inequality.

(c) It suffices to show $\|\boldsymbol{\sigma}_{(t)}^{\top}\bar{\boldsymbol{L}}_{(t)}\|\le\bar{\sigma}m\sqrt{I}=O(\sqrt{I})$ for any t.

Under Assumption 4, $\boldsymbol{P}_{(t)}$ is a block diagonal matrix, and so are $\boldsymbol{L}_{(t)}$ and $\bar{\boldsymbol{L}}_{(t)}$. Thus, each column of $\bar{\boldsymbol{L}}_{(t)}$ has at most m non-zero elements. Since $\boldsymbol{L}_{(t)}\boldsymbol{L}_{(t)}^{\top}=\boldsymbol{P}_{(t)}$, $\big|\big(\boldsymbol{L}_{(t)}\big)_{i,j}\big|\le 1$ for all i,j. Then,

$$
\|\boldsymbol{\sigma}_{(t)}^{\top}\bar{\boldsymbol{L}}_{(t)}\|\le\bar{\sigma}\|\boldsymbol{e}^{\top}\bar{\boldsymbol{L}}_{(t)}\|\le\bar{\sigma}\sqrt{m^2 I}=O(\sqrt{I}).
$$

□

1.7.2 A Truncated Normal Demand Model

Normal demands, while most frequently used in practice and in the inventory literature, have the undesirable property that one or more of their components can take on arbitrarily negative values. An alternative model that has demands non-negative *almost surely*, is the truncated Normal demand model, that is,

$$
\boldsymbol{u}_{(t)}^{Tr}=\boldsymbol{u}_{(t)}^{+},
$$

where $\boldsymbol{u}_{(t)} \sim \mathcal{N}(\boldsymbol{\mu}_{(t)}, \boldsymbol{\Sigma}_{(t)})$.

Our asymptotic optimality result can be extended to this truncated Normal model with only minor modifications needed to the proof of Lemma 3.

In part (a), we have

$$\begin{aligned}
|x^{*i}_{(t+1)} - y^{*i}_{(t+1)}| &= |\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)} - u^{i,+}_{(t)}(\omega_1) + u^{i,+}_{(t)}(\omega_2)| \\
&\leq |\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}| + |u^{i,+}_{(t)}(\omega_1) - u^{i,+}_{(t)}(\omega_2)| \\
&\leq |\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}| + |u^{i}_{(t)}(\omega_1) - u^{i}_{(t)}(\omega_2)|
\end{aligned}$$

where the second inequality follows from the fact that, for any two realizations of the Normal demand $u^i_{(t)}(\omega_1)$ and $u^i_{(t)}(\omega_2)$,

$$|{u^{i}}^{+}_{(t)}(\omega_1) - {u^{i}}^{+}_{(t)}(\omega_2)| \leq |u^i_{(t)}(\omega_1) - u^i_{(t)}(\omega_2)|. \tag{1.15}$$

Then the rest of (1.14) follows readily.

In part (b), $LHS \leq RHS$ holds for any random variable $\dot{u}^i_{(t)}$, and thus it holds when $\dot{u}^i_{(t)}$ is replaced by $\dot{u}^{i,+}_{(t)} := \sum_{\tau=t}^{t+\ell-1} u^{i,+}_{(t)}$. Then,

$$\begin{aligned}
\left|(\bar{x}^{*i}_{(t)} - \dot{u}^{i,+}_{(t)}(\omega_1))^+ - (\bar{y}^{*i}_{(t)} - \dot{u}^{i,+}_{(t)}(\omega_2))^+\right| &\leq \left|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\right| + \left|\dot{u}^{i,+}_{(t)}(\omega_1) - \dot{u}^{i,+}_{(t)}(\omega_2)\right| \\
&\leq \left|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\right| + \sum_{\tau=t}^{\tau=t+\ell-1} \left|u^{i,+}_{(t)}(\omega_1) - u^{i,+}_{(t)}(\omega_2)\right| \\
&\leq \left|\bar{x}^{*i}_{(t)} - \bar{y}^{*i}_{(t)}\right| + \sum_{\tau=t}^{\tau=t+\ell-1} \left|u^{i}_{(t)}(\omega_1) - u^{i}_{(t)}(\omega_2)\right|,
\end{aligned}$$

where the last inequality again follows from (1.15). Then the rest of part (b) follows readily.

Part (c) applies without modifications. Therefore, Lemma 3 applies in the truncated Normal demand model.

1.8 Numerical Study

In this section, we report the results of a numerical study comprising 240 instances to evaluate the performance of the Algorithm proposed in Section 1.6, in terms of optimality gap and computational time, as the number of items I is increased from $I = 50$ to $I = 1,000$.

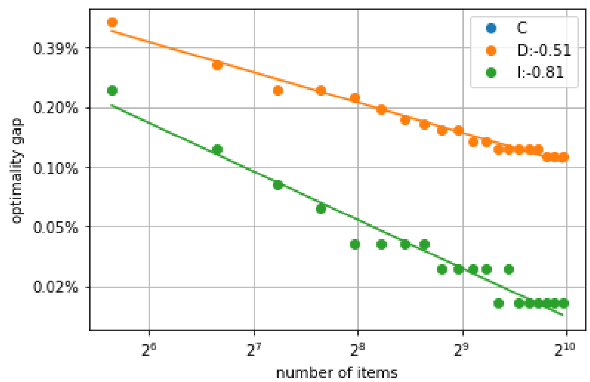
The instances all have $T = 20$ periods, $m = 10$ items per product line and Normal demands. The results of the numerical study are displayed in Figures 1.1 and 1.2. In Figure 1.1, the instances marked 'C' (for 'constant') have $\mu^i_{(t)} = 10$ for all i, t. Those marked as 'I' (resp. 'D') have $\mu^i_{(t)} = 5 + 10\frac{t-1}{T-1}$ (resp. $\mu^i_{(t)} = 5 + 10\frac{T-t}{T-1}$), i.e. the means are increasing (resp. decreasing) over time. Half of the instances assume independent demands, while the remaining half have a positive correlation of $\rho = 0.3$ for every pair of items within the same product line. Coefficients of variation are set at 0.3 for half of the instances and 0.4 for the remaining half.

The remaining cost parameters were set as follows. All holding cost rates $h^i_{(t)} = 10$, backlogging cost rates $p^i_{(t)} = 10h^i_{(t)}$, and variable procurement rates $c^i_{(t)} = 5h^i_{(t)}$. The lead time was set to $\ell = 4$ periods. Capacities $\chi_{(t)}$ were defined as

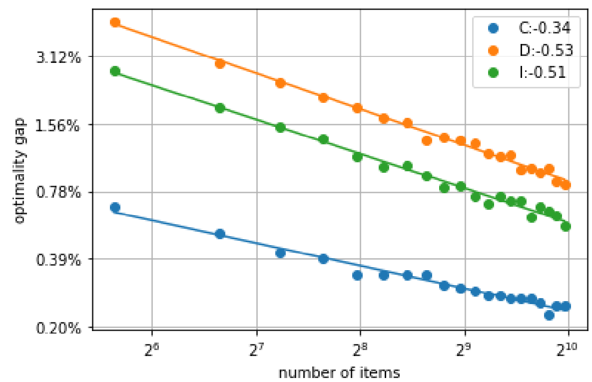
$$\chi_{(t)} \equiv \chi \equiv \frac{1}{T}\sum_{\tau=1}^{T}\sum_{i=1}^{I}\mu^i_{(\tau)} + 2\cdot\sqrt{\ell}\left[\frac{1}{T}\sum_{\tau=1}^{T}\sqrt{\sum_{i=1}^{I}(\sigma^i_{(\tau)})^2}\right].$$

Note that $\chi_{(t)} = O(I)$ because the first term grows linearly in I. The surplus above the mean demand (averaged over all periods) given by the second term is specified as twice the standard deviation of the lead time aggregate demand assuming independent demands. This is the case even for the 120 instances where the items within the same product line have a positive correlation $\rho = 0.3$ and the lead time aggregate demand standard deviation is considerably larger.

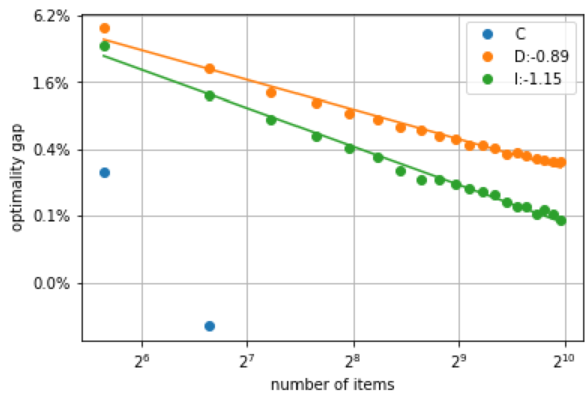
Finally, salvage costs are kept constant across all instances at $d = \max_i\left(\sum_{\tau=1}^{T} h^i_{(\tau)}\right)$. This value is chosen to ensure it is always cheaper to hold an item (even for the full planning horizon) than to salvage it. The salvage option is nevertheless used if it is absolutely necessary to meet the capacity constraint. Recall that salvage costs were included in our formulation to account for situations with merchandise return options but also to ensure feasibility.



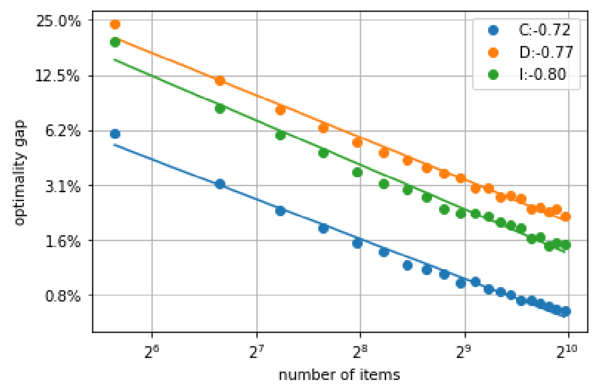
(a) $\rho = 0$, CVBase = 0.3



(b) $\rho = 0$, CVBase = 0.4



(c) $\rho = 0.3$, CVBase = 0.3



(d) $\rho = 0.3$, CVBase = 0.4

Figure 1.1: Asymptotic optimality (in logarithmic scale). In (a), gap is 0 for constant mean demand 'C'. Linear regression lines are displayed with the slopes in the legend.

Figure 1.1 exhibits the optimality gaps as I increases from 50 to 1,000 in logarithmic scale together with regression lines for different sets of instances. The gaps decline rapidly to zero throughout, confirming asymptotic optimality. Under independent demands, they decline to 0.1% at a rate of 0.51 for CVBase = 0.3 and to 1% at a rate of 0.53 for CVBase = 0.4 in the worst case when demands exhibit decreasing pattern. Under positively correlated demands, in the worst case, the gaps decline to below 0.5% and 2.5% at the rates of 0.89 and 0.77 for CVBase = 0.3 and 0.4, respectively. Throughout, optimality gaps decline almost monotonically and rapidly.

Note that the theoretically required value of the Sub-Gaussian parameter in Lemma 3(b) is often very large, and as a consequence so is γ, the amount by which the capacity needs to be reduced to ensure that the chance constraint is satisfied, see (1.12). We use a significantly smaller value $\tilde{\gamma}$ in our numerical studies:

$$\tilde{\gamma} \equiv \sigma\sqrt{-2\log\left(\frac{\beta}{2}\right)},$$

where $\sigma \equiv \max_{t=0,\cdots,T}(\tilde{\sigma}_{(t)})$, $\tilde{\sigma}^2_{(t)} \equiv \max_{\boldsymbol{a}\in[0,1]^I} \boldsymbol{a}^\top \dot{\boldsymbol{\Sigma}}_{(t)}\boldsymbol{a} \leq \boldsymbol{e}^\top|\dot{\boldsymbol{\Sigma}}_{(t)}|\boldsymbol{e}$, and $\boldsymbol{e}$ is a vector of ones.

This value is obtained by heuristically assuming that in every period, the aggregate inventories in the different product lines are independent even though this is not the case under arbitrarily inventory policies. We omit the details of the calculation. Our numerical results show that the smaller $\tilde{\gamma}$ is sufficient for asymptotic optimality to hold.

Figure 1.2 shows the computational time needed for calculating both the lower and upper bounds as the number of items I increases, averaged over all 12 instances for each value of I. This confirms that the practical computational complexity is linear in I, as is the theoretical complexity.

1.9 Conclusions and Generalizations

We have analyzed a periodic review stochastic inventory model with I items where demands are correlated and time-dependent. Joint chance constraints with respect to capacity limits for the aggregate inventory level at the beginning of each period, tie the different items together. We have designed a simple heuristic, and associated lower bound, which are, respectively, asymptotically optimal and asymptotically accurate. The

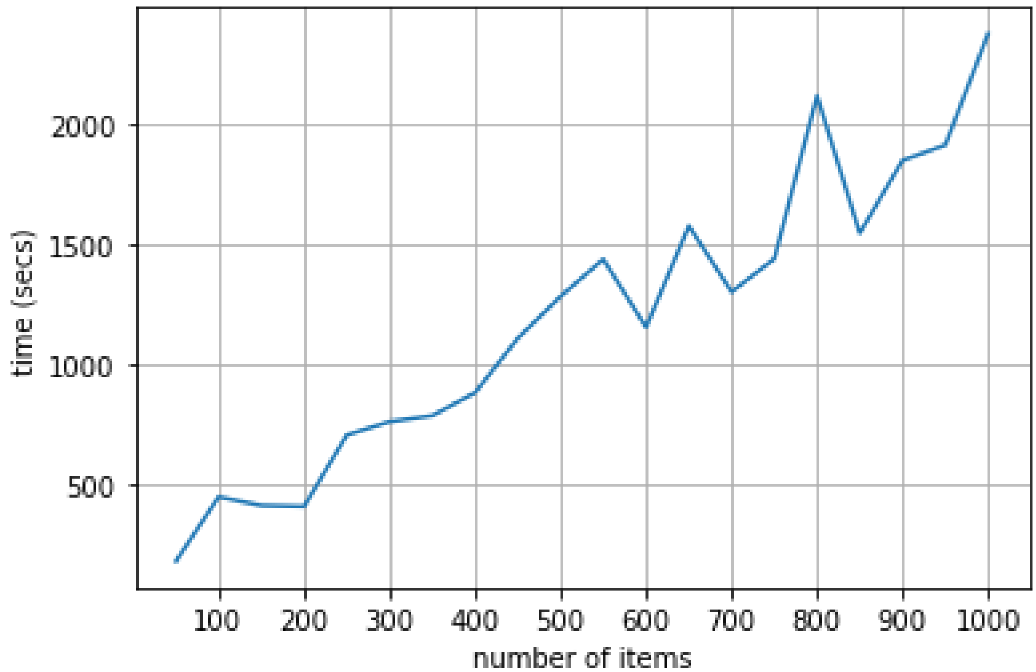


Figure 1.2: Computational time for evaluating both the lower and upper bounds.

policy manages each item independently via a double base stock policy, driven by T Lagrange multipliers from a common Lagrangian dual. The computational effort is $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$, to achieve an asymptotically ϵ-optimal heuristic. The optimality gap is in fact $O(\sqrt{I})$. All of this was achieved for demand distributions that are either multivariate Normal, or a family of multivariate truncated Normals, and lastly general distributions but of bounded support. Below we address two important generalizations: the long-run average cost criterion and the lost sales model.

1.9.1 The Long-Run Average Cost criterion

When the model is stationary, i.e., all model parameters, demand distributions and inventory constraints are time independent, the average long-run cost criterion could be used as an alternative to the total discounted costs over a finite horizon T. Emphasizing the dependence or the planning horizon T we have from Theorem 3, that

$$W^{*T}(\chi\boldsymbol{e} + R\boldsymbol{e}) \leq C^{*T} \leq W^{*T}(\chi\boldsymbol{e} - \gamma\boldsymbol{e}). \tag{1.16}$$

Let $\tilde{C}, \tilde{W}(\chi+\alpha\boldsymbol{e})$ and $\tilde{W}(\chi-\gamma\boldsymbol{e})$ denote the long-run average cost values in the original chance constrained problem and the problems with in each period an expected value constraint corresponding with a capacity level of $\chi+\alpha$ and $\chi-\gamma$, respectively. Dividing (1.16) by T and letting $T \to \infty$ we obtain:

$$\tilde{W}^*(\chi+R) \leq \tilde{C}^* \leq \tilde{W}^*(\chi-\gamma)$$

as a direct extension of Theorem 3 to the long-run average cost case. Similarly, we obtain as the analogue of Corollary 1 that

$$\tilde{W}(\chi-\gamma) \geq \tilde{C} \geq \tilde{W}(\chi+R) \geq \tilde{W}(\chi-\gamma) - \lambda^*_{(\chi-\gamma)}(R+\gamma), \tag{1.17}$$

where $\lambda^*_{(\chi-\gamma)}$ is the *single* optimal Lagrange multiplier in the infinite horizon model with an expected value cost constraint and capacity level of $(\chi-\gamma)$ in every single period. As before, we relax the expected value constraint through a Lagrangian relaxation with (a single) Lagrange multiplier λ. The resulting relaxed problem again decomposes into I separate single-item models, the infinite horizon long-run average cost analogue of $W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ in (1.7). (Under the long-run average costs criterion, the dependence on the initial state $\boldsymbol{x}_{(0)}$ vanishes.) For any given choice of the Lagrange multiplier λ, we again get double base stock policies to be optimal for any item $i = 1, \ldots, I$ as in Theorem 2. Since there is only one Lagrange multiplier, its optimal value λ^* can be found by a simple bisection method and the corresponding set of double base stock policies constitute again a heuristic that is easy to implement and asymptotically optimal as $I \to \infty$ when each of the items is correlated with a bounded number of other items, since again $\lambda^*_{(\chi-\gamma)} \leq p^* + c^*$, and R and γ are both sublinear in I. $\tilde{W}(\chi+R)$, which can be computed in the same way, provides again an asymptotically accurate lower bound.

1.9.2 Lost Sales

As is well known, when stockouts result in lost sales, the problem becomes intrinsically harder, even when there is only $I = 1$ item to be managed. The reason is that even in the single-item model, a $(\ell+1)$-dimensional state description is required, accounting e.g., for the inventory *level* at the beginning of each period, as well as the ℓ-dimensional pipeline of outstanding orders. As a consequence, it it prohibitively difficult to compute an optimal policy, unless ℓ is small, e.g., $\ell \leq 2$. Moreover, the structure of such a policy is complex, and the large literature on this model has confined itself mostly to heuristic policies, see e.g., Bijvank and Vis (2011).

Recently, Xin and Goldberg (2016) showed that for the long-run average cost criterion, a simple policy under which a constant size order is placed in every period has a worst-case optimality gap of $hK\xi^{\ell+1}$, where h represents the holding cost rate, and K and ξ are parameter-dependent constants, with $0 < \xi < 1$.

In a constrained multi-item model, the approach outlined in the previous section, can be used to obtain a policy ρ^* which uses for each item a constant order size policy, and which has a worst-case optimality gap which is asymptotically $(h + \lambda^*)K\xi^{\ell+1}$. Here λ^* is the optimal Lagrange multiplier of the (single) constraint corresponding with a reduced inventory capacity level of $(\chi - \gamma)$. (Once again, $\lambda^* \leq c^* + p^*$.) To verify this result, let $\tilde{C}(\rho^*)$ be the cost of policy ρ^*, with $\tilde{C}(\rho^*) \leq \tilde{W}(\chi - \gamma)\big(1 + (h + \lambda^*)K\xi^{\ell+1}\big) \leq \big(\tilde{C} + \lambda^*(\chi - \gamma)(R + \gamma)\big)\big(1 + (h + \lambda^*)K\xi^{\ell+1}\big)$, where the last inequality follows from (1.17). Thus,

$$\frac{\tilde{C}(\rho^*)}{\tilde{C}} \leq \big(1 + (h + \lambda^*)(\chi - \gamma)K\xi^{\ell+1}\big) + \frac{\lambda^*(\chi - \gamma)(R + \gamma)}{\tilde{C}}.$$

The last term is $o(1)$, as $I \to \infty$, since α and γ are $O(\sqrt{I})$ and $\tilde{C} = \Omega(I)$.

Chapter 2: Scalable Approximately Optimal Policies for Multi-Item Stochastic Inventory Problems with Conditional Chance Inventory Constraints

In the previous chapter, we analyzed a periodic review inventory system where multiple items share a common storage capacity constraints. Two types of capacity constraints are imposed, unconditional expected value constraints bounding the expected inventory level, and unconditional probability constraints bounding the expected overflow probability, both evaluated given the starting state of the planning horizon.

In this chapter, we consider a similar problem but with more complexity where the expected inventory level or overflow probability is evaluated conditioned on the starting state of each period. This is equivalent to having an infinite number of constraints in every period for each of the two types of constraints, as opposed to one constraint only in the unconditional case. Since the chance type constraints are much more complex than the expected value constraints as we have seen in the previous chapter, we will focus on how to handle conditional chance constrained problems.

2.1 Introduction and Summary

We analyze a single location, multi-item, multi-period, stochastic inventory model in which the items share a capacitated storage facility or in which inventory investments are limited by aggregate budget constraints. These joint inventory constraints introduce a complex interdependency among the different items – storing more of one item requires ordering less of another. The problem is especially challenging in the presence of lead times since the aggregate inventory level at the end of a given period depends not only on the inventory positions and orders placed in the past, but also on unknown stochastic demands that are realized over the lead time after the order. This relationship is, in addition, highly non-linear.

We impose chance constraints on the overflow in each period that limits the probability of the aggregate inventory exceeding the prevailing capacity level. These chance constraints are specified on *every* possible sample path; we refer to them as *conditional* chance constraints (conditional on the starting inventory level in that time period), as opposed to *unconditional* chance constraints, which limit only the *expected probability* of the overflow in any given period, given a specific starting state at the start of our time horizon.

Because of these constraints, the exact optimal policy for this model has a very complex structure and is prohibitively difficult to compute. In the absence of an optimal strategy, one might hope for an approximate procurement strategy, ideally one that meets one or more of the following criteria (i) a relatively simple structure, (ii) computability with a complexity which scales well with the number of items I, (iii) asymptotic optimality as the number of items I grows, (iv) empirical closeness to optimality in a wide variety of settings even for small values of I. Despite the importance and fundamental nature of this model, no effective approximate procurement strategies have been proposed to date, let alone strategies that meet any of these four desiderata. In this paper, we achieve all of these aims. In particular, we design a heuristic which is asymptotically optimal as I grows to infinity, and asymptotically efficient as well. In parallel, we design an alternative heuristic which will be shown to perform excellently for moderate as well as large values of I, based on extensive numerical studies.

Our model

Demands for the I items, in any given period in the planning horizon, are described by a general multivariate Normal distribution. As in most standard inventory models, we assume that the demand vectors are independent across time. (In Section 2.7 and Appendix B.2, we relax both assumptions.) The system managers may, at the beginning of each period, adjust the inventory position of each item upwards by placing an order at a given unit cost rate, or downwards by salvaging part of the inventory, again at a given cost rate or revenue per unit. In addition to the order and salvaging costs, there are standard holding and backlogging costs proportional to the inventory on hand or the backlog at the beginning of the period. All cost functions and parameters, as well as the joint demand distributions may be time dependent.

Procurement lead times aggravate the problem in a major way: when periodically selecting procurement quantities for the different items, the capacity limit needs to be accounted for at the time the orders arrive at the storage facility, by which time the demand realized over the intervening periods will have drawn down the aggregate inventory. Furthermore, during this lead time, some items may have run out of stock. As a result, the aggregate inventory level at the end of the lead time is not simply given by the aggregate inventory position at the start of the lead time minus the aggregate demand in the lead time interval.

The conditional chance constraints define a very complex non-convex set of possible actions in every state and period, in any dynamic programming approach to the problem. Even the verification of whether a given vector of inventory positions satisfies any of the chance constraints involves the evaluation of a complex I-dimensional integral or a Monte Carlo simulation.

An asymptotically optimal heuristic

We first design a heuristic which is asymptotically optimal as I grows to infinity. We believe this to be the relevant asymptotic result because in many applications the number of items can go into the ten thousands to millions, if not beyond, for large retailers, e.g. Amazon, Walmart or Alibaba and many others. (The heuristic is not asymptotically optimal in T, but the asymptotics with respect to T are unimportant since the length of a reasonable planning horizon is naturally bounded.)

The heuristic is designed by first computing the optimal policy when the (conditional) chance constraints are replaced by (unconditional) expected value constraints with capacity values reduced by an appropriately chosen quantity κ. The latter problem, as seen in Chapter 1, has only T constraints and can be solved exactly and efficiently via its Lagrangian dual without any duality gap. We have shown that for any $\epsilon > 0$, an ϵ-optimal policy can be computed in $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$ time. This policy is then transformed into our heuristic policy $\boldsymbol{\pi}$. We design policy $\boldsymbol{\pi}$ to satisfy *conditional* expected value constraints with capacity levels still reduced relative to the actual capacity levels, but by a quantity $\kappa_0 < \kappa$. The transformation involves, in any given period, the reduction in the inventory positions prescribed by the original policy, and this is done by minimizing a separable convex function subject to a single separable constraint. $\boldsymbol{\pi}$ is shown to satisfy the original chance constraints and we prove that its expected cost exceeds the optimal cost by $O(\sqrt{I \log I})$ when the items can be partitioned into product lines, such that only items within any product line may be correlated. Since the optimal cost is $\Omega(I)$, this shows the optimal gap grows slower than the optimal cost, thus proving asymptotic optimality.

A practical heuristic for moderate I

While the heuristic described in the previous section is asymptotically optimal and computable with modest effort, we recommend a different heuristic when the number of items I is small or moderate. In extensive and systematic numerical studies involving 28,800 instances, the average optimality gap for this second heuristic is 1.05%, and 98.2% of instances have gaps less than 5%.

Our approach, is to develop tight upper and lower bounds for the optimal cost value, along with an easily implementable procurement strategy which is anchored on the strategy that optimizes the lower bound problem. The worst case complexity of the lower bound problem is $O(I|\mathcal{A}|^2T^{\frac{3}{2}})$, and that of the upper bound problem is $O(TI^2 \log I)$, where $|\mathcal{A}|$ denotes the number of (linear) constraints required to obtain an adequate relaxation of the feasible chance-constrained action sets in any given state. The lower bound procedure relaxes these constrains in a Lagrangian manner, and finds the tightest possible bound of this type. Almost invariably, $|\mathcal{A}| = O(I)$; moreover, for the suggested sets $\mathcal{A}$, and under natural parallelization, the worst-case time complexity of the lower bound procedure reduces to $O(|\mathcal{A}|T^{\frac{3}{2}})$.

Managerial Insights

The high accuracy and scalability of the lower bound problem that drives the practical heuristic described in the previous section suggests that it can be effectively used in capacity planning models that determine optimal investments or reservations of storage capacity. As pointed out in a recent review paper by Song et al. (2020), very few studies integrate capacity planning (of any kind) with the determination of operational procurement strategies. Reviewing the MSOM literature over the past 20 years, they found no more than 7% of inventory or capacity related papers that qualify as such. We believe this paper does precisely that; indeed, the Lagrange multiplier vectors optimizing the Lagrangian Dual can be used directly to assess the (marginal) value of storage capacity expansions.

Storage utilization and the resulting supply chain costs are a major factor retailers need to consider in making fulfilment decisions. Perhaps even more importantly, retailers need to consider these costs when *designing* their procurement systems. For example, in deciding how much space to rent to retailers and for how much, Amazon needs a method to determine how valuable that space is. In deciding what assortment to stock, a retailer needs to take into consideration correlations among items and the assortment's impact on storage capacities. To our knowledge, the lack of even approximate procurement strategies for systems with these constraints makes it impossible to consider these costs in a principled fashion.

In section 2.6, we use our technique to demonstrate how retailers can use our technique to answer these important managerial questions. We show how to quantify the cost-impact of various important system parameters such as capacity and lead time, thus giving retailers a principled way to decide whether to invest in improving these parameters. We discuss how our technique can help a retailer decide on the optimal assortment to stock. Finally, we consider a simple example in which a retailer has to allocate a fixed number of products to one of two similar fulfilment facilities, and we show that the presence of a capacity constraint makes the task of deciding what product to fulfil from where nontrivial.

Prior work

We believe that the model considered in this paper is one of the most basic multi-item stochastic inventory model. It is surprising how little progress has been made to address this or related models. The model we consider was first proposed in a seminal paper by Veinott Jr (1965), later refined in Ignall and Veinott Jr (1969). Both focused on identifying conditions under which a myopic base stock policy is optimal in a setting with stationary demand distributions and parameters. Both papers assume, however, that orders arrive instantaneously, in which case the storage capacity constraints are easily expressed as a simple upper bound on the items' inventory levels. Ignall and Veinott Jr (1969) have a brief discussion of the case where there are positive lead times. Assuming a worst case demand scenario for all items, the inventory constraints once again reduce to simple upper bounds on the sum of the inventory positions. This is an extremely conservative representation that leaves much of the available storage space unused.

Later contributions, including DeCroix and Arreola-Risa (1998), Beyer et al. (2001, 2002), Shi et al. (2016) and Chen and Li (2019), have continued to confine themselves to the case of stationary models without lead times.

Planning models integrating assortment planning with inventory decisions have largely been confined to single period models without repeated replenishments or lead times, see e.g., Gaur and Honhon (2006), Mahajan and Van Ryzin (2001) and Rusmevichientong et al. (2010) and the references therein. Indeed Caro et al. (2020) in their review article write "Assortment optimization has been a popular area of research specific to retail, dealing with the critical question of which products to offer to customers. Combined with pricing and inventory decisions, the problem quickly becomes intractable."

Marklund and Rosling (2012), Nambiar et al. (2021), and Miao et al. (2022) consider a periodic review, multiple warehouse and retailer system in which the warehouses are replenished in common cycles. For any given cycle with an initial inventory at the warehouses – which cannot be replenished during the cycle – the problem is how to best procure the retailers over the course of the cycle. A simple unconditional expected value constraint ensures that the shipments from the warehouses (assumed to be instantaneous) do not exceed the given initial inventories there in expectation. The authors develop asymptotically optimal heuristics, based on Lagrangian relaxations of the constraints.

2.2 Model and Notation

We consider a periodic review system with a planning horizon of T periods storing I distinct items with random demands. At the start of each period, inventories may be increased or reduced through orders or salvages; these decisions take effect after ℓ periods. All stockouts are fully backlogged. A common storage facility or aggregate inventory budget has only a limited inventory capacity. Without loss of generality, we assume that a unit of every item occupies one unit of storage capacity (the budget). In addition to variable order and salvaging costs, there are holding and backlogging costs, specified as proportional to the beginning-of-period inventory levels. Demands for the I items are specified by a general multivariate, possibly time dependent, Normal distribution; as in almost all inventory models, demands are assumed to be independent across time, but not necessarily across products. See Section 2.7 for generalizations. Our objective is to minimize expected total (discounted) costs over the full planning horizon, subject to joint inventory constraints.

We now summarize our notation; for all variables, we will use a parenthesized subscript to denote time, and a superscript to denote the product. Bold symbols without superscripts will denote vectors over all products. For example, $\boldsymbol{c}_{(t)} \equiv [c^1_{(t)}, \cdots, c^I_{(t)}]$ will denote the vector of unit ordering costs for all products in period t.

- $c^i_{(t)}$: the per unit ordering cost for product i in period t, where $c^i_{(t)} \in [\underline{c}, \bar{c}]$.

- $d^i_{(t)}$: the per unit salvaging cost for product i in period t, where $d^i_{(t)} \in [\underline{d}, \bar{d}]$.

- $h^i_{(t)}$: the inventory holding cost per unit of product i held at the beginning of period t, assumed to be in $[\underline{h}, \bar{h}]$.

- $p^i_{(t)}$: the backlogging cost per unit of product i backlogged at the beginning of period t, assumed to be in $[\underline{p}, \bar{p}]$. To avoid an unbounded arbitrage opportunity, we assume, without loss of generality that for all i and t and $\tau \geq t$,

$$c^i_{(t)} + \gamma h^i_{(t+1)} + \ldots + \gamma^{\tau-t} h^i_{(\tau)} + \gamma^{\tau-t} d^i_{(\tau)} > 0. \tag{2.1}$$

- $\chi_{(t)}$: the inventory capacity at the end of period t.

- $u^i_{(t)}$: the demand for product i in period t. We assume $\boldsymbol{u}_{(t)}$ has a general multivariate Normal distribution, i.e., $\boldsymbol{u}_{(t)} \sim \mathcal{N}(\boldsymbol{\mu}_{(t)}, \boldsymbol{\Sigma}_{(t)})$. Letting $\mu^i_{(t)}$ and $\sigma^i_{(t)}$ denote the mean and standard deviation of $u^i_{(t)}$, we assume $\mu^i_{(t)} \in [\underline{\mu}, \bar{\mu}]$ and $\sigma^i_{(t)} \in [\underline{\sigma}, \bar{\sigma}]$ for all i, t.

  We use $\dot{u}^i_{(t)} \equiv \sum_{\tau=t}^{t+\ell-1} u^i_{(\tau)}$ to denote the demand for product i, during a lead time of ℓ periods starting with period t. We define $\dot{\boldsymbol{u}}_{(t)}$, $\dot{\boldsymbol{\mu}}_{(t)}$, $\dot{\boldsymbol{\Sigma}}_{(t)}$, $\dot{\mu}^i_{(t)}$, and $\dot{\sigma}^i_{(t)}$ accordingly.

- $x^i_{(t)}$: the inventory position for product i at the beginning of period t *before* placement of an order.

- $\bar{x}^i_{(t)}$: the inventory position for product i, at the beginning of period t, but *after* placement of an order. This is the variable the decision maker will have to choose in each period to minimize costs.

- γ: the discount factor. $0 \leq \gamma \leq 1$.

- $\boldsymbol{e}$: the vector of ones in $\mathbb{R}^I$.

The total inventory on hand at the beginning of period $t + \ell$ is given by $\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$, where $x^+ \equiv \max(x, 0)$ denotes the positive part of x (we will later use $x^- \equiv -\min(x, 0)$ to denote the negative part of x). We impose constraints on this inventory as chance constraints of the following type:

$$\mathbb{P}_{\dot{\boldsymbol{u}}_{(t)}}\left(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{(t+\ell)} \middle| \boldsymbol{x}_{(t)} \right) \leq \beta, \qquad t = 0, \ldots, T - \ell \text{ and all } x_{(t)}, \tag{2.2}$$

for a given (small) probability $0 < \beta < 1$. We emphasize that the probabilities for all periods $t = 1, \ldots, T$ are assessed with respect to any initial state $\boldsymbol{x}_{(t)}$.

Applying a standard accounting scheme, we charge to period t not the actual holding and backlogging costs incurred during that period, but the expected value of the costs incurred one lead time later, at the beginning of period $(t+\ell)$. Those costs are a function of $\bar{x}^i_{(t)}$ (the decision made in period t) and are denoted

$$Q^i_{(t)}(\bar{x}^i_{(t)}) \equiv \gamma^\ell \mathbb{E}_{\dot{u}^i_{(t)}} \left\{ h^i_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ + p^i_{(t+\ell)}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^- \right\},$$

and we assume $\lim\limits_{x \uparrow \infty} \frac{d}{dx} Q^i_{(t)}(x) > 0$.

In every period, the decision maker must choose $\bar{x}^i_{(t)}$ for every item i. The decision is lower-bounded by $\min(x^i_{(t)}, 0)$ (no 'artificial backlogs' can be created). It is also natural to upper-bound these decisions by $M^i_{(t)} = \chi_{(t+\ell)} + \dot{\mu}^i_{(t)} + F^{-1}(\beta)\dot{\sigma}^i_{(t)}$, where $F(\cdot)$ is the c.d.f of the standard Normal, since picking $\bar{x}^i_{(t)} > M^i_{(t)}$ would result in a violation of the chance constraint on account of item i, alone.

Thus, we let $V_{(t)}(\boldsymbol{x}_{(t)})$ denote the minimal expected present value of costs incurred in periods $t, t+1, \cdots, T$ when starting in state $\boldsymbol{x}_{(t)}$. With $V_{(T+1)}(\cdot) = 0$, we have

$$\text{(DP)} \quad V_{(t)}(\boldsymbol{x}_{(t)}) = \min_{\min\{\boldsymbol{x}_{(t)}, 0\} \leq \bar{\boldsymbol{x}}_{(t)} \leq \boldsymbol{M}_{(t)}} \left\{ \sum_{i=1}^{I} \left(c^i_{(t)}(\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)}(\bar{x}^i_{(t)} - x^i_{(t)})^- \right. \right. \tag{2.3}$$

$$\left. \left. + Q^i_{(t)}(\bar{x}^i_{(t)}) \right) + \gamma \mathbb{E}_{\boldsymbol{u}_{(t)}} \left[V_{(t+1)}(\bar{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}) \right] \right\}$$

$$\text{s.t.} \quad \mathbb{P}_{\dot{\boldsymbol{u}}_{(t)}} \left(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{(t+\ell)} \,\middle|\, \boldsymbol{x}_{(t)} \right) \leq \beta. \tag{2.2}$$

At this stage, it is worth noting that for certain combinations of parameters, the ability to salvage inventory is essential for this dynamic program to be feasible. To see this, consider a $T = 10$ period problem with lead time $\ell = 1$, very low holding costs, high expected demand in periods 1 through 5, and very low expected demand in periods 6 through 10. In period 4, given the lead time $\ell = 1$ and the expected high demand in period 5, the optimal solution might be to order a very large quantity – the high expected demand in period 5 would make it very likely that most of this large order is depleted in period 5, thus ensuring our large order does not violate our capacity constraints. Consider, however, a sample path (albeit a low-probability one) in which demand in period 5 ends up very low. We now find ourselves 'stuck' with a very large inventory at the end of period 5. Given the low expected demand in period 6, an overflow is almost sure to occur in that period. Since the chance constraint needs to be satisfied even in this unlikely starting state $\boldsymbol{x}_{(t)}$; only the option to salvage inventory can enable a feasible solution.

Except for very small values of I, (DP) is computationally intractable. Furthermore, even testing whether a given vector $\bar{x}$ satisfies one of the chance constraints is numerically very challenging.

2.3 An Asymptotically Optimal Heuristic

In this section, we design an asymptotically optimal heuristic. We do this for the following general correlation pattern and capacity model.

Assumption 4. *The items $\{1, \ldots, I\}$ can be clustered into disjoint product lines $\{G_1, \ldots, G_K\}$ such that any two items in different product lines have independent demands with $\max_k |G_k| \leq m$ for some constant m that does not grow with I.*

Thus, demands are allowed to have a fully general multivariate Normal distribution, but they are assumed to be correlated within product lines only. The heuristic developed in Section 2.4, allows, in contrast, for an arbitrary correlation pattern.

Assumption 5. *The storage capacity grows (at least) linearly in the number of items I, i.e., $\chi_{(t)} = \Omega(I)$ for all t.*

This assumption is an artifact of the Normal demands, having support on the negative half line; it is necessary to guarantee that when $\bar{\boldsymbol{x}} \equiv \mathbf{0}$, even the *expected* inventory on hand $\mathbb{E}[\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+] = \mathbb{E}[\sum_{i=1}^{I}(\dot{u}^i_{(t)})^-] = O(I)$ stays below the capacity level. For demand variables that are non-negative, the storage capacity only needs to grow as $\sqrt{I \log I}$, see Appendix B.2, reflecting economies of scope.

The chance constraints (2.2) can be written as

$$\mathbb{P}_{\dot{\boldsymbol{u}}_{(t)}}\left(\max_{\boldsymbol{a} \in \{0,1\}^I} \sum_{i=1}^{I} a^i (\bar{x}^i_{(t)} - \dot{u}^i_{(t)}) \geq \chi_{(t+\ell)} \middle| \boldsymbol{x}_{(t)} \right) \leq \beta. \tag{2.4}$$

Furthermore, we can write the overflow event as a function of a vector of *independent* Normals, by employing the Cholesky decomposition of the variance-covariance matrix of $\dot{\boldsymbol{u}}_{(t)}$. In particular, let $\dot{\boldsymbol{\Sigma}}_{(t)} = \dot{\boldsymbol{V}}_{(t)}\dot{\boldsymbol{V}}_{(t)}^\top$ so that $\dot{\boldsymbol{u}}_{(t)} = \dot{\boldsymbol{\mu}}_{(t)} + \dot{\boldsymbol{V}}_{(t)}\boldsymbol{z}$, with $\boldsymbol{z}$ a vector of *independent* standard Normal variables. The overflow probability can be written as the (right-hand) tail probability of a function of this $\boldsymbol{z}$ vector. In particular, let

$$f_{(t)}(\boldsymbol{z}) \equiv \max_{\boldsymbol{a}\in\{0,1\}^I} f_{(t)}^{\boldsymbol{a}}(\boldsymbol{z}),$$

where

$$f_{(t)}^{\boldsymbol{a}}(\boldsymbol{z}) = \sum_{i=1}^{I} a_i(\bar{x}_{(t)}^i - \dot{\boldsymbol{\mu}}_{(t)}^i - \dot{\boldsymbol{v}}_{(t)}^i\boldsymbol{z}),$$

and $\dot{\boldsymbol{v}}_{(t)}^i$ denotes the i-th row of the matrix $\dot{\boldsymbol{V}}_{(t)}^i$. The chance constraints (2.4) can be written as

$$\mathbb{P}(f_{(t)}(\boldsymbol{z}) \geq \chi_{(t+\ell)}|\boldsymbol{x}_{(t)}) \leq \beta. \tag{2.5}$$

We show that the chance constraints (2.5) can be restricted to expected value constraints corresponding with slightly reduced capacity levels. To establish this, we need the following lemma.

Lemma 4. *(a)*

$$\mathbb{P}\Big(f_{(t)}(\boldsymbol{z}) - \mathbb{E}[f_{(t)}(\boldsymbol{z})] \geq s\Big|\boldsymbol{x}_{(t)}\Big) \leq 2e^{-\frac{s^2}{2\sigma^2}}$$

where $\sigma^2 = \max_{t=0,\ldots,T} \tilde{\sigma}_{(t)}^2$ *and* $\tilde{\sigma}_{(t)}^2 = \max_{\boldsymbol{a}\in\{0,1\}^I} \boldsymbol{a}^\top\dot{\boldsymbol{\Sigma}}_{(t)}\boldsymbol{a} > 0.$

(b) Fix $t = 0, \ldots, T-\ell$. *If for a given strategy,* $\mathbb{E}\big[f_{(t)}(\boldsymbol{z})|\boldsymbol{x}_{(t)}\big] \leq \chi_{(t+\ell)} - \kappa_0$, *for*

$$\kappa_0 \equiv \sigma\sqrt{-2\log\left(\frac{\beta}{2}\right)}, \tag{2.6}$$

then $\mathbb{P}(f_{(t)}(\boldsymbol{z}) \geq \chi_{(t+\ell)}|\boldsymbol{x}_{(t)}) \leq \beta$.

Proof. (a) By Theorem 2.26 in Wainwright (2019), it is sufficient to show that the functions $f_{(t)}(\boldsymbol{z})$ of the vector of *independent* standard Normals are $\sigma_{(t)}$-Lipschitz continuous. (A function $g(\boldsymbol{z})$ is L-Lipschitz continuous, if for any $\boldsymbol{z}_1, \boldsymbol{z}_2, g(\boldsymbol{z}_2) - g(\boldsymbol{z}_1) \leq L\|\boldsymbol{z}_2 - \boldsymbol{z}_1\|_2$.) We can do this by noting that for any pair of vectors $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$,

$$\begin{aligned}
f(\boldsymbol{Z}_1) &= \max_{\boldsymbol{a}\in\{0,1\}^I}\left(\sum_{i=1}^{I} a^i(\bar{x}^i - \dot{\mu}^i - \dot{\boldsymbol{v}}^i\cdot\boldsymbol{Z}_1)\right)\\
&= \max_{\boldsymbol{a}\in\{0,1\}^I}\left(\sum_{i=1}^{I} a^i(\bar{x}^i - \dot{\mu}^i - \dot{\boldsymbol{v}}^i\cdot\boldsymbol{Z}_2) + \boldsymbol{a}^\top\dot{\boldsymbol{V}}(\boldsymbol{Z}_2-\boldsymbol{Z}_1)\right)\\
&\le \max_{\boldsymbol{a}\in\{0,1\}^I}\left(\sum_{i=1}^{I} a^i(\bar{x}^i - \dot{\mu}^i - \dot{\boldsymbol{v}}^i\cdot\boldsymbol{Z}_2)\right) + \max_{\boldsymbol{a}\in\{0,1\}^I}\left(\boldsymbol{a}^\top\dot{\boldsymbol{V}}(\boldsymbol{Z}_2-\boldsymbol{Z}_1)\right)\\
&= f(\boldsymbol{Z}_2) + \max_{\boldsymbol{a}\in\{0,1\}^I}\left(\boldsymbol{a}^\top\dot{\boldsymbol{V}}\right)(\boldsymbol{Z}_2-\boldsymbol{Z}_1)\\
&\le f(\boldsymbol{Z}_2) + \max_{\boldsymbol{a}\in\{0,1\}^I}\left(\|\boldsymbol{a}^\top\dot{\boldsymbol{V}}\|\right)\|\boldsymbol{Z}_2-\boldsymbol{Z}_1\|\\
&= f(\boldsymbol{Z}_2) + \left(\sqrt{\max_{\boldsymbol{a}\in\{0,1\}^I}\boldsymbol{a}^\top\dot{\boldsymbol{\Sigma}}\boldsymbol{a}}\right)\|\boldsymbol{Z}_2-\boldsymbol{Z}_1\|,
\end{aligned}$$

where the second-to-last line follows by the Cauchy-Schwartz inequality, and the last line makes use of the fact $\dot{\boldsymbol{\Sigma}} = \dot{\boldsymbol{V}}\dot{\boldsymbol{V}}^\top$. Thus, f is $\sigma_{(t)}$-Lipschitz continuous.

(b) Thus, if for a given strategy, $\mathbb{E}[f_{(t)}(\boldsymbol{z})|\boldsymbol{x}_{(t)}] \le \chi_{(t+\ell)} - \kappa_0$, the likelihood of an overflow beyond $\chi_{(t+\ell)}$ is bounded by $2e^{-\frac{\gamma^2}{2\sigma^2}}$. Equating $2e^{-\frac{\gamma^2}{2\sigma^2}} = \beta$ we get that by specifying $\kappa_0 = \sigma\sqrt{-2\ln(\frac{\beta}{2})}$, if a solution satisfies $\mathbb{E}[f_{(t)}(\boldsymbol{z})|\boldsymbol{x}_{(t)}] \le \chi_{(t+\ell)} - \gamma$, the overflow probability $\mathbb{P}(f_{(t)}(\boldsymbol{z}) \ge \chi_{(t+\ell)}|\boldsymbol{x}_{(t)}) \le \beta$.

□

Our first step is then to approximate our problem (DP) by a dynamic program ($\ddot{\text{DP}}$) in which the feasible action spaces described by the constraints (2.2) are replaced by convex sets described by the following (conditional) expected value constraints in each period

$$\ddot{g}_{(t|\boldsymbol{x}_{(t)})}(\bar{\boldsymbol{x}}_{(t)}) \equiv \mathbb{E}_{\dot{\boldsymbol{u}}_{(t)}}\left[\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+\middle|\boldsymbol{x}_{(t)}\right] \le \chi_{(t+\ell)} - \kappa_0, \qquad \forall \boldsymbol{x}_{(t)} \in \mathbb{R}^I \tag{$\ddot{7}$}$$

where κ_0 is defined in (2.6). In Appendix B.1, we will show that under Assumption 4, σ is $O(\sqrt{I})$, and because $\chi_{(t+\ell)}$ is $O(I)$ under Assumption 5, this implies that for large enough I, the RHS of ($\ddot{7}$) will be positive. By Lemma 4(b), the set of convex constraints ($\ddot{7}$) with our choice of κ_0 is a *restriction* of our original constraints (2.2).

Even though the dynamic program ($\ddot{\text{DP}}$) has a convex feasible region, it still comprises an infinite number of constraints (one for each possible starting state $\boldsymbol{x}_{(t)}$ in each period) and cannot easily be solved. As a consequence, we construct yet another dynamic program ($\check{\text{DP}}_\kappa$) which replaces ($\ddot{7}$) by T *unconditional* constraints

$$\check{g}_{(t|\boldsymbol{x}_{(0)})}(\bar{\boldsymbol{x}}_{(t)}) \equiv \mathbb{E}_{\{\boldsymbol{u}_{(0)},\cdots\boldsymbol{u}_{(t-1)}\}}\left(\overbrace{\mathbb{E}_{\dot{\boldsymbol{u}}_{(t)}}\left[\sum_{i=1}^{I}(\bar{x}^i_{(t)}-\dot{u}^i_{(t)})^+|\boldsymbol{x}_{(t)}\right]}^{=\ddot{g}_{(t|\boldsymbol{x}_{(t)})}(\bar{\boldsymbol{x}}_{(t)})}\middle|\boldsymbol{x}_{(0)}\right) \leq \chi_{(t+\ell)} - \kappa \tag{$\check{8}$}$$

with $\kappa \geq \kappa_0$. These constraints can be viewed as expectations over all sample paths and thus aggregations of the conditional constraints ($\ddot{7}$). In other words, if κ were equal to κ_0, constraints ($\check{8}$) would relax the conditional constraints ($\ddot{7}$). We will later compensate for this relaxation by setting $\kappa > \kappa_0$, thus making the capacity more restrictive.

Let $\check{\boldsymbol{\rho}}^\star_\kappa : \mathbb{R}^{I\times T} \to \mathbb{R}^{I\times T}$ denote the optimal policy for the $\check{\text{DP}}_\kappa$ dynamic program; specifically, $\check{\rho}^{\star,i}_{\kappa,(t)}(\boldsymbol{x}_{(t)}) : \mathbb{R}^I \to \mathbb{R}$ denotes the action that the policy prescribes for item i in period t when faced with state $\boldsymbol{x}_{(t)}$ at the start of that period. The cost of any given policy is denoted $C(\cdot)$, the optimal cost for (DP) in (2.3) is denoted $C^\star$, and the optimal cost for ($\check{\text{DP}}_\kappa$) is denoted $\check{C}^\star_\kappa \equiv C(\check{\boldsymbol{\rho}}^\star_\kappa)$.

In Chapter 1, we have shown that the dynamic program with unconditional expected value constraints ($\check{8}$) can be solved efficiently. In the following, we will give a brief review of how to compute $\check{C}^\star_\kappa$ and $\check{\boldsymbol{\rho}}^\star_\kappa$. Besdies, we will show that the optimal policy $\check{\boldsymbol{\rho}}^\star_\kappa$ can be transformed into a policy $\boldsymbol{\pi}$ that *is* feasible in the original chance-constrained dynamic program (DP). Moreover, the move from $\check{\boldsymbol{\rho}}^\star_\kappa$ to $\boldsymbol{\pi}$ only increases $\check{C}^\star_\kappa$ by an $O(1)$ term. Finally, we will show that the difference $\check{C}^\star_\kappa - C^\star$ is bounded by an $O(\sqrt{I\log I})$ term, and therefore, the gap between the cost of our constructed policy $C(\boldsymbol{\pi})$ and the optimal policy $C^\star$, grows as $O(\sqrt{I\log I})$. Since $C^\star$ itself is $\Omega(I)$, we conclude that $\boldsymbol{\pi}$ is asymptotically optimal in I.

Solving ($\check{\text{DP}}_\kappa$) **and finding** $\check{\boldsymbol{\rho}}^\star_\kappa$

The construction of policy $\boldsymbol{\pi}$ starts with the solution of ($\check{\text{DP}}_\kappa$) for a specific value of $\kappa \geq \kappa_0$. This dynamic program, as seen in Chapter 1, has only T convex constraints and can therefore be solved *exactly* via its Lagrangian dual: Lagrangian relaxation applies a Lagrange multiplier $\lambda_{(t)} \geq 0$ to the t-th constraint in ($\check{8}$), replacing it by the penalty term $\lambda_{(t)}\left[\check{g}_{(t)}(\bar{\boldsymbol{x}}_{(t)}) - (\chi_{(t+\ell)} - \kappa)\right]$ in the objective. For any $\boldsymbol{\lambda} \equiv$

$[\lambda_{(1)}, \cdots, \lambda_{(T)}] \geq \mathbf{0}$, this gives a new dynamic program $(\check{\mathrm{DP}}^{\boldsymbol{\lambda}}_{\kappa})$, with optimal value

$$\check{C}^{\boldsymbol{\lambda},\star}_{\kappa} = \min_{\min\{\boldsymbol{x}_{(t)},0\} \leq \bar{\boldsymbol{x}}_{(t)}} \mathbb{E}_{\{\boldsymbol{u}_{(0)},\cdots,\boldsymbol{u}_{(T)}\}} \Bigg[\sum_{t=1}^{T} \gamma^{t-1} \Bigg\{ \sum_{i=1}^{I} \Big[c^i_{(t)} (\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)} (\bar{x}^i_{(t)} - x^i_{(t)})^- \\ + Q^i_{(t)}(\bar{x}^i_{(t)}) \Big] + \lambda_{(t)} \Big[\check{g}_{(t|\boldsymbol{x}_{(0)})}(\bar{\boldsymbol{x}}_{(t)}) - \chi_{(t+\ell)} + \kappa \Big] \Bigg\} \Bigg], \quad (2.9)$$

and because of strong duality, $\check{C}^{\star}_{\kappa} = \max_{\boldsymbol{\lambda} \geq 0} \check{C}^{\boldsymbol{\lambda},\star}_{\kappa}$. We let $\boldsymbol{\lambda}^{\star}_{\kappa} \equiv \operatorname*{argmax}_{\boldsymbol{\lambda} \geq 0} \check{C}^{\boldsymbol{\lambda},\star}_{\kappa}$.

It is easily seen that this dynamic program can be decomposed into I single-item dynamic programs $(\check{\mathrm{DP}}^{i,\boldsymbol{\lambda}}_{\kappa})$ with optimal values

$$\check{C}^{i,\boldsymbol{\lambda},\star}_{\kappa} = \min_{\min\{x^i_{(t)},0\} \leq \bar{x}^i_{(t)}} \mathbb{E}_{\{\boldsymbol{u}_{(0)},\cdots,\boldsymbol{u}_{(T)}\}} \Bigg[\sum_{t=1}^{T} \gamma^{t-1} \Bigg\{ c^i_{(t)} (\bar{x}^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)} (\bar{x}^i_{(t)} - x^i_{(t)})^- \\ + (\gamma^{\ell} h^i_{(t+\ell)} + \lambda_{(t)})(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \\ + \gamma^{\ell} p^i_{(t+\ell)} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^- - \frac{\lambda_{(t)}}{I} \big(\chi_{(t+\ell)} - \kappa \big) \Bigg\} \Bigg]. \quad (2.10)$$

we then have $\check{C}^{\boldsymbol{\lambda},\star}_{\kappa} = \sum_{i=1}^{I} \check{C}^{i,\boldsymbol{\lambda},\star}_{\kappa}$.

Note that we have replaced $Q^i_{(t)}$ and $\check{g}_{(t|\boldsymbol{x}_{(0)})}$ by their definitions in (2.10) to emphasize the fact that the Lagrange multiplier $\lambda_{(t)}$ acts as an 'effective increase' in the holding cost per unit in each period, representing the shadow price of the constraint. We denote the value functions in each period of these dynamic programs by $\check{V}^{i,\boldsymbol{\lambda}}_{\kappa,(t)}(\boldsymbol{x}_{(t)})$.

$(\check{\mathrm{DP}}^{\boldsymbol{\lambda}}_{\kappa})$ is decomposable into I separate dynamic programs with one-dimensional state space. Theorem 1 in Chapter 1 shows that, for any item i, the optimal policy for the dynamic program (2.10) is of a "double base stock" structure. Solving $(\check{\mathrm{DP}}_{\kappa})$ amounts to maximizing an unconstrained concave function with respect to $\boldsymbol{\lambda}$, which can be done by any steepest ascent method, or more efficient methods such as FISTA, see Beck and Teboulle (2009). Lemma 1 in Chapter 1 shows that the overall complexity to achieve an ϵ-optimal solution for any $\epsilon > 0$ is $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$.

Transforming $\check{\rho}^\star_\kappa$ into a feasible policy $\boldsymbol{\pi}$

There is no guarantee that the optimal double base stock policy $\check{\boldsymbol{\rho}}^\star_\kappa$ for $(\check{\mathrm{DP}}_\kappa)$ will satisfy the *conditional* expected value constraints ($\ddot{7}$). We therefore transform this policy $\check{\boldsymbol{\rho}}^\star_\kappa$ into a policy $\boldsymbol{\pi}$ that satisfies ($\ddot{7}$) by reducing the inventory positions prescribed by $\check{\boldsymbol{\rho}}^\star_\kappa$, which, by applying Lemma 4(b) to every starting state in every period, also satisfies (2.2). We construct $\boldsymbol{\pi}$ as follows:

(a) At the start of every period, check whether ordering according to $\check{\boldsymbol{\rho}}^\star_{\kappa,(t)}$ would satisfy the constraints ($\ddot{7}$) in dynamic program $(\ddot{\mathrm{DP}})$. If yes, implement the policy without modification.

(b) If not, implement the following modified policy. To simply our notation, we make the dependence of $\check{\rho}^{\star,i}_{\kappa,(t)}(x^i_{(t)})$ on $x^i_{(t)}$ implicit in what follows, and write $\check{\rho}^{\star,i}_{\kappa,(t)}$. Let

$$\text{(P)}\quad \boldsymbol{\pi}_{(t)}(\boldsymbol{x}_{(t)}) = \operatorname*{argmin}_{\bar{\boldsymbol{\pi}}} \sum_{i=1}^{I} \Big\{ c^i_{(t)}(\bar{\pi}^i - x^i_{(t)})^+ + d^i_{(t)}(\bar{\pi}^i - x^i_{(t)})^- + Q^i_{(t)}(\bar{\pi}^i) + \delta\mathbb{E}[\check{V}^{i,\boldsymbol{\lambda}^\star}_{\kappa,(t+1)}(\bar{\pi}^i - u^i_{(t)})|x^i_{(t)}] \Big\} \qquad (2.11a)$$

$$\text{s.t. } \min\{\check{\boldsymbol{\rho}}^\star_{\kappa,(t)}, 0\} \le \bar{\boldsymbol{\pi}} \le \check{\boldsymbol{\rho}}^\star_{\kappa,(t)} \qquad (2.11b)$$

$$\ddot{g}_{(t|\boldsymbol{x}_{(t)})}(\bar{\boldsymbol{\pi}}) \le \chi_{(t+\ell)} - \kappa_0. \qquad (\ddot{7})$$

This mathematical program is almost identical to that appearing in the definition of $(\ddot{\mathrm{DP}})$ in periods $t, t+1, \cdots, T$, with three important differences. First, instead of using $\ddot{V}$ (which we cannot compute) in the objective (2.11a), we use $\check{V}$, which we compute in the process of determining the double base stock policy $\check{\boldsymbol{\rho}}^\star_\kappa$. Second, the first part of inequality (2.11b) ensures $\boldsymbol{\pi}$ does not create 'artificial backlogs', by insisting that $\bar{\pi}^i$ can only be negative if $\check{\rho}^{\star,i}_{\kappa,(t)}$ is negative as well (i.e. if the backlog already exists). Third, the second part of inequality (2.11b) ensures that $\boldsymbol{\pi}$ *reduces* the ordering quantity prescribed by $\check{\boldsymbol{\rho}}^\star_\kappa$, which will be essential in proving the cost-impact of moving to this policy is minimal.

Note that for sufficiently large I, $\bar{\boldsymbol{\pi}} = \min(\check{\boldsymbol{\rho}}^\star_{\kappa,(t)}, 0)$ is a feasible solution, because for that choice of $\bar{\boldsymbol{\pi}}$, $\ddot{g}_{(t|\boldsymbol{x}_{(t)})}(\bar{\boldsymbol{\pi}}) = 0$, and by Assumption 5, $\chi_{(t+\ell)}$ grows at least linearly in the number of items, so for I sufficiently large, $\chi_{(t+\ell)} - \kappa_0 \ge 0$.

Thus, in any given period t, determination of $\bar{\boldsymbol{x}}$ amounts to minimizing a separable convex function subject to a *single* convex constraint. This mathematical program can be easily solved by dualization of this constraint (resulting in an unconstrained separable convex function with only simple bounds on the variables) where the optimal Lagrange multiplier can be found via simple bisection. To prove the main theorem in this section, we will need one final and innocuous assumption.

Assumption 6. *There exists a constant X (independent of I) such that all initial inventory positions $x^i_{(0)} \leq X$ for all $i = 1, \dots, I$.*

Theorem 5. *The policy $\boldsymbol{\pi}$ is feasible in the original (DP), under the chance constraint* (2.2)*. Furthermore, if the capacity constriction κ in* ($\check{8}$) *is set to $\kappa = \kappa_0 + \sqrt{2}\bar{\sigma} m T \sqrt{I \log I}$ the cost-impact of moving from the optimal policy $\check{\boldsymbol{\rho}}^\star_\kappa$ to $\boldsymbol{\pi}$ is bounded by an $O(1)$ term.*

Proof. Note that $\boldsymbol{\pi}$ satisfies ($\ddot{7}$) by construction. The feasibility of $\boldsymbol{\pi}$ is therefore established by applying Lemma 4(b) to every starting state in every period. We prove the second part of the theorem in Appendix B.1. □

Complexity of computing policy $\boldsymbol{\pi}$

As mentioned earlier, computation of any ϵ-approximation of $\check{\boldsymbol{\rho}}^\star$ requires $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$ number of computations. The remaining computational effort is in transforming $\check{\boldsymbol{\rho}}^\star$ into $\boldsymbol{\pi}$, which in every period, requires the optimization of (P) above. Once again, this is best done by Lagrangian relaxation of the single (convex) constraint ($\ddot{7}$) and optimizing over the single associated Lagrange multiplier. For a given Lagrange multiplier, the problem reduces to minimizing a separable convex function subject to simple bounds on the variables which can be done in $O(I)$ time. Given a constant upper bound on the Lagrange multiplier (see the proof of Theorem 6 below), the number of Lagrange multipliers that need to be evaluated is $O(1)$ overall using, e.g., a bisection method. This implies the complexity of transforming $\check{\boldsymbol{\rho}}^\star$ into $\boldsymbol{\pi}$, is $O(IT)$ and the overall complexity of the heuristic is $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$.

Bounding $\check{C}^{\star}_{\kappa} - C^{\star}$

To complete our analysis, we show that the optimal value of $(\check{\mathrm{DP}})$ exceeds the optimal value $C^{\star}$ by a term that grows with I as κ does. Taking the value of κ as suggested in Theorem 5, we conclude that the gap is at most $O(\sqrt{I \log I})$, which concludes our proof of asymptotic optimality.

We first establish Lemma 5, which is a partial converse of Lemma 4, in that it shows that if a policy satisfies the conditional chance constraint, it also satisfies a conditional expected value constraint for a slightly expanded capacity.

Lemma 5. *If under a given strategy, chance constraints* (2.2) *are satisfied, then*

$$\mathbb{E}\Big[\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \Big| \boldsymbol{x}_{(t)}\Big] \leq \chi_{(t+\ell)} + \alpha, \qquad t = 0, \dots, T-\ell, \tag{2.12}$$

where

$$\alpha = \frac{\beta\sigma}{1-\beta}\left(\sqrt{-2\ln(\frac{\beta}{2})} + \sqrt{\frac{\pi}{2}}\right), \tag{2.13}$$

and σ is defined the same as in Lemma 4.

Proof. To keep our notation clearer, we will omit all time subscripts in this proof. We begin by denoting $Y = \sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$. In the notation of Lemma 4, $Y = f(\boldsymbol{z})$. For ease of notation, we write $\mathbb{E}[Y|\boldsymbol{x}_{(t)}]$ as $\mathbb{E}[Y]$. We denote the conditional c.d.f. of Y by $F(y) \equiv \mathbb{P}(Y \leq y|\boldsymbol{x}_{(t)})$, and the conditional complementary c.d.f. of Y by $\bar{F}(y) \equiv 1 - F(y)$.

Since the chance constraints (2.2) are satisfied, $\mathbb{P}(Y \geq \chi|\boldsymbol{x}_{(t)}) = \bar{F}(\chi) \leq \beta$. Furthermore, the results of Lemma 4(a) imply that

$$\mathbb{P}(Y \geq \mathbb{E}[Y] + w|\boldsymbol{x}_{(t)}) = \bar{F}(\mathbb{E}[Y] + w) \leq 2e^{-\frac{w^2}{2\sigma^2}} \tag{2.14}$$

Equating the right hand side of (2.14) to β, we get that $\bar{F}(s) \leq \beta$ for any $s \geq \bar{s} := \mathbb{E}[Y] + \sigma\sqrt{-2\ln(\frac{\beta}{2})}$. We now have two numbers – $\bar{s}$ and χ – beyond which we can guarantee $\bar{F}(s)$ will be less than or equal to β.

Note that $\bar{s}$ might be larger or smaller than χ. If $\bar{s} \leq \chi$, then $\mathbb{E}[Y] + \sigma\sqrt{-2\ln(\frac{\beta}{2})} \leq \chi \Rightarrow \mathbb{E}[Y] \leq \chi \Rightarrow$

$\mathbb{E}[Y] \leq \chi + \alpha$, trivially proving our lemma.

Let us now consider the case in which $\bar{s} \geq \chi$. Since Y is a non-negative random variable, we have

$$
\begin{aligned}
\mathbb{E}[Y] = \mathbb{E}\Big[\int_0^\infty \mathbf{1}\{Y \geq s\}\, \mathrm{d}s \Big| \boldsymbol{x}_{(t)}\Big] &= \int_0^\infty \bar{F}(s)ds \\
&= \int_0^\chi \bar{F}(s)ds + \int_\chi^{\bar{s}} \bar{F}(s)ds + \int_{\bar{s}}^\infty \bar{F}(s)ds
\end{aligned}
$$

Since $\bar{F}(s) \leq \beta$ for $s \geq \chi$, we get

$$
\begin{aligned}
\mathbb{E}[Y] &\leq \chi + \beta(\bar{s} - \chi) + \int_{\bar{s}}^\infty \bar{F}(s)ds \\
&= (1-\beta)\chi + \beta\mathbb{E}[Y] + \beta\sigma\sqrt{-2\ln(\frac{\beta}{2})} + \int_{\bar{s}}^\infty \bar{F}(s)ds
\end{aligned}
$$

Now, using (2.14) and changing the integration variable from s to $w = s - \bar{s}$, we have

$$
\begin{aligned}
\int_{\bar{s}}^\infty \bar{F}(s)ds &\leq \int_{\bar{s}}^\infty e^{-(s-\mathbb{E}[Y])^2/2\sigma^2} ds \\
&= \int_0^\infty e^{-(w+\bar{s}-\mathbb{E}[Y])^2/2\sigma^2} dw \\
&\leq e^{-(\bar{s}-\mathbb{E}[Y])^2/2\sigma^2} \int_0^\infty e^{-\frac{w^2}{2\sigma^2}} dw \\
&= \beta\sigma\sqrt{\frac{\pi}{2}},
\end{aligned}
$$

where the second inequality follows from the fact that $(\bar{s}-\mathbb{E}[Y]+w)^2 = (\bar{s}-\mathbb{E}[Y])^2+w^2+2w(\bar{s}-\mathbb{E}_{(t)}[Y]) \geq (\bar{s} - \mathbb{E}[Y])^2 + w^2$, since $w \geq 0$ on the entire integration interval, while $\bar{s} \geq \mathbb{E}[Y]$.

Collecting all terms and re-arranging, we obtain

$$
\mathbb{E}[Y] \leq \chi + \frac{\beta\sigma}{1-\beta}\Big(\sqrt{-2\ln(\frac{\beta}{2})} + \sqrt{\frac{\pi}{2}}\Big).
$$

□

Theorem 6. *Let* α *be defined in* (2.13)*, then* $\check{C}^\star_\kappa - C^\star \leq T(\bar{p} + \bar{c})(\alpha + \kappa)$.

Proof. Let $\check{C}^\star_{-\alpha}$ be the optimal value of the original problem with constraints (2.2) replaced by *unconditional* expected value constraints

$$\mathbb{E}\Big[\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \Big| \boldsymbol{x}_{(0)}\Big] \le \chi_{(t+\ell)} + \alpha$$

(or equivalently, the dynamic program ($\check{\mathrm{DP}}_\kappa$) with κ replaced by $-\alpha$).

These constraints are relaxation of the *conditional* expected value constraints (2.12), which by Lemma 5 are relaxation of the original *conditional* chance constraints (2.2). Thus, $C^\star \ge \check{C}^\star_{-\alpha} \ge \check{C}^\star_\kappa - \Big(\sum_{t=1}^{T}\lambda^\star_{\kappa,(t)}\Big)(\alpha + \kappa)$, where the last inequality follows from the fact $\check{C}^\star_\kappa$ is convex in κ (see (2.9) and the definition of $\check{C}^\star_\kappa$ that follows).

Finally, we note that $\lambda^\star_{\kappa,(t)} \le (\bar{p} + \bar{c})$ because the Lagrange multiplier in each period denotes the amount by which $\check{C}^\star_\kappa$ decreases if only the t-th period capacity level $(\chi_{(t+\ell)} - \kappa)$ were increased by one unit, and all other capacity levels remain unchanged. This capacity increment may allow us to eliminate one unit of backlog in the t-th or a future period, the value of which is bounded by $\bar{p}$. Additionally it may allow us to switch procurement time for a total of one unit, a benefit that is bounded by $\bar{c} - \underline{c} \le \bar{c}$. Thus, we conclude that

$$C^\star \ge \check{C}^\star_{-\alpha} \ge \check{C}^\star_\kappa - \left(\sum_{t=1}^{T}\lambda^\star_{\kappa,(t)}\right)(\alpha + \kappa) \ge \check{C}^\star_\kappa - T(\bar{p} + \bar{c})(\alpha + \kappa),$$

which concludes our proof. $\square$

2.4 An Alternative Heuristic

In this section, we develop an alternative heuristic that is shown to have excellent empirical performance for moderate as well as large values of I. However, we are unable to establish that *this* heuristic is asymptotically optimal. The heuristic is designed in 2 phases. First we develop a *lower bound* problem, and compute its optimal strategy, again a double base stock policy. This policy is then transformed into a feasible policy whose cost is an *upper bound*.

In this section, we allow for a general correlation pattern among *all* products as opposed to the more limited correlation pattern prescribed by Assumption 4.

2.4.1 A lower bound

The lower bound is obtained via three relaxation steps. The first replaces the complex feasible action regions by polyhedral sets only slightly larger than the actual feasible region. This is obtained by replacing the single chance constraint (for every product and starting state) by a set of upper bounds on the aggregate inventory position for every subset of the full product sets.

Next, we argue that removing all but a relatively small subset $\mathcal{A}$ of these 2^I constraints does not result in any significant loss of accuracy. (Typically, $|\mathcal{A}| = O(I)$, occasionally we need $|\mathcal{A}| = O(I^2)$.) The final relaxation step is to dualize these constraints via Lagrangian relaxation, employing $O(|\mathcal{A}|T)$ Lagrange multipliers. The best lower bound of this type is obtained by maximizing over the full vector of Lagrange multipliers, i.e. by solving the full Lagrangian dual.

The strategy that optimizes the Lagrangian relaxation for any Lagrange multipliers prescribes a separate so called "double base stock" policy for every item, similar to the policy designed in Section 2.3. We refer to this optimal policy for the Lagrangian dual as the Lower Bound policy.

We first use the identity $\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \equiv \max_{\boldsymbol{a}\in\{0,1\}^I} \sum_{i=1}^{I} a^i(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})$ to write the capacity constraint (2.2) as $\mathbb{P}\left(\max_{\boldsymbol{a}\in\{0,1\}^I} \sum_{i=1}^{I} a^i(\bar{x}^i_{(t)} - \dot{u}^i_{(t)}) \geq \chi_{(t+\ell)} \middle| \boldsymbol{x}_{(t)}\right) \leq \beta$. As a first relaxation, we replace this constraint by

$$\mathbb{P}\left(\sum_{i=1}^{I} a^i(\bar{x}^i_{(t)} - \dot{u}^i_{(t)}) \geq \chi_{(t+\ell)} \middle| \boldsymbol{x}_{(t)}\right) \leq \beta, \qquad \forall \boldsymbol{a} \in [0,1]^I \tag{2.15}$$

This is a relaxation because the original constraint allows for the possibility that a *different* value of $\boldsymbol{a}$ achieves the maximum depending on the realization $\dot{\boldsymbol{u}}_{(t)}$. In contrast, the relaxed constraint (2.15) can only consider violations with the same value of $\boldsymbol{a}$ for every realization $\dot{\boldsymbol{u}}_{(t)}$.

It is important to note that, frequently, even when the relaxation is relatively loose in some states, both the optimal value of the dynamic program and the optimal strategy are hardly affected. This occurs when vectors of $\bar{\boldsymbol{x}}_{(t)}$ for which the relaxation is loose would result in sub-optimal solutions that are therefore never or rarely adopted by the optimal strategy for the dynamic program. One example of where this happens is when backlogging costs are high, reflecting a high service level. In these situations, vectors of $\bar{\boldsymbol{x}}_{(t)}$ that result in significant backlogs are very unlikely to be chosen by an optimal strategy. All violations at state vectors $\bar{\boldsymbol{x}}_{(t)}$ that are close to the optimal trajectory are likely to occur when all components of $\boldsymbol{a} = \mathbf{1}$, where the original and relaxed constraints agree.

Since the demands $\dot{\boldsymbol{u}}_{(t)}$ follow a multivariate Normal distribution with mean $\dot{\boldsymbol{\mu}}_{(t)}$ and variance-covariance matrix $\dot{\Sigma}_{(t)}$, these constraints are equivalent to simple linear constraints:

$$\boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t)} \leq C^{\boldsymbol{a}}_{(t)} \equiv \chi_{(t)} + (\boldsymbol{a} \cdot \dot{\boldsymbol{\mu}}_{(t)}) - \Phi^{-1}(1-\beta)\sqrt{\boldsymbol{a}^\top \dot{\Sigma}_{(t)} \boldsymbol{a}} \qquad \forall \boldsymbol{a} \in \{0,1\}^I \tag{2.16}$$

Our first relaxation has therefore replaced our complex chance constraint with a set of simple linear constraints (albeit an exponential number of them).

As a second relaxation, we take the exponential set of constraints in (2.16), and choose only a subset $\mathcal{A} \subset \{0,1\}^I$ of constraints to consider. The final two relaxation steps thus replace the exact DP by

$$(\tilde{\text{DP}}) \quad V_{(t)}(\boldsymbol{x}_{(t)}) = \min_{\bar{\boldsymbol{x}}_{(t)}} \Big\{ \gamma \mathbb{E}\left[V_{(t+1)}(\bar{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}) \right] + \sum_{i=1}^{I} \Big(c^i_{(t)} (\bar{x}^i_{(t)} - x^i_{(t)})^+$$
$$+ d^i_{(t)} \left[(\bar{x}^i_{(t)} - x^i_{(t)})^- \right] + Q^i_{(t)}(\bar{x}^i_{(t)}) \Big) \Big\} \tag{2.17a}$$
$$\text{s.t.} \quad \boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t)} \leq C^{\boldsymbol{a}}_{(t)} \qquad \forall \boldsymbol{a} \in \mathcal{A} \tag{2.17b}$$

Selecting a Set of Constraints $\mathcal{A}$

In the following, we discuss which subset $\mathcal{A}$ of constraints among the exponentially large set $\{0,1\}^I$ we choose to include in our dynamic program.

We first note that in instances in which products are hardly ever back-ordered, the constraint in which every product is included (i.e., in which $\boldsymbol{a}$ is a vector of ones) is the only one we would require. Indeed, with no backorders, requiring the sum of all inventory positions to be bounded by $C^{\boldsymbol{a}}_{(t)}$, see (2.17b), is the strictest of all the constraints in (2.16), and the remaining constraints are redundant. Extending this logic further, we only expect a constraint $\boldsymbol{a} \in \mathcal{A}$ to be tight if there is a significant part of the sample space in which every product for which $a^i = 0$ is backordered at the end of the lead time. Our method partly relies on finding combinations of products for which this is likely, and constructing the corresponding $\boldsymbol{a}$ vectors.

Second, the constraint set $\mathcal{A}$ can be adjusted iteratively. Given a particular solution that is optimal for a version of the problem in which certain constraints have been omitted, we can easily check whether those omitted constraints are violated by simulating the solutions in the system, and observing the frequency of violations. We can then add *only* those constraints that are violated to the dynamic program to improve the solution.

Our solution makes use of these two observations to proceed iteratively as follows. We begin by including *only* the constraint with $\boldsymbol{a} = \boldsymbol{e}$, the vector of all ones, and solve the Lagrangian dual (2.20) to optimality. Next, we investigate whether the constraints corresponding to vectors $\boldsymbol{a}$ in the following three categories are violated:

- All vectors $\boldsymbol{a} = e^{-i}, i = 1, \ldots, I$, with $e^{-i} = e - e^{i}$, where e^{i} denotes the ith basis vector, i.e. consider sets obtained by omitting only one item from the complete product family. Since the system is designed to avoid stockouts with a high likelihood, many stockout scenarios involve a single back-ordered item.
- However, some scenarios do involve multiple items to run out of stock. In such scenarios, we are likely to designate the items with low backlogging cost rates, as the "run out" items. In each period, we therefore index the items in descending order of their backlogging cost rates, and consider the nested collection of sets of the form $A_j = \{1, 2, \ldots, j\}$. These constraints are constructed to ensure that the item with the *smallest* backorder cost rate (i.e., the item most likely to be backordered) is removed first, followed by the next cheapest item to backorder, and so on. Referring to the discussion above, these constraints prevent the dynamic program from selecting a solution in which items that are relatively cheap to backorder are indeed targeted for backordering to reserve more capacity for other items.
- Construct a collection of K clusters of items that are highly positively correlated together. These groups are constructed using K-means clustering or may be self-evident; in a storage facility for consumer electronics, consider all TV sets, cellular phones etc as such clusters. Within a cluster, if one of the items runs out, it is likely that the other items do as well. We therefore consider the K vectors $\boldsymbol{a}$ in which all components are 1 except for the items belonging to one of the clusters. The constraints reflect the reality that highly correlated items are likely to be backordered together.

Having determined what constraints in this $(T(2I+K))$-set of constraints are violated, we add these constraints into our dynamic program, and re-solve the corresponding Lagrangian dual to optimality. The number of constraints we need to check, and add to our dynamic program thus scales linearly with the number of products handled.

Finally, we evaluate the cost of the resulting upper bound (see Section 2.4.2) and evaluate the gap between the upper and lower bound. If this gap is greater than 5%, we repeat the process with the $I(I-1)$ constraints in which *two* of the entries in $\boldsymbol{a}$ are 0, and add those constraints that are violated to our dynamic program.

Lagrangian Relaxation

To decompose the dynamic program into I separate DPs with a one-dimensional state space, (again, as in Section 2.3) we use a Lagrangian relaxation to move each of the constraints in (2.17b) to the *objective* of our dynamic program. This effectively allows solutions in which these constraints are violated, but penalizes such violations at a rate $\lambda_{(t)}^{\boldsymbol{a}}$, the Lagrangian multiplier associated with that specific constraint:

$$V_{(t)}^{\lambda}(\boldsymbol{x}_{(t)}) = \min_{\min\{\boldsymbol{x}_{(t)},0\} \le \bar{\boldsymbol{x}}_{(t)}} \Bigg\{ \gamma \mathbb{E}\left[V_{(t+1)}^{\lambda}(\bar{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)})\right] + \sum_{i=1}^{I} \Big(c_{(t)}^{i}(\bar{x}_{(t)}^{i} - x_{(t)}^{i})^{+} \\ + d_{(t)}^{i}(\bar{x}_{(t)}^{i} - x_{(t)}^{i})^{-} + Q_{(t)}^{i}(\bar{x}_{(t)}^{i}) \Big) + \sum_{\boldsymbol{a}\in\mathcal{A}} \lambda_{(t)}^{\boldsymbol{a}} \left[\boldsymbol{a}\cdot\bar{\boldsymbol{x}}_{(t)} - C_{(t)}^{\boldsymbol{a}}\right] \Bigg\} \tag{2.18}$$

This dynamic program is readily seen to decompose into I separate, single-product dynamic programs, $V_{(t)}^{\lambda}(\boldsymbol{x}_{(t)}) = \left(\sum_{i=1}^{I} V_{(t)}^{i,\lambda}(x_{(t)}^{i})\right) - \left(\sum_{\boldsymbol{a}\in\mathcal{A}} \lambda_{(t)}^{\boldsymbol{a}} C_{(t)}^{\boldsymbol{a}}\right)$, where

$$V_{(t)}^{i,\lambda}(x_{(t)}^{i}) = \min_{\min\{x_{(t)}^{i},0\} \le \bar{x}_{(t)}^{i}} \Bigg\{ \gamma \mathbb{E}\left[V_{(t+1)}^{i,\lambda}(\bar{x}_{(t)}^{i} - u_{(t)}^{i})\right] + c_{(t)}^{i}(\bar{x}_{(t)}^{i} - x_{(t)}^{i})^{+} + d_{(t)}^{i}\left[(\bar{x}_{(t)}^{i} - x_{(t)}^{i})^{-}\right] \\ + Q_{(t)}^{i}(\bar{x}_{(t)}^{i}) + \sum_{\boldsymbol{a}\in\mathcal{A}} \lambda_{(t)}^{\boldsymbol{a}} a^{i} \bar{x}_{(t)}^{i} \Bigg\} \tag{2.19}$$

Furthermore, the optimal solution to each of these dynamic programs takes on a double base stock structure, as described in the following theorem, the proof of which is analogous to that of Theorem 1(a).

Theorem 7. *For any $\boldsymbol{\lambda} \ge \boldsymbol{0}$ and $i = 1, \dots, I$, the value functions $V_{(t)}^{i,\lambda}(\cdot)$ are convex, differentiable* *everywhere, except for at most two points,* $\lim_{x \uparrow \infty} \frac{d}{dx} V^{i,\lambda}_{(t)}(x) \geq 0$*, and* $\frac{d}{dx} V^{i,\lambda}_{(t)} = O(|x|^r)$*. The optimal policy* $\pi^{i,\lambda}_{(t)}$ *to dynamic program (2.19) follows a double base stock structure defined by two parameters* $-\infty \leq S^{i,\lambda}_{(t)} < B^{i,\lambda}_{(t)} \leq \infty$ *in each period t, with* $B^{i,\lambda}_{(t)} \geq 0$*. If the starting inventory position is below* $S^{i,\lambda}_{(t)}$*, an order should be placed to bring it up to* $S^{i,\lambda}_{(t)}$*, and if the starting inventory position is above* $B^{i,\lambda}_{(t)}$ *the salvage option should be exercised to bring it down to* $B^{i,\lambda}_{(t)}$*. Otherwise, it is optimal to leave the starting position unchanged.*

Proof. See Appendix B.3. □

Our three relaxations have therefore reduced a high-dimensional dynamic program with a highly intractable feasible action space to a series of simple one-dimensional dynamic programs, in which policies of simple structure are optimal. As before, V^λ is a *lower bound* on the initial set of value functions V, for any vector of Lagrange multipliers $\boldsymbol{\lambda} \geq \mathbf{0}$:

For any $\lambda \in \mathbb{R}_+^{|\mathcal{A}| \times T}$ *and any starting state* $(\boldsymbol{x}_{(t)})$*,* $V^\lambda_{(t)}(\boldsymbol{x}_{(t)}) \leq V_{(t)}(\boldsymbol{x}_{(t)})$*.*

To find the *best* lower bound, we solve, again as in Section 2.3, the complete Lagrangian dual

$$max_{\boldsymbol{\lambda} \geq \mathbf{0}} V^\lambda_{(1)}(\boldsymbol{x}_{(1)}) \tag{2.20}$$

It is well known and verifiable with standard arguments that this optimization problem is concave in $\boldsymbol{\lambda}$. Various subgradient methods are therefore guaranteed to converge to its unique maximizer $\boldsymbol{\lambda}^*$. The next Proposition derives the supergradients.

Proposition 1. *For* $t^0 = 1, \ldots, T - \ell$ *and* $\boldsymbol{a} \in \mathcal{A}$*,* $\frac{\partial V^\lambda_{(1)}(\boldsymbol{x}_{(1)})}{\partial \lambda^{\boldsymbol{a}}_{(t^0)}} = \gamma^{t^0 - 1} \mathbb{E}[\boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t^0)} - C^{\boldsymbol{a}}_{(t^0)}]$*.*

Proof. See Appendix B.4. □

To solve this convex program, we have used the FISTA gradient method, see (Beck and Teboulle, 2009). This method is considerably faster than more established methods such as ISTA, both in terms of worst case complexity and in terms of average computational time.

Theorem 8. *The FISTA procedure to solve the Lagrangian Dual (2.20) to* ϵ*-optimality is* $O(\frac{I|\mathcal{A}|^2 T^{\frac{3}{2}}}{\sqrt{\epsilon}})$*.*

Proof. See Appendix B.5 □

Thus, for a given number of items, the complexity scales exceptionally well with the planning horizon T. Most of our suggestions for the choice of the set $\mathcal{A}$ have $|\mathcal{A}| = O(I)$, resulting in an overall complexity bound which is $O(I^3)$ for a fixed planning horizon T.

Additionally, since $V^{\lambda}_{(1)}(\boldsymbol{x_{(1)}})$ is the expected cost value of the aggregate of I independent DPs, the procedure lends itself naturally to distributed parallel computations.

Note that the above complexity analysis only gives a theoretical worst case bound. Our experiments reveal that the average computational time grows roughly linearly with I.

2.4.2 An upper bound: a heuristic policy

Theorem 7 reveals that the solution to the lower bound dynamic programs is characterized by two sets of critical threshold inventory positions $\{(S^i_{(t)}, B^i_{(t)}) : i = 1, \cdots, I; t = 1, \cdots, T\}$ for the policy corresponding to the optimal Lagrangian multiplier vector $\boldsymbol{\lambda}^*$. Due to relaxations of the constraints, this policy is, often, not feasible since the inventory may violate the capacity constraint (2.2). In each period, we check whether the ordering decision dictated by the policy satisfies the capacity constraint. If it does, we implement it. If it does not, our heuristic carries out, in up to two stages, an iterative sequence of reductions of the order sizes to reach feasibility.

The expected cost, under this heuristic policy, is, of course, an *upper bound* on the optimal cost value. Since we cannot evaluate this strategy's cost analytically, we resort to estimating its performance using Monte Carlo simulations. Extensive numerical studies reported compare the *lower bound*, resulting from the approximate DP in the previous section, with this heuristic policy.

More specifically, in any given period t, assume the (lower bound) DP under λ^* prescribes the adoption of the vector of inventory positions $\bar{\boldsymbol{x}}_{(t)}$. The heuristic starts by checking whether $\bar{\boldsymbol{x}}_{(t)}$ is feasible with respect to the storage capacity constraint (2.2). As mentioned, even this single feasibility test is analytically intractable, and so we do this by simulation. We generate M samples from the joint demand distribution of $\dot{\boldsymbol{u}}_{(t)}$, denoted $\{\dot{\boldsymbol{u}}_{(t),n}\}_{n=1}^{M}$. We then calculate the empirical frequency with which the capacity constraint is violated in period $t+\ell$. To ensure that the true chance constraints (2.2) are satisfied with a very high probability, we will accept a solution only if, based on the simulation sample, an overflow occurs in a frequency, even smaller than β, i.e., $\beta - \Delta\beta$, i.e., a solution $\bar{\boldsymbol{x}}$ is accepted only if the sampling frequency

$$S_M := \frac{1}{M} \sum_{n=1}^{M} \mathbf{1}\Big(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t),n})^+ \geq \chi_{(t+\ell)} \Big) \leq \beta - \Delta\beta.$$

By the Hoeffding inequality for the random variables on the interval $[0,1]$, this ensures that

$$\mathbb{E}[S_M] = \mathbb{P}\Big(\sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ \geq \chi_{(t+\ell)} \Big) \leq \beta$$

with probability $\geq 1 - e^{-2(\Delta\beta)^2 M}$.

Then, if the constraint, with β reduced to $(\beta - \Delta\beta)$, is satisfied, this vector is implemented. If it is not, we reduce the order sizes, in two stages, while ensuring minimal cost increases.

In the first stage, the feasible region is relaxed to a tree-structured polymatroidal region and the vector of (reduced) inventory positions that minimizes the aggregate of current as well as a lower bound of all future expected costs is computed. This stage is described in section 2.4.2.

If the resulting inventory position remains infeasible, we apply our second stage, comprised of a greedy-drop procedure designed as a scaling algorithm. We describe this second stage in the supplementary material.

The worst case complexity of the first stage procedure is $O(I \log I \log(\chi_{(t)} + \sum_i \mu^i_{(t)}))$, that of the second stage procedure is roughly quadratic in I. More precisely, the complexity bound of the second stage depends on M, and B, the ($\lfloor \beta M \rfloor + 1$)-st largest overflow for the initial vector of inventory positions: $O(I \log(\frac{B}{I})(I \log I + M \log M + IM))$. As shown in Section 2.5, the average complexity in our simulations grows linearly with I. Both stages of our algorithm make extensive use of the lower-bound value functions.

Upper Bound Stage 1: A Math Programming Based Inventory Reduction

Let $\bar{\boldsymbol{x}}_{(t)}$ denote the vector of inventory positions prescribed by the dynamic program under λ^*. We ideally, seek to find the vector of feasible inventory positions $\bar{\boldsymbol{y}}_{(t)} \leq \bar{\boldsymbol{x}}_{(t)}$ that minimizes total expected discounted costs in this and future periods, assuming all future decisions are made using the optimal policy for the original DP. This problem is intractable; we therefore choose the feasible vector that minimizes cost under the assumption that the future costs are approximated by the Lower Bound value functions instead, i.e. by solving the following optimization problem

$$\min \sum_{i=1}^{I} C_i(\bar{y}^i_{(t)}) \text{ s.t. } \bar{\boldsymbol{y}}_{(t)} \text{ satisfies (2.2)}, \bar{y}^i_{(t)} \leq \bar{x}^i_{(t)}, i = 1, \ldots, I \tag{2.21}$$

where $C_i(\bar{y}^i_{(t)}) = Q^i_{(t)}(\bar{y}^i_{(t)}) + c^i_{(t)}(\bar{y}^i_{(t)} - x^i_t)^+ + d^i_{(t)}(x^i_t - \bar{y}^i_{(t)})^+ + \gamma \mathbb{E}\left[V^{i,\lambda^*}_{(t+1)}(\bar{y}^i_{(t)} - u^i_{(t)})\right]$ is convex.

As explained, even testing for feasibility of a given vector $\bar{\boldsymbol{y}}_{(t)}$ is intractable and one needs to resort to Monte Carlo simulation or I-fold numerical integration. It is, a fortiori, prohibitively difficult to solve (2.21). Instead, we use the first relaxation in Section 2.4, to extend the feasible region to the polyhedron:

$$\boldsymbol{a} \cdot \bar{\boldsymbol{y}}_{(t)} \leq C^{\boldsymbol{a}}_{(t)} \qquad\qquad \forall \boldsymbol{a} \in \{0,1\}^I \tag{2.22}$$

defined in (2.16) and (2.17b). Each vector $\boldsymbol{a} \in \{0,1\}^I$ corresponds with a specific set of products; in what follows, we abuse this notation, slightly, by sometimes treating these vectors as sets. Thus, for example 'the set of products $\boldsymbol{a}$' refers to $\{i : a_i = 1\}$, $i \in \boldsymbol{b}$ means $b_i = 1$, $\boldsymbol{a} \subseteq \boldsymbol{b}$ means $a_i \leq b_i$ for all products i, and $\boldsymbol{a} \cap \boldsymbol{b} = \emptyset \Leftrightarrow \boldsymbol{a} \cdot \boldsymbol{b} = 0$.

The feasible region is described by an exponential number of constraints. If the polyhedron were a polymatroid, various efficient and polynomially bounded algorithms have nevertheless been proposed to minimize a separable convex objective subject to an exponential number of constraints, see e.g. Groenevelt (1991), Federgruen and Groenevelt (1986), and Fujishige (2005).

Unfortunately, the polyhedron fails to be a polymatroid, because $C^{\boldsymbol{a}}_{(t)}$ fails to be submodular as a function of $\boldsymbol{a}$, in general. Indeed, in the special case where the items have independent demands, $C^{\boldsymbol{a}}_{(t)}$ is, in fact, supermodular. Let $\boldsymbol{a}$ and $\boldsymbol{b}$ denote two item sets with $j \notin \boldsymbol{b}$ and $\boldsymbol{a} \subseteq \boldsymbol{b}$. Then, denoting the variance of each product i over a lead time by $\dot{V}^i_{(t)} \equiv (\dot{\Sigma}_{(t)})_{i,i}$, we have $C^{\boldsymbol{b}\cup\{j\}}_{(t)} - C^{\boldsymbol{b}}_{(t)} = \dot{\mu}^j_{(t)} - \Phi^{-1}(1 - \beta)\left[\sqrt{(\boldsymbol{b} \cdot \dot{\boldsymbol{V}}_{(t)}) + \dot{V}^j_{(t)}} - \sqrt{\boldsymbol{b} \cdot \dot{\boldsymbol{V}}_{(t)}}\right] \geq \dot{\mu}^j_{(t)} - \Phi^{-1}(1-\beta)\left[\sqrt{(\boldsymbol{a} \cdot \dot{\boldsymbol{V}}_{(t)}) + \dot{V}^j_{(t)}} - \sqrt{\boldsymbol{a} \cdot \dot{\boldsymbol{V}}_{(t)}}\right] = C^{\boldsymbol{a}\cup\{j\}}_{(t)} - C^{\boldsymbol{a}}_{(t)}$, where the inequality follows from the concavity of the square root function, $\boldsymbol{b} \cdot \boldsymbol{V}_{(t)} > \boldsymbol{a} \cdot \boldsymbol{V}_{(t)}$, and $\Phi^{-1}(1 - \beta) \geq 0$, whenever $\beta \leq 0.5$.

We therefore expand the polyhedron to a polymatroid by selecting only a subset of the $(2^I - 1)$ constraints in (2.22). More specifically, we select a tree-structured collection of sets $\mathcal{T} \subseteq \{0,1\}^I$, i.e., for all $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{T}$ with $\boldsymbol{a} \cap \boldsymbol{b} \neq \emptyset$ either $\boldsymbol{a} \subseteq \boldsymbol{b}$ or $\boldsymbol{b} \subseteq \boldsymbol{a}$. This means that the sets in $\mathcal{T}$ can be represented as the nodes in a tree network with two sets (nodes) $\boldsymbol{a}$ and $\boldsymbol{b}$ linked if $\boldsymbol{a} \subseteq \boldsymbol{b}$.

A maximally-sized tree-structured collection of sets $\mathcal{T} \subseteq \{0,1\}^I$ is most easily selected when I is a power-of-two, say $I = 2^n, n \geq 1$. One possible choice for $\mathcal{T}$ consist of all singletons in $\{1, \cdots, I\}$, all consecutive pairs $\{1,2\}, \{3,4\}, \{5,6\}, \ldots$, all consecutive quadruples $\{1,2,3,4\}$, $\{5,6,7,8\}, \ldots$, along with all remaining partitions of $\{1, \cdots, I\}$ into consecutive sets with a power-of-two cardinality, terminating with the full set $\{1, \cdots, I\}$ as the final set in $\mathcal{T}$. Clearly $|\mathcal{T}| = I(1+\frac{1}{2}+\ldots+(\frac{1}{2})^n) = \frac{I(1-(\frac{1}{2})^{n+1})}{1-\frac{1}{2}} = 2I-1$. For a general value of I, the tree network is constructed as follows: start with an initial layer of nodes, each representing one of the singletons in $\{1, \ldots, I\}$. In moving from a layer to the next one, combine pairs of nodes and merge the associated item sets into one. When a layer has an odd number of nodes, the last node is passed on, by itself, to the next layer.

With $\mathcal{T}$ tree-structured, the polyhedron $\left\{\bar{\boldsymbol{y}}_{(t)} : \boldsymbol{a} \cdot \bar{\boldsymbol{y}}_{(t)} \leq C^{\boldsymbol{a}}_{(t)} \;\; \forall \boldsymbol{a} \in \mathcal{T}\right\}$ is a polymatroid, irrespective of the right hand side function, see Lawler and Martel (1982). Thus, the optimization problem

$$\min \sum_{i=i}^{I} C_i(\bar{y}^i_{(t)}) \text{ s.t. } \boldsymbol{a} \cdot \bar{\boldsymbol{y}}_{(t)} \leq C^{\boldsymbol{a}}_{(t)} \; \forall \boldsymbol{a} \in \mathcal{T}, \bar{y}^i_{(t)} \leq \bar{x}^i_{(t)} \; \forall i = 1, \cdots, I \tag{2.23}$$

can be solved using the following Bottom-Up Algorithm; see Theorem 3 in Groenevelt (1991).

Bottom-Up Algorithm: number the sets in $\mathcal{T} = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_K\}$, with $K \equiv |\mathcal{T}|$, such that $C^{\boldsymbol{a}_1}_{(t)} \leq C^{\boldsymbol{a}_2}_{(t)} \leq \cdots \leq C^{\boldsymbol{a}_K}_{(t)}$. Define $\boldsymbol{y}(r)$ recursively by and $\boldsymbol{y}(r) = \min_{\boldsymbol{y} \geq \boldsymbol{y}(r-1)} \sum_{i=1}^{I} C_i(\bar{y}^i) \;$ s.t. $\; \boldsymbol{a}_r \cdot \boldsymbol{y} \leq C^{\boldsymbol{a}_r}_{(t)}, \;\; \boldsymbol{y}(0) = \bar{\boldsymbol{x}}_{(t)}$. The optimal solution to (2.23) is simply $\boldsymbol{y}(K)$.

This algorithm is simple: it decouples the full polymatroid optimization problem with K constraints into K separate problems, each with a single budget constraint and simple lower bounds for the variables. In each of its iterations, the prior solution is used as the lower bound.

If the K optimization problems are solved as continuous optimization problems, this algorithm has complexity $O(I^2 \log I + I\xi)$ with ξ the time needed to find the root of a single variable increasing function, employing, for example the method of Zipkin (1980). This complexity bound follows from the fact that the Algorithm requires $K < 2I$ iterations, and each iteration takes $O(I \log I + \xi)$ time. If the single constraint problems are solved as a discrete optimization problem with Hochbaum's (1982) method, an alternative pseudo-polynomial bound of $O(I \log I \log C_{(t)}^{\boldsymbol{a}_K})$ arises.

Upper Bound Stage 2: the Greedy-Drop Procedure

If the optimal solution to the optimization problem (2.23) that results from stage 1 of our heuristic is feasible, i.e. satisfies the exact storage constraint (2.2), our heuristic implements this solution directly. If not, we further reduce the inventory positions with minimum expected cost increases, using a greedy procedure.

The basic idea behind this greedy procedure is simple; in each step, we choose a step size s and estimate the cost impact of reducing each of our I products by this step size, again assuming optimal future costs are approximated by the Lower Bound value functions:

$$\begin{aligned}\Delta V_{(t)}^{i} = Q_{(t)}^{i}(\bar{x}_{(t)}^{i} - s) - Q_{(t)}^{i}(\bar{x}_{(t)}^{i}) - c_{(t)}^{i} \min\{s, (\bar{x}_{(t)}^{i} - x_{(t)}^{i})^{+}\} + d_{(t)}^{i} \min\{s, (x_{(t)}^{i} - \bar{x}_{(t)}^{i} + s)^{+}\} \\ + \gamma \mathbb{E}\left[V_{(t+1)}^{i}(\bar{x}_{(t)}^{i} - s)\right] - \gamma \mathbb{E}\left[V_{(t+1)}^{i}(\bar{x}_{(t)}^{i})\right] - s \sum_{\boldsymbol{a} \in \mathcal{A}} \lambda_{(t)}^{\boldsymbol{a}} a^{i}\end{aligned}$$

We then choose the item with the *lowest* such cost and reduce it by the step-size s. If this reduction fails to make the allocation feasible, we re-calculate $\Delta V_{(t)}^{i}$ for this item, and repeat.

Unfortunately, this greedy procedure can be slow. Indeed, if the initial allocation is highly infeasible, the algorithm may need to cycle through many reductions to achieve feasibility, and each of these steps requires a full simulation to evaluate whether the capacity constraint is satisfied. In particular, the basic greedy procedure fails to be of polynomial complexity.

The final algorithm we propose below improves this simple greedy heuristic in a number of ways. (1) We use an adaptive step size directly related to the extent by which the incumbent solution violates the feasibility constraint – the larger the violation, the larger the step size. (2) We include an acceleration mechanism that, at every step, checks whether the simple unilateral reduction of a single item results in enough of an improvement. If it does not, the acceleration step uses linear knapsack problems to reduce allocations to multiple items at once. (3) Instead of running a simulation at every step of the algorithm to determine whether the new allocation is feasible (which requires generating sample paths of demand and checking every path to determine whether the allocation would result in a violation), we generate sample paths upfront and at each step only check those paths that might go from infeasible to feasible as a result of that change.

This procedure's complexity grows roughly quadratically with the number of items, making our complete algorithm highly scalable to large size instances, in particular since the average complexity grows linearly with the number of items.

We now describe this algorithm in detail.

Step 1 Initialize $\bar{\boldsymbol{y}}_{(t)}$ with the result of the Stage 1 optimization procedure. Throughout our description of the greedy-drop procedure, we shall let $\bar{\boldsymbol{y}}_{(t)}$ denote the current solution.

Step 2 Generate M demand sample paths. Let $u^i_{(t),m}$ denote the one-period demand for product i in period t on sample path m. As before, we let $\dot{u}^i_{(t),m} = \sum_{\tau=t}^{t+\ell-1} u^i_{(t),m}$ denote the demand for product i during a lead time following period t on sample path m.

On a given sample path m, the total on-hand inventory in time period $t+\ell$ is $\sum_{i=1}^{I}(\bar{y}^i_{(t)} - \dot{u}^i_{(t),m})^+$. Let $\mathcal{M}^+$ be the set of sample paths which violate the constraint:
$\mathcal{M}^+ = \left\{m : \sum_{i=1}^{I}(\bar{y}^i_{(t)} - \dot{u}^i_{(t),m})^+ > \chi\right\}$. If $|\mathcal{M}^+| \le \lfloor \beta M \rfloor$, then the proportion of sample paths with (and therefore the simulated probability of) an overflow is less than β, and the chance constraint is satisfied. Stop.

If not, the remaining steps of the algorithm will successively reduce $\bar{\boldsymbol{y}}_{(t)}$ to remove sample paths from $\mathcal{M}^+$. As we reduce $\bar{\boldsymbol{y}}_{(t)}$, sample paths will never *enter* $\mathcal{M}^+$ – they can only leave it. Instead of checking for overflows on all M sample paths, therefore, it will be sufficient to check for overflow on $|\mathcal{M}^+|$ sample paths only – this will result in considerable improvements in the speed of our algorithm.

Step 3 Set a step size proportional to the size of the overflow. This will of course be different for every sample path in $\mathcal{M}$, but since our aim is to reduce the number of paths with overflows to $\lfloor \beta M \rfloor$, the sample path with the $(\lfloor \beta M \rfloor + 1)$th largest overflow provides a good indication of the appropriate step size. Let this sample path be m_o, and denote the corresponding overflow B

$$B \equiv \left(\sum_{i=1}^{I} (\bar{y}^i_{(t)} - \dot{u}^i_{(t),m_o})^+ \right) - \chi \tag{2.24}$$

We set

$$s = \max \left\{ \epsilon, \frac{B}{I} \right\} \tag{2.25}$$

thus ensuring a minimum step size ϵ, e.g. $\epsilon = 0.1$.

We next reduce the inventory position of the item with the lowest value of $\Delta V^i_{(t)}$ by s; this is the *cheapest* item to reduce. However, to ensure sufficient progress towards the goal of reaching feasibility, we only consider items for which a reduction of its inventory positions by s units would result in a reduction of the $(\lfloor \beta M \rfloor + 1)$th overflow of at least αs, for some $0 < \alpha \leq 1$. (For example, an inventory reduction of an item with projected backlogs in all or the vast majority of scenarios does not alleviate the overflow frequency) The set of items that satisfy this requirement is given by:

$$\mathcal{I}^* = \left\{ i : \left| \left\{ m \in \mathcal{M} : \sum_{\iota=1}^{I} (\bar{y}^\iota_{(t)} - s\mathbf{1}_{\iota=i} - \dot{u}^\iota_{(t),m})^+ - \chi > (B - \alpha s)^+ \right\} \right| \leq \lfloor \beta M \rfloor \right\}$$

If $\mathcal{I}^*$ is empty, no single item reduction is able to adequately reduce the overflow; we then proceed to Step 4, to attempt a reduction of the inventory positions of multiple items at once. Otherwise, we set $\bar{y}^{i^*}_{(t)} \leftarrow \bar{y}^{i^*}_{(t)} - s$, and $i^* = \operatorname{argmin}_{i \in \mathcal{I}^*} \Delta V^i_{(t)}$. Return to Step 2, starting with a recalculation of B, see (2.24), which by the definition of I^*, has been reduced by at least αs units, and hence a corresponding reduction of the step size s, see (2.25).

Step 4 Step 4 is only activated if no unilateral reduction of magnitude s can result in a reduction of B by at least $\min\{B, \alpha s\}$. If this happens, Step 4 considers, in addition to sample path m_o, any other sample path with an overflow larger than $(B - \alpha s)^+$ and less than or equal to B, and solves a series of continuous linear knapsack problems to bring the overflows of these sample paths to or below $(B - \alpha s)^+$ by reducing *more than one* product. This ensures that all but at most $\lfloor \beta M \rfloor$ sample paths, now, have an overflow of $[B - \alpha s]^+$ or less, so that the new value of B is less than or equal to $[B - \alpha s]^+$.

In particular, we define the following set of linear knapsack problems

$$\kappa(m,\delta) = \text{argmin}_{\boldsymbol{z}\geq\mathbf{0}} \left(\sum_{i=1}^{I} \Delta V_{(t)}^{i} z^{i}\right) \text{ s.t. } \sum_{i=1}^{i} z^{i} = \delta, z^{i} \leq (\bar{y}_{(t)}^{i} - \dot{u}_{(t),m}^{i})^{+} \tag{2.26}$$

This program calculates the cheapest way to reduce total inventory on sample path m by δ. The second constraint ensures that, for every sample path, we do not create any backorders through our reduction. Given a ranked list $\{\Delta V_{(t)}^{i}\}$, this problem can be solved in $O(I)$ time for any m and δ.

Thus, we identify the sample paths $\mathcal{M}$ that result in an overflow larger than $(B-\alpha s)^{+}$ and less than or equal to B. Without loss of generality, we shall assume these paths are ordered in descending order of their overflow value:

$$\mathcal{M} = \left\{ m \in \mathcal{M}^{+} : (B-\alpha s)^{+} < \left(\sum_{i=1}^{I} (\bar{y}_{(t)}^{i} - \dot{u}_{(t),m}^{i})^{+}\right) \leq B \right\}$$

Go through each sample path in $\mathcal{M}$ starting with the one with the largest overflow B (sample path m_o), reduce the inventory by solving an instance of the linear knapsack problem (2.26), to ensure the overflow is not above $B-\alpha s$ on that sample path;

$$\bar{\boldsymbol{y}}_{(t)} \leftarrow \bar{\boldsymbol{y}}_{(t)} - \kappa\left(m, \left(\left[\sum_{i=1}^{I} (\bar{y}_{(t)}^{i} - \dot{u}_{(t),m}^{i})^{+}\right] - \chi - (B-\alpha s)^{+}\right)^{+}\right)$$

If after this set of linear knapsack problems is solved $(B-\alpha s)^{+} = 0$, the chance constraint is met so we can stop. Otherwise, return to Step 2 to re-calculate $\mathcal{M}^{+}$.

We now carry out a complexity analysis of this greedy-drop procedure. Let $B_{(t)}$ denote the overflow on the sample path with the $(\lfloor \beta M \rfloor + 1)$th largest overflow after the t-th iteration of the greedy-drop procedure. B denotes the initial overflow. Recall that when $B_{(t)}$ drops below 0, the chance constraint is met and the algorithm stops.

Each step of the algorithm sets a step size $s = \max(\epsilon, \frac{B_{(t)}}{I})$, and guarantees a decrease in $B_{(t)}$ of at least αs. Once $B_{(t)} \leq \epsilon I$ and the step size becomes ϵ, the algorithm will converge within $\frac{\epsilon}{\alpha}$ iterations. Thus, we only consider the number of iterations required to reach an overflow of ϵI, which is the smallest value of n for which $B\left(1-\frac{\alpha}{I}\right)^n \leq \epsilon I$. Solving, we find

$$n = \left\lceil \log\left(\frac{B}{\epsilon I}\right) \frac{1}{\log I - \log(I-\alpha)} \right\rceil \leq \left\lceil \log\left(\frac{B}{\epsilon I}\right) \frac{I}{\alpha} \right\rceil = O\left(\log\left(\frac{B}{\epsilon I}\right) I\right)$$

where the inequality follows from $\log I - \log(I-\alpha) \geq \frac{\alpha}{I}$, by the concavity of the log-function.

We now consider the complexity of each of these iterations, which involve the following steps. (1) Calculating $\Delta V^i_{(t)}$ for every i, and sorting the resulting list: $O(I + I\log I)$. (2) Determining the set of products in $\mathcal{I}^*$: $O(IM)$. (3) Updating the current overflow in each sample path, and sorting these overflow values to find the $(\lfloor \beta M \rfloor + 1)$th largest one: $O(M \log M)$. (4) Solving at most M linear continuous knapsack problems in step 4: $O(IM)$. Thus, the total time complexity of the greedy-drop procedure in stage 2 is $O\left(I\log\left(\frac{B}{\epsilon I}\right)[I\log I + M\log M + IM]\right)$.

2.5 Numerical Study

We describe an extensive numerical study to assess the optimality gap of our alternative heuristic policy. We also identify the impact of various factors on the overall cost and service performance. We divide our study into two parts. The first part covers a broad range of parameters representative of what might be encountered in practice, and has 28,800 instances.

In the vast majority of these settings, our heuristic performs remarkably well. In only 50 of the 28,800 instances do we observe optimality gaps larger than 10%. Furthermore, comparing our heuristic to a simpler one, we find that in the most constrained cases (those with `chiBase=1`; see Table 2.1), our heuristic performs on average 3.91% better. The second study evaluates the extent to which our methods scale with the number of products, and has 492 instances.

All instances consider a horizon of $T = 20$ periods and the total expected cost objective, with linear holding and backlogging costs. The cost performance of the proposed heuristic strategy is assessed via Monte Carlo Simulation, using 2,000 sample paths for each problem instance.

We first provide a detailed description of the parameters and demand distributions chosen in the numerical study.

Numerical Study Setting

Costs

Salvage costs are kept constant across all instances at $d = \max_i \left(\sum_{\tau=1}^{T} h^i_{(\tau)}\right)$. This value is chosen to ensure it is always cheaper to hold an item (even for the full planning horizon) than to salvage it. The salvage option is nevertheless used if it is absolutely necessary to meet the capacity constraint. Recall that salvage costs were included in our formulation to cover settings with merchandise return options but also to ensure feasibility.

Holding costs are set in two stages. First, we define a variable $\eta_{(t)}$ for each time period that determines the variation of these holding costs over time. if the meta-parameter `costRandomness` is equal to C, the costs are kept **C**onstant over time and we set $\eta_{(t)} = 10$ for all t. If `costRandomness` is equal to R, the $\eta_{(t)}$ are **R**andomly and uniformly drawn for each time period t in the interval [8, 12]. The `itemCostRatio` meta-parameter then determines the heterogeneity of these costs among the different items. If it is equal to F, costs are kept **F**ixed among all items in each time period, $h^i_{(t)} = \eta_{(t)}$. If it is equal to V, costs are allowed to **V**ary across items, $h^i_{(t)} = \frac{2}{3}\eta_{(t)}(1 + \frac{i-1}{I-1})$. Without loss of generality, items are indexed in increasing order of this expected holding cost rates. In addition, in every period, the holding cost rates are stochastically ordered in the same ranking.

Note that this scheme ensures that $\mathbb{E}\left[\sum_{i=1}^{I} h^i_{(t)}\right] = 10I$ across all instances. This allows different instances for a given number of items to be compared on a like-for-like basis.

Backlogging costs and **variable ordering costs** are set as a ratio of their holding costs; $p^i_{(t)} = \texttt{costRatio} \cdot h^i_{(t)}$ and $c^i_{(t)} = \texttt{orderToHoldRatio} \cdot h^i_{(t)}$.

Demand Distributions

In every one of our instances, the demands in each period, $\boldsymbol{u}_{(t)}$ follow a multivariate Normal distribution with mean $\boldsymbol{\mu}_{(t)}$ and covariance matrix $\boldsymbol{\Sigma}_{(t)}$.

The method used to generate the **means** is determined by the `demandRandomness` meta-parameter. If this parameter is set to C, the means are kept **C**onstant over time and items, at $\mu^i_{(t)} = 10$. If the parameter is set to I, the means are kept constant across items, but are allowed to **I**increase linearly over time; in this case, we set $\mu^i_{(t)} = 5 + 10\frac{t-1}{T-1}$. If the parameter is set to D, the means are kept constant across items, but are allowed to **D**ecrease linearly over time; we set $\mu^i_{(t)} = 5 + 10\frac{T-t}{T-1}$. If the parameter is set to R1, the means are allowed to vary over items *and* time and are set **R**andomly; more specifically, we generate each $\mu^i_{(t)}$ independently and uniformly on the interval $[8, 12]$. If the parameter is set to R2, each $\mu^i_{(t)}$ is **R**andomly set to 0 with probability $\frac{1}{4}$, and uniformly distributed on the interval $[10, \frac{50}{3}]$ with probability $\frac{3}{4}$. Finally, if the parameter is set to S, the mean demands vary in a **S**inusoidal pattern; we set $\mu^i_{(t)} = 10(1 + \frac{1}{2}\sin(2\pi\frac{t-1}{T}))$. Note that this scheme ensures $\mathbb{E}\left[\sum_{i=1}^{I}\sum_{\tau=1}^{T}\mu^i_{(\tau)}\right] = 10TI$ across all instances. This allows different instances for a given number of items to be compared on a like-for-like basis.

To generate the **covariance matrices**, we begin by setting the standard deviation of each individual single-product demand to $\sigma^i_{(t)} = \texttt{CVBase} \cdot \mu^i_{(t)}$. We then generate our covariance matrix as follows

$$[\mathbf{\Sigma}_{(t)}]_{p,q} = \sigma^p_{(t)}\sigma^q_{(t)} \cdot \begin{cases} 1 & \text{if } p = q \\ \rho & \text{if } p \neq q \text{ and } \texttt{covMethod} = \text{N} \\ |\rho| \cdot \text{sign}(\rho)^{p+q} & \text{if } p \neq q \text{ and } \texttt{covMethod} = \text{A} \end{cases}$$

When `covMethod` is equal to N, the covariance matrix reflects identical correlation coefficients for every pair of items. For the variance-covariance matrix to be positive definite, this implies a lower bound $\rho \geq -(I-1)^{-1}$. For negative correlations, this is, of course, quite restrictive. However, it is rare to have a facility in which all item pairs are negatively correlated. For this reason and to allow our numerical study to explore instances with more negative correlations, we introduce a second method for generating covariance matrices. When `covMethod` is set to A (for **A**lternating) the products are split into two groups – products *within* groups are positively correlated to each other, but negatively correlated to any product in the other group. This method can be shown to always lead to a valid, positive semidefinite covariance matrix for all values of ρ. Note that for positive correlations, the two methods are identical.

Storage Capacity

The storage capacity is assumed to be constant over the horizon and set as follows

$$\chi_{(t)} \equiv \chi \equiv \frac{1}{T}\sum_{\tau=1}^{T}\sum_{i=1}^{I}\mu^i_{(\tau)} + \texttt{chiBase}\cdot\sqrt{\ell}\left[\frac{1}{T}\sum_{\tau=1}^{T}\sqrt{\sum_{i=1}^{I}(\sigma^i_{(\tau)})^2}\right] \tag{2.27}$$

In order words, the capacity is set to the average aggregate mean demand in each period, plus a multiple of the average standard deviation of the aggregate demand in each period. When computing this standard deviation, we assume the items are uncorrelated – this is to ensure instances with different correlations and the same `chiBase` have the same capacity and can be compared.

There are 27,648 instances in our study with `covMethod` = A (4 values of I, 6 values of ρ, 4 values of `chiBase`, 2 values of `costRandomness`, 2 values of `costRatio`, 2 values of `itemCostRatio`, 2 values of `orderToHoldRatio`, 6 values of `demandRandomness`, and 3 values of `CVBase`). In addition when $I = 4$, it is possible to generate a valid positive semidefinite covariance matrix using `covMethod` = N when the correlation is $\rho = -0.3$. This adds 1,152 instances.

Numerical study I

In our first numerical study, 28,800 instances were chosen to cover a broad spectrum of potential settings, to ensure our heuristic performs well in these various situations, and to assess the impact of certain parameters on system costs. Table 2.1 summarizes the meta-parameters that are used to generate the various problem instances in our first study, as described in the supplementary material.

We evaluate the performance in any given instance with two key measures. The first is $(\mathbb{E}[UB] - LB)/LB$, where UB denotes the simulated cost of our proposed strategy, and LB the mean cost of the lower bound (calculated exactly using the relaxed DP). This ratio is a *conservative* upper bound for the true optimality gap.

The second performance measure quantifies the extent to which our strategy improves over a simple benchmark strategy, thus demonstrating the practical utility of our approach. We use two benchmarks; both begin by finding the optimal dual-base-stock policy for each product independently (ignoring capacity). In any period in which this policy is infeasible, the two benchmarks take a different approach

Parameter	Values	
I	4,10,20,40	The number of items.
ℓ	4	The lead time from the supplier to our facility.
ρ	-0.6, -0.3, 0, 0.3, 0.6, 0.9	The correlation between items.
`chiBase`	1,2,5,1000	A 'meta-parameter' used to specify the capacity at our facility.
`costRandomness`	C, R	The method used to generate holding costs – constant or random.
`costRatio`	10, 20	A parameter used to specify the ratio of backorder to holding costs.
`itemCostRatio`	F, V	A parameter used to specify the heterogeneity of costs among items.
`orderToHoldRatio`	5, 8	A parameter used to specify the ratio of variable ordering costs to holding costs.
`demandRandomness`	C, I, D, R1, R2, S	The method used to generate demand
`CVBase`	0.15, 0.3, 0.4	A parameter used to specify the coefficient of variation of the single-period demand at each retailer.
`covMethod`	N, A	A parameter used to specify the way the covariance matrices are generated in each period.

Table 2.1: Meta-parameters used in Numerical Study I. See the supplementary material for a description of the way these meta-parameters are used to generate the parameters in each instance.

1. Our first benchmark simply reduces the capacity consumed by each product successively by a small amount until the order is feasible.

2. Our second benchmark takes a less arbitrary approach and tries to identify the best products to reduce. We first calculate $p^i_{(t+\ell)} - h^i_{(t+\ell)} - c^i_{(t)}$ for each product to approximate the cost of reducing each item's inventory by one unit. We start with the cheapest item, and reduce its order-up-to quantity to the expected demand over the next ℓ periods. We then proceed with the next cheapest item. Finally, if all items have been reduced to their expected demand and the order is still infeasible, we proceed as in the first benchmark.

In the vast majority of instances, our first benchmark performs better, so we confine ourselves to it; all further references to 'the benchmark' will refer to this first benchmark. In particular, we report $\mathbb{E}[(Ben - UB)/Ben]$, where Ben is the benchmark on any given sample path. Given a specific implementation of the benchmark, this measure will tell him or her how much they might have saved by moving to our approach.

Finally, we evaluate the impact of various system parameters on the cost of managing inventories. This provides valuable managerial insights – for example, on the value of creating additional capacity. Figure 2.1 displays the values of the two evaluation measures.

Considering these plots, the optimality gaps $(\mathbb{E}[UB] - LB)/LB$ are almost always very small; the average gap is 1.05%, 99.83% of instances have gaps smaller than 10%, and 98.20% smaller than 5%.

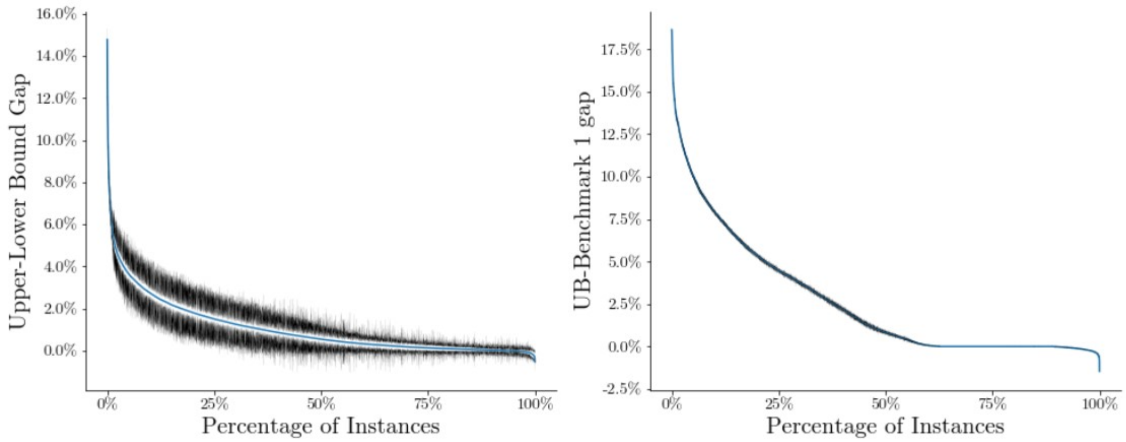


Figure 2.1: Plots summarizing the values of the two evaluation measures in our first numerical study. Gray lines represent 95% confidence intervals compared to these reported means induced by simulation errors. Note that around one quarter of instances in this study are unconstrained (`chiBase` = 1000); these instances all have a 0% gap by both measures.

We also find considerable improvement over our benchmark. Looking specifically at the most constrained instances (those with `chiBase` = 1), 87% of instances improve over our benchmark, and 33% beat it by at least 5%.

We note that even when the optimality gap is significant, significant improvements are obtained over our benchmark. Figure 2.2 graphs all instances with an optimality gap greater than 5%; even instances with larger gaps can significantly improve over our benchmark. This might indicate the gap stems from deficiencies in the lower bound rather than limitations of our strategy.

Referring to our discussion in the supplementary material referring to the subset of $\mathcal{A}$ we choose, we note that in 73% of cases, only one round of constraint generation is required to solve the lower bound. In 26% of cases, we need a second round which adds some of the more complex constraints described in the supplementary material but still leaves $|\mathcal{A}| = O(I)$. Finally, in the remaining 0.009% of instances, the gap after round 2 is greater than 5%, and additional constraints need to be included in a third round.

Unsurprisingly, those instances in which more than the single constraint are required tend to result in larger optimality gaps. Instances with one, two, and three rounds result in average gaps of 0.61%, 2.09%, and 7.04% respectively. Encouragingly, however, those instances also result in the greatest improvement over our benchmark, with average improvements of 1.72%, 5.08%, and 6.34% respectively. This further confirms

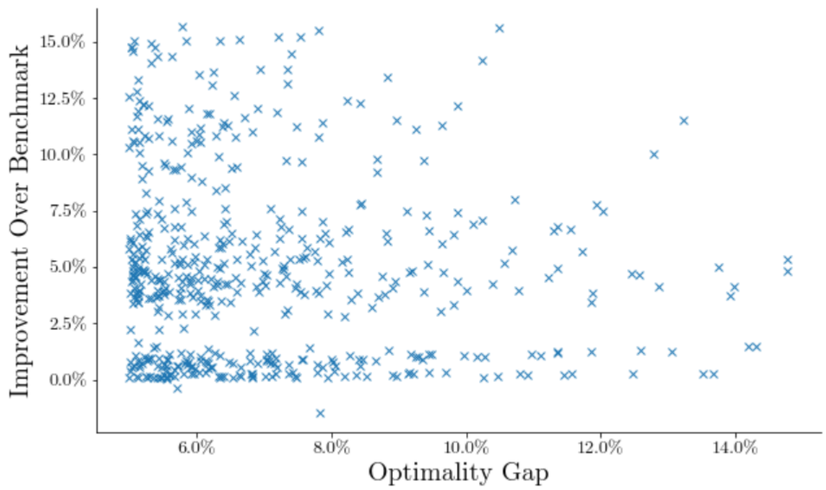


Figure 2.2: Scatter plot displaying all instances with optimality gaps greater than 5%. Even instances with larger optimality gap improve – sometimes significantly – on the benchmark.

the practical usefulness of our approach, even when a small optimality gap cannot be guaranteed.

Finally, we turn to the impacts of the system parameters on these evaluation measures; we summarize these impacts in Figure 2.3.

Our first observation is that our performance measures are surprisingly stable as the number of items is varied. This points to the stability of our approach across a broad range of parameters and number of products.

It is also interesting to note that the benefits of our technique are most apparent when faced with with non-stationary model parameters. Indeed, when `costRandomness=R` or `demandRandomness=R2`, our strategy improves over the benchmark by 4.3% on average. Looking at instances in which *both* parameters are set to these values simultaneously, the average improvement is 7.3%. By contrast instances in which *neither* parameters are set to these values, the average improvement is 0.33%. Our method therefore excels in the particularly difficult situation in which demands are highly non-stationary. We note that when `demandRandomness=R1` (still non-stationary, but more stable over time), the benefits are still apparent, but less pronounced).

The capacity at our facility unsurprisingly has a strong impact on both measures. When `chiBase` = 1000, capacity will never constrain our decisions; the optimality gap is very small and the improvement over the baseline is minimal. By contrast, when capacity is tight, the optimality gap is larger (albeit still small), and

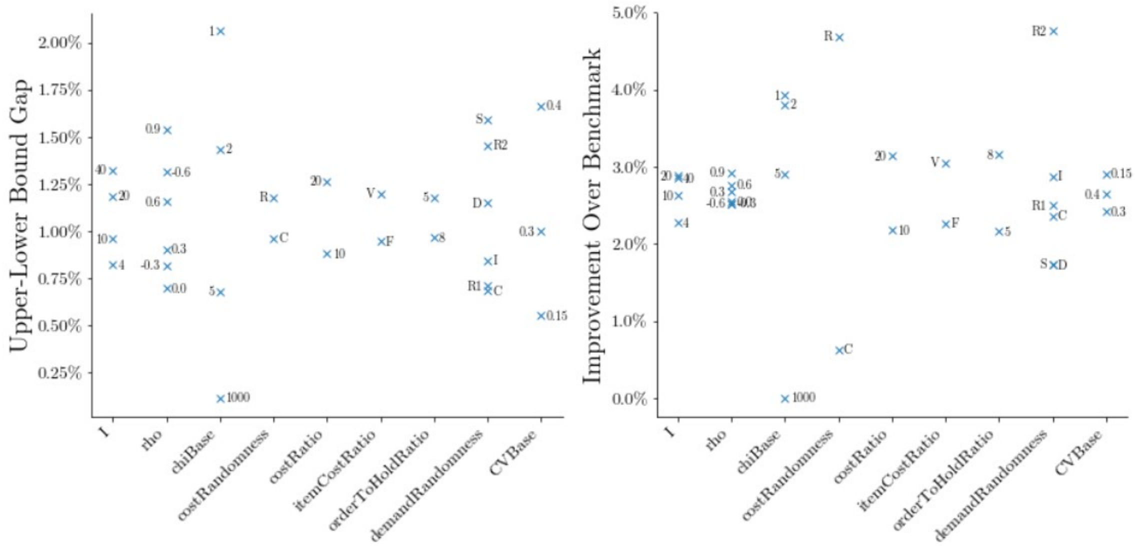


Figure 2.3: Each plot illustrates the impact of the system parameters. Each point corresponds to one value of a single parameter, and finds the average of the evaluation measure over all instances with this value of that parameter. This data excludes the 864 instances with $I = 4$, $\rho = -0.3$, and `covMethod`=N, to ensure that the remaining instances form a factorial design. This ensures that these averages correctly reflect the impact of each parameter value.

significant improvements over the benchmark are achieved.

We finally turn to those 11% of instances in which our strategy *underperforms* the benchmark. First, we note that in the *worst* of these instances, the benchmark only outperforms us by 1.46%, and on average, it only outperforms us by 0.15%. Figure 2.4 displays the parameter values for those instances. The figure strikingly demonstrates that almost all of these instances have one thing in common – the costs are identical across items. In those specific instances, symmetry among the products would suggest that all products should be treated identically, and our naive benchmark ensures this happens. In our more complex strategy, by contrast, symmetry may be broken depending on the first product that gets selected for reduction. As a result, our strategy might *very slightly* underperform the benchmark.

Numerical study II

Our second numerical study focuses on the runtime of the *upper bound* heuristic. We focus on the upper bound because the time taken to compute the *lower bound* is expended only *once*, when the system is set up. The upper bound, however, needs to be solved every period to determine order quantities, and therefore directly impacts the viability of the system.

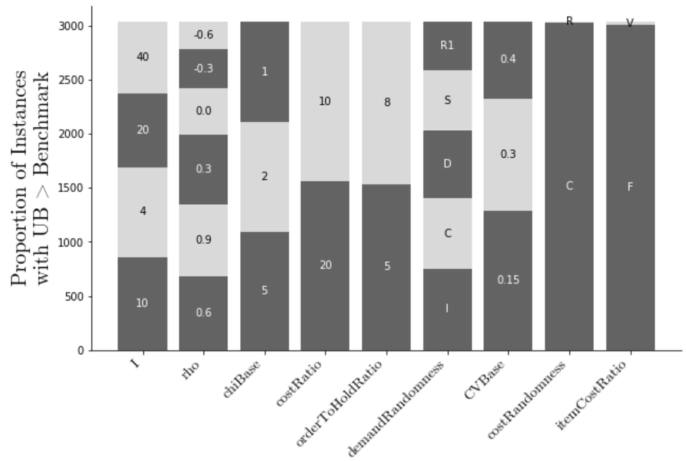


Figure 2.4: Parameter values for the 11% of instances in which our strategy underperformed the benchmark. Each bar corresponds to one parameter, and each segment corresponds to one value of that parameter. The height of each segment denotes the number of instances with that parameter. This data excludes the 864 instances with $I = 4$, $\rho = -0.3$, and `covMethod`=N, to ensure that the remaining instances form a factorial design.

To run this study, we consider all instances in our first numerical study with 40 items, and select the three instances that take longest to run, the three instances that run fastest, and the three instances at each of the 25th, 50th, and 75th percentile. We then run additional instances with each of these parameter combinations, with 40-200 items in 20 item increments, and also 300 and 400 items. The runtime of each of the resulting instances are summarized in Figure 2.5.

Our theoretical results indicate that the runtime is quadratic in the number of products. The results in Figure 2.5 reveal that in practice, the runtime scales almost linearly with the number of products, making it amenable to handling the larger number of products modern systems require.

2.6 Managerial insights

We have thus far focused on our ability to make near-optimal procurement decisions in complex, multi-item systems. There is, however, a second equally important benefit that can be derived from our technique.

The systems we consider in this paper are characterized by a large number of parameters. Some of these parameters are unlikely to be under the supply-chain planner's control. For example, the periodicity of demand is likely to be dictated by market conditions. Others, however, *will* be design parameters of the

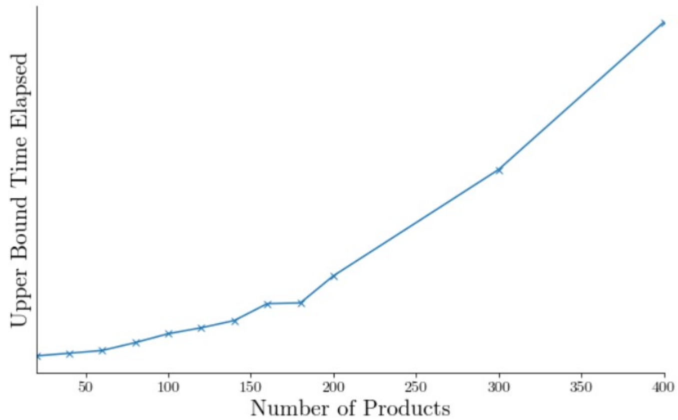


Figure 2.5: Upper bound runtimes as a function of products in the problem, in arbitrary units. Because we run this numerical study on a grid comprising many computers with different numbers of processors, we standardize all times by a 'benchmark time', obtained by running a large number of matrix multiplications on that machine.

system. For example, choosing to locate a fulfilment facility closer to a supplier would reduce the lead time ℓ, perhaps at the cost of smaller capacity or greater holding costs.

Faced with this bewildering number of choices, it is essential for a supply-chain planner to be able to simulate a myriad of scenarios to assess the impact of his or her decisions. Our technique allows these simulations to take place. In this subsection, we discuss the valuable managerial insights that result from understanding the impact of various cost parameters, and the implications for assortment planning and dual sourcing from different fulfillment centers.

Understanding the impact of cost parameters

At its most basic level, supply-chain planning requires a good understanding of how each of the system parameters affect the total cost of running the system. Does capacity impact costs linearly, or otherwise? How much of an impact does the correlation between items have on total costs?

To answer these questions, we display the cost of managing our inventories, per item in the system, in each of our instances. Figure 2.6a displays these costs as a function of the system parameters.

With the exception of `costRatio` and `orderToHoldRatio` the parameter values were specifically chosen so that their costs could be compared across instances. For example, demand parameters are chosen so the aggregate expected demand is invariant to variation among time periods or products.

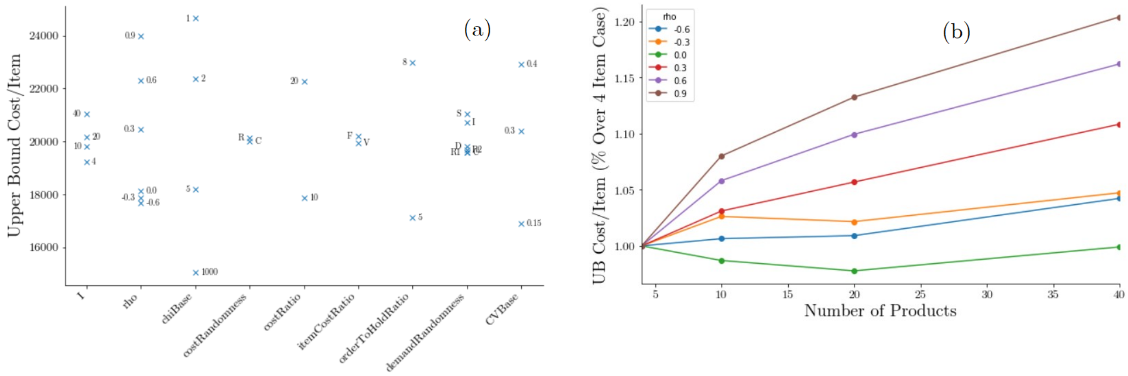


Figure 2.6: Managerial insights for Study I, excludes the 864 instances with $I = 4$, $\rho = -0.3$, and `covMethod`=N, to ensure that the remaining instances form a factorial design. (a) Upper bound cost per item as a function of the various system parameters. Each point corresponds to one value of a single parameter, and finds the average of the cost per item over all instances with this value of the parameter. (b) Dependence of the per-item cost of running our system as a function of number of items. The vertical axis represents the per-item cost of running our heuristic as a percentage of the cost for 4 items.

This plot provides valuable managerial insights. For example, focusing on the correlation parameter, we note that the larger the (positive) correlation between items, the higher the cost of running the system. This is unsurprising – when items are all positively correlated, surges and dips in demand are likely to be more extreme. It is interesting, however, to note that the decreased costs that result from negative correlation are not as pronounced as the increases are under high positive correlation.

Thus, in deciding how to assign inventory to a number of warehouses, minimizing average correlations is likely to reduce costs considerably. However, if positive correlations are due to similar items being *ordered* together, locating them in the same warehouse could result in savings in shipping costs; these considerations would have to be weighed against each other.

Looking at the first column of points in Figure 2.6a, it appears that the cost of running inventory systems of this kind scales superlinearly with the number of products. This would imply that based purely on fulfillment costs, it is likely to be cheaper to run several smaller facilities each handling a smaller number of products. However, this phenomenon bears closer scrutiny. Figure 2.6b demonstrates how the cost per item varies with I, by correlation level. When items are negatively correlated, per-unit costs are roughly flat and even sometimes decreasing as the number of items increases. It is only when correlations are positive that we see increases in per-unit costs as the number of items increases. The fact that under uniform positive correlations,

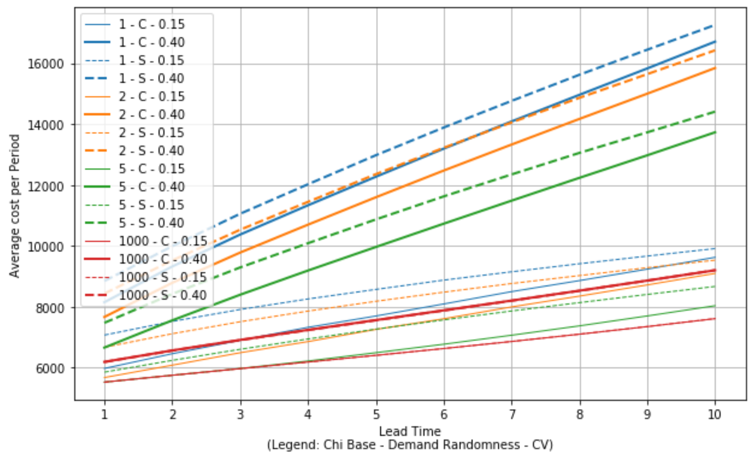


Figure 2.7: Analyzing the impact of lead time of system costs for a number of parameter configurations. The color of the line represents the capacity in the form of `chiBase`. The linestyle represents the seasonablity of the distribution. The width of the line represents the coefficient of variation of the demand distribution.

system-wide costs grow superlinearly with the number of items (see the upper curves of Figure 2.6b) appears to result from the way we have adjusted the allocated capacity as the number of items I increases. As described in equation (2.27) in the supplementary material, the surplus capacity beyond the aggregate mean demand for each instance was set to `chiBase` times the standard deviation of that demand, *calculated assuming no correlation between the items*. This choice was made to ensure correlated and uncorrelated instances could be compared on the same footing. However, when demands are positively correlated, the volatility grows faster with I than in the uncorrelated case ($\sqrt{I + I(I-1)\rho}$ assuming comparable standard deviations for all items, as compared to $\sqrt{I}$ in the uncorrelated case), and this disparity increases with ρ. Thus, for many items, the same `chiBase` might be more constraining in practice when correlation is positive. Our results therefore provide a valuable managerial insight – ignoring correlations among items can result in unpleasant and unexpected costs. Indeed, any system which provisions storage capacity based on aggregate mean demand and the items' coefficients of variation is likely to result in inadequate capacities, an effect that worsens as I and ρ increase.

It is also informative to consider the impact of lead times on procurement costs. Figure 2.7 investigates this effect in instances with a variety of system parameters. The larger the lead time, the greater the variance of the lead time demand. In light of our observation that large demand variability results in higher costs, we might expect instances with larger lead times to lead to higher costs, and this expectation is borne out. The

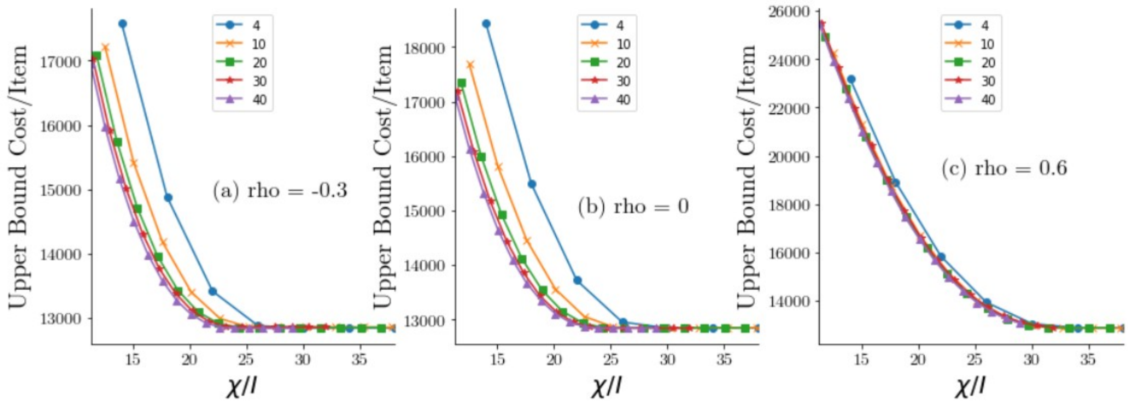


Figure 2.8: Cost per item for instances with various capacity levels and correlations, and `costRatio`=10, `itemCostRatio`=1, `orderToHoldRatio`=5, `demandRandomness`=C, `CVBase`=0.4 and `covMethod`=N. Each of the colored lines corresponds to instances with a different number of products.

dependence is most pronounced when the demand is highly uncertain (i.e., when the coefficient of variation is large), and highly constrained.

It is finally worth noting that, faced with a specific set of fixed parameters, a supply chain planner could use our technique to evaluate the cost impacts of the remaining parameters. Figure 2.8, for example, considers a specific set of instances and explores the impact of correlation, number of items, and capacity. A number of valuable insights can be garnered from these plots. First, we notice that when the correlation between items is negative and capacity is constrained, economies of scale are realized as the facility handles an increasing number of products. These economies all but disappear with positive correlations. Second, these plots allow us to quantify the value of increases in capacity, thus weighing them against the cost of building larger facilities.

Comparing the three plots with each other, we also notice that when correlations are positive, the per-item cost of running our system as a function of per-item capacity does not depend on the total number of items I. In other words, comparing a facility handling 10 items to one handling 100 items, increasing the capacity for *each* item by one unit would have the same effect in each case. This result is counter-intuitive (one might have expected economies of scale to make the extra storage per unit more valuable when many units are able to share that new space), and is a direct result of the strong positive correlation. Items are unable to ‘share’ space, because peaks in demand for one item are likely to occur at the same time as peaks in demand for other items. When the correlation is 0 or negative, we do observe the economies of scale our intuition would have suggested.

Assortment planning

One of the key characteristics of the system we study is its ability to handle multiple, competing products. One of the key decisions a practitioner is likely to be faced with in such a system is *which* assortment of products to stock. The problem has been studied extensively from the perspective of *demand*. The assortment chosen, however, will also impact the cost of achieving a certain service level from a procurement perspective. In this section, we use our technique to develop a number of general guidelines a practitioner might want to adopt in making this assortment decision from a supply perspective.

We consider a retailer faced with a menu of eight items, and the choice of *which* of these eight items to stock in his or her facility. These items are chosen to vary on three dimensions; the mean demand (low or high), the level of demand uncertainty (average coefficient of variation of 0.15 and 0.4), and the seasonality of demand (seasonal versus constant). To develop some general guidelines to help the retailer with this choice, we consider all $(2^8 - 1)$ possible assortment of these items, for 11 possible values of the capacity, leading to a total of $(2^8 - 1) \times 11 = 2,805$ individual instances. To compare each of these instances with different assortments, we consider a number of potential expected revenues per unit, and look at the ratio of expected revenue to procurement costs in each case.

These results allow us to identify – for any given value of the capacity – the optimal assortment. Figure 2.9 displays the optimal assortment size for a variety of capacities and per-unit revenues when selling the item. It is interesting to note that there is a revenue threshold below which no assortment leads to revenues, however large the capacity. Above that level, the larger the capacity and the larger the revenues per item, the larger the optimal assortment size.

To develop a set of guidelines a retailer might use to decide *which* items should be stocked, we observe the order in which items enter the assortment as capacity increases and larger assortments become optimal. We find that the most important factor is the demand uncertainty – in particular, low uncertainty items (with coefficient of variation 0.15) are the first to enter the assortment, followed by the more uncertain items (coefficient of variation 0.4). Constant demand also tends to result in better performance than seasonal

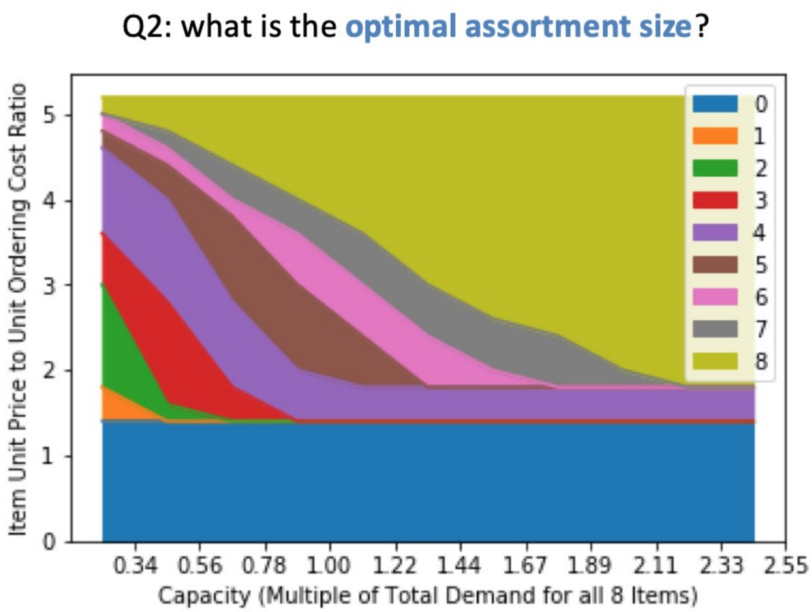


Figure 2.9: Analyzing the optimal assortment size for a variety of problem parameters. The x-axis represents the capacity, as a multiple of the total mean demand for all eight items, and the y-axis represents the ratio of the per-item revenue to the per-item holding cost. The color of each point in the plot represents the size of the optimal assortment.

demand. When demand is uncertain, *smaller* mean demands are preferable; when demand is more certain, *larger* mean demands are preferable. These guidelines might therefore be used to guide a retailer's assortment decision.

As discussed above, our model is able to handle very general demand distributions, including those with arbitrary correlations between items; our technique could therefore be used to simulate the cost of an assortment under *any* specific demand model that might prevail in a specific retail scenario.

Planning for dual-sourcing

We now consider one final managerial scenario in which a retailer stocks 10 items, and owns two fulfilment centers, each with equal capacity. Demand for each item is Normally distributed with the same mean and standard deviations, but the items exhibit a complex correlation structure, in which the items split into two clusters – a 7-item cluster and a 3-item cluster. Within each cluster, the items are completely uncorrelated. Each of the items in the 7-item cluster are correlated to each of the items in the 3-item cluster, and vice-versa. The retailer is faced with a simple choice – how to allocate these 10 items across the two fulfilment centers.

One might initially assume that the lowest-cost configuration would be to place five items in one facility and five in the other, to equalize the mean demand at each facility. There is, however, a second option – the correlation structure between the ten items means that placing each cluster in its own facility (3 items in one facility, and 7 in the other) will result in the lowest variance in demand at each facility; we have already seen variance in demand is a strong driver of increased costs.

To analyze this question, we simulate a number of instances with a variety of capacities. In each case, we consider every possible assignment of the 10 items across the two fulfilment centers, and identify those assignments that lead to the lowest costs. There are $\sum_{i=1}^{5} \binom{10}{i} = 637$ such combinations.

Figure 2.10 displays these results. The conclusions are striking. Unsurprisingly, when capacity is infinite, the exact assignment makes no difference (this serves as another proof of the importance of modelling capacity in fulfilment problems). When capacity is moderately constrained the cheapest combinations are the ones in which the mean demand is balanced at both fulfilment centers (5 products in each). However, when capacity is *severely* constrained, we find that reducing the variance is more importance than balancing mean demand.

2.7 Generalizations and Conclusions

Several variants of the base model may arise, and we expand their effective analysis in this section. We confine ourselves to the Alternative Heuristic of Section 2.4. The generalizations considered are (i) capacity limits for order sizes and/or salvage opportunities; (ii) intertemporal correlations for product demands; (iii) alternative capacity constraints such as conditional value at risk (CVaR) constraints and (iv) fixed order costs.

Capacity limits for order sizes and/or salvage opportunities

Our base model addresses only (joint) storage or inventory budget constraints. Other capacity constraints may curtail the feasible action sets in any state or period, most prominently bounds on the amount that can be procured for any given item in any given period, or the size of any salvage batch. Such constraints are easily incorporated into the analysis. The Lagrangian relaxed dynamic programs again decomposes on an item-by-item basis, and the optimal policy for each item continues to be of a very simple structure, again characterized by a pair of threshold values $\{S^i_{(t)}, B^i_{(t)}\}$, see Theorem 7. The only modification is that the

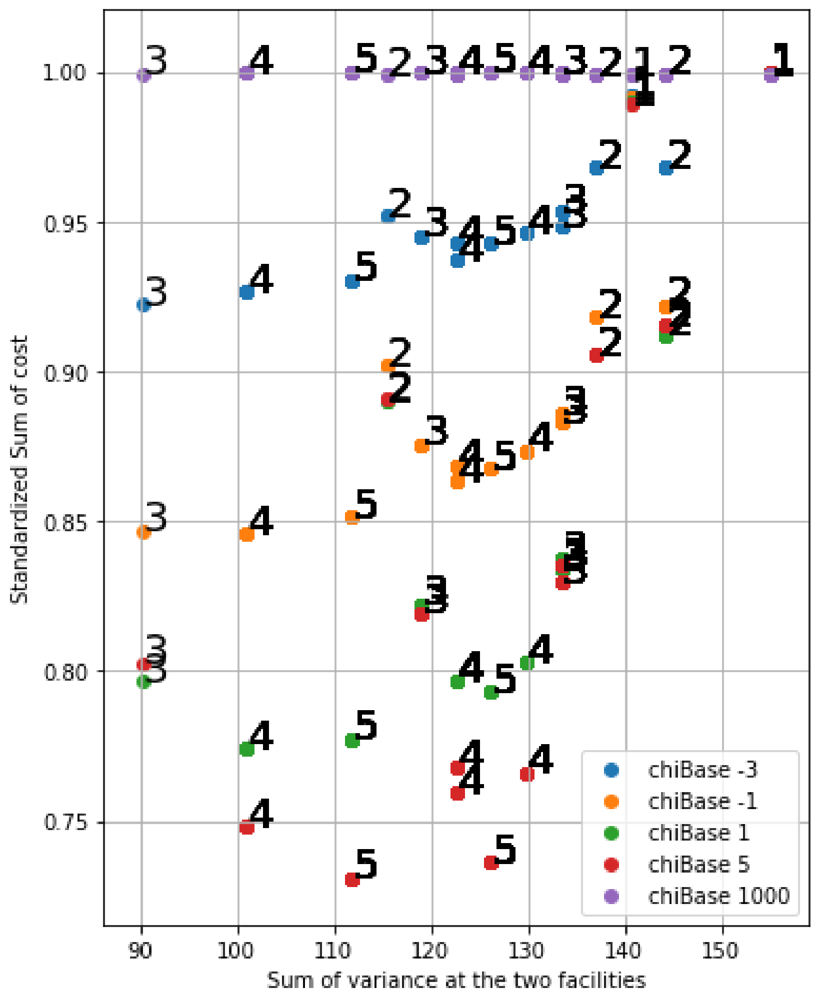


Figure 2.10: An analysis of the 637 possible assignments of ten items to two fulfilment facilities, for five values of the capacity, leading to $637 \times 5 = 3,185$ distinct combinations. The color of each point corresponds to the capacity simulated, and the number next to the point corresponds to the number of items in one of the two fulfilment facilities in that simulation. Many points are overlapping. The x-axis indicates the sum of the demand variance at each of the two facilities – because of the correlation structure between items, the lowest-variation configuration will always be to put the 3-item cluster in one facility, and the 7-item cluster in the order. The y-axis indicates the total cost of running the system, standardized so that – for each capacity – the maximum cost is 1.

optimal policy is a modified base-stock policy: when the inventory position is below $S^i_{(t)}$ [above $B^i_{(t)}$], an order is placed [a salvage batch is initiated] that brings the inventory position as close as possible to this threshold value as allowed by the new order/salvage constraints.

Finally, any joint order constraints imposing an upper bound on sums of order quantities over given order sets are of the exact same type as the constraint (2.2) and can thus be handled analogously. The upper bound heuristic can be implemented without any modification.

Intertemporal correlations for product demands

As in the vast majority of inventory models, our base model assumes that demands are independent across time (although with general correlation pattern across items). There are, however, several tractable and parsimonious ways to represent demands that are auto-correlated. One approach is to represent the demand process $\{u^i_{(t)} : i = 1, \ldots, I, t = 1, \ldots, T\}$ as modulated by an underlying state-of-the-world $W_{(t)}$ that evolves according to a finite state Markov chain. The state-of-the-world reflects environmental factors, such as fluctuating economic conditions (currency and interest rates, tariffs, labor cost indices, consumer confidence index etc.) or seasonality factors, among others. See Zipkin (2000) for a treatment of such state-of-the-world driven inventory models.

Our computational approach continues to apply without essential complications. The above relaxations, culminating with Lagrangian relaxation of capacity constraints, continues to generate a decomposition into I independent single-item DPs. It is easily shown that for each item and, for any choice of the vector of Lagrange multipliers, the optimal policy continues to be a double base stock policy, except the thresholds, $S^i_{(t)}, B^i_{(t)}$ now depend on the prevailing state-of-the-world.

Computing the Lagrangian dual, as well as the entire upper bound heuristic proceed without any modification. Note that the upper bound heuristic continues with the same complexity bound.

Simple time-series models provide an alternative approach to model auto-correlated demands. As an example, consider the exponential smoothing based model in Miller (1986). Let $\Delta^i_{(t)}$ be the "average demand factor for item i in period t" and assume that its dynamics satisfy the exponential smoothing formula $\Delta^i_{(1)} = \mu^i_{(1)}$ and $\Delta^i_{(t)} = (1 - e^i_{(t-1)})\Delta^i_{(t-1)} + e^i_{(t-1)} d^i_{(t-1)}$, where $\mu^i_{(1)}$ is an initial forecast of average demand, $0 \leq e^i_{(t)} \leq 1$

a smoothing factor, and $d^i_{(t)}$ the realized demand for item i in period t. Let $\{\boldsymbol{u}_{(t)}\}$ be as defined in Section 2.2, except that the vector is scaled so that $\mathbb{E}[\boldsymbol{u}_{(t)}] = \boldsymbol{e}_{(t)}$. Finally, define $D^i_{(t)} = \Delta^i_{(t)} u^i_{(t)}$. The Lower Bound DP under a given vector $\boldsymbol{\lambda}$ again decomposes by item. And while each of the single-item DPs now has a two-dimensional state space, Miller (1986) show that a double base stock policy continues to be optimal in every period, see Theorem 7. The critical base stock levels are of the form $\Delta^i_{(t)} S^i_{(t)}$ and $\Delta^i_{(t)} B^i_{(t)}$, with $S^i_{(t)}$ and $B^i_{(t)}$ easily determined from a one-dimensional DP.

The Upper Bound heuristic applies again without any modifications.

Alternative capacity constraints

A Conditional Value at Risk (CVaR) constraint may be used as an alternative to the chance constraint (2.2). Such a constraint takes the form $\mathbb{E}\left[\varphi\left(\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ - \chi\right)\right] \leq C$ for some convex, increasing function $\varphi(\cdot)$, with $\varphi(z) = 0$ for $z \leq 0$. We have treated the special case when $\varphi(z) = z$, for $z \geq 0$, in Section 2.3, and show its relationship to the chance constraints. For general (CVaR) constraints, it is easily verified that the function to the left of this constraint is jointly convex in the vector $\bar{\boldsymbol{x}}_{(t)}$. Thus, Lagrangian relaxation of this single constraint appears very attractive. However, the choice of the $\varphi(\cdot)$ function is rather arbitrary and implies very specific assumptions about how the "damage" associated with an overflow grows with its size. Moreover, to decompose the Lagrangian relaxed DP, it is necessary to find a separable approximation to the function to the left of this constraint. This would normally be done with convex conjugates, but the presence of $[\,]^+$-operators in the function's arguments makes this rather difficult. For these various reasons, we have focused on the chance constraint (2.2).

Fixed order costs

We conclude this section with a discussion of fixed ordering costs. The Lower bound DP could still be constructed as in Section 2.4; under any vector $\boldsymbol{\lambda}$ of Lagrangian multipliers, the optimal policy for any item i, in any period t, is still of a simple structure, now characterized by four thresholds $s^i_{(t)}, S^i_{(t)}, B^i_{(t)}, b^i_{(t)}$: it is optimal to elevate [reduce] the inventory position to $S^i_{(t)}$ [$B^i_{(t)}$] iff it is below [above] $s^i_{(t)}$ [$b^i_{(t)}$]. The upper bound heuristic requires modest adjustments. However, our more limited numerical experiments indicate that the gaps between LB and UB are typically significantly larger. Further work will address alternative lower bounds under fixed order costs.

Chapter 3: Stochastic Replenishment and Allocation Inventory Systems on DAG-Structured Networks

In the previous two chapters, we analyzed single-location inventory systems where multiple items share a common storage space. In reality, large retailer organizations, such as Amazon or Walmart, face the challenging problem of managing a complex inventory network so that a centralized decision involving all parties in the system needs to be made to maximize the overall corporate benefit. Any strategy under which each single facility location makes their own decision without coordination is suboptimal and can potentially incur huge costs that could have been avoided. Therefore, it is of great benefit to extend our inventory model by considering multi-location systems.

Our results in the previous chapters have shown that, a multi-item problem with shared capacity constraints can be decomposed into multiple single-item problems with modified storage costs, and depending on the type of the constraints (conditional or unconditional) imposed, a centralized problem may or may not be needed to solve. Thus, it is natural to first focus on single-item problems. Once the single-item multi-location problems have been solved, the decomposition technique can be naturally applied in the multi-item multi-location setting.

3.1 Introduction

In this chapter, we analyze a periodic review inventory replenishment and allocation system on a general supply network in which a single type of item can be procured, stored and distributed to satisfy random demands at downstream retailer locations. The network consists of: (i) a single source facility which accepts supplies of the item from the outside vendor; (ii) connecting facilities that can transship the item; (3) retailer stores facing random demands for the products. Such an inventory network can be described by a general directed acyclic graph (DAG), where each node of the graph represents a facility in the system, and each arc represents a direct route connecting two adjacent facilities. An example of DAG representation is in shown in Figure 3.1. Node 1 represents the source facility and nodes 5, 6, 7, 8 represent the retailer stores.

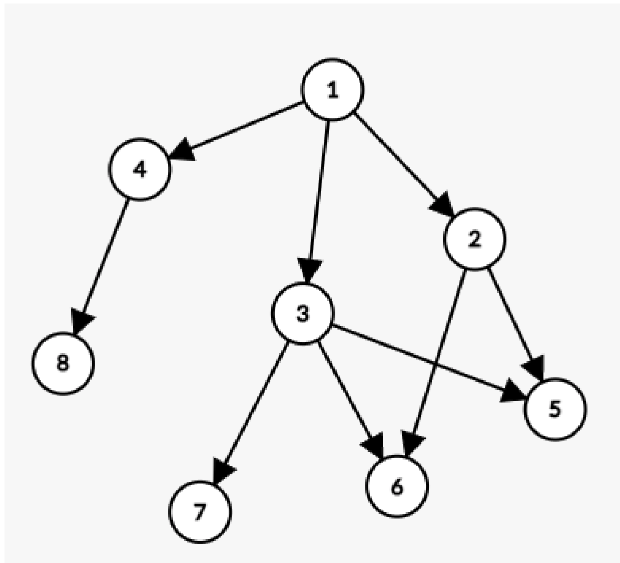


Figure 3.1: Supply network

In each period, a centralized decision is to make that includes: (i) an order to procure the item from the vendor that will arrive into the scoure facility a lead time later; (ii) a distribution of the current inventory at each facility to their connected downstream facilities. In each period, a random demand at each retailer store is observed that will consume the existing inventory at the store. Any demand exceeding the the inventory level at the retailer store is backlogged.

The goal of managing such an inventory system is to minimize the total expected cost over a fixed planning horizon while fulfilling the random demands. We consider four cost components. First, there is a unit ordering cost for each unit of the item procured from the vendor. Second, a facility-specific unit holding cost is charged on each unit of the item stored in the facilities. Third, we assume a universal unit transshipping cost for each unit of the item shipped within the system. Fourth, for each unit of the unfulfilled demand, a unit backlogging cost is incurred.

The problem can be formulated as a dynamic program. However, the dimension of the state space grows with the number of facilities in the network, which could be easily scaled up to hundreds or thousands. Thus, a direct solution to the dynamic program is computationally infeasible. Instead, a good approximate solution that is scalable is needed for practical interest.

3.1.1 Prior Work

One of the first papers addressing multi-location inventory models is the seminal paper by Clark and Scarf (1960). In their paper, they showed that an echelon base stock policy is optimal for the serial line problem. This optimality result was later extended by Federgruen and Zipkin (1984c) to an infinite horizon setting. Chen (2000) characterized optimal policies for serial system with batch ordering.

Clark and Scarf (1960) also discussed one-warehouse, multi-retailer systems and developed a lower bound for the problem based on a "balance assumption". Later important development on one-warehouse, multi-retailer problems includes Federgruen and Zipkin (1984a,b,c), Kunnumkal and Topaloglu (2008) and Federgruen et al. (2018). See Federgruen et al. (2018) for a detailed review on this type of problem.

An extension to the one-warehouse, multi-retailer model is the N-echelon distribution inventory model where the inventory network is generally described by a directed tree graph. Diks and De Kok (1998) showed that in this model, under a balance assumption, a base stock policy at each facility is optimal. A recent development on this model is by Maggiar et al. (2022) that considers lead time between installations. In their paper, a distribution problem is decomposed into serial line problems.

Opposite to the distribution inventory model is the assembly model where the inventory network can be described by a reverse directed tree graph. Rosling (1989) showed that under a "long-run balance" assumption, an assembly system can be interpreted as a serial system and thus an optimal policy can be derived. Chen (2000) derived a lower bound for assembly problems with batch ordering and developed a feasible policy that achieved this bound.

3.1.2 Our Work

In this chapter, we consider inventory systems that can be described by a general directed acyclic graph (DAG) which assembles the properties of both distribution and assembly systems. We derive a lower bound approximation for the dynamic program of the problem as well as a feasible heuristic which provides an upper bound on the optimal cost. To evaluate the accuracy of the lower bound approximation and the suboptimality of the proposed heuristic, we conduct an extensive numerical study to estimate the gap between the lower and upper bounds. The numerical study results suggest that our lower bound approximation is sufficiently accurate and our proposed heuristic is close to optimal.

The rest of this chapter is organized as follows. In Section 3.2, we introduce the model and the notation. In Section 3.3, the lower bound approximation to the original problem is presented. In Section 3.4, we propose a feasible heuristic which provides an upper bound on the optimal cost. In Section 3.5, we present an extensive numerical study to evaluate the gap between the lower and upper bounds. In Section 3.6, we conclude this chapter by discussing some generalizations.

3.2 Model and Notation

We consider a periodic review inventory network system over a finite planning horizon T. The network of the system can be described by a directed acyclic graph (DAG), where each facility location in the system is represented by a node in the graph. The products are shipped along the arcs of the network from the root facility to leaf facilities. One example of a such inventory network is shown in Figure 3.1 in Introduction.

We assume the DAG describing the system has a single root, which can procures products from an outside supplier. In other words, the system is single-sourced. Demand is observed and served only by the leaf facilities, and any unmet demand is back-logged at the leaf facilities.

A unit ordering cost is incurred for each unit of the ordered item. A unit shipping cost is charged for each unit of the item shipped within the system. At each facility, a holding cost proportional to the quantities of products stored in the facility is charged. At leaf facilities, a unit back-logging cost is incurred for each unit of back-orders.

The operation flow at a facility in a period is as follows. At the beginning of the period, products that were ordered or allocated into the facility in the last period arrive. Decisions regarding the ordering and allocation of the products are made. Demand during the period is observed. An ordering or back-logging cost depending on the end-of-period inventory is incurred.

3.2.1 Notation

Let $\mathbf{G} = \{\mathbf{V}, \mathbf{E}\}$ denote the DAG describing the inventory network, where $\mathbf{V}$ is the set of nodes and $\mathbf{E}$ is the set of arcs. Let $\mathcal{L} \subset \mathbf{V}$ denote the set of leaf nodes in $\mathbf{V}$, and $\mathcal{L}^c := \mathbf{V} \backslash \mathcal{L}$.

Let A denote the node-arc incidence matrix of the graph, where the m-th row and n-th column is 1 $[-1]$ if the n-th arc flows into [out of] the m-th node, and 0 otherwise. In other words, for an arc $e = (i, j)$ flowing from node i to j, we have $A_{i,e} = -1$, $A_{j,e} = +1$, and $A_{k,e} = 0$ for all $k \notin \{i, j\}$. For example, the node-arc incidence matrix of the DAG in Figure 3.1 is given by

$$A = \begin{bmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We assume a lead time of 1 for ordering into the source facility and shipping along any arc in the system. General procurement lead times and lead times of shipping into retailer stores (leaf facilities) can be easily handled by employing a forward accounting scheme. Transshipment lead times between non-leaf facilities can be handled by creating virtual facilities. We discuss these generalizations of lead times in section 3.6.1.

Throughout this chapter, we use the following notation for the problem.

- J: number of facilities in the network, i.e., $J = |\mathcal{V}|$. Each facility is index by $j = 1, 2, \ldots, J$, where 1 denotes the root facility.
- $c_{(t)}$: unit ordering cost in period t.
- $\gamma_{e,(t)}$: unit shipping cost on arc $e \in \mathbf{E}$ in period t.
- $h_{j,(t)}$: unit holding cost at facility j in period t.

- $p_{j,(t)}$: unit back-logging cost at facility j in period t.
- $u_{j,(t)}$: demand at facility j in period t following a distribution with mean $\mu_{j,(t)}$ and c.d.f $F_{j,(t)}(\cdot)$. For any $j \in \mathcal{L}^c$, $u_{j,(t)} = 0$ almost surely.
- $x_{j,(t)}$: inventory at facility j at the beginning of period t before inventory allocation and after order arrival (state of the system).

In addition, the decision variables are:

- $\bar{x}_{j,(t)}$: inventory at facility j at the beginning of the period t after inventory are distributed but before demand are observed.
- $\hat{x}_{j,(t)}$: inventory allocated to facility j that will arrive in period $t+1$.
- $\bar{w}_{(t)}$: order placed in period t that will arrive in period $t+1$.
- $f_{e,(t)}$: inventory that will be shipped along arc e in period t.

For each parameter or variable, we use a bold character to denote its vector form. For example, $\boldsymbol{x}_{(t)} := [x_{1,(t)}, \ldots, x_{J,(t)}]^\top$.

To facilitate the notation, we expand the ordering cost $c_{(t)}$ to a vector $\boldsymbol{c}_{(t)} := [c_{1,(t)}, \ldots, c_{J,(t)}]^\top$ where $c_{1,(t)} = c_{(t)}$ and $c_{j,(t)} = \infty$ for all $j > 1$. Similarly, we define $\bar{\boldsymbol{w}}_{(t)} := [\bar{w}_{1,(t)}, \ldots, \bar{w}_{J,(t)}]^\top$ where $\bar{w}_{1,(t)} = \bar{w}_{(t)}$ and $\bar{w}_{j,(t)} = 0$ for all $j > 1$.

In the following, we provide an exact formulation of the problem. The formulation itself, as we will see soon, is innovative with the inclusion of decision variables $\bar{\boldsymbol{x}}_{(t)}$ and $\hat{\boldsymbol{x}}_{(t)}$ in addition to the flow vector $\boldsymbol{f}_{(t)}$, which, as we will see in later sections, is of critical importance for deriving a tractable lower bound approximation and an upper bound heuristic.

3.2.2 Problem Formulation

In each period t, by the definitions of $\bar{x}_{j,(t)}$ and $x_{j,(t)}$, $x_{j,(t)} - \bar{x}_{j,(t)}$ is the inventory that will be shipped out of a facility j, and $\bar{w}_{j,(t)}$ or $\hat{x}_{j,(t)}$, (depending on whether j is a source facility or not), is the amount that will be shipped into the facility. Besides, $u_{j,(t)}$ is the demand that will be subtracted from the facility. Thus, the state transition of the system is given by

$$\begin{aligned}\boldsymbol{x}_{(t+1)} &= \boldsymbol{x}_{(t)} - (\boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)}) + \bar{\boldsymbol{w}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)} \\ &= \bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}.\end{aligned}$$

Let $A^+ := \max\{A, 0\}$ and $A^- := -\min\{A, 0\}$ denote the inflow and outflow matrix, respectively. In each period t, the sum of flows on all ingoing arcs into a facility j is given by $A_j^+ \boldsymbol{f}_{(t)}$, where A_j^+ is the j-th row of A^+. This is equal to the total inventory shipped into j, $\hat{x}_{j,(t)}$. Thus, we have the inflow constraints $A_j^+ \boldsymbol{f}_{(t)} = \hat{x}_{j,(t)}$, for $j = 1, \ldots, J$, or in its vector form

$$A^+ \boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)}. \tag{3.1}$$

Similarly, the sum of flows on all outgoing arcs from facility j, $A_j^- \boldsymbol{f}_{(t)}$, is equal to the total inventory shipped out of j, $x_{j,(t)} - \bar{x}_{j,(t)}$. This gives us the outflow constraints $A_j^- \boldsymbol{f}_{(t)} = x_{j,(t)} - \bar{x}_{j,(t)}$, for $j = 1, \ldots, J$, or in its vector form

$$A^- \boldsymbol{f}_{(t)} = \boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)}. \tag{3.2}$$

Note that for any leaf facility j, the outflow is 0, and this is forced by the outflow constraints since A_j^- is a zero vector. For each non-leaf facility j, the outflow from it cannot exceed the current inventory level, i.e., $x_{j,(t)} - \bar{x}_{j,(t)} \leq x_{j,(t)}$, or

$$\bar{x}_{j,(t)} \geq 0, \forall j \in \mathcal{L}^c. \tag{3.3}$$

In addition, we have the constraints

$$\bar{\boldsymbol{w}}_{(t)} \geq 0 \tag{3.4}$$

and

$$\boldsymbol{f}_{(t)} \geq 0 \tag{3.5}$$

which require that the ordering quantity and all flows must be non-negative.

Thus, the problem can be formulated as a dynamic program (DP)

$$
\begin{aligned}
V_{(t)}(\boldsymbol{x}_{(t)}) = \min_{\bar{\boldsymbol{x}}_{(t)}, \hat{\boldsymbol{x}}_{(t)}, \bar{\boldsymbol{w}}_{(t)}, \boldsymbol{f}_{(t)}} & \Big\{ \boldsymbol{c}_{(t)}^{\top} \bar{\boldsymbol{w}}_{(t)} + \sum_{j \in \mathcal{L}^c} h_{j,(t)} \bar{x}_{j,(t)} + \sum_{j \in \mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) \\
& + \boldsymbol{\gamma}_{(t)}^{\top} \boldsymbol{f}_{(t)} + \mathbb{E}\left[V_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}) \right] \Big\} \\
\text{s.t.} \quad & A^{+} \boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)} && (3.1) \\
& A^{-} \boldsymbol{f}_{(t)} = \boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)} && (3.2) \\
& \bar{x}_{j,(t)} \geq 0, \ \forall j \in \mathcal{L}^c && (3.3) \\
& \bar{\boldsymbol{w}}_{(t)} \geq 0 && (3.4) \\
& \boldsymbol{f}_{(t)} \geq 0, && (3.5)
\end{aligned}
$$

for $t = 0, \ldots, T-1$ and $V_{(T)}(\boldsymbol{x}_{(T)}) = 0$, where

$$
Q_{j,(t)}(y) \equiv \mathbb{E}[h_{j,(t+1)}(y - u_{j,(t)} - u_{j,(t+1)})^{+} + p_{j,(t+1)}(y - u_{j,(t)} - u_{j,(t+1)})^{-}]
$$

denotes the expected holding and backlogging costs in period $t+1$ given that the inventory position in period t is y.

It can be easily shown by a standard backwards induction argument that the optimization problem in each period of this dynamic program is a convex problem, since $Q(\cdot)$ is a convex function and all constraints are linear constraints. However, solving this DP directly is computationally intractable because the dimension of the state space is J, the number of facilities in the network, which can be easily scaled up. Instead, we seek a good approximate solution to this DP that is computationally efficient.

3.3 Lower Bound Relaxation

In this section, we provide a lower bound approximation for the original dynamic program which can be solved efficiently. The idea is similar to the one proposed by Federgruen et al. (2018), where one removes the dependence of the state on the components of the inventory vector $\boldsymbol{x}_{(t)}$, and makes the state depend only on the aggregate inventory positions so that the dynamic program becomes one-dimensional.

3.3.1 Lagrangian Relaxation

We dualize constraint (3.2) for each period t, which is the only constraint containing a state variable. More precisely, in each period, we remove constraint (3.2), and for any possible values of the state $\boldsymbol{x}_{(t)}$, we add to the objective a penalty term $\boldsymbol{\lambda}_{(t)}^{\top}\Big(A^{-}\boldsymbol{f}_{(t)} - (\boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)})\Big)$ with a fixed Lagrange multiplier $\boldsymbol{\lambda}_{(t)}$. Note that, this is different from using a state-dependent Lagrange multiplier $\boldsymbol{\lambda}_{(t)}(\boldsymbol{x}_{(t)})$ which could take different values in different states. Since the state space is continuous, there will be infinite number of constraints (3.2), each for a possible state. Thus, for a state dependent Lagrange multiplier, the dimension of the vector is infinite, and the corresponding Lagrange relaxation problem is intractable. As a result, we restrict ourselves to a fixed value of $\boldsymbol{\lambda}_{(t)}$ for all possible values of $\boldsymbol{x}_{(t)}$.

In addition, summing both sides of constraint (3.1) and (3.2), we have

$$\bar{x}_{+,(t)} + \hat{x}_{+,(t)} = x_{+,(t)} \tag{3.6}$$

where $x_{+,(t)} := \sum_{j\in\mathbf{V}} x_{j,(t)}$ represents the total inventory positions in all facilities in the system. Including constraint (3.6), we now have the relaxed DP given by

$$\begin{aligned}
V_{(t)}^{\lambda}(\boldsymbol{x}_{(t)}) = \min_{\bar{\boldsymbol{x}}_{(t)},\hat{\boldsymbol{x}}_{(t)},\bar{\boldsymbol{w}}_{(t)},\boldsymbol{f}_{(t)}} & \Bigg\{\boldsymbol{c}_{(t)}^{\top}\bar{\boldsymbol{w}}_{(t)} + \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) \\
& + \boldsymbol{\gamma}_{(t)}^{\top}\boldsymbol{f}_{(t)} + \boldsymbol{\lambda}_{(t)}^{\top}\Big(A^{-}\boldsymbol{f}_{(t)} - (\boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)})\Big) \\
& + \mathbb{E}\left[V_{(t+1)}^{\lambda}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)})\right]\Bigg\} \\
\text{s.t.} \quad & A^{+}\boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)} && (3.1) \\
& \bar{x}_{j,(t)} \geq 0,\ \forall j \in \mathcal{L}^c && (3.3) \\
& \bar{\boldsymbol{w}}_{(t)} \geq 0 && (3.4) \\
& \boldsymbol{f}_{(t)} \geq 0 && (3.5) \\
& \bar{x}_{+,(t)} + \hat{x}_{+,(t)} = x_{+,(t)}, && (3.6)
\end{aligned}$$

for $t = 0, \dots, T-1$ and $V^{\lambda}_{(T)}(\boldsymbol{x}_{(T)}) = 0$.

Now the inventory vector only appears in the objective. To further remove the dependence on the elements of $\boldsymbol{x}_{(t)}$, we define:

$$\bar{V}^{\lambda}_{(t)}(\boldsymbol{x}_{(t)}) := V^{\lambda}_{(t)}(\boldsymbol{x}_{(t)}) + \boldsymbol{\lambda}^{\top}_{(t)}\boldsymbol{x}_{(t)} \tag{3.7}$$

for $t = 0, \dots, T$, where $\boldsymbol{\lambda}_{(T)} = \mathbf{0}$. Replacing $V^{\lambda}_{(t)}(\cdot)$ with $\bar{V}^{\lambda}_{(t)}(\cdot)$, we have

$$\begin{aligned}
\bar{V}^{\lambda}_{(t)}(\boldsymbol{x}_{(t)}) = \min_{\bar{\boldsymbol{x}}_{(t)}, \hat{\boldsymbol{x}}_{(t)}, \bar{\boldsymbol{w}}_{(t)}, \boldsymbol{f}_{(t)}} & \Bigg\{ \sum_{j \in \mathcal{L}^c} h_{j,(t)} \bar{x}_{j,(t)} + \sum_{j \in \mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) + (\boldsymbol{c}_{(t)} - \boldsymbol{\lambda}_{(t+1)})^{\top} \bar{\boldsymbol{w}}_{(t)} \\
& + (\boldsymbol{\lambda}_{(t)} - \boldsymbol{\lambda}_{(t+1)})^{\top} \bar{\boldsymbol{x}}_{(t)} - \boldsymbol{\lambda}^{\top}_{(t+1)} \hat{\boldsymbol{x}}_{(t)} + (\boldsymbol{\gamma}_{(t)} + A^{-\top} \boldsymbol{\lambda}_{(t)})^{\top} \boldsymbol{f}_{(t)} \\
& + \boldsymbol{\lambda}^{\top}_{(t+1)} \mathbb{E}\left[\boldsymbol{u}_{(t)}\right] + \mathbb{E}\left[\bar{V}^{\lambda}_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)})\right] \Bigg\} \\
\text{s.t.} \quad & \text{constraints (3.1)}, (3.3) - (3.6),
\end{aligned}$$

for $t = 0, \dots, T-1$ and $\bar{V}^{\lambda}_{(T)}(\boldsymbol{x}_{(T)}) = 0$.

Now, the state vector $\boldsymbol{x}_{(t)}$ only enters this DP through constraint (3.6) in its aggregate form $x_{+,(t)}$. Thus, we can transform this dynamic program into a one-dimensional DP with aggregate inventory positions in the system $x_{+,(t)}$ being the state as follows.

$$\begin{aligned}
\bar{V}^{\lambda}_{(t)}(x_{+,(t)}) = \min_{\bar{\boldsymbol{x}}_{(t)}, \hat{\boldsymbol{x}}_{(t)}, \bar{\boldsymbol{w}}_{(t)}, \boldsymbol{f}_{(t)}} & \Bigg\{ \sum_{j \in \mathcal{L}^c} h_{j,(t)} \bar{x}_{j,(t)} + \sum_{j \in \mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) + (\boldsymbol{c}_{(t)} - \boldsymbol{\lambda}_{(t+1)})^{\top} \bar{\boldsymbol{w}}_{(t)} \\
& + (\boldsymbol{\lambda}_{(t)} - \boldsymbol{\lambda}_{(t+1)})^{\top} \bar{\boldsymbol{x}}_{(t)} - \boldsymbol{\lambda}^{\top}_{(t+1)} \hat{\boldsymbol{x}}_{(t)} + (\boldsymbol{\gamma}_{(t)} + A^{-\top} \boldsymbol{\lambda}_{(t)})^{\top} \boldsymbol{f}_{(t)} \\
& + \boldsymbol{\lambda}^{\top}_{(t+1)} \mathbb{E}\left[\boldsymbol{u}_{(t)}\right] + \mathbb{E}\left[\bar{V}^{\lambda}_{(t+1)}(\bar{w}_{(t)} + x_{+,(t)} - u_{+,(t)})\right] \Bigg\} \\
\text{s.t.} \quad & \text{constraints (3.1)}, (3.3) - (3.6),
\end{aligned}$$

for $t = 0, \dots, T-1$ and $\bar{V}^{\lambda}_{(T)}(\boldsymbol{x}_{(T)}) = 0$.

It is then easy to see that this DP can be computed by solving the following recursion for the value function

$\bar{V}^{\lambda}_{(t)}(x_{+,(t)})$.

$$
\begin{aligned}
\bar{V}^{\lambda}_{(t)}(x_{+,(t)}) &= R_{(t)}(x_{+,(t)}) \\
&\quad + \min_{\bar{w}_{(t)}\geq 0} \left\{ (c_{(t)} - \lambda_{1,(t+1)})\bar{w}_{(t)} + \mathbb{E}\left[\bar{V}^{\lambda}_{(t+1)}(\bar{w}_{(t)} + x_{+,(t)} - u_{+,(t)})\right] \right\} \\
&\quad + \boldsymbol{\lambda}^{\top}_{(t+1)}\mathbb{E}\left[\boldsymbol{u}_{(t)}\right]
\end{aligned}
$$

where

$$
\begin{aligned}
R_{(t)}(a) = \min_{\bar{\boldsymbol{x}}_{(t)},\hat{\boldsymbol{x}}_{(t)},\boldsymbol{f}_{(t)}} & \Bigg\{ \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) \\
& \quad + (\boldsymbol{\lambda}_{(t)} - \boldsymbol{\lambda}_{(t+1)})^{\top}\bar{\boldsymbol{x}}_{(t)} - \boldsymbol{\lambda}^{\top}_{(t+1)}\hat{\boldsymbol{x}}_{(t)} + (\boldsymbol{\gamma}_{(t)} + A^{-\top}\boldsymbol{\lambda}_{(t)})^{\top}\boldsymbol{f}_{(t)} \Bigg\} \\
\text{s.t.} \quad & A^{+}\boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)} \\
& \bar{x}_{j,(t)} \geq 0,\ \forall j \in \mathcal{L}^c \\
& \boldsymbol{f}_{(t)} \geq 0 \\
& \bar{x}_{+,(t)} + \hat{x}_{+,(t)} = a,
\end{aligned}
\tag{3.8}
$$

for $t = 0, \ldots, T-1$ and $\bar{V}_{(T)}(x_{+,(T)}) = 0$. That is, in each period t, the optimization problem is decomposed into two sub-problems. One is the ordering problem, which is one-dimensional, deciding on the amount of products bought into the root facility, $\bar{w}_{(t)}$. The other is the allocation problem deciding on the distribution of the current inventory at each facility into their immediate downstream facilities, with the decision variables $\boldsymbol{f}_{(t)}, \bar{\boldsymbol{x}}_{(t)}$ and $\hat{\boldsymbol{x}}_{(t)}$.

Since $Q(\cdot)$ is convex, it is obvious that the allocation problem is a convex problem. Moreover, again by a backwards induction argument, the ordering problem in each period is also a convex problem. Thus, the two sub-problems in each period can be efficiently solved to optimality.

3.3.2 Lagrangian Dual

To find the largest lower bound, we solve the Lagrangian dual problem

$$\max_{\boldsymbol{\lambda}} V_{(0)}^{\lambda}(\boldsymbol{x}_{(0)}) \tag{3.9}$$

where $V_{(0)}^{\lambda}(\boldsymbol{x}_{(0)}) \equiv \bar{V}_{(0)}^{\lambda}(\boldsymbol{x}_{(0)}) - \boldsymbol{\lambda}_{(0)}^{\top}\boldsymbol{x}_{(0)}$.

To solve this maximization problem, we need the following theorem.

Theorem 9. *(a)* $V_{(0)}^{\lambda}(\boldsymbol{x}_{(0)})$ *is a concave function of* $\boldsymbol{\lambda}$.

(b) For any t, $\nabla_{\boldsymbol{\lambda}_{(t)}} V_{(0)}^{\lambda}(\boldsymbol{x}_{(0)})$ *is a supergradient of* $V_{(0)}^{\lambda}(\boldsymbol{x}_{(0)})$ *with respect to* $\boldsymbol{\lambda}_{(t)}$*, where*

$$\nabla_{\boldsymbol{\lambda}_{(t)}} V_{(0)}^{\lambda}(\boldsymbol{x}_{(0)}) = \mathbb{E}\left[A^{-}\boldsymbol{f}_{(t)} - (\boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)})|\boldsymbol{x}_{(0)}\right].$$

Proof. See Appendix C.1. □

Thus, the maximization problem can be solved to within any precision of optimality by employing any gradient methods, e.g., FISTA (Beck and Teboulle, 2009).

3.4 Upper Bound Heuristic

In this section, we propose a feasible heuristic that can be computed efficiently. The expected cost under this heuristic provides an *upper* bound on the optimal cost value. In the next section, we conduct a numerical study to evaluate the suboptimality of this heuristic by comparing the gap between this upper bound and the lower bound in Section 3.3.

3.4.1 A Myopic Policy

Recall that the exact dynamic program is

$$
\begin{aligned}
V_{(t)}(\boldsymbol{x}_{(t)}) = \min_{\bar{\boldsymbol{x}}_{(t)}, \hat{\boldsymbol{x}}_{(t)}, \bar{\boldsymbol{w}}_{(t)}, \boldsymbol{f}_{(t)}} & \Bigg\{ \boldsymbol{c}_{(t)}^{\top} \bar{\boldsymbol{w}}_{(t)} + \sum_{j \in \mathcal{L}^c} h_{j,(t)} \bar{x}_{j,(t)} + \sum_{j \in \mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) \\
& + \boldsymbol{\gamma}_{(t)}^{\top} \boldsymbol{f}_{(t)} + \mathbb{E}\left[V_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}) \right] \Bigg\} \\
\text{s.t.} \quad & A^{+} \boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)} && (3.1) \\
& A^{-} \boldsymbol{f}_{(t)} = \boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)} && (3.2) \\
& \bar{x}_{j,(t)} \geq 0, \ \forall j \in \mathcal{L}^c && (3.3) \\
& \bar{\boldsymbol{w}}_{(t)} \geq 0 && (3.4) \\
& \boldsymbol{f}_{(t)} \geq 0, && (3.5)
\end{aligned}
$$

for $t = 0, \ldots, T-1$ and $V_{(T)}(\boldsymbol{x}_{(T)}) = 0$.

The exact value function $V_{(t+1)}(\cdot)$ cannot be computed. Then, a myopic policy is, in each period t, to ignore the expected future value $\mathbb{E}\left[V_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}) \right]$, and solve the following (convex) single-period optimization problem.

$$
\begin{aligned}
\min_{\bar{\boldsymbol{x}}_{(t)}, \hat{\boldsymbol{x}}_{(t)}, \bar{\boldsymbol{w}}_{(t)}, \boldsymbol{f}_{(t)}} & \left\{ \boldsymbol{c}_{(t)}^{\top} \bar{\boldsymbol{w}}_{(t)} + \sum_{j \in \mathcal{L}^c} h_{j,(t)} \bar{x}_{j,(t)} + \sum_{j \in \mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) + \boldsymbol{\gamma}_{(t)}^{\top} \boldsymbol{f}_{(t)} \right\} \\
\text{s.t.} \quad & \text{constraints (3.1)} - \text{(3.5)}.
\end{aligned}
$$

This policy minimizes the immediate cost in each period and ignores the future value of the state, and thus could find a solution that leads to huge costs in future periods. Next, we propose a heuristic that takes into account the value of the state by approximating the value function $V_{(t+1)}(\cdot)$.

3.4.2 An Approximate Look-Ahead Policy

Recall that in solving the relaxed DP, we have computed a relaxed value function $V_{(t)}^{\lambda}(\cdot)$. Thus, $V_{(t)}^{\lambda}(\cdot)$ can be used as an approximation for the true value function $V_{(t)}(\cdot)$. This gives us an algorithm to compute another feasible policy. That is, in each period t, we solve

$$\begin{aligned}\min_{\bar{\boldsymbol{x}}_{(t)},\hat{\boldsymbol{x}}_{(t)},\bar{\boldsymbol{w}}_{(t)},\boldsymbol{f}_{(t)}} \Bigg\{ & \boldsymbol{c}_{(t)}^\top \bar{\boldsymbol{w}}_{(t)} + \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) + \boldsymbol{\gamma}_{(t)}^\top \boldsymbol{f}_{(t)} \\ & + \mathbb{E}\left[V^{\lambda}_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)})\right]\Bigg\} \\ \text{s.t.} \quad & \text{constraints (3.1)} - \text{(3.5)}.\end{aligned}$$

Applying again (3.7) to replace $V^{\lambda}_{(t+1)}(\cdot)$ with $\bar{V}^{\lambda}_{(t+1)}(\cdot)$, we get

$$\begin{aligned}\min_{\bar{\boldsymbol{x}}_{(t)},\hat{\boldsymbol{x}}_{(t)},\bar{\boldsymbol{w}}_{(t)},\boldsymbol{f}_{(t)}} \Bigg\{ & \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) + (\boldsymbol{c}_{(t)} - \boldsymbol{\lambda}_{(t+1)})^\top \bar{\boldsymbol{w}}_{(t)} \\ & - \boldsymbol{\lambda}_{(t+1)})^\top \bar{\boldsymbol{x}}_{(t)} - \boldsymbol{\lambda}_{(t+1)}^\top \hat{\boldsymbol{x}}_{(t)} + \boldsymbol{\gamma}_{(t)}^\top \boldsymbol{f}_{(t)} \\ & + \boldsymbol{\lambda}_{(t+1)}^\top \mathbb{E}\left[\boldsymbol{u}_{(t)}\right] + \mathbb{E}\left[\bar{V}^{\lambda}_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)})\right]\Bigg\} \\ \text{s.t.} \quad & \text{constraints (3.1)} - \text{(3.5)}\end{aligned}$$

Recall that $\bar{V}^{\lambda}_{(t+1)}(\cdot)$ is only a function of the aggregate inventory positions in the system. Thus, $\bar{V}^{\lambda}_{(t+1)}(\bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{u}_{(t)}) \equiv \bar{V}_{(t+1)}(\bar{w}_{(t)} + x_{+,(t)} - u_{+,(t)})$, and does not depend on the decision variables $\bar{\boldsymbol{x}}_{(t)}$ and $\hat{\boldsymbol{x}}_{(t)}$ given the current state $x_{+,(t)}$. Thus, this problem can again be decomposed into two problems, an ordering problem

$$\min_{\bar{w}_{(t)}\geq 0} \left\{(c_{(t)} - \lambda_{1,(t+1)})\bar{w}_{(t)} + \mathbb{E}\left[\bar{V}_{(t+1)}(\bar{w}_{(t)} + x_{+,(t)} - u_{+,(t)})\right]\right\}, \tag{3.10}$$

which has been computed in Lower Bound Relaxation Section 3.3, and an allocation problem

$$\min_{\bar{\boldsymbol{x}}_{(t)},\hat{\boldsymbol{x}}_{(t)},\boldsymbol{f}_{(t)}} \left\{ \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) - \boldsymbol{\lambda}_{(t+1)}^\top \bar{\boldsymbol{x}}_{(t)} - \boldsymbol{\lambda}_{(t+1)}^\top \hat{\boldsymbol{x}}_{(t)} + \boldsymbol{\gamma}_{(t)}^\top \boldsymbol{f}_{(t)} \right\} \quad (3.11)$$

$$\begin{aligned} \text{s.t.}\quad & A^+ \boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)} \\ & A^- \boldsymbol{f}_{(t)} = \boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)} \qquad (3.2) \\ & \bar{x}_{j,(t)} \geq 0,\ \forall j \in \mathcal{L}^c \\ & \boldsymbol{f}_{(t)} \geq 0. \end{aligned}$$

Note that the only difference between this allocation problem (3.11) and the one in Lower Bound Relaxation Section is that we now add back the outflow constraints (3.2).

3.4.3 A Regularized Look-Ahead Policy

The look-ahead policy proposed in the previous section prescribes the ordering quantity in the relaxed problem that solves (3.10). This problem could admit multiple optima, or the objective function could be very flat curve relatively insensitive to the choice of the ordering quantity $\bar{w}_{(t)}$. The intuition is that, in the relaxed problem, by a proper choice of the Lagrange multiplier, one might be indifferent, for example, between two strategies:

(1) make a large order in the current period, and ship it downstream along valid routes in future periods; or

(2) make a zero or small order in the current period, and in future periods, make a large order and directly transfer it to the final destination without going through any intermediate facilities, at a penalty cost of violating the flow constraints.

However, strategy (2) is not feasible to the real problem since the products cannot be shipped along invalid routes. Therefore, making a small order in the current period is not a desired strategy. To rule out such suboptimal solutions, we consider adding regularization terms to both the ordering problem (3.10) and allocation problem (3.11).

The regularized ordering problem is given by

$$\min_{\bar{w}_{(t)}\geq 0}\left\{(c_{(t)}-\boldsymbol{\lambda}_{1,(t+1)})\bar{w}_{(t)}+\mathbb{E}\left[\bar{V}^{\lambda}_{(t+1)}(\bar{w}_{(t)}+x_{+,(t)}-u_{+,(t)})\right]+\alpha_1\left|\bar{w}_{(t)}-\bar{w}^{R}_{(t)}\right|\right\},$$

where $\bar{w}^{R}_{(t)}$ is the target ordering quantity to regularize with.

Similarly, the regularized allocation problem is given by

$$\begin{aligned}
\min_{\bar{\boldsymbol{x}}_{(t)},\hat{\boldsymbol{x}}_{(t)},\boldsymbol{f}_{(t)}} & \left\{\sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)}+\sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)}+\hat{x}_{j,(t)})-\boldsymbol{\lambda}^{\top}_{(t+1)}\bar{\boldsymbol{x}}_{(t)}-\boldsymbol{\lambda}^{\top}_{(t+1)}\hat{\boldsymbol{x}}_{(t)}\right. \\
& \left.+\boldsymbol{\gamma}^{\top}_{(t)}\boldsymbol{f}_{(t)}+\alpha_2\mathbf{1}^{\top}\left|\boldsymbol{f}_{(t)}-\boldsymbol{f}^{R}_{(t)}\right|\right\} \\
\text{s.t.}\quad & A^{-}\boldsymbol{f}_{(t)}=\boldsymbol{x}_{(t)}-\bar{\boldsymbol{x}}_{(t)} \\
& A^{+}\boldsymbol{f}_{(t)}=\hat{\boldsymbol{x}}_{(t)} \\
& \bar{x}_{j,(t)}\geq 0,\ \forall j\in\mathcal{L}^c \\
& \boldsymbol{f}_{(t)}\geq \mathbf{0},
\end{aligned}$$

where $\boldsymbol{f}^{R}_{(t)}$ is the target flow quantity to regularize with and $\mathbf{1}$ is a vector of ones.

α_1 and α_2 are some hyper-parameters that need to be tuned. The best values can be determined by simulating the policy with different values of parameters and picking the ones with the lowest average cost.

To determine the target quantities $\bar{w}^{R}_{(t)}$ and $\boldsymbol{f}^{R}_{(t)}$, we considered the robust optimization approach proposed by Ardestani-Jaafari and Delage (2016). Although the approach works for general demand distributions, we focus on the case when the demands follow a multivariate Normal distribution, since, correlated demands are of more practical interest, and multivariate Normal distribution is the most widely used distribution that describes correlated random variables.

Assume the random demand vector $\boldsymbol{u}_{(t)}\sim\mathcal{N}(\boldsymbol{\mu}_{(t)},\Sigma_{(t)})$. We write $\boldsymbol{u}_{(t)}=\boldsymbol{\mu}_{(t)}+C_{(t)}\boldsymbol{z}_{(t)}$, where $C_{(t)}=$

$\Sigma_{(t)}^{-\frac{1}{2}}$ and $\boldsymbol{z}_{(t)} \sim \mathcal{N}(0, \mathbf{I}_{|\mathcal{L}|})$. Then, the inventory level at facility location j at the end of period t is given by

$$
\begin{aligned}
x_{j,(t)} &= x_{j,(0)} + \sum_{\tau=0}^{t-1}(\bar{w}_{j,(\tau)} + \hat{x}_{j,(\tau)}) - \sum_{\tau=0}^{t} A_j^{-} f_{(\tau)}, && \forall j \in \mathcal{L}^c, \\
x_{j,(t)} &= x_{j,(0)} + \sum_{\tau=0}^{t-1}(\bar{w}_{j,(\tau)} + \hat{x}_{j,(\tau)}) - \sum_{\tau=0}^{t}(\mu_{j,(\tau)} + C_{j,(\tau)}\boldsymbol{z}_{(\tau)}), && \forall j \in \mathcal{L},
\end{aligned}
$$

where $C_{j,(\tau)}$ is the j-th row of $C_{(\tau)}$. The holding/back-ordering cost at a leaf facility j is thus given by

$$
\begin{aligned}
Y_{j,(t)}(\boldsymbol{z}) &= \max\left\{h_{j,(t)}x_{j,(t)}, -p_{j,(t)}x_{j,(t)}\right\} \\
&= \max\left\{\boldsymbol{a}_{1,j,(t)}^{\top}\boldsymbol{z} + b_{1,j,(t)}, \boldsymbol{a}_{2,j,(t)}^{\top}\boldsymbol{z} + b_{2,j,(t)}\right\}
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{z} &= \left(\boldsymbol{z}_{(0)}^{\top}, \ldots, \boldsymbol{z}_{(T-1)}^{\top}\right)^{\top} \\
\boldsymbol{a}_{1,j,(t)} &= -h_{j,(t)}\left(C_{j,(0)}, \ldots, C_{j,(t)}, \mathbf{0}^{\top}, \ldots, \mathbf{0}^{\top}\right) && (3.12) \\
b_{1,j,(t)} &= h_{j,(t)}\left(x_{j,(0)} + \sum_{\tau=0}^{t-1}(\bar{w}_{j,(\tau)} + \hat{x}_{j,(\tau)}) - \sum_{\tau=0}^{t}\mu_{j,(\tau)}\right) && (3.13) \\
\boldsymbol{a}_{2,j,(t)} &= p_{j,(t)}\left(C_{j,(0)}, \ldots, C_{j,(t)}, \mathbf{0}^{\top}, \ldots, \mathbf{0}^{\top}\right) && (3.14) \\
b_{2,j,(t)} &= -p_{j,(t)}\left(x_{j,(0)} + \sum_{\tau=0}^{t-1}(\bar{w}_{j,(\tau)} + \hat{x}_{j,(\tau)}) - \sum_{\tau=0}^{t}\mu_{j,(\tau)}\right). && (3.15)
\end{aligned}
$$

The worst-case total holding/back-ordering cost at the leaf facilities is given by

$$
\begin{aligned}
\max_{\boldsymbol{z}} \quad & \sum_{t=0}^{T-1}\sum_{j \in \mathcal{L}} Y_{j,(t)}(\boldsymbol{z}) && (3.16) \\
\text{s.t.} \quad & \boldsymbol{z} \in \mathcal{Z} := \left\{\boldsymbol{z} : -\mathbf{1} \le \boldsymbol{z} \le \mathbf{1}, \sum_{t=0}^{\tau}\|\boldsymbol{z}_{(t)}\|_1 \le \Gamma_{(\tau)}, \forall \tau \in \{0, \ldots, T-1\}\right\},
\end{aligned}
$$

where $\mathcal{Z}$ is the set of admissible deviations, and Γ is a parameter controlling the deviations. Note that the objective in (3.16) is a piecewise linear *convex* function, and therefore, it is not a convex optimization problem. However, the convex objective can be linearized by introducing binary variables as follows.

All the terms $Y_{j,(t)}(\boldsymbol{z})$ are of the form $\max\{\boldsymbol{a}_1^\top \boldsymbol{z} + b_1, \boldsymbol{a}_2^\top \boldsymbol{z} + b_2\}$. Let $\boldsymbol{z} := \boldsymbol{z}^+ - \boldsymbol{z}^-$ for $\boldsymbol{z}^\pm \geq \mathbf{0}$, and introduce binary variables $\boldsymbol{y} \in \{0,1\}^2$. Then

$$\begin{aligned}
Y(\boldsymbol{z}) &= \max\{\boldsymbol{a}_1^\top \boldsymbol{z} + b_1, \boldsymbol{a}_2^\top \boldsymbol{z} + b_2\}, \\
&= \max_{\{\boldsymbol{y}\in\{0,1\}^2 : y_1+y_2=1\}} \Big\{ \sum_{k=1}^{2} y_k(\boldsymbol{a}_k^\top(\boldsymbol{z}^+ - \boldsymbol{z}^-) + b_k) \Big\} \\
&= \max_{\{\boldsymbol{y}\in\{0,1\}^2 : y_1+y_2=1\}} \Big\{ \sum_{k=1}^{2} \boldsymbol{a}_k^\top \boldsymbol{z}^+ y_k - \boldsymbol{a}_k^\top \boldsymbol{z}^- y_k + b_k y_k \Big\}.
\end{aligned}$$

Let $\boldsymbol{\Delta}_k^+ := y_k \boldsymbol{z}^+$ and $\boldsymbol{\Delta}_k^- := y_k \boldsymbol{z}^-$. Then, the maximization problem (3.16) can be reformulated as the following MIP.

$$\max_{\boldsymbol{z}^+,\boldsymbol{z}^-,\boldsymbol{y},\boldsymbol{\Delta}^+,\boldsymbol{\Delta}^-} \sum_{t=0}^{T-1} \sum_{j\in\mathcal{L}} \sum_{k=1}^{2} \boldsymbol{a}_{k,j,(t)}^\top(\boldsymbol{\Delta}_{k,j,(t)}^+ - \boldsymbol{\Delta}_{k,j,(t)}^-) + b_{k,j,(t)} y_{k,j,(t)} \tag{3.17}$$

$$\begin{aligned}
\text{s.t.} \quad & \boldsymbol{z}^+ + \boldsymbol{z}^- \leq \mathbf{1}, && && (3.18)\\
& \mathbf{1}_{(0-t)}^\top(\boldsymbol{z}^+ + \boldsymbol{z}^-) \leq \Gamma_{(t)}, && \forall t && (3.19)\\
& \boldsymbol{y}_1 + \boldsymbol{y}_2 = \mathbf{1}, && && (3.20)\\
& \boldsymbol{\Delta}_{1,j,(t)}^+ + \boldsymbol{\Delta}_{2,j,(t)}^+ = \boldsymbol{z}^+, && \forall j,t && (3.21)\\
& \boldsymbol{\Delta}_{1,j,(t)}^- + \boldsymbol{\Delta}_{2,j,(t)}^- = \boldsymbol{z}^-, && \forall j,t && (3.22)\\
& \boldsymbol{\Delta}_{k,j,(t)}^+ + \boldsymbol{\Delta}_{k,j,(t)}^- \leq y_{k,j,(t)}, && \forall k,j,t && (3.23)\\
& \mathbf{1}_{(0-t')}^\top(\boldsymbol{\Delta}_{k,j,(t)}^+ + \boldsymbol{\Delta}_{k,j,(t)}^-) \leq \Gamma_{(t')} y_{k,j,(t)}, && \forall k,j,t,t' \leq t && (3.24)\\
& \boldsymbol{z}^+, \boldsymbol{z}^-, \boldsymbol{\Delta}^+, \boldsymbol{\Delta}^- \geq 0, && && (3.25)\\
& \boldsymbol{y} \in \{0,1\}^{2\times|\mathcal{L}|\times T}, && && (3.26)
\end{aligned}$$

where $\mathbf{1}_{(0-t)} := \left(\mathbf{1}_{(0)}^\top, \ldots, \mathbf{1}_{(t)}^\top, \mathbf{0}^\top, \ldots, \mathbf{0}^\top\right)^\top$. We remove the binary constraint (3.26) to obtain a linear relaxation of this MIP which admits a feasible solution where all variables takes 0. Thus, we can formulate the dual problem of the linear relaxation and strong duality holds. Defining dual variables $\boldsymbol{\xi}, \boldsymbol{\nu}, \boldsymbol{\pi}, \boldsymbol{\lambda}^+, \boldsymbol{\lambda}^-, \boldsymbol{\psi}, \boldsymbol{\theta}$

corresponding to the constraints (3.18) - (3.24), respectively, we have the dual problem:

$$\min_{\boldsymbol{\xi},\boldsymbol{\nu},\boldsymbol{\pi},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\boldsymbol{\psi},\boldsymbol{\theta}} \quad \mathbf{1}^\top\boldsymbol{\xi} + \boldsymbol{\Gamma}^\top\boldsymbol{\nu} + \mathbf{1}^\top\boldsymbol{\pi} \tag{3.27}$$

$$\text{s.t.} \quad \boldsymbol{\xi} + \sum_{t=0}^{T-1} \nu_{(t)}\mathbf{1}_{(0:t)} - \sum_{t=0}^{T-1}\sum_{j\in\mathcal{L}} \boldsymbol{\lambda}^+_{j,(t)} \geq 0, \tag{3.28}$$

$$\boldsymbol{\xi} + \sum_{t=0}^{T-1} \nu_{(t)}\mathbf{1}_{(0:t)} - \sum_{t=0}^{T-1}\sum_{j\in\mathcal{L}} \boldsymbol{\lambda}^-_{j,(t)} \geq 0, \tag{3.29}$$

$$\boldsymbol{\pi}_{j,(t)} - \mathbf{1}^\top\boldsymbol{\psi}_{k,j,(t)} - \sum_{t'=0}^{t} \Gamma_{(t')}\theta_{k,j,(t),t'} \geq b_{k,j,(t)}, \qquad \forall k,j,t \tag{3.30}$$

$$\boldsymbol{\lambda}^+_{j,(t)} + \boldsymbol{\psi}_{k,j,(t)} + \sum_{t'=0}^{t} \theta_{k,j,(t),(t')}\mathbf{1}_{(0:t')} \geq \boldsymbol{a}_{k,j,(t)}, \qquad \forall k,j,t \tag{3.31}$$

$$\boldsymbol{\lambda}^-_{j,(t)} + \boldsymbol{\psi}_{k,j,(t)} + \sum_{t'=0}^{t} \theta_{k,j,(t),(t')}\mathbf{1}_{(0:t')} \geq -\boldsymbol{a}_{k,j,(t)}, \qquad \forall k,j,t \tag{3.32}$$

$$\boldsymbol{\xi}, \nu, \boldsymbol{\psi}, \boldsymbol{\theta} \geq 0, \tag{3.33}$$

where constraints (3.28) - (3.32) correspond to the primal variables $\boldsymbol{z}^+, \boldsymbol{z}^-, \boldsymbol{y}, \boldsymbol{\Delta}^+, \boldsymbol{\Delta}^-$ of the MIP, respectively. Injecting this problem into the problem of minimizing all costs, we have the robust optimization formulation for the whole problem.

In summary, in any period τ, given the current state $\boldsymbol{x}_{(\tau)}$, the robust optimization approach solves the following linear programming problem.

$$
\begin{aligned}
\min_{\bar{\boldsymbol{w}},\hat{\boldsymbol{x}},\boldsymbol{f},\boldsymbol{\xi},\nu,\boldsymbol{\pi},\boldsymbol{\lambda}^+,\boldsymbol{\lambda}^-,\psi,\boldsymbol{\theta}} \quad & \sum_{t=\tau}^{T-1}\left\{\boldsymbol{c}_{(t)}^\top \bar{\boldsymbol{w}}_{(t)} + \boldsymbol{\gamma}^\top \boldsymbol{f}_{(t)} + \sum_{j\in\mathcal{L}^c} h_{j,(t)}\Big(x_{j,(\tau)} + \sum_{t'=\tau}^{t-1}(\bar{w}_{j,(t')} + \hat{x}_{j,(t')}) - \sum_{t'=\tau}^{t} A_j^- f_{(t')}\Big)\right\} \\
& + \mathbf{1}^\top \boldsymbol{\xi} + \Gamma\nu + \mathbf{1}^\top \boldsymbol{\pi} \\
\text{s.t.} \quad & A^+ \boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)}, \qquad \forall t \geq \tau \\
& \boldsymbol{\xi} + \sum_{t=\tau}^{T-1} \nu_{(t)} \mathbf{1}_{(\tau:t)} - \sum_{t=\tau}^{T-1}\sum_{j\in\mathcal{L}} \boldsymbol{\lambda}_{j,(t)}^+ \geq 0, \\
& \boldsymbol{\xi} + \sum_{t=\tau}^{T-1} \nu_{(t)} \mathbf{1}_{(\tau:t)} - \sum_{t=\tau}^{T-1}\sum_{j\in\mathcal{L}} \boldsymbol{\lambda}_{j,(t)}^- \geq 0, \\
& \boldsymbol{\pi}_{j,(t)} - \mathbf{1}^\top \boldsymbol{\psi}_{k,j,(t)} - \sum_{t'=0}^{t} \Gamma_{(t')}\theta_{k,j,(t)} \geq b_{k,j,(t)}, \qquad \forall k\in\{1,2\}, j\in\mathcal{L}, t\geq\tau \\
& \boldsymbol{\lambda}_{j,(t)}^+ + \boldsymbol{\psi}_{k,j,(t)} + \sum_{t'=\tau}^{t} \theta_{k,j,(t),(t')}\mathbf{1}_{(\tau:t')} \geq \boldsymbol{a}_{k,j,(t)}, \qquad \forall k\in\{1,2\}, j\in\mathcal{L}, t\geq\tau \\
& \boldsymbol{\lambda}_{j,(t)}^- + \boldsymbol{\psi}_{k,j,(t)} + \sum_{t'=\tau}^{t} \theta_{k,j,(t),(t')}\mathbf{1}_{(\tau:t')} \geq -\boldsymbol{a}_{k,j,(t)}, \qquad \forall k\in\{1,2\}, j\in\mathcal{L}, t\geq\tau \\
& \boldsymbol{\xi},\nu,\boldsymbol{\psi},\boldsymbol{\theta} \geq 0,
\end{aligned}
$$

where $\boldsymbol{a}_{k,j,(t)}$ and $\boldsymbol{b}_{k,j,(t)}$ are defined similarly to (3.12) - (3.15) except that the time index starts from τ instead of 0. Ardestani-Jaafari and Delage (2016) show that that the integrality gap of the fractional relaxation is small, and in fact, the fractional relaxation is exact in many cases.

3.5 Numerical Study

In this section, we describe the results of an extensive numerical study covering a total of 40,960 instances to assess the optimality gap between the lower bound in Section 3.3 and the upper bound given by the regularized look-ahead policy in Section 3.4.

3.5.1 Implementation Adjustments

In a practical business problem, the algorithms presented in both the lower bound and upper bound sections only need to be executed once, and thus can be implemented as exactly described. However, in a large-scale numerical study, some steps need to be compromised due to the limiting computing resource. In the following, we discuss how we make adjustments to the algorithms to reduce the needs for computing resource.

Optimizing over $\boldsymbol{\lambda}$

In the lower bound section, in order to find the largest lower bound, we need to maximize over the Lagrange multiplier $\boldsymbol{\lambda}$ by running gradient methods. In each iteration, a dynamic program needs to be solved, which involves plenty of optimization problems. In a practical setting, this can be achieved by parallelizing the optimization problems. However, in a numerical study with a large number of instances, the computing resource allocated to each instance is limited, and thus, large parallel computations within an instance is unavailable. To reduce the computation needs of calculating the lower bound for each instance, instead of running gradient methods to find the optimal Lagrangian multiplier, we find a good Lagrange multiplier by solving the following mean-path linear program.

$$\begin{aligned}
\min_{\bar{\boldsymbol{w}},\boldsymbol{x},\bar{\boldsymbol{x}},\hat{\boldsymbol{x}},\boldsymbol{f}} \sum_{t=0}^{T-1} \Bigg(& \boldsymbol{c}_{(t)}^{\top}\bar{\boldsymbol{w}}_{(t)} + \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} \Big(h_{j,(t+1)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)} - \mu_{j,(t)} - \mu_{j,(t+1)})^{+} \\
& + p_{j,(t+1)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)} - \mu_{j,(t)} - \mu_{j,(t+1)})^{-}\Big) + \boldsymbol{\gamma}_{(t)}^{\top}\boldsymbol{f}_{(t)}\Bigg) \\
\text{s.t.}\quad & \boldsymbol{x}_{(t+1)} = \bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{\mu}_{(t)},\ \forall t \\
& A^{+}\boldsymbol{f}_{(t)} = \hat{\boldsymbol{x}}_{(t)}, \forall t \\
& A^{-}\boldsymbol{f}_{(t)} = \boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)},\ \forall t \\
& \bar{x}_{j,(t)} \geq 0,\ \forall j \in \mathcal{L}^c,\ \forall t \\
& \boldsymbol{f}_{(t)} \geq \mathbf{0},\ \forall t.
\end{aligned} \tag{3.2}$$

That is, assuming that all future demands over the planning horizon are exactly at their means, we calculate the optimal ordering and allocation solutions for all future periods. Note that constraints (3.2) are the constraints we dualize in the Lagrangian relaxation. We set $\boldsymbol{\lambda}$ as the value of the dual variables corresponding to these constraints.

Tuning Hyper-Parameters α_1 and α_2

In the upper bound section, we tune the hyper-parameters α_1 and α_2 to find the best pair of values that gives the lowest cost. Again, this can be achieved in reality by parallelizing the computations which is not available in our extensive numerical study. Instead, we restrict ourselves to only two pairs of values, $(1, 1)$ and (∞, ∞). The latter choice means we are using the solution given by the robust optimization approach alone, which also enables us to compare our proposed policy with the robust optimization approach.

3.5.2 Numerical Study Design

In this section, we describe in details the design of our numerical study. We cover a wide spectrum of cost and distribution parameters that we believe are representative of what might be encountered in practice. In addition, we consider different structures of the graphs describing the inventory network.

Cost and Distribution Parameters

We consider a planning horizon of $T = 10$ periods. Table 3.1 summarizes the cost and distribution parameters used in our study of which we provide a detailed description as follows.

baseHoldCostRatio

Holding cost at each facility are likely to be different in reality. An upstream facility is usually closer to the product suppliers or factories which are commonly located in the outer suburbs, and thus have lower holding cost. On the contrary, a downstream facility tends to be closer to the customers, and thus are located in the city area where storage costs are higher. Thus, we consider varying holding cost at different facilities, with $baseHoldCostRatio$ denoting the ratio of unit holding cost at a facility to unit holding cost at its immediate upstream facility. $baseHoldCostRatio$ takes two values in our study, 1 and 1.3.

holdCostCV

For facilities at the same level, i.e., the same distance from the root facility, holding cost may be the same ($holdCostCV = 0$) or varying ($holdCostCV = 0.2$). If $holdCostCV = 0.2$, the unit holding cost at the same level follows a Normal distribution with coefficient of variation 0.2.

holdCostTimeRandomness

baseHoldCostRatio	$1, 1.3$	The ratio of holding cost at a node to holding cost at its parent.
holdCostCV	$0, 0.2$	The coefficient of variation of the holding cost across nodes at the same level.
holdCostTimeRandomness	C, S	Randomness of the holding cost across time.
backorderToHoldRatio	$10, 20$	The ratio of backordering cost to holding cost at a leaf node.
orderToHoldRatio	$5, 10$	The ratio of variable ordering cost to average holding cost at leaf nodes.
shipToHoldRatio	$1, 5$	The ratio of variable shipping cost to average holding cost at leaf nodes.
baseDemandCV	$0, 0.4$	The coefficient of variation of the demand means across leaf nodes.
demandTimeRandomness	C, S	Randomness of the demand means across time.
demandCV	$0.2, 0.4$	The coefficient of variation of the demand at each leaf node.
correlation	$-0.9, -0.5, 0, 0.5, 0.9$	The correlation coefficient between demands at different locations.

Table 3.1: Cost and distribution parameters used in numerical study.

The unit holding cost at each facility may stay the same ($holdCostTimeRandomness = C$) over time or be seasonal ($holdCostTimeRandomness = S$). If $holdCostTimeRandomness = S$, the holding cost at a facility j in period $t > 0$ is given by $h_{j,(t)} = h_{j,(0)}\big(1 + \frac{1}{2}\sin(2\pi\frac{t}{T})\big)$.

backorderToHoldRatio

This is the ratio of the unit backordering cost $p_{j,(t)}$ to unit holding cost $h_{j,(t)}$ at each facility j, taking two values 10 and 20.

orderToHoldRatio

This is the ratio of the unit ordering cost $c_{(t)}$ to the average unit holding cost at the leaf facilities $\frac{1}{|\mathcal{L}|}\sum_{j\in\mathcal{L}} h_{j,(t)}$, taking two values 5 and 10.

shipToHoldRatio

This is the ratio of the unit shipping cost $\gamma_{(t)}$ to the average unit holding cost at the leaf facilities $\frac{1}{|\mathcal{L}|}\sum_{j\in\mathcal{L}} h_{j,(t)}$, taking two values 1 and 5.

baseDemandCV

The mean demand $\mu_{j,(t)}$ at the leaf facilities may be the same ($baseDemandCV = C$) or varying ($baseDemandCV = S$). If $baseDemandCV = S$, mean demands follow a Normal distribution with coefficient of variation 0.4.

demandTimeRandomness

The mean demand $\mu_{j,(t)}$ may stay the same ($demandTimeRandomness = C$) over time or be seasonal ($demandTimeRandomness = S$). If $demandTimeRandomness = S$, the mean demand at a facility j in period $t > 0$ is given by $\mu_{j,(t)} = \mu_{j,(0)}\big(1 + \frac{1}{2}\sin(2\pi\frac{t}{T})\big)$.

demandCV

The demands at leaf facilities are assumed to follow a multivariate Normal distribution. The variance of the demand at a facility j in period t is given by $\mu_{j,(t)} \cdot demandCV$.

correlation

Parameters	Values	Explanation
numNodes (N)	$5, 10, 20$	The number of nodes.
numLeaves ($\|\mathcal{L}\|$)	$0.3N, 0.7N, N-1$	The number of leaf nodes.
height (H)	$1, 3, 6$	The height of the tree.
degree (D)	$0.3N, 0.7N, N-1$	The degree of the root and the tree.

Table 3.2: Tree graph parameters used in numerical study.

This is the correlation used to generate the variance-covariance matrix as follows.

$$[\boldsymbol{\Sigma}_{(t)}]_{p,q} = \sigma^p_{(t)}\sigma^q_{(t)} \cdot \begin{cases} 1 & \text{if } p = q, \\ |\rho| \cdot \text{sign}(\rho)^{p+q} & \text{if } p \neq q. \end{cases}$$

Graph Structure

We consider two types of directed acyclic graphs that are representative of practical inventory networks, tree graph and layered graph.

A tree graph represents a distributional system which have been widely studied. In a tree network, no facilities share common downstream facilities, and the path from the root facility to a leaf facility is unique. This is a commonly used inventory system in practice.

In a layered network, facilities are arranged in layers. There are direct paths only between facilities on adjacent layers. Facilities on the same layer may share common child facilities. This is a more complicated network that is also of practical interest.

In the following, we provide a detailed description of the graphs used in the study.

Tree Graph

We consider varying tree graphs in the study in four dimensions, number of nodes (N), number of leaf nodes ($|\mathcal{L}|$), height (H), and degree (D). Table 3.2 provides the values of the four parameters used in generating tree graphs in the study.

Given the four parameters, a tree graph is generated as follows. We denote the operation of setting a node a to be the parent of a node b as $a \to b$. Let $d(j)$ and $l(j)$ denote the degree and level of node j, respectively.

1. Generate N nodes labeled with $1, 2, \ldots, N$. Set $d(j) = 0, l(j) = 0$ for each node $j \in \{1, 2, \ldots, N\}$.
2. Set node 1 as the root of the tree. Set $\mathcal{L} = \{N - |\mathcal{L}| + 1, \ldots, N\}$.
3. Set $1 \to 2 \to \ldots \to H \to N$. For $j \in \{1, 2, \ldots, H\}$, set $d(j) = 1, l(j) = j - 1$.
4. For each node $j = H + 1, H + 2, \ldots, H + D - 1$, set $1 \to j, l(j) = 1$.
5. For each node $j = H + D, H + D + 1, \ldots, N - |\mathcal{L}|$, do
   (a) Let $\mathcal{S}_0 = \{k : k < j, d(k) = 0\}$.
   (b) If $|\mathcal{S}_0| \geq |\mathcal{L}|$, return `ERROR`; if $|\mathcal{S}_0| = |\mathcal{L}| - 1$, set $\mathcal{S} = \{k \in \mathcal{S}_0 : l(k) < H - 1\}$; else set $\mathcal{S} = \{k : k < j, d(k) < D, l(k) < H - 1\}$.
   (c) If $\mathcal{S} = \emptyset$, return `ERROR`; else randomly sample a node k from $\mathcal{S}$. Set $k \to j$. Set $d(k) = d(k) + 1, l(j) = l(k) + 1$.
6. Let $Q = \max\{H + D, N - |\mathcal{L}| + 1\}$. For each node $j = Q, Q + 1, \ldots, N - 1$, do
   (a) Let $\mathcal{S}_o = \{k \in \mathcal{L}^c : d(k) = 0\}$.
   (b) If $\mathcal{S}_o \neq \emptyset$, set $\mathcal{S} = \mathcal{S}_o$; else set $\mathcal{S} = \{k \in \mathcal{L}^c : d(k) < D\}$.
   (c) If $\mathcal{S} = \emptyset$, return `ERROR`; else randomly sample a node k from $\mathcal{S}$. Set $k \to j$. Set $d(k) = d(k) + 1, l(j) = l(k) + 1$.

Upon successful completion of the above algorithm, a tree with the specified parameters is generated. If the above algorithm returns `ERROR`, abandon this set of parameters.

Layered Graph

We consider layered graphs where any two facilities on adjacent layers are connected. The graphs vary in two dimensions, height (H) and degree (D). The first layer only includes the root facility, while all other layers, including the leaf layer, have the same number of facilities which is the degree. Thus, given H and D, a unique layered graph is determined.

(Height, Degree)	Number of Nodes	Number of Leaf Nodes
(2, 10)	21	10
(3, 3)	10	3
(3, 6)	19	6
(6, 3)	19	3

Table 3.3: Layered graph parameters used in numerical study.

We use four pairs of values for (H, D) in our study, $(2, 10), (3, 3), (3, 6)$ and $(6, 3)$. Table 3.3 summarizes the number of nodes and leaf facilities in each of the four cases.

3.5.3 Optimality Measure

For each instance, we calculate the relative optimality gap between the lower and upper bounds as $(UB - LB)/LB$, which gives us a more conservative estimate of the relative gap than calculating it as $(UB - LB)/UB$.

LB is computed by solving the dynamic program which provides an exact value of the lower bound. UB is estimated through Monte Carlo simulation. To reduce the variance of the upper bound estimator, we use control variate which is a standard variance reduction technique. In each sample path of the simulation, in addition to running the (feasible) upper bound heuristic to evaluate the upper bound cost, we also execute the (infeasible) lower bound policy given by the relaxed DP and evaluate the lower bound cost on the same sample path. The simulated lower bound cost is a random variable, denoted by L_s, whose expected value is LB. Thus $L_s - LB$ can serve as a control variate of the simulation. We provide the details of calculation in Algorithm 1.

3.5.4 Numerical Results

The numerical study covers a total of 40,960 instances. 30,720 of these instances have tree structured inventory networks, and 10,240 of them have layered structured inventory networks.

Algorithm 1: Upper bound cost simulation with control variate

for *each sample path* $s = 1, 2, \ldots, S$ **do**

  Run upper bound policy and get total cost U_s

  Run lower bound policy and get total cost L_s

  Set $\hat{U}_s := U_s - L_s + LB$

Set $UB := \frac{1}{S} \sum_{s=1}^{S} \hat{U}_s$.

Overall Results

The average gap between the lower and upper bounds of all 40,960 instances is 1.86% only, and the maximum gap is 13.3%. 94.7% of the instances have a gap smaller than 5%, and 99.88% of the instances have a gap smaller than 10%. Figure 3.2 shows the optimality gap of all instances, sorted from the largest to the smallest.

These results suggest that our lower bound is a close estimate of the true optimal cost, and our proposed upper bound heuristic is near optimal in a vast majority of the instances. Note that due to the limited computing resource for the numerical study, we have made adjustments in the actual implementation which will likely lead to a worse optimality result. However, these compromises should not be made in real business problems. One should optimize over the Lagrange multiplier, which will improve the lower bound and thus lead to a tighter optimality gap. Moreover, more combinations of the regularization parameters need to be considered to help find the policy with the lowest upper bound cost, which will further reduce the gap. Therefore, we have the reason to believe that the actual performance is even better in real business setting.

Policy Comparison

Our regularized policy also improves over the robust optimization approach alone. Recall that we have used two pairs of values for the regularization parameters (α_1, α_2) – $(1, 1)$ and (∞, ∞). The average gap for the policy with $(\alpha_1, \alpha_2) = (\infty, \infty)$, i.e., the robust optimization approach, is 2.2%. The average gap for the policy with $(\alpha_1, \alpha_2) = (1, 1)$ is 2.65%. While the gap is higher on average, the latter policy still outperforms the former in 42% of the instances. Therefore, combining the robust optimization approach with our proposed approximate look-ahead policy, and picking the best regularization parameters, is the best

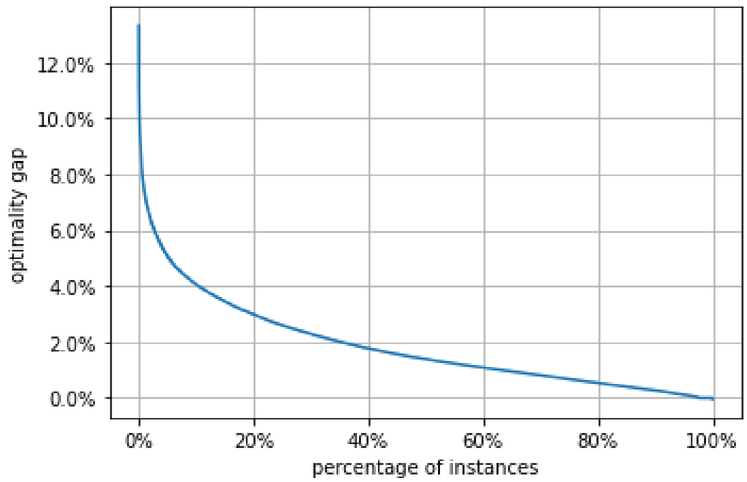


Figure 3.2: Optimality gaps of 40,960 instances, sorted from largest (left) to smallest (right).

strategy in real business problems. In the numerical study, by only including additional one set of values, we have succeeded in reducing the gap by 0.3% on average.

To further assess the conditions under which one policy is preferred to the other, we compare the performance of the two policies across each cost and distribution parameter. A significant finding is that the approximate look-ahead policy (with small regularization parameters) is preferred when the graph is broad and shallow, for example, in the classic one-warehouse multi-retailer system, while the robust optimization approach performs better when the graph is narrow and deep. This finding is not surprising. In the look-ahead policy, we approximate the future value by the value function from the relaxed lower bound DP which allows for "skip shippings", i.e., shippings between facilities that are not directly connected. Under this approximate value function, the policy may falsely choose an action in the current period that requires skip shippings in the future. When the graph is deep, skip shippings are more likely to occur in the relaxed problem and thus the policy tends to make bad decisions.

Figure 3.3 shows box plots of the gap differences between the two policies when the heights of tree graphs vary. A positive value means the policy with regularization parameters $(\alpha_1, \alpha_2) = (1, 1)$ has a larger gap. As the plots show, when the height of the tree graph is 1, this policy performs much better than the robust optimization policy. However, when the height is large, it is outperformed by the robust optimization policy.

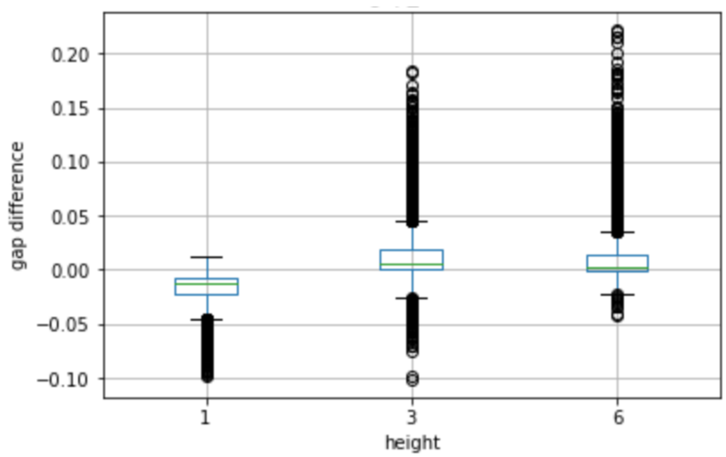


Figure 3.3: Box plots of the gap differences between the two policies, one with regularization parameters $(1, 1)$ and one with (∞, ∞). A positive value means the former has a larger gap.

3.6 Generalizations and Conclusion

In this chapter, we studied an inventory problem of replenishment and allocation systems on general directed acyclic graph (DAG) networks which assemble both the properties of distribution and assembly systems. We proposed an innovative formulation of the problem in which Lagrangian relaxation can be easily applied to formulate a tractable lower bound dynamic program. We used the information from the lower bound DP to help design the upper bound heuristic. We conducted an extensive numerical study consisting of 40,960 instances to assess the accuracy of the lower bound and suboptimality of the upper bond heuristic. The average gap is 1.86% only with gaps below 10% in 99.88% of the instances. We believe this performance is acceptable in any practical business problems.

In the following, we conclude this chapter by discussing some generalizations.

3.6.1 General Lead Times

So far we have assumed a lead time of one period for ordering into the source facility and shipping between any two adjacent facilities. In practice, larger lead times are very common. In the following, we extend our approach to accommodate general lead times.

Procurement Lead Time

For a problem with a general ordering lead time ℓ, the state of the dynamic program is represented by a vector of size $J+\ell-1$, $(x_{1,(t)},\ldots,x_{J,(t)},w_{(t-\ell+1)},w_{(t-\ell+2)},\ldots,w_{(t-1)})$, where $x_{j,(t)}$, for $j=1,\ldots,J$, is the inventory level at facility j at the beginning of period t, and $w_{(t-\tau)}$, for $\tau=\ell-1,\ldots,1$, corresponds to the order made in period $t-\tau$ that will arrive at root facility at the beginning of period $t+\ell-\tau$. After applying the Lagrangian relaxation step to aggregate the total inventory positions in all facilities, the state vector is collapsed to $(x_{+,(t)},w_{(t-\ell+1)},w_{(t-\ell+2)},\ldots,w_{(t-1)})$ which is of size ℓ.

By a standard substitution technique identified by Clark and Scarf (1960), the state of the dynamic program can be further collapsed to a single-dimensional one whose value is given by the total aggregate inventory positions in the system, including those in transit to the source facility. The relaxed DP then becomes

$$\begin{aligned}\bar{V}^{\lambda}_{(t)}(x_{+,(t)}+w_{+,(t)}) = \mathbb{E}&\left[R_{(t+\ell)}\Big(x_{+,(t)}+w_{+,(t)}-\sum_{\tau=t}^{t+\ell-1}u_{+,(\tau)}\Big)\right]\\ &+\min_{\bar{w}_{(t)}\geq 0}\left\{(c_{(t)}-\lambda_{1,(t+1)})\bar{w}_{(t)}+\mathbb{E}\left[\bar{V}^{\lambda}_{(t+1)}(\bar{w}_{(t)}+w_{+,(t)}+x_{+,(t)}-u_{+,(t)})\right]\right\}\\ &+\boldsymbol{\lambda}^{\top}_{(t+1)}\mathbb{E}\left[\boldsymbol{u}_{(t)}\right]\end{aligned}$$

where $R_{(t)}(\cdot)$ is defined as in (3.8). See Federgruen et al. (2018) for details of the derivation of this DP in a one-warehouse multi-retailer problem.

Retailer Stores Delivery Lead Time

General lead times of shipping into retailer stores (leaf facilities) can be easily handled by employing a standard forward accounting scheme. Let ℓ_j denote the lead time of shipping into retailer store j. We charge to period t the expected holding and backlogging costs at facility j at the beginning of period $t+\ell_j$, i.e.,

$$Q_{j,(t)}(\bar{x}_{j,(t)}) := \mathbb{E}\Big[h_{j,(t+\ell)}(\bar{x}_{j,(t)}-\sum_{\tau=t}^{t+\ell_j-1}u_{j,(\tau)})^{+}+p_{j,(t+\ell_j)}(\bar{x}_{j,(t)}-\sum_{\tau=t}^{t+\ell_j-1}u_{j,(\tau)})^{-}\Big].$$

Transshipment Lead Time

To handle general lead times of shipping between non-leaf facilities, we add virtual facilities between the two facilities. For a lead time of ℓ, $\ell - 1$ virtual facilities are created. Virtual facilities are no different from physical facilities in that the inflow constraints (3.1) and outflow constraints (3.2) are the same, except that for each virtual facility v, we add in each period t an additional constraint $\bar{x}_{v,(t)} = 0$, since virtual facilities cannot store any inventory.

3.6.2 Storage Capacity Constraints

For single-item problem, capacity constraints on inventory level at retailer stores of both conditional expected value and conditional probability type can be easily translated into upper bounds on inventory positions, i.e., linear constraints. Bound type constraints for non-leaf facilities are reasonable constraints to impose as well since inventory level is non-random at these facilities given the procurement/allocation decision.

With bound type constraints imposed, the only modifications needed in our approach is to add constraints $\bar{x}_{j,(t)} \leq U_{j,(t)}$ for each facility j, both in the lower bound DP and the upper bound heuristic. These bound type constraints add almost no additional complexity to the optimization problems.

Bibliography

Adelman, D. and Mersereau, A. J. (2008). Relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 56(3):712–727.

Alptekinoğlu, A., Banerjee, A., Paul, A., and Jain, N. (2013). Inventory pooling to deliver differentiated service. *Manufacturing & Service Operations Management*, 15(1):33–44.

Ardestani-Jaafari, A. and Delage, E. (2016). Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations research*, 64(2):474–494.

Baker, K. R. (1985). Safety stocks and component commonality. *Journal of Operations Management*, 6(1):13–22.

Baker, K. R., Magazine, M. J., and Nuttle, H. L. (1986). The effect of commonality on safety stock in a simple inventory model. *Management Science*, 32(8):982–988.

Balseiro, S. R., Brown, D. B., and Chen, C. (2021). Dynamic pricing of relocating resources in large networks. *Management Science*, 67(7):4075–4094.

Beck, A. and Teboulle, M. (2009). A Fast Iterative Shrinkage-Thresholding Algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1):183–202.

Beyer, D., Sethi, S. P., and Sridhar, R. (2001). Stochastic multiproduct inventory models with limited storage. *Journal of Optimization Theory and Applications*, 111(3):553–588.

Beyer, D., Sethi, S. P., and Sridhar, R. (2002). Average-cost optimality of a base-stock policy for a multi-product inventory model with limited storage. In *Decision & Control in Management Science*, pages 241–260. Springer.

Bijvank, M. (2014). Periodic review inventory systems with a service level criterion. *Journal of the Operational Research Society*, 65(12):1853–1863.

Bijvank, M. and Vis, I. F. (2011). Lost-sales inventory theory: A review. *European Journal of Operational Research*, 215(1):1–13.

Bijvank, M. and Vis, I. F. (2012). Lost-sales inventory systems with a service level criterion. *European Journal of Operational Research*, 220(3):610–618.

Boyaci, T. and Gallego, G. (2001). Serial production/distribution systems under service constraints. *Manufacturing & Service Operations Management*, 3(1):43–50.

Brown, D. B. and Smith, J. E. (2020). Index policies and performance bounds for dynamic selection problems. *Management Science*, 66(7):3029–3050.

Brown, D. B. and Zhang, J. (2021). Dynamic programs with shared resources and signals: Dynamic fluid policies and asymptotic optimality. *To appear in Operations Research*.

Caro, F., Kök, A. G., and Martínez-de Albéniz, V. (2020). The future of retail operations. *Manufacturing & Service Operations Management*, 22(1):47–58.

Chen, F. (2000). Optimal policies for multi-echelon inventory problems with batch ordering. *Operations research*, 48(3):376–389.

Chen, X. and Li, M. (2019). M-natural-convexity and its applications in operations. *Available at SSRN 3431474, to appear in Operations Research*.

Clark, A. J. and Scarf, H. (1960). Optimal policies for a multi-echelon inventory problem. *Management science*, 6(4):475–490.

DeCroix, G. A. and Arreola-Risa, A. (1998). Optimal production and inventory policy for multiple products under resource constraints. *Management Science*, 44(7):950–961.

Diks, E. B. and De Kok, A. (1998). Optimal control of a divergent multi-echelon inventory system. *European journal of operational research*, 111(1):75–97.

Federgruen, A. and Groenevelt, H. (1986). The greedy procedure for resource allocation problems: Necessary and sufficient conditions for optimality. *Operations Research*, 34(6):909–918.

Federgruen, A., Guetta, C. D., and Iyengar, G. (2018). Two-echelon distribution systems with random demands and storage constraints. *Naval Research Logistics (NRL)*, 65(8):594–618.

Federgruen, A., Guetta, D., Iyengar, G., and Liu, X. (2022). Scalable approximately optimal policies for multi-item stochastic inventory problems. *Available at SSRN 4223257*.

Federgruen, A. and Zipkin, P. (1984a). Allocation policies and cost approximations for multilocation inventory systems. *Naval Research Logistics Quarterly*, 31(1):97–129.

Federgruen, A. and Zipkin, P. (1984b). Approximations of dynamic, multilocation production and inventory problems. *Management Science*, 30(1):69–84.

Federgruen, A. and Zipkin, P. (1984c). Computational issues in an infinite-horizon, multiechelon inventory model. *Operations Research*, 32(4):818–836.

Fujishige, S. (2005). *Submodular functions and optimization*. Elsevier.

Gaur, V. and Honhon, D. (2006). Assortment planning and inventory decisions under a locational choice model. *Management Science*, 52(10):1528–1543.

Gerchak, Y., Magazine, M. J., and Gamble, A. B. (1988). Component commonality with service level requirements. *Management science*, 34(6):753–760.

Groenevelt, H. (1991). Two algorithms for maximizing a separable concave function over a polymatroid feasible region. *European Journal of Operational Research*, 54(2):227–236.

Hoeffding, W. (1963). Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, pages 13–30.

Hu, W. and Frazier, P. (2017). An asymptotically optimal index policy for finite-horizon restless bandits. *arXiv preprint arXiv:1707.00205*.

Ignall, E. and Veinott Jr, A. F. (1969). Optimality of myopic inventory policies for several substitute products. *Management Science*, 15(5):284–304.

Jackson, P. L. (1988). Stock allocation in a two-echelon distribution system or “what to do until your ship comes in”. *Management Science*, 34(7):880–895.

Jiang, J., Wang, S., and Zhang, J. (2019). Achieving high individual service-levels without safety stock? optimal rationing policy of pooled resources. *Working Paper (May 2, 2019)*.

Kunnumkal, S. and Topaloglu, H. (2008). A duality-based relaxation and decomposition approach for inventory distribution systems. *Naval Research Logistics (NRL)*, 55(7):612–631.

Kunnumkal, S. and Topaloglu, H. (2011). Linear programming based decomposition methods for inventory distribution systems. *European Journal of Operational Research*, 211(2):282–297.

Lawler, E. L. and Martel, C. U. (1982). Computing maximal “polymatroidal” network flows. *Mathematics of Operations Research*, 7(3):334–347.

Maggiar, A., Song, I., and Muharremoglu, A. (2022). Multi-echelon inventory management for a non-stationary capacitated distribution network. *Available at SSRN 4154780*.

Mahajan, S. and Van Ryzin, G. (2001). Stocking retail assortments under dynamic consumer substitution. *Operations Research*, 49(3):334–351.

Marklund, J. and Rosling, K. (2012). Lower bounds and heuristics for supply chain stock allocation. *Operations Research*, 60(1):92–105.

Miao, S., Jasin, S., and Chao, X. (2022). Asymptotically optimal lagrangian policies for multi-warehouse, multi-store systems with lost sales. *Operations Research*, 70(1):141–159.

Miller, B. L. (1986). Scarf’s state reduction method, flexibility, and a dependent demand inventory model. *Operations Research*, 34(1):83–90.

Mirchandani, P. and Mishra, A. K. (2002). Component commonality: Models with product-specific service constraints. *Production and Operations Management*, 11(2):199–215.

Nambiar, M., Simchi-Levi, D., and Wang, H. (2021). Dynamic inventory allocation with demand learning for seasonal goods. *Production and Operations Management*, 30(3):750–765.

Nemirovski, A. and Shapiro, A. (2006). Scenario approximations of chance constraints. *Probabilistic and randomized methods for design under uncertainty*, pages 3–47.

Nemirovski, A. and Shapiro, A. (2007). Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996.

Özer, Ö. and Xiong, H. (2008). Stock positioning and performance estimation for distribution systems with service constraints. *IIE Transactions*, 40(12):1141–1157.

Prékopa, A. (2003). Probabilistic programming. *Handbooks in operations research and management science*, 10:267–351.

Rosling, K. (1989). Optimal inventory policies for assembly systems under random demands. *Operations Research*, 37(4):565–579.

Rudin, W. et al. (1964). *Principles of mathematical analysis*, volume 3. McGraw-hill New York.

Rusmevichientong, P., Shen, Z.-J. M., and Shmoys, D. B. (2010). Dynamic assortment optimization with a multinomial logit choice model and capacity constraint. *Operations Research*, 58(6):1666–1680.

Shang, K. H. and Song, J.-S. (2006). A closed-form approximation for serial inventory systems and its application to system design. *Manufacturing & Service Operations Management*, 8(4):394–406.

Shi, C., Chen, W., and Duenyas, I. (2016). Nonparametric data-driven algorithms for multiproduct inventory systems with censored demand. *Operations Research*, 64(2):362–370.

Song, J.-S., van Houtum, G.-J., and Van Mieghem, J. A. (2020). Capacity and inventory management: Review, trends, and projections. *Manufacturing & Service Operations Management*, 22(1):36–46.

Swaminathan, J. M. and Srinivasan, R. (1999). Managing individual customer service constraints under stochastic demand. *Operations research letters*, 24(3):115–125.

Topkis, D. M. (1998). *Supermodularity and complementarity*. Princeton university press.

Veinott Jr, A. F. (1965). Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science*, 12(3):206–222.

Wainwright, M. J. (2019). *High-dimensional statistics: A non-asymptotic viewpoint*, volume 48. Cambridge University Press.

Wei, L., Jasin, S., and Xin, L. (2021). On a deterministic approximation of inventory systems with sequential service-level constraints. *Operations Research*, 69(4):1057–1076.

Xin, L. and Goldberg, D. A. (2016). Optimality gap of constant-order policies decays exponentially in the lead time for lost sales models. *Operations Research*, 64(6):1556–1565.

Zayas-Caban, G., Jasin, S., and Wang, G. (2019). An asymptotically optimal heuristic for general non-stationary finite-horizon restless multi-armed, multi-action bandits. *Advances in Applied Probability*, 51(3):745–772.

Zipkin, P. H. (1980). Simple ranking methods for allocation of one resource. *Management Science*, 26(1):34–43.

Zipkin, P. H. (2000). *Foundations of inventory management*. McGraw-hill.

Appendix A: Proofs of Chapter 1

A.1 Proof of Theorem 1

Proof. (a) Let ρ^i denote the optimal policy for the dynamic program (1.8). We prove that ρ^i has the following double base stock structure.

$$\rho^i(x^i_{(t)}) = \begin{cases} S^{i,\boldsymbol{\lambda}}_{(t)}, & x < S^{i,\boldsymbol{\lambda}}_{(t)} \quad \text{(A.1a)} \\ x, & x \in [S^{i,\boldsymbol{\lambda}}_{(t)}, B^{i,\boldsymbol{\lambda}}_{(t)}] \quad \text{(A.1b)} \\ B^{i,\boldsymbol{\lambda}}_{(t)}, & x > B^{i,\boldsymbol{\lambda}}_{(t)} \quad \text{(A.1c)} \end{cases}$$

Fix $i = 1, \dots, I$. To further simplify the notation, we omit the superscript indices i and λ. In view of (1.1), one easily verifies that it is never optimal to order and salvage in the same period; thus $W_{(t)}(x_{(t)}) = \min\{W^1_{(t)}(x_{(t)}), W^2_{(t)}(x_{(t)})\}$, where

$$W^1_{(t)}(x_{(t)}) = -c_{(t)}x_{(t)} + \min_{\bar{x}_{(t)} \geq x_{(t)}} \left\{ c_{(t)}\bar{x}_{(t)} + \bar{Q}_{(t)}(\bar{x}_{(t)}) + \delta\mathbb{E}[W_{(t+1)}(\bar{x}_{(t)} - u_{(t)})] \right\} - \frac{\lambda_{(t)}\chi_{(t+\ell)}}{I}, \tag{A.2}$$

and

$$W^2_{(t)}(x_{(t)}) = d_{(t)}x_{(t)} + \min_{\min\{x_{(t)},0\} \leq \bar{x}_{(t)} \leq x_{(t)}} \left\{ -d_{(t)}\bar{x}_{(t)} + \bar{Q}_{(t)}(\bar{x}_{(t)}) + \delta\mathbb{E}[W_{(t+1)}(\bar{x}_{(t)} - u_{(t)})] \right\} - \frac{\lambda_{(t)}\chi_{(t+\ell)}}{I}, \tag{A.3}$$

where $\bar{Q}_{(t)}(\bar{x}_{(t)}) = (\delta^\ell h_{(t+\ell)} + \lambda_{(t)})\mathbb{E}[(\bar{x}_{(t)} - \dot{u}_{(t)})^+] + \delta^\ell p_{(t+\ell)}\mathbb{E}[(\bar{x}_{(t)} - \dot{u}_{(t)})^-]$ is a strictly convex function since $\dot{u}_{(t)}$ has a continuous distribution on $(-\infty, \infty)$.

We prove the theorem by (backwards) induction with respect to t. The statement trivially holds for $t = T - \ell + 1$ with $W_{T-\ell+1}$ = 0. Assume it holds for some period $(t+1)$ with $t = 0, \ldots, T-\ell$. The minimand of (A.2) is a strictly convex function $H^1(\bar{x}_{(t)})$ of $\bar{x}_{(t)}$ only, since the function $W_{(t+1)}(\cdot)$ is convex, by the induction assumption, while $\bar{Q}_{(t)}$ is strictly convex. Moreover, $H^1(\cdot)$ is differentiable, with derivative:

$$H^{1\prime}(\bar{x}_{(t)}) = c_{(t)} + \bar{Q}'_{(t)}(\bar{x}_{(t)}) + \delta\mathbb{E}[W'_{(t+1)}(\bar{x}_{(t)} - u_{(t)})], \tag{A.4}$$

where both the existence of the derivative and the interchange of the expectation and derivative operators are justified, by the fact that $|W'_{(t+1)}(y)| \leq A|y|$, for some constant $A > 0$. Thus $\mathbb{E}\left[W'_{(t+1)}(\bar{x}_{(t)} - u_{(t)})\right] \leq A\mathbb{E}\left[|\bar{x}_{(t)} - u_{(t)}|\right] \leq A\mathbb{E}\left[|\bar{x}_{(t)}| + |w_{(t)}|\right] < \infty$. The interchange of the expectation and derivative operators is now justified by a standard arguement, see e.g. Theorem 9.42 in Rudin et al. (1964).

Moreover, $\lim\limits_{x\uparrow\infty} H^{1\prime}(x) > 0$, since $c_{(t)} > 0, \lim\limits_{x\uparrow\infty} \bar{Q}'_{(t)}(x) > 0$ and $\lim\limits_{x\uparrow\infty} W'_{(t+1)}(x) \geq 0$. Thus, the strictly convex function $H^1{}_{(t)}(\cdot)$ achieves its unique minimum at

$$-\infty \leq S_{(t)} = \inf\{x : H^{1\prime}{}_{(t)}(x) > 0\} < \infty$$

since the infimum is taken over a non-empty set. It follows that for $x_{(t)} < S_{(t)}$, it is optimal to place an order to elevate $\bar{x}_{(t)}$ to $S_{(t)}$, thus verifying (A.1a).

By the same argument, the function $H^2_{(t)}(\cdot)$ is a strictly convex function with

$$H^{2\prime}(\bar{x}_{(t)}) = -d_{(t)} + \bar{Q}'_{(t)}(\bar{x}_{(t)}) + \delta\mathbb{E}[W'_{(t+1)}(\bar{x}_{(t)} - u_{(t)})], \tag{A.5}$$

a strictly increasing function which is positive for $\bar{x}_{(t)}$ sufficiently large. Thus $H^2_{(t)}(\cdot)$ achieves its minimum at $\tilde{B}_{(t)}$, defined as

$$\tilde{B}_{(t)} = \inf\{x : H^{2\prime}{}_{(t)}(x) > 0\} \leq \infty.$$

Moreover, by (1.1) $H^{2\prime}{}_{(t)}(\cdot)$ is pointwise smaller than $H^{1\prime}{}_{(t)}(\cdot)$ so that $S_{(t)} < \tilde{B}_{(t)} \leq \max\{\tilde{B}_{(t)}, 0\}$.

Thus, for $x_{(t)} \geq B_{(t)}$, it is optimal to reduce the inventory position to $B_{(t)}$, thus veryfying (A.1b). (It is infeasible to reduce the inventory position below the zero-level.)

Finally, for $S_{(t)} < x < \tilde{B}$, it is optimal not to order or to salvage, and the same applies for $\tilde{B}_{(t)} \leq x < B_{(t)}$ since salvaging is infeasible on this interval.

In addition,

$$W_{(t)}(x_{(t)}) = \begin{cases} c_{(t)}(S_{(t)} - x_{(t)}) + \bar{Q}_{(t)}(S_{(t)}) + \delta \mathbb{E}\left[W_{(t+1)}(S_{(t)} - u_{(t)})\right] \\ \qquad - \dfrac{\lambda_{(t)} \chi_{(t+\ell)}}{I}, \text{ if } x < S^{i,\lambda}_{(t)} & \text{(A.6a)} \\ \bar{Q}_{(t)}(x_{(t)}) + \delta \mathbb{E}\left[W_{(t+1)}(x_{(t)} - u_{(t)})\right] \\ \qquad - \dfrac{\lambda_{(t)} \chi_{(t+\ell)}}{I}, \text{ if } x \in [S^{i,\lambda}_{(t)}, B^{i,\lambda}_{(t)}] & \text{(A.6b)} \\ d_{(t)}(x_{(t)} - B_{(t)}) + \bar{Q}_{(t)}(B_{(t)}) + \delta \mathbb{E}\left[W_{(t+1)}(B_{(t)} - u_{(t)})\right] \\ \qquad - \dfrac{\lambda_{(t)} \chi_{(t+\ell)}}{I}, \text{ if } x > B^{i,\lambda}_{(t)} & \text{(A.6c)} \end{cases}$$

showing that $W_{(t)}(\cdot)$ is differentiable everywhere, with the possible exception of the points $x_{(t)} = S_{(t)}$ and $x_{(t)} = B_{(t)}$, while $W'_{(t)}(x) = O(|x|)$, and $\lim_{x \uparrow \infty} W'_{(t)}(x) \geq 0$. This completes the induction step.

(b) To simplify our notation, we fix i and omit the superscript in this part of the proof. Consider a dynamic program analogous to (1.8), but that maximizes profits instead of minimizing costs. This program can be decomposed over products; let $W^{\nu}_{(t)}(x_{(t)})$ denote the value function for this dynamic program where the Lagrange multiplier $\boldsymbol{\lambda}$ is replaced by $\boldsymbol{\nu} \leq 0$. That is, we have the recursion

$$W^{\nu}_{(t)}(x_{(t)}) = \max_{\bar{x}_{(t)} \geq \min\{x_{(t)}, 0\}} \Big\{ - c\big(\bar{x}_{(t)} - x_{(t)}\big)^{+} - d\big(x_{(t)} - \bar{x}_{(t)}\big)^{+} - Q_{(t)}(\bar{x}_{(t)}) \qquad \text{(A.7)}$$
$$+ \nu_{(t)} \mathbb{E}\big[(\bar{x}_{(t)} - \dot{u}_{(t)})^{+}\big] + \mathbb{E}\big[W^{\nu}_{(t+1)}(\bar{x}_{(t)} - u_{(t)})\big] - \frac{\nu_{(t)} \chi_{(t+\ell)}}{I} \Big\}.$$

We first use induction to establish that, for each t, the value function $W^{\nu}_{(t)}(x_{(t)})$ is supermodular in $(\boldsymbol{\nu}, x_{(t)})$. To that end, suppose $W^{\nu}_{(t+1)}(x_{(t+1)})$ is supermodular in $(\boldsymbol{\nu}, x_{(t+1)})$. (Since $W^{\nu}_{(T+1)}(x_{(T+1)}) = 0$, this is true for $t = T$.)

We interpret (A.7) as a mathematical program in which $\bar{x}_{(t)}$ is the decision variable and $\boldsymbol{\theta} \equiv (\boldsymbol{\nu}, x_{(t)})$ the set of parameters. We claim that the objective in (A.7) is a supermodular function of the $(T+2)$-dimensional vector $(\bar{x}_{(t)}, \boldsymbol{\theta})$. This follows from the following observations:

- The sum of the first two terms $-c(\bar{x}_{(t)} - x_{(t)})^+ - d(x_{(t)} - \bar{x}_{(t)})^+$ is a concave function of $(\bar{x}_{(t)} - x_{(t)})$, and therefore supermodular in $(\bar{x}_{(t)}, \boldsymbol{\theta})$.
- The third term $-Q_{(t)}(\bar{x}_{(t)})$ is modular in $(\bar{x}_{(t)}, \boldsymbol{\theta})$.
- The fourth term $\nu_{(t)}(\bar{x}_{(t)} - \dot{u}_{(t)})^+$ is supermodular in $(\bar{x}_{(t)}, \boldsymbol{\theta})$ since $(\nu_{(t)}^{(2)} - \nu_{(t)}^{(1)})\mathbb{E}[(\bar{x}_{(t)} - \dot{u}_{(t)})^+]$ is increasing in $\bar{x}_{(t)}$ for all $\nu_{(t)}^{(2)} \geq \nu_{(t)}^{(1)}$.
- By the induction assumption, for any realization of $u_{(t)}$, $W_{(t+1)}^{\nu}(\bar{x}_{(t)} - u_{(t)})$ is supermodular in $(\bar{x}_{(t)}, \boldsymbol{\theta})$, and this property is maintained after taking the expectation.
- The last term is modular in $(\bar{x}_{(t)}, \boldsymbol{\theta})$.

Therefore, by Theorem 2.7.6 in Topkis (1998), $W_{(t)}^{\nu}(x_{(t)})$ is a supermodular function of $\boldsymbol{\theta} = (x_{(t)}, \boldsymbol{\nu})$, completing the induction proof.

Since we have shown that the objective of (A.7) is supermodular in $(\bar{x}_{(t)}, \boldsymbol{\theta})$, and the feasible set $\{\bar{x}_{(t)} : \bar{x}_{(t)} \geq \min\{x_{(t)}, 0\}\}$ is increasing in $x_{(t)}$ and therefore in the entire vector $\boldsymbol{\theta}$, it follows from Theorem 2.8.1 in Topkis (1998) that the optimal decision $\bar{x}_{(t)}$ is monotone in $\boldsymbol{\theta}$, in particular, $\bar{x}_{(t)}$ is increasing in $\boldsymbol{\nu}$, and therefore, decreasing in $\boldsymbol{\lambda}$.

Given the optimality of a double base stock policy this implies in particular that (restoring i in our notation) $S_{(t)}^{i,\boldsymbol{\lambda}} \leq S_{(t)}^{i,\mathbf{0}}$ and $B_{(t)}^{i,\boldsymbol{\lambda}} \leq B_{(t)}^{i,\mathbf{0}}$ for all t and all $\boldsymbol{\lambda} \geq \mathbf{0}$.

(c) By a proof similar to that of part (b), we can show that on every sample path, the optimal $\bar{x}_{(t)}^i$ decreases as the cost-rate vectors $\{c_{(t)}^i, d_{(t)}^i : t = 0, \ldots, T\}$ increase. Thus, replacing $c_{(t)}^i$ and $d_{(t)}^i$ by their minimum possible values $\underline{c}$ and $\underline{d}$, we *increase* $\{S_{(t)}^{i,\mathbf{0}}\}$. By a slight adaptation of the proof of Theorem 9.4.2(a) in Zipkin (2000) to allow for the salvaging option, we obtain that

$$S_{(t)}^{i,\mathbf{0}} \leq S_{(t)}^{i,+} := F^{-1}\Big(\frac{p_{(t+\ell)}^i - (1-\gamma)\underline{c}}{p_{(t+\ell)}^i + h_{(t+\ell)}^i}\Big) \leq F^{-1}\Big(\frac{\bar{p} - (1-\gamma)\underline{c}}{\bar{p} + \underline{h}}\Big) =: \bar{S},$$

where the second inequality follows from the fact the fractile $\frac{p^i_{(t+\ell)}-(1-\gamma)\underline{c}}{p^i_{(t+\ell)}+h^i_{(t+\ell)}}$ is increasing in $p^i_{(t+\ell)} \leq \bar{p}$ and decreasing in $h^i_{(t+\ell)} \geq \underline{h}$.

□

A.2 Proof of Lemma 1

Proof. Let $f^i_{(t)}(\cdot)$ and $F^i_{(t)}(\cdot)$ denote the p.d.f. and c.d.f. of $u^i_{(t)}$, respectively, and define $M^i_{(t)} := \min_{L\leq x\leq U} f^i_{(t)}(x) > 0$ because continuous function achieves its minimum on any closed interval. Without of loss of essential generality, assume $M^{min} := \min_{i,t} M^i_{(t)} > 0$. Define $\dot{f}^i_{(t)}(\cdot)$, $\dot{F}^i_{(t)}(\cdot)$ and $\dot{M}^{min} > 0$, similarly.

(a) The Lagrange multiplier $\lambda_{(t)}$ in each period denotes the amount by which $W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ decreases if only the t-th period capacity level $\chi_{(t+\ell)}$ were increased by one unit, and all other capacity levels remain unchanged. This capacity increment may allow us to eliminate one unit of backlog in the t-th or a future period, the value of which is bounded by $\bar{p}$. Additionally it may allow us to switch procurement time for a total of one unit, a benefit that is bounded by $\bar{c} - \underline{c} \leq \bar{c}$.

(b) We separately assess (1) the number of operations to be performed in any given iteration of FISTA, and (2) a worst-case bound for the number of iterations.

As far as (1) is concerned, in each iteration I separate dynamic programs need to be solved. With, for numerical purposes, a bounded state and action space, a total of $O(IT)$ operations is required. Since the gradient has T components, and evaluation of each component requires $O(I)$ operations, the total number of operations to evaluate the entire gradient is $O(IT)$ as well. Thus, the total complexity for any given iteration is $O(IT)$.

As far as (2) is concerned, Beck and Teboulle (2009) show that the number of iterations to achieve ϵ-absolute optimality is $O(\frac{\sqrt{L}D}{\sqrt{\epsilon}})$, where $D^2 := \|\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}^*\| = \|\boldsymbol{\lambda}^*\|$ which is $O(T)$ by part (a).

Let H^{λ} denote the Hessian of the concave function $W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ with respect to the vector $\boldsymbol{\lambda}$. Then L is a Lipschitz constant which is bounded by $\sup_{0\leq\boldsymbol{\lambda}\leq\bar{p}+\bar{c}} \nu(H^{\lambda})$, where $\nu(H^{\lambda})$ denotes the largest absolute eigenvalue of H^{λ}. Because H^{λ} is the Hessian of a concave function, it is symmetric and negative semi-definite. Therefore all of its eigenvalues are non-positive and $\nu(H^{\lambda}) \leq -Tr[H^{\lambda}]$, the negative trace of the Hessian matrix. We will show that

$$-Tr[H^{\lambda}] = \sum_{t=0}^{T} -\frac{\partial^2 W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})}{\partial \lambda^2_{(t)}} = O(I) \tag{A.8}$$

so that the number of iterations is $O(\frac{\sqrt{T}}{\sqrt{\epsilon}})$, and the total complexity is $O(\frac{IT^{\frac{3}{2}}}{\sqrt{\epsilon}})$.

To show (A.8), it is easily shown that the value function $W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ is differentiable with respect to $\boldsymbol{\lambda}$. Moreover, for any vector $\boldsymbol{\lambda} \geq 0$, it follows from (1.7) that

$$\begin{aligned}
\frac{\partial W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})}{\partial \lambda_{(t)}} &= \delta^{t-1}\mathbb{E}\Big[\sum_{i=1}^{I}(\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ - \chi_{(t+\ell)}\Big] \\
&= \delta^{t-1}\Big(\sum_{i=1}^{I}\int_0^{\bar{x}^i_{(t)}}(\bar{x}^i_{(t)} - u)\dot{f}^i_{(t)}(u)du - \chi_{(t+\ell)}\Big).
\end{aligned}$$

Applying Leibniz's rules, this implies that

$$\begin{aligned}
\Big|\frac{\partial^2 W^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})}{\partial \lambda^2_{(t)}}\Big| &= \delta^{t-1}\Big|\sum_{i=1}^{I}\int_0^{\bar{x}^i_{(t)}}\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}}\dot{f}^i_{(t)}(u)du\Big| \\
&= \delta^{t-1}\Big|\sum_{i=1}^{I}\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}}\dot{F}^i_{(t)}(\bar{x}^i_{(t)})\Big| \\
&\leq \delta^{t-1}\sum_{i=1}^{I}\Big|\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}}\dot{F}^i_{(t)}(\bar{x}^i_{(t)})\Big| \\
&\leq \delta^{t-1}\sum_{i=1}^{I}\Big|\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}}\Big|.
\end{aligned}$$

Since $\sum_{t=0}^{T}\delta^{t-1} \leq \frac{1}{1-\delta} < \infty$, to prove (A.8), it suffices to show that there is a uniform upper bound for $\Big|\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}}\Big|$ which applies for all i and t on the entire interval $\lambda_{(t)} \in [0, \bar{c} + \bar{p}]$. See part (a).

Recall that $\bar{x}^i_{(t)}$ is the unique root of (1.9a) or (1.9b). Either way:

$$\left|\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}}\right| = -\frac{\partial \bar{x}^i_{(t)}}{\partial \lambda_{(t)}} = \frac{\dot{F}^i_{(t)}(\bar{x}^i_{(t)})}{Q''^i_{(t)}(\bar{x}^i_{(t)}) + \mathbb{E}[W'^{\lambda,i}_{(t+1)}(\bar{x}^i_{(t)} - u^i_{(t)})] + \lambda_{(t)} \dot{f}^i_{(t)}(\bar{x}^i_{(t)})} \leq \frac{1}{Q''^i_{(t)}(\bar{x}^i_{(t)})}$$
$$\leq \frac{1}{(\underline{h} + \underline{p})\dot{M}^{min}} < \infty,$$

independent of I and T, where $\underline{h}$ and $\underline{p}$ are defined in the notation section. The first inequality follows from the second and third term in the denominator of the fraction to its left, being non-negative.

□

Appendix B: Proofs of Chapter 2

B.1 Proof of Theorem 5

In this section, we prove that moving from policy $\check{\boldsymbol{\rho}}^{\star}_{\kappa}$ (henceforth $\check{\boldsymbol{\rho}}$ to simplify notation) to policy $\boldsymbol{\pi}$ only adds at most $O(\sqrt{I})$ to the cost of the policy; i.e., $C(\boldsymbol{\pi}) - C(\check{\boldsymbol{\rho}}) = O(\sqrt{I})$.

The crux of our proof will focus on those parts of the state space in which $\boldsymbol{\pi}$ and $\check{\boldsymbol{\rho}}$ disagree – in other words, those states in which the LHS of constraint $(\ddot{7})$ exceeds its mean in the LHS of constraint $(\check{8})$ by at least $(\kappa - \kappa_0)$. Our proof will proceed in two steps. First, we note that the LHS of $(\ddot{7})$ is a function of the underlying demand random variables, and we show that this function is Lipschitz-continuous in those variables. This allows us to use well-known concentration inequalities to bound the probability of such deviations to be $O(1/I)$. We then prove the fact that when deviations *do* occur, they result in an *expected* impact of at most $O(I)$ to our total cost. Unfortunately, this does not quite prove our result, because these deviations are correlated with the event of the two policies disagreeing in the first place. In the last step of our proof, we show that despite this potential correlation, our result holds.

To prove these statements, we will require some new notation. Throughout this section, we fix γ and drop it from our notation.

- Assume all policies start from the same $\boldsymbol{x}_{(0)}$. As they progress through the time horizon, they visit a sequence of (random) states. $\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)}$ denotes the state at the start of period t when following policy $\boldsymbol{\pi}$. $\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)}$ is defined analogously for policy $\check{\boldsymbol{\rho}}$.
- In each time period, let $\mathcal{B}_{(t)} \equiv \left\{\boldsymbol{x} \in \mathbb{R}^I : \boldsymbol{\pi}(\boldsymbol{x}) = \check{\boldsymbol{\rho}}(\boldsymbol{x})\right\}$ denote the set of starting states in which both policies make the same decision.
- Let $C_{(t)}(\boldsymbol{x} \to \boldsymbol{y}) \equiv \sum_{i=1}^{I} c^i_{(t)}(y^i_{(t)} - x^i_{(t)})^+ + d^i_{(t)}(y^i_{(t)} - x^i_{(t)})^- + Q^i_{(t)}(y^i_{(t)})$ denote the cost incurred

at time t when the state is $\boldsymbol{x}$ and the action is $\boldsymbol{y}$.

- Let the value function $V_{(t)}^{\boldsymbol{\pi}}(\boldsymbol{x})$ $[V_{(t)}^{\check{\boldsymbol{\rho}}}(\boldsymbol{x})]$ denote the cost incurred in our system from period $t, \cdots, T$ when we start period t in state $\boldsymbol{x}$ and follow policy $\boldsymbol{\pi}$ $[\check{\boldsymbol{\rho}}]$ thereafter.

We are now ready to state the following lemma.

Lemma 6.

$$C(\boldsymbol{\pi}) - C(\check{\boldsymbol{\rho}}) = \sum_{t=0}^{T} \gamma^t \mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}_{(t)}^{\boldsymbol{\pi}} \notin \mathcal{B}_{(t)})} \cdot \epsilon_{(t)}(\boldsymbol{X}_{(t)}^{\boldsymbol{\pi}})\right] \leq \sum_{t=0}^{T} \gamma^t \mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}_{(t)}^{\boldsymbol{\pi}} \notin \mathcal{B}_{(t)})} \cdot |\epsilon_{(t)}(\boldsymbol{X}_{(t)}^{\boldsymbol{\pi}})|\right] \tag{B.1}$$

where

$$\begin{aligned}\epsilon_{(t)}(\boldsymbol{x}) = C_{(t)}\Big(\boldsymbol{x} \to \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})\Big) - C_{(t)}\Big(\boldsymbol{x} \to \boldsymbol{\pi}_{(t)}(\boldsymbol{x})\Big)& \\ + \gamma \mathbb{E}\left[V_{(t+1)}^{\check{\boldsymbol{\rho}}}\Big(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big) - V_{(t+1)}^{\check{\boldsymbol{\rho}}}\Big(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big)\right]&\end{aligned} \tag{B.2}$$

denotes the expected additional cost of deviating from policy $\check{\boldsymbol{\rho}}$ to $\boldsymbol{\pi}$ in period t for one period only. Intuitively, this lemma states that as we move through our time horizon, only those states in which the two policies disagree need be considered in calculating the difference in their costs.

Proof. We begin by writing

$$V_{(t)}^{\check{\boldsymbol{\rho}}}(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{x} \in \mathcal{B}_{(t)}) V_{(t)}^{\check{\boldsymbol{\rho}}}(\boldsymbol{x}) + \mathbb{P}(\boldsymbol{x} \notin \mathcal{B}_{(t)}) V_{(t)}^{\check{\boldsymbol{\rho}}}(\boldsymbol{x})$$

Note that when $\boldsymbol{x} \in \mathcal{B}_{(t)}$, $\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) = \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})$, and so we can write

$$
\begin{aligned}
=&\mathbb{P}(\boldsymbol{x} \in \mathcal{B}_{(t)})\Big(C_{(t)}(\boldsymbol{x} \to \boldsymbol{\pi}_{(t)}(\boldsymbol{x})) + \gamma\mathbb{E}[V^{\check{\boldsymbol{\rho}}}_{(t+1)}(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)})]\Big) \\
&\qquad + \mathbb{P}(\boldsymbol{x} \notin \mathcal{B}_{(t)})\Big(C_{(t)}(\boldsymbol{x} \to \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})) + \gamma\mathbb{E}[V^{\check{\boldsymbol{\rho}}}_{(t+1)}(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)})]\Big) \\
=&C_{(t)}(\boldsymbol{x} \to \boldsymbol{\pi}_{(t)}(\boldsymbol{x})) + \gamma\mathbb{E}[V^{\check{\boldsymbol{\rho}}}_{(t+1)}(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)})]\mathbb{P}(\boldsymbol{x} \notin \mathcal{B}_{(t)})\Big(\\
&\underbrace{C_{(t)}(\boldsymbol{x} \to \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})) - C_{(t)}(\boldsymbol{x} \to \boldsymbol{\pi}_{(t)}(\boldsymbol{x})) + \gamma\mathbb{E}[V^{\check{\boldsymbol{\rho}}}_{(t+1)}(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}) - V^{\check{\boldsymbol{\rho}}}_{(t+1)}(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)})]}_{=\epsilon_{(t)}(\boldsymbol{x})}\Big)
\end{aligned}
\tag{B.3}
$$

Repeatedly substituting for the second term, and noting that $V^{\check{\boldsymbol{\rho}}}_{(T+1)}(\cdot) = 0$ proves our lemma.

□

Since for any pair of random variables A, B: $\mathbb{E}[AB] = \mathbb{E}[A]\mathbb{E}[B] + \mathbb{C}\mathrm{orr}[A,B]\mathbb{S}\mathrm{td}[A]\mathbb{S}\mathrm{td}[B] \le \mathbb{E}[A]\mathbb{E}[B] + \mathbb{S}\mathrm{td}[A]\mathbb{S}\mathrm{td}[B]$, Lemma 6 implies

$$
C(\boldsymbol{\pi}) - C(\check{\boldsymbol{\rho}}) \le \sum_{t=0}^{T} \gamma^t \left(\mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)})}\right]\mathbb{E}\left[\left|\epsilon_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)})\right|\right] + \mathbb{S}\mathrm{td}\left[\mathbf{1}_{(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)})}\right] \cdot \mathbb{S}\mathrm{td}\left[\left|\epsilon_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)})\right|\right]\right)
$$

Noting that the standard deviation of a Bernoulli variable with probability of success p is $\sqrt{p(1-p)} \le \sqrt{p}$, we get

$$
C(\boldsymbol{\pi}) - C(\check{\boldsymbol{\rho}}) \le \sum_{t=0}^{T} \gamma^t \left(\mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)})}\right]\mathbb{E}\left[\left|\epsilon_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)})\right|\right] + \sqrt{\mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)})}\right]} \cdot \mathbb{S}\mathrm{td}\left[\left|\epsilon_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)})\right|\right]\right)
$$

The rest of our proof proceeds as follows. First, in Section B.1, we prove that $\mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)})}\right] = O\left(\frac{1}{I}\right)$. Then, in Section B.1, we prove that both the mean and variance of $\left|\epsilon_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)})\right|$ are $O(I)$. Taken together, these yield

$$
C(\boldsymbol{\pi}) - C(\check{\boldsymbol{\rho}}) \le O\left(\frac{1}{I}\right)O(I) + O\left(\frac{1}{\sqrt{I}}\right)O(\sqrt{I}) = O(1).
$$

In other words, we show that even though the expected cost divergence between the two policies in any time period starting from any given state is $O(I)$, these divergences occur seldom enough to make the overall impact $O(1)$ only.

Proving that $\mathbb{E}\left[\mathbf{1}_{(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)})}\right] = \mathbb{P}\left(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)}\right) = O\left(\frac{1}{I}\right)$

By the construction of policy $\boldsymbol{\pi}$ from $\check{\boldsymbol{\rho}}$, it is immediate that

$$\mathcal{B}_{(t)} = \left\{\boldsymbol{x} \in \mathbb{R}^I : \ddot{g}_{(t|\boldsymbol{x})}(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})) \leq \chi_{(t+\ell)} - \kappa_0\right\}.$$

We first show that

$$\mathbb{P}\left(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \notin \mathcal{B}_{(t)}\right) \leq \mathbb{P}\left(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)}\right)$$

To do this, it suffices to show that $\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)} \leq \boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)}$ almost surely for every t. We do this by induction. Note $\boldsymbol{X}^{\boldsymbol{\pi}}_{(0)} = \boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(0)} = \boldsymbol{X}_{(0)}$. Suppose the inequality is true for a general period t. Then,

$$\boldsymbol{X}^{\boldsymbol{\pi}}_{(t+1)} = \boldsymbol{\pi}_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)}) - \boldsymbol{u}_{(t)} \leq \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)}) - \boldsymbol{u}_{(t)} \leq \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)}) - \boldsymbol{u}_{(t)} = \boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t+1)} \text{ a.s.}$$

where the first inequality follows from the construction of the policy $\boldsymbol{\pi}$ (see constraint (2.11b) in particular) and the second inequality from the monotonicity of $\check{\boldsymbol{\rho}}$ in its starting state (as a double base stock policy), see Theorem 1 in Chapter 1, combined with the induction assumption.

We now show that provided $(\kappa - \kappa_0)$ is large enough, $\mathbb{P}\left(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)}\right) = O\left(\frac{1}{I}\right)$, thus proving the desired result.

To do this, we will need some additional notation. Recall $\boldsymbol{u}_{(t)} \sim \mathcal{N}(\mu_{(t)}, \boldsymbol{\Sigma}_{(t)})$, and write $\boldsymbol{\Sigma}_{(t)} = \boldsymbol{S}^\top_{(t)}\boldsymbol{P}_{(t)}\boldsymbol{S}_{(t)}$, where $\boldsymbol{S}_{(t)} \equiv \operatorname{diag}(\boldsymbol{\sigma}_{(t)})$ and $\boldsymbol{P}_{(t)}$ is the correlation matrix. Let $\boldsymbol{L}_{(t)}$ be the Cholesky decomposition of the correlation matrix (i.e., $\boldsymbol{P}_{(t)} = \boldsymbol{L}_{(t)}\boldsymbol{L}^\top_{(t)}$), and let $\boldsymbol{L}^i_{(t)}$ denote the i-th row of $\boldsymbol{L}_{(t)}$. We can then write $\boldsymbol{u}_{(t)} = \boldsymbol{\mu}_{(t)} + \boldsymbol{S}_{(t)}\boldsymbol{L}_{(t)}\boldsymbol{Z}_{(t)}$, where $\{\boldsymbol{Z}_{(t)}\}_{t=0}^T$ is a process of standard Normal vectors that generates the demand process. We will write $\boldsymbol{Z}_{(0\to t)} \equiv \left(\boldsymbol{Z}^\top_{(0)}, \ldots, \boldsymbol{Z}^\top_{(t)}\right)^\top$. We now state and prove the following lemma.

Lemma 7. *Assume a given starting state $\boldsymbol{x}_{(0)}$, and consider two sequences of standard Normal vectors $\{\boldsymbol{Z}^1_{(t)}\}_{t=0}^T$ and $\{\boldsymbol{Z}^2_{(t)}\}_{t=0}^T$. Let $\{\boldsymbol{X}^1_{(t)}\}_{t=0}^T$ and $\{\boldsymbol{X}^2_{(t)}\}_{t=0}^T$ be the trajectories of the starting inventory positions under policy $\check{\boldsymbol{\rho}}$ under the demands generated by each of these two sequences respectively. Finally, let*

$\ddot{G}_{(t)}(\boldsymbol{Z}^1_{(0\to t-1)}) \equiv \ddot{g}_{(t|\boldsymbol{X}^1_{(t)})}(\check{\boldsymbol{\rho}}(\boldsymbol{X}^1_{(t)}))$ denote the trajectory of the LHS of constraints $(\ddot{7})$ after making a decision in each period employing policy $\check{\boldsymbol{\rho}}$ as a function of the underlying standard Normal random variables $\{\boldsymbol{Z}^1_{(t)}\}_{t=0}^T$, with $\ddot{G}_{(t)}(\boldsymbol{Z}^2_{(0\to t-1)})$ defined accordingly for $\{\boldsymbol{Z}^2_{(t)}\}_{t=0}^T$.

(a) For any i and t, we can bound the divergence between the two sample paths as follows:

$$|X^{1,i}_{(t)} - X^{2,i}_{(t)}| \le \sum_{\tau=0}^{t-1} \left|\sigma^i_{(\tau)} \boldsymbol{L}^i_{(\tau)}(\boldsymbol{Z}^1_{(\tau)} - \boldsymbol{Z}^2_{(\tau)})\right|. \tag{B.4}$$

(b) Let $\bar{\boldsymbol{L}}_{(t)} \equiv |\boldsymbol{L}_{(t)}|$ be the element-wise absolute value of $\boldsymbol{L}_{(t)}$. Then for any t,

$$\left|\ddot{G}_{(t)}(\boldsymbol{Z}^1_{(0\to t-1)}) - \ddot{G}_{(t)}(\boldsymbol{Z}^2_{(0\to t-1)})\right| \le \left(\sum_{\tau=0}^{t-1} \|\boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)}\|\right) \|\boldsymbol{Z}^1_{(0\to t-1)} - \boldsymbol{Z}^2_{(0\to t-1)}\|, \tag{B.5}$$

i.e. when applying policy $\check{\boldsymbol{\rho}}$, the left-hand-side of conditional constraint $(\ddot{7})$ after making decisions in each period is Lipschitz continuous in the underlying sample path $\boldsymbol{Z}_{(0\to t-1)}$ with Lipschitz constant $\sum_{\tau=0}^{t-1} \|\boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)}\|$.

(c) Under Assumption 4, the Lipschitz constant in the previous part satisfies

$$\sum_{\tau=0}^{t-1} \|\boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)}\| \le \bar{\sigma} t m \sqrt{I} = O(\sqrt{I}), \tag{B.6}$$

where m is the constant defined in Assumption 4, and $\bar{\sigma}$ is the upper bound on $\sigma^i_{(t)}$ for all i, t as defined in the notation section.

Proof. (a) We prove this by forward induction in t. At $t = 0$, $|X^{1,i}_{(0)} - X^{2,i}_{(0)}| = 0$ by assumption, and the inequality trivially holds. Assume the inequality holds for a general t.

$$\begin{aligned}
|X^{1,i}_{(t+1)} - X^{2,i}_{(t+1)}| &= \left| \left(\check{\rho}^i_{(t)}(X^{1,i}_{(t)}) - u^{1,i}_{(t)}\right) - \left(\check{\rho}^i_{(t)}(X^{2,i}_{(t)}) - u^{2,i}_{(t)}\right) \right| \\
&= |\check{\rho}^i_{(t)}(X^{1,i}_{(t)}) - \check{\rho}^i_{(t)}(X^{2,i}_{(t)}) - \sigma^i_{(t)} \boldsymbol{L}^i_{(t)}(\boldsymbol{Z}^1_{(t)} - \boldsymbol{Z}^2_{(t)})| \\
&\le |\check{\rho}^i_{(t)}(X^{1,i}_{(t)}) - \check{\rho}^i_{(t)}(X^{2,i}_{(t)})| + |\sigma^i_{(t)} \boldsymbol{L}^i_{(t)}(\boldsymbol{Z}^1_{(t)} - \boldsymbol{Z}^2_{(t)})| \\
&\le |X^{1,i}_{(t)} - X^{2,i}_{(t)}| + |\sigma^i_{(t)} \boldsymbol{L}^i_{(t)}(\boldsymbol{Z}^1_{(t)} - \boldsymbol{Z}^2_{(t)})|
\end{aligned}$$

This last inequality follows from the double base stock structure of $\check{\boldsymbol{\rho}}$. Indeed, applying $\check{\rho}^i_{(t)}$ could only grow the difference between the two paths if the action on both paths were different, and if the lesser of the two inventory positions were salvaged, and/or the greater of the two positions were replenished (thus increasing the gap). The structure of the policy, however, makes this impossible – if the lesser of the two inventory positions were above the salvage limit, the upper level would be too, and if the greater of the two inventory positions were below the replenishment limit, the lesser would be too. In both cases, they would both be salvaged or replenished to the same level, thus making the difference 0. The final possibility has the greater [lower] of the two inventory positions above [below] the salvage [order-up-to] level, and the difference of the after-ordering inventory positions is again reduced. Thus, the inequality holds. Finally, we can use the induction assumption to conclude

$$\leq \sum_{\tau=0}^{t} \left|\sigma^i_{(\tau)} \boldsymbol{L}^i_{(\tau)} (\boldsymbol{Z}^1_{(\tau)} - \boldsymbol{Z}^2_{(\tau)})\right| \tag{B.7}$$

This completes the induction proof.

(b) Begin by using the definition of $\ddot{g}$ in ($\ddot{7}$) to write

$$\begin{aligned} &\left|\ddot{G}_{(t)}(\boldsymbol{Z}^1_{(0\to t-1)}) - \ddot{G}_{(t)}(\boldsymbol{Z}^2_{(0\to t-1)})\right| \\ &\qquad = \mathbb{E}\left[\sum_{i=1}^{I} \left(\check{\rho}^i_{(t)}(X^1_{(t)}) - \dot{u}^i_{(t)}\right)^+ - \left(\check{\rho}^i_{(t)}(X^2_{(t)}) - \dot{u}^i_{(t)}\right)^+ \middle| \boldsymbol{Z}^1_{(0\to t-1)}, \boldsymbol{Z}^2_{(0\to t-1)}\right] \end{aligned}$$

Note that the expectation is taken with respect to the next lead time demand, conditional on having reached specific values of $X^i_{(t)}$ at time t. By considering the three possible rankings of $\left\{\check{\rho}^i_{(t)}(X^1_{(t)}), \check{\rho}^i_{(t)}(X^2_{(t)}), \dot{u}^i_{(t)}\right\}$, we can bound this as follows:

$$\leq \sum_{i=1}^{I} \left|\check{\rho}^i_{(t)}(X^{1,i}_{(t)}) - \check{\rho}^i_{(t)}(X^{2,i}_{(t)})\right|$$

Using the logic from part (a), we can further bound this as follows:

$$\leq \sum_{i=1}^{I} \left|X^{1,i}_{(t)} - X^{2,i}_{(t)}\right|$$

Letting $\boldsymbol{e}$ denote a vector of 1s and $\|\cdot\|$ denote the L_2-norm of a vector, we can write this as

$$= \boldsymbol{e}^\top \left| \boldsymbol{X}^1_{(t)} - \boldsymbol{X}^2_{(t)} \right|$$

Directly using part (a), on each i, we can write

$$\begin{aligned}
&\leq \sum_{\tau=0}^{t-1} \boldsymbol{e}^\top \left| \boldsymbol{S}_{(\tau)} \boldsymbol{L}_{(\tau)} (\boldsymbol{Z}^1_{(\tau)} - \boldsymbol{Z}^2_{(\tau)}) \right| \\
&\leq \sum_{\tau=0}^{t-1} \boldsymbol{e}^\top \boldsymbol{S}_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)} \left| \boldsymbol{Z}^1_{(\tau)} - \boldsymbol{Z}^2_{(\tau)} \right|
\end{aligned}$$

We now use the Cauchy-Schwartz inequality

$$\begin{aligned}
&\leq \sum_{\tau=0}^{t-1} \left\| \boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)} \right\| \cdot \left\| \boldsymbol{Z}^1_{(\tau)} - \boldsymbol{Z}^2_{(\tau)} \right\| \\
&\leq \left(\sum_{\tau=0}^{t-1} \left\| \boldsymbol{\sigma}^\top_{(\tau)} \bar{\boldsymbol{L}}_{(\tau)} \right\| \right) \cdot \left\| \boldsymbol{Z}^1_{(0\to t-1)} - \boldsymbol{Z}^2_{(0\to t-1)} \right\|
\end{aligned}$$

(c) It suffices to show $\|\boldsymbol{\sigma}^\top_{(t)} \bar{\boldsymbol{L}}_{(t)}\| \leq \bar{\sigma} m \sqrt{I} = O(\sqrt{I})$ for any t. Under Assumption 4, $\boldsymbol{P}_{(t)}$ is a block diagonal matrix, and so are $\boldsymbol{L}_{(t)}$ and $\bar{\boldsymbol{L}}_{(t)}$. Thus, each column of $\bar{\boldsymbol{L}}_{(t)}$ has at most m non-zero elements. Since $\boldsymbol{L}_{(t)} \boldsymbol{L}^\top_{(t)} = \boldsymbol{P}_{(t)}$, and the values of $\boldsymbol{P}_{(t)}$ are correlations bounded between 0 and 1, we conclude $\bar{\boldsymbol{L}}_{(t)} \leq 1$. Thus,

$$\|\boldsymbol{\sigma}^\top_{(t)} \bar{\boldsymbol{L}}_{(t)}\| \leq \bar{\sigma} \|\boldsymbol{e}^\top \bar{\boldsymbol{L}}_{(t)}\| \leq \bar{\sigma} \sqrt{m^2 I} = O(\sqrt{I}).$$

□

We are now ready to bound $\mathbb{P}\left(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)} \right)$.

$$\begin{aligned}
\mathbb{P}(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)}) &= \mathbb{P}\left(\ddot{G}_{(t)}(\boldsymbol{Z}_{(0\to t-1)}) \geq \chi_{(t+\ell)} - \kappa_0 \right) \\
&= \mathbb{P}\left(\ddot{G}_{(t)}(\boldsymbol{Z}_{(0\to t-1)}) \geq (\chi_{(t+\ell)} - \kappa) + (\kappa - \kappa_0) \right)
\end{aligned}$$

We now use the fact that policy $\check{\boldsymbol{\rho}}$ is constructed to satisfy $\check{g}_{(t|\boldsymbol{x}_{(0)})}\left(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)})\right) = \mathbb{E}\left[\ddot{G}_{(t)}(\boldsymbol{Z}_{(0\to t-1)})\right] \leq \chi_{(t+\ell)} - \kappa$. Thus, we can write

$$\mathbb{P}\left(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)}\right) \leq \mathbb{P}\left(\ddot{G}_{(t)}(\boldsymbol{Z}_{(0\to t-1)}) \geq \mathbb{E}\left[\ddot{G}_{(t)}(\boldsymbol{Z}_{(0\to t-1)})\right] + (\kappa - \kappa_0)\right)$$

We have therefore reduced our problem to that of finding the probability that a function $\ddot{G}_{(t)}$ of Normal random variables exceeds its mean by at least $\kappa - \kappa_0$. Furthermore, we have shown that the function is Lipschitz continuous in these variables, with a Lipschitz constant that does not exceed $\bar{\sigma}tm\sqrt{I}$. We can therefore directly apply Theorem 2.26 in Wainwright (2019) to obtain

$$\mathbb{P}\left(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)}\right) \leq 2\exp\left(-\frac{(\kappa-\kappa_0)^2}{2(\bar{\sigma}tm\sqrt{I})^2}\right)$$

Provided we set $\kappa - \kappa_0 \geq \sqrt{2}\bar{\sigma}mT\sqrt{I\log I}$, as prescribed in the statement of Theorem 5, we have

$$\begin{aligned}\mathbb{P}\left(\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t)} \notin \mathcal{B}_{(t)}\right) &\leq e^{-\log I}\\ &= O\left(\frac{1}{I}\right)\end{aligned}$$

Proving that the mean and variance of $|\epsilon_{(t)}(\boldsymbol{X}^{\boldsymbol{\pi}}_{(t)})|$ are $O(I)$

We begin with the following Lemma.

Lemma 8. *(a) Let $S^i_{(t)}$ be the order-up-to level under policy $\check{\boldsymbol{\rho}}$ (see Theorem 1 in Chapter 1). Let $\bar{S}_i = \max(X^{\check{\rho},i}_{(0)}, S^i_{(0)}, \cdots, S^i_{(T)})$, and $\bar{S} = \max_i \bar{S}^i$. Then*

$$X^{\check{\boldsymbol{\rho}},i}_{(t)} \leq \bar{S} + \sum_{\tau=0}^{t-1}(u^i_{(\tau)})^-$$

and

$$\check{\rho}^i_{(t)}(X^{\check{\boldsymbol{\rho}},i}_{(t)}) \leq \bar{S} + \sum_{\tau=0}^{t-1}(u^i_{(\tau)})^-$$

(b) Fix $t = 0, \ldots, T-1$. Given two different starting states $X^{\check{\boldsymbol{\rho}},i}_{(t+1)}$ and $Y^{\check{\boldsymbol{\rho}},i}_{(t+1)}$, let $\{X^{\check{\boldsymbol{\rho}},i}_{(\tau)}\}^T_{\tau=t+1}$ and

$\{Y^{\check{\boldsymbol{\rho}},i}_{(\tau)}\}_{\tau=t+1}^{T}$ *be the ensuing trajectories when following policy* $\check{\boldsymbol{\rho}}$*. Then, for any* $\tau \geq t+1$,

$$\check{\rho}^i_{(\tau)}(X^{\check{\boldsymbol{\rho}},i}_{(\tau)}) - \check{\rho}^i_{(\tau)}(Y^{\check{\boldsymbol{\rho}},i}_{(\tau)}) \leq X^{\check{\boldsymbol{\rho}},i}_{(\tau)} - Y^{\check{\boldsymbol{\rho}},i}_{(\tau)} \leq X^{\check{\boldsymbol{\rho}},i}_{(t+1)} - Y^{\check{\boldsymbol{\rho}},i}_{(t+1)}.$$

Proof. (a) We show this by induction. At $t = 0$, the two inequalities hold. Assume the two inequalities hold for general t. Then,

$$X^{\check{\boldsymbol{\rho}},i}_{(t+1)} = \check{\rho}^i_{(t)}\left(X^{\check{\boldsymbol{\rho}},i}_{(t)}\right) - u^i_{(t)} \leq \max\left[\bar{S}, X^{\check{\boldsymbol{\rho}},i}_{(t)}\right] + \sum_{\tau=0}^{t-1}(u^i_{(\tau)})^- + (u^i_{(t)})^- \leq \bar{S} + \sum_{\tau=0}^{t}(u^i_{(\tau)})^-,$$

and

$$\check{\rho}^i_{(t+1)}\left(X^{\check{\boldsymbol{\rho}},i}_{(t+1)}\right) \leq \max\left[\bar{S}, X^{\check{\boldsymbol{\rho}},i}_{(t+1)}\right] \leq \bar{S} + \sum_{\tau=0}^{t}(u^i_{(\tau)})^-.$$

This completes the induction step.

(b) We prove this again by induction. For $\tau = t+1$, the statement is true since $\check{\rho}^i_{(\tau)}(X^{\check{\boldsymbol{\rho}},i}_{(\tau)}) - \check{\rho}^i_{(\tau)}(Y^{\check{\boldsymbol{\rho}},i}_{(\tau)}) \leq X^{\check{\boldsymbol{\rho}},i}_{(\tau)} - Y^{\check{\boldsymbol{\rho}},i}_{(\tau)}$ as argued in the proof of Lemma 7(a). Assume it is true for an arbitrary value of $\tau \geq t+1$. Then,

$$\begin{aligned}
\check{\rho}^i_{(\tau+1)}(X^{\check{\boldsymbol{\rho}},i}_{(\tau+1)}) - \check{\rho}^i_{(\tau+1)}(Y^{\check{\boldsymbol{\rho}},i}_{(\tau+1)}) &\leq X^{\check{\boldsymbol{\rho}},i}_{(\tau+1)} - Y^{\check{\boldsymbol{\rho}},i}_{(\tau)} \\
&= \left(\check{\rho}^i_{(\tau)}(X^{\check{\boldsymbol{\rho}},i}_{(\tau)}) - u^i_{(\tau)}\right) - \left(\check{\rho}^i_{(\tau)}(Y^{\check{\boldsymbol{\rho}},i}_{(\tau)}) - u^i_{(\tau)}\right) \\
&= \check{\rho}^i_{(\tau)}(X^{\check{\boldsymbol{\rho}},i}_{(\tau)}) - \check{\rho}^i_{(\tau)}(Y^{\check{\boldsymbol{\rho}},i}_{(\tau)}) \\
&\leq X^{\check{\boldsymbol{\rho}},i}_{(t+1)} - Y^{\check{\boldsymbol{\rho}},i}_{(t+1)}.
\end{aligned}$$

$\square$

Recall the definition of $\epsilon_{(t)}(\boldsymbol{x})$ in (B.2). Let's begin with the first part; as before, let $\boldsymbol{e}$ be a vector of 1s

$$\begin{aligned}
\left|C_{(t)}\Big(\boldsymbol{x} \to \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})\Big) - C_{(t)}\Big(\boldsymbol{x} \to \boldsymbol{\pi}_{(t)}(\boldsymbol{x})\Big)\right| &\leq (\bar{h} + \bar{p} + \bar{c} + \bar{d}) \cdot \boldsymbol{e}^\top |\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{\pi}_{(t)}(\boldsymbol{x})| \\
&\leq (\bar{h} + \bar{p} + \bar{c} + \bar{d}) \cdot \boldsymbol{e}^\top |\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})|.
\end{aligned}$$

The last equality is justified component-wise thanks to the second part of equation (2.11b) in the construction of $\boldsymbol{\pi}$. Indeed, if $\check{\rho}^i_{(t)}(x^i) \leq 0$, then $\check{\rho}^i_{(t)}(x^i) = \pi^i_{(t)}(x^i)$, and the inequality holds. If $\check{\rho}^i_{(t)}(x^i) \geq 0$, then $\pi^i_{(t)}(x^i) \leq \check{\rho}^i_{(t)}(x^i)$ and the inequality holds as well.

Let us now consider the second part of $\epsilon_{(t)}(\boldsymbol{x})$, i.e. $\mathbb{E}\left[V^{\check{\boldsymbol{\rho}}}_{(t+1)}\Big(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big) - V^{\check{\boldsymbol{\rho}}}_{(t+1)}\Big(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big)\right]$. Applying Lemma 8(b) with $\boldsymbol{X}^{\check{\boldsymbol{\rho}}}_{(t+1)} = \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}$ and $\boldsymbol{Y}^{\check{\boldsymbol{\rho}}}_{(t+1)} = \boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}$, we can be sure that the sample paths followed in $V^{\check{\boldsymbol{\rho}}}_{(t+1)}\Big(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big)$ and $V^{\check{\boldsymbol{\rho}}}_{(t+1)}\Big(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big)$ under the policy $\check{\boldsymbol{\rho}}$ will never deviate by more than $\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{\pi}_{(t)}(\boldsymbol{x})$. As such, accounting for the costs in every period from $t+1$ onward, we can write

$$
\begin{aligned}
&\left|\mathbb{E}\left[V^{\check{\boldsymbol{\rho}}}_{(t+1)}\Big(\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big) - V^{\check{\boldsymbol{\rho}}}_{(t+1)}\Big(\boldsymbol{\pi}_{(t)}(\boldsymbol{x}) - \boldsymbol{u}_{(t)}\Big)\right]\right| \\
&\qquad\qquad \leq (T-t-1)(\bar{h}+\bar{p}+\bar{c}+\bar{d}) \cdot \boldsymbol{e}^\top |\check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x}) - \boldsymbol{\pi}_{(t)}(\boldsymbol{x})| \\
&\qquad\qquad \leq (T-t-1)(\bar{h}+\bar{p}+\bar{c}+\bar{d}) \cdot \boldsymbol{e}^\top \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})
\end{aligned}
$$

Together, these imply that

$$
\left|\epsilon_{(t)}(\boldsymbol{x})\right| \leq (T-t)(\bar{h}+\bar{p}+\bar{c}+\bar{d}) \sum_{i=1}^{I} \check{\boldsymbol{\rho}}_{(t)}(\boldsymbol{x})
$$

By Lemma 8(a), we can then write

$$
\left|\epsilon_{(t)}(\boldsymbol{x})\right| \leq (T-t)(\bar{h}+\bar{p}+\bar{c}+\bar{d}) \sum_{i=1}^{I} \left(\bar{S} + \sum_{\tau=0}^{t-1} (u^i_{(\tau)})^-\right)
$$

Thus, $\left|\epsilon_{(t)}(\boldsymbol{x})\right|$ is bounded by a multiple of a sum of tI truncated Normal random variables with means in $[t\underline{\mu}, t\bar{\mu}]$ and standard deviations in $[\sqrt{t}\underline{\sigma}, \sqrt{t}\bar{\sigma}]$, and thus has finite first and second moments, which, given the correlation pattern specified in Assumption 4, grow as $O(I)$. This concludes our proof.

B.2 An Asymptotically Optimal Heuristic Under General Demand Distributions with Bounded Support

In this section, we consider demands with general distributions. Instead of Normality, as assumed in the body of the paper, we merely assume that the demands have general continuous distributions with bounded support, WLOG in $[0, \nu]$. (Actual demands are, generally, non-negative.) The heuristic is constructed as in the Normal case and its asymptotic optimality proof maintained with minor adaptions.

First, the capacity deduction κ_0 in equation ($\ddot{7}$) requires a different choice of the parameter σ, still with $\sigma = O(\sqrt{I})$. Recall that in the Normal case, σ arose from a concentration inequality that pertains to Lipschitz continuous functions of Normal random variables, as shown in Lemma 7. Instead we prove an analogue to Lemma 7.

Lemma 7′. *Let* $U := \max\{X, \bar{S}\}$ *and* $\sigma := \frac{m\sqrt{I}U}{2}$. *Let* $\{\bar{x}^i_{(t)}\}$ *be the inventory process after ordering under policy* $\boldsymbol{\pi}$. *Let* $Y_{(t)} = \sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$ *be the total inventory on hand at the beginning of period* $(t+\ell)$ *under policy* $\boldsymbol{\pi}$. *Then*

$$\mathbb{P}\big(Y_{(t)} - \mathbb{E}[Y_{(t)}] \geq y\big) \leq e^{-\frac{y^2}{2\sigma^2}}.$$

Proof. Let $x^{*i}_{(t)}$ and $\bar{x}^{*i}_{(t)}$ denote the inventory position of item i before and after ordering in period t, under policy $\check{\boldsymbol{\rho}}^{\star}_{\kappa}$. We first show that

$$\bar{x}^{*i}_{(t)} \leq U \text{ a.s. for all } i \text{ and } t = 1, \ldots, T. \tag{B.8}$$

We prove this inequality by forward induction with respect to t. $x^{*i}_{(0)} \leq X \leq U$, so the inequality holds for $t = 0$. Assume (B.8) holds for a general period t. Let $S^{i,\boldsymbol{\lambda}^*}_{(t)}$ be the base stock order-up-to level of policy $\check{\rho}^{\star i}_{\kappa,(t)}$ in period t. If $x^{*i}_{(t+1)} \leq S^{i,\boldsymbol{\lambda}^*}_{(t+1)}$, then $\bar{x}^{*i}_{(t+1)} = S^{i,\boldsymbol{\lambda}^*}_{(t+1)} \leq S^{i,\mathbf{0}}_{(t+1)} \leq \bar{S} \leq U = \max\{X, \bar{S}\}$, where the first and second inequality follow from Theorem 1(a) and 1(b) in Chapter 1, respectively. Alternatively, $x^{*i}_{(t+1)} > S^{i,\boldsymbol{\lambda}^*}_{(t+1)}$, in which case $\bar{x}^{*i}_{(t+1)} \leq x^{*i}_{(t+1)} = \bar{x}^{*i}_{(t)} - u^i_{(t)} \leq \bar{x}^{*i}_{(t)} \leq U$, where the first inequality follows from the double base stock structure of the policy $\check{\rho}^{\star i}_{\kappa,(t+1)}$, the second inequality from the fact that the demand is non-negative, and the last inequality from the induction assumption. This completes the induction proof of (B.8).

By the construction of policy $\boldsymbol{\pi}$, we have $\bar{x}^i_{(t)} \leq \bar{x}^{*i}_{(t)}$ a.s. so that (B.8) implies:

$$\bar{x}^i_{(t)} \leq U \text{ a.s. for all } i \text{ and } t = 1, \ldots, T. \tag{B.9}$$

We have $Y_{(t)} = \sum_{i=1}^{I} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+ = \sum_{k=1}^{K} Z_k$, where $Z_k \equiv \sum_{i \in G_k} (\bar{x}^i_{(t)} - \dot{u}^i_{(t)})^+$ and $0 \leq Z_k \leq mU$.

Applying the Hoeffding bound to the independent and bounded random variables $\{Z_k\}$ (see e.g. Proposition 2.5 in Wainwright (2019)), we have

$$\mathbb{P}(Y_{(t)} - \mathbb{E}[Y_{(t)}] \geq y) = \mathbb{P}(\sum_{k=1}^{K} (Z_k - \mathbb{E}[Z_k]) \geq y) \leq e^{-\frac{y^2}{2K(mU/2)^2}} \leq e^{-\frac{y^2}{2I(mU/2)^2}} = e^{-\frac{y^2}{2\sigma^2}},$$

where $\sigma = \frac{\sqrt{I}mU}{2}$. $\square$

The solution of the ($\check{\text{DP}}_\kappa$), and the construction of the policy $\check{\boldsymbol{\rho}}^*_\kappa$ and $\boldsymbol{\pi}$ proceed without any modifications, with the new choice of the σ given by the specification of κ and κ_0. The same applies to Theorem 1 in Chapter 1. To prove the feasibility of policy $\boldsymbol{\pi}$ in the original DP under the original chance constraints (2.2), we apply Lemma 7$'$, instead of Lemma 4(b). The remainder of the proof is identical, except for the justification of the inequality $C^\star \geq C^\star_{-\alpha}$, where we need to make slight modifications to Lemma 5. With random demand variables $u^i_{(t)}$ that have bounded support on an interval $[0, \nu]$, clearly $\bar{x}^i_{(t)} \leq (T-t)\nu \leq T\nu$ since at most $(T-t)\nu$ units will be needed over the planning horizon. This implies that the random variables $\{Z_k\}$ are still independent and bounded by $(mT\nu)$. Thus, the Hoeffding inequality (2.14) continues to apply, merely replacing σ by $\sigma' \equiv \frac{1}{2}\sqrt{I}(mT\nu)$ which is still $O\sqrt{I}$. Thus, replacing α in Theorem 6 by $\bar{\alpha} \equiv \frac{\beta\sigma'}{1-\beta}\left(\sqrt{-2\log\left(\frac{\beta}{2}\right)} + \sqrt{\pi/2}\right)$, Theorem 6 and its proof continue to apply.

B.3 Proof of Theorem 7

In this section, we prove Theorem 7, which states that the optimal policy $\pi^{i,\lambda}_{(t)}$ to dynamic program (2.19) follows the following double base stock structure

$$\pi^{i,\boldsymbol{\lambda}}_{(t)}(x) = \begin{cases} S^{i,\boldsymbol{\lambda}}_{(t)}, & x < S^{i,\boldsymbol{\lambda}}_{(t)} \quad \text{(B.10a)} \\ x, & x \in [S^{i,\boldsymbol{\lambda}}_{(t)}, B^{i,\boldsymbol{\lambda}}_{(t)}] \quad \text{(B.10b)} \\ B^{i,\boldsymbol{\lambda}}_{(t)}, & x > B^{i,\boldsymbol{\lambda}}_{(t)} \quad \text{(B.10c)} \end{cases}$$

The proof is very similar to that of Theorem 1 in Chapter 1.

Fix $i = 1, \dots, I$. To simplify the notation, we omit the superscript indices i and λ. In view of (2.1), one easily verifies that it is never optimal to order and salvage in the same period; thus $V_{(t)}(x_{(t)}) = \min\{V^1_{(t)}(x_{(t)}), V^2_{(t)}(x_{(t)})\}$, where

$$V^1_{(t)}(x_{(t)}) = -c_{(t)}x_{(t)} + \min_{\bar{x}_{(t)} \geq x_{(t)}} \left\{ c_{(t)}\bar{x}_{(t)} + Q_{(t)}(\bar{x}_{(t)}) + \gamma\mathbb{E}[V_{(t+1)}(\bar{x}_{(t)} - u_{(t)})] + \sum_{\boldsymbol{a}\in\mathcal{A}} \lambda^{\boldsymbol{a}}_{(t)} a(\bar{x}_{(t)}) \right\}, \quad \text{(B.11)}$$

and

$$V^2_{(t)}(x_{(t)}) = d_{(t)}x_{(t)} + \min_{\min\{x_{(t)},0\} \leq \bar{x}_{(t)} \leq x_{(t)}} \Bigg\{ - d_{(t)}\bar{x}_{(t)} + Q_{(t)}(\bar{x}_{(t)}) + \gamma\mathbb{E}[V_{(t+1)}(\bar{x}_{(t)} - u_{(t)})] + \sum_{\boldsymbol{a}\in\mathcal{A}} \lambda^{\boldsymbol{a}}_{(t)} a(\bar{x}_{(t)}) \Bigg\}. \quad \text{(B.12)}$$

We prove the theorem by (backwards) induction with respect to t. The statement trivially holds for $t = T - \ell + 1$ with $V_{T-\ell+1} = 0$. Assume it holds for some period $(t+1)$ with $t = 1, \dots, T - \ell$. The minimand of (B.11) is a strictly convex function $H^1(\bar{x}_{(t)})$ of $\bar{x}_{(t)}$ only, since the function $V_{(t+1)}(\cdot)$ is convex, by the induction assumption, while $Q_{(t)}$ is strictly convex. Moreover, $H^1(\cdot)$ is differentiable, with derivative:

$$H^{1'}(\bar{x}_{(t)}) = c_{(t)} + Q'_{(t)}(\bar{x}_{(t)}) + \gamma\mathbb{E}[V'_{(t+1)}(\bar{x}_{(t)} - u_{(t)})] + \sum_{\boldsymbol{a}\in\mathcal{A}} \lambda^{\boldsymbol{a}}_{(t)} a, \quad \text{(B.13)}$$

where both the existence of the derivative and the interchange of the expectation and derivative operators are justified, by the fact that $|V'_{(t+1)}(y)| \leq A|y|^r$, for some constant $A > 0$. Thus $\mathbb{E}\left[V'_{(t+1)}(\bar{x}_{(t)} - u_{(t)})\right] \leq A\mathbb{E}\left[|\bar{x}_{(t)} - u_{(t)}|^r)\right] \leq A\mathbb{E}\left[(|\bar{x}_{(t)}| + |w_{(t)}|)^r\right] \leq A\sum_{l=0}^r \binom{r}{l} |\bar{x}_{(t)}|^r \mathbb{E}\left[|u_{(t)}|^{r-l}\right] < \infty$.

Moreover, $\lim\limits_{x\uparrow\infty} H^{1'}(x) > 0$, since $c_{(t)} > 0$, $\lim\limits_{x\uparrow\infty} Q'_{(t)}(x) > 0$, $\lim\limits_{x\uparrow\infty} V'_{(t+1)}(x) \geq 0$ and $\sum_{\boldsymbol{a}\in\mathcal{A}} \lambda^{\boldsymbol{a}} a \geq 0$. Thus, the strictly convex function ${H^1}_{(t)}(\cdot)$ achieves its unique minimum at

$$-\infty \leq S_{(t)} = \inf\{x : {H^{1}}'_{(t)}(x) > 0\} < \infty$$

since the infimum is taken over a non-empty set. It follows that for $x_{(t)} < S_{(t)}$, it is optimal to place an order to elevate $\bar{x}_{(t)}$ to $S_{(t)}$, thus verifying (B.10a).

By the same argument, the function $H^2_{(t)}(\cdot)$ is a strictly convex function with

$$H^{2'}(\bar{x}_{(t)}) = -d_{(t)} + Q'_{(t)}(\bar{x}_{(t)}) + \gamma \mathbb{E}[V'_{(t+1)}(\bar{x}_{(t)} - u_{(t)})] + \sum_{\boldsymbol{a} \in \mathcal{A}} \lambda^{\boldsymbol{a}}_{(t)} a, \tag{B.14}$$

a strictly increasing function which is positive for $\bar{x}_{(t)}$ sufficiently large. Thus $H^2_{(t)}(\cdot)$ achieves its minimum at $\tilde{B}_{(t)}$, defined as

$$\tilde{B}_{(t)} = \inf\{x : {H^{2}}'_{(t)}(x) > 0\} < \infty.$$

Moreover, by (2.1) ${H^{2}}'_{(t)}(\cdot)$ is pointwise smaller than ${H^{1}}'_{(t)}(\cdot)$ so that $S_{(t)} < \tilde{B}_{(t)} \leq \max\{\tilde{B}_{(t)}, 0\}$. Thus, for $x_{(t)} \geq B_{(t)}$, it is optimal to reduce the inventory position to $B_{(t)}$, thus verifying (B.10b). (It is infeasible to reduce the inventory position below the zero-level.)

Finally, for $S_{(t)} < x < \tilde{B}_{(t)}$, it is optimal not to order or to salvage, and the same applies for $\tilde{B}_{(t)} \leq x < B_{(t)}$ since salvaging is infeasible on this interval.

In addition,

$$V_{(t)}(x_{(t)}) = \begin{cases} c_{(t)}(S_{(t)} - x_{(t)}) + Q_{(t)}(S_{(t)}) + \gamma \mathbb{E}\left[V_{(t+1)}(S_{(t)} - u_{(t)})\right] \\ \qquad + \sum_{\boldsymbol{a} \in \mathcal{A}} \lambda^{\boldsymbol{a}} a S_{(t)}, \text{ if } x < S^{i,\lambda}_{(t)} & \text{(B.15a)} \\ Q_{(t)}(x_{(t)}) + \gamma \mathbb{E}\left[V_{(t+1)}(x_{(t)} - u_{(t)})\right] \\ \qquad + \sum_{\boldsymbol{a} \in \mathcal{A}} \lambda^{\boldsymbol{a}} a x_{(t)}, \text{ if } x \in [S^{i,\lambda}_{(t)}, B^{i,\lambda}_{(t)}] & \text{(B.15b)} \\ d_{(t)}(x_{(t)} - B_{(t)}) + Q_{(t)}(B_{(t)}) + \gamma \mathbb{E}\left[V_{(t+1)}(B_{(t)} - u_{(t)})\right] \\ \qquad + \sum_{\boldsymbol{a} \in \mathcal{A}} \lambda^{\boldsymbol{a}} a B_{(t)}, \text{ if } x > B^{i,\lambda}_{(t)} & \text{(B.15c)} \end{cases}$$

showing that $V_{(t)}(\cdot)$ is differentiable everywhere, with the possible exception of the points $x_{(t)} = S_{(t)}$ and $x_{(t)} = B_{(t)}$, while $V'_{(t)}(x) = O(|x|^r)$, and $\lim_{x\uparrow\infty} V'_{(t)}(x) \geq 0$. This completes the induction step.

B.4 Proof of Proposition 1

Proof. We prove this by showing that for $\tau = 1, \ldots, t^0$,

$$\frac{\partial V^\lambda_{(\tau)}(\boldsymbol{x}_{(\tau)})}{\partial \lambda^{\boldsymbol{a}}_{(t^0)}} = \gamma^{t^0-\tau}\mathbb{E}[\boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t^0)} - C^{\boldsymbol{a}}_{(t^0)}] \tag{B.16}$$

and that for any $\tau = t^0 + 1, \ldots, T - \ell$,

$$\frac{\partial V^\lambda_{(\tau)}(\boldsymbol{x}_{(\tau)})}{\partial \lambda^{\boldsymbol{a}}_{(t^0)}} = 0. \tag{B.17}$$

The proof of (B.17) is immediate. The proof of (B.16) follows by backwards induction for $\tau = t^0, t^0 - 1, \cdots, 1$. As shown in the proof of Theorem 7, the minimand in (2.19) for $t = t^0$ is a differentiable function of $\bar{\boldsymbol{x}}_{(t)}$. Moreover, for all $i = 1, \ldots, I$ and $x^i_{(t)}$, there exists a unique value $\bar{x}^i_{(t)}$ which achieves the minimum in (2.19). It then follows from Danskin's Theorem, that

$$\frac{\partial V^\lambda_{(t^0)}(\boldsymbol{x}_{(t)})}{\partial \lambda^{\boldsymbol{a}}_{(t^0)}} = \boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t^0)} - C^{\boldsymbol{a}}_{(t^0)}, \text{ verifying (B.16) for } \tau = t^0.$$

Assume (B.16) holds for some $2 \leq \tau \leq t^0$. As above, the minimand in (2.19) for $t = \tau - 1$ is a differentiable function of $\bar{\boldsymbol{x}}_{(t)}$, and for any $\boldsymbol{x}_{(t)}$, there exists a unique vector $\bar{\boldsymbol{x}}_{(t)}$ which achieves the minimum in (2.19). It then follows from Danskin's Theorem that

$$\begin{aligned}\frac{\partial V^\lambda_{(t)}(\boldsymbol{x}_{(t)})}{\partial \lambda^{\boldsymbol{a}}_{(t^0)}} &= \gamma\gamma^{t^0-\tau}\mathbb{E}[\boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t^0)} - C^{\boldsymbol{a}}_{(t^0)}|\boldsymbol{x}_{(t)}] \\ &= \gamma^{t^0-(\tau-1)}\mathbb{E}[\boldsymbol{a} \cdot \bar{\boldsymbol{x}}_{(t^0)} - C^{\boldsymbol{a}}_{(t^0)}|\boldsymbol{x}_{(t)}]\end{aligned}$$

thus completing the induction proof. $\square$

B.5 Proof of Theorem 8

We provide the proof for the case where the backlogging and holding cost functions $p^i_{(t)}(\cdot)$ and $h^i_{(t)}(\cdot)$ are linear with unit cost rates $p^i_{(t)}$ and $h^i_{(t)}$, i.e. the degree of these polynomially bounded function $r = 1$. The proof for general $r \geq 1$ is analogous.

We first need the following lemma.

Lemma 9. *Let* $p^* = \max_{i,t} p^i_{(t)}, c^* = \max_{i,t} c^i_{(t)}$.

(a) $\lambda^{*\boldsymbol{a}}_{(t)} \leq p^* + c^*$ *under linear holding and backlogging cost rates.*

(b) The Hessian $\frac{\partial^2 V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})}{\partial \lambda^{\boldsymbol{a}}_{(t)} \partial \lambda^{\boldsymbol{b}}_{(s)}}$ *exists almost everywhere. Moreover,* $\frac{\partial^{2+} V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})}{\partial \lambda^{\boldsymbol{a}}_{(t)} \partial \lambda^{\boldsymbol{b}}_{(s)}}$ *exists everywhere, where right hand second derivative denotes the right hand derivative of* $\frac{\partial V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})}{\partial \lambda^{\boldsymbol{a}}_{(t)}}$ *with respect to* $\lambda^{\boldsymbol{b}}_{(s)}$.

(c)

$$Tr[-\frac{\partial^{2+} V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})}{\partial \lambda^{\boldsymbol{a}}_{(t)} \partial \lambda^{\boldsymbol{b}}_{(s)}}] = \begin{cases} O(I|\mathcal{A}|), & \text{if } \gamma < 1, \quad \text{(B.18a)} \\ O(I|\mathcal{A}|T), & \text{if } \gamma = 1 \quad \text{(B.18b)} \end{cases}$$

Proof. (a) $\lambda^{\boldsymbol{a}}_{(t)}$ denotes the amount by which $V^{\lambda}_{(1)}$ decreases if only the capacity size $C^{\boldsymbol{a}}_{(t)}$ were increased by one unit leaving all other capacity levels, associated with all other constraints, unaltered. The 1 unit increase of $C^{\boldsymbol{a}}_{(t)}$ may allow for a 1 unit in the-end-of-period t inventory level, assigned to a single or a combination of items. The benefit of the availability of these additional inventory increments is to eliminate one unit of backlog, during the current or some future period, a benefit which is bounded by p^*. Additionally, the increase of $C^{\boldsymbol{a}}_{(t)}$ by 1 unit, may save some a future procurement cost rate, an additional benefit bounded by c^*.

(b) It follows from Proposition 1 that

$$\frac{\partial^{+} V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})}{\partial \lambda^{\boldsymbol{a}}_{(t)}} = \gamma^{t-1}\mathbb{E}[\boldsymbol{a}\bar{x}_{(t)} - C^{\boldsymbol{a}}_{(t)}] = \gamma^{t-1}\mathbb{E}[\sum_{j:\boldsymbol{a}^j=1} \bar{x}^j - C^{\boldsymbol{a}}_{(t)}].$$

Justifying an interchange of differentiation and expectation operators as in the proof of Proposition 1, we get

$$\frac{\partial^2 V^{\boldsymbol{a}}_{(1)}}{\partial {\lambda^{\boldsymbol{a}}_{(t)}}^2} = \gamma^{t-1}\mathbb{E}[\sum_{j:\boldsymbol{a}^j=1} \frac{\partial \bar{x}^j_{(t)}}{\partial \lambda^{\boldsymbol{a}}_{(t)}}]. \tag{B.19}$$

If $\bar{x}^j_{(t)}$ is in the interior of the feasible region $[L^j_{(t)}, U^j_{(t)}]$, where $L^j_{(t)}$ is an imposed lower bound on the system inventory position after ordering and $U^j_{(t)} = \chi_{(t)} + \dot{\mu}^j_{(t)} + \dot{\sigma}^j_{(t)}\Phi^{-1}(\beta)$, it is the unique root of (B.13), so that, by the Implicit Function Theorem,

$$\frac{\partial \bar{x}^j_{(t)}}{\partial \lambda^{\boldsymbol{a}}_{(t)}} = \frac{\partial^+ \bar{x}^j_{(t)}}{\partial \lambda^{\boldsymbol{a}}_{(t)}} = \begin{cases} -\dfrac{1}{Q''^j(\bar{x}^j_{(t)}) + \gamma\mathbb{E}[V''^j_{(t+1)}(\bar{x}^j_{(t)} - \dot{u}^j_{(t)})]}, & \text{if } \boldsymbol{a}^j = 1, \quad \text{(B.20a)} \\ 0, & \text{if } \boldsymbol{a}^j \neq 1. \quad \text{(B.20b)} \end{cases}$$

On the other hand, if $\bar{x}^j_{(t)}$ is at the boundary of the feasible region, an infinitesimal change of $\lambda^{\boldsymbol{a}}_{(t)}$ may either has no impact on $\bar{x}^j_{(t)}$ or $\frac{\partial^+ \bar{x}^j_{(t)}}{\partial \lambda^{\boldsymbol{a}}_{(t)}}$ is given by the expression in (B.20a).

(c) Since the value function is convex, we show that the second term in the denominator of (B.20a) is non-negative so that

$$\frac{\partial \bar{x}^j_{(t)}}{\partial \lambda^{\boldsymbol{a}}_{(t)}} \leq \left[(p^* + h^*)\phi(\frac{\bar{x}^j_{(t)} - \dot{\mu}^j_{(t)}}{\dot{\sigma}^j_{(t)}})\right]^{-1} \leq (p^* + h^*)^{-1}\left[\min_{L^j_{(t)} \leq x \leq U^j_{(t)}} \phi(\frac{x - \dot{\mu}^j_{(t)}}{\dot{\sigma}^j_{(t)}})\right]^{-1},$$

where $\dot{\mu}^j_{(t)}$ and $\dot{\sigma}^j_{(t)}$ denote the mean and standard deviation of item j's lead time demand.

Since the standard Normal density function $\phi(\cdot)$ is quasi-convex, the minimum within square brackets is $O(1)$. It follows that each of the expression in (B.20a) is bounded by a constant κ, independent of i and t. Summing the diagonal elements of all $\boldsymbol{a} \in \mathcal{A}$, for a fixed value of t, this sum is bounded by $\kappa I|\mathcal{A}|\gamma^{t-1}$. Finally, the trace of this Hessian, is itself bounded by

$$Tr[\mathcal{H}] \leq \textstyle\sum_{t=1}^{T} \kappa I|\mathcal{A}|\gamma^{t-1} \leq \begin{cases} \dfrac{\kappa I|\mathcal{A}|}{1-\gamma}, & \text{if } \gamma < 1, \quad \text{(B.21a)} \\ \kappa I|\mathcal{A}|T, & \text{if } \gamma = 1. \quad \text{(B.21b)} \end{cases}$$

□

Proof of Theorem (8): we first consider the case where $\gamma < 1$. Note, first that, $V^{\lambda^*}_{(1)}(\boldsymbol{x}_{(1)}) \geq V^0_{(1)}(\boldsymbol{x}_{(1)}) = \Omega(I)$, since $V^0_{(1)}(\boldsymbol{x}_{(1)}) = \sum_{i=1}^{I} V^{0,i}_{(1)}(x_{(1)}) \geq \frac{I\nu}{1-\gamma}$, where ν represents minimum across all items i and periods t of the expected cost value in a news vendor problem assuming each period could be started with the inventory position which is optimal to cover the next lead time interval. Since the expected cost in incurred period t are weighted by a discount factor γ^{t-1}, this proves the lower bound $\frac{I\nu}{1-\gamma}$.

Thus, the complexity to achieve $\epsilon - relative$ optimality is guaranteed to the complexity of obtaining an absolute ϵI-optimal lower bound $V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})$:

$$V^{\lambda^*}_{(1)}(\boldsymbol{x}_{(1)}) - V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)}) \leq \epsilon I$$

For ease of notation, define $f(\lambda) = -V^{\lambda}_{(1)}(\boldsymbol{x}_{(1)})$. Let L be a Lipschitz constant of ∇f, i.e., $\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$, where for any vector z, $\|z\|$ denotes its Euclidean norm.

To access the worst case overall complexity of the FISTA-based Lagrangian Dual optimization procedure, we bound both the complexity of executing a single FISTA iteration, as well as the number of iterations required.

Complexity of a Single Iteration

In each iteration, I separate dynamic programs need to be solved. With a bounded state- and action-space, each of this dynamic program requires $O(|\mathcal{A}|T)$ operations. In addition, the gradient has $|\mathcal{A}|T$ components, and evaluation of each component requires $O(I)$ time. Thus, the complexity of each FISTA iteration is $O(I|\mathcal{A}|T)$.

Worst Case Bound for Number of Iterations

Beck and Teboulle (2009) show that the number of iterations to achieve ϵ-absolute optimality is $O(\frac{\sqrt{L}R}{\sqrt{\epsilon}})$, hence to achieve ϵI-absolute optimality is $O(\frac{\sqrt{L}R}{\sqrt{I\epsilon}})$, where

$$R^2 = \|\lambda_0 - \lambda^*\| = \|\lambda^*\| \leq \sum_{t=1}^{T} |\mathcal{A}| \max_{\boldsymbol{a},t} [\lambda^{*\boldsymbol{a}}_{(t)}]^2 = O(|\mathcal{A}|T)$$

by Lemma 9(a). The Lipschitz constant L is bounded by the $\sup_{\mathbf{0} \leq \lambda \leq p^* + c^*} \rho(\mathcal{H}^{\lambda})$, the supremum over all feasible λ of the largest eigenvalue of the Hessian.

Since V^{λ} is concave, $-V^{\lambda}$ is convex and its Hessian $\mathcal{H}$ is therefore symmetric and semi-positive definite. This implies that its eigenvalues are real and non-negative so that $L \leq \sup_{\mathbf{0} \leq \lambda \leq p^*+c^*} Tr[\mathcal{H}] = O(I|\mathcal{A}|)$. Thus, the worst case bound for the number of iterations is $O(\frac{\sqrt{L}R}{\sqrt{I\epsilon}}) = O(\frac{|\mathcal{A}|T^{\frac{1}{2}}}{\sqrt{\epsilon}})$. Multiplying this bound with the complexity of the computations in each iteration, i.e. $O(I|\mathcal{A}|T)$, we get the bound $O(\frac{I|\mathcal{A}|^2T^{\frac{3}{2}}}{\sqrt{\epsilon}})$.

When $\gamma = 1$, the analysis remains the same except that $V_{(1)}^{\lambda^*}(\boldsymbol{x}_{(1)}) = \Omega(IT)$, so ϵ-relative optimality corresponds with ϵIT-absolute optimality. At the same time, Lemma 9(c) shows that, in this case $L = O(|\mathcal{A}|IT)$. Thus the number of iterations remains $O(\frac{|\mathcal{A}|T^{\frac{1}{2}}}{\sqrt{\epsilon}})$, while the complexity of each iteration is unaffected by the value of $\gamma \leq 1$. In other words, the overall complexity bound $O(\frac{I|\mathcal{A}|^2T^{\frac{3}{2}}}{\sqrt{\epsilon}})$ continues to apply.

Appendix C: Proofs of Chapter 3

C.1 Proof of Theorem 9

Proof. For any given $\boldsymbol{\lambda}$, we can write the relaxed DP defining $V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ as

$$\begin{aligned} V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}) = \min_{\bar{\boldsymbol{x}},\hat{\boldsymbol{x}},\bar{\boldsymbol{w}},\boldsymbol{f}} \mathbb{E}\Big[\sum_{t=0}^{T}\Big\{ & c_{(t)}\bar{w}_{(t)} + \sum_{j\in\mathcal{L}^c} h_{j,(t)}\bar{x}_{j,(t)} + \sum_{j\in\mathcal{L}} Q_{j,(t)}(\bar{x}_{j,(t)} + \hat{x}_{j,(t)}) + \boldsymbol{\gamma}_{(t)}^{\top}\boldsymbol{f}_{(t)} \\ & + \boldsymbol{\lambda}_{(t)}^{\top}(A^{-}\boldsymbol{f}_{(t)} - (\boldsymbol{x}_{(t)} - \bar{\boldsymbol{x}}_{(t)}))\Big\}\Big|\boldsymbol{x}_{(0)}\Big] \\ \text{s.t.} \quad & \boldsymbol{x}_{(t+1)} = \bar{\boldsymbol{w}}_{(t)} + \bar{\boldsymbol{x}}_{(t)} + \hat{\boldsymbol{x}}_{(t)} - \boldsymbol{\mu}_{(t)} \\ & \text{constraints (3.1)} - \text{(3.5)}. \end{aligned}$$

Define $\boldsymbol{y} := (\bar{\boldsymbol{x}}, \hat{\boldsymbol{x}}, \bar{\boldsymbol{w}}, \boldsymbol{f})$ and write $V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}) := \min_{\boldsymbol{y}} g(\boldsymbol{y}, \boldsymbol{\lambda})$ for some function g.

(a) For any $\boldsymbol{y}$, $g(\boldsymbol{y}, \boldsymbol{\lambda})$ is a linear function of $\boldsymbol{\lambda}$. Thus, $\min_{\boldsymbol{y}} g(\boldsymbol{y}, \boldsymbol{\lambda})$ is a concave function of $\boldsymbol{\lambda}$.

(b) Consider any $\boldsymbol{\lambda}, \boldsymbol{\lambda}'$ such that $\boldsymbol{\lambda}_{(t)} \neq \boldsymbol{\lambda}'_{(t)}$ and $\boldsymbol{\lambda}_{(\tau)} = \boldsymbol{\lambda}'_{(\tau)}$ for all $\tau \neq t$. Let $\boldsymbol{y}^{\lambda} := \operatorname{argmin}_{\boldsymbol{y}} g(\boldsymbol{y}, \boldsymbol{\lambda})$ and $\boldsymbol{y}^{\lambda'} := \operatorname{argmin}_{\boldsymbol{y}} g(\boldsymbol{y}, \boldsymbol{\lambda}')$, respectively. Then,

$$\begin{aligned} V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}) &= g(\boldsymbol{y}^{\lambda}, \boldsymbol{\lambda}) \\ &= g(\boldsymbol{y}^{\lambda}, \boldsymbol{\lambda}') - \mathbb{E}\Big[(\boldsymbol{\lambda}'_{(t)} - \boldsymbol{\lambda}_{(t)})^{\top}(A^{-}\boldsymbol{f}^{\lambda}_{(t)} - (\boldsymbol{x}^{\lambda}_{(t)} - \bar{\boldsymbol{x}}^{\lambda}_{(t)}))\Big|\boldsymbol{x}_{(0)}\Big] \\ &\geq g(\boldsymbol{y}^{\lambda'}, \boldsymbol{\lambda}') - (\boldsymbol{\lambda}'_{(t)} - \boldsymbol{\lambda}_{(t)})^{\top}\mathbb{E}\Big[(A^{-}\boldsymbol{f}^{\lambda}_{(t)} - (\boldsymbol{x}^{\lambda}_{(t)} - \bar{\boldsymbol{x}}^{\lambda}_{(t)}))\Big|\boldsymbol{x}_{(0)}\Big] \\ &= V^{\lambda'}_{(0)}(\boldsymbol{x}_{(0)}) - (\boldsymbol{\lambda}'_{(t)} - \boldsymbol{\lambda}_{(t)})^{\top}\nabla_{\boldsymbol{\lambda}_{(t)}} V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)}), \end{aligned}$$

where the inequality follows from that $\boldsymbol{y}^{\lambda'}$ is a minimizer of $g(\cdot, \boldsymbol{\lambda}')$. Thus, $\nabla_{\boldsymbol{\lambda}_{(t)}} V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ is a supergradient of $V^{\lambda}_{(0)}(\boldsymbol{x}_{(0)})$ with respect to $\boldsymbol{\lambda}_{(t)}$. $\square$