Approximation Algorithms for Stochastic Inventory Control Models

Retsef Levi* Martin Pal † Robin Roundy‡ David B. Shmoys§

Abstract

We consider stochastic control inventory models in which the goal is to coordinate a sequence of orders of a single commodity, aiming to supply stochastic demands over a discrete finite horizon with minimum expected overall ordering, holding and backlogging costs. In this paper, we address the long-standing problem of finding computationally efficient and provably good inventory control policies to these models in the presence of correlated and non-stationary (time-dependent) stochastic demands. This problem arises in many domains and has many practical applications in supply chain management. We consider two classical models, the *periodic-review stochastic inventory control problem* and the *stochastic lot-sizing problem* with correlated and non-stationary demands. Here the correlation is inter-temporal, i.e., what we observe in period s changes our forecast for the demand in future periods. We provide what we believe to be the first computationally efficient policies with constant worst-case performance guarantees; that is, there exists a constant C such that, for any instance of the problem, the expected cost of the policy is at most C times the expected cost of an optimal policy.

The dominant paradigm in almost all of the existing literature has been to formulate these models using a dynamic programming framework. This approach has turned out to be very successful in characterizing the structure of the optimal policies, which follow simple forms of state-dependent base-stock policies and state-dependent (s, S) policies. However, in case the demands are non-stationary and correlated over time, computing these optimal policies is likely to be intractable.

We present a new approach that leads to general approximation algorithms with constant performance guarantee for these classical models. Our approach is based on several novel ideas: we present a new (marginal) cost accounting for stochastic inventory models; we use cost-balancing techniques; and we consider non base-stock (order-up-to) policies that are extremely easy to implement on-line. Our results are valid for all of the currently known approaches in the literature to model correlation and non-stationarity of demands over time.

More specifically, we provide a general 2-approximation algorithm for the periodic-review stochastic inventory control problem and a 3-approximation algorithm for the stochastic lot-sizing problem. That is, the constant guarantees are 2 and 3, respectively. For the former problem, we show that the classical myopic policy can be arbitrarily more expensive compared to the optimal policy. We also present an extended class of myopic policies that provides both upper and lower bounds on the optimal base-stock levels.

*`rl227@cornell.edu`. School of ORIE, Cornell University, Ithaca, NY 14853. Research supported partially by a grant from Motorola and NSF grant CCR-9912422.

†`mpal@cs.cornell.edu` Dept. of Computer Science, Cornell University, Ithaca, NY 14853.

‡`robin@orie.cornell.edu`. School of ORIE, Cornell University, Ithaca, NY 14853. Research supported partially by a grant from Motorola, NSF grant DMI-0075627, and the Querétaro Campus of the Instituto Tecnológico y de Estudios Superiores de Monterrey.

§`shmoys@cs.cornell.edu`. School of ORIE and Dept. of Computer Science, Cornell University, Ithaca, NY 14853. Research supported partially by NSF grant CCR-9912422.

1 Introduction

In this paper we address the long-standing problem of finding computationally efficient and provably good inventory control policies in supply chains with correlated and non-stationary (time-dependent) stochastic demands. This problem arises in many domains and has many practical applications (see for example [3, 6]). We consider two classical models, the *periodic-review stochastic inventory control problem* and the *stochastic lot-sizing problem* with correlated and non-stationary demands. Here the correlation is inter-temporal, i.e., what we observe in period s changes our forecast for the demand in future periods. We provide what we believe to be the first computationally efficient policies with constant worst-case performance guarantees; that is, there exists a constant C such that, for any instance of the problem, the expected cost of the policy is at most C times the expected cost of an optimal policy.

A major domain of applications in which demand correlation and non-stationarity are commonly observed is where dynamic demand forecasts are used as part of the supply chain. Demand forecasts often serve as an essential managerial tool, especially when the demand environment is highly dynamic. The problem of how to use a demand forecast that evolves over time to devise an efficient and cost-effective inventory control policy is of great interest to managers, and has attracted the attention of many researchers over the years. However, it is well known that such environments often induce high correlation between demands in different periods that makes it very hard to compute the optimal inventory policy. Another relevant and important domain of applications is for new products and/or new markets. These scenarios are often accompanied by an intensive promotion campaign and involve many uncertainties, which create high levels of correlation and non-stationarity in the demands over time. Correlation and non-stationarity also arise for products with strong cyclic demand patterns, and as products are phased out of the market.

The two classical stochastic inventory control models considered in this paper capture many if not most of the application domains in which correlation and non-stationarity arise. More specifically, we consider single-item models with one location and a finite planning horizon of T discrete periods. The demands over the T periods are random variables that can be non-stationary and correlated. In the periodic-review stochastic inventory control problem, the cost consists of per-unit, time-dependent ordering cost, holding cost for carrying excess inventory from period to period and backlogging cost, which is a penalty we incur for each unit of unsatisfied demand (where all shortages are fully backlogged). In addition, there is a lead time between the time an order is placed and the time that it actually arrives. In the stochastic lot-sizing problem, we consider, in addition, a fixed ordering cost that is incurred in each period in which an order is placed (regardless of its size), but with no lead time. In both models, the goal is to find a policy of orders with minimum expected overall discounted cost over the given planning horizon. The assumptions that we make on the demand distributions are very mild and generalize all of the currently known approaches in the literature to model correlation and non-stationarity of demands over time. This includes classical approaches like the martingale modulated forecast evolution model (MMFE), exogenous Markovian demand, time series, order-one auto-regressive demand and random walks. For an overview of the different approaches and models, and for relevant references, we refer the reader to [4, 7]. Moreover, we believe that the models we consider are general enough to capture almost any other reasonable way of modelling correlation and non-stationarity of demands over time.

These models have attracted the attention of many researchers over the years and there exists a huge body of related literature. The dominant paradigm in almost all of the existing literature has been to formulate these models using a dynamic programming framework. The optimization problem is defined recursively over time using subproblems for each possible state of the system. The state usually consists of a given time period, the level of the echelon inventory at the beginning of the period, a given conditional distribution on the future demands over the rest of the horizon, and possibly more information that is available by time t. For each subproblem, we compute an optimal solution to minimize the expected overall discounted cost from time t until the end of the horizon.

This framework has turned out to be very effective in characterizing the optimal policy of the overall system. Surprisingly, the optimal policies for these rather complex models follow simple forms. In the models with only per-unit ordering cost, the optimal policy is a *state-dependent base-stock policy*. In each period, there exists an optimal target base-stock level that is determined only by the given conditional distribution (at that period) on future demands and possibly by additional information that is available, but it is independent of the starting inventory level at the beginning of the period. The optimal policy aims to keep the inventory level at each period as close as possible to the target base-stock level. That is, it orders up to the target level whenever the inventory level at the beginning of the period is below that level, and orders nothing otherwise. We note that Iida and Zipkin have shown the optimality of state-dependent base-stock policies only for the special cases of the MMFE model [4]. However, it seems that their results can be generalized to show the optimality of state-dependent base-stock levels in more general cases.

For the models with fixed ordering cost, the optimal policy follows a slightly more complicated pattern. Now, in each period, there are lower and upper thresholds that are again determined only by the given conditional distribution (at that period) on future demands. The optimal policy places an order in a certain period if and only if the inventory level at the beginning of the period has dropped below the lower threshold. Once an order is placed, the inventory level is increased up to the upper threshold. This class of policies is usually called *state-dependent* (s, S) *policies*. We note that the optimality of state-dependent (s, S) policies was proven for the case of non-stationary but independent demand (see [17]). We are not aware of such a proof for the case where demand in different periods can be correlated. We refer the reader to [7, 4, 17] for the details on some of the results along these lines, as well as a comprehensive discussion of relevant literature.

Unfortunately, these rather simple forms of policies do not always lead to efficient algorithms for computing the optimal policies. This is especially true in the presence of correlated and non-stationary demands which cause the state space of the relevant dynamic programs to grow exponentially and explode very fast. The difficulty essentially comes from the fact that we need to solve 'too many' subproblems. This phenomena is known as *the curse of dimensionality*. Moreover, because of this phenomenon, it seems unlikely that there exists an efficient algorithm to solve these huge dynamic programs. This gap between the excellent knowledge on the structure of the optimal policies and the inability to compute them efficiently provides the stimulus for future theoretical interest in these problems.

For the periodic-review stochastic inventory control problem, Muharremoglu and Tsitsiklis [12] have proposed an alternative approach to the dynamic programming framework. They have observed that this problem can be decoupled into a series of *unit supply-demand subproblems*, where each subproblem corresponds to a single unit of supply and a single unit of demand that are matched together. This novel approach enabled them to substantially simplify some of the dynamic programming based proofs on the structure of optimal policies, as well as to prove several important new structural results. Using this unit decomposition, they have also suggested new methods to compute the optimal policies. However, their computational methods are essentially dynamic programming approaches applied to the unit subproblems, and hence they suffer from similar problems in the presence of correlated and non-stationary demand. Although our approach is very different than theirs, we use some of their ideas as technical tools in some of the proofs in the paper.

As a result of this apparent computational intractability, many researchers have attempted to construct computationally efficient (but suboptimal) heuristics for these problems. However, we are aware of very few attempts to analyze the worst-case performance of these heuristics (see for example [8]). Moreover, we are aware of no computationally efficient policies for which there exist constant performance guarantees. For details on some of the proposed heuristics and a discussion of others, see [7, 8, 4]. One specific class of suboptimal policies that has attracted a lot of attention is the class of *myopic policies*. In a myopic policy, in each period we attempt to minimize the expected cost for that period, ignoring the impact on the cost in future periods. The myopic policy is attractive since it yields a base-stock policy that is easy to compute on-line, that is, it does not require information on the control policy in the future periods. In each period, we need to solve a one-variable convex minimization problem. In many cases, the myopic policy seems to perform well. However, in many other cases, especially when the demand can drop significantly from period to period, the myopic policy performs poorly. Myopic policies were extensively explored by Vienott [16], Zipkin [17], Iida and Zipkin [4] and Lu, Song and Regan [8]. In [4, 8], they have focused on the martingale modulated evolution forecast model and shown necessary conditions and rather strong sufficient conditions for myopic policies to be optimal. They have also used myopic policies to compute upper and lower bounds on the optimal base-stock levels, as well as bounds on the relative difference between the optimal cost and the cost of different heuristics. However, the bounds they provide on this relative error are not constants.

Chan and Muckstadt [1] have considered a different way for approximating huge dynamic programs that arise in the context of inventory control problems. More specifically, they have considered un capacitated and capacitated multi-item models. Instead of solving the one period problem (as in the myopic policy) they have added to the one period problem a penalty function which they call Q-function. This function accounts for the holding cost incurred by the inventory left at the end of the period over the entire horizon. Their look ahead approach with respect to the holding cost is somewhat related to our approach, though significantly different.

We note that our work is also related to a huge body of approximation results for stochastic and on-line combinatorial problems. The work on approximation results for stochastic combinatorial problems goes back to the work of Mohring, Radermacher and Weiss [9, 10] and the more recent work of Mohring, Schulz and Uetz [11]. They have considered stochastic scheduling problems. However, their performance guarantees are dependent on the specific distributions (namely on second moment information). Recently, there is a growing stream of approximation results for several 2-stage stochastic combinatorial problems. For a comprehensive literature review we refer the reader to [15, 2, 13]. We note that the problems we consider in this paper are by nature *multi-stage* stochastic problems, which are usually much harder.

Our work is distinct from the existing literature in several significant ways, and is based on three novel ideas:

Marginal cost accounting. We introduce a novel approach for cost accounting in stochastic inventory control problems. The key observation is that once we place an order of a certain number of units in some period, then the expected ordering and holding cost that these units are going to incur over the rest of the planning horizon is a function only of the realized demands over the rest of the horizon, not of future orders. Hence, with each period, we can associate the overall expected ordering and holding cost that is incurred by the units ordered in this period, over the entire horizon. This new way of marginal cost accounting is significantly different from the dynamic programming approach, which, in each period, accounts only for the costs that are incurred in that period. We believe that this new approach will have more applications in the future in analyzing stochastic inventory control problems.

Cost balancing. The idea of cost balancing was used in the past to construct heuristics with constant performance guarantees for deterministic inventory problems. The most well-known examples are the Silver-Meal heuristic for the lot-sizing problem (see [14]) and the Cost-Covering heuristic of Joneja for the joint-replenishment problem [5]. We are not aware of any application of these ideas to stochastic inventory control problems. For the periodic-review stochastic inventory control problem, we use the marginal cost accounting approach to construct a policy that, in each period, balances the expected (marginal) ordering and holding cost against the expected backlogging cost in that period. For the stochastic lot-sizing problem, we construct a policy that balances the expected fixed ordering cost, holding cost and backlogging cost over each interval between consecutive orders.

Non base-stock policies. Our policies are not state-dependent base-stock policies. This enable us to use, in each period, the distributional information about the future demands beyond the current period (unlike the myopic policy), without the burden of solving huge dynamic programs. Moreover, our policies can be easily implemented on-line and are simple, both conceptually and computationally.

Using these ideas we provide what is called a 2-approximation algorithm for the periodic-review stochastic inventory control problem; that is, the expected cost of our policy is no more than twice the expected cost of an optimal policy. Note that this is not the same requirement as stipulating that, for each realization of the demands, the cost of our policy is at most twice the optimal cost, which is a much more stringent requirement. We also note that this guarantee refers only to the worst-case performance and it is likely that on average the performance would be significantly better. We then use a standard cost transformation to achieve significantly better guarantees if the ordering cost is the dominant part in the overall cost, as it is the case in many real life situations. Our result is valid for all known approaches used to model correlated and non-stationary demands. For the periodic-review stochastic inventory control problem, we also present an extended class of myopic policies that provides easily computed upper bounds and lower bounds on the optimal base-stock levels.

An interesting question that is left open in the current literature is whether the myopic policy has a constant worst-case performance guarantee. We provide a negative answer to this question, by showing a family of examples in which the expected cost of the myopic policy can be arbitrarily more expensive than the expected cost of an optimal policy. Our example provides additional insight into situations in which the myopic policy performs poorly.

For the stochastic lot-sizing problem we provide a 3-approximation algorithm. This is again a worst-case analysis and we would expect the typical performance to be much better.

The rest of the paper is organized as follows. In Section 2 we present a mathematical formulation of the periodic-review stochastic inventory control problem. Then in Section 3 we explain the details of our new marginal cost accounting approach. In Section 4 we describe a 2-approximation algorithm for the periodic-review stochastic inventory control problem. In Section 5 we present an extended class of myopic policies for this problem, develop upper and lower bounds on the optimal base-stock levels, and discuss the example in which the performance of the myopic policy is arbitrarily bad. The stochastic lot-sizing problem is discussed in Section 6, where we present a 3-approximation algorithm for the problem. We then conclude with some remarks and open research questions.

2 The Periodic-Review Stochastic Inventory Control Problem

In this section, we provide the mathematical formulation of the periodic-review stochastic inventory problem and introduce some of the notation used throughout the paper. As a general convention throughout the paper, we distinguish between a random variable and its realization using capital letters and lower case letters, respectively. Script font is used to denote sets. We consider a finite planning horizon of T periods numbered $t = 1, \ldots, T$. The demands over these periods are random variables, denoted by $D_1, \ldots, D_T$.

As part of the model, we will assume that at the beginning of each period s, we are given what we call an *information set* that is denoted by f_s. The information set f_s contains all of the information that is available at the beginning of time period s. More specifically, the information set f_s consists of the realized demands $(d_1, \ldots, d_{s-1})$ over the interval $[1, s)$, and possibly some more (external) information denoted by $(w_1, \ldots, w_s)$. The information set f_s in period s is one specific realization in the set of all possible realizations of the random vector $(D_1, \ldots, D_{s-1}, W_1, \ldots, W_s)$. This set is denoted by $\mathcal{F}_s$. In addition, we assume that in each period s there is a known conditional joint distribution of the future demands $(D_s, \ldots, D_T)$, denoted by $I_s := I_s(f_s)$, which is determined by f_s (i.e., knowing f_s, we also know $I_s(f_s)$). For ease of notation, D_t will always denote the random demand in period t according to the conditional joint distribution I_s for some $s \leq t$, where it will be clear from the context to which period s we refer. We will use t as the general index for time, and s will always refer to the period we are currently in.

The only assumption on the demands is that for each $s = 1, \ldots, T$, and each $f_s \in \mathcal{F}_s$ the conditional expectation $E[D_t|f_s]$ is well defined and finite for each period $t \geq s$. In particular, we allow non-stationarity and correlation between the demands of different periods. We note again that by allowing correlation we let I_s be dependent on the realization of the demands over the periods $1, \ldots, s-1$ and possibly on some other information that becomes available by time s (i.e., I_s is a function of f_s). However, the conditional joint distribution I_s is assumed to be independent of the specific inventory control policy being considered.

In the periodic-review stochastic inventory control problem our goal is to supply each unit of demand while attempting to avoid ordering it either too early or too late. At the end of period t ($t = 1, \ldots, T$) three types of costs are incurred, a per-unit ordering cost c_t for ordering any number of units in period t, a per-unit holding cost h_t for holding excess inventory from period t to $t+1$, and a unit backlogging penalty p_t that is incurred for each unsatisfied unit of demand at the end of period t. Unsatisfied units of demand are usually called *back orders*. The assumption is that back orders fully accumulate over time until they are satisfied. That is, each unit of unsatisfied demand will stay in the system and will incur a backlogging penalty in each period until it is satisfied. In addition, we consider a model with a lead time of L periods between the time an order is placed and the time at which it actually arrives. We first assume that the lead time is a known integer L. In Section 4, we will show that our policy can be modified to handle stochastic lead times under the assumption of no order crossing (i.e., any order arrives no later than orders placed later in time).

There is also a discount factor $\alpha \leq 1$. The cost incurred in period t is discounted by a factor of α^t. Since the horizon is finite and the cost parameters are time-dependent, we can assume without loss of generality that $\alpha = 1$. We also assume that there are no speculative motivations for holding inventory or having back orders in the system. To enforce this, we will assume that for each $t = 1, \ldots, T-L$, the inequalities $c_t + h_{t+L} \geq c_{t+1}$ and $c_t \leq c_{t+1} + p_{t+L}$ are maintained (where $C_{T+1} = 0$). We also assume that the parameters $h_t,\ p_t$ and c_t are all non-negative. We note that the parameters h_T and p_T can be defined to take care of excess inventory and back orders at the end of the planning horizon. In particular, p_T can be set to be high enough to ensure that there are very few back orders at the end of time period T. In Section 4, we will show how to relax the non-negativity requirement and incorporate a salvage value at the end of the horizon (i.e., excess inventory at the end of the horizon can be sold back).

The goal is to find a policy that minimizes the overall expected discounted ordering cost, holding cost and backlogging cost. We consider only policies that are *non-anticipatory*, i.e., at time s, the information that a feasible policy can use consists only of f_s.

Throughout the paper we will use $D_{[s,t]}$ to denote the accumulated demand over the interval $[s,t]$, i.e., $D_{[s,t]} := \sum_{j=s}^{t} D_j$. We will also use superscripts P and OPT to refer to a given policy P and the optimal policy respectively.

2.1 System Dynamics

Given a feasible policy P, we describe the dynamics of the system using the following terminology. We let NI_t denote the *net inventory* at the end of period t, which can be either positive (in the presence of physical on-hand inventory) or negative (in the presence of back orders). Since we consider a lead time of L periods, we also consider the orders that are on the way. The sum of the units included in these orders, added to the current net inventory is referred to as the *inventory position* of the system. We let X_t be the inventory position at the beginning of period t *before* the order in period t is placed, i.e., $X_t := NI_{t-1} + \sum_{j=t-L}^{t-1} Q_j$ (for $t = 1, \ldots, T$), where Q_j denotes the number of units ordered in period j (we will sometime denote $\sum_{j=t-L}^{t-1} Q_j$ by $Q_{[t-L,t-1]}$). Similarly, we let Y_t be the inventory position *after* the order in period t is placed, i.e., $Y_t = X_t + Q_t$. Note that once we know the policy P and the information set $f_s \in \mathcal{F}_s$, we can easily compute ni_{s-1}, x_s and y_s, where again these are the realizations of NI_{s-1}, X_s and Y_s, respectively.

Since time is discrete, we next specify the sequence of events in each period s:

1. The order placed in period $s-L$ of q_{s-L} units arrives and the net inventory level increases accordingly to $ni_{s-1} + q_{s-L}$.

2. The decision of how many units to order in period s is made, i.e., following a given policy P, q_s units are ordered and consequently the inventory position is raised by q_s units (from x_s to y_s). This incurs a linear cost $c_s q_s$.

3. We observe the the realized demand in period s which is realized according to the conditional joint distribution I_s. We also observe the new information set $f_{s+1} \in \mathcal{F}_{s+1}$, and hence we also know the updated conditional joint distribution I_{s+1}. The net inventory and the inventory position each decrease by d_s units. In particular, we have $x_{s+1} = x_s + q_{s-L} - d_s$.

4. If $ni_s > 0$, then we incur a holding cost $h_s ni_s$ (this means that there is excess inventory that needs to

be carried to time period $s+1$). On the other hand, if $ni_s < 0$ we incur a backlogging penalty $p_t|ni_s|$ (this means that there are currently unsatisfied units of demand).

3 Marginal Cost Accounting

In this section, we will present a new approach to the cost accounting of stochastic inventory control problems. Our approach differs from the traditional dynamic programming based approach. In particular, we account for the holding cost incurred by a feasible policy in a different way, which enables us to design and analyze new approximation algorithms. We believe that this approach will be useful in other stochastic inventory models.

3.1 Dynamic Programming Framework

Traditionally, stochastic inventory control problems of the kind described in Section 2 are formulated using a dynamic programming framework. For simplicity, we discuss the case with $L = 0$ (for a detailed discussion see Zipkin [17]). For each period s we consider a given state, which usually consists of the initial inventory position x_s at the beginning of period s and the given information set $f_s \in \mathcal{F}_s$, and as a function of f_s the joint conditional distribution of future demands, I_s. Note that given an information set f_s the inventory position x_s can be computed for each given policy P. The space of possible decisions consists of the number of units to be ordered at time s or equivalently the level $y_s \geq x_s$ to which the inventory position is increased. This decision incurs a cost that is traditionally divided into two parts. The first part is the *immediate cost* incurred in period s, i.e., the ordering cost and the expected backlogging (in case of shortage) or holding cost (in case of excess inventory) at the end of period s. The second part is the *future cost* that accounts for the overall expected cost over the rest of the horizon. The decision that was made in period s will impact the starting inventory position in period $s+1$, namely $X_{s+1} = y_s - D_s$. For each possible combination of a period $t = 1, \ldots, T$ and an information set $f_t \in \mathcal{F}_t$, we seek to find an optimal policy for the interval $[t, T]$. In other words, we wish to order q_t^{OPT} units to minimize the expected discounted cost over $[t, T]$, assuming that in future periods we are going to make optimal decisions.

Observe that the cost accounting in the dynamic programming framework is done in an additive manner, period by period. In each period t, we account for the ordering cost and expected holding and backlogging costs that are incurred in period t and the cost over the interval $(t, T+1]$ (where $T+1$ is a dummy period). In other words, we associate with the decision in period t the cost incurred in period t.

As was noted in Section 1, this yields an optimal base-stock policy, $\{R(f_t) : f_t \in \mathcal{F}_t\}$. Given that the information set at time s is f_s, then the optimal base-stock level is $R(f_s)$. The optimal policy then follows the following pattern. In case the inventory position level at the beginning of period s is lower than $R(f_s)$ (i.e., $x_s < R(f_s)$), then the inventory position is increased to $y_s = R(f_s)$ by placing an order of the appropriate number of units. In case $x_s \geq R(f_s)$, the inventory position is kept the same (i.e., nothing is ordered) and $y_s = x_s$. However, the set $\mathcal{F}_s$ can be exponentially large or infinite. Thus, computing the optimal policy involves solving recursively exponentially many or even an infinite number of subproblems, which is intractable.

3.2 Marginal Accounting of Cost

We take a different approach for accounting for the holding cost associated with each period. Observe that once we decide to order q_s units at time s (where $q_s = y_s - x_s$), then the holding cost they are going to incur from period s until the end of the planning horizon is independent of any future decision in subsequent time periods. It is dependent only on the demand to be realized over the time interval $[s, T]$.

To make this rigorous, we use a *ground distance-numbering scheme* for the units of demand and supply, respectively. More specifically, we think of two infinite lines, each starting at 0, the *demand line* and the *supply line*. The demand line $\mathcal{L}_D$ represents the units of demands that can be potentially realized over the planning horizon, and similarly, the supply line $\mathcal{L}_S$ represents the units of supply that can be ordered over the planning horizon. Each 'unit' of demand, or supply, now has a *distance-number* according to its respective distance from the origin of the demand line and the supply line, respectively. If we allow continuous demand (rather then discrete) and continuous order quantities the unit and its distance-number are defined infinitesimally. We can assume without loss of generality that the units of demands are realized according to increasing distance-number. For example, if the accumulated realized demand up to time t is $d_{[1,t)}$ and the realized demand in period t is d_t, we then say that the demand units numbered $(d_{[1,t)}, d_{[1,t)} + d_t]$ were realized in period t. Similarly, we can describe each policy P in terms of the periods in which it orders each supply unit, where all unordered units are "ordered" in period $T + 1$. It is also clear that we can assume without loss of generality that the supply units are ordered in increasing distance-number. Specifically, the supply units that are ordered in period t are numbered
$(ni_0 + q_{[1-L,t)}, ni_0 + q_{[1-L,t]}]$, where ni_0 and $q_j,\ 1 - L \leq j \leq 0$ are the net inventory and the sequence of the last L orders, respectively, given as an input at the beginning of the planning horizon (in time 0). We can further assume (again without loss of generality) that as the demand is realized, the units of supply are consumed on a *first-ordered-first-consumed basis*. Therefore, we can *match* each unit of supply that is or-

dered to a certain unit of demand that has the same number. We note that Muharremoglu and Tsitsiklis have used the idea of matching units of supply to units of demand in a novel way to characterize and compute the optimal policy in different stochastic inventory models. However, their computational method is based on applying dynamic programming to the single-unit problems. Therefore, their cost accounting within each single-unit problem is still additive, and differs fundamentally from ours.

Suppose now that at the beginning of period s with observed information set f_s. Assume that the inventory position is x_s and q_s additional units are ordered. Then the expected additional (marginal) holding cost that these q_s units are going to incur from time period s until the end of the planning horizon is equal to $\sum_{j=s+L}^{T} E[h_j(q_s - (D_{[s,j]} - x_s)^+)^+|f_s]$ (recall that we assume without loss of generality that $\alpha = 1$), where $x^+ = \max(x, 0)$. Here we assume again that in time s we know a given joint distribution I_s of the demands $(D_s, \dots, D_T)$.

Using this approach, consider any feasible policy P and let $H_t^P := H_t^P(Q_t^P)$ ($t = 1, \dots, T$) be the discounted ordering and holding cost incurred by the additional Q_t^P units ordered in period t by policy P. Thus, $H_t^P = H_t^P(Q_t^P) := c_t Q_t^P + \sum_{j=t+L}^{T} h_j(Q_t^P - (D_{[t,j]} - X_t)^+)^+$. Now let B_t^P be the discounted backlogging cost incurred in period $t+L$ ($t = 1-L, \dots, T-L$). In particular, $B_t^P := p_{t+L}(D_{[t,t+L]} - (X_{t+L} + Q_t^P))^+$ (where $D_j := 0$ with probability 1 for each $j \leq 0$, and $Q_t^P = q_t$ for each $t \leq 0$). Let $\mathcal{C}(P)$ be the cost of the policy P. Clearly,

$$\mathcal{C}(P) := \sum_{t=1-L}^{0} B_t^P + H_{[1,L]} + \sum_{t=1}^{T-L} (H_t^P + B_t^P),$$

where $H_{[1,L]}$ denotes the total holding cost incurred over the interval $[1, L]$ (by units ordered before period 1). We note that the first two expressions $\sum_{t=1-L}^{0} B_t^P$ and $H_{[1,L]}$ are not affected by our decisions (i.e., they are the same for any feasible policy and each realization of the demand), and therefore we will omit them. Since they are non-negative, this will not effect our results. Also observe that without loss of generality, we can assume that $Q_t^P = H_t^P = 0$ for any policy P and each period $t = T-L+1, \dots, T$, since nothing that is ordered in these periods can be used within the given planning horizon. We now can write $\mathcal{C}(P) = \sum_{t=1}^{T-L}(H_t^P + B_t^P)$. In some sense, we change the accounting of the holding cost from periodical to marginal. As we will demonstrate in the sections to come, this new approach serves as a powerful tool for designing simple approximation algorithms that can be analyzed with respect to their worst-case expected performance.

4 Dual-Balancing Policy

In this section, we consider a new policy for the periodic-review stochastic inventory control problem, which we call a *dual-balancing policy*. In this policy we aim to balance the expected marginal ordering and holding cost against the expected marginal backlogging cost. In each period $s = 1, \ldots, T-L$, we focus on the units that we order in period s only, and balance the expected ordering cost and holding cost they are going to incur over $[s, T]$ against the expected backlogging cost in period $s+L$. We do that using the marginal accounting of the holding cost as introduced in Section 3.

We next describe the details of the policy, which is very simple to implement, and then analyze its expected performance. In particular, we will show that for any input of demand distributions and cost parameters, the expected cost of the dual-balancing policy is at most twice the expected cost of an optimal policy. A superscript B will refer to the dual-balancing policy described below. At the end of this section we will show how a simple transformation of the costs can yield a better worst-case performance guarantee and certainly a better typical (average) performance in many cases in practice.

4.1 The Algorithm

We first describe the algorithm and its analysis in case the demands have density and fractional orders are allowed. Later on, we will show how to extend the algorithm and the analysis to the case in which the demands and the order sizes are integer-valued. In each period $s = 1, \ldots, T-L$, we consider a given information set f_s (where again $f_s \in \mathcal{F}_s$) and the resulting pair (x_s^B, I_s) of the inventory position at the beginning of period s and the conditional joint distribution I_s of the demands $(D_s, \ldots, D_T)$. We then consider the following two functions:

(i) The expected ordering cost and holding cost over $[s, T]$ that is incurred by the *additional* q_s units ordered in period s, conditioning on f_s. We denote this function by $l_s^B(q_s)$, where $l_s^B(q_s) := E[H_s^B(q_s)|f_s]$ (recall the definition in Section 3 that $H_t^B(Q_t) := c_t q_t + \sum_{j=t+L}^{T} h_j(Q_t - (D_{[t,j]} - X_t)^+)^+$).

(ii) The expected backlogging cost incurred in period $s+L$ as a function of the additional q_s units ordered in period s, conditioning again on f_s. We denote this function by $b_s^B(q_s)$, where $b_s^B(q_s) := E[B_s^B(q_s)|f_s]$ (recall the definition in Section 3 that $B_t^B := p_t(D_{[t,t+L]} - (X_t^B + Q_t))^+ = p_t(D_{[t,t+L]} - Y_t^B)^+$). We note that conditioned on some

$f_s \in \mathcal{F}_s$ and given any policy P, we already know x_s, the starting inventory position in time period s. Hence, the backlogging cost in period s, $B_s^B | f_s$, is indeed only a function of q_s and future demands.

The dual-balancing policy now orders q_s^B units in period s, where q_s^B is such that $l_s^B(q_s^B) = b_s^B(q_s^B)$. In other words, we set q_s^B so that the expected holding cost incurred over the time interval $[s, T]$ by the additional q_s^B units we order at s is equal to the expected backlogging cost in period $s + L$, i.e., $E[H_s^B(q_s^B)|f_s] = E[B_s^B(q_s^B)|f_s]$. Since we assume that the demands are continuous we know that the functions $l_t^P(q_t)$ and $b_t^P(q_t)$ are continuous in q_t for each $t = 1, \dots, T - L$ and each feasible policy P.

Note again that for any given policy P, once we condition on some information set $f_s \in \mathcal{F}_s$, we already know x_s^P deterministically. It is then straightforward to verify that both $l_s^P(q_s)$ and $b_s^P(q_s)$ are convex functions of q_s. Moreover, the function $l_s^P(q_s)$ is equal to 0 for $q_s = 0$ and is an increasing function in q_s, which goes to infinity as q_s goes to infinity. In addition, the function $b_s^P(q_s)$ is non-negative for $q_s = 0$ and is a decreasing function in q_s, which goes to 0 as q_s goes to infinity. Thus, q_s^B is well-defined and we can indeed balance the two functions.

We also point out that q_s^B can be computed as the minimizer of the function $g_s(q_s^B) := \max\{l_s^B(q_s), b_s^B(q_s)\}$. Since $g_s(q_s)$ is the maximum of two convex functions of q_s, it is also a convex function of q_s. This implies that in each period s we need to solve a single-variable convex minimization problem and this can be solved efficiently. In particular, if for each $j \geq s$, $D_{[s,j]}$ has any of the distributions that are commonly used in inventory theory, then it is extremely easy to evaluate the functions $l_s^P(q_s)$ and $b_s^P(q_s)$ (observe that x_s is known at time s). More generally, the complexity of the algorithm is of order T (number of time periods) times the complexity of solving the single variable convex minimization defined above. The complexity of this minimization problem can vary depending on the level of information we assume on the demand distributions and their characteristics. In all of the common scenarios there exist straightforward methods to solve this problem efficiently.

We end this discussion by pointing out that the dual-balancing policy is not a state-dependent base-stock policy. However, it can be implemented on-line, free from the burden of solving large dynamic programming problems. This concludes the description of the algorithm for continuous-demand case. Next we describe the analysis of the worst-case expected performance of this policy.

4.2 Analysis

We start the analysis by expressing the expected cost of the dual-balancing policy.

Lemma 4.1 *Let $\mathcal{C}(B)$ denote the cost incurred by the dual-balancing policy. Then $E[\mathcal{C}(B)] = 2\sum_{t=1}^{T-L} E[Z_t]$, where $Z_t := E[H_t^B|\mathcal{F}_t] = E[B_t^B|\mathcal{F}_t]$ $(t = 1, \ldots, T-L)$.*

Proof : In Section 3, we have already observed that the cost $\mathcal{C}(B)$ of the dual-balancing policy can be expressed as $\sum_{t=1}^{T-L}(H_t^B + B_t^B)$. Using the linearity of expectations and conditional expectations, we can express $E[\mathcal{C}(B)]$ as $\sum_{t=1}^{T-L} E[E[(H_t^B(q_t^B) + B_t^B(q_t^B))|\mathcal{F}_t]]$. However, by the construction of the policy, we know that for each $t = 1, \ldots, T-L$, we have that $E[H_t^B|\mathcal{F}_t] = E[B_t^B|\mathcal{F}_t] = Z_t$. Note that Z_t is a random variable and a function of the realized information set in period t. We then conclude that the expected cost of the solution provided by the dual-balancing policy is $E[\mathcal{C}(B)] = 2\sum_{t=1}^{T-L} E[Z_t]$, where for each t, the expectation $E[Z_t]$ is taken over the possible realizations of information sets in period t, i.e., over the set $\mathcal{F}_t$.
■

Next we wish to show that the expected cost of any feasible policy is at least $\sum_{t=1}^{T-L} E[Z_t]$. In each period $t = 1, \ldots, T-L$ let $\mathcal{Q}_t \subseteq \mathcal{L}_S$ be the set of supply units that were ordered by the dual-balancing policy in period t.

Given an optimal policy OPT and the dual-balancing policy B, we define the following random variables Z'_t for each $t = 1, \ldots, T-L$. In case $Y_t^{OPT} \leq Y_t^B$, we let Z'_t be equal to the backlogging cost incurred by OPT in period $t+L$, denoted by B_t^{OPT}. In case $Y_t^{OPT} > Y_t^B$, we let Z'_t be the ordering and holding cost that the supply units in $\mathcal{Q}_t$ incur in OPT, denoted by $\bar{H}_t^{OPT}$. Note that by our assumption each of the supply units in $\mathcal{Q}_t$ was ordered by OPT in some period t' such that $t' \leq t$. Moreover, for each period s, if we condition on the some information set $f_s \in \mathcal{F}_s$ and given the two policies OPT and B, then we already know y_s^{OPT} and y_s^B deterministically, and hence we know which one of the above cases applies to Z'_s, but we still do not know its value.

We now show that $\sum_{t=1}^{T-L} E[Z'_t]$ is at most the expected cost of OPT, denoted by opt. In other words, it provides a lower bound on the expected cost of an optimal policy. Observe that this lower bound is closely related to the dual-balancing policy through the definition of the variables Z'_t.

Lemma 4.2 *Given an optimal policy OPT, we have $\sum_{t=1}^{T-L} E[Z'_t] \leq E[\mathcal{C}(OPT)] =: opt$.*

Proof : In fact, we will prove a stronger statement, that is $\sum_{t=1}^{T-L} Z'_t \leq \mathcal{C}(OPT)$ with probability 1. Let $\mathcal{T}_B$ be the set of periods $t = 1, \ldots, T-L$ such that $Z'_t = B_t^{OPT}$, and similarly let $\mathcal{T}_H$ be the set of periods $t = 1, \ldots, T-L$ such that $Z'_t = \bar{H}_t^{OPT}$. Clearly, $\mathcal{T}_B$ and $\mathcal{T}_H$ induce a partition of the periods $1, \ldots, T-L$. Now by the definition of Z'_t, we know that $\sum_{t\in\mathcal{T}_B} Z'_t \leq \sum_{t=1}^{T-L} B_t^{OPT}$.

In addition, for each $t \in \mathcal{T}_H$ we know that $Y_t^{OPT} > Y_t^B$, and in particular we know that each of the units in $\mathcal{Q}_t$ was ordered by OPT in some period $t' \leq t$. It is also clear that all of the sets $\{\mathcal{Q}_t : t \in \mathcal{T}_H\}$ are disjoint, since the dual-balancing policy has ordered them in different periods. It now follows that $\sum_{t \in \mathcal{T}_H} Z'_t \leq H^{OPT}$ (where H^{OPT} denotes the overall ordering and holding costs incurred by the by OPT over the planning horizon). This concludes the proof of the lemma. ■

Next we would like to show that for each $s = 1, \ldots, T - L$, we have that $E[Z_s] \leq E[Z'_s]$.

Lemma 4.3 *For each $s = 1, \ldots, T - L$, we have $E[Z_s] \leq E[Z'_s]$.*

Proof : First observe again that if we condition on some information set $f_s \in \mathcal{F}_s$, then we already know whether $Z'_s = B_s^{OPT}$ or $Z'_s = \bar{H}_s^{OPT}$. It is enough to show that for each possible information set $f_s \in \mathcal{F}_s$, conditioning on f_s, we have $z_s \leq E[Z'_s | f_s]$.

In case $Z'_s := B_s^{OPT}$, we know that $y_s^{OPT} \leq y_s^B$. Since for each period t and any given policy P, $B_t^P := p_t(D_{[t,t+L]} - Y_t)^+$, this implies that for each possible realization $d_s, \ldots, d_T$ of the demands over $[s, T]$, we have that the backlogging cost B_s^B that the dual-balancing policy will incur in period $s + L$ is at most the backlogging cost $B_s^{OPT} = Z'_s$ that OPT will incur in that period. That is, with probability 1, we have $B_s^B | f_s \leq Z'_s | f_s$. The claim for this case then follows immediately.

We now consider the case where $Z'_s := \bar{H}_s^{OPT}$. Observe again that each unit in $\mathcal{Q}_s$ was ordered by OPT in some period $t' \leq s$. It is then clear that for each realization of demands $d_s, \ldots, d_T$ over the interval $[s, T]$, the ordering and holding costs $H_s^B | f_s$ that the units in $\mathcal{Q}_s$ will incur in B will be at most the ordering and holding costs $\bar{H}_s^{OPT} | f_s = Z'_s | f_s$ that they will incur in OPT. Here we use the assumption that there is no speculative motive to hold inventory. The lemma then follows. ■

As a corollary of Lemmas 4.1, 4.2 and 4.3, we conclude the following theorem.

Theorem 4.4 *The dual-balancing policy provides a 2-approximation algorithm for the periodic-review stochastic inventory control problem with continuous demands and orders, i.e., for each instance of the problem, the expected cost of the dual-balancing policy is at most twice the expected cost of an optimal solution.*

4.3 Integer-Valued Demands

We now discuss the case in which the demands are integer-valued random variables, and the order in each period is also assumed to be integer. In this case, in each period s, the functions $l_s^B(q_s)$ and $b_s^B(q_s)$ are originally defined only for integer values of q_s. We define these functions for any value of q_s by interpolating piecewise linear extensions of the integer values. It is clear that these extended functions preserve the properties of convexity and monotonicity discussed in the previous (continuous) case. However, it is still possible (and even likely) that the value q_s^B that balances the functions l_s^B and b_s^B is not an integer. Instead we consider the two consecutive integers q_s^1 and $q_s^2 := q_s^1 + 1$ such that $q_s^1 \le q_s^B \le q_s^2$. In particular, $q_s^B := \lambda q_s^1 + (1-\lambda) q_s^2$ for some $0 \le \lambda \le 1$. We now order q_s^1 units with probability λ and q_s^2 units with probability $1-\lambda$. This constructs what we call a *randomized dual-balancing policy*.

Observe that now at the beginning of time period s the order quantity of the dual-balancing policy is still a random variable Q_s^B with support $\{q_s^1, q_s^2\}$. It clear that in each period s we have:

$$E[H_s^B(Q_s^B)|\mathcal{F}_s] = E[B_s^B(Q_s^B)|\mathcal{F}_s] = E[H_s^B(q_s^B)|\mathcal{F}_s] = E[B_s^B(q_s^B)|\mathcal{F}_s] := Z_s.$$

Here, the expectation $E[Z_s|\mathcal{F}_s]$ is taken over Q_s and the future demands $(D_s, \ldots, D_T)$. It is then clear that Lemma 4.1 holds for the randomized dual-balancing policy.

For each $t = 1, \ldots, T-L$, we again define the random variable Z_t'. In case $Y_t^{OPT} \le X_t^B + Q_t^1$, we define Z_t' to be equal to B_t^{OPT}. Observe that now at the beginning of time t, Y_t^B is still a random variable, but with probability 1 it is either $x_t^B + q_t^1$ or $x_t^B + q_t^2$. Otherwise Z_t' is again equal to the ordering and holding costs incurred in OPT by the units in $\mathcal{Q}_t$. Note that $\mathcal{Q}_t$ is now a random set at the beginning of period t because the size of the order Q_t^B is still a random variable at the beginning of time t.

For each realization of demands $d_1, \ldots, d_T$ over the interval $[1, T]$, the output of the randomized dual-balancing policy is now random. In each period $t = 1, \ldots, T-L$, the dual-balancing policy flips a coin with the appropriate probabilities λ and $1-\lambda$, respectively, in order to decide how many units to order. This induces a tree of different possible outcomes that result from the possible realizations of these coin flips. There is a one-to-one correspondence between the leaves of this tree and the possible outcomes, where each root-leaf path corresponds to a particular realization of the $T-L$ coin flips that generated this outcome. The main observations are that for each path, the sets $\{\mathcal{Q}_t \,:\, t \in \mathcal{T}_H\}$ are disjoint. Also note that for each $t \in \mathcal{T}_H$ we still have $y_t^{OPT} \ge y_t^B$, and for each $t \in \mathcal{T}_B$ we have that $y_t^{OPT} \le y_t B$. This implies that Lemma 4.2 still holds (since the sum of the probabilities of all root-leaf paths is exactly 1). Finally, note that for each period s, if we condition on some $f_s \in \mathcal{F}_s$, then we still have $z_s \le E[Z_s'|f_s]$, where again $E[Z_s'|f_s]$ is taken over Q_s^B and future demands $(D_s, \ldots, D_T)$. Hence Lemma 4.3 holds too. We now conclude the following theorem.

Theorem 4.5 *The randomized dual-balancing policy provides a 2-approximation algorithm for the periodic-review stochastic inventory control problem, i.e., for each instance of the problem, the expected cost of the dual-balancing policy is at most twice the expected cost of an optimal solution.*

4.4 Stochastic Lead Times

In this section, we consider the more general model, where the lead time of an order placed in period s is some integer-valued random variable L_s. However, we assume that the random variables $L_1, \dots, L_T$ are correlated, and in particular, that $s + L_s \leq t + L_t$ for each $s \leq t$. In other words, we assume that any order placed at time s will arrive no later than any other order placed after period s. This is a very common assumption in the inventory literature, usually described as ”no order crossing”.

W next describe a version of the dual-balancing that provides a 2-approximation algorithm for this more general model. Let $\mathcal{A}_s$ be the set of all periods $t \geq s$ such that an order placed in s is the latest order to arrive by time period t. More precisely, $\mathcal{A}_s := \{t \geq s : s + L_s \leq t$ and $t' + L_{t'} > t,\ \forall t' \in (s, t]\}$. Clearly, $\mathcal{A}_s$ is a random set of demands. Observe that the sets $\{\mathcal{A}_s\}_{s=1}^{T}$ induce a partition of the planning horizon. Hence, we can write the cost of each feasible policy P in the following way:

$$\mathcal{C}(P) = \sum_{s=1}^{T}(H_s^P + (\sum_{t \in \mathcal{A}_s} B_t^P))$$

Now let $\tilde{B}_s^P := \sum_{t \in \mathcal{A}_s} B_t^P$ and write $\mathcal{C}(P) = \sum_{s=1}^{T}(H_s^P + \tilde{B}_s^P)$. Similar to the previous case, we consider in each period s the two functions $E[H_s^B|f_s]$ and $E[\tilde{B}_s^B|f_s]$, where again f_s is the information set observed in period s. Here the expectation is with respect to future demands as well as future lead times. Finally we order q_s^B units to balance these two functions. By arguments identical to those in Lemmas 4.1, 4.2 and 4.3 we conclude that this policy yields a worst-case performance guarantee of 2.

Observe that in order to implement the dual-balancing policy in this case, we have to know in period s the conditional distributions of the lead times of future orders (as seen from period s conditioned on some $f_s \in \mathcal{F}_s$). This is required in order to evaluate the function $E[\tilde{B}_s^B|f_s]$.

4.5 Enhancing the Performance Guarantee

In this section we use a simple transformation of the cost parameters that will improve the performance guarantee of the dual-balancing policy in cases where the ordering cost is the dominant part of the overall cost. In practice this is often the case. For example, consider the following extreme case, where the demand in each period is deterministic and equal to 10, the per-unit ordering cost is 1, the per-unit holding cost is 1 and the per-unit backlogging penalty is 1. It clear that the optimal policy will order 10 units in each period and will incur only ordering cost. However, the dual-balancing policy, as we have discussed it, will attempt to balance the costs, and thus will do something suboptimal.

We solve this problem using a cost transformation to get an equivalent problem to the original problem. By applying the dual-balancing to the modified problem we are likely to improve the performance guarantee.

We now describe the transformation for the case with no lead time ($L = 0$); the extension to the case of arbitrary lead time is straightforward. Recall that any feasible policy P satisfies the following equations. For each $t = 1, \ldots, T$, we have that $Q_t = NI_t - NI_{t-1} + D_t$ (for ease of notation we omit the superscript P). Using these equations we can express the ordering cost in each period t as $c_t(NI_t - NI_{t-1} + D_t)$ (we again assume without loss of generality that $\alpha = 1$). Using the positive and negative parts of the net inventory at the end of each period, we can replace NI_t with $NI_t^+ - NI_t^-$.

This leads to the following transformation of cost parameters. We let $\hat{c}_t := 0$, $\hat{h}_t := h_t + c_t - c_{t+1}$ $(c_{T+1} = 0)$ and $\hat{p}_t := p_t - c_t + c_{t+1}$. Note assumption discussed in Section 2 on the cost parameters c_t, h_t, and p_t, and in particular, the assumption that there is no speculative motivation to hold inventory of back orders, imply that $\hat{h}_t$ and $\hat{b}_t$ are non-negative ($t = 1, \ldots, T$). Observe that the parameters $\hat{h}_t$ and $\hat{b}_t$ will still be non-negative even if the parameters c_t, h_t, and p_t are negative and as long and the above assumption holds. Moreover, this enables us to incorporate into the model a negative salvage cost at the end of the planning horizon (after the cost transformation we will have non-negative cost parameters). It is readily verified that the induced problem is equivalent to the original one. However, for each realization of the demands, the cost of each policy P in the modified input decreases by exactly $\sum_{t=1}^{T} c_t d_t$ (compared to its cost in the original input). Therefore, any optimal policy for the modified input is also optimal for the original input.

We now apply the dual-balancing policy to the modified problem. We have seen that the assumptions on c_t, h_t and p_t ensure that $\hat{h}_t$ and $\hat{p}_t$ are non-negative and hence the analysis presented above is still valid. Let opt and $\bar{opt}$ be the optimal expected cost of the original and modified inputs, respectively. Clearly, $opt = \bar{opt} + E[\sum_{t=1}^{T} c_t D_t]$. Now the expected cost of the dual balancing policy in the modified input is at most $2\bar{opt}$. Its cost in the original input is then at most $2\bar{opt} + E[\sum_{t=1}^{T} c_t D_t] = 2opt - E[\sum_{t=1}^{T} c_t D_t]$. This implies that if $E[\sum_{t=1}^{T} c_t D_t]$ is a large fraction of opt, then the performance guarantee of the expected cost of the dual-balancing policy might be significantly better than 2. For example, in case $E[\sum_{t=1}^{T} c_t D_t] \geq 0.5opt$ we can conclude that the expected cost of the dual-balancing policy would be at most $1.5opt$. It is indeed the case in many real life problems that a major fraction of the total cost is due the ordering cost.

The intuition of the above transformation is that $\sum_{t=1}^{T} c_t D_t$ is a cost that any feasible policy must pay. As a result, we treat it is an invariant in the cost of any policy and apply the approximation algorithm to the rest of the cost.

In the case where we have a lead time L, we use the equations $Q_t := NI_{t+L} - NI_{t+L-1} + D_{t+L}$ for $t = 1, \ldots, T-L$ to get the same

5 A Class of Myopic Policies

No computationally tractable procedure is known for finding the optimal base-stock inventory levels for the periodic-review inventory control problem with correlated demands. As a result, various simpler heuristics have been considered in the literature. In particular, many researchers have considered a *myopic policy*. In the myopic policy, we follow a base-stock policy $\{R^{my}(f_t) : \ f_t \in \mathcal{F}_t\}$. For each period t and possible information set in period t, the target inventory level $R^{my}(f_t)$ is computed as the minimizer of a one-period problem. More specifically, in period $s = 1, \ldots, T-L$ we focus only on minimizing the expected immediate cost that is going to be incurred in this period (or in $s+L$ in the presence of a lead time L). In other words, the target inventory level $R^{my}(f_s)$ minimizes the expected ordering, holding and backlogging costs in period $s+L$, while ignoring the cost over the rest of the horizon (i.e., the cost over $(s+L, T+1]$). This optimization problem has been proven to be convex and hence easy to solve (see [17]). It is then possible to implement the myopic policy on-line, where in each period s, we compute the base-stock level based on the current observed information set f_s. For each period t and each $f_t \in \mathcal{F}_t$, the myopic base-stock level provides an upper bound on the optimal base-stock level (see [17] for a proof). The intuition is that the myopic policy underestimates the holding cost, since it considers only the one-period holding cost. Therefore, it always orders more units than the optimal policy. Clearly, this policy might not be optimal in general, though in many cases it seems to perform extremely well. Under rather strong conditions it might even be optimal (see [16, 4, 8]). A natural question to ask is whether the myopic policy yields a constant performance guarantee for the periodic-review inventory control problem, i.e., is its expected cost always bounded by some constant times the optimal expected cost.

In this section, we provide a negative answer to this question. We show that the expected cost of the myopic policy can be arbitrarily more expensive than the expected optimal cost, even for the case when the demands are independent and the costs are stationary. The example that we construct provides important intuition concerning the cases for which the myopic policy performs poorly. In addition, we describe an extended class of myopic policies that generalizes the myopic policy discussed above. It is interesting that this class of policies also provides a lower bound on the optimal base-stock levels.

5.1 Myopic Policy - Bad example

Consider the following set of instances parameterized by T, the number of periods. We have a per-unit ordering cost of $c = 0$, a per-unit holding cost $h = 1$ and a unit backlogging penalty $p = 2$. The demands are specified as follows, $D_1 \in \{0, 1\}$ with probability 0.5 for 0 and 1, respectively. For $t = 2, \ldots, T-1,\ D_t := 0$ with probability 1, and $D_T := 1$ with probability 1. The lead time is considered to be equal 0, and $\alpha = 1$.

It is easy to verify that the myopic policy will order 1 unit in period 1 and that this will result an expected cost of $0.5T$. On the other hand, if we do not order in period 1, then the expected cost is 1. This implies that as T becomes larger the expected cost of the myopic policy is $\Omega(T)$ times as expensive as the expected cost of the optimal policy.

The above example indicates that the myopic policy may perform poorly in cases where the demand from period to period can vary a lot, and forecasts can go down. There are indeed many real-life situations, when this is exactly the case, including new markets, volatile markets or end-of-life products.

5.2 A Class of Myopic Policies

As we mentioned before, by considering only the one-period problem, the myopic policy described above underestimates the actual holding cost that each unit ordered in period t is going to incur. This results in base-stock levels that are higher than the optimal base-stock levels.

We now describe an alternative myopic base-stock policy that we call *a minimizing policy*. Recall the functions $l_s^P(q_s),\ b_s^P(q_s)$ defined in Section 4 for each period $s = 1, \ldots, T-L$, where $q_s \geq 0$. Since at each period s we know x_s, we can equivalently write $l_s^P(y_s - x_s),\ b_s^P(y_s - x_s)$, where $y_s \geq x_s$. We now consider in each period s the problem: minimize $(l_s^P(y_s - x_s) + b_s^P(y_s - x_s))$ subject to $y_s \geq x_s$, i.e., minimizing the expected ordering and holding costs incurred by the units ordered in period s over $[s, T]$ and the backlogging cost incurred in period $s + L$, conditioned on some $f_s \in \mathcal{F}_s$. We have already seen that this function is convex in y_s. Observe that $l_s^P(y_s - x_s) - l_s^P(y_s)$ and $b_s^P(y_s - x_s) - b_s^P(y_s)$ do not depend on y_s for $y_s \geq x_s$. This gives rise to the following equivalent one-period problem: $\min_{y_s \geq x_s}(l_s^P(y_s) + b_s^P(y_s))$. That is, both problems have the same minimizer. It is also clear that the new minimization problem is also convex in y_s and is easy to solve, in many cases as easy as the one-period problem solved by the myopic policy described above. We note that the function we minimize was used by Chan and Muckstadt in [1].

For each $t = 1, \ldots, T$ and $f_t \in \mathcal{F}_t$, let $R^M(f_t)$ be the base-stock level resulting from the minimizing policy in period t, for a given observed information set f_t. We now show that for each period t and $f_t \in \mathcal{F}_t$, we have $R^M(f_t) \leq R^{OPT}(f_t)$, where $R^{OPT}(f_t)$ is the optimal base-stock level.

Theorem 5.1 *For each period t and $f_t \in \mathcal{F}_t$, we have $R^M(f_t) \leq R^{OPT}(f_t)$.*

Proof : Recall the dynamic programming based framework described in Section 3. Observe that for each state (x_t, f_t), we know that $R^{OPT}(f_t)$ is the optimal base-stock level that results from the optimal solution for the corresponding subproblem defined over the interval $[t, T]$. It is enough to show that the optimal solution for each such problem must be at least $R^M(f_t)$.

Assume otherwise, i.e., $R^{OPT}(f_s) < R^M(f_s)$ for some period s and for all optimal policies. Consider now the base-stock policy P with base-stock level $R^P(f_s) = R^M(f_s)$ for period s, and $R^P(f_t) := R^{OPT}(f_t)$ for each $t = s+1, \ldots, T$ and $f_t \in \mathcal{F}_t$. We will show that P, starting from period s with observed information set f_s, has an expected cost that is at most the expected cost of the optimal solution. From Section 3 we know that the expected cost of each policy P can be expressed as
$\sum_{t=s}^{T-L} E[H_t^P + B_t^P]$. Now by the definition of $R^M(f_s)$ we know that

$$E[(H_s^P + B_s^P)|f_s] \leq E[(H_s^{OPT} + B_s^{OPT})|f_s].$$

Moreover, for each $t \in (s, T]$, the inventory position Y_t^P will always be at least Y_t^{OPT}, and therefore $E[B_t^P|f_s] \leq E[B_t^{OPT}|f_s]$. It is also clear that in each period $t \in (s, T]$, the Q_t^P units ordered by policy P in period t will always be a subset of the units ordered by OPT in this period. Therefore, for each $t = s+1, \ldots, T$, we have that $E[H_t^P|f_s] \leq E[H_t^{OPT}|f_s]$. This concludes the proof. ■

We now define a generalization that captures the myopic policy and the minimization policy as two special cases. For each $t = 1, \ldots, T-L$, we define a sequence of one-period problems for each $k = 0, \ldots, T-t$, each generates a corresponding base-stock policy. Given k, we define the one-period problem that aims to minimize the expected ordering and holding cost incurred by the units ordered in period t over the interval $[t, t+L+k]$, and the expected backlogging cost in period $t+L$. In other words, the parameter k defines the length of the horizon considered in the one-period problem being solved. It is clear that for $k = 0$ we get the myopic policy and for $k = T-t$ we get the minimizing policy. Note again that the myopic and the minimizing policies provide an upper bound and lower bound, respectively, on the optimal base-stock levels.

6 The Stochastic Lot-Sizing Problem

In this section, we change the previous model and in addition to the per-unit ordering cost, consider a fixed ordering cost K that is incurred in each period t with positive order (i.e., when $Q_t > 0$). For ease of notation, we will assume without loss of generality that $c_t = 0$ and that $\alpha = 1$. We call this model the *stochastic lot-sizing problem*. The goal is again to find a policy that minimizes the expected discounted overall ordering, holding and backlogging costs. Naturally, this model is more complicated. Here we will assume that $L = 0$ and that in each period $t = 1, \ldots, T$, the conditional joint distribution I_t of $(D_t, \ldots, D_T)$ is such that the demand D_t is known deterministically (i.e., with probability 1). The underlying assumption here is that at the beginning of period t our forecast for the demand in that period is sufficiently accurate, so that we can assume it is given deterministically. A primary example is make-to-order systems.

As noted in Section 1, if the demands are independent it is known that the optimal solution can be described as a set $\{(s_t, S_t)\}_t$. In each period t place an order if and only if the current inventory level is below s_t. If we place an order in period t, we will increase the inventory level up to S_t. We next describe a policy which we call the *triple-balancing policy* and denote by TB, and analyze its worst-case expected performance. More specifically, we show that its expected cost is at most 3 times the expected cost of the optimal solution. We note that in this case the policy and its analysis are identical for discrete and continuous demands.

6.1 The Triple-Balancing Policy

The policy follows two rules that specify when to place an order and how many units to order once an order is placed:

Rule 1: In each period s, we let s^* be the period in which the triple-balancing policy has last placed an order, i.e., s^* is the latest order placed so far. Thus, $s^* < s$, where $s^* = 0$ if no order has been placed yet. We place an order in period s if and only if, by not placing it in period s, the accumulated backlogging cost over the interval $(s^*, s]$ will exceed K. If we place an order, we update s^* and set it equal to s. Observe that since, in each period s, the conditional joint distribution I_s is such that D_s is known deterministically, this procedure is well-defined.

Rule 2: If we place an order in period $s < T$, then we focus on the holding cost incurred by the units ordered in s over the interval $[s, T]$, again using marginal cost accounting. We then order q_s^B units such that $q_s^B := \max\{q_s : E[H_s^B(q_s)|f_s] \leq K\}$, where again $f_s \in \mathcal{F}_s$ is the current information set. That is, we order the maximum number of units as long as the conditional expectation of the holding cost that these units will incur over $[s, T]$, as seen from time period s, is at most K. In case $s = T$, we just order enough to cover all current back orders and the demand d_T. Observe that q_s^B must always be large enough to cover all of the backlogged units of demand over $(s^*, s]$. Hence, at the end of a period s in which an order was placed, there are no unsatisfied units of demand. We note that since for each $f_s \in \mathcal{F}_s$, the function $E[H_s^B(q_s)|f_s]$ is convex in q_s, it is relatively easy to compute q_s^B.

This concludes the description of the algorithm. Next we describe the analysis of the worst-case expected performance.

Analysis. Let N be the random variable of the number of orders placed by the triple-balancing policy. We next define a sequence of random variables $S_0, \dots, S_{T+1}$. We let $S_0 = 0,\ S_{T+1} = T + 1$, and let S_i (for $i = 1, \dots, T$) be the time period in which the i^{th} order of the triple-balancing policy was placed, or $T + 1$ if $N < i$ (i.e., the triple-balancing policy has placed fewer than i orders). Observe that $S_1, \dots, S_T$ are random variables, which induce a partition of the time horizon. Consequently, we let Z_i, for each $i = 0, \dots, T$, be the following random variable. If $S_i < T$, then Z_i is equal to the holding cost that the triple balancing policy incurs over $[S_i, S_{i+1})$ (denoted by H_i) plus the backlogging and ordering costs it incurs over $(S_i, S_{i+1}]$. If $S_i \geq T$, then $Z_i = 0$. Similarly, we define the set of variables $Z'_0, \dots, Z'_{T+1}$ with respect to the cost of OPT over the corresponding intervals induced by the orders of the triple-balancing policy. It is clear that $\mathcal{C}(B) = \sum_{i=0}^{T} Z_i \cdot 1(S_i < T)$ and $C(OPT) = \sum_{i=0}^{T} Z'_i \cdot 1(S_i < T)$. We first develop a lower bound on the expected cost of OPT using the expectation of the random variable N.

Lemma 6.1 *For each instance of the stochastic lot-sizing problem with correlated demand the expected cost of an optimal policy OPT is at least $KE[N]$.*

Proof : We have already observed that $\mathcal{C}(OPT) = \sum_{i=0}^{T} Z'_i \cdot 1(S_i < T)$. Using again the linearity of expectation and conditional expectation, we can write,

$$E[\mathcal{C}(OPT)] = \sum_{i=0}^{T} E[1(S_i < T)E[Z'_i|S_i, \mathcal{F}_{S_i}]] \geq$$

$$\sum_{i=0}^{T} E[1(S_i < T)E[Z'_i \cdot 1(S_{i+1} \leq T)|S_i, \mathcal{F}_{S_i}]]$$

Next we show that for each $i = 0, \dots, T$, we have that,

$$E[Z'_i \cdot 1(S_{i+1} \leq T)|S_i, \mathcal{F}_{S_i}] \geq K \cdot Pr(S_{i+1} \leq T|S_i, \mathcal{F}_{S_i})$$

Conditioned on some $S_i = s_i$ and $f_{s_i} \in \mathcal{F}_{s_i}$, we know d_{s_i} (where s_i is the realization of S_i). As a result, we also know the inventory levels of OPT and the triple-balancing policy at the end of period s_i deterministically. Therefore, exactly one of the following 2 cases must apply:

Case 1: At the end of period s_i, the inventory level of OPT is at most the inventory level of the triple-balancing policy, i.e., $y_{s_i}^{OPT} \leq y_{s_i}^{TB}$. Now either OPT places an order over $(s_i, S_{i+1}]$ and hence incurs a cost of at least K over this interval, or it does not; then, unless s_i is the last order of the triple-balancing policy, it will incur backlogging cost of at least K.

Case 2: At the end of period s_i, the inventory level of OPT is strictly larger than the inventory level of the triple-balancing policy, i.e., $y_{s_i}^{OPT} > y_{s_i}^{TB}$. However, by the construction of the triple-balancing policy, we know that if OPT has more physical inventory, then the expected holding cost it will incur over $[s_i, S_{i+1})$ is at least K.

We conclude that in both cases, $Z_i' \cdot 1(S_{i+1} \leq T)|s_i, f_{s_i})] \geq K \cdot 1(S_{i+1} \leq T|s_i, f_{s_i})$. Taking expectation we have $E[Z_i' \cdot 1(S_{i+1} \leq T)|S_i, \mathcal{F}_{S_i}] \geq K \cdot Pr(S_{i+1} \leq T|S_i, \mathcal{F}_{S_i})$.

This implies that

$$E[\mathcal{C}(OPT)] \geq K \cdot E[\sum_{i=0}^{T} 1(S_i < T) \cdot Pr(S_{i+1} \leq T|S_i, F_{S_i})] =$$

$$K \cdot E[\sum_{i=0}^{T} E[1(S_i < T) \cdot 1(S_{i+1} \leq T)|S_i, F_{S_i}]] = K \cdot E[N]$$

■

To finish the analysis we next show that the expected difference between the cost of the triple-balancing policy (denoted by TB) and the cost of the optimal policy is at most $2KE[N]$.

Lemma 6.2 *For each instance of the problem, we have $E[\mathcal{C}(TB) - \mathcal{C}(OPT)] \leq 2KE[N]$.*

Proof : Clearly,

$$E[\mathcal{C}(TB) - \mathcal{C}(OPT)] = E[\sum_{i=0}^{T}(Z_i - Z_i') \cdot 1(S_i < T)] = \sum_{i=0}^{T} E[1(S_i < T) \cdot E[(Z_i - Z_i')|S_i, \mathcal{F}_{S_i}]]$$

We next bound $E[(Z_i - Z_i')|S_i, \mathcal{F}_{S_i}]$ for each $i = 0, \ldots, T$. For $i = 0$, it is clear that the holding costs that the TB policy and OPT incur over $[s_0, S_1)$ are identical (this cost is due initial inventory that exists at the beginning of the horizon). Also observe that the backlogging and ordering costs of the TB policy over $(S_0, S_1]$ are at most K if $S_1 = T+1$ and at most $2K$ otherwise. In the latter case, we conclude that OPT either placed an order on the interval $(S_0, S_1]$ or incurred backlogging cost of at least K. Hence, $Z_0 - Z'_0 \leq K$.

For each $i = 1, \ldots, T$, we condition on some s_i and $f_{s_i} \in \mathcal{F}_{s_i}$. We then know what are $y^{TB}_{s_i}$ and $y^{OPT}_{s_i}$ deterministically. We now claim that:

$$(Z_i - Z'_i)|s_i, f_{s_i} \leq 1(y^{TB}_{s_i} \leq y^{OPT}_{s_i}) \cdot (K + 1(S_{i+1} \leq T|s_i, f_{s_i}) \cdot K) +$$
$$1(y^{TB}_{s_i} > y^{OPT}_{s_i}) \cdot (H_i|s_i, f_{s_i} + 1(S_{i+1} \leq T|s_i, f_{s_i}) \cdot K)$$

In first case where $y^{TB}_{s_i} \leq y^{OPT}_{s_i}$ we know that OPT will incur over $[s_i, S_{i+1})$ at least as much holding cost as the TB policy. By the construction of the algorithm we know that the TB policy will not incur more than K backlogging cost and will place at most one order over $(s_i, S_{i+1}]$. In the second case where $y^{TB}_{s_i} > y^{OPT}_{s_i}$ we know that the ordering cost and backlogging costs of OPT over $(s_i, S_{i+1}]$ are at least K, which is more than the backlogging cost the TB policy incurs on that interval. In addition, TB will incur holding cost $H_i|s_i, f_{s_i}$ over $[s_i, S_{i+1})$ and will place at most one order over $(s_i, S_{i+1}]$. Taking expectation of both sides we conclude that:

$$E[(Z_i - Z'_i)|S_i, \mathcal{F}_{S_i}] \leq E[1(y^{TB}_{s_i} \leq y^{OPT}_{S_i}) \cdot (K + 1(S_{i+1} \leq T) \cdot K)|S_i, \mathcal{F}_{s_i}] +$$
$$E[1(y^{TB}_{S_i} > y^{OPT}_{S_i}) \cdot (H_i + 1(S_{i+1} \leq T) \cdot K)|S_i, \mathcal{F}_{S_i}] \leq E[K + 1(S_{i+1} \leq T)|S_i, \mathcal{F}_{S_i}]$$

The last inequality is by the construction of the algorithm ($E[H_i|s_i, f_{s_i}] \leq K]$ for each $S_i = s_i$ and $f_{s_i} \in \mathcal{F}_{s_i}$).

This implies that for each $i = 2, \ldots, T$, we have:

$$E[(Z_i - Z'_i) \cdot 1(S_i < T)] = E[1(S_i < T) \cdot E[Z_i - Z'_i|S_i, \mathcal{F}_{S_i}]] \leq E[K + 1(S_{i+1} \leq T)]$$

Finally, we have that:

$$E[\sum_{i=0}^{T}(Z_i - Z'_i) \cdot 1(S_i < T)] \leq K + K \cdot E[\sum_{i=1}^{T} 1(S_i < T)] + K \cdot E[\sum_{i=1}^{T+1} 1(S_i < T) \cdot 1(S_{i+1} \leq T)] =$$
$$K + K \cdot E[N] + K \cdot (E[N] - 1) = 2KE[N]$$

■

As a corollary of Lemmas 6.1 and 6.2, we get the following theorem.

Theorem 6.3 *For each instance of the stochastic lot-sizing problem, the expected cost of the triple-balancing policy is at most 3 times the expected cost of an optimal policy.*

7 Conclusions

In this paper we have proposed a new approach for devising provably good policies for stochastic inventory control models with time dependent and correlated demand. These models are known to be hard, in the sense that computing optimal policies is usually intractable. In turn our approach leads to policies that are simple computationally and conceptually and provides constant performance guarantees on the worst-case expected behavior of these policies.

We note that all of the results described in the paper can be extended under rather mild conditions to the counterpart models with infinite horizon, where the goal is to minimize the expected average or discounted cost.

In a subsequent paper, we consider the periodic-review stochastic inventory control problem with correlated demands and with hard capacity restrictions on the amount of units ordered in each period. We use and extend some of the ideas introduced in this paper to construct policies that provide worst-case performance guarantees. We think it would be an interesting challenge to extend the ideas introduced in this paper to more complicated inventory models, such as multi-echelon and/or multi-item models. These issues will be addressed in future work.

It would also be important to establish a more rigorous analysis of the computational hardness of these models. As far as we know there does not exist any rigorous proof of that kind.

Acknowledgments The first author thanks Shane Henderson for stimulating discussions and helpful suggestions that made the paper significantly better.

The authors thank Jack Muckstadt for his valuable comments that lead to the extension to the stochastic lead times.

References

[1] E. Chan. *Markov Chain Models for multi-echelon supply chains*. PhD thesis, School of OR&IE, Cornell University, Ithaca, NY, January 1999.

[2] S. Dye, L. Stougie, and A. Tomasgard. The stochastic single resource service provision problem. *Naval Research Logistics*, 50:869–887, 2003.

[3] N. Erkip, W. Hausman, and S. Nahmias. Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands. *Mangement Science*, 36:381–392, 1990.

[4] T. Iida and P. Zipkin. Approximate solutions of a dynamic forecast-inventory model. Working paper, 2001.

[5] D. Joneja. The joint replenishment problem: new heuristics and worst case performance bounds. *Operations Research*, 38:723–771, 1990.

[6] H. Lee, K. So, and C. Tang. The value of information sharing in two-level supply chains. *Mangement Science*, 46:626–643, 1999.

[7] D. Lingxiu and H. Lee. Optimal policies and approximations for a serial multiechelon inventory system with time-correlated demand. *Operations Research*, 51, 2003.

[8] X. Lu, J. Song, and A. Regan. Inventory planning with forecast updates: approximate solutions and cost error bounds. Working paper, 2003.

[9] R. Mohring, F. Radermacher, and G. Weiss. Stochastic scheduling problems I: general strategies. *ZOR-Zeitschrift fur Operations Research*, 28:193–260, 1984.

[10] R. Mohring, F. Radermacher, and G. Weiss. Stochastic scheduling problems II: set strategies. *ZOR-Zeitschrift fur Operations Research*, 29:65–104, 1984.

[11] R. Mohring, A. Schulz, and M. Uetz. Approximation in stochastic scheduling: the power of LP-based priority policies. *Journal of the ACM (JACM)*, 46:924–942, 1999.

[12] A. Muharremoglu and J. Tsitsiklis. A single-unit decomposition approach to multi-echelon inventory systems. Working paper, 2001.

[13] D. Shmoys and C. Swamy. Stochastic optimization is (almost) as easy as deterministic optimization. To appear in FOCS 2004, 2004.

[14] E. Silver and H. Meal. A heuristic selecting lot-size requirements for the case of a deterministic time varying demand rate and discrete opportunities for replenishment. *Production and Inventory Management*, 14:64–74, 1973.

[15] L. Stougie and M. der Vlerk. Approximation in stochastic integer programming. Technical Report SOM Research Report 03A14, Eindhoven University of Technology, 2003.

[16] A. Veinott. Optimal policy for a multi-product, dynamic, non-stationary inventory problem. *Operations Research*, 12:206–222, 1965.

[17] P. Zipkin. *Foundation of inventory management*. The McGraw-Hill Companies, Inc, 2000.