



## Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:  
<http://pubsonline.informs.org>

### Approximation Algorithms for Stochastic Inventory Control Models

Retsef Levi, Martin Pál, Robin O. Roundy, David B. Shmoys,

To cite this article:

Retsef Levi, Martin Pál, Robin O. Roundy, David B. Shmoys, (2007) Approximation Algorithms for Stochastic Inventory Control Models. Mathematics of Operations Research 32(2):284-302. <https://doi.org/10.1287/moor.1060.0205>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact [permissions@informs.org](mailto:permissions@informs.org).

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2007, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

# Approximation Algorithms for Stochastic Inventory Control Models

Retsef Levi

Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139, [retsef@mit.edu](mailto:retsef@mit.edu)

Martin Pál

Google, Inc., 1440 Broadway, 21st Floor, New York, New York 10018, [martin@palenica.com](mailto:martin@palenica.com)

Robin O. Roundy

School of Operations Research and Information Engineering, Cornell University,  
Ithaca, New York 14853, [robin@orie.cornell.edu](mailto:robin@orie.cornell.edu)

David B. Shmoys

School of Operations Research and Information Engineering and Department of Computer Science, Cornell University,  
Ithaca, New York 14853, [shmoys@cs.cornell.edu](mailto:shmoys@cs.cornell.edu)

We consider two classical stochastic inventory control models, the *periodic-review stochastic inventory control problem* and the *stochastic lot-sizing problem*. 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 important problem of finding computationally efficient and provably good inventory control policies for these models in the presence of correlated, nonstationary (time-dependent), and evolving stochastic demands. This problem arises in many domains and has many practical applications in supply chain management.

Our approach is based on a new marginal cost accounting scheme for stochastic inventory control models combined with novel cost-balancing techniques. Specifically, in each period, we balance the expected cost of overordering (i.e., costs incurred by excess inventory) against the expected cost of underordering (i.e., costs incurred by not satisfying demand on time). This leads to what we believe to be the first computationally efficient policies with constant worst-case performance guarantees for a general class of important stochastic inventory control models. 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. In particular, we provide a worst-case guarantee of two for the periodic-review stochastic inventory control problem and a worst-case guarantee of three for the stochastic lot-sizing problem. Our results are valid for all of the currently known approaches in the literature to model correlation and nonstationarity of demands over time.

**Key words:** inventory; approximation; dual-balancing; algorithms

**MSC2000 subject classification:** Primary: 90B05; secondary: 68W25

**OR/MS subject classification:** Primary: inventory/production, approximation/heuristics; secondary: production/scheduling, approximation/heuristics

**History:** Received October 6, 2004; revised September 12, 2005, and February 1, 2006.

**1. Introduction.** In this paper, we address the fundamental problem of finding computationally efficient and provably good inventory control policies in supply chains with correlated, nonstationary (time-dependent), and evolving stochastic demands. This problem arises in many domains and has many practical applications (see, for example, Erkip et al. [10] and Lee et al. [17]). We consider two classical models: the *periodic-review stochastic inventory control problem* and the *stochastic lot-sizing problem* with correlated, nonstationary, and evolving demands. Here, the correlation is intertemporal, 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 nonstationarity 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 (see, for example, Iida and Zipkin [15] and Dong and Lee [8]). 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 nonstationarity in the demands over time. Correlation and nonstationarity also arise for products with strong cyclic demand patterns and for products as they 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 nonstationarity 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 nonstationary and correlated. In the periodic-review stochastic inventory control problem, the cost consists of a per-unit, time-dependent ordering cost, a per-unit holding cost for carrying excess inventory from period to period, and a per-unit 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 nonstationarity of demands over time. This includes classical approaches like the *martingale model of forecast evolution model* (MMFE), exogenous Markovian demand, time series, order-one autoregressive demand and random walks. For an overview of the different approaches and models and for relevant references, we refer the reader to Iida and Zipkin [15] and Dong and Lee [8]. Moreover, we believe that the models we consider are general enough to capture almost any other reasonable way of modelling correlation and nonstationarity 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 *not* a function 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. The optimality of base-stock policies has been proven in many settings, including models with Markov-modulated demands and correlated demand and forecast evolution (see, for example, Iida and Zipkin [15], Özer and Gallego [25], and Zipkin [31]).

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. Following an optimal policy, an order is placed 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*. The optimality of state-dependent  $(s, S)$  policies has been proven for the case of nonstationary but independent demand (Zipkin [31]). Gallego and Özer [11] have established their optimality for models with correlated demands. We refer the reader to Dong and Lee [8], Iida and Zipkin [15], Zipkin [31], and Gallego and Özer [11] for the details on some of the results along these lines as well as a comprehensive discussion of relevant literature.

Unfortunately, the rather simple forms of these policies do not always lead to efficient algorithms for computing the optimal policies. The corresponding dynamic programs are relatively straightforward to solve if the demands in different periods are independent. The standard dynamic programming approach can still be tractable for uncapacitated models with exogenous Markov-modulated demand, but under rather strong assumptions on the structure and the size of the state space of the underlying Markov process (see, for example, Song and Zipkin [28] and Chen and Song [6]). However, in many scenarios with more complex demand structure, the state space of the corresponding dynamic programs grows exponentially in the size of the time horizon  $T$  and explodes very fast. Thus, solving the corresponding dynamic programs becomes practically (and often also theoretically) intractable (see Iida and Zipkin [15] and Dong and Lee [8] for relevant discussions on the MMFE model). This is especially true in the presence complex demand structures where demands in different periods are correlated and evolve over time. The difficulty essentially comes from the fact that we need to solve “too many” subproblems. This phenomenon is known as *the curse of dimensionality*. Moreover, because of this

phenomenon, it seems very challenging to find efficient algorithms that solve these huge dynamic programs to optimality. 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 [24] 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. In particular, they have established the optimality of state-dependent base-stock policies for the uncapacitated model with general Markov-modulated demand. 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 nonstationary 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, Lu et al. [20]). 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 Dong and Lee [8], Lu et al. [20], and Iida and Zipkin [15]. 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 potential effect on the cost in future periods. The myopic policy is attractive since it yields a base-stock policy that is easy to compute online, that is, it does not require information on the control policy in the future periods. In each period, we need to solve a single-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. Veinott [30] and Ignall and Veinott [14] have shown that the myopic policy can be optimal even in models with nonstationary demand as long as the demands are stochastically increasing over time. Iida and Zipkin [15] and Lu et al. [20] have focused on the martingale model of forecast evolution (MMFE) 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 [4] and Chan and Muckstadt [5] have considered a different way for approximating huge dynamic programs that arise in the context of inventory control problems. More specifically, they have considered uncapacitated 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 online combinatorial problems. The work on approximation results for stochastic combinatorial problems goes back to the work of Möhring et al. [21, 22] and the more recent work of Möhring et al. [23]. 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 two-stage stochastic combinatorial problems. For a comprehensive literature review, we refer the reader to Stougie and van der Vlerk [29], Dye et al. [9], and Shmoys and Swamy [26]. We note that the problems we consider in this work are by nature *multistage* stochastic problems, which are usually much harder (see Dean et al. [7] for a recent result on the stochastic knapsack problem).

Another approach that has been applied to these models is the robust optimization approach (see Bertsimas and Thiele [3]). Here, the assumption is of a distribution-free model where instead the demands are assumed to be drawn from some specified uncertainty set. Each policy is then evaluated with respect to the worst possible sequence of demands within the given uncertainty set. The goal is to find the policy with the best worst case (i.e., a min-max approach). This objective is very different from the objective of minimizing expected (average) cost discussed in most of the existing literature, including this work.

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

*Marginal Cost Accounting Scheme.* We introduce a novel approach for cost accounting in uncapacitated stochastic inventory control problems. The standard dynamic programming approach directly assigns to the decision of how many units to order in each period only the expected holding and backlogging costs incurred in that period, although this decision might affect the costs in future periods. Instead, our new cost accounting scheme assigns to the decision in each period *all* the expected costs that, once this decision is made, become independent of any decision made in future periods and are dependent only on the future demands. Specifically, we introduce a *marginal holding cost accounting approach*. This approach is based on the key observation that once we place an order for 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 over the entire horizon by the units ordered in this period. We note that similar ideas of holding cost accounting were previously used in the context of models with continuous time, infinite horizon, and stationary (Poisson distributed) demand (see, for example, the work of Axsäter and Lundell [2] and Axsäter [1]). In addition, in an uncapacitated model the decision of how many units to order in each period affects the expected backlogging cost in only a single future period, namely, a lead time ahead. Thus, our cost accounting approach is marginal, i.e., it associates with each period the overall expected costs that become independent of any future decision. 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 best known examples are the Silver-Meal part-period balancing heuristic for the lot-sizing problem (see Silver and Meal [27]) and the cost-covering heuristic of Joneja for the joint-replenishment problem (Joneja [16]). We are not aware of any application of these ideas to stochastic inventory control problems. The key observation is that any policy in any period incurs potential expected costs due to overordering (namely, expected holding costs of carrying excess inventory) and underordering (namely, expected backlogging costs incurred when demand is not met on time). 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) holding cost against the expected (marginal) backlogging cost. 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. As we shall show, the simple idea of balancing is powerful and leads to policies that have constant expected worst-case performance guarantees. As already stated we believe that the balancing idea will have more applications in constructing and analyzing algorithms for other stochastic inventory control models (see Levi et al. [19] and Levi et al. [18] for follow-up work).

*Nonbase-Stock Policies.* Our policies are not state-dependent base-stock policies in that the order up-to level order of the policy in each period *does* depend on the inventory control in past periods, i.e., it depends on the inventory position at the beginning of the period. However, this enables us to use, in each period, the distributional information about 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 online (like the myopic policy) and are simple both conceptually and computationally (see Hurley et al. [13]).

Using these ideas, we provide what is called a 2-approximation algorithm for the uncapacitated periodic-review stochastic inventory control problem, that is, the expected cost of our policies 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 expected cost of an optimal policy, which is a much more stringent requirement. We also note that these guarantees refer only to the worst-case performance and it is likely that the typical performance would be significantly better (see Hurley et al. [13]). 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 is the case in many real-life situations. Our results are valid for all known approaches used to model correlated and nonstationary demands. We note that the analysis of the worst-case performance is tight. In particular, we describe a family of examples for which the ratio between the expected cost of the balancing policy and the expected cost of the optimal policy is asymptotically two. We also present an extended class of myopic policies that provides easily computed upper bounds and lower bounds on the optimal base-stock levels. As shown in Hurley et al. [13], these bounds combined with the balancing techniques lead to *improved balancing policies*. These policies have a worst-case performance guarantee of two and they seem to perform significantly better in practice.

An interesting question that is not addressed 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 §2 we present a mathematical formulation of the periodic-review stochastic inventory control problem. Then, in §3 we explain the details of our new marginal cost accounting approach. In §4 we describe a 2-approximation algorithm for the periodic-review stochastic inventory control problem. In §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 §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, 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  $t = 1, \dots, T$  (note that  $t$  and  $T$  are both deterministic unlike the convention above). The demands over these periods are random variables, denoted by  $D_1, \dots, D_T$ .

As part of the model, we 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  $(d_1, \dots, d_{s-1})$  over the interval  $[1, s)$ , and possibly some more (external) information denoted by  $(w_1, \dots, w_s)$ . The information set  $f_s$  in period  $s$  is one specific realization in the set of all possible realizations of the random vector  $F_s = (D_1, \dots, D_{s-1}, W_1, \dots, W_s)$ . 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  $(D_s, \dots, D_T)$ , 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  $t$  conditioning on some information set  $f_s \in \mathcal{F}_s$  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 current period.

The only assumption on the demands is that, for each  $s = 1, \dots, T$ , 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 nonstationarity and correlation between the demands in different periods. We note again that by allowing correlation we let  $I_s$  be dependent on the realization of the demands over the periods  $1, \dots, s-1$  and possibly on some other information that becomes available by time  $s$  (i.e.,  $I_s$  is a function of  $f_s$ ). However, the information set  $f_s$  as well as the conditional joint distribution  $I_s$  are 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. In period  $t$ , ( $t = 1, \dots, T$ ) three types of costs are incurred: a per-unit ordering cost  $c_t$  for ordering any number of units at the beginning of period  $t$ , a per-unit holding cost  $h_t$  for holding excess inventory from period  $t$  to  $t+1$ , and a per-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 *backorders*. The assumption is that backorders 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 §4.5, 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  $\alpha^t$ . 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 backorders in the system. To enforce this, we assume that for each  $t = 2, \dots, T-L$ , the inequalities  $c_t \leq c_{t-1} + h_{t+L-1}$  and  $c_t \leq c_{t+1} + p_{t+L}$  are maintained (where  $c_{T+1} = 0$ ). (In case the discount factor is smaller than one, we require that  $\alpha c_t \leq c_{t-1} + \alpha^L h_{t+L-1}$  and  $c_t \leq \alpha c_{t+1} + \alpha^L p_{t+L}$ .) We also assume that the parameters  $h_t$ ,  $p_t$ , and  $c_t$  are all nonnegative. In §4.6, we will show how to incorporate end-of-horizon salvage cost into the model.

The goal is to find an ordering policy that minimizes the overall expected discounted ordering cost, holding cost, and backlogging cost. We consider only policies that are *nonanticipatory*, i.e., at time  $s$ , the information that a feasible policy can use consists only of  $f_s$  and the current inventory level. In particular, given any feasible policy  $P$  and conditioning on a specific information set  $f_s$ , we know the inventory level  $x_s^P$  deterministically.

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.

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 backorders). 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  $t = 1, \dots, T$ ), where  $Q_j$  denotes the number of units ordered in period  $j$  (we will sometimes 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  $s$ :

- (i) The order placed in period  $s - L$  of  $q_{s-L}$  units arrives and the net inventory level increases accordingly to  $ni_{s-1} + q_{s-L}$ .
- (ii) The decision of how many units to order in period  $s$  is made. Following a given policy  $P$ ,  $q_s$  units are ordered ( $0 \leq q_s$ ). Consequently, the inventory position is raised by  $q_s$  units (from  $x_s$  to  $y_s$ ). This incurs a linear cost  $c_s q_s$ .
- (iii) We observe the 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  $x_{s+1} = x_s + q_s - d_s$  and  $ni_{s+1} = ni_s + q_{s-L} - d_s$ .
- (iv) If  $ni_{s+1} > 0$ , then we incur a holding cost  $h_s ni_{s+1}$  (this means that there is excess inventory that needs to be carried to time period  $s + 1$ ). On the other hand, if  $ni_{s+1} < 0$ , we incur a backlogging penalty  $p_t |ni_{s+1}|$  (this means that there are currently unsatisfied units of demand).

**3. Marginal cost accounting.** In this section, we present a new approach to the holding cost accounting in stochastic inventory control problems, which leads to what we call a *marginal cost accounting scheme*. 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 §2 are formulated using a dynamic programming framework. For simplicity, we discuss the case with  $L = 0$ , where  $x_s = ni_s$  (for a detailed discussion, see Zipkin [31]).

In a dynamic programming framework, the problem is defined recursively over time through subproblems that are defined for each possible state. A state usually consists of a time period  $t$ , an information set  $f_t \in \mathcal{F}_t$ , and the inventory position at the beginning of period  $t$ , denoted by  $x_t$ . For each subproblem, let  $V_t(x_t, f_t)$  be the optimal expected cost over the interval  $[t, T]$ , given that the inventory position at the beginning of period  $t$  was  $x_t$  and the observed information set was  $f_t$ . We seek to compute an optimal policy in period  $t$  that minimizes the expected cost over  $[t, T]$  (i.e., minimizes  $V_t(x_t, f_t)$ ) under the assumption that we are going to make optimal decisions in future periods. The space of feasible decisions consists of all orders of size  $q_t \geq 0$  or, alternatively, the level  $y_t$  to which the inventory position is raised, where  $x_t \leq y_t$  (and  $q_t = y_t - x_t$ ). Assuming that the optimal policy for all subproblems of states with periods  $t + 1, \dots, T$  has been already computed, the dynamic programming formulation for computing the optimal policy for the subproblem of period  $t$  is

$$V_t(x_t, f_t) = \min_{x_t \leq y_t} \{c_t(y_t - x_t) + E[h_t(y_t - D_t)^+ + p_t(D_t - y_t)^+ | f_t] + E[V_{t+1}(y_t - D_t, F_{t+1}) | f_t]\}.$$

As can be seen, the cost of any feasible decision  $x_t \leq y_t$  is divided into two parts. The first part is the *period cost* associated with period  $t$ , namely, the ordering cost incurred by the order placed in period  $t$  and the resulted expected holding cost and backlogging cost in this period, i.e.,

$$c_t(y_t - x_t) + E[h_t(y_t - D_t)^+ + p_t(D_t - y_t)^+ | f_t].$$

In addition, there are the future costs over  $[t + 1, T]$  (again, assuming that optimal decisions are made in future periods). The impact of the decision in period  $t$  on the future costs is captured through the state in the next period, namely,  $y_t - D_t$ . In particular, in a standard dynamic programming framework, the cost accounted directly in each period  $t$  is only the expected period cost, although the decision made in this period might imply additional costs in the future periods. We note that if  $L > 0$ , then the period cost is always computed a lead time ahead. That is, the period cost associated with the decision to order up to  $y_t$  in period  $t$  is

$$c_t(y_t - x_t) + E[h_{t+L}(y_t - D_{[t, t+L]})^+ + p_{t+L}(D_{[t, t+L]} - y_t)^+ | f_t],$$

where  $D_{[t, t+L]}$  is the accumulated demand over the lead time.

The dynamic programming approach has turned out to be very effective in characterizing the structure of optimal policies. As was noted in §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  $f_s$ , the optimal base-stock level is  $R(f_s)$ . The optimal policy then has the following pattern. In case the inventory position 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.

The above dynamic program can be solved efficiently in case the demands in different periods are independent of each other. This approach might still be tractable in cases where the demand is Markov modulated, as long as the underlying Markov chain has a relatively small number of states or there are some other structural properties that enables us to reduce the number of states being considered in each period. Unfortunately, in many scenarios where the demands in different periods are correlated, obtaining the optimal policy using this dynamic programming formulation is likely to be intractable. To compute the optimal policy, we usually need to consider a subproblem for every possible period and possible state of the system in that period. However, the set  $\mathcal{F}_s$  can be very large or even infinite, which makes solving the corresponding dynamic program practically intractable. The theoretical complexity of this problem is determined by the way the sets  $\mathcal{F}_t$  are specified. For example, in the MMFE model (Iida and Zipkin [15], Lu et al. [20]), the sets  $\{\mathcal{F}_t: t = 1, \dots, T\}$  lie in  $\mathbb{R}^T$  and are of size exponential in the size of the input. This phenomenon is known as the “curse of dimensionality.”

**3.2. Marginal holding cost accounting.** 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 unaffected by any future decision in subsequent time periods. It is affected only by 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 demand 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 than 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 demand 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 further assume (again without loss of generality) that as 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 ordered to a certain unit of demand that has the same number (see Figure 1). We note that Muharremoglu and Tsitsiklis [24] 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 methods are 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.



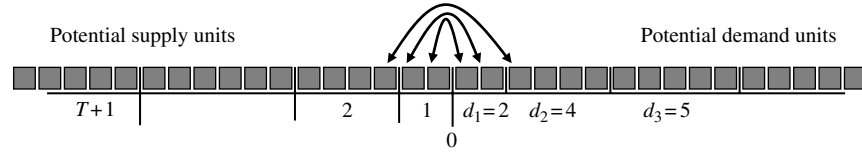


FIGURE 1. The demand and supply lines.

Suppose now that at the beginning of period  $s$  we have observed an 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  $x^+ = \max(x, 0)$ . Recall that at time  $s$  we assume to 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  $H_t^P := H_t^P(Q_t^P)$  (for  $t = 1, \dots, T$ ) be the discounted ordering and expected holding cost incurred by the additional  $Q_t^P$  units ordered in period  $t$  by policy  $P$ . Thus,

$$H_t^P = H_t^P(Q_t^P) := c_t Q_t^P + \sum_{j=t+L}^T h_j(Q_t^P - (D_{[t,j]} - X_t)^+)^+ \quad (1)$$

(assume again  $\alpha = 1$ ). This is very different than the traditional (period-by-period) holding cost accounting approach (see also Figure 2). We note again that similar ideas of holding cost accounting have been used in settings with continuous time and infinite horizon (see, for example, Axsäter and Lundell [2] and Axsäter [1]).

Now let  $\Pi_t^P$  be the discounted expected backlogging cost incurred in period  $t + L$  ( $t = 1 - L, \dots, T - L$ ). That is,

$$\Pi_t^P := p_{t+L}(D_{[t,t+L]} - (X_{t+L} + Q_t^P))^+, \quad (2)$$

where  $D_j := 0$  with probability one 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,

$$\mathcal{C}(P) := \sum_{t=1-L}^0 \Pi_t^P + H_{(-\infty, 0]} + \sum_{t=1}^{T-L} (H_t^P + \Pi_t^P), \quad (3)$$

where  $H_{(-\infty, 0]}$  denotes the total holding cost incurred by units ordered before period 1. We note that the first two expressions  $\sum_{t=1-L}^0 \Pi_t^P$  and  $H_{(-\infty, 0]}$  are not affected by our decisions (i.e., they are the same for any feasible policy and each realization of the demands) and, therefore, we will omit them. Since they are nonnegative, this will not affect our approximation results. Also, observe that without loss of generality, we can assume that  $Q_t^P = H_t^P = 0$

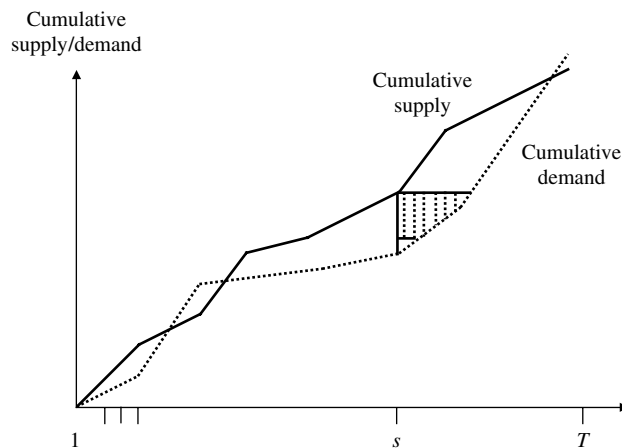


FIGURE 2. The solid line is cumulative supply and the dashed line is cumulative demand. Assume that  $h_t = h$  for each  $t$ . At  $s$ , we measure the period (traditional) holding cost ( $h$  times the vertical solid line) and marginal holding cost  $H_s^P$  ( $h$  times the dotted vertical lines).

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

$$\mathcal{C}(P) = \sum_{t=1}^{T-L} (H_t^P + \Pi_t^P). \quad (4)$$

The cost accounting scheme in Equation (4) is marginal, i.e., in each period, we account for all the expected costs that become independent of any future decision (i.e., costs that become inevitable). In the next section, we shall demonstrate that this new cost accounting 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 uncapacitated periodic-review stochastic inventory control problem, which we call a *dual-balancing policy*. In this model, the goal is to attempt and satisfy future demands just on time. The key intuition is that since the exact future demands are not known (only their distribution), any policy incurs two opposing expected costs. That is, each policy incurs an expected holding cost resulting from overordering and an expected backlogging cost resulting from underordering. The dual-balancing policy balances, in each period, the expected marginal holding cost against the expected marginal backlogging penalty cost. In each period  $s = 1, \dots, T - L$ , we focus on the units that we order in period  $s$  only, and balance the expected 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 that was introduced in §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. We will then show that the worst-case guarantee of two is tight. Specifically, we will show that there exists a set of instances for which the ratio between the expected cost of the dual-balancing policy and the expected cost of the optimal policy converges to two asymptotically.

Recall the assumption discussed in §2 that the cost parameters imply no speculative motivation for holding inventory or backorders. Using this assumption and a standard cost transformation from inventory theory, we can assume, without loss of generality, that  $c_t = 0$  and  $h_t, p_t \geq 0$ , for each  $t = 1, \dots, T$ . Moreover, we first describe the algorithm and its analysis under the latter assumption. Then, in §4.6 we discuss in detail the generality of this assumption. In that subsection, we will also show how a simple cost transformation can yield a better worst-case performance guarantee and certainly a better typical (average) performance in many cases in practice.

In the rest of the paper, we will use superscripts  $B$  and  $OPT$  to refer to the dual-balancing policy described below and an optimal policy, respectively.

**4.1. The algorithm.** We first describe the algorithm and its analysis in the case where fractional orders are allowed. Later, 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  $s = 1, \dots, T - L$ , we consider a given information set  $f_s$  (where again  $f_s \in \mathcal{F}_s$ ) and the resulting pair  $(x_s^B, I_s)$ , where  $x_s^B$  is the inventory position of the dual-balancing policy at the beginning of period  $s$  and  $I_s$  is the conditional joint distribution  $I_s$  of the demands  $(D_s, \dots, D_T)$ . We then consider the following two functions:

(i) The expected holding cost over  $[s, T]$  incurred by the *additional*  $q_s$  units ordered in period  $s$ , conditioned on  $f_s$ . We denote this function by  $l_s^B(q_s)$ , where  $l_s^B(q_s) := E[H_s^B(q_s) | f_s]$ . As we have seen in Equation (1),

$$H_t^B(Q_t) := \sum_{j=t+L}^T h_j(Q_t - (D_{[t,j]} - X_t)^+)$$

(recall that  $c_t = 0$ ).

(ii) The expected backlogging cost incurred in period  $s + L$  as a function of the additional  $q_s$  units ordered in period  $s$ , conditioned again on  $f_s$ . We denote this function by  $\pi_s^B(q_s)$ , where  $\pi_s^B(q_s) := E[\Pi_s^B(q_s) | f_s]$ . In Equation (2), we have defined

$$\Pi_t^B := p_t(D_{[t,t+L]} - (X_t^B + Q_t))^+ = p_t(D_{[t,t+L]} - Y_t^B)^+.$$

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

The dual-balancing policy now orders  $q_s^B = q'_s$  units in period  $s$ , where  $q'_s$  is such that  $l_s^B(q'_s) = \pi_s^B(q'_s)$ . In other words, we set  $q'_s$  so that the expected holding cost incurred over the time interval  $[s, T]$  by the additional  $q'_s$  units we order at  $s$  is equal to the expected backlogging cost in period  $s + L$ , i.e.,  $E[H_s^B(q'_s) | f_s] = E[\Pi_s^B(q'_s) | f_s]$ . Since we assume that fractional orders are allowed, we know that the functions  $l_t^P(q_t)$  and  $\pi_t^P(q_t)$  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 a specific 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  $\pi_s^P(q_s)$  are convex functions of  $q_s$ . Moreover, the function  $l_s^P(q_s)$  is equal to zero 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  $\pi_s^P(q_s)$  is nonnegative for  $q_s = 0$  and is a decreasing function in  $q_s$ , which goes to zero as  $q_s$  goes to infinity. Thus,  $q'_s$  is well-defined and we can indeed balance the two functions.

Observe the difference between the marginal holding cost function  $l_s$  that accounts for costs over an entire time interval and the backlogging cost function  $\pi_s$  that accounts for costs incurred in a single period. The intuitive explanation is that in an uncapacitated model, underordering (i.e., ordering “too little”) can always be fixed in the next period to avoid further costs. On the other hand, since we cannot order a negative number of units, overordering (i.e., ordering “too many” units) can not be fixed by any decision made in future periods, and the resulting costs are only a function of future demands, not of future orders. We also point out that  $q'_s$  can be computed as the minimizer of the function

$$g_s(q_s^B) := \max\{l_s^B(q_s), \pi_s^B(q_s)\}.$$

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 given that the functions  $l_s^B$  and  $\pi_s^B$  can be evaluated efficiently. 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. Note that  $q'_s$  lies at the intersection of two monotone convex functions, which suggests that bisection methods can be effective in computing  $q'_s$ .

In particular, the functions  $l_s$  and  $\pi_s$  consist of a sum of partial expectations. Observe that  $x_s$  is known at time  $s$  and these are expectations of simple piecewise linear functions. If the accumulated demand  $D_{[s,j]}$  (for each  $j \geq s$ ) has any of the distributions that are commonly used in inventory theory (e.g., normal, gamma, lognormal, Laplace, etc.) (Zipkin [31]), then it is extremely easy to evaluate the functions  $l_s^P(q_s)$  and  $\pi_s^P(q_s)$ . If the distribution of  $D_{[s,j]}$  is discrete, these functions can be computed recursively in efficient ways using the cumulative distribution functions. More generally, the complexity of evaluating the functions  $l_s^P(q_s)$  and  $\pi_s^P(q_s)$  and minimizing  $g_s(q_s)$  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 (see also Hurley et al. [13] for a discussion on implementation issues and the performance of the dual-balancing policy when demand follows an MMFE model). Note that in the presence of positive lead times, even computing a simple myopic policy requires the same knowledge on the distribution of the accumulated demand over the lead time. Thus, it seems that the computational effort involved with implementing the dual-balancing policy is of the same order of magnitude as the myopic policy.

Finally, observe that the dual-balancing policy is not a state-dependent base-stock policy. That is, the control of the dual-balancing policy does depend on the inventory control policy in past periods, namely, on  $x_s^B$ . However, it can be implemented online, i.e., it does not require any knowledge of the control policy in future periods. Thus, we avoid the burden of solving large dynamic programming problems. Moreover, unlike the myopic policy, the dual-balancing policy is using in each period available distributional information about the future demands.

This concludes the description of the algorithm for the continuous-demand case. Next, we describe the analysis of the worst-case expected performance of this policy.

**4.2. Analysis.** Next, we shall show that, for each instance of the problem, the expected cost of the dual-balancing policy described above is at most twice the expected cost of an optimal policy. We will use the marginal cost accounting approach described in §3 (see Equation 4), and amortize the period cost of the dual-balancing policy with the cost of the optimal policy. That is, we will show that the optimal policy must incur on expectation at least half of the period expected cost of the dual-balancing policy.

Using the marginal holding cost accounting approach discussed in §3, the expected cost of the dual-balancing policy can be expressed as

$$E[\mathcal{C}(B)] = \sum_{t=1}^{T-L} E[H_t^B + \Pi_t^B]. \quad (5)$$

For each  $t = 1, \dots, T - L$ , let  $Z_t$  be the *random balanced cost* by the dual-balancing policy in period  $t$ , i.e.,  $Z_t = E[H_t^B | F_t] = E[\Pi_t^B | F_t]$ . Note that  $Z_t$  is realized in period  $t$  as a function of the observed information set  $f_t$  (we will denote its realization by  $z_t$ ). By the construction of the dual-balancing policy, we know that, with probability one,  $E[H_t^B | F_t] = E[\Pi_t^B | F_t]$  for each period  $t = 1, \dots, T - L$ . This implies that, for each period  $t$ ,  $E[H_t^B + \Pi_t^B | F_t] = 2Z_t$ , which proves the following lemma.

**LEMMA 4.1.** *The expected cost of the dual-balancing policy is equal to twice the expected sum of the  $Z_t$  variables, i.e.,  $E[\mathcal{C}(B)] = 2 \sum_{t=1}^{T-L} E[Z_t]$ .*

**PROOF.** Using the marginal cost accounting scheme discussed in §3 and a standard argument of conditional expectations, we express

$$E[\mathcal{C}(B)] = \sum_{t=1}^{T-L} E[H_t^B + \Pi_t^B] = \sum_{t=1}^{T-L} E[E[H_t^B + \Pi_t^B | F_t]] = 2 \sum_{t=1}^{T-L} E[Z_t]. \quad \square$$

Next, we will state and prove two lemmas which imply that the expected cost of an optimal policy is at least  $\sum_{t=1}^{T-L} E[Z_t]$ . For each realization of the demands  $D_1, \dots, D_T$ , let  $\mathcal{T}_H$  be the set of periods in which the optimal policy had more inventory than the dual-balancing policy, i.e., the set of periods  $t$  such that  $Y_t^B < Y_t^{OPT}$ . Let  $\mathcal{T}_\Pi$  be the set of periods in which the dual-balancing policy had at least as much inventory as *OPT*, i.e., the set of periods  $t$  such that  $Y_t^B \geq Y_t^{OPT}$ . Observe that  $\mathcal{T}_H$  and  $\mathcal{T}_\Pi$  are random sets that induce a random partition of the planning horizon. The next lemma shows that, with probability one, the marginal holding cost incurred by the dual-balancing policy in periods  $t \in \mathcal{T}_H$  is at most the overall holding cost incurred by *OPT*, denoted by  $H^{OPT}$ , i.e.,  $\sum_{t \in \mathcal{T}_H} H_t^B \leq H^{OPT}$  with probability one.

Recall the concepts of  $\mathcal{L}_D$ , the line of potential units of demand to be realized over the horizon, and  $\mathcal{L}_S$ , the line of supply units to be ordered over the planning horizon, discussed in §3. Since the demand is independent from the inventory policy, we can compare between any two feasible policies by looking at the respective periods, in which each supply unit in  $\mathcal{L}_S$  was ordered. The proof technique in the next lemma will be based on such comparison between the dual-balancing policy and an optimal policy.

**LEMMA 4.2.** *For each realization  $f_T \in \mathcal{F}_T$ , the marginal holding cost incurred by the dual-balancing policy in all periods  $t \in \mathcal{T}_H$  is at most the overall holding cost incurred by *OPT*, denoted by  $H^{OPT}$ , i.e.,  $\sum_{t \in \mathcal{T}_H} H_t^B \leq H^{OPT}$  with probability one.*

**PROOF.** Consider an information set  $f_T \in \mathcal{F}_T$  which corresponds to a complete evolution over the planning horizon and some period  $s \in \mathcal{T}_H$ . We slightly abuse the notation and let  $\mathcal{T}_H$  denote the deterministic set of periods that corresponds to the specific information set  $f_T$ . Let  $\mathcal{Q}_s \subseteq \mathcal{L}_S$  be the set of supply units ordered by the dual-balancing policy in period  $s$ , where clearly  $|\mathcal{Q}_s| = q'_s$ , the balancing order quantity in period  $s$ . By the definition of  $\mathcal{T}_H$ , we know that in period  $s$  we had  $y_s^B < y_s^{OPT}$ . This implies that the units in  $\mathcal{Q}_s$  were ordered by *OPT* either in period  $s$  or even prior to  $s$ . Since we assume that  $c_s = 0$  and that  $h_t \geq 0$  for each period  $t$ , we conclude that the holding cost that these units have incurred in *OPT* is at least as much as the holding cost they have incurred in the dual-balancing policy.

We conclude the proof by observing that the sets  $\{\mathcal{Q}_s : s \in \mathcal{T}_H\}$  are of disjoint supply units since they consist of units ordered by the dual-balancing policy in different periods. This implies that indeed  $\sum_{t \in \mathcal{T}_H} H_t^B \leq H^{OPT}$ , with probability one.  $\square$

The next lemma shows that, with probability one, the marginal backlogging penalty cost of the dual-balancing policy associated with periods  $t \in \mathcal{T}_\Pi$  is at most the overall backlogging penalty incurred by *OPT*, denoted by  $\Pi^{OPT}$ .

**LEMMA 4.3.** *For each realization  $f_T \in \mathcal{F}_T$ , the marginal backlogging penalty cost of the dual-balancing policy associated with all periods  $t \in \mathcal{T}_\Pi$  is at most the overall backlogging penalty incurred by *OPT*, denoted by  $\Pi^{OPT}$ , i.e.,  $\sum_{t \in \mathcal{T}_\Pi} \Pi_t^B \leq \Pi^{OPT}$  with probability one.*

**PROOF.** Consider a realization  $f_T \in \mathcal{F}_T$  and some period  $s \in \mathcal{T}_\Pi$  (where again we abuse the notation and use  $\mathcal{T}_\Pi$  to denote a deterministic set). Note that period  $s$  is associated with the backlogging cost incurred in period  $s + L$ . By definition of  $\mathcal{T}_\Pi$ , we know that  $y_s^B \geq y_s^{OPT}$ . However, this implies that, with probability one, the backlogging cost incurred by the dual-balancing policy in period  $s + L$  is no greater than the respective backlogging cost incurred by the optimal policy in period  $s + L$ . The proof then follows.  $\square$

As a corollary of Lemmas 4.1, 4.2, and 4.3, we get the following theorem:

**THEOREM 4.1.** *The dual-balancing policy for the uncapacitated periodic-review stochastic inventory control problem has a worst-case performance guarantee of two, 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, i.e.,  $E[\mathcal{C}(B)] \leq 2E[\mathcal{C}(OPT)]$ .

PROOF. From Lemma 4.1, we know that the expected cost of the dual-balancing policy is equal to twice the expected cost of the sum of the  $Z_t$  variables, i.e.,

$$E[\mathcal{C}(B)] = 2 \sum_{t=1}^{T-L} E[Z_t].$$

From Lemmas 4.2 and 4.3, we know that, with probability one, the cost of  $OPT$  is at least as much as the holding cost incurred by units ordered by the dual-balancing policy in periods  $t \in \mathcal{T}_H$  plus the backlogging cost of the dual-balancing policy that is associated with periods  $t \in \mathcal{T}_\Pi$ . In other words, with probability one,  $H^{OPT} + \Pi^{OPT} \geq \sum_{t \in \mathcal{T}_H} H_t^B + \sum_{t \in \mathcal{T}_\Pi} \Pi_t^B$ . Using again conditional expectations and the definition of  $Z_t$ , this implies that, indeed,

$$\begin{aligned} E[\mathcal{C}(OPT)] &\geq E\left[\sum_{t \in \mathcal{T}_H} H_t^B + \sum_{t \in \mathcal{T}_\Pi} \Pi_t^B\right] = \sum_t E[H_t^B \cdot \mathbb{1}(t \in \mathcal{T}_H) + \Pi_t^B \cdot \mathbb{1}(t \in \mathcal{T}_\Pi)] \\ &= \sum_t E[E[H_t^B \cdot \mathbb{1}(t \in \mathcal{T}_H) + \Pi_t^B \cdot \mathbb{1}(t \in \mathcal{T}_\Pi) \mid F_t]] \\ &= \sum_t E[(\mathbb{1}(t \in \mathcal{T}_H) + \mathbb{1}(t \in \mathcal{T}_\Pi))Z_t] = \sum_t E[Z_t]. \end{aligned}$$

We note that if the optimal policy is deterministic (i.e., it makes deterministic decisions in each period  $t$  given the observed information set  $f_t$ ), then if we condition on  $F_t$ ,  $y_t^B$  and  $y_t^{OPT}$  are known deterministically and so are the indicators  $\mathbb{1}(t \in \mathcal{T}_H)$  and  $\mathbb{1}(t \in \mathcal{T}_\Pi)$ . Suppose that the optimal policy is a randomized policy, i.e., it selects an order of size  $Q_t^{OPT}$  as a random function of  $f_t$ , then the same arguments above still work. We now need to condition not only on  $F_t$  but also on  $Q_t^{OPT}$ . Since the inventory control policy does not have any effect on the evolution of the future demands, the arguments above are still valid. This concludes the proof of the theorem.  $\square$

**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 restricted to an integer. In this case, in each period  $s$ , the functions  $l_s^B(q_s^B)$  and  $\pi_s^B(q_s^B)$  are originally defined only for integer values of  $q_s^B$ . We define these functions for any value of  $q_s^B$  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  $q_s'$  that balances the functions  $l_s^B$  and  $\pi_s^B$  is not an integer. Instead, we consider the two consecutive integers  $q_s^1$  and  $q_s^2 := q_s^1 + 1$  such that  $q_s^1 < q_s' < q_s^2$ . In particular,  $q_s' := \lambda q_s^1 + (1 - \lambda)q_s^2$  for some  $0 < \lambda < 1$ . In period  $s$ , we now order  $q_s^1$  units with probability  $\lambda$  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  $Q_s^B = Q_s'$  with support consisting of two points  $\{q_s^1, q_s^2\} = \{q_s^1(f_s), q_s^2(f_s)\}$ , which is a function of the observed information set  $f_s$ . We would like to show that this policy admits the same performance guarantee of two. For each  $t = 1, \dots, T - L$ , let  $Z_t$  be again the random balanced cost of the dual-balancing policy in period  $t$ . Focus now on some period  $s$ . For a given observed information set  $f_s \in \mathcal{F}_s$ , we have for some  $0 \leq \lambda = \lambda(f_s) \leq 1$ ,

$$z_s = E[H_s^B(Q_s') \mid f_s] = \lambda E[H_s^B(q_s^1) \mid f_s] + (1 - \lambda)E[H_s^B(q_s^2) \mid f_s] = E[H_s^B(\lambda q_s^1 + (1 - \lambda)q_s^2) \mid f_s],$$

and

$$z_s = E[\pi_s^B(Q_s') \mid f_s] = \lambda E[\pi_s^B(q_s^1) \mid f_s] + (1 - \lambda)E[\pi_s^B(q_s^2) \mid f_s] = E[\pi_s^B(\lambda q_s^1 + (1 - \lambda)q_s^2) \mid f_s].$$

The second equality (in each of the two expressions above) follows from the fact that we consider piecewise linear functions. By the definition of the algorithm, we also have

$$\lambda E[H_s^B(q_s^1) \mid f_s] + (1 - \lambda)E[H_s^B(q_s^2) \mid f_s] = \lambda E[\pi_s^B(q_s^1) \mid f_s] + (1 - \lambda)E[\pi_s^B(q_s^2) \mid f_s]. \quad (6)$$

It is now readily seen that, for each period  $s$  and each  $f_s \in \mathcal{F}_s$ , we again have  $E[H_s^B(Q_s') + \pi_s^B(Q_s') \mid f_s] = 2z_s$ , i.e.,  $E[H_s^B(Q_s') + \pi_s^B(Q_s') \mid F_s] = 2Z_s$ . This implies that Lemma 4.1 is still valid.

Now, define the sets  $\mathcal{T}_H$  and  $\mathcal{T}_\pi$  in the following way. Let  $\mathcal{T}_H = \{t: X_t^B + Q_t^2 \leq Y_t^{OPT}\}$  and  $\mathcal{T}_\pi = \{t: X_t^B + Q_t^2 > Y_t^{OPT}\}$ . Observe that for each period  $s$ , conditioned on some  $f_s \in \mathcal{F}_s$ , we know deterministically  $x_s^B$ ,  $q_s^B$  and, if the optimal policy is deterministic, we also know  $y_s^{OPT}$ . Therefore, we know whether  $s \in \mathcal{T}_H$  or  $s \in \mathcal{T}_\pi$ . If the optimal policy is also a randomized policy, we condition not only on  $f_s$  but also on the decision made by the optimal policy in period  $s$ . Moreover, if  $s \in \mathcal{T}_H$ , then, with probability one,  $Y_s^B \leq Y_s^{OPT}$  and, if  $s \in \mathcal{T}_\pi$ , then, with probability one,  $Y_s^B \geq Y_s^{OPT}$ . This implies that Lemmas 4.2 and 4.3 are also still valid. The following theorem is now established (the proof is identical to that of Theorem 4.1).

**THEOREM 4.2.** *The randomized dual-balancing policy has a worst-case performance guarantee of two, i.e., for each instance of the uncapacitated periodic-review stochastic inventory control problem, the expected cost of the randomized dual-balancing policy is at most twice the expected cost of an optimal solution, i.e.,  $E[\mathcal{C}(B)] \leq 2E[\mathcal{C}(OPT)]$ .*

**4.4. Dual-balancing policy bad example.** We shall show that the above analysis is tight. Consider an instance with  $h > 0$ ,  $p = h\sqrt{L}$  where  $L > 0$  is again a positive integer that denotes the lead time between the time an order is placed and the time it arrives. Also assume that  $T = 1 + 2L$  and  $\alpha = 1$ . The random demands have the following structure. There is one unit of demand that is going to occur with equal probability either in period  $L + 1$  or in period  $2L + 1$ . For each  $t \neq L + 1, 2L + 1$ , we have  $D_t = 0$  with probability one. Fractional orders are allowed.

It is readily verified that the optimal policy orders one unit in period 1 and incurs expected cost of  $\frac{1}{2}hL$ . On the other hand, the dual-balancing policy will order in each one of the periods  $1, \dots, \sqrt{L} + 1$  just a small amount of the commodity. In particular, in period 1, the dual-balancing policy orders  $1/(\sqrt{L} + 1)$  of a unit (this can be calculated by equating  $\frac{1}{2}(\sqrt{L}h)(1 - q) = \frac{1}{2}Lhq$ , where  $q$  is the size of the order). It can be easily verified that when  $L$  goes to  $\infty$ , the ratio between the expected cost of the dual-balancing policy and the expected cost of the optimal policy converges to two. The calculations are rather messy to present but can be easily coded.

**4.5. Stochastic lead times.** Next, we consider the more general model, where the lead time of an order placed in period  $s$  is some integer-valued random variable  $L_s$  (here we deviate from our convention in that  $l_s$  is not the realization of  $L_s$  but the function defined above). 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.”

We next describe a version of the dual balancing that provides a two-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,  $\mathcal{A}_s := \{t \geq s: s + L_s \leq t \text{ and } t' + L_{t'} > t, \forall t' \in (s, t]\}$ . 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:

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

Now, let  $\tilde{\Pi}_s^P := \sum_{t \in \mathcal{A}_s} \Pi_t^P$  and write  $\mathcal{C}(P) = \sum_{s=1}^T (H_s^P + \tilde{\Pi}_s^P)$ . Similar to the previous case, we consider in each period  $s$  the two functions  $E[H_s^B | f_s]$  and  $E[\tilde{\Pi}_s^B | f_s]$ , 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 two.

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  $E[\tilde{\Pi}_s^B | f_s]$ .

**4.6. Cost transformation.** In this section, we discuss in detail the cost transformation that enables us to assume, without loss of generality, that for each period  $t = 1, \dots, T$ , we have  $c_t = 0$  and  $h_t, p_t \geq 0$ . Consider any instance of the problem with cost parameters that imply no speculative motivation for holding inventory or backorders (as discussed in §2). We use a simple standard transformation of the cost parameters (see Zipkin [31]) to construct an equivalent instance, with the property that for each period  $t = 1, \dots, T$ , we have  $c_t = 0$  and  $h_t, p_t \geq 0$ . The modified instance has the same set of optimal policies. Applying the dual-balancing policy

to that instance will provide a policy that is feasible and also has a performance guarantee of at most two with respect to the original problem. We shall also show that this cost transformation can 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.

We now describe the transformation for the case with no lead time ( $L = 0$ ) and  $\alpha = 1$ ; the extension to the case of arbitrary lead time and  $\alpha < 1$  is straightforward. Recall that any feasible policy  $P$  satisfies, for each  $t = 1, \dots, T$ ,  $Q_t = NI_t - NI_{t-1} + D_t$  (for ease of notation, we omit the superscript  $P$ ). Using these equations we can express the ordering cost in each period  $t$  as  $c_t(NI_t - NI_{t-1} + D_t)$ . Now, replace  $NI_t$  with  $NI_t^+ - NI_t^-$ , its respective positive and negative parts.

This leads to the following transformation of cost parameters: We let  $\hat{c}_t := 0$ ,  $\hat{h}_t := h_t + c_t - c_{t+1}$ , and  $\hat{p}_t := p_t - c_t + c_{t+1}$  ( $c_{T+1} = 0$ ). Note that the assumptions on the cost parameters  $c_t$ ,  $h_t$ , and  $p_t$  discussed in §2 and, in particular, the assumption that there is no speculative motivation for holding inventory or backorders, imply that  $\hat{h}_t$  and  $\hat{p}_t$  above are nonnegative ( $t = 1, \dots, T$ ). Observe that the parameters  $\hat{h}_t$  and  $\hat{p}_t$  will still be nonnegative even if the parameters  $c_t$ ,  $h_t$ , and  $p_t$  are negative and as long as 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 nonnegative cost parameters). It is readily verified that the induced problem is equivalent to the original one. More specifically, for each realization of the demands, the cost of each feasible 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.

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 nonnegative and hence the analysis presented above is valid. Let  $opt$  and  $\overline{opt}$  be the optimal expected cost of the original and modified inputs, respectively. Clearly,  $opt = \overline{opt} + E[\sum_{t=1}^T c_t D_t]$ . Now, the expected cost of the dual-balancing policy in the modified input is at most  $2\overline{opt}$ . Its cost in the original input is then at most  $2\overline{opt} + E[\sum_{t=1}^T c_t D_t] = 2opt - E[\sum_{t=1}^T c_t D_t]$ . 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 two. 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 some real life problems that a major fraction of the total cost is due to 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 as 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 each  $t = 1, \dots, T - L$  to get the same cost transformation. The transformation for  $\alpha > 1$  is also straightforward.

**5. A class of myopic policies.** For many periodic-review inventory control problems with intertemporal correlated demands, there is no known computationally tractable procedure for finding the optimal base-stock inventory levels. 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  $s = 1, \dots, T - L$ , 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 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 Zipkin [31]). It is then possible to implement the myopic policy online, where in each period  $s$ , we compute the base-stock level based on the currently 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 Zipkin [31] 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 Veinott [30], Ignall and Veinott [14], Iida and Zipkin [15], and Lu et al. [20]). 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 zero and one, respectively. For  $t = 2, \dots, T-1$ ,  $D_t := 0$  with probability one and  $D_T := 1$  with probability one. The lead time is considered to be equal to zero, and  $\alpha = 1$ .

It is easy to verify that the myopic policy will order one unit in period 1 and that this will result in an expected cost of  $0.5T$ . On the other hand, if we do not order in period 1, then the expected cost is one. 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*. We shall show that the base-stock levels of this policy are always smaller than the corresponding optimal base-stock levels. Thus, combined with the classical myopic base-stock levels we derive both lower and upper bounds on the optimal base-stock levels. These bounds can be used in computing good inventory control policies. In particular, in follow-up work (see Hurley et al. [13]), we have shown that these bounds can be used to derive improved balancing policies. These policies have a worst-case guarantee of two and their typical performance is significantly better than the dual-balancing policy described in this paper.

Recall the functions  $l_s^P(q_s)$ ,  $\pi_s^P(q_s)$  defined in §4 for each period  $s = 1, \dots, T-L$ , where  $q_s \geq 0$ . Since at each period  $s$  we know  $x_s$ , we can equivalently write  $l_s^P(y_s - x_s)$ ,  $\pi_s^P(y_s - x_s)$ , where  $y_s \geq x_s$ . We now consider in each period  $s$  the problem: minimize  $(l_s^P(y_s - x_s) + \pi_s^P(y_s - x_s))$  subject to  $y_s \geq x_s$ , i.e., minimize 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  $\pi_s^P(y_s - x_s) - \pi_s^P(y_s)$  do not depend on  $y_s$  for  $y_s \geq x_s$ . This gives rise to the following equivalent one-period problem:  $\min_{y_s \geq x_s} (l_s^P(y_s) + \pi_s^P(y_s))$ . 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 [5].

For each  $t = 1, \dots, T$  and  $f_t \in \mathcal{F}_t$ , let  $R^M(f_t)$  be the smallest base-stock level resulted 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 §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  $t = s+1, \dots, T$  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 smaller than the expected cost of the optimal solution. From §3, we know that the expected cost of each policy  $P$  can be expressed as  $\sum_{t=s}^{T-L} E[H_t^P + \Pi_t^P]$ . Now, by the definition of  $R^M(f_s)$ , we know that

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

Moreover, for each  $t \in (s, T]$ , the inventory position  $Y_t^P$  will always be at least  $Y_t^{OPT}$  and, therefore,  $E[\Pi_t^P | f_s] \leq E[\Pi_t^{OPT} | f_s]$ . 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  $t = s+1, \dots, T$ , we have that  $E[H_t^P | f_s] \leq E[H_t^{OPT} | f_s]$ . This concludes the proof.  $\square$



We now define a generalization that captures the myopic policy and the minimization policy as two special cases. For each  $t = 1, \dots, T - L$ , we define a sequence of one-period problems for each  $k_t = 0, \dots, T - t$ ; each generates a corresponding base-stock level. 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  $[t, t + L + k_t]$  and the expected backlogging cost in period  $t + L$ . In other words, the parameter  $k_t$  defines the length of the horizon considered in the one-period problem being solved in period  $t$ . For each sequence of  $k_1, \dots, k_T$ , we get a corresponding  $k$ -minimizing policy. It is clear that if  $k_t = 0$  for each  $t$ , we get the myopic policy, and if  $k_t = T - t$ , 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. We again refer to Hurley et al. [13] for discussion on the performance of these policies.

**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 again, without loss of generality, that  $c_t = 0$ . 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  $L = 0$ ,  $\alpha = 1$  and that in each period  $t = 1, \dots, T$ , the conditional joint distribution  $I_t$  of  $(D_t, \dots, D_T)$  is such that the demand  $D_t$  is known deterministically (i.e., with probability one). 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 §1, for many settings it is known that the optimal solution can be described as a set  $\{(s_t, S_t) = (s_t(f_t), S_t(f_t))\}_t$ . 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* denoted by  $TB$ , and analyze its worst-case expected performance. Specifically, we show that its expected cost is at most three times the expected cost of an optimal policy. 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: When to Order.* At the beginning of 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  $(s^*, s]$  exceeds  $K$ . 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: How Much to Order.* 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 backorders 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.

**6.2. 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). Now, for each  $i = 0, \dots, T$ , let  $Z_i$  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  $Z_i = 0$ . Observe that  $[S_i, S_{i+1})$  and  $(S_i, S_{i+1}]$  are random intervals induced by the triple-balancing policy. 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  $\mathcal{C}(B) = \sum_{i=0}^T Z_i \cdot \mathbb{1}(S_i < T)$  and

$C(OPT) = \sum_{i=0}^T Z'_i \cdot \mathbb{1}(S_i < T)$ . 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  $K \cdot E[N]$ .*

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

$$E[\mathcal{C}(OPT)] = \sum_{i=0}^T E[\mathbb{1}(S_i < T) E[Z'_i | S_i, F_{S_i}]] \geq \sum_{i=0}^T E[\mathbb{1}(S_i < T) E[Z'_i \cdot \mathbb{1}(S_{i+1} \leq T) | S_i, F_{S_i}]].$$

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

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

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 two 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,  $[Z'_i \cdot \mathbb{1}(S_{i+1} \leq T) | s_i, f_{s_i}] \geq K \cdot \mathbb{1}(S_{i+1} \leq T | s_i, f_{s_i})$ . Taking expectations, we have  $E[Z'_i \cdot \mathbb{1}(S_{i+1} \leq T) | S_i, F_{S_i}] \geq K \cdot \Pr(S_{i+1} \leq T | S_i, F_{S_i})$ .

This implies that

$$\begin{aligned} E[\mathcal{C}(OPT)] &\geq K \cdot E\left[\sum_{i=0}^T \mathbb{1}(S_i < T) \cdot \Pr(S_{i+1} \leq T | S_i, F_{S_i})\right] \\ &= K \cdot E\left[\sum_{i=0}^T E[\mathbb{1}(S_i < T) \cdot \mathbb{1}(S_{i+1} \leq T) | S_i, F_{S_i}]\right] = K \cdot E[N]. \quad \square \end{aligned}$$

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  $2K \cdot E[N]$ .

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

PROOF. Clearly,

$$E[\mathcal{C}(TB) - \mathcal{C}(OPT)] = E\left[\sum_{i=0}^T (Z_i - Z'_i) \cdot \mathbb{1}(S_i < T)\right] = \sum_{i=0}^T E[\mathbb{1}(S_i < T) \cdot E[(Z_i - Z'_i) | S_i, F_{S_i}]].$$

We next bound  $E[(Z_i - Z'_i) | S_i, F_{S_i}]$  for each  $i = 0, \dots, T$ . For  $i = 0$ , it is clear that the holding costs that the triple-balancing policy and  $OPT$  incur over  $[s_0, S_1)$  are identical (this cost is due to initial inventory that exists at the beginning of the horizon). Also, observe that the backlogging and ordering costs of the triple-balancing 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  $i = 1, \dots, T$ , we condition on some  $s_i$  and  $f_{s_i} \in \mathcal{F}_{s_i}$ . We then know what are  $y_{s_i}^{TB}$  and  $y_{s_i}^{OPT}$ , deterministically. We now claim that:

$$\begin{aligned} (Z_i - Z'_i) | s_i, f_{s_i} &\leq \mathbb{1}(y_{s_i}^{TB} \leq y_{s_i}^{OPT}) \cdot (K + \mathbb{1}(S_{i+1} \leq T | s_i, f_{s_i}) \cdot K) \\ &\quad + \mathbb{1}(y_{s_i}^{TB} > y_{s_i}^{OPT}) \cdot (H_i | s_i, f_{s_i} + \mathbb{1}(S_{i+1} \leq T | s_i, f_{s_i}) \cdot K). \end{aligned}$$

In the first case where  $y_{s_i}^{TB} \leq y_{s_i}^{OPT}$ , we know that  $OPT$  will incur over  $[s_i, S_{i+1})$  at least as much holding cost as the triple-balancing policy. By the construction of the algorithm we know that the triple-balancing 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  $y_{s_i}^{TB} > y_{s_i}^{OPT}$ , 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 triple-balancing policy incurs on that interval. In addition, the triple-balancing policy 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:

$$\begin{aligned} E[(Z_i - Z'_i) | S_i, F_{s_i}] &\leq E[\mathbb{1}(y_{s_i}^{TB} \leq y_{s_i}^{OPT}) \cdot (K + \mathbb{1}(S_{i+1} \leq T) \cdot K) | S_i, F_{s_i}] \\ &\quad + E[\mathbb{1}(y_{s_i}^{TB} > y_{s_i}^{OPT}) \cdot (H_i + \mathbb{1}(S_{i+1} \leq T) \cdot K) | S_i, F_{s_i}] \leq E[K + \mathbb{1}(S_{i+1} \leq T) | S_i, F_{s_i}]. \end{aligned}$$

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  $i = 2, \dots, T$ , we have

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

Finally, we have that:

$$\begin{aligned} E\left[\sum_{i=0}^T (Z_i - Z'_i) \cdot \mathbb{1}(S_i < T)\right] &\leq K + K \cdot E\left[\sum_{i=1}^T \mathbb{1}(S_i < T)\right] + K \cdot E\left[\sum_{i=1}^{T+1} \mathbb{1}(S_i < T) \cdot \mathbb{1}(S_{i+1} \leq T)\right] \\ &= K + K \cdot E[N] + K \cdot (E[N] - 1) = 2K \cdot E[N]. \quad \square \end{aligned}$$

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

**THEOREM 6.1.** *For each instance of the stochastic lot-sizing problem, the expected cost of the triple-balancing policy is at most three 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. Specifically, it is sufficient to have an efficient oracle that can compute the functions  $l_s$  and  $\pi_s$ .

In subsequent papers (Levi et al. [19], Levi et al. [18]), we consider two generalizations of the model considered in this paper: the extension to the case where there exists a hard capacity constraint on the number of units ordered in each period, and the extension to multiechelon supply chains. 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 other supply chain related models.

It would also be important to establish a more rigorous analysis of the computational hardness of these models (see, for example, the recent work by Halman et al. [12]).

**Acknowledgments.** This research was conducted while Retsef Levi was a Ph.D. student in the School of ORIE and Martin Pál was a Ph.D. student in the Department of Computer Science, at Cornell University. Levi's research was partially supported by a grant from Motorola and National Science Foundation (NSF) Grants CCR-9912422 and CCR-0430682, and Pál's research was supported by U.S. Office of Naval Research Grant N00014-98-1-0589. The research of Robin Roundy was partially supported by grants from Motorola, NSF Grant DMI-0075627, and the Querétaro Campus of the Instituto Tecnológico y de Estudios Superiores de Monterrey, and the research of David Shmoys was partially supported by NSF grants CCR-9912422, CCR-0430682, and CCF-0635121. The authors thank Shane Henderson for stimulating discussions and helpful suggestions that made the paper significantly better. They thank Jack Muckstadt for his valuable comments that led to the extension to the stochastic lead times, and deeply thank John Tsitsiklis who suggested the bad example discussed in §4. The authors thank the anonymous referees for numerous valuable comments that made the paper more precise and improved its exposition.

## References

- [1] Axsäter, S. 1990. Simple solution procedures for a class of two-echelon inventory problems. *Oper. Res.* **38** 64–69.
- [2] Axsäter, S., P. Lundell. 1984. In process safety stock. *Proc. 23rd IEEE Conf. Decision Control, Lecture Notes in Computer Science*, Springer, Heidelberg, Germany, 839–842.
- [3] Bertsimas, D., A. Thiele. 2004. A robust optimization approach to supply chain management. *Proc. 14th IPCO, Lecture Notes in Computer Science*, Springer, Heidelberg, Germany, 86–100.
- [4] Chan, E. W. M. 1999. Markov chain models for multi-echelon supply chains. Ph.D. thesis, Cornell University, Ithaca, NY.
- [5] Chan, E. W. M., J. Muckstadt. 1999. Markov chain models for multi-echelon supply chains. Unpublished manuscript.
- [6] Chen, F., J. Song. 2001. Optimal policies for multi-echelon inventory problems with Markov-modulated demand. *Oper. Res.* **49** 226–234.
- [7] Dean, B. C., M. X. Goemans, J. Vondrák. 2004. Approximating the stochastic knapsack problem: The benefit of adaptivity. *Proc. 45th Annual IEEE Sympos. Foundations Comput. Sci.*, IEEE Computer Society, Washington, D.C., 208–217.
- [8] Dong, L., H. L. Lee. 2003. Optimal policies and approximations for a serial multiechelon inventory system with time-correlated demand. *Oper. Res.* **51** 969–980.
- [9] Dye, S., L. Stougie, A. Tomasgard. 2003. The stochastic single resource service provision problem. *Naval Res. Logist.* **50** 869–887.
- [10] Erkip, N., W. H. Hausman, S. Nahmias. 1990. Optimal centralized ordering policies in multi-echelon inventory systems with correlated demands. *Management Sci.* **36** 381–392.
- [11] Gallego, G., Ö. Özer. 2001. Integrating replenishment decisions with advanced demand information. *Management Sci.* **47** 1344–1360.
- [12] Halman, N., D. Klabjan, M. Mostagir, J. Orling. 2006. A fully polynomial time approximation scheme for single-item stochastic lot-sizing problems with discrete demands. Working paper, MIT, Boston, MA.
- [13] Hurley, G., P. Jackson, R. Levi, R. O. Roundy, D. B. Shmoys. 2005. New policies for stochastic inventory control models—theoretical and computational results. Submitted.
- [14] Ignall, E., A. F. Veinott. 1969. Optimality of myopic inventory policies for several substitute products. *Management Sci.* **15** 284–304.
- [15] Iida, T., P. H. Zipkin. 2006. Approximate solutions of a dynamic forecast-inventory model. *Manufacturing Service Oper. Management* **8** 407–425.
- [16] Joneja, D. 1990. The joint replenishment problem: New heuristics and worst-case performance bounds. *Oper. Res.* **38** 723–771.
- [17] Lee, H. L., K. C. So, C. S. Tang. 1999. The value of information sharing in two-level supply chains. *Management Sci.* **46** 626–643.
- [18] Levi, R., R. O. Roundy, V. A. Truong. 2005. Provably near-optimal balancing policies for multi-echelon stochastic inventory control models. Submitted.
- [19] Levi, R., R. O. Roundy, D. B. Shmoys, V. A. Truong. 2004. Approximation algorithms for capacitated stochastic inventory control models. Submitted.
- [20] Lu, X., J. S. Song, A. C. Regan. 2003. Inventory planning with forecast updates: Approximate solutions and cost error bounds. *Oper. Res.* Forthcoming.
- [21] Möhring, R. H., F. J. Radermacher, G. Weiss. 1984. Stochastic scheduling problems I: General strategies. *ZOR-Zeitschrift Oper. Res.* **28** 193–260.
- [22] Möhring, R. H., F. J. Radermacher, G. Weiss. 1984. Stochastic scheduling problems II: Set strategies. *ZOR-Zeitschrift Oper. Res.* **29** 65–104.
- [23] Möhring, R. H., A. Schulz, M. Uetz. 1999. Approximation in stochastic scheduling: The power of LP-based priority policies. *J. ACM* **46** 924–942.
- [24] Muharremoglu, A., J. N. Tsitsiklis. 2001. A single-unit decomposition approach to multi-echelon inventory systems. Working paper, Columbia University, New York.
- [25] Özer, Ö., G. Gallego. 2001. Integrating replenishment decisions with advance demand information. *Management Sci.* **47** 1344–1360.
- [26] Shmoys, D. B., C. Swamy. 2004. Stochastic optimization is (almost) as easy as deterministic optimization. *Proc. 45th Annual IEEE Sympos. Foundations Comput. Sci.*, IEEE Computer Society, Washington, D.C., 228–237.
- [27] Silver, E. A., H. C. Meal. 1973. A heuristic selecting lot-size requirements for the case of a deterministic time varying demand rate and discrete opportunities for replenishment. *Production Inventory Management* **14** 64–74.
- [28] Song, J., P. Zipkin. 1993. Inventory control in a fluctuating demand environment. *Oper. Res.* **41** 351–370.
- [29] Stougie, L., M. H. van der Vlerk. 2003. Approximation in stochastic integer programming. Technical report, SOM Research Report 03A14, Eindhoven University of Technology, Eindhoven, The Netherlands.
- [30] Veinott, A. F. 1965. Optimal policy for a multi-product, dynamic, non-stationary inventory problem. *Management Sci.* **12** 206–222.
- [31] Zipkin, P. H. 2000. *Foundations of Inventory Management*. McGraw-Hill, New York.