Approximating the Nonlinear Newsvendor and Single-Item Stochastic Lot-Sizing Problems When Data Is Given By an Oracle

N. Halman J. Orlin D. Simchi-Levi
MIT, Cambridge, MA, E-mail: {halman,jorlin,dslevi}@mit.edu

April 30, 2011

Abstract

The single-item stochastic lot-sizing problem is to find an inventory replenishment policy in the presence of discrete stochastic demands under periodic review and finite time horizon. A closely related problem is the single-period newsvendor model. It is well known that the newsvendor problem admits a closed formula for the optimal order quantity whenever the revenue and salvage values are linear increasing functions and the procurement (ordering) cost is fixed plus linear. The optimal policy for the single-item lot-sizing model is also well known under similar assumptions.

In this paper we show that the classical (single-period) newsvendor model with fixed plus linear ordering cost cannot be approximated to any degree of accuracy when either the demand distribution or the cost functions are given by an oracle. We provide a fully polynomial time approximation scheme for the nonlinear single-item stochastic lot-sizing problem, when demand distribution is given by an oracle, procurement costs are provided as nondecreasing oracles, holding/backlogging/disposal costs are linear, and lead time is positive. Similar results exist for the nonlinear newsvendor problem. These approximation schemes are designed by extending the technique of K-approximation sets and functions.

1 Introduction

Inventory control plays a significant role in operations management. In a typical inventory system, a facility, e.g., a retail outlet or a warehouse, maintains an inventory of a particular product. Since demand is random, the facility only has information regarding its distribution. The facility's objective is to decide at what point to reorder a new batch of products, and how much to order so as to minimize the expected cost of ordering and holding inventory. In many such systems, ordering costs consist of two components: a fixed amount, independent of the size of the order, e.g., the cost of sending a vehicle from the supplier to the facility, and a variable amount which is linearly dependent on the number of products ordered. Inventory holding cost, typically linear with the amount of inventory kept at the end of each time period, is incurred at every time period.

Newsvendor problem (NV). A fundamental single-period problem in stochastic inventory theory is the newsvendor problem. A vendor needs to decide how many units x of an item with short life cycle (such as newspapers, fashion items and electrical circuits) to order based on the known demand distribution, the costs of ordering, and the revenues from sales and salvage. Let the cost of ordering x units of the item in the beginning of the period be $c(x)$, the revenue of selling x units throughout the period be $r(x)$, and the salvage value of returning x units in the end of the period be $s(x)$, where all these functions are nonnegative and nondecreasing. The stochastic demand D for the item is a discrete random variable with support that is contained in $[0, \dots, M]$ for a given M, and is described by a cumulative distribution function (CDF) $F(y) = F_D(y) = \text{Prob}(D \leq y)$. Having x units of inventory in the beginning of the period, the vendor decides on the number y of items to order in order to maximize her profit, i.e.,

$$z(x) \quad = \quad \max_{y} \quad E_D[r(\min(D, x+y)) + s(x+y-D) - c(y)], \tag{1}$$

where the expectation is with respect to the random variable D. Arrow, Harris and Marschak [AHM51] showed some 60 years ago that if all costs and revenues are linear, i.e., $c(x) = cx,\ r(x) = rx$, and $s(x) = sx$ (clearly, these cost parameters must satisfy $r > c > s$, otherwise the problem can trivially be solved), then this problem permits a simple solution: Determine the maximum value S such that $\text{Prob}(D \le S) \le \frac{r-c}{r-s}$ and order enough (i.e., $\max\{0, S-x\}$) to bring the stock level to S, the so-called *base stock policy*. See also page 120 in [SCB05] or any basic book on inventory management. Interest in the NV has greatly increased over the past 50 years. This interest stems from the fact that the problem serves as a building block in many inventory models as well as its relevance to practice.

Nonlinear newsvendor problem (NNV). Assuming that the ordering cost $c(\cdot)$, revenue $r(\cdot)$ and salvage value $s(\cdot)$ are all linear functions in the number of units is often not realistic, e.g., when the ordering cost includes a setup cost and quantity discounts. Indeed, during the last half century a large body of work has been focused on the *structure* of the optimal policy as a function of the initial inventory level, under various assumptions. In the following paragraph we give a few examples. Please refer to the excellent extensive survey of Porteus [Por90] and the references therein for more detail.

We first address the case of convex revenues and holding costs[1]. It is well known that then a base stock policy is optimal. In the case of quadratic revenues and holding costs, base stock policy is known to be optimal, where S equals the mean demand. If in addition there are order setup costs, then (s, S) policy is optimal, i.e., one orders enough to bring the stock level up to S if the initial stock level is bellow s, and does not order otherwise. When there are minimum and maximum order quantities and piecewise-linear order costs then the optimal order level is a piecewise-linear function of the initial stock level. If the order costs are convex, then generalized base stock policy is optimal, i.e., the stock level after ordering is an increasing function of the initial stock level, and the optimal amount ordered is a decreasing function of the initial stock level; if in addition the ordering cost is also piecewise-linear, a finite generalized base stock policy is optimal, i.e., there is a finite number of distinct base stock levels. In the case of concave order costs, a generalized (s, S) policy is optimal, i.e., the level $y(x)$ up to which one orders, as a function of the initial inventory level x has the following form: there are two parameters, s and S, such that $y(x)$=0 if $x \ge s$, and $y(z) \ge y(x) \ge S \ge s$ for $z \le x \le s$. When $x \le s$, then a positive order is made. The lower the initial level of inventory, the higher the level of inventory after ordering. If, on the other hand, the order costs are piecewise-linear and concave, a finite generalized (s, S) policy is optimal, i.e., there exist $s_1 \le s_2 \le \ldots \le s_n \le S_n \le \ldots \le S_2 \le S_1$ such that the optimal policy, as a function of the initial inventory level x is to order up to S_1 if $x < s_1$, order up to S_2 if $s_1 \le x < s_2$, and so on, order up to S_n if $s_{n-1} \le x < s_n$, and do not order otherwise. We note that in the general case where all these functions are arbitrarily nondecreasing, an optimal policy does not have any structure. Perhaps for this reason there is almost no research about the general case.

Single-item stochastic lot-sizing problem (SLS). We also address the single-item stochastic economic lot-sizing problem. This problem can be described as follows. Let T be the length of the planning horizon. At the beginning of each period t–e.g., each week or every month–the inventory of a certain item at a warehouse is reviewed, the inventory level is noted, and an order of x_t units is placed. If $x_t > 0$, the order arrives after L time periods, i.e., there is a lead time of L time periods. Just after the replenishment decision is made, the demand D_t is observed. The demand is either immediately satisfied or (partially) backlogged, depending on the inventory on hand. Backlogging is represented as a negative inventory level. Last, a disposal decision is made. The holding/backloogging cost is accounted for at the end of the time period. The random variables $D_1, \ldots, D_T$ are independent, and are not necessarily identically distributed. We assume without loss of generality that there is no demand in the last $L+1$ time periods, i.e., $D_T = \ldots = D_{T-L} = 0$.) We summarize below the functions and variables involved $(t = 1 \ldots, T)$:

[1] for the ease of presentation we refer to the negative of the salvage value, $-s(\cdot)$, as holding cost.

x_t :	procurement quantity in time period t (if $L > 0$ then $x_0 = \ldots = x_{1-L} = 0$);
I_t :	inventory level at the beginning of time period t, just before the arrival of an order;
$c_t(x)$:	procurement cost in time period t, given an order of size $x > 0$;
y_t :	disposal quantity in time period t;
$d_t(y)$:	disposal cost in time period t, given a disposal of size $y > 0$;
$h_t(x)$:	holding cost in time period t, given positive inventory level x at the end of the time period;
$b_t(x)$:	backlogging cost in time period t, given negative inventory level $-x$ at the end of the time period.

(For ease of notation we define $c_t(0) = d_t(0) = b_t(0) = h_t(0) = 0$ and $h_t(x) = b_t(-x)$ for $x < 0$). We assume functions $c_t(\cdot), d_t(\cdot), h_t(\cdot), b_t(\cdot)$ are all nonnegative rational valued, and are computed in polynomial time. We denote by D the random vector of demands, i.e., $D = (D_1, \ldots, D_T)$. The procurement and disposal cost functions $c_t(\cdot), d_t(\cdot)$ are nondecreasing nonnegative over $\mathbb{Z}^+$, and the holding cost function $h_t(\cdot)$ is nonnegative and unimodal over $\mathbb{Z}$ and attains a minimum at $x = 0$.
The objective is to minimize the total expected cost. The problem can be formulated as finding

$$z^*(I_1) = \min_{x_t, y_t} E_D \left(\sum_{t=1}^{T} c_t(x_t) + d_t(y_t) + h_t(I_t + x_{t-L} - y_t - D_t) \right), \tag{2}$$

where the expectation is taken with respect to the joint distribution of $D_1, \ldots, D_t$, and subject to the system dynamics

$$I_{t+1} = I_t + x_{t-L} - y_t - D_t, \qquad t = 1, ..., T. \tag{3}$$

The action space requirement is $x_t, y_t \in \mathbb{Z}^+$ for $t = 1, ..., T$ and the initial inventory level is I_1. The boundary condition is that the values of $h_T(x), b_T(x)$ are very high for positive x. Since the demand in the last $L + 1$ periods is zero, this condition implies that any optimal solution will end time period T with zero inventory.

In the case of fixed plus linear procurement costs and linear holding and backlogging costs it is well known that the optimal policy is (s_t, S_t), i.e., in time period t, if the observed inventory level is below the reorder level s_t, then an order is placed so that the inventory level will reach the base-stock level S_t, otherwise nothing is done. See [SCB05, Por02, Zip00] for an in-depth coverage of the topic. We note that while the structure of the optimal policy is known, it is #P-hard to calculate the reorder points and base-stock levels even in the special case where the cost functions are all linear [HKM+09]. The hardness proof of [HKM+09] relies on the hardness of evaluating CDF of convolutions of discrete random variables. We also note that the NP-hardness of the *deterministic* lot-sizing problem, proved some 3 decades ago by Florian, Lenstra and Rinnoy Kan [FLR80] (even for the special case of zero holding costs, fixed plus linear production costs and capacity limits, i.e., $c_t(x) = \delta_{x>0} c'_t + c''_t x, \;\; c'_t, c''_t \in \mathbb{Z}^+$ for production quantity x up to the capacity limit for period t, and $c_t(x) = M$ for x above that capacity limit and a sufficiently large integer M), implies NP-hardness of SLS as well. (If in the deterministic lot-sizing problem backlogging is not allowed, we set the backlogging costs in the corresponding SLS problem to be arbitrary large.)

Inventory systems when information is given as an oracle. To the best of our knowledge all past work about the newsvendor problem assumed that either all functions are given to the vendor *explicitly* as formulae, or that additional structure about the revenue, salvage and order cost functions is known. But this is not always a realistic assumption. For example, in some cases, the supplier does not reveal the order cost function $c(\cdot)$ to the vendor, and instead gives quotes $c(x)$ for every query x submitted by the vendor. This scenario applies, for example, when the vendor purchases in the spot market, as well as situations where orders are placed over the Internet. For instance, suppose the vendor is a tourism agency that books a block of seats in a specific flight. It is not realistic to assume that the airline provides the vendor with the function $c(\cdot)$, revealing in this way the number of seats it allocates in each of the various booking classes. We believe the aforementioned quotes model is more appropriate in these settings, i.e., the various functions are given to the vendor as "black boxes", or oracle functions.

The same holds for demand forecast. Indeed, firms typically maintain a database that includes historical customer demand information and update it with daily or weekly point of sale (POS) data. In such an environment, there may not be a function representing the demand distribution. Rather, the database provides the probability that demand (or more precisely, sales) is smaller than a certain value, for any value inspected by the user.

Another reason to use oracle functions for representing non-linear cost functions is that oracle functions do not restrict the non-linear function to be given in any particular form. Thus, a fully polynomial time approximation scheme (FPTAS) that relies on an oracle function will be an FPTAS for any function that can be computed in polynomial time. Oracle functions also permit strong negative results, such as proving that an exponential number of steps are required to solve a problem.

An alternative inventory control approach. In this paper we show that the optimal expected profit of the NNV model cannot be approximated to within any given constant factor. This begs the question, what can the firm do to identify effective inventory control policies? Motivated by common firm behavior in the market, we propose to focus on two dimensions when maximizing business performance: expected *profit* and *profit-to-cost ratio*. For example, the expected profit-to-cost ratio of the newsvendor problem when having x units of inventory in the beginning of the period and ordering y units is

$$\nu(y) = \frac{E_D[r(\min(D, x+y)) + s(x+y-D)]}{c(y)} - 1.$$

We claim that a reasonable strategy is to look for a policy that maximizes expected profit while providing a given minimum (expected) profit-to-cost ratio. Indeed, this was the case for IBM when the firm decided to sell its PC business in 2004 to Lenovo–it was not because of lack of profit; rather, it was because of low profit margins (low relative to margins in other IBM businesses). In other words, many businesses avoid investments if the profit-to-cost ratio is close to 0 or negative. Rather, they only invest if the profit-to-cost ratio is strictly positive, and possibly more than some target level ν. (The profit-to-cost ratio is often referred to in the business literature as return on investment or ROI. This ratio has a 1-to-1 correspondence with the profit margin, i.e, the profit-to-revenue ratio.)

This idea is nicely illustrated in Figure 1 where we present expected profit for the linear NV model as a function of order quantity (A) or of profit-to-cost ratio (B). As the reader observes, expected profit decreases with profit-to-cost ratios. Thus, a minimum requirement on (expected) profit-to-cost ratio corresponds to a specific level of expected profit. Of course, the order quantity maximizing expected profit is not the same as the order quantity maximizing expected profit-to-cost ratio–which in the case of the linear NV is equal to ordering a single unit. In Figure 1A, the maximum expected profit is $29,200 and is achieved by ordering 32,020 units thus returning an expected profit-to-cost ratio of 9.13%. However, ordering only 29,000 units achieves an expected profit of $28,130 (which is 97% of maximal expected profit) but increases expected profit-to-cost ratio to 9.7%: this is a cut in the gap between maximal profit-to-cost ratio (10%) and the one associated with optimal profit (9.13%) by 66%.

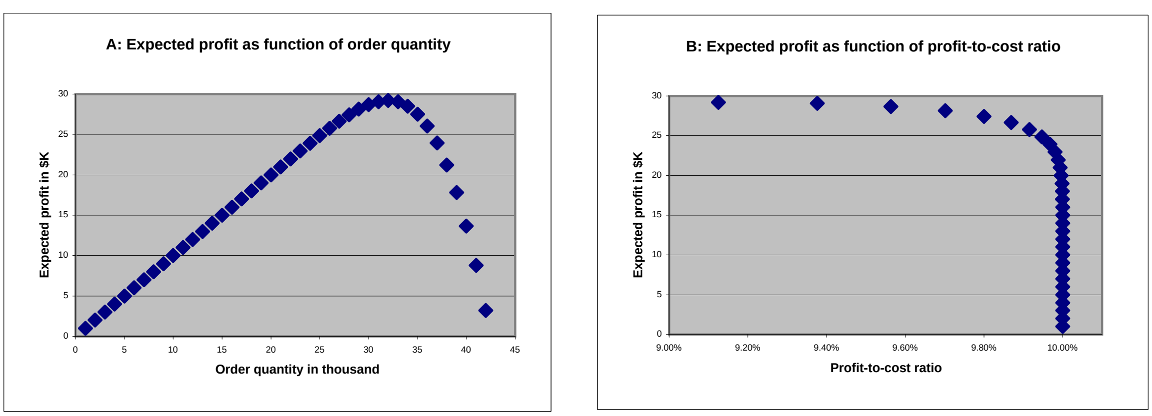

Figure 1: Expected profit for NV with unit cost $10, unit revenue $11, and $D \sim N(\mu = 40K, \sigma = 6K)$

Following this discussion, we suggest that firms plot their expected profit as a function of expected profit-to-cost ratio (Figure 1B) and apply it to determine their strategy. This approach allows us to overcome the challenge associated with the inability to approximate the optimal expected profit of the NNV model as explained below.

Our results. We establish that the non-linear version of the NV problem is provably hard, even in the case that the revenues are linear, and the costs are fixed plus linear. In fact, we show that determining whether there is a solution with positive profit is hard. In such a case, there can be no polynomial time approximation algorithm with a bounded error.

We show that even though the linear newsvendor problem is easy to solve–i.e., it is easy to compute the optimal order quantity–calculating the expected profit takes an exponential number of queries in the worst case under the assumption that the demand probability distribution is given as an oracle. This demonstrates that the gap between the complexity of finding the *argument* (i.e., order quantity) of the optimal solution and finding the *value* (i.e., profit) of this solution is exponential. The proof of this result also implies that the newsvendor problem with fixed plus linear costs requires exponential number of queries to determine the optimal order quantity and optimal expected profit. The reverse is also true. That is, if the demand distribution is given explicitly, but at least one of the cost (or revenue) functions is given as an oracle, then the problem is intractable.

Not only inventory systems when information is given as an oracle are intractable, we also show that these models cannot be approximated to within any given constant. We can still give positive results by following the alternative approach discussed above. Specifically, we develop an FPTAS for NNV that approximates the function plotted in Figure 1.B: For every positive ϵ, δ and ν our algorithm determines a solution with profit-to-cost ratio of at least ν, and with profit at least $\frac{1}{1+\epsilon}$ times the profit of an optimal solution that has profit-to-cost ratio of at least $\nu(1+\delta)$. The running time is polynomial in the size of the problem and in $1/\epsilon + 1/\delta + 1/\nu$. Our algorithm approximates all non-linear newsvendor problems so long as the purchase costs are monotonically non-decreasing in the amount purchased, and the revenues are monotonically non-decreasing in the amount purchased. We show that in general it is not possible to set either ϵ, δ or ν to 0.

Using a similar approach we give an FPTAS for SLS under a wide range of assumptions on the data. We assume that ordering, holding, backlogging and disposal costs are all nonnegative and monotonically non-decreasing. We also need a somewhat technical (but realistic) assumption on the cost of carrying an excess unit of inventory (i.e., on its holding and disposal costs). This allows us the relaxation of the convexity assumption made in [HKM$^+$09]. In addition, it enables us to give FPTASs for various SLS models with a positive lead-time.

Relevance to existing literature. To the best of our knowledge, no FPTAS has been reported in the literature for NNV. Recently, [CKP06, CNC10] have developed FPTASs for the deterministic capacitated lot-sizing problem. [HKM$^+$09] provide an FPTAS for the special case of SLS where the procurement cost is convex. To the best of our knowledge, no FPTAS is known for SLS even for the special case where the procurement cost is fixed plus linear. Moreover, to the best of our knowledge no approximations with worst-case guarantees for SLS with a positive lead-time exist in the literature. Last, we note in passing that approximations for stochastic inventory control problems with bounded error are not common, see [LPRS07, LRST08, HKM$^+$09] for a few examples.

Technique used. To develop the approximation schemes for NNV and SLS we build upon and extend the technique of K-approximation sets and functions, introduced by Halman *et al.* [HKM$^+$09]. For every constant $K > 1$, the idea is to approximate a *monotone* nonnegative function f within a ratio K, by a piecewise-constant function $\tilde{f}$ with "small" number of break-points (i.e., polynomial in the size of the input and $\log_K(\max f/\min f)$). The K-approximation piecewise-constant function $\tilde{f}$ can be minimized efficiently, e.g., by enumerating over all the break-points that belong to the K-approximation set.

Our contribution. This paper makes four contributions. First, we provide new hardness results for finding the maximum expected profit for various NNV models or approximating these values. Second, we provide FPTASs for NNV models in the case that the profit-to-cost ratio is bounded away from 0. In fact, this is the first paper to relate expected profit to profit-to-cost ratio as an approach to suggest an effective strategy. Third, this is the first paper to focus on inventory models where the data is given in a form of a query to an oracle. Our use of oracle functions makes our approximations schemes valid for a wide range of cost functions, revenue functions, and distributions. Fourth, we provide more general conditions that guarantee the existence of an FPTAS. Indeed, in [HKL+08] Halman et *al.* develop a framework for deriving FPTASs for certain stochastic dynamic programs with additive objective functions consisting of monotone or convex *nonnegative* functions (e.g., minimizing costs or maximizing revenues). Here we extend it to problems with additive functions of monotone or convex (not necessarily nonnegative) functions, thus enabling us to approximate maximization of *profits*. We also extend the aforementioned framework to deal with implicitly-defined random variables. Thus, instead of requiring the random variables to be represented explicitly as sequences of values and probabilities, the extended framework can deal with implicitly-defined random variables such as Poisson or Normal distributions.

Organization of the paper. In Section 2 we present the hardness results. In Section 3 we first review K-approximation sets and functions and then give new results about approximations of cumulative distribution functions and subtraction of functions. Using the material given in this section we provide an FPTAS for NNV in Section 4. In Section 5 we review nonincreasing stochastic dynamic programming, which we use in Section 6 to develop FPTASs for various models of SLS with disposal. In Section 7 we deal with several extensions such as SLS with a positive lead time, implicitly-described random variables and non-exact evaluation of CDF and cost functions. We conclude the paper with a discussion and open problems.

2 Hardness results

In this section we show that NNV is intractable and does not admit a constant-factor approximation algorithm in general. Our hardness results rely on the following trivial observation:

Observation 2.1. *Finding a minimum of an arbitrary integer-valued function* $f : [0, ..., N] \rightarrow \mathbb{Z}$*, or even deciding whether such a minimum realizes in either* $[1, ..., \lfloor N/2 \rfloor]$ *or* $[\lceil N/2 \rceil, \ldots, N]$*, requires* $N + 1$ *queries in the worst case.*

Of course, if we have additional information about the function, the number of queries needed may be reduced. For example, if the function is monotone, only two queries are needed. If the function is convex, only $O(\log N)$ queries are needed. But if the function is unimodal with a unique minimum (e.g., a function which is zero everywhere except for one point), the number of queries needed is N.

Theorem 2.2. *The following problems regarding the nonlinear newsvendor problem require exponential number of function evaluations:*

1. *Deciding whether there exists a profit that is strictly positive, even if the revenue and salvage functions are linear and the demand is fixed,*
2. *Deciding whether there exists a profit-to-cost ratio that is strictly positive, even if the revenue and salvage functions are linear and the demand is fixed,*
3. *Calculating the expected optimal profit, even if the ordering cost, revenue and salvage functions are all linear,*
4. *Calculating the order quantity that maximizes expected profit, even if revenue and salvage functions are linear and ordering cost is fixed plus linear.*

Moreover, none of problems 1, 2 and 4 is approximable within any given constant factor.

Proof. Considering the first two problems, let the fixed demand be $D = N$, the revenue for each item sold be 1, the salvage value is 0, and the cost of each item purchased be 1 except items i^* and $i^* + 1$. Item i^* costs 0, and item $i^* + 1$ costs 2. The index i^* (which minimizes the function $f(x)$ = cost of item x) is unknown to the newsvendor and must be determined by evaluations of $f(\cdot)$. If $i^* \leq N$ and if one orders exactly i^* items, then one obtains a profit of 1. Otherwise the optimum profit is 0. Observation 2.1 tells us that deciding whether $i^* \leq N$ requires $\Omega(N)$ queries in the worst case. Note that these two problems cannot be approximated to any degree of accuracy in polynomial time.

We next consider the following instance of the linear NV problem. Let the per-unit cost and per-unit salvage value be identical and equal to 1. Let the per-unit revenue be 2. The support of the demand D is $\{0, 1, \ldots, N\}$ and its distribution is $P(D = i) = 1/N$ for all indices $i = 0, \ldots, N$ but one i^*, for which $P(D = i^*) = 0$. Note that the distribution function of D is unimodal with a unique minimum, and that the input size is $O(\log N)$. In this case an optimal policy is to order N units, and the resulting profit is:

$$E(D) = \frac{N+1}{2} - \frac{i^*}{N}.$$

Therefore, computing the expected profit is equivalent to finding i^*, which by Observation 2.1 requires $O(N)$ queries in the worst case.

We last look at the instance of the NV problem considered above, where the value of N is odd and with an ordering setup cost of $N/2$. If $i^* < N/2$, then every optimal policy must place an order and results in a strictly positive expected profit. Otherwise, the optimal policy is to order nothing and it yields zero profit. This implies that approximating the problem within any constant ratio is equivalent to solving it, i.e., to deciding whether i^* is in either $[1, ..., \lfloor N/2 \rfloor]$ or $[\lceil N/2 \rceil, \ldots, N]$, which by Observation 2.1 requires $O(N)$ queries in the worst case. □

An alternative to using oracle functions is to require that the non-linear function be computable in polynomial time by some Turing Machine. This model is slightly less general than the oracle model.

Problem Q.
Instance: A Turing machine M that computes a real-valued function $f(x) : [0, \ldots, U] \rightarrow \mathbb{R}$ in polynomial time for a given value of x in the domain $[0, \ldots, U]$, and an integer number L.
Question: Is $f(x) \neq x$ for at least one $x \geq L$?

Observation 2.3. *Problem Q is NP-hard.*

Proof. Problem Q is easily shown to be in the class NP. Suppose that we consider integers in binary. For each binary integer x with exactly n bits (the leading bits may be 0), we associate the following subset of $\{1, 2, \ldots, n\}$. $\text{SET}(x) = \{i \mid \text{ the } i\text{-th bit of } x \text{ is } 1\}$.

We carry out a transformation from determining whether there is an independent set of cardinality K on a graph $G = (V, E)$ with n vertices. Let $f(x) = x - 1$ if $\text{SET}(x)$ has at least K vertices and is an independent set. Otherwise, let $f(x) = x$. Note that $f(x)$ can be computed in polynomial time. We conclude the proof by noting that the independent set problem instance has an affirmative answer if and only if Problem Q with f as described above and $U = 2^n$, $L = 0$ has a positive answer. □

Using the observation above, we get the following result.

Observation 2.4. *The problems stated in Theorem 2.2 are all NP-hard whenever all cost functions and the CDF are all computable in polynomial time.*

Proof. We give a proof for the first problem. Proofs for the remaining problems are similar. Let $f(x)$ be the function defined in the proof of Observation 2.3. Note that $f(x)$ can be computed in polynomial time. Note also that $f(x)$ is monotonically non-decreasing.

Suppose that the demand is 2^n, and the revenue per unit is 1. Under these circumstances, the optimum profit for the newsvendor problem is 0 or 1 according to whether $f(x) = x$ over $[0, \ldots, 2^n]$ or not. □

Remark: Computing the optimal order quantity of the linear NV is already #P-hard if the population (i.e., the set of Bernoulli random variables representing the demand of each customer) is subdivided into n subpopulations and if a piecewise-constant CDF is given for the total demand of each of these subpopulations. The reason for this is that it is #P-complete to determine the maximum value x such that Prob(Total Demand $\leq x) \leq d$ [HKM$^+$09]. So, it does not take much for newsvendor problems to become difficult.

3 K-approximation sets and functions and calculus of approximation

Halman *et al.* [HKM$^+$09] introduced the notions of K-approximation sets and functions explained in the Introduction. They used these notions in order to solve a single-item inventory control problem. Halman *et al.* [HLS09] used these notions in order to give an FPTAS for time-cost tradeoff problems in series-parallel project networks. Halman *et al.* [HKL$^+$08] provide a set of general computational rules of K-approximation functions, which they call the *calculus of K-approximation functions*, and which we review below. In addition to developing the calculus of approximation, they develop a framework for deriving FPTASs for certain stochastic dynamic programs and apply their framework on several basic problems in inventory control, economics, and finance. In this section we review K-approximation sets and functions, as well as the calculus of K-approximation functions, and expand the calculus to deal with cumulative distribution functions and subtraction of functions. To simplify the discussion, we modify Halman *et al.*'s definition of the K-approximation function by restricting its domain D to be an interval of integers.

Let $K \geq 1$ and let $\varphi : D \rightarrow \mathbb{R}^+$ be a nonnegative function. We say that $\tilde{\varphi} : D \rightarrow \mathbb{R}$ is a *K-approximation function of* φ (K-approximation of φ, in short) if for all $x \in D$ we have $\varphi(x) \leq \tilde{\varphi}(x) \leq K\varphi(x)$. The following proposition provides a set of general computational rules of K-approximation functions. Its validity follows directly from the definition of K-approximation functions.

Proposition 3.1 (**Calculus of K-approximation Functions** [HKL$^+$08])**.** *For $i = 1, 2$ let $K_i \geq 1$, let $\varphi_i : D \rightarrow \mathbb{R}^+$ be an arbitrary function over domain D, and let $\tilde{\varphi}_i : D \rightarrow \mathbb{R}$ be a K_i-approximation of φ_i. Let $\psi_1 : D \rightarrow D$, and let $\alpha, \beta \in \mathbb{R}^+$. The following properties hold:*

1. *φ_1 is a 1-approximation of itself,*
2. *(linearity of approximation) $\alpha + \beta\tilde{\varphi}_1$ is a K_1-approximation of $\alpha + \beta\varphi_1$,*
3. *(summation of approximation) $\tilde{\varphi}_1 + \tilde{\varphi}_2$ is a $\max\{K_1, K_2\}$-approximation of $\varphi_1 + \varphi_2$,*
4. *(composition of approximation) $\tilde{\varphi}_1(\psi_1)$ is a K_1-approximation of $\varphi_1(\psi_1)$,*
5. *(minimization of approximation) $\min\{\tilde{\varphi}_1, \tilde{\varphi}_2\}$ is a $\max\{K_1, K_2\}$-approximation of $\min\{\varphi_1, \varphi_2\}$,*
6. *(maximization of approximation) $\max\{\tilde{\varphi}_1, \tilde{\varphi}_2\}$ is a $\max\{K_1, K_2\}$-approximation of $\max\{\varphi_1, \varphi_2\}$,*
7. *(approximation of approximation) If $\varphi_2 = \tilde{\varphi}_1$ then $\tilde{\varphi}_2$ is a K_1K_2-approximation of φ_1.*

Let $K > 1$. Let $\varphi : [L, U] \rightarrow \mathbb{Z}^+$ be a monotone function over the contiguous interval $[L, U] = \{L, L+1, \ldots, U-1, U\}$. (Note that the minimal positive value of φ is 1.) We say that an ordered set $S = \{i_1 < \cdots < i_r\}$ of integers is a *K-approximation set of* φ if $L, U \in S \subseteq \{L, \ldots, U\}$ and for each $k = 1$ to $r-1$, if $i_{k+1} > i_k + 1$, then $\frac{\varphi(i_k)}{K} \leq \varphi(i_{k+1}) \leq K\varphi(i_k)$.

Lemma 3.2 ([HKL$^+$08])**.** *Let $K > 1$ and $\varphi : [L, U] \rightarrow \mathbb{Z}^+$ be a monotone function. There exists a K-approximation set of φ with cardinality $O(\log_K(1 + \varphi^{\max}))$, where $\varphi^{\max} = \max\{\varphi(L), \varphi(U)\}$. Furthermore, this set can be constructed in $O\big((1 + \tau(\varphi)) \log_K \varphi^{\max} \log(U - L + 1)\big)$ time, where $\tau(\varphi)$ is the amount of time required to evaluate φ.*

K-approximation sets are very useful for getting succinct approximations for functions that have large domains:

Theorem 3.3 ([HKL$^+$08])**.** *For $i = 1, 2$ let $K_i > 1$, $L_i \geq 1$ and let $\varphi_i : D \rightarrow \mathbb{R}^+$ be a function over domain D. Let $\tilde{\varphi}_i : D \rightarrow \mathbb{R}$ be a L_i-approximation of φ_i. For every fixed $x \in D$, let $\psi_i : D \times E \rightarrow D$ be a function such that $\tilde{\varphi}_i(\psi_i(x, \cdot))$ is monotone over the totally ordered domain E. If $S_i(x) \subseteq E$ is a K_i-approximation set of $\tilde{\varphi}_i(\psi_i(x, \cdot))$, then*

$$\min_{y \in S_1(x) \cup S_2(x)} \{\tilde{\varphi}_1(\psi_1(x, y)) + \tilde{\varphi}_2(\psi_2(x, y))\}$$

is a $\max\{L_1, L_2, \min\{K_1L_1, K_2L_2\}\}$-approximation of

$$\min_{y \in E} \{\varphi_1(\psi_1(x, y)) + \varphi_2(\psi_2(x, y))\}.$$

Halman *et al.* use K-approximation sets to construct approximation functions in the following way:

Definition 3.4 ([HKM$^+$09])**.** *Let $K > 1$ and let $\varphi : [L, U] \to \mathbb{Z}^+$ be a monotone function. Let S be a K-approximation set of φ. A function $\hat{\varphi}$ defined as follows is called the* approximation of φ corresponding to *S. For any integer $L \leq x \leq U$ and successive elements $i_k, i_{k+1} \in S$ with $i_k < x \leq i_{k+1}$ let*

$$\hat{\varphi}(x) = \begin{cases} \varphi(x) & \textit{if } x \in S; \\ \max\{\varphi(i_k), \varphi(i_{k+1})\} & \textit{otherwise.} \end{cases}$$

Note that $\hat{\varphi}$ is a K-approximation of φ. Suppose S is computed according to Lemma 3.2. If we calculate the values of φ on S in advance and store them in a sorted array $(x, \varphi(x))$, then any query for the value of $\hat{\varphi}(x)$, for any x, can be calculated in $O(\log |S|) = O(\log \log_K (1 + \varphi^{\max}))$ time. This is done by performing binary search over S to find the consecutive elements $i_k, i_{k+1} \in S$ such that $i_k < x \leq i_{k+1}$.

3.1 On approximating cumulative distribution functions

In this subsection we expand the Calculus of K-approximation to deal with cumulative distribution functions. One of the assumptions Halman *et al.* [HKL$^+$08] make in their framework is that the CDF of each random variable D is given *explicitly* as a set of ordered pairs $(d, \text{Prob}(D = d))$. In this section we show a way of using CDFs instead of discrete distributions in the analysis, hence enabling us to handle random variables with finite support of size exponential in the size of the remaining input (i.e., not containing the description of the random variables).

Let D be a random variable whose support set consists of nonnegative integers bounded by M, and let $F(\cdot)$ be its CDF. We assume that we have access to an oracle for $F(\cdot)$. Let $\psi(\cdot)$ be a monotone nondecreasing real-valued step function with breakpoints at $a_1 < \ldots < a_n$. Let $f(\cdot, \cdot)$ be a real-valued function, e.g., $f(x, D) = x - D$ or $f(x, D) = \min(x, D)$. We are interested in computing the following function approximately

$$\chi(x) \quad = \quad E_D(\psi(f(x, D))) \quad = \quad \sum_{d=0}^{M} \psi(f(x, d)) \text{Prob}(D = d) \quad = \quad \sum_{d=0}^{M} \psi(f(x, d)) \left(F(d) - F(d-1)\right).$$

It turns out that this becomes easier if one represents $\psi(\cdot)$ in an unusual manner. We represent it as the sum of 2-step functions $\psi_i(\cdot)$, where a 2-step function has the property that $\psi_i(x) = 0$ for $x < a_i$, and $\psi_i(x)$ is constant for $x \geq a_i$. More specifically, we write

$$\psi(x) = \sum_{i=1}^{n} \psi_i(x),$$

where

$$\psi_i(x) = \begin{cases} \psi(a_i) - \psi(a_{i-1}) & \text{if } x \geq a_i; \\ 0 & \text{otherwise,} \end{cases}$$

and where $\psi(a_0) = 0$. It turns out that this representation is very useful for our purposes. Then

$$\begin{aligned} \chi(x) = E_D(\psi(f(x, D))) \quad &= \textstyle\sum_{d=0}^{M} \psi(f(x, d)) \text{Prob}(D = d) \\ &= \textstyle\sum_{d=0}^{M} \sum_{i=1}^{n} \psi_i(f(x, d)) \text{Prob}(D = d) \\ &= \textstyle\sum_{i=1}^{n} \left(\sum_{d=0}^{M} \psi_i(f(x, d)) \text{Prob}(D = d) \right) \\ &= \textstyle\sum_{i=1}^{n} (\psi(a_i) - \psi(a_{i-1})) \text{Prob}(f(x, D) \geq a_i). \end{aligned} \tag{4}$$

If $\text{Prob}(f(x, D) \geq a_i)$ is monotone in x then we get that $\chi(x)$ is expressible as the sum of n functions, each of which is a constant times a monotone function. So, to get an approximated value for χ, we first approximate $\text{Prob}(f(x, D) \geq a_i)$, and then sum the n functions.

Proposition 3.5. *Let D be a nonnegative integer-valued random variable and suppose $\text{Prob}(f(x, D) \geq a_i)$ is monotone in x, $i = 1, \ldots, n$. Let $\xi : [L, \ldots, U] \to \mathbb{Z}^+$ be a nonnegative nondecreasing function. Let*

$K_1, K_2 \geq 1$, $\xi(a_0) = 0$, and let $S = \{a_1 < \ldots < a_n\}$ be a K_1-approximation set of ξ. Finally, let $\eta_i(x)$ denote Prob$(f(x, D) \geq a_i)$ and let $\tilde{\eta}_i(\cdot)$ be a K_2-approximation of $\eta_i(\cdot)$, $i = 1, \ldots, n$. Then

$$\tilde{\xi}_0(x) = \sum_{i=1}^{n} (\xi(a_i) - \xi(a_{i-1}))\tilde{\eta}_i(x)$$

is a K_1K_2-approximation of

$$E_D(\xi(f(x, D))).$$

Moreover, if $\tilde{\eta}_i(\cdot)$ are monotone, then so is $\tilde{\xi}_0(\cdot)$.

Proof. Let ψ be the approximation of ξ corresponding to S (see Definition 3.4) and let χ be defined as in (4). Then

$$E_D(\xi(f(x, D))) = \sum_{d=0}^{M} \xi(f(x, d))\text{Prob}(D = d) \leq \chi(x) \leq \sum_{i=1}^{n} (\xi(a_i) - \xi(a_{i-1}))\tilde{\eta}_i(x),$$

where the first inequality is due to ψ being a K_1-approximation of ξ, and the second inequality is due to (4), $\tilde{\eta}_i$ being a K_2-approximation of η_i, and since ψ and ξ coincide on S. On the other hand, by using similar arguments we have

$$E_D(\xi(f(x, D))) = \sum_{d=0}^{M} \xi(f(x, d))\text{Prob}(D = d) \geq \frac{1}{K_1}\chi(x) \geq \frac{1}{K_1K_2} \sum_{i=1}^{n} (\xi(a_i) - \xi(a_{i-1}))\tilde{\eta}_i(x).$$

We conclude the proof by noting that if $\tilde{\eta}_i$ are all monotone, then so is $\tilde{\xi}_0$. □

By applying the above proposition with $f(x, D) = x - D$ and noting that Prob$(x - D \geq a_i) =$ Prob$(D \leq x - a_i) = F(x - a_i)$ we get

Corollary 3.6. *Let D be a nonnegative integer-valued random variable and let F be its cumulative distribution function. Let $\xi : [L, \ldots, U] \rightarrow \mathbb{Z}^+$ be a nonnegative nondecreasing function. Let $K_1, K_2 \geq 1$, $\xi(a_0) = 0$, and let $S = \{a_1 < \ldots < a_n\}$ be a K_1-approximation set of ξ. Finally, let $\tilde{F}$ be a K_2-approximation of F. Then*

$$\tilde{\xi}_1(x) = \sum_{i=1}^{n} (\xi(a_i) - \xi(a_{i-1}))\tilde{F}(x - a_i)$$

is a K_1K_2-approximation of

$$E_D(\xi(x - D)).$$

Moreover, if $\tilde{F}(\cdot)$ is nondecreasing, then so is $\tilde{\xi}_1(\cdot)$.

We conclude this section by applying Proposition 3.5 with $f(x, D) = \min(x, D)$ and noting that Prob$(\min(x, D) \geq a_i) =$ Prob$(D \geq a_i)\delta_{x \geq a_i} = (1 - F(a_i - 1))\delta_{x \geq a_i}$ (recall that δ_A is 1 if the expression A is true and is 0 otherwise) we get

Corollary 3.7. *Let D be a nonnegative integer-valued random variable and let F be its cumulative distribution function. Let $\xi : [L, \ldots, U] \rightarrow \mathbb{Z}^+$ be a nonnegative nondecreasing function. Let $K_1, K_2 \geq 1$, $\xi(a_0) = 0$, and let $S = \{a_1 < \ldots < a_n\}$ be a K_1-approximation set of ξ. Finally, let $\tilde{F}^c$ be a K_2-approximation of $1 - F$. Then*

$$\tilde{\xi}_2(x) = \sum_{i=1}^{n} (\xi(a_i) - \xi(a_{i-1}))\tilde{F}^c(a_i - 1)\delta_{x \geq a_i}$$

is a nondecreasing K_1K_2-approximation of

$$E_D(\xi(\min(x, D))).$$

3.2 Subtraction of approximation

In this section we expand the Calculus of K-approximation functions to deal with the substraction of functions. Let $z(x)$ denote the value of an optimal solution of an optimization problem, starting at an initial state x, for every initial state x (say, the inventory in the system). If such a function is hard to compute, one aims at approximating it. Since Halman *et al.* [HKM+09] deal with a minimization problem, they define K-approximation functions so that the error is one-sided and is realized by a feasible policy: For every $K \geq 1$ they construct a function $\tilde{z}$ that K-approximates z, i.e., $z(x) \leq \tilde{z}(x) \leq Kz(x)$, for every x. Moreover, for every initial state x they design a feasible policy $\tilde{P}(x)$ that realizes the value given by $\tilde{z}$. If one draws the graph of z and $\tilde{z}$, then $\tilde{z}$ lies "above" z. To stress this point we will say that $\tilde{z}$ K-approximates z *from above.*

If we have a maximization problem on hand, we would like to construct an approximation function $\tilde{z}$ so that the error is still one-sided, but of the other side. In other words, $\tilde{z}$ is a K-approximation of z *from below* if $\frac{z}{K} \leq \tilde{z} \leq z$. Clearly, if $\tilde{z}$ K-approximates z from above then $\frac{\tilde{z}}{K}$ K-approximates z from below. Similarly, if $\tilde{z}$ K-approximates z from below then $K\tilde{z}$ K-approximates z from above.

All the problems dealt by [HKM+09, HKL+08, HLS09] are either for minimizing costs, or maximizing revenues. If one wants to maximize profit, i.e., the difference between revenues and costs, having a rule in the calculus of approximation that deals with subtraction is desirable. Note that such a rule cannot be analogous to "summation of approximation" (Property 3 in Proposition 3.1): it is easy to see that while the ratio between $\varphi_1 + \varphi_2$ and $\tilde{\varphi}_1 + \tilde{\varphi}_2$ is bounded, it is not necessarily so between $\varphi_1 - \varphi_2$ and $\tilde{\varphi}_1 - \tilde{\varphi}_2$ (e.g., whenever φ_1 and φ_2 are very close to each other). The next easy proposition which we prove in the Appendix shows that by imposing the restriction that $\varphi_2 \leq c\varphi_1$ for any given constant $c > 0$, the aforementioned ratio is bounded.

Proposition 3.8 (subtraction of approximation from below)**.** *Let $\varphi_i : D \to \mathbb{R}^+$ be a nonnegative function over domain D and $K_i \geq 1$ be arbitrary, $i = 1, 2$. Let $\tilde{\varphi}_1 : D \to \mathbb{R}^+$ be a K_1-approximation of φ_1 from below, and $\tilde{\varphi}_2 : D \to \mathbb{R}^+$ be a K_2-approximation of φ_2 from above. Let $c < \frac{1}{K_1 K_2}$ be an arbitrary positive real number. If $x \in D$ satisfies $\varphi_2(x) \leq c\varphi_1(x)$ then $(\tilde{\varphi}_1 - \tilde{\varphi}_2)(x)$ is a $\frac{(1-c)K_1}{1-cK_1K_2}$-approximation of $(\varphi_1 - \varphi_2)(x)$ from below.*

Proof. For the ease of presentation let us fix x and write φ_i instead of $\varphi_i(x)$, for $i = 1, 2$. From the definition of K-approximation functions and since $\varphi_2 \leq c\varphi_1$ we get that

$$\frac{\varphi_1}{K_1} \leq \tilde{\varphi}_1 \leq \varphi_1, \qquad \varphi_2 \leq \tilde{\varphi}_2 \leq K_2\varphi_2, \qquad \varphi_2 \leq \frac{c}{1-c}(\varphi_1 - \varphi_2),$$

so

$$\tilde{\varphi}_1 - \tilde{\varphi}_2 \leq \varphi_1 - \varphi_2.$$

On the other hand,

$$\tilde{\varphi}_1 - \tilde{\varphi}_2 \geq \frac{\varphi_1}{K_1} - K_2\varphi_2 = \frac{\varphi_1 - \varphi_2}{K_1} - \frac{K_1K_2 - 1}{K_1}\varphi_2 \geq \frac{\varphi_1 - \varphi_2}{K_1}\left(1 - \frac{c(K_1K_2 - 1)}{1-c}\right) = \frac{1 - cK_1K_2}{(1-c)K_1}(\varphi_1 - \varphi_2).$$

□

We note that whenever φ_1 represents revenues and φ_2 represents costs, then the expression $\varphi_1 - \varphi_2$ represents profit. In this case c is an upper bound on the cost-to-revenue ratio. This ratio has a 1-to-1 correspondence with the profit-to-cost ratio ν in the following way:

$$c = \frac{1}{1+\nu}, \quad \text{and} \quad \nu = \frac{1}{c} - 1. \tag{5}$$

We can also deal with φ_2 that is not necessarily nonnegative. In this case instead of approximating φ_2 we will use it itself. The proof of the proposition below is similar to the proof of Proposition 3.8:

Proposition 3.9 (subtraction of approximation from below)**.** *Let $\varphi_1 : D \to \mathbb{R}^+$ be a nonnegative function over domain D and $K_1 \geq 1$ be arbitrary. Let $\tilde{\varphi}_1 : D \to \mathbb{R}^+$ be a K_1-approximation of φ_1 from below. Let $\varphi_2 : D \to \mathbb{R}$ be an arbitrary function. Let $c < \frac{1}{K_1}$ be a nonnegative real number. If $x \in D$ satisfies $\varphi_2(x) \leq c\varphi_1$ then $(\tilde{\varphi}_1 - \varphi_2)(x)$ is a $\frac{(1-c)K_1}{1-cK_1}$-approximation of $(\varphi_1 - \varphi_2)(x)$ from below.*

We note that when $c = 0$ we get that $\varphi_2(\cdot) \leq 0$, so the proposition above coincides with summation of approximation in the Calculus of K-approximation Functions.

The following proposition is similar to Proposition 3.9, is intended for getting approximations from above for minimization problems, and is used for approximating the single-item lot-sizing problem.

Proposition 3.10 (subtraction of approximation from above)**.** *Let $\varphi_1 : D \rightarrow \mathbb{R}^+$ be a nonnegative function over domain D and $K_1 \geq 1$ be arbitrary. Let $\tilde{\varphi}_1 : D \rightarrow \mathbb{R}^+$ be a K_1-approximation of φ_1 from above. Let $\varphi_2 : D \rightarrow \mathbb{R}$ be an arbitrary function, and let $c < 1$ be an arbitrary nonnegative real number. If $x \in D$ satisfies $\varphi_2(x) \leq c\varphi_1(x)$ then $\tilde{\varphi}_1 - \varphi_2$ is a $\frac{K_1 - c}{(1-c)}$-approximation of $(\varphi_1 - \varphi_2)(x)$ from above.*

We next prove the following theorem which is the analogue of Theorem 3.3 for maximization of a difference of functions.

Theorem 3.11. *Let $K_i > 1$, $L_i \geq 1$ and let $\varphi_i : D \rightarrow \mathbb{R}^+$ be a function over domain D, $i = 1, 2$. Let $\tilde{\varphi}_1 : D \rightarrow \mathbb{R}$ be an L_1-approximation of φ_1 from below and $\tilde{\varphi}_2 : D \rightarrow \mathbb{R}$ be an L_2-approximation of φ_2 from above. For every fixed $x \in D$, let $\psi_i : D \times E \rightarrow D$ be a function such that both $\tilde{\varphi}_1(\psi_1(x, \cdot))$, $\tilde{\varphi}_2(\psi_2(x, \cdot))$ are monotone in the same direction over a totally ordered domain E. Let $S_i(x) \subseteq E$ be a K_i-approximation set of $\tilde{\varphi}_i(\psi_i(x, \cdot))$. Let $c < \frac{1}{K_1 L_1 L_2}$ be an arbitrary positive real number. Then for every $x \in D$, the value of*

$$\tilde{\zeta}(x) = \max_{y \in S_1(x) \cup S_2(x) \mid \tilde{\varphi}_2(\psi_2(x,y)) \leq cK_1 L_1 L_2 \tilde{\varphi}_1(\psi_1(x,y))} \{\tilde{\varphi}_1(\psi_1(x, y)) - \tilde{\varphi}_2(\psi_2(x, y))\} \tag{6}$$

is at least $\frac{1 - cK_1 L_1 L_2}{(1-c)K_1 L_1}$-times the value of

$$\zeta(x) = \max_{y \in E \mid \varphi_2(\psi_2(x,y)) \leq c\varphi_1(\psi_1(x,y))} \{\varphi_1(\psi_1(x, y)) - \varphi_2(\psi_2(x, y))\}. \tag{7}$$

Moreover, if $y^{@}$ is an argmax of $\tilde{\zeta}(x)$, then

$$\varphi_1(\psi_1(x, y^{@})) - \varphi_2(\psi_2(x, y^{@})) \geq \zeta(x), \qquad \textit{and} \qquad \frac{\varphi_2(\psi_2(x, y^{@}))}{\varphi_1(\psi_1(x, y^{@}))} \leq cK_1 L_1 L_2.$$

Proof. Due to symmetry arguments we can assume without loss of generality that both $\tilde{\varphi}_1(\psi_1(x, \cdot))$, $\tilde{\varphi}_2(\psi_2(x, \cdot))$ are nondecreasing. Let x be fixed, let y^* be the minimal argmax of ζ, and let $x_i^* = \psi_i(x, y^*)$, for $i = 1, 2$. From the definition of K-approximation functions and since $\varphi_2(x_2^*) \leq c\varphi_1(x_1^*)$ we get that

$$\frac{\varphi_1}{L_1} \leq \tilde{\varphi}_1 \leq \varphi_1, \quad \varphi_2 \leq \tilde{\varphi}_2 \leq L_2 \varphi_2, \quad \varphi_2(x_2^*) \leq \frac{c}{1-c}(\varphi_1(x_1^*) - \varphi_2(x_2^*)). \tag{8}$$

By the definition of $S_1(x)$ and $S_2(x)$, we have that $y' = \max\{y \leq y^* \mid y \in S_1(x) \cup S_2(x)\}$ satisfies

$$\tilde{\varphi}_1(\psi_1(x, y')) \geq \frac{\tilde{\varphi}_1(\psi_1(x, y^*))}{K_1} \geq \frac{\varphi_1(\psi_1(x, y^*))}{K_1 L_1} = \frac{\varphi_1(x_1^*)}{K_1 L_1}, \tag{9}$$

where the last inequality is due to (8). Since $\tilde{\varphi}_2(\psi_2(x, \cdot))$ is nondecreasing and by using (8) again we get

$$\tilde{\varphi}_2(\psi_2(x, y')) \leq \tilde{\varphi}_2(\psi_2(x, y^*)) \leq L_2 \varphi_2(\psi_2(x, y^*)) = L_2 \varphi_2(x_2^*). \tag{10}$$

We next show that y' satisfies the constraint of the maximization in (6). Indeed,

$$\frac{\tilde{\varphi}_2(\psi_2(x, y'))}{\tilde{\varphi}_1(\psi_1(x, y'))} \leq \frac{K_1 \tilde{\varphi}_2(x_2^*)}{\tilde{\varphi}_1(x_1^*)} \leq K_1 L_1 L_2 \frac{\varphi_2(x_2^*)}{\varphi_1(x_1^*)} \leq cK_1 L_1 L_2, \tag{11}$$

where the first inequality is since $S_1(x) \cup S_2(x)$ is a K_1-approximation set of $\tilde{\varphi}_1(\psi_1(x, \cdot))$ and due to the monotonicity of $\tilde{\varphi}_2(\psi_2(x, \cdot))$. The second inequality is due to (8), and the last inequality is due to the constraint of the maximization in (7). We conclude the first part of the proof by using (8), (9), and (10) to get

$$\begin{aligned} \tilde{\zeta}(x) &\geq \tilde{\varphi}_1(\psi_1(x, y')) - \tilde{\varphi}_2(\psi_2(x, y')) \geq \tfrac{\varphi_1(x_1^*)}{K_1 L_1} - L_2 \varphi_2(x_2^*) = \tfrac{\varphi_1(x_1^*) - \varphi_2(x_2^*)}{K_1 L_1} + (\tfrac{1}{K_1 L_1} - L_2)\varphi_2(x_2^*)) \\ &\geq \left(\tfrac{1}{K_1 L_1} + \tfrac{c(1 - K_1 L_1 L_2)}{(1-c)K_1 L_1}\right)(\varphi_1(x_1^*) - \varphi_2(x_2^*)) = \tfrac{1 - cK_1 L_1 L_2}{(1-c)K_1 L_1} \zeta(x). \end{aligned} \tag{12}$$

The (exact) value of the approximated solution is

$$\varphi_1(\psi_1(x,y^@))-\varphi_2(\psi_2(x,y^@) \geq \tilde{\varphi}_1(\psi_1(x,y^@))-\tilde{\varphi}_2(\psi_2(x,y^@)) \geq \tilde{\varphi}_1(\psi_1(x,y'))-\tilde{\varphi}_2(\psi_2(x,y')) \geq \frac{1-cK_1L_1L_2}{(1-c)K_1L_1}\zeta(x),$$

where the first inequality is due to (8), and the third inequality is due to (12). Moreover, its (exact) cost-to-revenue ratio is

$$\frac{\varphi_2(\psi_2(x,y^@))}{\varphi_1(\psi_1(x,y^@))} \leq \frac{\tilde{\varphi}_2(\psi_2(x,y^@))}{\tilde{\varphi}_1(\psi_1(x,y^@))} \leq cK_1L_1L_2,$$

where the first inequality is due to (8) and the second inequality is since $y^@$ satisfies the constraint of the maximization in (6). □

4 An FPTAS for the nonlinear newsvendor problem

In this section we design a 3-parameter FPTAS for approximating profit functions as follows: Let $\nu > 0$ be a lower threshold for the profit-to-cost ratio of the solution obtained by the algorithm. Let $\delta > 0$ refer to a relative deviation of the profit-to-cost ratio of the solution (to be explained in more detail below). Let $\epsilon > 0$ refer to a relative deviation of the profit. The algorithm produces a solution whose profit-to-cost ratio is at least ν. Moreover, the profit of this solution is at least as large as $\frac{1}{1+\epsilon}$ times the maximum profit under the restriction that the profit-to-cost ratio is at least $\nu(1+\delta)$. In other words, the profit is almost as large as the optimal profit for a slightly perturbed problem. Moreover, the running time is polynomial in the size of the problem and in $1/\epsilon + 1/\delta + 1/\nu$. We conclude this section by showing that in general it is not possible to set either ν, ϵ or δ to 0.

Theorem 4.1. *Let $M \in \mathbb{Z}$ be an arbitrary positive number and let $F(\cdot)$ be the cumulative distribution function of a discrete random variable D with support that is contained in $[0,M]$. Let $\epsilon,\delta,\nu > 0$ be arbitrary positive parameters. Let q^* be the order quantity of a minimal-cost optimal solution of the nonlinear newsvendor with stochastic demand D and profit-to-cost ratio of at least $\nu(1+\delta)$, and let $z(q^*)$ be its value. Then for every $\nu,\delta,\epsilon > 0$ one can compute in $O(\frac{\log M \log^2 r(M)}{\min(\epsilon^2,\delta^2)\nu^2})$ time an approximate solution with order quantity q' and expected profit $\tilde{z}(q')$, such that its value satisfies $z(q^*) \geq \tilde{z}(q') \geq \frac{z(q^*)}{1+\epsilon}$ and its profit-to-cost ratio is at least ν.*

Remark: It is interesting to note that whenever it is known that there exists an optimal solution with a profit-to-cost ratio greater than a given constant $\alpha > 0$, then the above 3-parameter FPTAS collapses into an “ordinary” FPTAS–with a single parameter ϵ–by setting, e.g., $\nu = \alpha/2$ and $\delta = 0.1$.

Proof. Let $1 < K < 2$ be an arbitrary number (we will fix it soon). For every $0 < \epsilon < 1$ we approximate $z(\cdot)$ in (1) by applying Theorem 3.11 with

$$\varphi_1(t) = E_D[r(\min(D,t))+s(t-D)],\ \psi_1(x,y) = x+y,\ \varphi_2(t) = c(t),\ \psi_2(x,y) = y,\ L_1 = K_1 = K_2 = K,\ L_2 = 1.$$

We note that $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$ are nondecreasing. We compute for them approximations in the following way. Since $L_2 = 1$ we take $\tilde{\varphi}_2 = \varphi_2$ and $\tilde{\varphi}_2(\psi_2(x,y)) = c(y)$, which is nondecreasing. Hence a K-approximation set S_2 for it is well defined. As for φ_1, by linearity of expectation we decompose it into two functions $\varphi_1 = f_1 + f_2$, where

$$f_1(t) = E_D[r(\min(D,t))], \quad f_2(t) = E_D[s(t-D)].$$

Note that both f_1 and f_2 are nondecreasing. We next approximate f_1 from below by applying Corollary 3.7 with $\xi = r, K_1 = K$ and $K_2 = 1$. Note that the resulting $\tilde{f}_1$ is a nondecreasing K-approximation of f_1 from below. As for f_2, we approximate it by applying Corollary 3.6 with $\xi = s, K_1 = K$ and $K_2 = 1$, and therefore the resulted $\tilde{f}_2$ is a nondecreasing K-approximation of f_2 from below. Let $\tilde{\varphi}_1 = \tilde{f}_1 + \tilde{f}_2$. By summation of approximation (third property in Proposition 3.1) we get that $\tilde{\varphi}_1$ is a nondecreasing K-approximation of φ_1 from below. Hence a K-approximation set S_1 for it is well defined as well.

By applying Theorem 3.11 (with $K_1 = K_2 = L_1 = K$ and $L_2 = 1$) we get that the value of the approximated solution is at least $\frac{1-cK^2}{(1-c)K^2}$-times the value of the optimal solution $z(x)$, and that the cost-to-revenue ratio of the approximated solution is at most K^2c. We need to fix K and c such that all of the following statements hold:

1. The profit-to-cost ratio of a minimal-cost optimal solution is at least $\nu(1+\delta)$,
2. (regarding a condition in Theorem 3.11) $c < \frac{1}{K^2}$,
3. (regarding the value of the approximated solution) $\frac{1-cK^2}{(1-c)K^2} \geq \frac{1}{1+\epsilon}$,
4. the profit-to-cost ratio of the approximated solution is at least ν.

In order to satisfy 1, and by using (5), we set $\frac{1}{c} - 1 = \nu(1+\delta)$, hence

$$c = \frac{1}{1+\nu(1+\delta)}.$$

Note that $c < 1$. We set

$$K = \sqrt{\frac{1+\epsilon}{1+c\epsilon}} = \sqrt{1+\frac{\epsilon\nu(1+\delta)}{1+\epsilon+\nu(1+\delta)}}, \tag{13}$$

thus statement 2 holds as well. It is easy to check that statement 3 holds for every nonnegative $c < 1$, and that statement 4 holds whenever $\epsilon \leq \delta$.

It remains to analyze the running time of the algorithm. Suppose first that $\epsilon \leq \delta$. In order to find an optimal solution, because $r(x) \geq c(x) \geq s(x)$, it suffices to set the domain of the functions involved to be $[0, M]$. Also note that the largest value computed throughout the algorithm is $r(M)$. In our algorithm we compute approximation sets for r, s, c and $\tilde{\varphi}_1$. By Lemma 3.2 each of these approximation sets is of cardinality $O(\log_K r(M))$, and is calculated in time $O((1+\tau_f)\log_K r(M)\log M)$, where τ_f is the time needed to perform a query of the corresponding function. For the ease of presentation we assume the query time for r, s, c, F is a constant, so the most time-consuming operation is to compute the approximation set for $\tilde{\varphi}_1$, which takes $O(\log_K^2 r(M)\log M)$ time, since by Corollaries 3.6 and 3.7 each evaluation of $\tilde{\varphi}_1$ takes $O(n) = O(\log_K r(M))$ time. In order to work with logarithm of base 2 we recall that for $K = 1+\theta$ with $0 < \theta < 1$ we have $\frac{1}{\log K} = O(\frac{1}{\theta})$. Due to (13) $\log_K r(M) = O(\frac{\log r(M)}{\epsilon\nu})$. If $\epsilon > \delta$ we set $\epsilon = \delta$. Therefore the running time of the algorithm is

$$O\left(\frac{\log M \log^2 r(M)}{\min(\epsilon^2, \delta^2)\nu^2}\right).$$

□

We conclude this section by showing that Theorem 4.1 is in a sense the strongest one can hope for:

Observation 4.2. *One cannot relax the requirement of Theorem 4.1 that each one of ν, ϵ and δ must be a strictly positive real number.*

Proof. We first show that we cannot have $\nu = 0$. Indeed, considering the example in the proof for the 2 first problems in Theorem 2.2 and running the FPTAS with $\nu = 0$ and $\epsilon = \delta = 0.1$ will result in a decision whether there exists a positive profit.

We next show that we cannot allow $\delta = 0$. We change the example in the proof for the 2 first problems in Theorem 2.2 as follows. Let the fixed demand and the cost of each item purchased remain the same, and the revenue for each item sold be now 2. Note that the profit-to-cost ratio when ordering i^* units is $\frac{i^*+1}{i^*-1} \geq 1 + \frac{2}{N-1}$ and is 100% for any other positive order quantity. In this way, running the FPTAS with $\epsilon = 0.1$, $\delta = 0$ and $\nu = 1 + \frac{1}{N}$ will reveal the value of i^*.

Last, we show that we cannot have $\epsilon = 0$. We change the example in the proof for the 2 first problems in Theorem 2.2 as follows. Let the fixed demand remain N. Let the revenue for each item be 3 and the salvage value of each item be 2. Let the cost of each item purchased be 2 except for items i^* and i^*+1. Item i^* costs 1, and item i^*+1 costs 3. While the newsvendor knows that $2N \geq i^* > N$, the exact value of index i^*

is unknown to her, and must be determined by function evaluations. The optimal policy is to order $N + i^*$ (at a cost of $2(N + i^*) - 1$), which results in a profit of $N + 1$. Note that ordering N units yields a profit of only N. Note also that a minimal-cost optimal solution has profit-to-cost ratio of at least 25%. Hence, running the FPTAS with $\epsilon = 0$ and $\delta = \nu = 0.1$ determines the value of i^*. □

5 Nonincreasing stochastic dynamic programming

In this section we review the model for finite horizon stochastic dynamic programs that is studied by [HKL+08]. The model has two principal features: an underlying *discrete time dynamic system*, and a *cost function that is additive over time.* The system dynamics are of the form

$$I_{t+1} = f_t(I_t, x_t, D_t), \quad t = 1, \ldots, T, \tag{14}$$

where

t	is the discrete time index,
I_t	is the state of the system,
x_t	is the action or decision to be selected in time period t,
D_t	is a discrete random variable,
T	is the number of time periods.

The cost function is additive in the sense that the cost incurred at time period t, denoted by $g_t(I_t, x_t, D_t)$, is accumulated over time. Let I_1 be the initial state of the system. Given a realization d_t of D_t, for $t = 1, \ldots, T$, the total cost is

$$g_{T+1}(I_{T+1}) + \sum_{t=1}^{T} g_t(I_t, x_t, d_t),$$

where $g_{T+1}(I_{T+1})$ is the terminal cost incurred at the end of the process. The problem is to find

$$z^*(I_1) = \min_{x_1,\ldots,x_T} E\left\{ g_{T+1}(I_{T+1}) + \sum_{t=1}^{T} g_t(I_t, x_t, D_t) \right\}, \tag{15}$$

where the expectation is taken with respect to the joint distribution of the random variables involved. The optimization is over the actions $x_1, \ldots, x_T$ which are selected with the knowledge of the current state.

The state I_t is an element of a given *state space* $\mathcal{S}_t$, the action x_t is constrained to take values in a given *action space* $\mathcal{A}_t(I_t)$, and the discrete random variable D_t takes values in a given set $\mathcal{D}_t$. The state space and the action space are one-dimensional. We state now the well known DP recursion for this model.

Theorem 5.1 (The DP Recursion [Bel57])**.** *For every initial state I_1, the optimal cost $z^*(I_1)$ of the DP is equal to $z_1(I_1)$, where the function z_1 is given by the last step of the following recursion, which proceeds backward from period T to period 1:*

$$z_{T+1}(I_{T+1}) = g_{T+1}(I_{T+1}),$$
$$z_t(I_t) = \min_{x_t \in \mathcal{A}_t(I_t)} E_{D_t} \left\{ g_t(I_t, x_t, D_t) + z_{t+1}(f_t(I_t, x_t, D_t)) \right\}, \quad t = 1, \ldots, T, \tag{16}$$

where the expectation is taken with respect to the probability distribution of D_t.

The input data of the problem consists of the number of time periods T, the initial state I_1, the function g_{T+1}, and for each time period $t = 1, \ldots, T$, the functions g_t and f_t, and the discrete random variable D_t given explicitly as a set of ordered pairs $(d, \text{Prob}(D_t = d))$. Halman *et al.* assume the following three conditions hold:

Condition 1. *$\mathcal{S}_{T+1}, \mathcal{S}_t, \mathcal{A}_t(I_t) \subset \mathbb{Z}$ for $I_t \in \mathcal{S}_t$ and $t = 1, \ldots, T$. For any set X among these sets, $\log \max_{x \in X} |x + 1|$ is bounded polynomially by the (binary) input size, and the kth largest element in X can be identified in constant time for any $1 \le k \le |X|$. Moreover, $\mathcal{D}_t \subset \mathbb{Q}$ for $t = 1, \ldots, T$.*

Condition 2. *For every $t = 1, \dots, T+1$, the values of function g_t are nonnegative rational numbers. In addition, the logarithm of the ratio between the maximal single-period cost and the maximum of 1 and the minimal positive single-period cost is polynomially bounded by the (binary) size of the input.*

Condition 3 (Nonincreasing DP)**.** *g_{T+1} is nonincreasing. For every $t = 1, \dots, T$, and fixed $d \in \mathcal{D}_t$, $f_t(\cdot,\cdot,d)$ is nondecreasing in its first variable and is monotone in its second variable, and $g_t(\cdot,\cdot,d)$ is a nonnegative function monotone in its second variable. Moreover, either z_t is nonincreasing for $t = 1, \dots, T$, or $g_t(\cdot,\cdot,d)$ is nonincreasing in its first variable and $\mathcal{A}_t(I') \subseteq \mathcal{A}_t(I)$ for all $I', I \in \mathcal{S}_t$ with $I' \leq I$.*

The DP formulation (16) that satisfies Conditions 1-3 is called *nonincreasing.*

Theorem 5.2 ([HKL$^+$08])**.** *Every stochastic nonincreasing DP admits an FPTAS.*

6 Single-item stochastic lot-sizing with disposal

In this section we derive FPTASs for various versions of SLS with disposal by transforming them into nonincreasing DPs and applying the framework of [HKL$^+$08].

We transform SLS into a nonincreasing DP as follows. We split period t in the original problem into two periods in the transformed problem, denoted as periods $2t-1$ and $2t$. Period $2t-1$ corresponds to procurement decisions and procurement costs in period t of the original problem. Period $2t$ corresponds to disposal decisions and disposal and holding costs in period t of the original problem. While production in period $2t-1$ in the transformed problem corresponds to procurement in period t of the original problem, production of $-y$ units in period $2t$ in the transformed problem corresponds to disposal of y units in period t of the original problem.

It is convenient to consider the *stochastic network flow (SNF) minimization problem corresponding to the lot-sizing problem* as follows. The network $G = (V, E)$ consists of $2T + 1$ vertices labeled $0, \dots, 2T$. Vertices $1, \dots, 2T$ are connected in series, and vertex 0 is connected to each of the vertices $1, \dots, 2T$. For each odd-numbered vertex $2t-1$ there is an ingoing edge $e_{0,2t-1}$ with cost function $c_{0,2t-1}(\cdot) = c_t(\cdot)$ representing production. There is also an ingoing edge $e_{2t,2t-1}$ and an outgoing edge $e_{2t-1,2t}$, both connecting to vertex $2t$ with no cost (i.e., $c_{2t-1,2t}(\cdot) = c_{2t,2t-1}(\cdot) = 0$). If $t > 1$, then there is also an ingoing edge $e_{2(t-1),2t-1}$ with cost function $c_{2(t-1),2t-1}(\cdot) = h_{t-1}(\cdot)$. Each even-numbered vertex $2t > 0$ has an outgoing edge $e_{2t,0}$ with cost function $c_{2t,0}(\cdot) = d_t(\cdot)$ representing disposal. If $t < T$, then it has also an ingoing edge $e_{2t+1,2t}$ with cost function $c_{2t+1,2t}(\cdot) = b_t(\cdot)$ representing backlogging, see Figure 2.

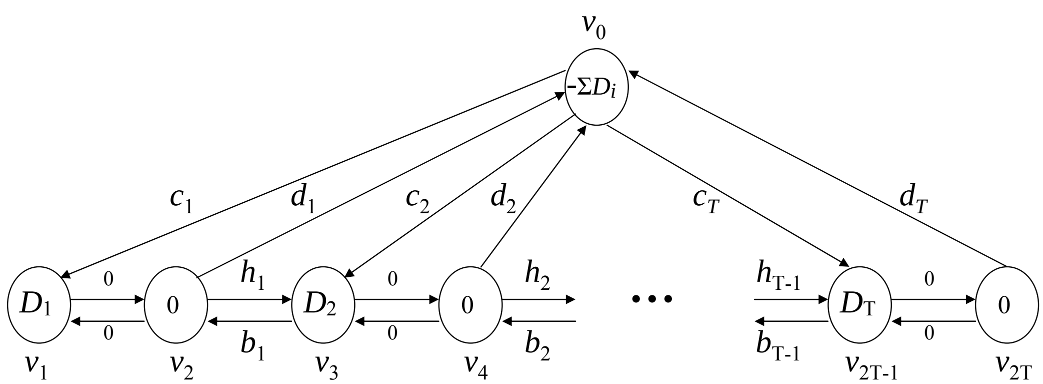

Figure 2: Transforming SLS to SNF

Let

$$A(I) = (A_0, \dots, A_{2T+1}) = (I - \sum_{t=1}^{T} D_t, D_1, 0, D_2, 0, \dots, D_T, 0)$$

be a random vector consisting of the random variables $D_1, \dots, D_T$. The *minimum stochastic cost flow problem* is the optimization model formulated as follows:

$$\text{Minimize } E_{D,P} \sum_{(i,j)\in A} c_{i,j}(x_{i,j}) \tag{17}$$

subject to

$$\sum_{\{j:(j,i)\in E\}} x_{j,i} - \sum_{\{j:(i,j)\in E\}} x_{i,j} = A_i - I\delta_{i=1}, \quad i = 0, 1, \ldots, 2T,$$

where the expectation is taken over the joint distribution of $D_1, \ldots, D_T$ and a dynamic flow policy P. A *dynamic flow policy* is a decision rule for assigning nonnegative values for the variables $x_{i,j}$, that is determined with the gradual realization of the the random variables $D_1, \ldots, D_T$. A policy P first determines the value of $x_{0,1}$ (which is the number of units to produce in time period 1). Then the value of D_1 is revealed, and based upon this information, the policy determines the values of $x_{1,2}, x_{2,1}, x_{2,0}, x_{2,3}, x_{3,2}$ and $x_{0,3}$ (which determine the disposal amount in period 1 and the production amount in period 2). In general, just after the value of D_t is revealed, the policy determines the values of $x_{2t-1,2t}, x_{2t,2t-1}, x_{2t,0}, x_{2t,2t+1}, x_{2t+1,2t}$ and $x_{0,2t+1}$. It is easy to see that each feasible solution for SLS with cost C is transferred to a feasible solution for the corresponding SNF with the same cost and vice versa.

In order to formulate the SLS (2) as a monotone DP (16) we set the time horizon of the transformed problem to consist of $2T$ time periods and define the cost functions $g_t(\cdot)$ as follows (recall that if $y < 0$ then $h_t(y) = b_t(-y)$):

$$g_{2t-1}(I, x, D) = c_t(x), \qquad g_{2t}(I, x, D) = d_t(-x) + h_t(I + x), \quad t = 1, \ldots, T, \qquad g_{2T+1}(I) = 0.$$

We define the transition functions $f_t(\cdot)$ and the random variables $D'_1, \ldots, D'_{2T}$ as follows:

$$f_{2t-1}(I, x, D) = f_{2t}(I, x, D) = I + x - D, \qquad D'_{2t-1} = D_t, \qquad D'_{2t} = 0, \quad t = 1, \ldots, T.$$

Let D^* be an upper bound (polynomially bounded by the input size) on the maximum possible aggregated demand over the entire time horizon. We define the state space and action space as:

$$\mathcal{S}_{2t-1} = \mathcal{S}_{2t} = [-D^*, \ldots, D^*], \qquad \mathcal{A}_{2t-1}(I) = [0, \ldots, D^*], \qquad \mathcal{A}_{2t}(I) = \begin{cases} [-I, \ldots, 0] & \text{for } I > 0; \\ \{0\} & \text{otherwise.} \end{cases}$$

Note that the transformed problem satisfies Conditions 1-2. Indeed, Condition 1 is satisfied since both the state space and the action space are intervals of length at most $2D^*$, for every time period. Condition 2 is satisfied by the assumptions in the problem description.

As for Condition 3, we note that indeed $f_{2t-1}(\cdot, \cdot, D), f_{2t}(\cdot, \cdot, D)$ are nondecreasing in their first variable and monotone in their second variable as needed. Moreover, $g_{2t-1}(\cdot, \cdot, D)$ are nonnegative functions nonincreasing in their first variable and nondecreasing in their second variable. Furthermore, $\mathcal{A}_{2t-1}(I) = \mathcal{A}_{2t-1}(I')$ for all $I', I \in \mathcal{S}_{2t-1}$ with $I' < I$. However, $g_{2t}(I, x, D)$ are not necessarily monotone in their second variable and z_t is not necessarily nonincreasing. In order to satisfy Condition 3 we make additional assumptions as stated in the next subsections.

6.1 Disposal at no cost

In this section we deal with the case where disposal of inventory is free of charge.

Assumption 1 (Free disposal)**.** *Inventory can be disposed at no cost at any time period.*

Theorem 6.1. *The single-item stochastic lot-sizing problem under the free disposal assumption (Assumption 1) admits an FPTAS.*

Proof. Due to Theorem 5.1 and the discussion in the end of the section above, it suffices to prove that $g_{2t}(I, \cdot, D)$ are monotone and that the cost functions $z_t(\cdot)$ are nonincreasing. We note that by the free disposal assumption $d_t(\cdot) \equiv 0$. We also note that $h_t(I + x)$ is nondecreasing in $x \in \mathcal{A}_{2t}(I)$ for every fixed I. Therefore, $g_{2t}(I, \cdot, D)$ are nondecreasing.

We last prove by backward induction that $z_t(\cdot)$ are nonincreasing. The base case of $z_{2T+1}(I) = g_{2T+1}(I) = 0$ is trivial. The induction hypothesis is that $z_{2t+1}(\cdot)$ is nonincreasing. We distinguish between 3 cases: $I < 0$, $I > 0$ and $I = 0$. If $I < 0$ we have $\mathcal{A}_{2t} = \{0\}$ so:

$$z_{2t}(I) = b_t(-I) + z_{2t+1}(I) \geq b_t(-I-1) + z_{2t+1}(I+1) = z_{2t}(I+1),$$

where the inequality is due to the monotonicity of the backlogging cost function and the induction hypothesis. If, on the other hand, $I > 0$ then:

$$\begin{aligned} z_{2t}(I) &= \min\{h_t(I) + z_{2t+1}(I), \cdots, h_t(0) + z_{2t+1}(0)\} \\ &\geq \min\{h_t(I+1) + z_{2t+1}(I+1), h_t(I) + z_{2t+1}(I), \cdots, h_t(0) + z_{2t+1}(0)\} = z_{2t}(I+1). \end{aligned}$$

Last, if $I = 0$ then

$$z_{2t}(0) = z_{2t+1}(0) \geq \min\{h_t(1) + z_{2t+1}(1), z_{2t+1}(0)\} = z_{2t}(1).$$

Hence $z_{2t}(\cdot)$ is nonincreasing. It remains to show that $z_{2t-1}(\cdot)$ is nonincreasing as well:

$$\begin{aligned} z_{2t-1}(I) &= \min\{z_{2t}(I - D_t), c_t(1) + z_{2t}(I+1-D_t), \cdots, c_t(D^*) + z_{2t}(I + D^* - D_t)\} \\ &\geq \min\{z_{2t}(I+1-D_t), c_t(1) + z_{2t}(I+2-D_t), \cdots, c_t(D^*) + z_{2t}(I+1+D^*-D_t)\} = z_{2t-1}(I+1). \end{aligned}$$

The inequality is due to the monotonicity of z_{2t}. □

6.2 Disposal at a cost

When disposal of inventory incurs a cost, we will make the following three assumptions:

Assumption 2 (Bounded disposal cost). *There exists a positive constant κ such that for every $t = 1, \ldots, T$, and for every random vector $D^t = (D_t, \ldots, D_T)$ and $I \in \mathbb{Z}$, and for every feasible solution for the lot-sizing problem from time-period t onwards, starting with initial inventory I, the expected cost of that solution is at least κ times the total cost of disposing $E(D_t) - I, E(D_{t+1}), \ldots, E(D_T)$ units of inventory in time periods $t, t+1, \ldots, T$, respectively (either directly, or indirectly by holding it a few more time periods and then disposing it, i.e., the cost of disposing x units in time period t is $\min\{d_t(x), h_t(x) + d_{t+1}(x), h_t(x) + h_{t+1}(x) + d_{t+2}(x), \ldots, \sum_{i=t}^{T-1} h_i(x) + d_T(x)\}$).*

Assumption 3 (Linear holding and disposal costs). *For every time period, each of the holding, backlogging and disposal costs is linear.*

For each time period $t = 1, \ldots, T$, we denote the per-unit disposal cost by d_t, the per-unit holding cost by h_t, and the per-unit backlogging cost by b_t.
Note that this assumption allows the procurement cost functions to be nonlinear.

Assumption 4 (No backward disposal). *For every time period $t = 2, \ldots, T$, it is not beneficial to dispose of inventory in the previous time period, i.e., $b_{t-1} + d_{t-1} > \min\{d_t, h_t + d_{t+1}, h_t + h_{t+1} + d_{t+2}, \ldots, \sum_{i=t}^{T-1} h_i + d_T\}$.*

An easy special case where this assumption holds is when the per-unit disposal costs are nonincreasing with time (or even stationary).

We are now ready to give an FPTAS for SLS with disposal costs. The idea is to first transform the stochastic network flow minimization problem corresponding to the lot-sizing problem to be monotone by adding a constant C to any feasible solution. In this way the resulted transformed problem is a monotone DP, and therefore by the framework of [HKL⁺08] admits an FPTAS. We retrieve the value of the approximated original problem by subtracting C from the value of the transformed problem. By appropriate choices of C and K, and by using subtraction of approximation this value is at most $1 + \epsilon$ times the optimal value.

Theorem 6.2. *The single-item stochastic lot-sizing problem with bounded disposal cost, linear holding and disposal costs, and without backward disposal (Assumptions 2-4) admits an FPTAS.*

Proof. We first note that due to Assumption 3, $g_{2t}(I,x,d)$ are linear in $x \in \mathcal{A}_{2t}(I)$, for every fixed I and d, and therefore are monotone.

We next transform (17), the stochastic network flow minimization problem corresponding to the lot-sizing problem, in the following way. We assign to each vertex $t = 1, \dots, 2T$, a nonnegative number π_t that we call a *potential*. The assignment of potentials goes in backward as follows. We first set $\pi_{2T} = \pi_{2T-1} = d_T$. Assuming π_{2t+1} is determined, we set $\pi_{2t} = \pi_{2t-1} = \min\{h_t + \pi_{2t+1}, d_t\}$. We continue iterating until setting all potentials $\pi_{2T}, \dots, \pi_1$. (Note that all potentials assigned in this way are indeed nonnegative.) By its construction, the potential π_t is the minimal total cost of disposing a unit of inventory from vertex t via some vertex $j \geq t$ in the network, i.e.,

$$\pi_{2t} = \pi_{2t-1} = \min\left\{d_t, h_t + d_{t+1}, h_t + h_{t+1} + d_{t+2}, \dots, \sum_{i=t}^{T-1} h_i + d_T\right\}. \tag{18}$$

We set $\pi_0 = 0$ and $\Pi = (\pi_0, \dots, \pi_{2T})$.

We next change the cost functions as follows. We change the cost of flow in each edge $e_{i,i+1}$ to be $c'_{i,i+1}(x) \leftarrow c_{i,i+1}(x) + (\pi_{i+1} - \pi_i)x$. We change the cost of flow in each edge $e_{i+1,i}$ to be $c'_{i+1,i}(x) \leftarrow c_{i+1,i}(x) - (\pi_{i+1} - \pi_i)x$. We increase the cost of flow in edge $e_{0,2t-1}$ to $c'_{0,2t-1}(x) \leftarrow c_{0,2t-1}(x) + \pi_{2t-1}x$. Last, we decrease the cost of flow in each edge $e_{2t,0}$ to $c'_{2t,0}(x) \leftarrow c_{2t,0}(x) - \pi_{2t}x$. In this way the marginal cost of net inflow to vertex t is increased by π_t, while the marginal cost of net outflow from vertex t is decreased by π_t.

We now show that the DP formulation corresponding to the transformed stochastic network flow minimization problem is a nonincreasing DP. Since the state and access spaces remain the same as in the original SLS problem, Condition 1 is satisfied. As for Condition 2, the values of the transformed functions differ from the original ones by combinations of πs, so by (18) they remain polynomially bounded by the input size. It remains to show that the single-period cost functions, which are sums of costs of flows over edges are nonnegative. It is easy to verify that the cost of flow from vertex $2t-1$ to vertex $2t$ and vice versa is zero. The cost of flow of x units from vertex 0 (which has potential 0) to vertex $2t-1$ is $c_t(x) + \pi_{2t-1}x$. This cost is indeed nonnegative because $c_t(\cdot)$ is a nondecreasing nonnegative function and π_{2t-1} is a nonnegative number. The per-unit cost of flow from vertex $2t$ to vertex $2t+1$ is $h_t + \pi_{2t+1} - \pi_{2t}$, so by the recursive definition of the potentials it is nonnegative. The per-unit cost of flow from vertex $2t$ to vertex 0 is $d_t - \pi_{2t}$, which by the definition of the potentials is a nonnegative number. Lastly, the per-unit cost of flow from vertex $2t+1$ to vertex $2t$ is $b_t + (\pi_{2t} - \pi_{2t+1})$. If $\pi_{2t} = h_t + \pi_{2t+1}$ then this last term is nonnegative as well. Else $\pi_{2t} = d_t$, and by the no backward disposal assumption (Assumption 4) and (18) this term is again nonnegative.

It remains to show that Condition 3 is satisfied as well. Since the transition functions in the transformed problem are the same as in the original SLS problem, and since we showed above that the single-period cost functions are nonnegative nondecreasing functions in the amount of flow, it suffices to prove that the transformed problem is nonincreasing in the amount of inventory. It suffices to show that the choice of the potentials implies that for every vertex i, the cost of flow to either vertex 0 (if edge $e_{i,0}$ exists) or vertex $i+1$ is zero. This implies that the transformed problem is nonincreasing - the more inventory we have on hand - the less expenses we have to satisfy the demand. (A formal proof for this is via backward induction, similarly to the proof of Theorem 6.1.) Suppose first that $i = 2t-1$. Then the cost of flow of x units on edge $e_{2t-1,2t}$ is $c'_{2t-1,2t}(x) = 0 + (\pi_{2t} - \pi_{2t-1})x = 0$. If on the other hand $i = 2t$, then edge $e_{2t,0}$ exists and with cost $c'_{2t,0}(x) = d_t x - \pi_{2t}x$ per x units. If this cost is not zero, i.e., if $\pi_{2t} \neq d_t$, then we must have $\pi_{2t} = h_t + \pi_{2t+1}$, i.e., $\pi_{2t+1} - \pi_{2t} = -h_t$. But then the cost of flow of x units on edge $e_{2t,2t+1}$ is $c'_{2t,2t+1} = h_t x + (\pi_{2t+1} - \pi_{2t})x = 0$. To summarize, the transformed problem satisfies Conditions 1-3, and therefore is a nonincreasing DP. Due to Theorem 5.2, it admits an FPTAS.

Note that by the above transformation, for every random vector D, initial inventory level I and policy P, we get that

$$\begin{aligned} z'^*(I) = E_{D,P} \textstyle\sum_{(i,j)\in E} c'_{i,j}(x_{i,j}) &= \textstyle\sum_{\omega\in\Omega} \mathrm{Prob}(\omega) \sum_{(i,j)\in E} c'_{i,j}(x_{i,j}) \\ &= \textstyle\sum_{\omega\in\Omega} \mathrm{Prob}(\omega) [\sum_{(i,j)\in E} c_{i,j}(x_{i,j}|\omega) + (A(I)\Pi^T|\omega)] \\ &= E_{D,P} \textstyle\sum_{(i,j)\in E} c_{i,j}(x_{i,j}) + E_D A(I)\Pi^T = z^*(I) + E_D A(I)\Pi^T, \end{aligned}$$

where the second equality is from Property 2.4 on page 43 of [AMO93]. This means that the difference between $z'^*(I)$–i.e., the value of an optimal solution to the transformed problem–and $z^*(I)$–i.e., the value of an optimal solution of the original problem–is fixed to be $\Pi E_D A^T(I)$. So if we can find the optimal value of the transformed problem, then by subtracting from it $E_D A(I)\Pi^T$ we get the optimal value of the original problem. If we cannot efficiently compute the exact value $z'^*(I)$ of the transformed problem, by the discussion above we can design for it an FPTAS and K_1-approximate it for every $K_1 > 1$. By the bounded disposal cost assumption (Assumption 2), we get that $z^*(I) \geq \kappa E_D A(I)\Pi^T$, so $z'^*(I) = z^*(I) + E_D A(I)\Pi^T \geq (1+\kappa)E_D A(I)\Pi^T$. By subtraction of approximation from above (Proposition 3.10 applied with $c = 1/(1+\kappa), \varphi_1 = z'^*, \varphi_2 = E_D A(I)\Pi^t$ and $K_1 = 1 + \frac{\kappa\epsilon}{1+\kappa}$) we get a $(1+\epsilon)$-approximation for $\varphi_1 - \varphi_2 = z^*$. □

7 Extensions

7.1 Implicitly-described random variables

In this section we show how to deal with a more general setting of nonincreasing stochastic dynamic programming where the random variables are given implicitly by oracles to their CDFs. In this way we can handle distributions with support of exponential size (in the binary input size), such as truncated Poisson with given rate λ and upper bound M, or truncated discrete normal with parameters μ, σ and lower and upper bounds m, M.

Considering SLS with implicitly-described random variables $D_1, \ldots, D_T$, the DP recursion (16) specialized for this problem reads

$$z_{2t-1}(I) = \min_{x \geq 0}\{c_t(x) + E_{D_t} z_{2t}(I + x - D_t)\}, \tag{19}$$

and

$$z_{2t}(I) = \min_{x \leq 0}\{d_t(-x) + h_t(I + x) + z_{2t+1}(I + x)\}. \tag{20}$$

While (20) is a deterministic recursion which can be approximated directly via Theorem 3.3, (19) is a stochastic one. Since we assume the random variable is given implicitly by its CDF, the way the stochastic DP model studied in [HKL+08] calculates expectations does not apply. But we can bypass this difficulty by applying Corollary 3.6 in order to compute $\tilde{\xi}_{2t}$ that approximates $E_{D_t} z_{2t}(I + x - D_t)$. (Note that since we are given the CDF as an oracle function, we apply this proposition with $K_2 = 1$.) We then approximate (19) via Theorem 3.3 by setting $\tilde{\varphi}_2 = \tilde{\xi}_{2t}$, and iterate the recursion similarly to the way it is done in [HKL+08]. This gives us an FPTAS for SLS with implicitly-described random variables.

7.2 Positive lead times

Under general lead times, the value function of SLS (with explicitly-described random variables) is multivariate. It is well known that this dynamic program can be transformed into a single-variable dynamic program [Zip00] (the state corresponds to inventory position, which is defined as the inventory on-hand and all outstanding inventory). It is easy to show that this transformation preserves the approximation ratio and as a result it suffices to find an FPTAS for this single variate dynamic program. If $L > 0$ is an arbitrary lead time, then the underlying demand distribution of the transformed problem is $\bar{D}_t = \sum_{\hat{t}=t}^{t+L-1} D_{\hat{t}}$. The FPTAS in [HKM+09] requires that we know $\mathrm{Prob}[\bar{D}_t = \bar{d}_{t,i}]$, which is a convolution of L distributions. As a result, computing these probabilities takes $(n^*)^L$ time, where n^* is the maximal cardinality of the supports of the various D_i's. If L is 2 or 3 (or any other constant value), then the term $(n^*)^L$ is polynomial, and the algorithm is an FPTAS. If L is not constrained to be small (e.g., $L = T/4$), then the running time is exponentially large. In the latter case, the algorithm in [HKM+09] is not an FPTAS. An open question was raised in [HKM+09] whether one can modify the approach and create an FPTAS for the problem in which the lead times are permitted to be a fraction of T.

We give a positive answer to this question and design an FPTAS in the following way. For $0 \leq j \leq L$ and $1 \leq i \leq T-j$, let F_i^j be the CDF of the convolution of $D_i, \ldots, D_{i+j}$, i.e., $F_i^j(x) = \mathrm{Prob}(D_i + \cdots + D_{i+j} \leq x)$.

We compute F_i^j exactly for $j = 0, 1$ and $1 \le i \le T - j$. For $2 \le j \le L$ and $1 \le i \le T - j$ we build a K^{j-1}-approximation function $\tilde{F}_i^j$ for F_i^j via K-approximation sets (see Lemma 3.2 and Definition 3.4) in a recursive way by using the calculus of approximation and the equality

$$F_i^j(x) = P(D_i + \cdots + D_{i+j} \le x) = \sum_{y \le x \text{ and } y \text{ is in the support of } D_i} \mathrm{Prob}(D_i = y) F_{i+1}^{j-1}(x - y).$$

(Since CDF is a monotone function, a K-approximation set for it is well defined.) We then proceed as described in Section 7.1 with the only difference that instead of having oracles that compute the CDFs exactly, we use approximations, i.e., we apply Corollary 3.6 with $K_2 = K^{L-1}$).

7.3 Non-exact evaluation of CDF and cost functions

In the problem formulation we require that there exist oracles that compute the CDF and cost functions *exactly*. We can weaken this requirement as follows.

Assumption 5. *For every $\epsilon \ge 0$, there exist cost functions $\tilde{f}^\epsilon$ and CDF functions $\tilde{F}^\epsilon$ such that*

$$\frac{|\tilde{f}^\epsilon(x) - f(x)|}{f(x)} \le \epsilon, \quad \frac{|\tilde{F}^\epsilon(x) - F(x)|}{F(x)} \le \epsilon,$$

for every x, and these functions can be evaluated in polynomial time in the input size and $1/\epsilon$.

This assumption is equivalent to the statement that the cost functions and the CDF have an FPTAS. This assumption is useful when the population is divided into n subpopulations, each of which is provided with its CDF. This is true since the sum of n discrete distributions can be computed approximately. Also, any cost function that requires simulation can be computed approximately with high probability when the lowest and greatest non-zero probability is bounded away from 0 and 1. It can be shown that by performing minor modifications all the results presented in this paper hold under this assumption as well.

8 Conclusion and future research

In this paper we show that NNV requires exponential number of queries to solve and provide an FPTAS in the case that the profit-to-cost ratio is bounded away from 0. We can design FPTASs for variants of NNV in a similar way. For instance, when there is a penalty $p(\cdot)$ for lost sales, we will add $p((D - (x + y)))$ to the right hand side of equation (1). When there is a possibility for expedited ordering and shipping at cost $p(\cdot)$, we will add $p((D - (x + y)))$ to the right hand side of equation (1), and replace $r(\min(D, x + y))$ with $r(D)$.

Previous researchers also designed worst-case approximation algorithms for certain families of instances of otherwise inapproximable optimization problems (see, e.g., [AKK99, KPR04, FIMN09]). Kleinberg et *al.* [KPR04] study a novel genre of optimization problems that they call segmentation problems. They analyze a greedy algorithm for the variable catalogue segmentation problem when the number of catalogues is not set in advance, and show lower and upper bounds on the approximation ratio of the algorithm, which depends on the profit-to-cost ratio of the minimal-cost optimal solution of the specific instance. They also present a general greedy scheme, which can be specialized to approximate any segmentation problem. Feige *et al.* [FIMN09] introduce a framework for designing and analyzing algorithms. They design guarantees for classes of instances, parameterized according to properties of the optimal solution (which they call *signature* of the solution). They consider greedy algorithms as well as LP-based algorithms to derive approximation algorithms, some of which strictly improve over the previous results of Kleinberg *et al.* concerning the approximation ratio of the greedy algorithm.

Kleinberg *et al.* and Feige *et al.* deal with a constant-factor approximation without any guarantee on the profit-to-cost ratio of the approximated solution. Moreover, it can be shown that the greedy algorithms stated in their works may have an arbitrary-low profit-to-cost ratio. Similarly to Kleinberg *et al.* and Feige *et al.* we use the profit-to-cost ratio as a parameter in the analysis of the approximation ratio. But we deal with an arbitrarily-good approximation (FPTAS) that has an arbitrarily-good guarantee on the profit-to-cost ratio of the approximated solution. It may be of interest, in the context of the works of [KPR04, FIMN09], to develop 2-parameter (K, δ) algorithms that provide solutions that approximate the values of the optimal solutions of the problems they consider within a factor of K, and have profit-to-cost ratios of at least $1 - \delta$ times the profit-to-cost ratio of the minimal-cost optimal solutions.

In this paper we also extend previous results of Halman *et al.* [HKL$^+$08, HKM$^+$09, HLS09] in various ways. One of the assumptions they make in the analysis of their FPTASs is that the probability distribution function of each random variable D is given *explicitly* as a set of ordered pairs $(d, \text{Prob}(D = d))$. In this paper we show a way of using cumulative distribution functions (CDFs) instead of discrete distributions in the analysis, hence enabling us to handle random variables with support of size exponential in the (binary) size of their description. This also extends the framework of [HKL$^+$08] to models where the random variables are given implicitly (e.g., truncated Poisson with rate λ and upper bound M). We also relax the convexity assumption made by [HKM$^+$09]. This enables us to give FPTASs for SLS with a positive lead-time.

All the problems dealt by [HKM$^+$09, HKL$^+$08, HLS09] are either for minimizing costs, or maximizing revenues. All these works used some set of general computational rules of K-approximation functions, which [HKL$^+$08] called the calculus of K-approximation functions. If the objective is to maximize profit, i.e., the difference between revenues and costs, having a rule in the calculus of approximation that deals with subtraction is desirable. In this paper we extend the calculus of approximation to deal with subtraction of functions and use it to develop an FPTAS for NNV and SLS.

A natural extension of our models and the approach introduced in this paper is in the context of *revenue management,* where profit maximization is typically the objective. For example, in the stochastic inventory-pricing model (see Chapter 9 in [SCB05]), the objective is to coordinate inventory replenishment and pricing decisions so as to maximize expected profit. This extension is presented in a follow up work [HOS11].

References

[AHM51] K. Arrow, T. Harris, and J. Marschak. Optimal inventory policy. *Econometrica*, 19:250–272, 1951.

[AKK99] S. Arora, D. R. Karger, and M. Karpinski. Polynomial time approximation schemes for dense instances of np-hard problems. *J. Comput. Syst. Sci.*, 58(1):193–210, 1999.

[AMO93] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network Flows: Theory, Algorithms, and Applications.* Prentice Hall, 1993.

[Bel57] R. E. Bellman. *Dynamic Programming*. Princeton University Press, Princeton, NJ, 1957.

[CKP06] S. Chubanov, M.Y. Kovalyov, and E. Pesch. An FPTAS for a single-item capacitated economic lot-sizing problem with monotone cost structure. *Mathematical Programming*, 106:453–466, 2006.

[CNC10] M.Y. Kovalyov C.T. Ng and T.C.E. Cheng. A simple FPTAS for a single-item capacitated economic lot-sizing problem with a monotone cost structure. *European Journal of Operational Research*, 200:621–624, 2010.

[FIMN09] U. Feige, N. Immorlica, V. Mirrokni, and H. Nazerzadeh. Pass approximation: A framework for analyzing and designing heuristics. In *Proceedings of the 19th International Workshop Approximation Algorithms for Combinatorial Optimization (APPROX)*, pages 111–124, Berkeley, CA, 2009.

[FLR80] M. Florian, J.K. Lenstra, and A.H.G. Rinnooy Kan. Deterministic production planning: Algorithms and complexity. *Management Science*, 26:669–679, 1980.

[HKL$^+$08] N. Halman, D. Klabjan, C.-L. Li, J. Orlin, and D. Simchi-Levi. Fully polynomial time approximation schemes for stochastic dynamic programs. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 700–709, 2008.

[HKM$^+$09] N. Halman, D. Klabjan, M. Mostagir, J. Orlin, and D. Simchi-Levi. A fully polynomial time approximation scheme for single-item stochastic inventory control with discrete demand. *Mathematics of Operations Research*, 34:674–685, 2009.

[HLS09] N. Halman, C.-L. Li, and D. Simchi-Levi. Fully polynomial time approximation schemes for time-cost tradeoff problems in series-parallel project networks. *Operations Research Letters*, 37:239–244, 2009.

[HOS11] N. Halman, J. Orlin, and D. Simchi-Levi. Approximating the nonlinear newsvendor problem with joint price and order size decisions. Working paper, Massachusetts Institute of Technology, 2011.

[KPR04] J. Kleinberg, C. Papadimitriou, and P. Raghavan. Segmentation problems. *Journal of the ACM*, 51:263–280, 2004.

[LPRS07] R. Levi, M. Pal, R.O. Roundy, and D.B. Shmoys. Approximation algorithms for stochastic inventory control models. *Mathematics of Operations Research*, 32:284–302, 2007.

[LRST08] R. Levi, R.O. Roundy, D.B. Shmoys, and V.A. Truong. Approximation algorithms for capacitated stochastic inventory control models. *Operations Research*, 56:1184–1199, 2008.

[Por90] E. Porteus. Stochastic inventory theory. In D.P. Heyman and M.J. Sobel, editors, *Handbook in OR & MS*, volume 2. Elsevier Science Publishers B.V., North-Holland, 1990.

[Por02] E.L. Porteus. *Foundations of Stochastic Inventory Theory*. Stanford Business Books, Stanford, CA, 2002.

[SCB05] D. Simchi-Levi, X. Chen, and J. Bramel. *The Logic of Logistics: Theory, Algorithms, and Applications for Logistics and Supply Chain Management*. Springer-Verlag, NY, NY, second edition, 2005.

[Zip00] P.H. Zipkin. *Foundations of Inventory Management*. McGraw-Hill, Boston, 2000.