Sandwich functions for the lot-sizing problem

Master thesis Operations Research and Quantitative Logistics, Erasmus School of Economics, Erasmus University Rotterdam

Student: Adriaan L. Molendijk (409865)

Supervisor: W. van den Heuvel

Second reader: A.P.M. Wagelmans

July 2018

Abstract

In this thesis we study the use of so-called sandwich functions for lot-sizing problems. A sandwich function is a function that replaces the objective function of an optimisation problem in such a way that the original objective function is bounded by this function and a scalar multiple of this function from below and above, respectively. We start by pointing out some places in the literature where this method has been used before. We then provide an analysis of sandwich functions for two well known cost functions, the modified all-unit discount cost function and the stepwise cost function. Finally, several applications of this method to existing lot-sizing problems are presented. Amongst these applications is a 2-approximation algorithm for the lot-sizing problem with demand time windows and stepwise cost, which is strongly $\mathcal{NP}$-hard when order splitting is not allowed. To the best of our knowledge, this is the first constant factor approximation algorithm for this problem.

Preface

What you are about to read is a project submitted in partial fulfillment of the requirements for the degree of Master of Science, provided that the grade g of this project is such that $\sqrt{g - 5.5}$ exists over the real numbers.

This project is my Master thesis, which started a couple of months ago. Among the choices of available thesis topics then, I noticed one which discussed *sandwich functions for lot-sizing problems*. I thought: sandwiches! That must be great, aren't they delicious?

Anyway, I would like to take the opportunity to thank Wilco van den Heuvel for being my supervisor throughout the process. I am grateful for the meetings we had, my productivity increased infinitely after those. Moreover, a special thanks to Albert P.M. Wagelmans for being the second reader of this project. Besides the gentlemen mentioned above, the presence of the coffee machine at the Erasmus University Rotterdam also helped a bit.

A comment about this document is in order. You might notice that in the document, I refer to 'we' whenever I make a statement. This is not because the author of this document is socially awkward or anything. At least, that was not the case before undertaking this project. It is because I would like to include you, the reader, in the thinking process. To this same reader, Congratulations! You almost managed to read at least one page of this report. I challenge you to read the rest...

Just like this awesome sandwich project, all things must come to an end. The four years in Rotterdam were an amazing experience, partly due to all my friends, family and God. I can't wait to continue my journey next year in England, where I will study Applicable Mathematics at the London School of Economics.

Contents

1 Introduction

In the classical lot-sizing problem (LSP) we are given a discrete and finite time horizon, with known demands for each period. The problem is to determine when and how much to produce each time period, such that the demand is met in time and the cost of production and holding inventory over time is minimised.

The problem was introduced in Wagner and Whitin (1958), and has seen many extensions ever since. Extensions of the LSP include production capacities, generalised cost structures, backlogging, lost sales, perishability, stochastic demand, multi-item or multi-mode considerations, and many more. One extension that will be central throughout this thesis is the transportation decision that is part of the production process. The importance of including inbound transportation cost cannot be overstated, see Carter and Ferrin (1996), who provide a wide range of companies and numerical examples to illustrate this. We will look at two common discount structures in detail, these being the Full Truck Load (FTL) structure and Less Than Truck Load (LTL) structure. The FTL structure is common in situations where costs are paid per shipped batch or container, while the LTL structure is common in situations where the shipment sizes are less than truck capacity.

Although the classical lot-sizing problem is considered a relatively easy problem, many of its extensions are not. In particular, while the LSP is solvable in polynomial time, using various cost structures makes the problem $\mathcal{NP}$-hard in some scenarios. The intractibility of these problems implies that in some cases we have to resort to inexact approaches like approximation algorithms or heuristics.

Amongst these approximation algorithms are construction approaches (see Van Den Heuvel and Wagelmans (2009)), primal dual methods (Levi et al. (2006)) and LP-based rounding methods (Levi et al. (2008)). Remarkably, very recently Akbalik and Rapine (2017) suggest an approach where the objective function of their problem is "sandwiched" by an alternative cost function.

When we say “sandwiched”, we mean that the original objective function is bounded by this alternative function and a scalar multiple of this function from below and above, respectively. They then show that they obtain an approximation algorithm for their problem, because of the sandwich relationship between the two cost functions and the fact that the problem with the alternative cost function is much easier to solve. We will refer to this alternative function as the sandwich function and studying the method of using sandwich functions will be the central topic of this thesis. To be more specific, this thesis is dedicated to 1) studying existing sandwich approaches in the lot-sizing literature 2) studying the use of non-linear sandwich functions, in particular the effect of introducing discontinuities in the sandwich function 3) applying this sandwich approach to extensions of the LSP. We will now state the results that can be found in this thesis.

- We apply the concept of sandwich functions to the modified all-unit discount function. In particular, we show that no linear sandwich function for this function exists. This result can be modified in such a way that it states that an existing linear approximation from Hill and Galbreth (2008) can perform arbitrarily bad. We also revisit the upper concave envelope sandwich approach by Chan et al. (2002a). We elaborate on their work by providing the complexity of the construction of the upper concave envelope and a more sophisticated bound on the approximation guarantee. We also present limitations on the introduction of discontinuities in this sandwich function.

- We apply the concept of sandwich functions to the stepwise cost function. We show that we can sandwich this cost function arbitrarily close, using only a finite number of discontinuities. We conclude that for certain sets of problems, sandwich functions can lead to polynomial time approximation schemes (PTAS).

- We present a 2-approximation algorithm for the lot-sizing problem with demand time windows and stepwise cost, which is $\mathcal{NP}$-hard in the strong sense when order splitting is not allowed. To the best of our knowledge, this is the first constant factor approximation algorithm for this problem.

· We present an application of sandwich functions to mathematical programming formulations. We show that some formulations are related to the formulation that is obtained when replacing the objective function by its sandwich function. This gives the opportunity to model sandwich functions that cannot be used directly in a dynamic programming environment.

This thesis is organised as follows. We start with some notation in Section 1.1. In Section 2, we will present a literature review on the lot-sizing problem and on the concept of sandwich functions, especially on its applications to the LSP. The formal definition and properties of sandwich functions are presented in Section 3. Two common cost functions, the stepwise cost function and the modified all-unit discount cost function, which represent FTL and LTL cost structures respectively, are discussed in Section 4. Then, in Section 5, we present applications of the sandwich function approach to the LSP. A conclusion is provided in Section 6.

1.1 Notation

For a function $f : \mathbb{R} \to \mathbb{R}$ we write $\Delta f(a, b)$ for the slope between the points $(a, f(a))$ and $(b, f(b))$ in the Euclidean plane, assuming $a < b$. We use the shorthand notation (x_i) for the list of numbers $(x_1, x_2, \ldots, x_n)$. We let $\mathbb{N}_{\geq 1} = \mathbb{N} \backslash \{0\}$ be the set of positive integers. The notation $\lceil x \rceil$ is used for the smallest integer $n \geqslant x$. We write $x \sim [a, b], a < b$, to indicate that x is chosen uniformly from the integers $\{a, \ldots, b\}$. Finally, we write $x = 0^+$ for $x \to 0^+$ and $f(0^+) = \lim_{x \to 0^+} f(x)$.

2 Literature review

In this section we will present a literature review on both the lot-sizing problem and the concept of sandwich functions.

2.1 The lot-sizing problem

Since the seminal paper by Wagner and Whitin (1958), a lot of research has been done on the classical lot-sizing problem, and many extensions have been proposed since then. The LSP is well known to be polynomially solvable (e.g. Wagelmans et al. (1992), Federgruen and Tzur (1991)), but some extensions may not be necessarily. We present some extensions of the LSP that are the most closely related to this research. For other extensions or a more detailed state of the art review of the LSP we refer to the recent surveys Brahimi et al. (2006) and Brahimi et al. (2017).

The lot-sizing problem with modified all-unit discount cost structure (LSP-M) has been introduced primarily in Chan et al. (2002a). Their main result is the complexity result that this problem is $\mathcal{NP}$-hard when the cost function is time dependent or the number of breakpoints is part of the input. They also present a 4/3-approximation algorithm for this problem running in $\mathcal{O}(T^2)$, and show that the approximation guarantee reduces to 5.6/4.6 whenever the cost function is stationary over time. Here, T denotes the length of the time horizon. They leave the complexity of the problem where the cost function is time-invariant and the breakpoints are stationary as an open question. This question is later answered in Li et al. (2012), who show that the problem is solvable in $\mathcal{O}(T^{m+3})$, where m is the number of price breakpoints. This result is later generalised by Ou (2017a), who presents a $\mathcal{O}(T^{m+2}\log T)$ algorithm for the case of piecewise linear production cost, with m the number of breakpoints in this cost function.

The LSP-M has also been analysed in a more general setting. In particular, only shortly after the paper Chan et al. (2002a), Chan et al. (2002b) extend their methodology by considering the one-warehouse multiretailer problem (OWMR) with a modified all-unit discount cost structure. They show that the set cover problem (see Feige (1998)) reduces to this problem, hence it cannot be approximated within a factor $\gamma \log T$ in polynomial time unless $\mathcal{P} = \mathcal{NP}$, for any $\gamma > 0$. They also show that the optimal zero inventory policy has a cost of at most 4/3 from the optimal solution, but finding this policy is an $\mathcal{NP}$-hard problem itself. Later, Hill and Galbreth (2008) propose a heuristic for the OWMR where the modified all-unit discount functions are approximated with linear functions. We show that these linear approximations can perform arbitrarily poor.

The lot-sizing problem with batch procurement (LSP-B) has been researched for quite some time, see for example Lippman (1969), Pochet and Wolsey (1993) or Li et al. (2004). The problem is polynomially solvable, in $\mathcal{O}(T^6)$ time to be exact, whenever capacities are considered (Akbalik and Rapine (2012)). Yet, a thorough analysis of the complexity of the LSP-B where the batches are time dependent (LSP-B$_t$) was performed only recently by Akbalik and Rapine (2013). In their paper, they show that the LSP-B$_t$ is $\mathcal{NP}$-hard whenever any of the cost parameters is nonstationary over time. A couple of years later, Akbalik and Rapine (2017) show that the lot-sizing problem with multi-mode replenishment and batches cost (LSP-MMB) is a special case of the LSP-B$_t$. It can then be shown that this problem has $\mathcal{NP}$-hard complexity too, even when the time horizon is restricted to a single period.

The lot-sizing problem with demand time windows (LSP-TW), first introduced in Lee et al. (2001), is also discussed in this thesis. In the LSP-TW, every demand order has a set of consecutive time periods, called its demand time window, within it can be satisfied without incurring holding or backlogging cost. Lee et al. (2001) present two algorithms to solve this problem. First an $\mathcal{O}(T^2)$ algorithm for the case without backlogging and next an $\mathcal{O}(T^3)$ algorithm for the case with backlogging, both under a nonspeculative cost structure. Hwang (2007) shows that the latter problem can be solved in $\mathcal{O}(\max\{T^2, nT\})$ time, where n is the number of different demands in the time horizon. In a later paper, Jaruphongsa and Lee (2008) consider the LSP-TW where the production cost function is given by the stepwise cost function (LSP-TWS). They show that the problem is strongly $\mathcal{NP}$-hard when split delivery is not allowed, by reducing the 3-partition problem to it. They go on to present polynomial time algorithms for the case when split delivery is allowed. We show that the LSP-TWS admits a 2-approximation algorithm running in $\mathcal{O}(T^2)$ time, by sandwiching the stepwise function by a linear sandwich function.

Finally, most of the extensions of the classical lot-sizing problem fall under a specific category, which is the class of lot-sizing problem with general cost structures. In Van Hoesel and Wagelmans (2001), a fully polynomial time approximation scheme (FPTAS) is developed for the capacitated LSP with a concave cost structure. A more general FPTAS is presented later in Chubanov et al. (2006), only requiring monotonicity of the cost structure. We can therefore conclude that many extensions of the LSP like the LSP-M and the LSP-B_t admit FPTASs. However, these algorithms are often more of theoretical interest, as their running time makes them prohibitive to be used in practice.

2.2 Sandwich functions

In this section we discuss the use of sandwich functions in the lot-sizing literature. We will start with the history of sandwich functions in mathematical programming in general, and next look at the applications of this concept to the LSP. We should point out that, to the best of our knowledge, no literature review on the concept of sandwich functions in the lot-sizing literature exists yet. We have tried to find all the instances in the literature where this technique has been used and include them in the literature review below. In none of the found instances there is a reference to an earlier application of this concept of sandwich functions.

The first occurrence of sandwich functions in mathematical programming seems to be in the form of additive sandwich functions. These are sandwich functions for which the deviation from the original cost function is measured in an absolute, instead of a relative way. Geoffrion (1977) claims that the idea of sandwiching functions is a practice that already exists (e.g. in non-linear optimisation), but that his paper is the first to consider it in a mathematical programming setting. In fact, he points out the application of sandwich functions in separable programming. This application was later analysed in Meyer (1979).

We will now present several occasions where sandwich functions have been used to construct approximation algorithms for extensions of the lot-sizing problem. One place where sandwich functions have been used in an LSP setting is in Chan et al. (2002a). Recall that in this paper they present a 4/3-approximation algorithm for the LSP-M. However, as an alternative solution they consider the upper concave envelope as a sandwich function for the modified all-unit discount cost function. They show that this function is a 2-sandwich function for the original cost function. We elaborate on their result with a couple of findings. First, we show that no linear sandwich function exists for the modified all-unit discount cost function. Secondly, we present the complexity of the construction of the upper concave envelope of the modified all-unit discount cost function, which is not mentioned in the paper of Chan et al. (2002a). Thirdly, we show that their bound of 2 can be refined and finally, we show that even when we introduce a finite number of discontinuities to this sandwich function (independent of the input), the approximation guarantee of the sandwich function does not improve.

Another example of where sandwich functions have found their use in the lot-sizing literature is in Hu (2016). In Chapter 4 of this doctoral thesis, many sandwich functions are suggested, amongst which are a 2-sandwich function for the stepwise function and a (2γ)-sandwich function for regular modified all-unit discount cost function, where this sandwich function can only be constructed under certain conditions and $\gamma \geqslant 1$ is data dependent. Note that the regular modified all-unit discount cost function, introduced in Archetti et al. (2014), is a special case of the modified all-unit discount cost function where the sections of the cost function with positive slope have identical length and the flat sections have identical length too. Hence this setting is more restricted than the general modified all-unit discount cost setting considered in this thesis.

Another instance where sandwich functions have been used, as mentioned earlier, is in Akbalik and Rapine (2017). Besides showing that the LSP-MMB was a special case of the LSP-B$_t$, they also presented a 2-sandwich function for the stepwise cost function. We elaborate on their result by showing that when we allow discontinuities in the sandwich function, the approximation guarantee improves.

The linear sandwich function for the stepwise cost function can also be used for extensions of the LSP-B. In Goisque (2017), they show that this sandwich approach can be used to construct a 2-approximation algorithm for the multi-level lot-sizing problem with batch deliveries.

3 The sandwich function

In this section we introduce the concept of a β-sandwich function. Consider the optimisation problem

$$\text{(P)} \quad \min \quad f(x)$$
$$\text{s.t.} \quad x \in X,$$

where X is a nonempty set of feasible solutions and $f(x)$ is an objective function. We assume that problem (P) is bounded and has optimal solution value z_{P}. It could be that (P) is difficult to solve, due to for example a complicated objective function $f(x)$. A natural solution would then be to replace $f(x)$ by a much simpler function, say $r(x)$, that approximates $f(x)$ sufficiently. This gives rise to another optimisation problem, namely

$$\text{(R)} \quad \min \quad r(x)$$
$$\text{s.t.} \quad x \in X.$$

If problem (R) is much easier to solve, one might consider solving this problem instead of (P). If, in addition, $r(x)$ does not deviate to much from the original objective $f(x)$, we can guarantee that the value of the solution obtained from solving (R) does not deviate too much from the value z_{P}. We formalise this concept as follows.

Definition 3.1. A function $r(x)$ is called a *β-sandwich function* of $f(x)$ if $f(x) \leqslant r(x) \leqslant \beta f(x)$ for all feasible $x \in X$, for some $\beta > 1$. We call β the *approximation guarantee*. Morerover, the approximation guarantee is said to be *tight* if there are $x_1, x_2 \in X$ such that $f(x_1) = r(x_1)$ and $r(x_2) = \beta f(x_2)$.

If the approximation guarantee is clear from the context we might refer to $r(x)$ simply as a sandwich function of $f(x)$, and drop the β prefix. Also, we almost always assume or show that the approximation guarantee of a sandwich function is tight, else we can introduce another sandwich function that has an approximation guarantee with this property.

Recall that an α-approximation algorithm for a problem (P), with $\alpha \geqslant 1$, is an algorithm that provides a solution with cost at most αz_{P} and runs in polynomial time in terms of the input of the problem. The following property is a direct consequence of the definition of a β-sandwich function for a problem (P).

Property 3.1. If $f(x)$ has a β-sandwich function $r(x)$ that can be constructed in polynomial time and problem (R) admits an α-approximation algorithm, then (P) admits an $\alpha\beta$-approximation algorithm.

Proof. Let x^* be the optimal solution to problem (P) with value z_{P} and let z_{R} be the value of the optimal solution to (R). Suppose that running the approximation algorithm on problem (R) gives solution $\tilde{x}$. Because the feasible region remained unchanged when sandwiching $f(x)$ by $g(x)$, $\tilde{x}$ is also feasible to (P). Hence, $f(\tilde{x}) \leqslant r(\tilde{x}) \leqslant \alpha z_{\mathrm{R}} \leqslant \alpha r(x^*) \leqslant \alpha\beta z_{\mathrm{P}}$, where the first inequality follows from the sandwich function $r(x)$, the second inequality from the approximation algorithm for (R), the third from definition of z_{R} and the fourth from the sandwich function $r(x)$ again. Because $r(x)$ can be constructed in polynomial time and the approximation algorithm runs in polynomial time, the solution to (P) can be obtained in polynomial time. □

Note that as soon as we can solve (R) in polynomial time to optimality, i.e. $\alpha = 1$, the derivation above reduces to $f(\tilde{x}) \leqslant r(\tilde{x}) \leqslant r(x^*) \leqslant \beta z_{\mathrm{P}}$, which means that we obtained a β-approximation algorithm for (P).

In the discussion above we restricted ourselves to sandwiching a single function $f(x)$ by a single function $g(x)$. Two important properties that we will use are the following, which state that we can also sandwich a function $f(x)$ that is either separable or defined piecewise.

Property 3.2. (Separable sandwich function) If $r_i(x)$ is a β_i-sandwich function of $f_i(x)$ for $i = 1, \ldots, n$, then $r(x) = \sum_{i=1}^n r_i(x)$ is a β-sandwich function of $f(x) = \sum_{i=1}^n f_i(x)$, where $\beta = \max_{i=1,\ldots,n} \beta_i$.

For the following property, let $0 \leqslant x_0 < x_1 < \cdots < x_n$.

Property 3.3. (Piecewise defined sandwich function) If $r_i(x)$ is a β_i-sandwich function of $f(x)$ on the interval $[x_i, x_{i+1}]$, for $i = 0, \ldots, n-1$, then the piecewise defined function $r(x) = \{r_i(x)\}$ is a β-sandwich function of $f(x)$ on $[x_0, x_n]$, where $\beta = \max_{i=1,\ldots,n} \beta_i$.

We end this section by mentioning that most of these concepts can be extended to include more general situations one might encounter when modelling sandwich functions. For example, Property 3.2 can be extended to linear combinations of functions, Property 3.3 can be extended to any collection of disjoint intervals, and the concept of tightness of the approximation guarantee of a sandwich function can be refined by saying that for all $\varepsilon > 0$ there exists $x_1, x_2 \in X$ such that $r(x_1) - f(x_1) < \varepsilon$ and $\beta f(x_2) - r(x_2) < \varepsilon$. Also, the sandwich function only has to actually sandwich the objective function over the feasible region $X' \subset X$, as long as $x^*, \tilde{x} \in X'$. Property 3.1 still applies in this case.

4 Common cost functions

In this section we analyse sandwich functions for two common cost functions, the modified all-unit discount function (Section 4.1) and the stepwise cost function (Section 4.2). The former is often used to model an LTL cost structure, while the latter is used to model an FTL cost structure.

4.1 The modified all-unit discount cost function

In this section we will use the notation from Chan et al. (2002a) on the input parameters of the modified all-unit discount cost function. The modified all-unit discount cost function $f(x)$ takes as an input a list of n constants (α_i) and breakpoints (M_i), and is given by

$$f(x) = \begin{cases} 0 & \text{if } x = 0 \\ \alpha_1 M_1 & \text{if } 0 < x \leqslant M_1 \\ \alpha_1 x & \text{if } M_1 \leqslant x \leqslant M_1' \\ \alpha_2 M_2 & \text{if } M_1' \leqslant x \leqslant M_2 \\ \alpha_2 x & \text{if } M_2 \leqslant x \leqslant M_2' \\ \vdots & \\ \alpha_n M_n & \text{if } M_{n-1}' \leqslant x \leqslant M_n \\ \alpha_n x & \text{if } M_n \leqslant x. \end{cases}$$

This input (α_i) and (M_i) has to satisfy three conditions: (i) $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n > 0$ (ii) $M_n > \cdots > M_2 > M_1 > 0$ (iii) $\alpha_{i+1} M_{i+1} > \alpha_i M_i$ for $i = 1, \ldots, n-1$. The breakpoints (M_i') are calculated using the identity $M_i' = \alpha_{i+1} M_{i+1} / \alpha_i$, for $i = 1, \ldots, n-1$. We let $M_n' = \sum_t d_t$ be an artificial breakpoint, representing an upper bound on the total demand to be satisfied. We will assume w.l.o.g. that $M_n' > M_n$, extending the last input condition.

The modified all-unit discount function $f(x)$ is an extension of the all-unit discount cost function $\tilde{f}(x)$. The all-unit discount function is given by

$$\tilde{f}(x) = \begin{cases} 0 & \text{if } x = 0 \\ \alpha_1 M_1 & \text{if } 0 < x \leqslant M_1 \\ \alpha_1 x & \text{if } M_1 \leqslant x < M_2 \\ \alpha_2 x & \text{if } M_2 \leqslant x < M_3 \\ \vdots & \\ \alpha_n x & \text{if } M_n \leqslant x. \end{cases}$$

The all-unit discount function is used in situations where a supplier encourages large shipments, and offers all-unit discounts at a specific set of shipment sizes. Chan et al. (2002a) mention that in practice, the cost of the order size of the shipper is calculated using $\min\{\alpha_i x, \alpha_{i+1} M_{i+1}\}$ whenever $M_i \leqslant x < M_{i+1}$. This means that after a certain threshold, M_i' to be exact, the shipper orders x units but pays as if they were shipping M_{i+1} units. In the industry this is called shipping x but declaring M_{i+1}. When modelling this cost structure, we obtain the modified all-unit discount cost function. For this function, the term $\alpha_1 M_1$ acts as a minimum charge for a small shipping volume. The all-unit discount cost function is shown in Figure 1 and the modified all-unit discount cost function is shown in Figure 2.

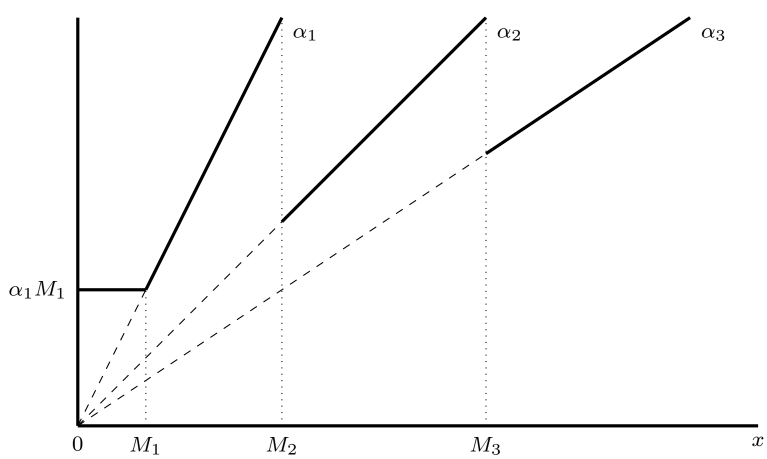

Figure 1: The all-unit discount cost function for $n = 3$.

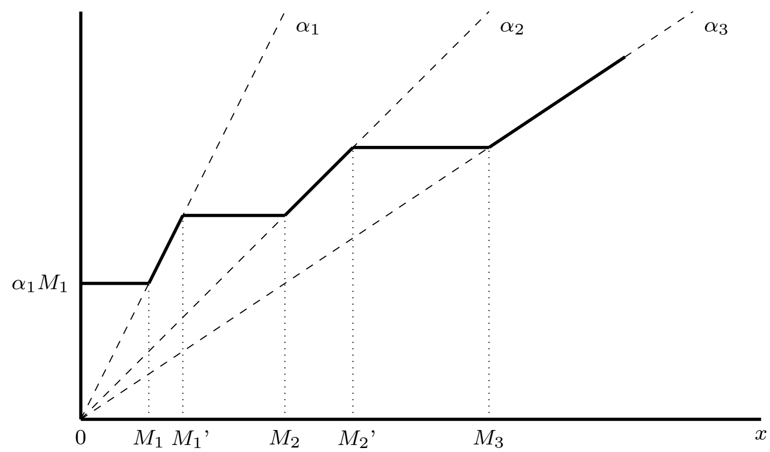

Figure 2: The modified all-unit discount cost function for $n = 3$.

We will now study the modified all-unit discount cost function from a sandwiching point of view. We will first show that no linear sandwich function for the modified all-unit discount cost function exists.

Theorem 4.1. For any $\beta > 1$, there exists an input (α_i) and (M_i) to the modified all-unit discount cost function $f(x)$ for which any linear function satisfying $g(x) \geqslant f(x)$ for $x \geqslant 0$ has $\frac{g(x_1)}{f(x_1)} > \beta$ for some $x_1 \geqslant 0$.

Proof. For simplicity, consider $n = 2$ breakpoints. If we let $\alpha_1 = 2\beta$, $M_1 = 1/(4\beta^2)$, $\alpha_2 = M_2 = 1$, then the input is valid because it satisfies input conditions (i) - (iii). We will prove this theorem by showing that either $x_1 = 0^+$ or $x_1 = M_2$ for this class of modified all-unit discount functions. Suppose a function $g(x) = a + bx$ exists for which $g(x) \geqslant f(x)$ for all $x \geqslant 0$. This holds in particular for $x = 0^+$, from which we obtain $\alpha_1 M_1 \leqslant a$. If $g(0^+) > \beta f(0^+)$ then we are done by setting $x_1 = 0^+$, else $a < \beta\alpha_1 M_1$. We also know that $g(x) \geqslant f(x)$ for $x = M_1'$, from which we obtain $g(M_1') = a + bM_1' \geqslant f(M_1') = \alpha_2 M_2$. Hence $b \geqslant (\alpha_2 M_2 - a)/M_1' \geqslant (\alpha_2 M_2 - \beta\alpha_1 M_1)/M_1' = (\alpha_2 M_2 - \beta\alpha_1 M_1)/(\alpha_2 M_2/\alpha_1)$. Now for $x_1 = M_2$ we get $g(x_1) = a + M_2 b > M_2 b \geqslant M_2(\alpha_2 M_2 - \beta\alpha_1 M_1)/(\alpha_2 M_2/\alpha_1) = (1 - \beta(2\beta)(1/(4\beta^2)))/(1/2\beta) = \beta = \beta\alpha_2 M_2 = \beta f(x_1)$. $\square$

Simply put, there does not exist a linear sandwich function for the modified all-unit discount cost function, because the relative deviation from any linear function that lies above the modified all-unit discount function to this function is not bounded by a constant. We will now consider the use of piecewise linear concave sandwich functions. We start with the construction of such a function.

Lemma 4.1. We can construct the upper concave envelope of the modified all-unit discount cost function $f(x)$ in $\mathcal{O}(n)$ time.

Once we notice that the upper concave envelope of $f(x)$ is the upper convex hull of the points $(M_i', f(M_i'))$, the work simplifies rather much. This is because the breakpoints are trivially sorted (from a lexicographical ordering point of view), hence we can use an existing convex hull algorithm to construct this function, see Andrew (1979).

Now that we have constructed the upper concave envelope, we would like to see if this is in fact a β-sandwich function of $f(x)$, and possibly derive an expression for β. We will do this with the help of the following lemma.

Lemma 4.2. Consider the interval $[\underline{M}, M] \cup [M, \overline{M}]$, $\underline{M} < M < \overline{M}$. Let $f(x)$ be piecewise defined as $f(x) = \alpha M$ for $x \in [\underline{M}, M]$ and $f(x) = \alpha x$ for $x \in [M, \overline{M}]$, and let $r(x) = \alpha M + \alpha_g(x - \underline{M})$, with $\alpha_g \leqslant \alpha$, be the line through the points $(\underline{M}, \alpha M)$ and $(\overline{M}, \alpha\overline{M})$. Then $r(x)$ is a β-sandwich function of $f(x)$ with $\beta = 1 + \frac{\alpha_g(M - \underline{M})}{\alpha M}$. Moreover, this approximation guarantee is tight.

Proof. Clearly we have that $r(x) \geqslant f(x)$ for all $x \in [\underline{M}, \overline{M}]$. One key observation is that the ratio $r(x)/f(x)$ is increasing on the interval $[\underline{M}, M]$ and decreasing on the interval $[M, \overline{M}]$. Hence, $\beta \leqslant r(M)/f(M) = (\alpha M + \alpha_g(M - \underline{M}))/(\alpha M) = 1 + (\alpha_g/\alpha)(M - \underline{M})/M$. Tightness of the approximation guarantee trivially follows from $r(\underline{M}) = f(\underline{M})$ and $r(M) = \beta f(M)$. □

We will now derive the approximation guarantee of the piecewise linear concave sandwich function $r(x)$ of $f(x)$. We assume that $r(x)$ has $k \leqslant n$ pieces, where each piece $r_i(x)$ of $r(x)$ is linear. We denote the approximation guarantee of $r_i(x)$ for that piece of the original function by β_i. We start by bounding the approximation guarantee β_i of the piece $r_i(x)$.

Lemma 4.3. Let $r(x)$ be the upper concave envelope of the modified all-unit discount cost function $f(x)$. For each piece $r_i(x)$ of $r(x)$, $i = 1, \ldots, k$, it holds that $\beta_i \leqslant 2 - \frac{M_1}{\sum_t d_t}$.

Proof. Let $r_i(x)$ be a linear piece of the upper concave envelope $r(x)$, with approximation guarantee β_i. We assume that the piece is defined over the interval $[\underline{M}, M] \cup [M, M'_j]$, $\underline{M} < M < M'_j$, where M'_j is one the breakpoints from (M'_i). We assume that $\underline{M} > 0$, but the case of $\underline{M} = 0$ is similar. Notice that the slope of the function at M'_j is equal to α_j. We introduce the new function $\tilde{f}(x)$ on the interval $[\underline{M}, M] \cup [M, M'_j]$, which is defined piecewise as $\tilde{f}(x) = f(\underline{M})$ and $\tilde{f}(x) = \alpha_j x$ on those two intervals. An important observation is that $\tilde{f}(x) \leqslant f(x) \leqslant r_i(x)$ on this interval, and $r_i(x) = f(\underline{M}) + \alpha_g(x - \underline{M})$. Hence, if $r_i(x)$ β-sandwiches $\tilde{f}(x)$, then it β-sandwiches $f(x)$ also. By Lemma 4.2 we know that it does, and $\beta_i = 1 + (\alpha_g/\alpha_j)(M - \underline{M})/M$. We will bound α_g, and show that it leads to the desired result.

One observation on α_g is that it is defined as the slope between the points $(\underline{M}, f(\underline{M}))$ and $(M'_j, \alpha_j M'_j)$. Since $M'_j \leqslant \sum_t d_t$, α_g is smaller or equal to the slope between the points $(\underline{M}, f(\underline{M}))$ and $(\sum_t d_t, \alpha_j \sum_t d_t)$. We will use this last slope to derive the bound. We obtain

$$
\begin{aligned}
\beta_i &= 1 + \frac{\alpha_g}{\alpha_j}\left(\frac{M - \underline{M}}{M}\right) \\
&\leqslant 1 + \frac{1}{\alpha_j}\left(\frac{\alpha_j \sum_t d_t - f(\underline{M})}{\sum_t d_t - \underline{M}}\right)\left(\frac{M - \underline{M}}{M}\right) \\
&\leqslant 1 + \frac{1}{\alpha_j}\left(\frac{\alpha_j \sum_t d_t - f(\underline{M})}{\sum_t d_t - \underline{M}}\right)\left(\frac{\sum_t d_t - \underline{M}}{\sum_t d_t}\right) \\
&= 1 + \frac{1}{\alpha_j}\left(\frac{\alpha_j \sum_t d_t - f(\underline{M})}{\sum_t d_t}\right) \\
&\leqslant 1 + \frac{1}{\alpha_j}\left(\frac{\alpha_j \sum_t d_t - \alpha_1 M_1}{\sum_t d_t}\right)
\end{aligned}
$$

$$= 2 - \frac{\alpha_1 M_1}{\alpha_j \sum_t d_t}$$
$$\leqslant 2 - \frac{M_1}{\sum_t d_t}.$$

Here, the first inequality follows from the bound on α_g, the second inequality follows from the fact that the function $f(x) = (x-a)/x$ is increasing on (a, ∞), the third inequality follows from the fact that the modified-all unit discount function is nondecreasing and the last inequality follows from the first input condition. □

Hence, we obtain the following result.

Theorem 4.2. The modified all-unit discount cost function $f(x)$ has a β-sandwich function $r(x)$ with approximation guarantee $\beta = 2 - \frac{M_1}{\sum_t d_t}$, which can be constructed in $\mathcal{O}(n)$ time. Moreover, if $n = 1$, that is, the function has a single breakpoint, then this approximation guarantee is tight.

Proof. The approximation guarantee follows from Lemma 4.3 and the fact that the approximation guarantee of this upper concave envelope $r(x)$ is given by (Property 3.3) $\beta = \max_{i=1,\ldots,k}\{\beta_i\}$, where k is the number of pieces of this function and the β_is are the approximation guarantees of these pieces. The tightness follows from Lemma 4.2. □

One of the main research questions of this thesis is whether introducing discontinuities to the sandwich function will improve the approximation guarantee. We show that the answer is negative for this sandwich function for the modified all-unit discount cost function. That is, introducing even a finite number of discontinuities (independent of the input) to the piecewise linear concave sandwich function does not seem to improve the approximation guarantee. This is stated more explicitly in the following theorem.

Theorem 4.3. For any $0 < \varepsilon < 1$ and $m \in \mathbb{N}_{\geq 1}$, there exists an input (α_i) and (M_i) to the modified all-unit discount cost function $f(x)$ for which the piecewise linear concave β-sandwich function $r(x)$ has at least m pieces with tight approximation guarantee of at least $2 - \varepsilon$.

Proof. We consider an input (α_i) and (M_i) where the upper concave envelope is the collection of lines between the (except for possibly the first) consecutive breakpoints M'_i and M'_{i+1}. By then choosing the appropriate input size, breakpoints and slopes we get $\beta_i \geqslant 2 - \varepsilon$ for each piece (except for possibly the first), and we obtain the desired result.

Let $r > 1$ and assume $n \in \mathbb{N}_{\geq 1}$ general for now. We set $\alpha_i = r^{n-i}$, $M_1 = 1$ and $M_{i+1} = r^2 M_i$. Clearly, the first two input conditions are satisfied. The last input condition is also satisfied, because $\alpha_{i+1}/\alpha_i M_{i+1} = (1/r)M_{i+1} = rM_i > M_i$. We now calculate the slope $\Delta_i = \Delta f(M'_{i+1}, M'_i)$ between two consecutive breakpoints:

$$\begin{aligned}
\Delta_i &= \Delta f(M'_{i+1}, M'_i) \\
&= \frac{f(M'_{i+1}) - f(M'_i)}{M'_{i+1} - M'_i} \\
&= \frac{r^{n-i-2}M_{i+2} - r^{n-i-1}M_{i+1}}{\frac{1}{r}(M_{i+2} - M_{i+1})} \\
&= r^{n-i-1}\left(\frac{M_{i+2} - rM_{i+1}}{M_{i+2} - M_{i+1}}\right) \\
&= r^{n-i-1}\left(1 - \frac{(r-1)M_{i+1}}{M_{i+2} - M_{i+1}}\right) \\
&= r^{n-i-1}\left(1 - \frac{r-1}{r^2-1}\right) \\
&= r^{n-i-1}\left(1 - \frac{1}{r+1}\right).
\end{aligned}$$

Because the sequence $\{\Delta_i\}$ is decreasing, the upper concave envelope will be the collection of lines between the (except for possibly the first) consecutive breakpoints M'_i and M'_{i+1}. In fact, the piece of $r(x)$ between the breakpoints M'_i and M'_{i+1} has slope Δ_i and approximation guarantee (using Lemma 4.2 and setting $r = 1/(1 - \sqrt{1-\varepsilon})$)

$$\begin{aligned}
\beta_i &= 1 + \frac{\Delta_i}{\alpha_{i+1}}\left(\frac{M - M'_i}{M}\right) \\
&= 1 + \frac{r^{n-i-1}}{r^{n-i-1}}\left(1 - \frac{1}{r+1}\right)\left(\frac{M - \frac{1}{r}M}{M}\right)
\end{aligned}$$

$$
\begin{aligned}
&= 1 + \left(1 - \frac{1}{r+1}\right)\left(1 - \frac{1}{r}\right) \\
&\geqslant 1 + \left(1 - \frac{1}{r}\right)^2 \\
&= 1 + (1 - (1 - \sqrt{1-\varepsilon}))^2 \\
&= 2 - \varepsilon.
\end{aligned}
$$

If we let $n = m + 2$, then the upper concave envelope $r(x)$ of $f(x)$ has at least m pieces with tight approximation guarantee of at least $2 - \varepsilon$. $\square$

4.2 The stepwise cost function

In this section we will use the notation from Akbalik and Rapine (2013) on the input parameters of the stepwise cost function. This is possible if we bear in mind that $f(\cdot)$ denotes a function and f a parameter. The stepwise cost function or multiple setup cost function $f(x)$ takes as an input the quadruple (f, p, k, B), and is given by $f(x) = 0$ for $x = 0$ and $f(x) = f + px + k\lceil x/B \rceil$ for $x > 0$. Here, f is the fixed setup or ordering cost, p is the unit production or procurement cost, k is the batch production or ordering cost and $B \in \mathbb{N}_{\geq 1}$ is the batch size.

Because the term px is trivially sandwiched for any linear sandwich function, we will from now on omit it from further consideration, so that the cost function reduces to $f(x) = f + k\lceil x/B \rceil$. When we decide to use the sandwich function to do computations for example, we can always add the term px back to the model. When trying to fit a linear β-sandwich function, we obtain the following result.

Theorem 4.4. The stepwise cost function $f(x)$ has a linear β-sandwich function $r(x)$ with approximation guarantee $\beta = 2 - \frac{f}{f+k}$. Moreover, this approximation guarantee is tight.

Proof. This theorem is a special case of Lemma 4.4. $\square$

This theorem is illustrated in Figure 3.

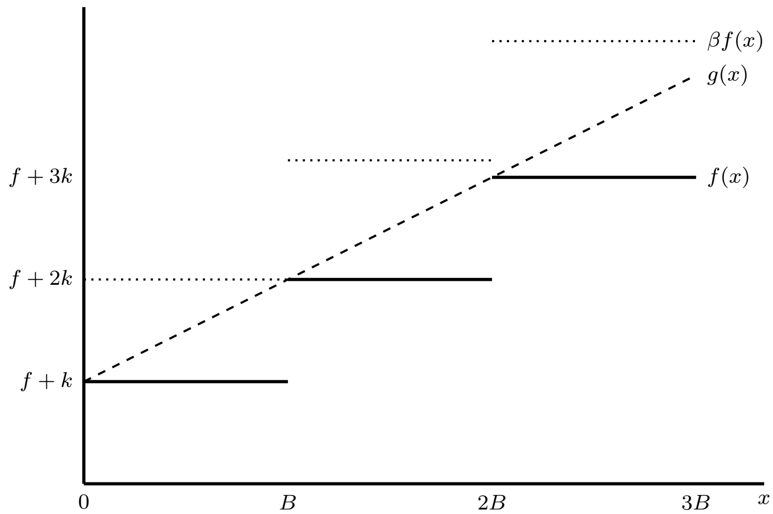

Figure 3: The stepwise cost function and its β-sandwich function.

However, when we allow for discontinuities we can do better. This is illustrated by the following lemma.

Lemma 4.4. The stepwise cost function $f(x)$ has a β-sandwich function $r(x)$ consisting of m pieces with approximation guarantee $\beta = 1 + \frac{1}{m} - \frac{f}{m(f+mk)}$. Moreover, this approximation guarantee is tight.

Proof. Let $r(x)$ be a piecewise linear function defined as

$$r(x) = \begin{cases} 0 & \text{if } x = 0 \\ f + k\lceil \frac{x}{B} \rceil & \text{if } 0 < x \leqslant (m-1)B \\ f + k + k\frac{x}{B} & \text{if } x > (m-1)B \end{cases}$$

Clearly we have that $r(x) \geqslant f(x)$ for all x. One key observation is that $r(x)/f(x)$ is increasing on the interval $((m-1)B, mB]$. Hence, $\beta \leqslant r(mB)/f(mB) = (f+(m+1)k)/(f+mk) = 1+1/m-f/(m(f+mk))$. Tightness of the approximation guarantee follows from $r(0^+) = f(0^+)$ and $r(mB) = \beta f(mB)$. $\square$

While the sandwich function for the modified all-unit discount function cannot be improved even when introducing a finite number of discontinuities, quite the opposite is true for the stepwise cost function. This function can be approximated arbitrarily close with a finite number of pieces.

Theorem 4.5. For any $\varepsilon > 0$, there is a β-sandwich function $r(x)$ of the stepwise cost function $f(x)$ consisting of a finite number of pieces having a tight approximation guarantee of at most $1 + \varepsilon$.

Proof. Considering a sandwich function consisting of $m = \lceil 1/\varepsilon \rceil$ pieces gives $\beta \leqslant 1 + 1/m = 1 + 1/(\lceil 1/\varepsilon \rceil) \leqslant 1 + 1/(1/\varepsilon) = 1 + \varepsilon$. $\square$

5 Applications to literature

In this section we will look at applications of sandwich functions to lot-sizing problems. In Section 5.1, we will show that the linear approximation of Hill and Galbreth (2008) can perform arbitrarily poor. In Section 5.2, we look at applications amongst the lot-sizing problem with modified all-unit discount cost, the lot-sizing problem with batch cost and the lot-sizing problem with demand time windows and stepwise cost. We end this section with Section 5.3, where we provide integer programming formulations for sandwich functions that cannot be used efficiently in a dynamic programming environment.

5.1 Modified all-unit discount cost function linear approximation

Hill and Galbreth (2008) propose to approximate the modified all-unit discount cost function $f(x)$ instead of sandwiching it. The approximation function $g(x)$ they suggest satisfies three conditions. First of all, the function is linear. Secondly, the minimum charge for both functions are the same, i.e. $g(0^+) = f(0^+) = \alpha_1 M_1$. Finally, the total area under the linear function $\mathcal{A}$ equals to total area under the modified all-unit discount function. In the calculation of this area they ignore the area from the rectangle $(0,0)$ to $(M_n, \alpha_1 M_1)$, because both functions share this area under their curves. The linear approximation they obtain is given by $g(x) = (2\mathcal{A}/M_n^2)x + \alpha_1 M_1$, where

$$\mathcal{A} = \sum_{i=2}^{n} \left[\frac{1}{2}(\alpha_i M_i - \alpha_{i-1} M_{i-1})(M'_{i-1} - M_{i-1}) + (M_n - M'_{i-1})(\alpha_i M_i - \alpha_{i-1} M_{i-1}) \right] + \frac{1}{2}(M_n - M_{n-1})^2 \alpha_{n-1}.$$

We will show that the approximation function $g(x)$ can perform arbitrarily bad. This statement is similar to Theorem 4.1 from Section 4, which says that no linear β-sandwich function for the modified all-unit discount cost function exists. The only difference is that for this approximation, the ratio between the approximation function and the true function can be arbitrarily small.

Theorem 5.1. The linear approximation from Hill and Galbreth (2008) for the modified all-unit discount cost function can perform arbitrarily bad.

Proof. Proving this theorem is equivalent to showing that for any $0 < \varepsilon < 1$, there exists an input (α_i) and (M_i) to the modified all-unit discount cost function $f(x)$ for which the linear approximation $g(x)$ from Hill and Galbreth (2008) has $\frac{g(x_1)}{f(x_1)} < \varepsilon$ for some $x_1 \geqslant 0$. We will do this by providing a bound on $\mathcal{A}$ for a certain class of modified all-unit discount functions, and show that this leads to the ratio $g(x_1)/f(x_1)$ being sufficiently small for suitably chosen x_1.

We will start with providing a bound on $\mathcal{A}$ for a set of modified all-unit discount cost functions. Similar as in the proof of Theorem 4.1, consider $n = 2$ breakpoints. The expression for $\mathcal{A}$ reduces to $\mathcal{A} = (\alpha_2 M_2 - \alpha_1 M_1)(M_1' - M_1)/2 + (M_2 - M_1')(\alpha_2 M_2 - \alpha_1 M_1) + (M_2 - M_1)^2\alpha_1/2 = (\alpha_2 M_2 - \alpha_1 M_1)(M_2 - M_1'/2 - M_1/2) + (M_2 - M_1)^2\alpha_1/2 = (\alpha_2 M_2 - \alpha_1 M_1)(M_2 - \alpha_2 M_2/(2\alpha_1) - M_1/2) + (M_2 - M_1)^2\alpha_1/2$. Now consider the input $\alpha_1 = 1/2 + 1/(2\max\{1/3, 1 - k/4\})$, $M_1 = \max\{1/3, 1 - k/4\}$, $\alpha_2 = M_2 = 1$ for some constant $0 < k < 1$. One can check that this input is valid. Moreover, we obtain $\mathcal{A} = (1 - \alpha_1 M_1)(1 - 1/(2\alpha_1) - M_1/2) + (1 - M_1)^2\alpha_1/2 \leqslant (1 - \alpha_1 M_1)(1 - M_1) + (1 - M_1)^2\alpha_1/2 = (1 - M_1)(1 + \alpha_1/2 - 3\alpha_1 M_1/2) \leqslant (1 - M_1)(1 + \alpha_1/2) \leqslant 2(1 - M_1) \leqslant 2(1 - (1 - k/4)) = k/2$. Here, the first inequality follows from the third input condition, the second inequality follows from the nonnegativity of the modified all-unit discount cost function and the last two inequalities from the fact that $\max\{1/3, 1 - k/4\} \geqslant 1/3$ and $\max\{1/3, 1 - k/4\} \geqslant 1 - k/4$. Setting $k = \varepsilon/2$ and $x_1 = 2/\varepsilon$ gives $g(x_1) = 2\mathcal{A}x_1 + \alpha_1 M_1 < (\varepsilon/2)(2/\varepsilon) + 1 = 2 = \varepsilon(2/\varepsilon) = \varepsilon f(x_1)$. □

5.2 The lot-sizing problem

In the classical lot-sizing problem we are given a discrete time horizon $t = 1, \ldots, T$. For each period t, there is a known demand $d_t \in \mathbb{N}$. The problem asks to find the optimal production quantities (x_t) and inventory levels (I_t) such that the demand is met and the cost of production and holding inventory over the time period is minimised. The problem can be formulated as a mathematical program. This mathematical program, following the notation from Brahimi et al. (2017), is given by

$$\begin{aligned} \min \quad & \sum_{t=1}^{T} \left[f_t^p(x_t) + f_t^h(I_t) \right] \\ \text{s.t.} \quad & I_t = I_{t-1} + x_t - d_t && \text{for } t = 1, \ldots, T, \\ & x_t, I_t \geqslant 0 && \text{for } t = 1, \ldots, T. \end{aligned}$$

Here, $f_t^p(x)$ denote the production cost functions, $f_t^h(x)$ denote the holding cost functions, the first constraint is the inventory balance constraint and the last constraint specifies the domain of the decision variables. A common assumption made in the literature is that initial and ending inventory are zero, that is, $I_0 = I_T = 0$.

We have discussed sandwich functions in a general setting but not in a lot-sizing setting. However, one can show that if we can sandwich the production cost functions $f_t^p(x)$ by β_t-sandwich functions $r_t^p(x)$ each, then we can sandwich the objective function of this formulation with approximation guarantee $\beta = \max_{t=1,\ldots,T} \beta_t$. This claim is a direct application of Property 3.2. Hence, we can use the sandwich functions discussed earlier to construct approximation algorithms for the lot-sizing problem.

5.2.1 Modified all-unit discount cost structure

The lot-sizing problem with the modified all-unit discount cost structure (LSP-M) is the classical lot-sizing problem where the production cost function $f_t^p(x)$ is the modified all-unit discount cost function. Chan et al. (2002a) show that this problem is $\mathcal{NP}$-hard when the cost functions $f_t^p(x)$ are either time dependent or the number of breakpoints of the cost functions is not bounded by a constant. When we sandwich the production cost functions with their upper concave envelope β-sandwich functions, we obtain the following result.

Theorem 5.2. The lot-sizing problem with the modified all-unit discount cost structure admits a 2-approximation algorithm running in $\mathcal{O}(nT^3)$ time.

The running time $\mathcal{O}(nT^3)$ comes from the fact that the sandwich functions are piecewise linear concave. This allows us to use the algorithm from Ou (2017b), which solves the capacitated lot-sizing problem with piecewise linear concave costs in $\mathcal{O}(nT^3)$ time, where n is the average number of line segments of the cost functions. We should point out that this algorithm is developed for the capacitated problem, thus better algorithms for this problem could exist.

5.2.2 Batch cost structure

The lot-sizing problem with batch procurement (LSP-B) is the classical lot-sizing problem where the production cost function $f_t^p(x)$ is the stepwise cost function. Akbalik and Rapine (2013) show that this problem is $\mathcal{NP}$-hard when the batches are time dependent and any of the cost parameters is nonstationary over time. When we sandwich the production cost functions with linear β-sandwich functions, we obtain the following result.

Theorem 5.3. (Akbalik and Rapine (2017)) The lot-sizing problem with batch procurement admits a 2-approximation algorithm running in $\mathcal{O}(T \log T)$ time.

The running time $\mathcal{O}(T \log T)$ comes from the fact that the sandwich functions are linear, allowing us to use the algorithm from Wagelmans et al. (1992).

We would like to end this subsection with another application of sandwich functions. In particular, we show that because we can sandwich the stepwise cost function arbitrary close, we obtain a polynomial time approximation scheme (PTAS) for the CLSP-B. Recall that a PTAS for a problem is an algorithm (or family of algorithms) such that for every $\varepsilon > 0$, this algorithm is a $(1 + \varepsilon)$-approximation algorithm for this problem.

Theorem 5.4. The lot-sizing problem with batch procurement admits a PTAS.

This result follows from the fact that we can sandwich the stepwise cost function by a sandwich function having an approximation guarantee of at most

$(1+\varepsilon)$, see Theorem 4.5. Because for a PTAS the value ε is fixed, so is the number of pieces of the sandwich function that is used. We can therefore apply the algorithm from Ou (2017a), which solves the lot-sizing with piecewise linear cost in polynomial time in T. We do have to point out that in this specific scenario, a PTAS is not particularly useful. This is because the problem can be solved in polynomial time in $\mathcal{O}(T^6)$ by an existing algorithm (Akbalik and Rapine (2012)). Yet, it remains nice to see that a sandwich approach cannot only lead to constant factor approximation algorithms, but also to approximation schemes.

5.2.3 2-approximation algorithm for the time windows and batch cost structure

The lot-sizing problem with demand time windows and stepwise cost (LSP-TWS) is introduced in Jaruphongsa and Lee (2008). This problem is an extension of the classical LSP where each demand order d_i, $i = 1, \dots, n$, has a demand time window $[e_i, \ell_i]$, $e_i \leqslant \ell_i$. If the demand is satisfied in this time period, no penalty cost in terms of holding inventory or backlogging is incurred. We assume w.l.o.g. that each demand order is unique, i.e. if demand i and i' have the same demand time window ($e_i = e_{i'}$ and $\ell_i = \ell_{i'}$) they can be aggregated to a single demand order. This way, the number of demands is bounded by $n \leqslant \frac{1}{2}T(T+1)$. When we set $e_i = \ell_i$ for each demand i, we obtain the classical lot-sizing problem.

The production cost function in the LSP-TWS is the stepwise cost function $f_t(x) = f_t + p_t x + k_t\lceil x/B\rceil$ from Section 4.2, assuming stationary batch size B, and (f_t), (p_t) and (k_t) all to be nonincreasing over time. This is the same setting as the one used in Jaruphongsa and Lee (2008). Note that they do not consider backordering in their model, while this could have been included.

They show that the problem is strongly $\mathcal{NP}$-hard when split delivery, also known as order splitting, is not allowed. Order splitting refers to the phenomenon of satisfying a demand order from a single order period, not from multiple. We show that we can obtain a 2-approximation algorithm for this problem when we sandwich the production cost function by a linear β-sandwich function. The most important ingredient for this approximation algorithm is the following property.

Property 5.1. (Lee et al. (2001)) There exists an optimal solution to the lot-sizing problem with demand time windows where delivery is not split. This solution can be found in $\mathcal{O}(T^2)$ time.

Property 5.1 can be proven by considering an optimal solution to the LSP-TW and any demand d_i. Suppose that order i is split, that is, $d_i = d_{it} + d_{it'}$ for two periods t and t'. One can check that satisfying d_i fully from period t or t' does not increase the cost, hence there is an optimal solution to the LSP-TW where order is not split. The runtime $\mathcal{O}(T^2)$ follows from Lee et al. (2001), who present a dynamic programming algorithm for this problem. We obtain the following result.

Theorem 5.5. The lot-sizing problem with demand time windows and step-wise cost admits a 2-approximation algorithm running in $\mathcal{O}(T^2)$ time, when order splitting is not allowed.

This theorem follows from the fact that when we replace the objective function by a 2-sandwich function, we obtain a lot-sizing problem with demand time windows where order splitting is not allowed. But, in this case the cost are linear, so we can use the algorithm from Lee et al. (2001) to obtain an optimal solution to this problem. Because this solution is also feasible to the LSP-TWS and the objective function is 2-sandwiched, we have obtained a 2-approximation algorithm running in $\mathcal{O}(T^2)$ time.

5.3 Integer programming formulations

One of the issues with the sandwich approach taken thus far is that although some sandwich functions allow for a good approximation guarantee, they cannot be used efficiently in a dynamic programming environment due to their complexity. One example of such a sandwich function is the piecewise linear sandwich function for the stepwise function described in Section 4.2. However, in some cases these sandwich functions that are too evolved to be used in a dynamic programming approach can be modelled by a mathematical programming formulation. We show that for some mathematical programming formulations, solving an alternative formulation leads to a solution whose quality with respect to the optimal solution is known a priori. Moreover, any feasible solution to this alternative formulation can be used to construct a feasible solution to the original formulation, often with only little time resource.

In this section we will give a couple of examples where sandwich functions can be used to construct alternative programming formulations. In particular, we provide a family of formulations to model the piecewise linear sandwich function for the LSP-B and a formulation to model the piecewise linear sandwich function for the lot-sizing problem with batch ordering and capacity reservation (LSP-BCR).

5.3.1 Batch cost structure

The lot-sizing problem with batches is the lot-sizing problem where the production cost functions are stepwise cost functions. In Section 4.2 we showed that the stepwise cost function has a piecewise linear sandwich function consisting of m pieces having an approximation guarantee of $1 + \frac{1}{m}$. We will present a mathematical programming formulation where the stepwise cost function is replaced by this sandwich function.

The LSP-B can be modelled using an integer programming formulation provided by Akbalik and Pochet (2009). It is given by

$$(F) \quad \min \quad \sum_{t=1}^{T} [f_t y_t + p_t x_t + k_t z_t + h_t I_t] \tag{1}$$

$$\text{s.t.} \quad I_t = I_{t-1} + x_t - d_t \qquad \text{for } t = 1, \ldots, T, \tag{2}$$

$$x_t \leqslant M y_t \qquad \text{for } t = 1, \ldots, T, \tag{3}$$

$$x_t \leqslant B_t z_t \qquad \text{for } t = 1, \ldots, T, \tag{4}$$

$$x_t, I_t \geqslant 0 \qquad \text{for } t = 1, \dots, T, \tag{5}$$

$$y_t \in \mathbb{B} \qquad \text{for } t = 1, \dots, T, \tag{6}$$

$$z_t \in \mathbb{N} \qquad \text{for } t = 1, \dots, T. \tag{7}$$

Here, (1) is the objective function to be minimised, consisting of a setup cost, unit procurement cost, batch procurement cost and holding cost. Constraints (2) are the classical inventory balance constraints, and constraints (3) and (4) relate the setup and batch variables (y_t) and (z_t) to the production quantities (x_t). Constraints (5) - (7) specify the domain of the production quantities (x_t), inventory levels (I_t), setup indicators (y_t) and batch sizes (z_t).

We now present a formulation where we replace the objective functions by their sandwich functions, which are piecewise linear functions consisting of m pieces. In order to do so we will split the production quantity in $x = x^B + x^\ell$, where the first quantity x^B models the production quantity $0 \leqslant x_t \leqslant (m-1)B$ and the second quantity models the production quantity $x > (m-1)B$. Now that we have the split the production quantity into these two quantities, we can assign different cost structures to each, these being batch cost and linear cost, respectively. The integer programming formulation is given by (F_m). Note that when $m = 1$ we simply use a linear sandwich function, and obtain the classical lot-sizing problem formulation.

$$(F_m) \quad \min \quad \sum_{t=1}^{T} \left[f_t y_t + p_t x_t + k_t z_t + h_t I_t + k_t y_t^\ell + \left(\frac{k_t}{B_t} \right) x_t^\ell \right] \tag{8}$$

$$\text{s.t.} \quad I_t = I_{t-1} + x_t - d_t \qquad \text{for } t = 1, \dots, T, \tag{9}$$

$$x_t \leqslant M y_t \qquad \text{for } t = 1, \dots, T, \tag{10}$$

$$x_t = x_t^B + x_t^\ell \qquad \text{for } t = 1, \dots, T, \tag{11}$$

$$x_t^B \leqslant B_t z_t \qquad \text{for } t = 1, \dots, T, \tag{12}$$

$$z_t \leqslant (m-1) \qquad \text{for } t = 1, \dots, T, \tag{13}$$

$$(m-1) - z_t \leqslant M \delta_t \qquad \text{for } t = 1, \dots, T, \tag{14}$$

$$x_t^\ell \leqslant M(1 - \delta_t) \qquad \text{for } t = 1, \dots, T, \tag{15}$$

$$x_t^\ell \leqslant M y_t^\ell \qquad \text{for } t = 1, \ldots, T, \tag{16}$$

$$x_t, x_t^B, x_t^\ell, I_t \geqslant 0 \qquad \text{for } t = 1, \ldots, T, \tag{17}$$

$$y_t, y_t^\ell, \delta_t \in \mathbb{B} \qquad \text{for } t = 1, \ldots, T, \tag{18}$$

$$z_t \in \mathbb{N} \qquad \text{for } t = 1, \ldots, T. \tag{19}$$

Constraints (9) and (10) are the same constraints as the ones from (F), and constraints (11) split the production quantities (x_t) into (x_t^B) and (x_t^ℓ) as described earlier. Constraints (12) - (16) model the cost structure of the piecewise linear sandwich function. In short, if $x_t \leqslant (m-1)B_t$, then the production costs are $p_t x_t + k_t \lceil x_t/B_t \rceil$, and if $x_t > (m-1)B_t$ then the costs are $p_t x_t + k_t + (k_t/B_t)x_t$. Finally, (17) - (19) specify the domain of the decision variables.

We should point out that it is only worthwhile to consider solving an alternative formulation if this alternative formulation has some desirable properties. It turns out that this is the case, as we can see from the following theorem.

Theorem 5.6. Let OPT_F and OPT_{F_m} be the optimal objective values of formulations (F) and (F_m), respectively. Then

(1) Any solution to (F_m) can be transformed in $\mathcal{O}(T)$ time to a solution of (F) with an at least as good objective value.

(2) $\mathrm{OPT}_F \leqslant \mathrm{OPT}_{F_m} \leqslant (1 + \frac{1}{m})\mathrm{OPT}_F$.

Proof. The first claim follows from the fact that given a feasible solution $(x_t, x_t^B, x_t^\ell, I_t, y_t, y_t^\ell, \delta_t, z_t)$ for (F_m), the tuple (x_t, I_t, y_t, z_t') is feasible for (F), where $z_t' = \lceil x_t/B_t \rceil$. Because the objective function of (F_m) is a $(1 + \frac{1}{m})$-sandwich function for the objective function of (F), the objective value cannot increase. This also proves (2). $\square$

Note that the quality of the solution obtained from solving (F_m) depends on m, the number of pieces used in the sandwich function of the stepwise cost function. This gives the modeller the opportunity to tune this parameter, with a potential trade-off between solution quality and runtime of solving the formulation.

5.3.2 Batch cost and capacity reservation cost structure

Another example of a problem where the sandwich approach by dynamic programming seems to fail is the lot-sizing problem with batch ordering and a capacity reservation contract (LSP-BCR). This problem was recently discussed in Akbalik et al. (2017), where $\mathcal{NP}$-hard and polynomial cases of the problem are identified. We will use their notation throughout this section. Simply put, the problem is the same as the LSP-B, but now at a certain capacity, R, the batch costs increases. This corresponds to a situation where a long-term contract is established between a manufacturer and an external supplier in which the costs and the reserved capacities are specified. Formally, the function is given by

$$f(x) = \begin{cases} K + a\lceil x/V \rceil & \text{if } 0 \leqslant x \leqslant RV, \\ K + aR + b(\lceil x/V \rceil - R) & \text{if } x \geqslant RV, \end{cases}$$

where K is the fixed setup cost, a is the batch cost when production is below capacity, b is the batch cost when production exceeds capacity, V is the batch size and R is the set capacity. It is generally assumed that $a < b$, that is, the suppliers offers a discount whenever production is below capacity.

An obvious 2-sandwich function for the production cost function is a piecewise linear function consisting of 2 pieces. The first piece being the linear sandwich function on $[0, RV]$ and the second piece being the linear sandwich function on (RV, ∞). Unfortunately, because of the potential discontinuity at the point RV, this sandwich function cannot be used efficiently in a dynamic programming environment. However, this function can be modelled using a mathematical programming formulation. We will present this formulation below.

The LSP-BCR can be modelled using an integer programming formulation provided by Akbalik et al. (2017). In this formulation, the inventory holding costs (h_t) have been substituted out using the inventory balance constraints. The formulation is given by

$$
\begin{aligned}
(FCR) \quad \min \quad & \sum_{t=1}^{T} [K_t y_t + a_t A_t + b_t B_t + p_t x_t] && && (20)\\
\text{s.t.} \quad & \sum_{s=1}^{T} x_s = \sum_{s=1}^{T} d_s && && (21)\\
& \sum_{s=1}^{t} x_s \geqslant \sum_{s=1}^{t} d_s && \text{for } t = 1, \ldots, T-1, && (22)\\
& x_t \leqslant M y_t && \text{for } t = 1, \ldots, T, && (23)\\
& x_t \leqslant V_t(A_t + B_t) && \text{for } t = 1, \ldots, T, && (24)\\
& A_t \leqslant R_t && \text{for } t = 1, \ldots, T, && (25)\\
& x_t \geqslant 0 && \text{for } t = 1, \ldots, T, && (26)\\
& y_t \in \mathbb{B} && \text{for } t = 1, \ldots, T, && (27)\\
& A_t, B_t \in \mathbb{N} && \text{for } t = 1, \ldots, T. && (28)
\end{aligned}
$$

The objective function (20) consists of the setup cost, batch cost (both below and above production capcacity) and unit procurement cost. The demand is met due to constraints (21) and (22), and the setup variables (y_t) are related to the production quantities (x_t) in (23). The batch variables (A_t) and (B_t) are calculated using (24) and (25), and the domain of the production quantities (x_t), setup indicators (y_t), and batch sizes (A_t) and (B_t) are specified in the constraint set (26) - (28).

When we sandwich the production cost function by a piecewise linear sandwich function consisting of two pieces, we obtain the following formulation.

$$
(FCR') \quad \min \quad \sum_{t=1}^{T} \left[K_t y_t + p_t x_t + a_t y_t^A + \left(\frac{a_t}{V_t}\right) x_t^A + b_t y_t^B + \left(\frac{b_t}{V_t}\right) x_t^B \right] \quad (29)
$$

$$
\text{s.t.} \quad \sum_{s=1}^{T} x_s = \sum_{s=1}^{T} d_s \qquad (30)
$$

$$\sum_{s=1}^{t} x_s \geqslant \sum_{s=1}^{t} d_s \qquad \text{for } t = 1, \ldots, T-1, \tag{31}$$

$$x_t \leqslant M y_t \qquad \text{for } t = 1, \ldots, T, \tag{32}$$

$$x_t = x_t^A + x_t^B \qquad \text{for } t = 1, \ldots, T, \tag{33}$$

$$x_t^A \leqslant V_t R_t y_t^A \qquad \text{for } t = 1, \ldots, T, \tag{34}$$

$$V_t R_t - x_t^A \leqslant M \delta_t \qquad \text{for } t = 1, \ldots, T, \tag{35}$$

$$x_t^B \leqslant M(1 - \delta_t) \qquad \text{for } t = 1, \ldots, T, \tag{36}$$

$$x_t^B \leqslant M y_t^B \qquad \text{for } t = 1, \ldots, T, \tag{37}$$

$$x_t, x_t^A, x_t^B \geqslant 0 \qquad \text{for } t = 1, \ldots, T, \tag{38}$$

$$y_t, y_t^A, y_t^B, \delta_t \in \mathbb{B} \qquad \text{for } t = 1, \ldots, T. \tag{39}$$

Constraints (30) - (32) are the same constraints as the ones from (FCR), and constraints (33) split the production quantities (x_t^A) and (x_t^B). Here x_t^A is the production quantity under the capacity reservation and x_t^B is the production quantity exceeding the capacity reservation. The piecewise linear cost structure is captured in constraints (34) - (37). Shortly put, if the production quantity $x_t \leqslant V_t R_t$ then the production costs are $p_t x_t + a_t + (a_t/V_t)x_t$, and if $x_t > V_t R_t$ then the costs are $p_t x_t + a_t R_t + b_t + (b_t/V_t)(x_t - V_t R_t)$. Finally, the domain of the decision variables is specified in (38) - (39).

Similar as for the sandwich formulation for the LSP-B, we can derive certain relationships between the formulation (FCR) and its alternative formulation (FCR').

Theorem 5.7. Let OPT_{FCR} and $\text{OPT}_{FCR'}$ be the optimal objective values of formulations (FCR) and (FCR'), respectively. Then

(1) Any solution to (FCR') can be transformed in $\mathcal{O}(T)$ time to a solution of (F) with an at least as good objective value.

(2) $\text{OPT}_{FCR} \leqslant \text{OPT}_{FCR'} \leqslant 2\text{OPT}_{FCR}$.

Proof. The first claim follows from the fact that given a feasible solution $(x_t, x_t^A, x_t^B, y_t, y_t^A, y_t^B, \delta_t)$ for (FCR'), the tuple (x_t, y_t, A_t', B_t') is feasible for

(FCR), where $A'_t = \lceil x_t^A / V_t \rceil$ and $B'_t = \lceil x_t^B / V_t \rceil$. Because the objective function of (FCR') is a 2-sandwich function for the objective function of (FCR), the objective value cannot increase. This also proves (2). $\square$

5.3.3 Computational results

In this section we perform a computational study on the family of formulations provided in Section 5.3.1. That is, we consider an integer programming approach for the lot-sizing problem with batch procurement (LSP-B). We compare the solution times and qualities of solving formulations (F) and (F_m) for different values of m.

The instances are generated with the following parameter values (recall the notation from Section 4.2): $f_t \sim [1, 25]$, $p_t = 20$, $k_t \sim [26, 50]$, $h_t = 5$, $d_t \sim [1, 100]$, $B_t \sim [3, 9]$, for each period t. The parameters are chosen in such a way that a typical solution consists of several periods in which production takes place, and several periods in which inventory is held only. The time horizon is chosen $T = 300$ and the output statistics are averaged over 100 executions. All computations are done using Gurobi 8.0 on an Intel Core 2.30 Gigahertz, 8 Gigabyte RAM.

The results are displayed in Table 5.3.3. For a given instance, the formulations that are solved are (F), the exact formulation for the LSP-B, and the formulations (F_m), which use the sandwich function consisting of m pieces as the objective function. The theoretical gaps are calculated using

$$\max_{t=1,\ldots,T} \left\{ 1 + \frac{1}{m} - \frac{f_t}{m(f_t + mk_t)} \right\},$$

and are known in advance. The real gaps for an alternative formulation F_i are given by the objective value (in terms of objective from F) of the solution obtained from solving this formulation divided by the value of the optimal solution of formulation F.

Table 1: Computational results of the integer programming formulations for the lot-sizing problem with batch procurement cost. We use the abbreviation std for the standard deviation.

Formulation	Theoretical gap	Real gap	Runtime (sec)	Runtime std (sec)
F	NA	NA	12.0180	8.9294
F_1	1.9797	1.0053	0.0465	0.0136
F_2	1.4949	1.0052	0.1787	0.0605
F_3	1.3310	1.0049	0.2983	0.1094
F_4	1.2487	1.0046	0.3870	0.1343
F_5	1.1992	1.0044	0.5894	0.2884
F_6	1.1661	1.0041	0.8370	0.3294
F_7	1.1424	1.0037	1.0961	0.3732
F_8	1.1247	1.0034	1.3471	0.2800
F_9	1.1109	1.0032	1.8607	0.8116
F_{10}	1.0998	1.0029	3.7314	1.9023
F_{11}	1.0907	1.0026	5.7440	2.1929
F_{12}	1.0832	1.0022	7.7609	2.7968
F_{13}	1.0768	1.0019	10.3222	5.0122
F_{14}	1.0713	1.0016	14.1193	7.1578
F_{15}	1.0666	1.0014	19.8360	12.2098

There are several observations that can be made here. One of these is that there is a clear trade-off between solution quality and runtime. That is, when the sandwich functions consists of more pieces, the theoretical and real gap decrease while the overall running time increases. Remarkably, for this parameter setting, the linear sandwich function ($m = 1$) performs really well, resulting in an average gap of 0.53% and a runtime much faster than the one obtained from the exact approach. On the other hand, it is interesting to see that when the number of pieces in the sandwich function becomes high enough, in this case around 13 pieces, the runtime of solving the alternative formulation is higher than the one from the exact approach, and the quality of the obtained solution is worse.

We should mention that this computational study is not very exhaustive. The LSP-B is quite an easy problem in the sense that it is solvable to optimality in a couple of seconds, even with a time horizon of $T = 300$, and the parameter setting is rather restricted. The standard deviations are also quite high, see for example the 8.92 seconds for solving the exact approach, while the average solving time of this formulation 12.02 seconds.

These results suggest that it might only be worthy to use the alternative formulations to obtain a decent solution within a short time span, rather than for finding near optimal solutions. The best setting under where this can be achieved could be a problem that is very hard to solve, or a problem where even finding a feasible solution is very difficult.

6 Conclusion

In this thesis we considered sandwich functions for the lot-sizing problem. We presented instances where this method has been used before, and theoretically studied the modified-all unit discount cost function and the stepwise cost function. We showed that the stepwise cost function can be sandwiched arbitrarily close when we introduce discontinuities in the sandwich functions. We also presented limitations on the sandwich function approach for the modified all-unit discount cost function. These limitations include the nonexistence of a linear sandwich function for this cost function, and the absence of improvement in the approximation guarantee when introducing discontinuities in the sandwich function.

Applications to the lot-sizing literature were also presented. Although some of the approximation algorithms obtained from this method were not competitive with existing approximation algorithms in terms of runtime or approximation guarantee, we did obtain a 2-approximation algorithm for the lot-sizing problem with demand time windows and stepwise cost. Recall that this problem is strongly $\mathcal{NP}$-hard when order splitting is not allowed. To the best of our knowledge, this is the first constant factor approximation algorithm for this problem.

Several research suggestions are as follows. First of all, a more detailed literature review on the sandwich function approach would be interesting. This is mainly because to the best of our knowledge, this method is not systematically used yet. Secondly, it would also be worthwhile to find other lot-sizing problems where the sandwich function approach can be used. Although the method can only be used in a limited setting, the method itself is extremely simple and general, only relying on Property 3.1. Finally, more detailed computational experiments could be of interest as well.

References

Ayse Akbalik and Yves Pochet. Valid inequalities for the single-item capacitated lot sizing problem with step-wise costs. *European Journal of Operational Research*, 198(2):412–434, 2009.

Ayse Akbalik and Christophe Rapine. Polynomial time algorithms for the constant capacitated single-item lot sizing problem with stepwise production cost. *Operations Research Letters*, 40(5):390–397, 2012.

Ayse Akbalik and Christophe Rapine. The single item uncapacitated lot-sizing problem with time-dependent batch sizes: $\mathcal{NP}$-hard and polynomial cases. *European Journal of Operational Research*, 229(2):353–363, 2013.

Ayse Akbalik and Christophe Rapine. Lot sizing problem with multi-mode replenishment and batch delivery. *Omega*, 2017.

Ayse Akbalik, Atidel B Hadj-Alouane, Nathalie Sauer, and Houcem Ghribi. Np-hard and polynomial cases for the single-item lot sizing problem with batch ordering under capacity reservation contract. *European Journal of Operational Research*, 257(2):483–493, 2017.

Alex M Andrew. Another efficient algorithm for convex hulls in two dimensions. *Information Processing Letters*, 9(5):216–219, 1979.

Claudia Archetti, Luca Bertazzi, and M Grazia Speranza. Polynomial cases of the economic lot sizing problem with cost discounts. *European Journal of Operational Research*, 237(2):519–527, 2014.

Nadjib Brahimi, Stephane Dauzere-Peres, Najib M Najid, and Atle Nordli. Single item lot sizing problems. *European Journal of Operational Research*, 168(1):1–16, 2006.

Nadjib Brahimi, Nabil Absi, Stéphane Dauzère-Pérès, and Atle Nordli. Single-item dynamic lot-sizing problems: An updated survey. *European Journal of Operational Research*, 263(3):838–863, 2017.

Joseph R Carter and Bruce G Ferrin. Transportation costs and inventory management: Why transportation costs matter. *Production and Inventory Management Journal*, 37(3):58, 1996.

Lap Mui Ann Chan, Ana Muriel, Zuo-Jun Shen, and David Simchi-Levi. On the effectiveness of zero-inventory-ordering policies for the economic lot-sizing model with a class of piecewise linear cost structures. *Operations Research*, 50(6):1058–1067, 2002a.

Lap Mui Ann Chan, Ana Muriel, Zuo-Jun Max Shen, David Simchi-Levi, and Chung-Piaw Teo. Effective zero-inventory-ordering policies for the single-warehouse multiretailer problem with piecewise linear cost structures. *Management Science*, 48(11):1446–1460, 2002b.

Sergei Chubanov, Mikhail Y Kovalyov, and Erwin Pesch. An FPTAS for a single-item capacitated economic lot-sizing problem with monotone cost structure. *Mathematical programming*, 106(3):453–466, 2006.

Awi Federgruen and Michal Tzur. A simple forward algorithm to solve general dynamic lot sizing models with n periods in $\mathcal{O}(n \log n)$ or $\mathcal{O}(n)$ time. *Management Science*, 37(8):909–925, 1991.

Uriel Feige. A threshold of $\ln n$ for approximating set cover. *Journal of the ACM (JACM)*, 45(4):634–652, 1998.

Arthur M Geoffrion. Objective function approximations in mathematical programming. *Mathematical Programming*, 13(1):23–37, 1977.

Guillaume Goisque. *Optimization methods for multi-level lot-sizing problems*. PhD thesis, Université de Lorraine, 2017.

James Hill and Michael Galbreth. A heuristic for single-warehouse multiretailer supply chains with all-unit transportation cost discounts. *European Journal of Operational Research*, 187(2):473–482, 2008.

Weihong Hu. *Modeling and solution of some multi-period supply chain optimization problems*. PhD thesis, Georgia Institute of Technology, 2016.

Hark-Chin Hwang. An efficient procedure for dynamic lot-sizing model with demand time windows. *Journal of Global Optimization*, 37(1):11–26, 2007.

Wikrom Jaruphongsa and Chung-Yee Lee. Dynamic lot-sizing problem with demand time windows and container-based transportation cost. *Optimization Letters*, 2(1):39–51, 2008.

Chung-Yee Lee, Sila Çetinkaya, and Albert PM Wagelmans. A dynamic lot-sizing model with demand time windows. *Management Science*, 47(10): 1384–1395, 2001.

Retsef Levi, Robin O Roundy, and David B Shmoys. Primal-dual algorithms for deterministic inventory problems. *Mathematics of Operations Research*, 31(2):267–284, 2006.

Retsef Levi, Andrea Lodi, and Maxim Sviridenko. Approximation algorithms for the capacitated multi-item lot-sizing problem via flow-cover inequalities. *Mathematics of Operations Research*, 33(2):461–474, 2008.

Chung-Lun Li, Vernon Ning Hsu, and Wen-Qiang Xiao. Dynamic lot sizing with batch ordering and truckload discounts. *Operations Research*, 52(4): 639–654, 2004.

Chung-Lun Li, Jinwen Ou, and Vernon N Hsu. Dynamic lot sizing with all-units discount and resales. *Naval Research Logistics (NRL)*, 59(3-4):230–243, 2012.

Steven A Lippman. Optimal inventory policy with multiple set-up costs. *Management Science*, 16(1):118–138, 1969.

RR Meyer. Two-segment separable programming. *Management Science*, 25 (4):385–395, 1979.

Jinwen Ou. Improved exact algorithms to economic lot-sizing with piecewise linear production costs. *European Journal of Operational Research*, 256(3): 777–784, 2017a.

Jinwen Ou. A polynomial time algorithm to the economic lot sizing problem with constant capacity and piecewise linear concave costs. *Operations Research Letters*, 45(5):493–497, 2017b.

Yves Pochet and Laurence A Wolsey. Lot-sizing with constant batches: Formulation and valid inequalities. *Mathematics of Operations Research*, 18(4): 767–785, 1993.

Wilco Van Den Heuvel and Albert PM Wagelmans. A holding cost bound for the economic lot-sizing problem with time-invariant cost parameters. *Operations Research Letters*, 37(2):102–106, 2009.

CPM Van Hoesel and Albert PM Wagelmans. Fully polynomial approximation schemes for single-item capacitated economic lot-sizing problems. *Mathematics of Operations Research*, 26(2):339–357, 2001.

Albert Wagelmans, Stan Van Hoesel, and Antoon Kolen. Economic lot sizing: an $\mathcal{O}(n \log n)$ algorithm that runs in linear time in the wagner-whitin case. *Operations Research*, 40(1-supplement-1):S145–S156, 1992.

Harvey M Wagner and Thomson M Whitin. Dynamic version of the economic lot size model. *Management science*, 5(1):89–96, 1958.