INFORMS Journal on Computing
Vol. 28, No. 4, Fall 2016, pp. 766–780
ISSN 1091-9856 (print) | ISSN 1526-5528 (online)
http://dx.doi.org/10.1287/ijoc.2016.0712

Local Cuts and Two-Period Convex Hull Closures for Big-Bucket Lot-Sizing Problems

Kerem Akartunalı
Department of Management Science, University of Strathclyde, Glasgow G4 0GE, Scotland, kerem.akartunali@strath.ac.uk
Ioannis Fragkos
Technology and Operations Management, Rotterdam School of Management, Rotterdam 3062, Netherlands, fragkos@rsm.nl
Andrew J. Miller
Advanced Analytics Manager, UPS, Atlanta, Georgia 30328, foresomenteneikona@gmail.com
Tao Wu
Advanced Analytics Department, Dow Chemical, Midland, Michigan 48642, danielwu9999@gmail.com

Despite the significant attention they have drawn, big-bucket lot-sizing problems remain notoriously difficult to solve. Previous literature contained results (computational and theoretical) indicating that what makes these problems difficult are the embedded single-machine, single-level, multiperiod submodels. We therefore consider the simplest such submodel, a multi-item, two-period capacitated relaxation. We propose a methodology that can approximate the convex hulls of all such possible relaxations by generating violated valid inequalities. To generate such inequalities, we separate two-period projections of fractional linear programming solutions from the convex hulls of the two-period closure we study. The convex hull representation of the two-period closure is generated dynamically using column generation. Contrary to regular column generation, our method is an outer approximation and can therefore be used efficiently in a regular branch-and-bound procedure. We present computational results that illustrate how these two-period models could be effective in solving complicated problems.

Keywords: lot-sizing; integer programming; local cuts; convex hull closure; quadratic programming; column generation
History: Accepted by Karen Aardal, Area Editor for Design and Analysis of Algorithms; received February 2015; revised October 2015; accepted March 2016. Published online October 12, 2016.

1. Introduction

Lot-sizing is an important part of the planning process in many manufacturing environments. It has therefore been the subject of extensive study by researchers and practitioners for decades. Since the seminal paper of Wagner and Whitin (1958) addressing the simplest version of the problem—the uncapacitated single-item lot-sizing problem—various types of lot-sizing problems have been investigated. Only some special cases of these problems can be solved in polynomial time (e.g., Zangwill 1969, Florian and Klein 1971, Federgruen and Tzur 1991), and even the capacitated version with a single item is $\mathcal{NP}$-hard (Florian et al. 1980).

Solution approaches for lot-sizing problems have varied from heuristic methods to exact approaches based on mathematical programming. A variety of heuristics can be found in Stadtler (2003), Pochet and Van Vyve (2004), Federgruen et al. (2007), and Akartunalı and Miller (2009). Mathematical programming approaches have mainly involved adding valid inequalities (e.g., Barany et al. 1984; Constantino 1996; Pochet and Wolsey 1988, 1994) and extended reformulations of the problem (e.g., Krarup and Bilde 1977, Eppen and Martin 1987, Rardin and Wolsey 1993), although few studies facilitate other techniques, such as Lagrangian relaxation (e.g., Billington et al. 1986) and Dantzig-Wolfe decomposition (e.g., Bitran and Matsuo 1986, Degraeve and Jans 2007). The interested reader is also referred to Belvaux and Wolsey (2001) for modeling and reformulation issues and to Pochet and Wolsey (2006) for an excellent, thorough review of lot-sizing problems and solution methods used.

In spite of this extensive research, the mathematical programming community has focused mainly on single-item problems, and results for multi-item problems are rather limited. The research in Pochet and Wolsey (1991) and Belvaux and Wolsey (2000) extends some of the single-item problem results to multi-item problems, and the recent studies of Anily et al. (2009) and Levi et al. (2008) provide insight into some versions of capacitated multi-item problems. Most recently, some insightful polyhedral results on multilevel relaxations, such as the valid inequalities of Zhang et al. (2012) and the compact formulations of Van Vyve et al. (2014), have shown great potential for problems with *small-bucket* capacities, i.e., items that do not share resources. However, even these references do not explicitly address the structural complications caused by the presence of multiple items competing for limited capacity. Research that explicitly analyzes this structure is limited, and, to the best of our knowledge, includes Miller et al. (2000, 2003), Jans and Degraeve (2004), and Van Vyve and Wolsey (2006).

Previous computational results in the literature have indicated high duality gaps for *big-bucket* lot-sizing problems; i.e., multiple items share the same resource, even though some strategies can be partially efficient for generating lower bounds and feasible solutions. Some notable decompositions, such as the simultaneous period and item decomposition of Pimentel et al. (2010) and the single-period decomposition of the shortest path formulation of Jans and Degraeve (2004), show evidence that considering period-based decompositions, in addition to item-based decompositions, can lead to improved lower bounds. The study by Akartunalı and Miller (2012) has provided important insights about why big-bucket lot-sizing problems are still very hard to solve. More specifically, better approximations for the convex hull of the *single-machine, single-level, multi-period capacitated* problems are necessary to get better results on general lot-sizing problems. In this paper, we investigate the potentials of the simplest such model, a relaxation of a two-period model. We propose a methodology that exactly separates over the convex hull of this model by dynamically generating extreme points of the hull. It is important to note that, to the best of our knowledge, the structure of these subproblems has not been investigated yet, so our computational framework can give insights toward characterizing certain classes of valid inequalities.

The work of Atamtürk and Muñoz (2004) formulated the single-item capacitated lot-sizing problem as a bottleneck flow network problem, enabling the authors to define a rich family of facet-defining inequalities for this problem. The specific two-period relaxation that we exploit can be seen as a multi-item extension of the bottleneck flow problem. It can also be seen as the intersection of two mixed knapsack constraints (the capacity constraints) linked by the demand and inventory for each item. For these and other reasons, the polyhedral structure of this model is, in general, rich and complicated. However, solving such small problems to optimality (i.e., solving the pricing problem in our framework) is computationally tractable, as attested by authors who have used such submodels in primal heuristics (e.g., Stadtler 2003, Federgruen et al. 2007, Akartunalı and Miller 2009). In this paper, although we do not characterize new families of inequalities, the methodology we develop is capable of providing information concerning how effective such results could be.

In the last 15 years, a number of researchers have investigated the "closures" of general cutting planes and some particular polyhedra (e.g., Letchford 2001, Andersen et al. 2005, Dash et al. 2010, Balas and Margot 2013). Even partially achieving some elementary closures has helped researchers be able to close duality gaps efficiently and solve some problems that were never solved before (Fischetti and Lodi 2005, Balas and Saxena 2008). The term "closure" can be defined as the smallest possible polyhedron that satisfies all the valid inequalities of a given type. In our framework, we generate all violated valid inequalities for each two-period relaxation using the characteristics of the convex hull of the two-period relaxation in question (rather than using predefined families of valid inequalities). Applying this procedure to all possible two-period relaxations of a given problem, we approximate the "two-period convex hull closure," which is the intersection of the convex hulls of all possible two-period relaxations. We note that column generation is used to generate the extreme points of these two-period relaxation convex hulls, and Farkas' lemma provides a proof of validity of these cutting planes.

To the best of our knowledge, such a framework has not been used before to strengthen the formulation of lot-sizing problems. There have been a few relevant approaches for generating cutting planes for other problems: in the work of Ralphs et al. (2003), violated inequalities for capacitated vehicle-routing problems (VRPs) are generated using submodels based on small traveling salesman problem (TSP) instances, where the extreme points of these small TSP polyhedrons are generated using column generation. In Applegate et al. (2003), "local cuts" are defined as mapping a fractional solution into a lower dimension and searching for a cut separating it, and this is applied to TSP instances using the so-called "tangled tours." The work of Fukasawa and Goycoolea (2011) employed a subroutine in which localized inequalities are mapped to the original space when generating exact mixed-integer knapsack cuts. A more general approach applicable to mixed-integer programming (MIP) problems is first suggested by Boyd (1994), and the more recent work of Chvátal et al. (2013) extended the concept of "local cuts" to general MIP problems through a sophisticated methodology including tilting the cuts to increase their effectiveness and addressing some of the issues inherent in the precision of coefficients.

We continue this line of research by investigating how efficient a local cuts approach is in the context of multi-item capacitated lot-sizing problems. Contrary to earlier works, we do not try to generate known inequalities, but rather consider a relaxation of a two-period substructure whose polyhedral characterization is not known yet. Our computational study sheds light on the strength of cuts generated by two-period relaxations and paves the way toward their integration in an automated framework.

In the next section, we present the formulation for the multi-item, big-bucket lot-sizing problem. In Section 3, we give a detailed overview of the two-period convex hull closure methodology, including a discussion of the strength of cuts generated. In Section 4, we discuss how to define two-period relaxations in case of a multiperiod problem. Then we present computational results varying from simple two-period problems to more general test problems from the literature. We conclude with a discussion of possible extensions and generalizations.

2. Problem Formulation

We consider the general multi-item lot-sizing problem, with the objective of minimizing the total cost by abiding by big-bucket capacity limitations and demand satisfaction. The decisions to be made for a production plan consist of production and inventory quantities in each period, as well as setup decisions. Next, we present our notation.

Indices and Sets:
NT: Number of periods
NI: Number of items
NK: Number of machines

Variables:
x_t^i: Production quantity of item $i \in \{1, \ldots, NI\}$ in period $t \in \{1, \ldots, NT\}$
y_t^i: Setup of item $i \in \{1, \ldots, NI\}$ in period $t \in \{1, \ldots, NT\}$ ($=1$ if production occurs, $=0$ otherwise)
s_t^i: Inventory held of item $i \in \{1, \ldots, NI\}$ at the end of period $t \in \{1, \ldots, NT\}$

Parameters:
f_t^i: Fixed cost per setup of item $i \in \{1, \ldots, NI\}$ in period $t \in \{1, \ldots, NT\}$
h_t^i: Holding cost per unit of item $i \in \{1, \ldots, NI\}$ from period $t \in \{1, \ldots, NT\}$ to period $t+1$
d_t^i: Demand for item $i \in \{1, \ldots, NI\}$ in period $t \in \{1, \ldots, NT\}$
$d_{t,t'}^i$: Total demand from period $t \in \{1, \ldots, NT\}$ to $t' \in \{t, \ldots, NT\}$, i.e., $d_{t,t'}^i = \sum_{\bar{t}=t}^{t'} d_{\bar{t}}^i$
a_k^i: Processing time per item $i \in \{1, \ldots, NI\}$ on machine $k \in \{1, \ldots, NK\}$
ST_k^i: Setup time for item $i \in \{1, \ldots, NI\}$ on machine $k \in \{1, \ldots, NK\}$
C_t^k: Capacity of machine $k \in \{1, \ldots, NK\}$ in period t

Then the formulation of the problem is as follows:

$$\min \sum_{t=1}^{NT}\sum_{i=1}^{NI} f_t^i y_t^i + \sum_{t=1}^{NT}\sum_{i=1}^{NI} h_t^i s_t^i \tag{1}$$

$$\text{s.t.} \quad x_t^i + s_{t-1}^i - s_t^i = d_t^i, \quad t \in \{1, \ldots, NT\}, i \in \{1, \ldots, NI\}; \tag{2}$$

$$\sum_{i=1}^{NI} (a_k^i x_t^i + ST_k^i y_t^i) \le C_t^k, \quad t \in \{1, \ldots, NT\}, k \in \{1, \ldots, NK\}; \tag{3}$$

$$x_t^i \le M_t^i y_t^i, \quad t \in \{1, \ldots, NT\}, i \in \{1, \ldots, NI\}; \tag{4}$$

$$y \in \{0,1\}^{NT \times NI}; \; x, s \ge 0. \tag{5}$$

The constraints (2) are production balance equations for all items. The constraints (3) are the big-bucket capacity constraints, and constraints (4) guarantee the setup variable set to 1 whenever production is positive, where M_t^i represents maximum numbers of item i that can be produced in period t. Finally, constraints (5) provide the integrality and nonnegativity requirements. We assume that each item is processed by one preassigned machine. We also note that this formulation can be easily extended to problems with multiple levels using echelon demands and stock variables (see, e.g., Pochet and Wolsey 2006); however, for the sake of simplicity, we present this single-level problem with multiple machines instead.

3. Separation over the Two-Period Convex Hull

In this section, we first explain our proposed framework conceptually. A detailed description, along with the theoretical results that prove the validity of the framework, follows. We elaborate on the use of column generation and the nonconventional way that it is used in our framework.

3.1. Overall View of the Framework

First, we define the function ϕ: $\{1, \ldots, NT-1\} \times \{1, \ldots, NI\} \to \{2, \ldots, NT\}$, which, for a given period $t \in \{1, \ldots, NT-1\}$ and item $i \in \{1, \ldots, NI\}$, indicates a forward period $\phi(i,t) > t$. Equipped with function ϕ, we define $\tilde{d}_{t'}^i := \sum_{u=t'}^{\phi(i,t)} d_{t'}^i$ for each $t' \in \{t, t+1\} \subseteq \{1, \ldots, NT\}$. Then we define $X_{t\phi}^{2PL}$, the feasible region of a two-period relaxation, defined over periods t and $t+1$:

$$x_{t'}^i \le M_{t'}^i y_{t'}^i, \quad i \in \{1, \ldots, NI\}, t' \in \{t, t+1\}; \tag{6}$$

$$x_{t'}^i \le \tilde{d}_{t'}^i y_{t'}^i + s_{\phi(i,t)}^i, \quad i \in \{1, \ldots, NI\}, t' \in \{t, t+1\}; \tag{7}$$

$$x_t^i + x_{t+1}^i \le \tilde{d}_t^i y_t^i + \tilde{d}_{t+1}^i y_{t+1}^i + s_{\phi(i,t)}^i, \quad i \in \{1, \ldots, NI\}; \tag{8}$$

$$x_t^i + x_{t+1}^i \le \tilde{d}_t^i + s_{\phi(i,t)}^i, \quad i \in \{1, \ldots, NI\}; \tag{9}$$

$$\sum_{i=1}^{NI}(a^i x^i_{t'} + ST^i y^i_{t'}) \le C_{t'}, \quad t' \in \{t, t+1\}; \tag{10}$$

$$x, s \ge 0, y \in \{0,1\}^{2\times NI}. \tag{11}$$

For notational brevity, we will drop the dependency on t and ϕ and use the standard notation of $\text{conv}(X^{2PL})$ in the remainder of the paper to indicate the convex hull of the extreme points and extreme rays of X^{2PL}. Likewise, we use s^i to refer to the inventory variable of item i at the end of time period $\phi(i,t)$. We note that this is a valid relaxation of any two-period subproblem (rather than an exact formulation of it), where production balance Equations (2) are replaced by (7)–(9). Since the production balance constraints are omitted, there is no need to define an intermediate inventory variable s^i_t. Furthermore, note that since we are looking at a single-machine problem, we omitted the subscripts k representing machines. We also note that each item has an inventory variable of a period represented by $\phi(i,t)$, which can be different from item to item. The demand parameter is defined in a similar fashion to the original parameters, with $\tilde{d}^i_t$ representing the cumulative demand between periods t and $\phi(t,i)$, inclusive. A discussion on how to define $\phi(i,t)$ can be found in Section 4.

One can easily observe the similarity between the constraints of X^{2PL} and of the original lot-sizing problem, noting that constraints (7) and (8) are simply the (ℓ, S) inequalities of Barany et al. (1984), which can be defined in general form as follows:

$$\sum_{t\in S} x^i_t \le \sum_{t\in S} d^i_{t,\ell} y^i_t + s^i_\ell,$$
$$\ell \in \{1,\dots,NT\},\ i \in \{1,\dots,NI\},\ S \subseteq \{1,\dots,\ell\}. \tag{12}$$

This formulation is a multi-item extension of the bottleneck flow formulation studied by Atamtürk and Muñoz (2004) when $NT = 2$. It also extends the single-period study of Miller et al. (2000, 2003). Next, we remark the following basic polyhedral property of $\text{conv}(X^{2PL})$.

Proposition 1. *Without loss of generality, we make the following assumptions for X^{2PL}:*
1. *$0 < M^i_{t'}$ for all $i \in \{1,\dots,NI\}$, $t' \in \{t,t+1\}$,*
2. *$ST^i < C_{t'}$ for all $i \in \{1,\dots,NI\}$, $t' \in \{t,t+1\}$.*

Then $\text{conv}(X^{2PL})$ is full-dimensional.

The proof for this proposition is straightforward and is hence omitted. It is trivial to note that if either assumption is not satisfied, one can simply remove the associated setup and production variable from the problem.

Column generation is used to generate the favorable extreme points of $\text{conv}(X^{2PL})$, since the number of extreme points can grow exponentially. Using these favorable extreme points, we check whether a given fractional solution can be written as their convex combination. If not, we can generate a valid inequality using a theory based on Farkas' lemma that cuts off the fractional point. This cut approximates the convex hull closure of this two-period relaxation.

One important point is that this framework is not based on predefining a family of valid inequalities, which is one of its advantages. An inequality will be generated in all cases when the fractional solution is not in the convex hull of a two-period relaxation. This is also why we expect this framework to provide an adequate approximation of the bottleneck of the general lot-sizing problems, as this focuses on the capacitated single-machine problems with an approach providing exact solutions for the subproblems. Next, we describe more details of our methodology, including the key theoretical results, and elaborate on important details.

3.2. Details of the Cut-Generation Methodology

To describe the methodology, first we let $(\bar{x}, \bar{y}, \bar{s}) \in \mathbb{R}^{5\times NI}_+$ be any point (e.g., it can be a projection of a solution obtained from the LPR of X^{2PL}). We use the infinity-norm distance ($\mathcal{L}_\infty$) of this point to $\text{conv}(X^{2PL})$ to define the *distance problem* as follows:

$$\min \quad z_\infty \tag{13}$$

$$\text{s.t.} \quad \sum_k \lambda_k (x_k)^i_{t'} - z_\infty \le \bar{x}^i_{t'},$$
$$i \in \{1,\dots,NI\},\ t' \in \{t,t+1\} \quad (\alpha^{-i}_{t'}); \tag{14}$$

$$\bar{x}^i_{t'} \le \sum_k \lambda_k (x_k)^i_{t'} + z_\infty,$$
$$i \in \{1,\dots,NI\},\ t' \in \{t,t+1\} \quad (\alpha^{+i}_{t'}); \tag{15}$$

$$\sum_k \lambda_k (y_k)^i_{t'} - z_\infty \le \bar{y}^i_{t'},$$
$$i \in \{1,\dots,NI\},\ t' \in \{t,t+1\} \quad (\beta^{-i}_{t'}); \tag{16}$$

$$\bar{y}^i_{t'} \le \sum_k \lambda_k (y_k)^i_{t'} + z_\infty,$$
$$i \in \{1,\dots,NI\},\ t' \in \{t,t+1\} \quad (\beta^{+i}_{t'}); \tag{17}$$

$$\sum_k \lambda_k (s_k)^i - z_\infty \le \bar{s}^i, \quad \forall i \quad (\gamma^i); \tag{18}$$

$$\sum_k \lambda_k \le 1, \quad (\eta); \tag{19}$$

$$\lambda_k \ge 0,\ z_\infty \ge 0. \tag{20}$$

Note that (x_k, y_k, s_k) is the vector representing the kth extreme point of $\text{conv}(X^{2PL})$. Variable z_∞ represents the ∞-norm distance, which measures the distance between two points as the biggest absolute deviation across their coordinates. Variables λ are multipliers used for the convex combination of the extreme points, where convexity is assured by (19). Note that in (19) we use an inequality rather than an equality for a simpler discussion in the remainder of this section, where the inequality is indeed valid due to the fact that the origin is an extreme point of $\text{conv}(X^{2PL})$, and hence any point satisfying this inequality will also satisfy an equality. Also note that the formulation above has the associated dual variables written next to all constraints in the parentheses to assist explanations in the forthcoming discussion. We have only one set of inequalities for s variables because of the following property.

Proposition 2. $\text{conv}(X^{2PL})$ *has NI extreme rays, each of which for an* $i \in \{1, \ldots, NI\}$ *is in the form* $s^i = 1$, $s^j = 0$, $x = 0$, $y = 0\ \forall j \neq i$.

This property ensures that any point in $\text{conv}(X^{2PL})$ is indeed written as a convex combination of its extreme points and a conical combination of its extreme rays. This distance problem is always feasible, since we can assign 0 to all λ variables and take $z_\infty = \max_{i,t'}(x^i_{t'}, y^i_{t'})$. Moreover, the problem is bounded since $z_\infty \geq 0$. Therefore, we will always have an optimal solution (as well as an optimal dual solution).

If the optimal solution of this problem for a given $(\bar{x}, \bar{y}, \bar{s})$ has an objective function value $z^*_\infty = 0$, then we know that $(\bar{x}, \bar{y}, \bar{s}) \in \text{conv}(X^{2PL})$, since this point is simply written as a convex combination of the extreme points and a conical combination of its extreme rays. However, if $z^*_\infty > 0$, then $(\bar{x}, \bar{y}, \bar{s}) \notin \text{conv}(X^{2PL})$, and this allows us to generate a valid inequality to cut off the fractional point, as stated in the next theorem. We next present the dual of the distance problem for an easier understanding of the forthcoming results:

$$\max \sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\left(\bar{x}^i_{t'}({\alpha^+}^i_{t'} + {\alpha^-}^i_{t'}) + \bar{y}^i_{t'}({\beta^+}^i_{t'} + {\beta^-}^i_{t'})\right) + \sum_{i=1}^{NI}\bar{s}^i\gamma^i + \eta \tag{21}$$

$$\text{s.t.} \sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\left((x_k)^i_{t'}({\alpha^+}^i_{t'} + {\alpha^-}^i_{t'}) + (y_k)^i_{t'}({\beta^+}^i_{t'} + {\beta^-}^i_{t'})\right) + \sum_{i=1}^{NI}(s_k)^i\gamma^i + \eta \leq 0 \quad \forall k; \tag{22}$$

$$\sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\left({\alpha^+}^i_{t'} - {\alpha^-}^i_{t'} + {\beta^+}^i_{t'} - {\beta^-}^i_{t'}\right) - \sum_{i=1}^{NI}\gamma^i \leq 1 \tag{23}$$

$$\alpha^+ \geq 0,\ \beta^+ \geq 0,\ \alpha^- \leq 0,\ \beta^- \leq 0,\ \gamma \leq 0,\ \eta \leq 0. \tag{24}$$

Theorem 1. *Let* $z^*_\infty > 0$ *for* $(\bar{x}, \bar{y}, \bar{s})$, *and* $(\alpha^{*+}, \alpha^{*-}, \beta^{*+}, \beta^{*-}, \gamma^*, \eta^*)$ *be the optimal dual values. Then*

$$\sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\left(({\alpha^{*+}}^i_{t'} + {\alpha^{*-}}^i_{t'})x^i_{t'} + ({\beta^{*+}}^i_{t'} + {\beta^{*-}}^i_{t'})y^i_{t'}\right) + \sum_{i=1}^{NI}\gamma^{*i}s^i + \eta^* \leq 0 \tag{25}$$

is a valid inequality for $\text{conv}(X^{2PL})$ *that cuts off* $(\bar{x}, \bar{y}, \bar{s})$.

Proof. The validity of (25) follows from the fact that (22) is valid for every extreme point of $\text{conv}(X^{2PL})$ and that $\gamma \leq 0$ holds. In contrast, since $z^*_\infty > 0$, the violation follows simply from the optimal value of (21) being strictly positive for $(\bar{x}, \bar{y}, \bar{s})$. □

As we show in the next section, we only generate a small subset of the extreme points to ensure that our approach is computationally efficient. The general framework of the separation procedure over the convex hull of the two-period model is summarized in Algorithm 1. It is worth noting that extreme points are added dynamically, and only while z_P, the objective function of the column generation subproblem, is negative. The next part elaborates further on the column generation procedure.

Algorithm 1 (Two-period separation algorithm)
Input: A point $(\bar{x}, \bar{y}, \bar{s})$; a 2-period problem X^{2PL}; $\epsilon \geq 0$
Output: A cutting plane or inclusion certificate

```
repeat
    Solve the distance problem for conv(X^{2PL});
    if z_∞ ≤ ε, then
        break
    else
        Solve column generation problem;
        if z_P ≤ 0, then
            break
        else
            Add new extreme point
        end
    end
until z_∞ ≤ ε or z_P ≤ 0;
if z_∞ ≤ ε, then
    (x̄, ȳ, s̄) ∈ conv(X^{2PL})
else
    Add the violated cut (25)
else.
```

3.3. Column Generation

All but one of the variables of the distance problem are associated with an extreme point of $\text{conv}(X^{2PL})$. Since the number of extreme points is exponential to the problem inputs, but only a small subset of their corresponding variables is basic at an optimal solution of the distance problem, we use column generation to generate the favorable extreme points, as explained below. Recall that there are only as many extreme rays as the number of items.

When we solve the distance problem with only a subset of extreme points, we obtain a dual optimal solution $(\alpha^{*+}, \alpha^{*-}, \beta^{*+}, \beta^{*-}, \gamma^*, \eta^*)$. For any λ_k variable associated with the extreme point (x_k, y_k, s_k), the reduced cost is defined as follows:

$$\sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\left((x_k)^i_{t'}({\alpha^{*+}}^i_{t'} + {\alpha^{*-}}^i_{t'}) + (y_k)^i_{t'}({\beta^{*+}}^i_{t'} + {\beta^{*-}}^i_{t'})\right) + \sum_{i=1}^{NI}(s_k)^i\gamma^{*i} + \eta^*. \tag{26}$$

Note that for all the extreme points added so far, (26) is less than or equal to 0. If this condition holds for all the extreme points not yet included in the problem, there does not exist any extreme point that will improve the solution of the previous distance problem; hence the current solution is optimal. Therefore, we define the following *pricing problem*:

$$\begin{aligned} \max \quad z_P = & \sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\big(x_{t'}^i(\alpha^{*+}{}_{t'}^i+\alpha^{*-}{}_{t'}^i)+y_{t'}^i(\beta^{*+}{}_{t'}^i+\beta^{*-}{}_{t'}^i)\big) \\ & +\sum_{i=1}^{NI} s^i\gamma^{*i}+\eta^* \\ \text{s.t.} \quad & (x,y,s)\in X^{2PL}. \end{aligned}$$

Corollary 1. *If the optimal value $z_P^* \leq 0$, then the solution of the distance problem is optimal. Otherwise, the optimal (x, y, s) values should be added as a new column to the distance problem.*

Note that the pricing problem is an MIP, which, because of its small size, can be solved to optimality efficiently. However, it may still be helpful to ensure that the method converges as fast as possible to the real distance, i.e., that the *number* of generated extreme points does not grow unnecessarily large, especially when the sequence of distance values converges very slowly. Therefore, it would be beneficial to terminate the column generation prematurely, especially if a cut can be generated (even weaker than the original one). The following result, adapted from Theorem 1, is crucial for this computational aspect.

Corollary 2. *Let $z_\infty^* > 0$ for $(\bar{x}, \bar{y}, \bar{s})$, $(\alpha^{*+}, \alpha^{*-}, \beta^{*+}, \beta^{*-}, \gamma^*, \eta^*)$ be the optimal dual values, and $z_P^* > 0$. Then,*

$$\begin{aligned} &\sum_{i=1}^{NI}\sum_{t'=t}^{t+1}\big((\alpha^{*+}{}_{t'}^i+\alpha^{*-}{}_{t'}^i)x_{t'}^i+(\beta^{*+}{}_{t'}^i+\beta^{*-}{}_{t'}^i)y_{t'}^i\big) \\ &\quad +\sum_{i=1}^{NI}\gamma^{*i}s^i+\eta^*\leq z_P^* \end{aligned} \tag{27}$$

is a valid inequality for $\text{conv}(X^{2PL})$.

Proof. Note that (27) holds for every extreme point of $\text{conv}(X^{2PL})$, since z_P^* is the maximum value attained by any extreme point. Since $\eta < 0$ holds, any point of $\text{conv}(X^{2PL})$ written as a convex combination of extreme points and a conical combination of its extreme rays will satisfy this inequality. □

We also note that by using the reduced cost information, and because $\sum_k \lambda_k \leq 1$, the distance function value can be at most reduced by z_P in each iteration. We conclude this section with the note that this "reduced cost cut" is implemented in our computational tests for better efficiency.

3.4. Alternative Distance Functions

Here we discuss the use of alternative norms instead of $\mathscr{L}_\infty$. The first and obvious candidate is the Manhattan distance (or $\mathscr{L}_1$), since it can be linearly modelled; this is discussed in detail in Akartunalı (2007). Formulating an $\mathscr{L}_1$-based distance problem is straightforward; for the sake of brevity the particular details are omitted here. However, we consider $\mathscr{L}_1$-based distance problems in our computational tests for the sake of completeness.

Next, we discuss how to use Euclidean distance, i.e., 2-norm or $\mathscr{L}_2$, in our framework. The main motivation for using the Euclidean distance is that it has a faster convergence rate than the linear approach of Manhattan distance, when a sequence of points is expected to converge to a specific point (in our case, this sequence of points consists of the closest point of $\text{conv}(X^{2PL})$ to $(\bar{x}, \bar{y}, \bar{s})$ in each iteration of the algorithm, since the more extreme points are added with column generation, the more we converge to the real distance).[1] In addition, contrary to the $\mathscr{L}_\infty$-based distance formulation discussed earlier, the minimized Euclidean objective involves the individual distance variables associated with each element of the $(\bar{x}, \bar{y}, \bar{s})$ vector. As a result, the optimal solutions have often more binding constraints than the $\mathscr{L}_\infty$ problem, for which an optimal solution with one binding constraint always exists. This implies that the cuts generated from the Euclidean formulation are likely to be more dense that those produced by the $\mathscr{L}_\infty$ formulation. On a more practical note, it is also important to remark that the quadratic programming (QP) solvers have achieved significant developments similar to linear programming (LP) and IP solvers, which allow fast solutions. For the LP relaxation (LPR) solution $(\bar{x}, \bar{y}, \bar{s})$, we define the Euclidean distance problem as follows:

$$\min_{\Delta,\lambda} \quad z_2=\sum_i\bigg([(\Delta_s)^i]^2+\sum_{t'=t}^{t+1}([(\Delta_x)_{t'}^i]^2+[(\Delta_y)_{t'}^i]^2)\bigg) \tag{28}$$

$$\text{s.t.} \quad \bar{x}_{t'}^i=\sum_k\lambda_k(x_k)_{t'}^i+(\Delta_x)_{t'}^i, \quad i\in\{1,\dots,NI\},\ t'\in\{t,t+1\} \quad (\alpha_{t'}^i); \tag{29}$$

$$\bar{y}_{t'}^i=\sum_k\lambda_k(y_k)_{t'}^i+(\Delta_y)_{t'}^i, \quad i\in\{1,\dots,NI\},\ t'\in\{t,t+1\} \quad (\beta_{t'}^i); \tag{30}$$

$$\bar{s}^i\geq\sum_k\lambda_k(s_k)^i-(\Delta_s)^i, \quad \forall i \quad (\gamma^i); \tag{31}$$

$$\sum_k\lambda_k\leq 1, \quad (\eta); \tag{32}$$

$$\lambda_k\geq 0,\ \Delta_s\geq 0,\ \Delta_x,\ \Delta_y, \quad \text{free}. \tag{33}$$

The distance variables Δ_x and Δ_y are defined as free, whereas Δ_s variables can be restricted to nonnegative

[1] Personal communication with S. Robinson.

because of the Proposition 2, ensuring that any point in $\mathrm{conv}(X^{2PL})$ can indeed be written as a convex combination of its extreme points and a conical combination of its extreme rays. Dual variables are highlighted in parentheses next to associated constraints. Note that this is a QP problem with linear constraints, and the objective function has quadratic terms with positive coefficients only (i.e., if we write z_2 in the form $\frac{1}{2}x^T Qx$ with x indicating the variable vector, then the matrix Q is positive semidefinite). Therefore, the dual of the Euclidean distance problem can be stated as follows (for QP duality, see, e.g., Mangasarian 1994, pp. 123–124):

$$\max_{\Delta,\alpha,\beta,\gamma,\eta} \; z_D = -\sum_i \left([(\Delta_s)^i]^2 + \sum_{t'=t}^{t+1} [(\Delta_x)^i_{t'}]^2 + [(\Delta_y)^i_{t'}]^2 \right) - \left(\sum_{i=1}^{NI} \sum_{t'=t}^{t+1} (\bar{x}^i_{t'} \alpha^i_{t'} + \bar{y}^i_{t'} \beta^i_{t'}) + \sum_{i=1}^{NI} \bar{s}^i \gamma^i + \eta \right) \quad (34)$$

$$\text{s.t.} \; \sum_{i=1}^{NI} \sum_{t'=t}^{t+1} ((x_k)^i_{t'} \alpha^i_{t'} + (y_k)^i_{t'} \beta^i_{t'}) + \sum_{i=1}^{NI} (s_k)^i \gamma^i + \eta \geq 0, \quad \forall k; \quad (35)$$

$$\alpha^i_{t'} = -2(\Delta_x)^i_{t'}, \; \beta^i_{t'} = -2(\Delta_y)^i_{t'}, \; -\gamma^i \geq -2(\Delta_s)^i, \quad \forall i, t'; \quad (36)$$

$$\gamma \geq 0, \eta \geq 0, \Delta_s \geq 0, \alpha, \beta, \Delta_x, \Delta_y, \text{ free.} \quad (37)$$

Next we establish the following theorem, which allows us to generate inequalities if $(\bar{x}, \bar{y}, \bar{s}) \notin \mathrm{conv}(X^{2PL})$.

Theorem 2. *Let $z_2^* > 0$ for $(\bar{x}, \bar{y}, \bar{s})$, with optimal dual values $(\alpha^*, \beta^*, \gamma^*, \eta^*)$. Then*

$$\sum_i \sum_{t'=t}^{t+1} (\alpha^{*i}_{t'} x^i_{t'} + \beta^{*i}_{t'} y^i_{t'}) + \sum_i \gamma^{*i} s^i + \eta^* \geq 0 \quad (38)$$

is a valid inequality for $\mathrm{conv}(X^{2PL})$ that cuts off $(\bar{x}, \bar{y}, \bar{s})$.

Proof. The validity of (38) follows from the fact that (35) is valid for every extreme point of $\mathrm{conv}(X^{2PL})$ and that $\gamma \geq 0$ holds. To observe the violation, let the associated optimal primal values be Δ^*. Then

$$\left(\sum_{i=1}^{NI} \sum_{t'=t}^{t+1} (\bar{x}^i_{t'} \alpha^{*i}_{t'} + \bar{y}^i_{t'} \beta^{*i}_{t'}) + \sum_{i=1}^{NI} \bar{s}^i \gamma^{*i} + \eta^* \right) = -2((\Delta^*_s)^2 + (\Delta^*_x)^2 + (\Delta^*_y)^2) < 0$$

holds, where the equality follows from the QP duality, and the inequality follows from $z_2^* > 0$. Hence, (38) cuts off $(\bar{x}, \bar{y}, \bar{s})$. □

As in the $\mathcal{L}_\infty$ norm case, column generation can be used to generate only a small subset of extreme points, where the pricing problem is stated as follows:

$$\min \; z_P = \sum_{i=1}^{NI} \sum_{t'=t}^{t+1} (\alpha^{*i}_{t'} x^i_{t'} + \beta^{*i}_{t'} y^i_{t'}) + \sum_i \gamma^{*i} s^i + \eta^*$$

$$\text{s.t.} \; (x, y, s) \in X^{2PL}.$$

Corollary 3. *If the optimal value $z_P^* \geq 0$, then the solution of the distance problem is optimal. Otherwise, the optimal (x, y, s) values of the pricing problem indicate a favorable column.*

Note that the framework presented in Algorithm 1 remains valid for 2-norm as well, with the exception of changing the condition $z_P \leq 0$ to $z_P \geq 0$ and replacing z_∞ with z_2. Similarly, a parallel result to Corollary 2 is also easy to extend to the case of z_2. As we illustrate in Section 5 in our computational results, using the Euclidean distance function often seems to result in better convergence than using the infinity norm. Moreover, we make a remark on the recent work of Cadoux (2010), who presented a proof that using 2-norm can provide "deepest disjunctive cuts" (rather than using standard linear norms). Although that work focuses on separating fractional points from a disjunctive polyhedron, this assertion seems in line with our computational experience, where the 2-norm often achieves the fastest converging cuts. However, we also observed that the 2-norm has the biggest potential for causing numerical issues.

We present the following example to illustrate a simple two-dimensional case comparing the use of different norms to generate cuts.

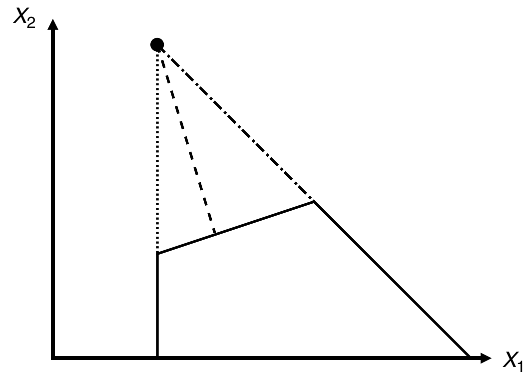

Figure 1 Graphical View of Different Norms (Example 1)

Example 1. Consider the convex hull defined by the corners $(1,0)$, $(1,1)$, $(2.5,1.5)$, and $(4,0)$, as seen in Figure 1. Assume we have the point $\bar{x} = (1,3)$ we want to cut off. Using Manhattan distance, one would obtain minimal distance $z_1 = 2$ with $\Delta = \{0,0,0,2\}$ and the cut $-x_1 + x_2 \leq 0$, which is only a face of the polyhedron. Using $\mathcal{L}_\infty$, the minimal distance is $z_\infty = 1.5$ with $\Delta = \{0, 1.5, 1.5, 0\}$ and we obtain the cut $-0.25x_1 + 0.75x_2 - 0.5 \leq 0$. Finally, using Euclidean distance, we obtain the minimal distance $z_2 = \sqrt{3.6}$ with $\Delta = \{-0.6, 1.8\}$ and the cut $1.2x_1 - 3.6x_2 + 2.4 \geq 0$, which is the same as the cut obtained by infinity norm, i.e., a facet of the polyhedron.

3.5. Strength of Cuts Generated

The distance problems discussed in the previous sections are defined using standard norms, i.e., $\mathscr{L}_2$ and $\mathscr{L}_\infty$. However, a potential disadvantage of the formulation used is that the inequalities that are generated are possibly not facets or even high dimensional faces. If possible, it would be beneficial to define a "distance" problem for which the resulting primal/dual pair is guaranteed to yield faces of high dimension. Therefore, we discuss how to define such a distance problem in this section.

Consider the following "distance" problem (P):

$$\min \; z \tag{39}$$

$$\text{s.t. } \bar{x}^i_{t'} = \sum_k \lambda_k (x_k)^i_{t'}, \quad i \in \{1, \ldots, NI\},\; t' \in \{t, t+1\}, \quad (\alpha^i_{t'}); \tag{40}$$

$$\bar{y}^i_{t'} = \sum_k \lambda_k (y_k)^i_{t'}, \quad i \in \{1, \ldots, NI\},\; t' \in \{t, t+1\}, \quad (\beta^i_{t'}); \tag{41}$$

$$\bar{s}^i \geq \sum_k \lambda_k (s_k)^i - z, \quad \forall\, i, \quad (\gamma^i); \tag{42}$$

$$\sum_k \lambda_k - z \leq 1, \quad (\eta); \tag{43}$$

$$\lambda_k \geq 0,\; z \geq 0. \tag{44}$$

While (P) is similar to the ∞-norm primal distance problem, there are differences as well. In particular, it is not obvious that a feasible solution exists. To address this issue, we define a priori $2NI$ extreme points in the form of $x^i_{t'} = C_{t'} - ST^i$, $y^i_{t'} = 1$, $s^i = (C_{t'} - ST^i - \tilde{d}_{t'})^+$, for each $t' \in \{t, t+1\}$ and $i \in \{1, \ldots, NI\}$, and $2NI$ extreme points in the form $y^i_{t'} = 1$, $x^i_{t'} = 0$, $s^i_{t'} = 0$ for each $t' \in \{t, t+1\}$ and $i \in \{1, \ldots, NI\}$, with all variables not explicitly listed in these extreme points to be 0.

Lemma 1. *Given the $4NI$ extreme points previously defined, (P) is always feasible.*

To prove this condition, it is easy to see that for any given point $(\bar{x}, \bar{y}, \bar{s})$, since $x^i_t \leq M^i_t$ holds, the equalities (40) and (41) will hold with the correct choice of λ variables.

Proposition 3. *When (P) is solved, either $z^* = 0$ or there exists a violated inequality for $(\bar{x}, \bar{y}, \bar{s})$ in the form*

$$\sum_{i=1}^{NI}\sum_{t'=t}^{t+1}((x_k)^i_{t'}\alpha^{*i}_{t'} + (y_k)^i_{t'}\beta^{*i}_{t'}) + \sum_{i=1}^{NI}(s_k)^i\gamma^{*i} + \eta^* \leq 0.$$

The proposition is quite straightforward and can be proved in a similar fashion to previous inequalities, in particular when considering the dual of (P), which we refer to as (D) and present as follows:

$$\max \left\{ \sum_{i=1}^{NI}\sum_{t'=t}^{t+1}(\bar{x}^i_{t'}\alpha^i_{t'} + \bar{y}^i_{t'}\beta^i_{t'}) + \sum_{i=1}^{NI}\bar{s}_k{}^i\gamma^i + \eta \right\} \tag{45}$$

$$\text{s.t. } \sum_{i=1}^{NI}\sum_{t'=t}^{t+1}((x_k)^i_{t'}\alpha^i_{t'} + (y_k)^i_{t'}\beta^i_{t'}) + \sum_{i=1}^{NI}(s_k)^i\gamma^i + \eta \leq 0, \quad \forall\, k; \tag{46}$$

$$-\sum_{i=1}^{NI}\gamma^i - \eta \leq 1, \quad \gamma \leq 0,\; \eta \leq 0. \tag{47}$$

Finally, we conclude this section with the main result regarding the strength of the cuts generated.

Theorem 3. *If $z^* > 0$, then the generated inequality has a dimension of at least $4NI - 1$.*

Proof. We first note that the dual problem has $5NI + 1$ dual variables and its dimension is also $5NI + 1$. Therefore, when this LP is optimized, $5NI + 1$ dual constraints will be satisfied as equality for every extreme point solution. Since $z^* > 0$, the dual constraint $-\sum_{i=1}^{NI}\gamma^i - \eta \leq 1$ is satisfied at equality and therefore, at most NI of γ^i, η variables can be zero (i.e., at least one of them has to be nonzero). Therefore, at least $4NI$ of the dual constraints (46) are satisfied as equation for the optimal solution. Hence, at least $4NI$ extreme points of the primal problem, which are affinely independent, lie on the hyperplane defining the cut. □

4. Defining Two-Period Relaxations from a Multiperiod Problem

In this section, we discuss how X^{2PL} can be used to define two-period relaxations of a generic, multiperiod problem. Considering the lot-sizing problems we have investigated with multiple periods and items, the first decision is at which two periods to run the separation algorithm. For a problem with NT periods, we can look at all the two-period problems; i.e., we can create $NT - 1$ two-period problems and run the separation routine we discussed in the previous sections, which we apply to the remainder of the paper.

Next, we recall that X^{2PL} assigns one inventory variable to each item, namely s^i, which represents the ending inventory of period $\phi(i, t)$, for a given function ϕ. This leads to the question, "Which period's stock is represented by s^i?" As one can easily note, the obvious choice for the horizon parameter would be $\phi(t, i) = t + 1$, for all $i \in \{1, \ldots, NI\}$. The main disadvantage of this strategy is that the demand of later periods is not taken into consideration in the formulation of X^{2PL}. For example, consider a case where the algorithm tries to separate a fractional point in which, for some item i, no production occurs in periods $t + 2, \ldots, l$, for some $l \gg t + 2$. Then the inventory s^i_{t+1} of that fractional point will be large, because it needs to cover the demand of periods $t+2, \ldots, l$, and therefore a subproblem with $\phi(t, i) = t+1$ might not be able to separate that point, because the production variables at the corresponding extreme points do not incorporate the cumulative demand of periods $t+2, \ldots, l$. Therefore, given a fractional point, our intention is to select $\phi(t, i)$ such that the extreme points of the underlying polytope are *dissimilar* to the point we try to separate.

As Miller et al. (2000) noted for their single-period relaxations, one key observation is that if several periods have no setups following the period $t+1$, their demands should be incorporated to obtain the smallest amount of inventory carried from period $t+1$ without weakening the (ℓ, S) inequalities. Another observation is that if a setup occurs in a period after $t+1$, the (ℓ, S) inequalities will be weakened if that period is included in the horizon and hence it should be avoided. Therefore, Miller et al. (2000) propose the following definition of horizon parameter:

$$\phi(t,i) = \max\left\{ u \mid u \geq t+1, \sum_{t''=t+1}^{u} y^i_{t''} \leq y^i_{t+1} + \Theta \right\}, \quad (48)$$

where Θ is a random number between 0 and 1; they argue that this assignment is computationally efficient in the case of their single-period relaxation. In lieu of adopting a randomized approach, we have experimented with different levels of Θ and identified that $\Theta = 0$ generates the deepest cuts. Therefore, we use $\phi(t, i) = \max\{u \mid u \geq t+1, \sum_{t''=t+1}^{u} y^i_{t''} \leq y^i_{t+1}\}$ in our computational experiments. This choice ensures that cumulative demand of later periods with zero production is captured in the extreme points of X^{2PL}_t. One final note is that the values $\phi(t, i)$ are recalculated every time the separation procedure is called, so X^{2PL} is updated both when a new pair of periods is considered and when a new fractional point is at hand. The overall framework for multiperiod problems can be seen in Algorithm 2.

Algorithm 2 (Two-period convex hull closure framework)

Update (ℓ, S) inequalities;
Solve LPR of the original problem;
$\rightarrow (\bar{x}, \bar{y}, \bar{s})$;
for $t = 1$ *to* $t = NT - 1$ **do**
  Define $\phi(t, i)\ \forall i \in \{1, \ldots, NI\}$ and update X^{2PL}_t;
  Apply two-period separation algorithm
    for $(\bar{x}, \bar{y}, \bar{s})$, X^{2PL}_t;
end.

5. Computational Results

In this section, we present our computational experience regarding the two-period convex hull closure framework. All alternative distance approaches discussed earlier are implemented, and FICO® Xpress Optimization Suite (2015 version) is used as the solver. We first present the results for two-period problems and focus on the efficacy of each of the proposed distance norms, and then follow with the results for some multi-period problems, including a discussion of computational issues and considerations. We note that a limited version of preliminary tests were presented in Fragkos and Akartunalı (2014).

5.1. Two-Period Problems

To provide a thorough investigation, we first generated 20 problems with two periods only and with two to six items. The detailed data of the instances of this set, called 2PCLS (2-Period Capacitated Lot-Sizing) can be found in Akartunalı (2007). One of the advantages of having such small problems is that we might actually obtain the full description of the convex hull using software like cdd (Fukuda 2014), which is currently investigated in a companion paper (Doostmohammadi et al. 2016). Another important remark is on the number of items: the more items share a resource, the more the structure of the optimal solutions tends to resemble that of an uncapacitated problem, as noted by many others, including Manne (1957). This is our motivation not to generate problems with too many items.

Next we present results of the 2PCLS instances using three different distance approaches: the L_1, L_2, and L_∞ norms. Table 1 summarizes the results for these instances (I indicates the number of items), where the LP bound obtained by separating the (ℓ, S) inequalities (see, e.g., Pochet and Wolsey 2006) is indicated by XLP, and IP shows the optimal integer solution for different instances. The same 2PL value is attained for all problem instances by all the three approaches, indicating that the 2PL bound closed the gap of all two-period instances that we tested. The number of cuts needed for each different norm (indicated by #C) is also provided for comparison, as well as the average number of columns generated per iteration (indicated by #col).

As the number of cuts indicates, the Euclidean norm is often more efficient than the linear norms, in the sense that it generates a reduced number of cuts, especially for bigger instances. The superiority of the Euclidean norm can also be seen from the rate of convergence; i.e., Euclidean generates approximately half the number of columns per iteration compared to the infinity norm, which is the clearly more efficient linear approach. In addition, it is worth noticing that although X^{2PL} is only a relaxation of a two-period capacitated lot-sizing problem (CLSP), we are able to completely close the gap for the above two-period instances. Finally, all approaches generate more columns on average when the number of items increases, which happens because the number

Table 1 Separation of 2PCLS Instances Using All Three Approaches

					L^1		L^∞		L^2	
Instance	I	XLP	IP	2PL	#C	#col	#C	#col	#C	#col
2pcls01	3	17.033	25	25	11	38.08	8	41.44	19	27.06
2pcls02	3	12.6253	19	19	13	40.29	7	43.88	7	24.38
2pcls03	3	76.5345	104	104	5	51	3	26.25	1	14
2pcls04	2	14.7674	19	19	4	14	2	10	1	25
2pcls05	3	38.39	52	52	8	26.11	6	43.14	4	20.8
2pcls06	3	117.375	173	173	5	26.33	6	43.71	5	20.17
2pcls07	2	36.5	43	43	2	14	1	15	1	7
2pcls08	2	21.45	26	26	7	14.63	2	17.67	2	8.67
2pcls09	2	129	153	153	2	17	3	19.5	3	8.75
2pcls10	3	17.6539	24	24	1	48	3	27.5	1	14
2pcls11	3	71.7209	102	102	4	19.2	1	23.5	1	3.5
2pcls12	3	46.68	69	69	4	12	1	22	2	13.67
2pcls13	4	85.6256	113	113	7	41	7	72.75	9	35.27
2pcls14	4	70.2961	81	81	6	52.71	8	81.89	5	40
2pcls15	4	54.1848	74	74	6	49.43	3	65	1	25.5
2pcls16	4	34.0844	39	39	6	61	6	69.29	4	38.2
2pcls17	5	164.858	211	211	39	79.87	19	135.35	14	63.73
2pcls18	5	57.0825	97	97	34	97.97	10	81.18	6	48.14
2pcls19	6	115.131	150	150	23	96.91	6	105	1	45
2pcls20	6	59.2412	89	89	34	133.03	11	137.83	5	64.5

of extreme points increases exponentially with the number of items.

5.2. Multiperiod Problems

The computational results of the previous section indicate a significant potential for improving the lower bounds of CLSPs. In this section, we demonstrate computationally that the gap closed by the two-period closure algorithm in multiperiod problems can be substantial, and competitive or superior to the gap closed by other state-of-the-art approaches. Similar to other approaches that investigate the lower bound improvement by optimizing over a closure (Balas and Saxena 2008) or, more generally, by employing a computationally heavy algorithm (Bergner et al. 2015), our framework needs to reach further computational maturity until it can be time efficient enough to be embedded in modern solvers. We therefore focus on obtaining the best lower bounds possible, possibly at the expense of CPU times, to gain a thorough understanding of multiperiod, multi-item, single-level, big-bucket relaxations. For the sake of completeness, we report average CPU times for each experiment and detailed computational results in the online supplement (available as supplemental material at http://dx.doi.org/10.1287/ijoc.2016.0712) that accompanies this paper. Although a comparison with other lower bounds found in the literature can indicate the theoretical strength of each methodology, benchmarking CPU times across different implementations has to be taken with a grain of salt, as it might lead to incorrect conclusions. Before presenting numerical results, we first discuss some important implementation details and potential numerical issues pertinent to our approach.

5.2.1. Computational Considerations. The frequent generation of cuts with fractional coefficients that may not have an exact representation in floating point arithmetic can cause numerical issues. This is a pertinent issue to the generation of deep cuts and is also the reason commercial solvers refrain from generating many rounds of MIR cuts (Cook et al. 2009). To circumvent this problem, Chvátal et al. (2013) use the rational solver they developed in Applegate et al. (2007). In addition, they provide a floating point implementation of their method to compare their results with other studies. In this paper we develop a floating point implementation, as the development of an exact rational solver is beyond the scope of our research. This is also in line with our primary aim of this paper, i.e., to show the effectiveness of the cuts generated using the framework.

In experiments with two-period instances it was found that the Euclidean norm L_2 exhibits important numerical issues, e.g., being very sensitive to some control parameters used, although it converges faster than the other norms. In contrast, the L_1 norm exhibits the slowest convergence but has the most stable numerical performance. To strike a balance between computational convergence and numerical accuracy we utilize the L_∞ norm in the remainder of the paper; it has both a fast convergence and overall a stable numerical performance.

5.2.2. Implementation Details. Given a fractional point, we call Algorithm 2 that generates $NT-1$

two-period cuts from subproblems $\{1,2\}, \{2,3\}, \dots, \{NT-1, NT\}$, and then apply the (ℓ, S) inequalities. We iteratively apply Algorithm 2 and (ℓ, S) inequalities until the resulting fractional point can no longer be separated, i.e., the two-period projections of the fractional solution belong to the corresponding two-period closures, or until a time limit is reached. Note that even if some two-period closure does not generate a cut during a particular iteration, it might generate a cut in a subsequent iteration. This is because the definition of X^{2PL} depends on the setup variables of the point to be separated, which is updated after each round of two-period and (ℓ, S) cuts. We have noticed however that the cuts coming from two-period closures that have not generated cuts in previous iterations tend to be weak. Therefore, when a two-period subproblem cannot separate a point we abort the separation subroutine for that particular subproblem. Initially, we solve the column generation subproblems to feasibility instead of optimality and apply cuts in the form (27) instead of (25). When no more cuts can be generated, we solve the subproblems to optimality. This two-mode strategy offers an improved convergence when compared to the textbook column-generation implementation. An advantage of our framework is that we can generate valid cuts without solving the subproblem to optimality.

Regular column-generation algorithms have to add all the columns that price out in order to guarantee that the resulting relaxation is valid. This is because column-generation algorithms work with inner approximations of the relaxed feasible region, whereas cutting planes are outer approximations (Bergner et al. 2015). We also add all subproblem columns that are found to price out in each iteration. Finally, to keep numerical issues to a minimum, we change the default scaling settings of Xpress to include row, column, and Curtis-Reid scaling (Curtis and Reid 1972).

5.2.3. Trigeiro Instances. First we compare the lower bounds obtained by the two-period closure with other approaches using six instances taken from the data set of Trigeiro et al. (1989), which are often used by researchers in the area as benchmark problems. Although a comparison based on six instances offers limited conclusions, the fact that these instances have been widely used (Miller et al. 2000, Jans and Degraeve 2004, Van Vyve and Wolsey 2006, de Araujo et al. 2015) allows us to obtain an indication of how the strength of the two-period closure lower bound compares to that of other approaches.

The results of Table 2 show that the two-period closure can close a considerable amount of gap, especially when it is combined with cuts generated by the Xpress solver. In particular, it seems that the obtained lower bound is stronger when the number of items is small relative to the number of periods (instances G30, G62, and G53). To interpret this finding, we note that a result from Manne (1957) implies that the solution of the LP relaxation of the per-item Dantzig-Wolfe decomposition of CLSP is a good approximation of the optimal solution when the number of items is large compared to the number of capacity constraints. Since the lower bound obtained from the per item decomposition formulation of Manne (1957) and from the use of (ℓ, S) inequalities is the same, the application of (ℓ, S) inequalities in problems with a large number of items leads to an LP relaxation that is a good approximation of the optimal solution. Therefore, separating the two-period projections of a fractional point, which is already a good approximation of the optimal solution, does not improve the lower bound as much as it does in instances with fewer items, where the improvement is more profound. We note that the average CPU time for instances with a small number of items is 333 seconds; this increases considerably as the number of items increases, to 2,772 and 10,800 seconds for 12 and 24 items, respectively.

5.2.4. Süral Instances. Next, we report results on a subset of instances utilized from Süral et al. (2009).

Table 2 Trigeiro Instances: 2PL Results Without (2PL) and with Xpress Cuts (X2PL) Compared to Preceding Inventory (PI) Relaxation of Miller et al. (2000), Dantzig-Wolfe (DW) Decomposition Based on Single-Period Relaxations of Jans and Degraeve (2004), Approximate Extended Formulation with XPRESS Cuts (XAEF) of Van Vyve and Wolsey (2006), and Optimal IP Solutions (OPT)

	G30 (6–15)	G62 (6–30)	G53 (12–15)	G69 (12–30)	G57 (24–15)	G72 (24–30)
XLP	37,201	60,946	73,848	130,177	136,366	287,753
PI	37,319	61,150	73,929	130,292	136,388	287,811
DW	37,382	61,205	73,945	130,338	136,418	287,824
XAEF	37,469	61,294	74,230	130,335	136,417	287,828
2PL	37,329	61,081	74,183	130,251	136,372^t	287,771^t
X2PL	37,432	61,232	74,295	130,337	136,374^t	287,810^t
OPT	37,721	61,746	74,634	130,596	136,509	287,929

Notes. The values of (DW) of Jans and Degraeve (2004) for instances G57 (24–15) and G72 (24–30) are obtained through Lagrangean relaxation. t = Terminated due to time limit. Time limit = 10,800 seconds.

Table 3 **Average Integrality Gap Calculated as** $((z_{\text{UB}} - z_{\text{LB}})/z_{\text{LB}})\%$**, Using the Best Known Upper Bound Across All Three Methods**

Homogeneous					Heterogeneous				
NI–NT	Instances	LR	PD	X2PL	*NI–NT*	Instances	LR	PD	X2PL
12–10	10	42	22	4	12–10	10	29	15	3
12–15	9	29	19	9	12–15	9	22	14	6
12–30	5	24	20	12	12–30	5	20	15	10
24–10	10	20	12	7	24–10	10	16	8	4
24–15	10	20	13	8	24–15	10	18	9	7
24–30	5	30	20	17	24–30	5	31	21	17
Average		28	17	8	Average		22	13	7

Notes. LR denotes the Lagrange relaxation approach of Süral et al. (2009) and PD denotes the period decomposition bound of de Araujo et al. (2015). A time limit of 10,800 seconds was imposed.

The authors constructed new instances by modifying the instances of Trigeiro et al. (1989). In particular, they consider problems without setup costs and divide the data set into instances with unit inventory cost (called homogeneous), and into instances with nonunit inventory cost (called heterogeneous). The integrality gaps reported in their paper are significantly larger than those of the original problems and therefore constitute a good test bed for lower bounding techniques. The lower bounds of Süral et al. (2009) are obtained by solving the Lagrange dual of the facility location formulation using subgradient optimization. We compare the strength of the lower bound obtained by the two-period closure with their approach, and also the period decomposition approach of de Araujo et al. (2015). Table 3 summarizes the comparison of these three different methods by presenting the root integrality gaps. We refer the interested reader to the online supplement for detailed results for all instances.

The results presented in Table 3 suggest that the gap closed by the two-period closure cuts can be quite considerable in many cases, although superior lower bounds come at the cost of higher CPU time (7,900 seconds for X2PL and a few seconds for the other two methods). In particular, we attain better average integrality gaps than the LR for any tested number of items and periods, and a better overall average integrality gap than the period decomposition approach (PD), which generates stronger bounds than LR. Since integrality gaps are calculated using the best upper bound found from all algorithms, our approach generates the most competitive lower bounds for homogeneous and heterogeneous instances. Although the other two approaches, LR and PD, in contrast, return improved lower bounds in heterogeneous instances, our method still delivers better lower bounds thanthese methods. The consistent performance of X2PL indicates that the lower bound quality is not affected by the input structure and that it is more robust than the two other methods considered.

5.2.5. More Trigeiro Instances. To further investigate the strength of the two-period closure lower bound, we performed additional computational experiments on the X data set of Trigeiro et al. (1989). This data set consists of 180 instances of 10 products and 20 periods each, with varying levels of demand variability, EOQ capacity utilization, time between orders, and average setup times. More information on this data set can be found in Trigeiro et al. (1989). We excluded the instances for which the gap was simply closed by (ℓ, S) inequalities. Table 4 presents the integrality gap obtained by the two-period closure, compared to the gap obtained by Pimentel et al. (2010) and de Araujo et al. (2015), the two most recent approaches that have considered this data set. We refer the interested reader to the online supplement for detailed results for all instances.

We see that the difference of the branch-and-price-based methods of Pimentel et al. (2010) and de Araujo et al. (2015) and the strength of the cuts generated by the two-closure procedure is even more profound for this data set. The average gap is just above 1%, more than 40% improvement over Pimentel et al. (2010) and 17% improvement over de Araujo et al. (2015). More importantly, our approach seems to be the most effective in some instances with large gaps. In particular, in sets X11419 and X11429, which have the top average gaps across all methods (with better gaps of PD, 7.07% and 4.99%, respectively), our algorithm delivers the best gaps of 5.70% and 4.85%, respectively. In terms of CPU times, PD is the fastest approach, as it needs an average 42 seconds of CPU time to complete when using a time limit of 150 seconds; Pimentel and X2PL are slower, with 2,309 and 1,787 seconds, respectively. We note that a direct comparison of CPU times is not meaningful, as different software platforms and technologies were used.

Table 4 Average Integrality Gaps of the Decomposition Approaches of Pimentel et al. (2010), de Araujo et al. (2015), and X2PL for the Trigeiro X Data Set

Instance group	Pimentel	PD	X2PL	Instance group	Pimentel	PD	X2PL
X11119	1.54	3.13	1.59	X12119	1.73	3.46	1.77
X11128	0.1	0.11	0	X12127	0.02	0.06	0
X11129	2.53	2.87	1.41	X12128	0.29	0.55	0.16
X11217	0.08	0.07	0.02	X12129	1.14	0.92	0.47
X11218	0.41	0.24	0.16	X12217	0.17	0.17	0.06
X11219	2.38	2.47	1.85	X12218	0.63	0.59	0.37
X11227	0.09	0.03	0.03	X12219	2.75	2.51	1.75
X11228	0.33	0.29	0.23	X12227	0.15	0.11	0.04
X11229	3.05	2.07	2.09	X12228	0.46	0.28	0.17
X11417	0.47	0.22	0.21	X12229	3.24	1.92	1.73
X11418	2.25	0.98	0.90	X12417	1.09	0.56	0.52
X11419	10.37	7.07	5.70	X12418	1.82	0.86	0.78
X11427	0.98	0.25	0.28	X12419	5.97	4.25	5.04
X11428	4.78	0.95	1.13	X12427	0.59	0.27	0.37
X11429	9.6	4.99	4.85	X12428	3.91	1.39	1.41
X12117	0.05	0.05	0	X12429	8	3.51	4.04
X12118	0.2	0.18	0	X11118	0.03	0.07	0
				Average	2.09	1.45	1.20

Notes. Each X row reports the average gap of five instances, excluding those whose gap closed by the (ℓ, S) inequalities. The gap is calculated as $(z_{\mathrm{UB}} - z_{\mathrm{LB}})/z_{\mathrm{UB}}$, where the upper bound is calculated by each method invidually. We note that Pimentel et al. (2010) do not report individual lower or upper bounds but gaps, and therefore we cannot use a common upper bound to calculate an integrality gap across all methods.

6. Conclusions

We have presented a new methodology that can significantly improve traditional lower bounds for the lot-sizing problems by generating cuts from two-period subproblem relaxations. An important advantage of the framework is that is does not require the study of families of valid inequalities or reformulations, and to our knowledge, this is an original approach in the lot-sizing literature from this perspective. A side benefit of our methodology is that the automatic generation of valid inequalities is an invaluable tool toward the study of their structure and of their strength. This is currently investigated in a companion paper (Doostmohammadi et al. 2016). From a practical viewpoint, our computational results show that the lower bound improvement resulting from two-period subproblem cuts is comparable or superior to methodologies such as column generation (de Araujo et al. 2015) and LR (Süral et al. 2009).

Different distance approaches have proven useful to generate cuts and improve lower bounds significantly, particularly for small problems of the test set 2PCLS. From the aspect of computational efficiency, the Euclidean approach achieves significant convergence rates compared to linear norms studied, although it might easily cause numerical issues. As the use of floating arithmetic might be limiting for cut generation processes, an interesting future direction of research is the improvement of the computational stability of our approach. Moreover, it would also be interesting to experiment with various computational strategies, e.g., have a pool of extreme points that could be used in subsequent iterations.

Although the application context of our methodology is capacitated lot sizing, the same principle can be applied readily to any other MIP problems. A matter of ongoing research is the development of an algorithm that automatically selects substructures of MIP formulations that are expected to generate deep cuts. An interesting relevant study is the work of Chvátal et al. (2013), which investigates generating local cuts for general MIP problems. Although the impact of these cuts was not always obvious, the paper discusses a number of effective computational strategies that could provide significant improvements. This provides a motivation for future research investigating extending our framework to general MIP problems.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/ijoc.2016.0712.

Acknowledgments

The authors thank Laurence Wolsey and Stephen Robinson for earlier discussions that led to significant improvements of the paper. The research of the first author was supported in part by the EPSRC [Grant EP/L000911/1] entitled "Multi-Item Production Planning: Theory, Computation and Practice," and the research of the third author was supported in part by the NSF [Grants CMMI-0323299 and CMMI-0521953].

References

Akartunalı K (2007) Computational methods for big bucket production planning problems: Feasible solutions and strong formulations. Unpublished doctoral thesis, Industrial Engineering, University of Wisconsin-Madison, Madison, WI.

Akartunalı K, Miller AJ (2009) A heuristic approach for big bucket multi-level production planning problems. *Eur. J. Oper. Res.* 193(2):396–411.

Akartunalı K, Miller AJ (2012) A computational analysis of lower bounds for big bucket production planning problems. *Comput. Optim. Appl.* 53(3):729–753.

Andersen K, Cornuéjols G, Li Y (2005) Split closure and intersection cuts. *Math. Program.* 102(3):457–493.

Anily S, Tzur M, Wolsey LA (2009) Multi-item lot-sizing with joint set-up costs. *Math. Program.* 119(1):79–94.

Applegate D, Bixby R, Chvátal V, Cook W (2003) Implementing the Dantzig-Fulkerson-Johnson algorithm for large traveling salesman problems. *Math. Program.* 97(1):91–153.

Applegate DL, Cook W, Dash S, Espinoza DG (2007) Exact solutions to linear programming problems. *Oper. Res. Lett.* 35(6):693–699.

Atamtürk A, Muñoz JC (2004) A study of the lot-sizing polytope. *Math. Program.* 99(3):443–465.

Balas E, Margot F (2013) Generalized intersection cuts and a new cut generating paradigm. *Math. Program.* 137(1):19–35.

Balas E, Saxena A (2008) Optimizing over the split closure. *Math. Program.* 113(2):219–240.

Barany I, Van Roy TJ, Wolsey LA (1984) Strong formulations for multi-item capacitated lot-sizing. *Management Sci.* 30(10): 1255–1261.

Belvaux G, Wolsey LA (2000) bc—prod: A specialized branch-and-cut system for lot-sizing problems. *Management Sci.* 46(5):724–738.

Belvaux G, Wolsey LA (2001) Modelling practical lot-sizing problems as mixed-integer programs. *Management Sci.* 47(7): 993–1007.

Bergner M, Caprara A, Ceselli A, Furini F, Lübbecke ME, Malaguti E, Traversi E (2015) Automatic Dantzig-Wolfe reformulation of mixed integer programs. *Math. Program.* 149(1):391–424.

Billington PJ, McClain JO, Thomas LJ (1986) Heuristics for multilevel lot-sizing with a bottleneck. *Management Sci.* 32(8): 989–1006.

Bitran GR, Matsuo H (1986) The multi-item capacitated lot size problem: Error bounds of Manne's formulations. *Management Sci.* 32(3):350–359.

Boyd EA (1994) Fenchel cutting planes for integer programs. *Oper. Res.* 42(1):53–64.

Cadoux F (2010) Computing deep facet-defining disjunctive cuts for mixed-integer programming. *Math. Program.* 122(2):197–223.

Chvátal V, Cook W, Espinoza D (2013) Local cuts for mixed integer programming. *Math. Program. Comput.* 5(2):171–200.

Constantino M (1996) A cutting plane approach to capacitated lot-sizing with start-up costs. *Math. Program.* 75(3):353–376.

Cook W, Dash S, Fukasawa R, Goycoolea M (2009) Numerically safe Gomory mixed-integer cuts. *INFORMS J. Comput.* 21(4): 641–649.

Curtis AR, Reid JK (1972) On the automatic scaling of matrices for Gaussian elimination. *IMA J. Appl. Math.* 10(1):118–124.

Dash S, Günlük O, Lodi A (2010) MIR closures of polyhedral sets. *Math. Program.* 121(1):33–60.

de Araujo SA, De Reyck B, Degraeve Z, Fragkos I, Jans R (2015) Period decompositions for the capacitated lot sizing problem with setup times. *INFORMS J. Comput.* 27(3):431–448.

Degraeve Z, Jans R (2007) A new Dantzig-Wolfe reformulation and branch-and-price algorithm for the capacitated lot-sizing problem with setup times. *Oper. Res.* 55(5):909–920.

Doostmohammadi M, Fragkos I, Akartunalı K (2016) A polyhedral study of and a branch-and-cut framework for two-period relaxations of big-bucket lot-sizing problems with setup times. Working paper, University of Strathclyde Business School, Glasgow.

Eppen GD, Martin RK (1987) Solving multi-item capacitated lot-sizing problems using variable redefinition. *Oper. Res.* 35(6): 832–848.

Federgruen A, Tzur M (1991) A simple forward algorithm to solve general dynamic lot sizing models with n periods in $0(n \log n)$ or $0(n)$ time. *Management Sci.* 37(8):909–925.

Federgruen A, Meissner J, Tzur M (2007) Progressive interval heuristics for multi-item capacitated lot-sizing problems. *Oper. Res.* 55(3):490–502.

Fischetti M, Lodi A (2005) Optimizing over the first Chvátal closure. *Integer Programming and Combinatorial Optimization (IPCO) XI*, Lecture Notes in Computer Science, Vol. 3509 (Springer, Berlin), 12–22.

Florian M, Klein M (1971) Deterministic production planning with concave costs and capacity constraints. *Management Sci.* 18(1):12–20.

Florian M, Lenstra JK, Rinnooy Kan AHG (1980) Deterministic production planning: Algorithms and complexity. *Management Sci.* 26(7):669–679.

Fragkos I, Akartunalı K (2014) A computational study of the local cuts from two-period convex hull closures for big-bucket lot-sizing problems. *5th Internat. Workshop Lot Sizing, Porto, Portugal*, 41–45.

Fukasawa R, Goycoolea M (2011) On the exact separation of mixed integer knapsack cuts. *Math. Program.* 128(1):19–41.

Fukuda K (2014) cdd and cddplus. Accessed October 19, 2015, http://www.inf.ethz.ch/personal/fukudak/cdd_home/.

Jans R, Degraeve Z (2004) Improved lower bounds for the capacitated lot sizing problem with setup times. *Oper. Res. Lett.* 32(2):185–195.

Krarup J, Bilde O (1977) Plant location, set covering and economic lotsizes: An $O(mn)$ algorithm for structured problems. Collatz L, et al., eds. *Optimierung bei Graphentheoretischen und Ganzzahligen Probleme* (Birkhauser Verlag, Basel, Switzerland), 155–180.

Letchford AN (2001) On disjunctive cuts for combinatorial optimization. *J. Comb. Optim.* 5(3):299–315..

Levi R, Lodi A, Sviridenko M (2008) Approximation algorithms for the capacitated multi-item lot-sizing problem via flow-cover inequalities. *Math. Oper. Res.* 33(2):461–474.

Mangasarian O (1994) *Nonlinear Programming* (SIAM Classics in Applied Mathematics, Philadelphia).

Manne AS (1957) Programming of economic lot sizes. Cowles Foundation Discussion Papers 23, Cowles Foundation for Research in Economics, Yale University, New Haven, CT.

Miller AJ, Nemhauser GL, Savelsbergh MWP (2000) Solving the multi-item capacitated lot-sizing problem with setup times by branch-and-cut. Technical Report CORE DP 2000/39, Université Catholique de Louvain, Louvain-la-Neuve, France.

Miller AJ, Nemhauser GL, Savelsbergh MWP (2003) On the polyhedral structure of a multi-item production planning model with setup times. *Math. Program.* 94(2):375–405.

Pimentel CMO, Alvelos FP, Valério de Carvalho JM (2010) Comparing Dantzig–Wolfe decompositions and branch-and-price algorithms for the multi-item capacitated lot sizing problem. *Optim. Method. Softw.* 25(2):299–319.

Pochet Y, Van Vyve M (2004) A general heuristic for production planning problems. *INFORMS J. Comput.* 16(3):316–327.

Pochet Y, Wolsey LA (1988) Lot-size models with backlogging: Strong reformulations and cutting planes. *Math. Program.* 40(1):317–335.

Pochet Y, Wolsey LA (1991) Solving multi-item lot-sizing problems using strong cutting planes. *Management Sci.* 37(1):53–67.

Pochet Y, Wolsey LA (1994) Polyhedra for lot-sizing with Wagner-Whitin costs. *Math. Program.* 67(1):297–323.

Pochet Y, Wolsey LA (2006) *Production Planning by Mixed Integer Programming* (Springer, New York).

Ralphs TK, Kopman L, Pulleyblank WR, Trotter LE (2003) On the capacitated vehicle routing problem. *Math. Program.* 94(2): 343–359.

Rardin RL, Wolsey LA (1993) Valid inequalities and projecting the multicommodity extended formulation for uncapacitated fixed charge network flow problems. *Eur. J. Oper. Res.* 71(1):95–109.

Stadtler H (2003) Multilevel lot sizing with setup times and multiple constrained resources: Internally rolling schedules with lot-sizing windows. *Oper. Res.* 51(3):487–502.

Süral H, Denizel M, Van Wassenhove LN (2009) Lagrangean relaxation based heuristics for lot sizing with setup times. *Eur. J. Oper. Res.* 194(1):51–63.

Trigeiro WW, Thomas LJ, McClain JO (1989) Capacitated lot sizing with setup times. *Management Sci.* 35(3):353–366.

Van Vyve M, Wolsey L (2006) Approximate extended formulations. *Math. Program.* 105(2):501–522.

Van Vyve M, Wolsey LA, Yaman H (2014) Relaxations for two-level multi-item lot-sizing problems. *Math. Program.* 146(1):495–523.

Wagner HM, Whitin TM (1958) Dynamic version of the economic lot size model. *Management Sci.* 5(1):89–96.

Zangwill WI (1969) A backlogging model and a multi-echelon model of a dynamic economic lot size production system—A network approach. *Management Sci.* 15(9):506–527.

Zhang M, Küçükyavuz S, Yaman H (2012) A polyhedral study of multiechelon lot sizing with intermediate demands. *Oper. Res.* 60(4):918–935.