Ann Oper Res (2015) 235:233–258
DOI 10.1007/s10479-015-1915-4

The single-item lot-sizing polytope with continuous start-up costs and uniform production capacity

Mariana Escalante[1,2] · **Javier Marenco**[3] ·
María del Carmen Varaldo[2]

Published online: 7 June 2015

Abstract In this work we consider the uniform capacitated single-item single-machine lot-sizing problem with continuous start-up costs. A continuous start-up cost is generated in a period whenever there is a nonzero production in the period and the production capacity in the previous period is not saturated. This concept of start-up does not correspond to the standard (discrete) start-up considered in previous models, thus motivating a polyhedral study of this problem. In this work we explore a natural integer programming formulation for this problem. We consider the polytope obtained as convex hull of the feasible points in this problem. We state some general properties, study whether the model constraints define facets, and present an exponentially-sized family of valid inequalities for it. We analyze the structure of the extreme points of this convex hull, their adjacency and bounds for the polytope diameter. Finally, we study the particular case when the demands are high enough in order to require production in all the periods. We provide a complete description of the convex hull of feasible solutions in this case and show that all the inequalities in this description are separable in polynomial time, thus proving its polynomial time solvability.

Keywords Lot-sizing · Continuous start-up · Polyhedral combinatorics

✉ Mariana Escalante
mariana@fceia.unr.edu.ar

Javier Marenco
jmarenco@ungs.edu.ar

María del Carmen Varaldo
mvaraldo@fceia.unr.edu.ar

1 CONICET, Buenos Aires, Argentina

2 Departamento de Matemática, FCEIA, Universidad Nacional de Rosario, Rosario, Argentina

3 Instituto de Ciencias, Universidad Nacional de General Sarmiento, Buenos Aires, Argentina

1 Introduction

In this work we consider the capacitated single-item single-machine lot-sizing problem with *continuous start-up* costs. In this problem we have a planning horizon consisting of p periods, $T = \{1, \ldots, p\}$, having nonnegative demand on each period. We assume that the production of the single item in each period can be any real number between 0 and 1. A continuous start-up cost is incurred in period $t > 1$ if the production at period t is nonzero and the production at period $t - 1$ is not saturated (i.e., strictly less than 1). The *lot-sizing problem with continuous start-up costs* asks for a production plan (i.e., the quantity to be produced at each period) satisfying the demands and minimizing the total costs.

This concept of continuous start-up was first presented in Toledo et al. (2008) in the context of planning the operation of water pumps in a real setting in Brazil. Water pumps usually operate at a fixed speed (i.e., there are no controls allowing the pump to run at different speeds) and must be turned off if the operator needs just a fraction of the production capacity in a certain period. In this case, the operator must run the water pump at full speed for part of the time, and then the pump must be shut off (as opposed to running the pump at partial speed for the entire period). Since the water pump must be turned off and restarted in the next period needing production, the particular concept of so-called continuous start-up considered in this work is generated. Besides water pumps, this setting is relevant for any continuous single-velocity machine.

An important assumption in this model is the uniformity of the production capacity. The periods are assumed to span similar time intervals, hence at each period the maximum production is constant. Furthermore, we normalize this maximum production so the production in each period is any real number in $[0, 1]$. This normalization helps in the reductions presented in this work.

The capacitated lot-sizing problem with start-up costs can be modeled using the following variables. For $t \in T$, we introduce the *production variable* $x_t \in [0, 1]$ representing the production in period t. For $t \in T$, we employ a binary *set-up variable* $y_t \in \{0, 1\}$ representing whether there is production in period t or not (i.e., $y_t = 1$ if $x_t > 0$). Finally, for $t \in T$, we introduce a binary *start-up variable* $a_t \in \{0, 1\}$ asserting whether the (continuous) production starts at period t or not (i.e., $a_t = 1$ if $x_t > 0$ and $x_{t-1} < 1$). Note that we treat the variable a_1 as if there were an initial production $x_0 = 0$. Throughout this work we assume a null initial inventory.

For each $t \in T$, we denote by d_t the nonnegative demand of the single item in the period t. For $i, j \in T$, $i \leq j$, we define $d_{ij} = \sum_{t=i}^{j} d_t$. A feasible solution for the capacitated lot-sizing problem with continuous start-up costs is an assignment of values to the variables satisfying the following set of constraints:

$$0 \leq x_t \leq y_t \quad t \in T, \tag{1a}$$
$$a_1 = y_1, \tag{1b}$$
$$a_{t+1} \geq y_{t+1} - x_t \quad t \in T \setminus \{p\}, \tag{1c}$$
$$d_{1t} \leq \sum_{k=1}^{t} x_k \quad t \in T, \tag{1d}$$
$$y_t \in \{0, 1\} \quad t \in T, \tag{1e}$$
$$a_t \in \{0, 1\} \quad t \in T. \tag{1f}$$

Constraints (1a) assert that the set-up variable y_t must take value $y_t = 1$ if there is production in the period t, for $t \in T$. Constraints (1b) and (1c) define the start-up variables, and constraints (1d) ask for the demand to be satisfied. We write $(x, y, a) \in \mathbb{R}^{3p}$ for $x, y, a \in \mathbb{R}^p$

and define $P = \text{conv}(\mathcal{S})$ where $\mathcal{S}$ is the set of feasible solutions of the formulation (1a)–(1f). We denote by P_{LR} the linear relaxation of P, that is, the polyhedron obtained by allowing the binary variables in $\mathcal{S}$ to be any real number between 0 and 1.

In Constantino (1996) and Hoesel et al. (1994), the authors study the *discrete* start-up model where the start-up constraint (1c) is replaced by $a_{t+1} \geq y_{t+1} - y_t$ for $t \in T \setminus \{p\}$. This discrete start-up cost is incurred in a period $t+1$ if there is no production in period $t \in T$ (i.e., $x_t = 0$) and there is positive production in period $t+1$ (i.e., $x_{t+1} > 0$). In contrast, a continuous start-up cost is generated at period $t+1$ if $x_t < 1$ and $x_{t+1} > 0$. In Hoesel et al. (1994) the authors present a complete characterization of the convex hull of feasible solutions for the *uncapacitated* discrete start-up model.

Some lot-sizing models have been shown to admit an interesting polyhedral structure, and this motivates us to explore in this work the polytope P. We are interested in studying how the particular start-up considered in this work affects the structure of the polytope. In particular, we continue the search for general valid inequalities, and we present for the first time a family of valid inequalities composed by an exponential number of elements. We are also interested in general properties of this polytope, including the structure of its extreme points and bounds on its diameter.

Finally, we consider the particular case where the demands force production in all the periods, and we provide a complete characterization by linear inequalities of the polytope in this case. We also show that the inequalities in this characterization are separable in polynomial time, hence showing that it can be solved in polynomial time for *any* linear objective function. To the best of our knowledge, this is the first case of a *capacitated* lot-sizing problem with arbitrary (production and continuous start-up) costs for which this goal is attained. Furthermore, this result provides another example of a combinatorial problem for which a nice polyhedral characterization leads to a polynomial-time algorithm.

The paper is organized as follows. In Sect. 2 we first present general properties of the polytope P, such as its dimension and a minimal system of equations. We identify the model constraints that induce facets of P and we also introduce two new families of valid inequalities. A preliminary version of some of these results appeared without proof in the conference paper (Escalante et al. 2011). In Sect. 3 we explore the combinatorial structure of this polytope. In particular, we study under which assumptions a feasible solution is an extreme point and we explore conditions ensuring that two extreme points are neighbors in P (i.e., the segment joining them is an edge of the polytope). These results allow us to find bounds on the diameter of the polytope P. In Sect. 4 we study the particular case of high demands forcing production in all the periods. We provide a complete description of P in this particular setting by introducing a set of valid inequalities that includes all the facets of P, and we show that this new family of inequalities can be separated in polynomial time. In order to clarify the presentation, we have added an "Appendix" where we present a particular instance of the continuous start-up lot-sizing problem with high demands.

2 The polytope P for general demands

In this section we present some general results on the structure of P. For $w \in P$ and $A \subseteq \{1, \ldots, 3p\}$, we define $w_A = (w_i)_{i \in A}$ to be the vector obtained from w by projecting out the variables outside A. Let $\mathbf{1}$ be the all-ones vector, $\mathbf{0}$ be the all-zeros vector and e_i be the i-th unit vector for $i = 1, \ldots, p$, all of them of appropriate dimension. For simplicity, we write $x_{ij} = \sum_{t=i}^{j} x_t$ for $i, j \in T,\ i \leq j$.

By combining constraint (1d) with $x_i \leq 1$ for every $i \in T$, we have the following result.

Proposition 1 *The polytope P is nonempty if and only if $d_{1k} \leq k$ for every $k \in T$.*

Throughout this work we assume that P is nonempty and $d_1 > 0$. If there exists a period $k \in T$ with $d_{1k} = k$, then $x_t = 1$ for $t \in \{1, \dots, k\}$. We can preprocess this case; hence in this work we assume w.l.o.g. $d_{1k} < k$ for every period $k \in T$. Define $k_{\text{prod}} \in T$ to be the maximum period k such that $d_{1k} > k - 1$ if such period exists, and $k_{\text{prod}} = 0$ otherwise. Note that $x_t > 0$ for $t = 1, \dots, k_{\text{prod}}$.

A *minimal equation system* for a polytope $Q \in \mathbb{R}^n$ is a set of linearly independent equations such that any equation satisfied by every point in Q is a linear combination of them. If such a system has q equations, then $\dim(Q) = n - q$.

Theorem 1 *A minimal equation system for P is given by*

(i) $y_k = 1$ for $k \in \{1, \dots, k_{\text{prod}}\}$,
(ii) $a_1 = 1$ if $k_{\text{prod}} > 0$.

Proof If $k_{\text{prod}} > 0$, then any feasible solution $(x, y, a) \in \mathbb{R}^{3p}$ has nonzero production in the first k_{prod} periods, hence $y_k = 1$ for $k \in \{1, \dots, k_{\text{prod}}\}$ and, furthermore, $a_1 = 1$.

Let $\Lambda = (\mu, \gamma, \delta) \in \mathbb{R}^{3p}$, where $\mu, \gamma, \delta \in \mathbb{R}^p$ and $\Lambda_0 \in \mathbb{R}$ such that $\Lambda w = \Lambda_0$ for every $w = (x, y, a) \in P$. We verify that $\Lambda w = \Lambda_0$ is a linear combination of (i)–(ii).

Clearly, $\mathbf{1} \in P$. Let $0 < \varepsilon < \min\{1, t - d_{1t} : t \in T\}$ and define $w^k = (x^k, \mathbf{1}, \mathbf{1})$ with $x^k = \mathbf{1} - \varepsilon e_k$, for each $k \in T$. We have that $w^k \in P$, since $x^k_{1t} = t - \varepsilon > d_{1t}$ for $t \in T$. The points $\mathbf{1}$ and w^k only differ in their x_k-variable which implies that $\mu_k = 0$.

For each $k \in \{k_{\text{prod}} + 1, \dots, p\}$, let $\bar{w}^k = (\bar{x}^k, \mathbf{1}, \mathbf{1})$ with $\bar{x}^k = \mathbf{1} - e_k$. Since $\bar{x}^k_{1t} = t - 1 \geq d_{1t}$ for every $t \geq k > k_{\text{prod}}$, $\bar{w}^k \in P$. Moreover, $\hat{w}^k = (\hat{x}^k, \hat{y}^k, \mathbf{1})$ with $\hat{x}^k = \hat{y}^k = \mathbf{1} - e_k$ is also feasible since $\hat{x}^k_k = 0$. The points $\bar{w}^k$ and $\hat{w}^k$ only differ in their y_k-variable, implying that $\gamma_k = 0$.

Finally, for $k = 2, \dots, p$, let $\widetilde{w}^k = (\mathbf{1}, \mathbf{1}, \widetilde{a}^k)$ with $\widetilde{a}^k = \mathbf{1} - e_k$, which is feasible since $\widetilde{a}_t \geq \widetilde{y}_t - \widetilde{x}_{t-1} = 0$ for every period t. By the definition of $\widetilde{w}^k$ we have that $\delta_k = 0$.

In this way we have shown that the only nonzero coordinates in Λ are those in (i)–(ii). Furthermore, each such equation has exactly one nonzero coefficient, hence $\Lambda w = \Lambda_0$ is a linear combination of (i)–(ii) which, therefore, defines a minimal equation system for P. □

2.1 Facet-inducing inequalities for P

In Pulleyblank (1989), Pulleyblank presents the following result that characterizes facet-defining inequalities of a polyhedron, we rewrite it here for completeness using our notation.

Theorem 2 [Pulleyblank (1989)] *Let F be a proper face of $P = \{w \in \mathbb{R}^{3p} : Aw \leq b\}$. If A_I is the row submatrix of A indexed in the set I (similarly for b_I) and $A_I w = b_I$ is a minimal equation system for P, then the following statements are equivalent:*

(i) F is a facet of P.
(ii) $dim(F) = dim(P) - 1$.
(iii) Let $\alpha, \ \Lambda \in \mathbb{R}^{3p}$ and $\beta, \ \Lambda_0 \in \mathbb{R}$ be such that $\alpha w \leq \beta$ and $\Lambda w \leq \Lambda_0$ are valid inequalities for P. If $F = \{w \in P : \alpha w = \beta\}$ and $B = \{w \in P : \Lambda w = \Lambda_0\}$ satisfy $F = B$, then there exist $v \in \mathbb{R}^I$ and $u \in \mathbb{R}^+$ such that $\Lambda = u\alpha + vA_I$ and $\Lambda_0 = u\beta + vb_I$.

Theorem 2 allows us to characterize the model constraints that induce facets of P.

In the proofs of the following theorems we shall need to distinguish between different kinds of points. We denote by $w^{(i,j)} \in \mathbb{R}^{3p}$ for $i \in \{1, 2, 3\}$ and $j \in T$ a point considered for studying the x-variable ($i = 1$), the y-variable ($i = 2$) and the a-variable ($i = 3$), respectively, corresponding to the period j. For instance, $w^{(1,j)}$ corresponds to the analysis of the variable x_j.

Theorem 3 *The following model constraints define facets of P:*

(i) $\sum_{j=1}^{p} x_j \geq d_{1p}$.
(ii) $x_s \leq 1$ for $s \in \{1, \ldots, k_{\text{prod}}\}$.
(iii) $x_s \leq y_s$ and $y_s \leq 1$ for $s \in \{k_{\text{prod}} + 1, \ldots, p\}$.
(iv) $a_s \leq 1$ for $s \in \{2, \ldots, p\}$.

Proof Since the arguments are similar for all constraints, we prove that the inequality (i) induces a facet and omit the rest of the proofs.

Let $\Lambda = (\mu, \gamma, \delta) \in \mathbb{R}^{3p}$ with $\mu_1 \neq 0$ and $\Lambda_0 \in \mathbb{R}$ such that $\Lambda w \leq \Lambda_0$ is a valid inequality for P. If $\alpha = (-\mathbf{1}, \mathbf{0}, \mathbf{0})$ and $\beta = -d_{1p}$, the inequality $\alpha w \leq \beta$ corresponds to the constraint (i) for $w = (x, y, a) \in \mathbb{R}^{3p}$.

With this choice of α and β, if $F = \{w \in P\colon \alpha w = \beta\}$ and $B = \{w \in P\colon \Lambda w = \Lambda_0\}$ we show that (i) defines a facet of P by using (iii) in Theorem 2. Assuming that $F = B$, we prove that there exist $v \in \mathbb{R}^I$ and $u \in \mathbb{R}^+$ such that $\Lambda = u(-\mathbf{1}, \mathbf{0}, \mathbf{0}) + vA_I$ and $\Lambda_0 = u(-d_{1p}) + vb_I$ where I is the set of indices of the minimal equation system presented in Theorem 1.

Let $r \in T$ such that $\lceil d_{1p} \rceil = r$. Since $d_1 < 1$, $d_{1p} < p$. If $k_{\text{prod}} = p$ then $r = p$ else $r < p$. If $r < p$ we define $\hat{w} = (\hat{x}, \mathbf{1}, \mathbf{1})$ where

$$\hat{x}_i = \begin{cases} 1 & i \in \{1, \ldots, r-1\}, \\ d_{1p} - r + 1 & i = r, \\ 0 & i \in \{r+1, \ldots, p\} \end{cases}$$

and, if $r = p$, we consider $\hat{w} = (\hat{x}, \mathbf{1}, \mathbf{1})$ such that

$$\hat{x}_i = \begin{cases} 1 & i \in \{1, \ldots, p-1\}, \\ d_{1p} - p + 1 & i = p. \end{cases}$$

In any case it is clear that $\hat{w} \in F$. Using this definition we subdivide the rest of the proof into five cases.

Case (a) $\mu_j = \mu_1$ for $j \in \{1, \ldots, p\}$.

Assume that $r < p$.
Let $j \in \{1, \ldots, r-1\}$. Let us consider

$$\varepsilon = \min\{\hat{x}_{1i} - d_{1i} \,:\, i = 1, \ldots, r\},$$

or, by the definition of $\hat{x}$ for $r < p$,

$$\varepsilon = \min\left\{d_{1p} - d_{1r} + 1, \min\{i - d_{1i} \,:\, i = 1, \ldots, r-1\}\right\}.$$

Note that $0 < \varepsilon < 1$ since $d_1 > 0$. We consider $w_i^{(1,j)} = \hat{w}_i$ for every i except for $x_j^{(1,j)} = \hat{x}_j - \varepsilon$ and $x_{r+1}^{(1,j)} = \varepsilon$. Then, $w^{(1,j)} \in F$. Using the fact that $F = B$, we have $\Lambda\hat{w} = \Lambda w^{(1,j)}$ which implies that $\mu_j \hat{x}_j = \mu_j(\hat{x}_j - \varepsilon) + \mu_{r+1}\varepsilon$. Then, $\mu_j = \mu_{r+1}$. Let $j \in \{r+1, \ldots, p\}$ and $\varepsilon_j = \min\{\hat{x}_r, d_{1p} - d_{1(j-1)}\}$. Again, $\varepsilon_j \leq 1$. Let $w_i^{(1,j)} = \hat{w}_i$ for every i except for $x_r^{(1,j)} = \hat{x}_r - \varepsilon_j$ and $x_j^{(1,j)} = \varepsilon_j$. Then, $w^{(1,j)} \in F$.

Using the fact that $F = B$, we have $\Lambda\hat{w} = \Lambda w^{(1,j)}$ which implies that $\mu_r \hat{x}_r = \mu_r(\hat{x}_r - \varepsilon_j) + \mu_j \varepsilon_j$. Then, $\mu_j = \mu_r$.
Now, assume that $r = p$.
Let $j \in \{1, \ldots, p-1\}$. Define

$$\varepsilon = \min\{i - d_{1i}\colon \quad i = 1, \ldots, p\},$$

clearly, $0 < \varepsilon < p - d_{1p} < 1$. We consider $w_i^{(1,j)} = \hat{w}_i$ for every i except for $x_j^{(1,j)} = \hat{x}_j - \varepsilon = 1 - \varepsilon$ and $x_p^{(1,j)} = \hat{x}_p + \varepsilon$. Then, $w^{(1,j)} \in F$. Using the fact that $F = B$, we have $\Lambda\hat{w} = \Lambda w^{(1,j)}$ which implies that $\mu_j \hat{x}_j + \mu_p \hat{x}_p = \mu_j(\hat{x}_j - \varepsilon) + \mu_p(\hat{x}_p + \varepsilon)$. Then, $\mu_j = \mu_p$.

In both cases we conclude that $\mu_j = \mu_1$ for $j \in \{1, \ldots, p\}$. $\lozenge$
Case (b) If $r < p$, $\gamma_j = 0$ for $j \in \{r+1, \ldots, p\}$.
Let $j \in \{r+1, \ldots, p\}$. Observe that $k_{\text{prod}} < r + 1$. Define $w_i^{(2,j)} = \hat{w}_i$ for every i except for $y_j^{(2,j)} = 0$. Clearly, $w^{(2,j)} \in F$. Again, using the fact that $F = B$, we have $\Lambda\hat{w} = \Lambda w^{(2,j)}$ which implies $\gamma_j = 0$. $\lozenge$
Case (c) If $r < p$, $\gamma_j = 0$ for $j \in \{1, \ldots, r\} \cap \{k_{\text{prod}} + 1, \ldots, p\}$.
Let $j \in \{1, \ldots, r\} \cap \{k_{\text{prod}} + 1, \ldots, p\}$. Consider $w_i^{(2,j)} = \hat{w}_i$ for every i except for $x_j^{(2,j)} = y_j^{(2,j)} = 0$, $x_r^{(2,j)} = 1$ and $x_{r+1}^{(2,j)} = \hat{x}_r$. Since $j > k_{\text{prod}}$, it is easy to see that $w^{(2,j)} \in F$. Again, using the fact that $F = B$ we have $\Lambda\hat{w} = \Lambda w^{(2,j)}$ which implies $\mu_r \hat{x}_r + \mu_j + \gamma_j = \mu_r + \mu_{r+1}\hat{x}_r$. After Case (a), $\gamma_j = 0$. $\lozenge$
Case(d) $\delta_j = 0$ for $j \in \{2, \ldots, k_{\text{prod}}\}$.
Let $j \in \{1, \ldots, k_{\text{prod}}\}$. Define $w_i^{(3,j)} = \hat{w}_i$ for every i except for $a_j^{(3,j)} = 0$. Clearly, $w^{(3,j)} \in F$ and using the fact that $F = B$ we have $\Lambda\hat{w} = \Lambda w^{(3,j)}$ which implies that $\delta_j = 0$. $\lozenge$
Case (e) If $k_{\text{prod}} < p$, $\delta_j = 0$ for $j \in \{k_{\text{prod}} + 1, \ldots, p\}$.
Let $j \in \{k_{\text{prod}} + 1, \ldots, p\}$. Define $w_i^{(3,j)} = w_i^{(2,j)}$ for every i except for $a_j^{(3,j)} = 0$. Clearly, $w^{(3,j)} \in F$ and using the fact that $F = B$ we have $\Lambda w^{(2,j)} = \Lambda w^{(3,j)}$ which implies that $\delta_j = 0$. $\lozenge$

Now, by denoting $u = |\mu_1|$ and $v = (\gamma_1, \ldots, \gamma_{k_{\text{prod}}}, \delta_1)$, the previous analyzed cases imply that $\Lambda w = u(-\sum_{t=1}^{p} x_j) + v(y_1, \ldots, y_{k_{\text{prod}}}, a_1)$ and $\Lambda_0 = u(-d_{1p}) + v\mathbf{1}$. According to Theorem 2, the inequality (i) defines a facet for P. □

Several families of facet-inducing valid inequalities were presented in the conference paper (Escalante et al. 2011), all of them composed by at most p inequalities. The following result presents an additional family of facet-defining inequalities of P for $s \in \{1, \ldots, k_{\text{prod}} - 1\}$. These new inequalities dominate the model constraints (1c) when $s - d_{1s} < 1$, implying that the original constraints (1c) do not induce facets of P in this case.

Theorem 4 *For $i \in T$ let $d_i \leq 1$ be such that $k_{\text{prod}} > 1$. Then, for each $s \in \{1, \ldots, k_{\text{prod}} - 1\}$, the inequality*

$$x_s + (s - d_{1s})a_{s+1} \geq 1 \tag{2}$$

defines a facet of P.

Proof We first prove that (2) is a valid inequality. Since P is a polytope it is enough to prove the validity over $\mathcal{S}$. Let us consider the following two cases for $(x, y, a) \in \mathcal{S}$.

Case 1 If $a_{s+1} = 0$, then $s < k_{\text{prod}}$ implies $x_s > 0$, hence either $x_s = 1$ or $y_{s+1} = 0$. We cannot have $y_{s+1} = 0$ as $s + 1 \le k_{\text{prod}}$ and then $x_{s+1} > 0$. Therefore, $x_s = 1$ and (2) is satisfied. $\diamondsuit$

Case 2 If $a_{s+1} = 1$, then the demand satisfaction constraint (1d) for $t = s$ implies

$$x_s + (s - d_{1s})a_{s+1} \ge d_{1s} - \sum_{t=1}^{s-1} x_t + (s - d_{1s})$$
$$= s - \sum_{t=1}^{s-1} x_t \;=\; 1 + \sum_{t=1}^{s-1} (1 - x_t) \;\ge\; 1.\ \diamondsuit$$

We conclude that (2) is valid for P.

We now address the facetness of the constraint in (2). Let $\Lambda = (\mu, \gamma, \delta) \in \mathbb{R}^{3p}$ with $\mu_s \neq 0$ and $\Lambda_0 \in \mathbb{R}$ such that $(\Lambda, \Lambda_0) \in \mathbb{R}^{3p+1}$ is a valid inequality for P. If $\alpha = (e_s, \mathbf{0}, (s - d_{1s})e_{s+1}) \in \mathbb{R}^{3p}$ and $\beta = 1 \in \mathbb{R}$ the inequality $\alpha w \ge \beta$ corresponds to the constraint (2) for $w = (x, y, a) \in \mathbb{R}^{3p}$. With this choice of α and β, if $F = \{w \in P : \alpha w = \beta\}$ and $B = \{w \in P : \Lambda w = \Lambda_0\}$, we show that (2) defines a facet of P by using (iii) in Theorem 2.

Assuming that $F = B$, we prove that there exist $v \in \mathbb{R}^I$ and $u \in \mathbb{R}^+$ such that $\Lambda = u\alpha + vA_I$ and $\Lambda_0 = u\beta + vb_I$ where I is the set of indices of the minimal equation system presented in Theorem 1.

Define $\hat{w} = (\mathbf{1}, \mathbf{1}, \hat{a})$ with $\hat{a}_t = 1$ for every $t \in T \setminus \{s+1\}$ and $\hat{a}_{s+1} = 0$. Clearly, $\hat{w} \in F$. We divide the proof into four cases.

Case (a) $\mu_j = 0$ for $j \in T \setminus \{s\}$.

Let $j \in \{1, \ldots, k_{\text{prod}}\} \setminus \{s\}$. Consider $w_i^{(1,j)} = \hat{w}_i$ for every i except for $x_j^{(1,j)} = 1 - \varepsilon_j$ where $\varepsilon_j = \min\{i - d_{1i}, i = 1, \ldots, k_{\text{prod}}\}$. Note that $0 < \varepsilon_j < 1$. Then, $w^{(1,j)} \in F$ and using the fact that $F = B$ we have that $\Lambda\hat{w} = \Lambda w^{(1,j)}$ which implies that $\mu_j = \mu_j(1 - \varepsilon_j)$. Then, $\mu_j = 0$.

Let $j \in \{k_{\text{prod}} + 1, \ldots, p\}$. Consider $w_i^{(1,j)} = \hat{w}_i$ for every i except for $x_j^{(1,j)} = 0$. Since $j > k_{\text{prod}}$ we have that $d_{1j} \le j - 1$ and clearly $w^{(1,j)} \in F$. Using the fact that $F = B$ we arrive to $\Lambda\hat{w} = \Lambda w^{(1,j)}$ which implies that $\mu_j = 0$. $\diamondsuit$

Case (b) $\delta_{s+1} = (s - d_{1s})\mu_s$.
Let $w_i^{(3,s+1)} = \hat{w}_i$ for every i except for $x_s^{(3,s+1)} = d_{1s} - (s - 1)$ and $a_{s+1}^{(3,s+1)} = 1$. It is clear that $w^{(3,s+1)} \in F$ and using the fact that $F = B$ we have $\Lambda\hat{w} = \Lambda w^{(3,s+1)}$ which implies that $\mu_s = \mu_s(d_{1s} - (s-1)) + \delta_{s+1}$. By simplifying we arrive to $\delta_{s+1} = (s - d_{1s})\mu_s$. $\diamondsuit$

Case (c) $\delta_j = 0$ for $j \in T \setminus \{1, s+1\}$.
Let $j \in T \setminus \{1, s+1\}$ and $w_i^{(3,j)} = \hat{w}_i$ for every i except for $a_j^{(3,j)} = 0$. Again, clearly $w^{(3,j)} \in F$ and using the fact that $F = B$ we have $\Lambda\hat{w} = \Lambda w^{(3,j)}$ which implies that $\delta_j = 0$. $\diamondsuit$

Case (d) $\gamma_j = 0$ for $j \in \{k_{\text{prod}} + 1, \ldots, p\}$.
Let $j \in \{k_{\text{prod}} + 1, \ldots, p\}$ and $w_i^{(2,j)} = \hat{w}_i$ for every i except for $x_j^{(2,j)} = y_j^{(2,j)} = 0$. Then, $w^{(2,j)} \in F$ and using the fact that $F = B$ we have $\Lambda\hat{w} = \Lambda w^{(2,j)}$ which implies that $\mu_j + \gamma_j = 0$. Since $\mu_j = 0$ after Case (a), we obtain $\gamma_j = 0$. $\diamondsuit$

Now, by denoting $u = |\mu_s|$ and $v = (\gamma_1, \ldots, \gamma_{k_{\text{prod}}}, \delta_1)$, the previous analyzed cases imply $\Lambda w = u(x_s + (s - d_{1s})a_{s+1}) + v(y_1, \ldots, y_{k_{\text{prod}}}, a_1)$ and $\Lambda_0 = u + v\mathbf{1}$. According to Theorem 2, (2) defines a facet of P. $\square$

In the remainder of this section we make the following additional assumptions

$$d_i \le 1 \quad \text{for every } i \in T \setminus \{1\} \quad \text{and} \tag{3}$$

$$k_{\text{prod}} > 1. \tag{4}$$

These conditions ensure that it is feasible to produce exactly the demand at each period, since the production can take any value in [0, 1]. Since $d_1 < 1$, the production may not be saturated at the first period. Note, furthermore, that we may have over-production in some periods and zero production in subsequent periods, but this is not mandatory as $d_i \le 1$ for every $i \in T$. In this case we are able to present the following exponentially-sized family of valid inequalities for P.

Definition 1 Let $k_{\text{prod}} > 1, k_0 \ge 1$ and $M \subseteq \{k_0, \dots, k_{\text{prod}}-1\}$. If $A = \{i \in \{k_0, \dots, k_{\text{prod}}-1\}\colon\ i \notin M\}$ and $B = \{i \in \{1, \dots, k_{\text{prod}} - 1\}\colon\ i \notin M\}$, we define

$$\sum_{i\in B} x_i + \sum_{i\in A}(i - k_0 + 1 - d_{k_0 i})a_{i+1} \ \ge \sum_{i\in B\setminus A} d_i + |A| \tag{5}$$

to be the (k_0, M)-*inequality* associated with the sets A and B, and with the period k_0.

Theorem 5 *Under the assumptions* (3) *and* (4) *the* (k_0, M)-*inequality (5) is valid for* P.

Proof Since $k_{\text{prod}} > 1$ we have $x_t > 0$ for every $t \in \{1, \dots, k_{\text{prod}}\}$. Let $w = (x, y, a)$ be an arbitrary feasible solution, and define $\{I_1(w), I_2(w), I_3(w)\}$ to be the following partition of $\{1, \dots, k_{\text{prod}} - 1\}$:

$$\begin{aligned} I_1(w) &= \{i \in \{1, \dots, k_{\text{prod}} - 1\}\colon a_{i+1} = 1,\ 0 < x_i < 1\},\\ I_2(w) &= \{i \in \{1, \dots, k_{\text{prod}} - 1\}\colon a_{i+1} = 0,\ x_i = 1\},\\ I_3(w) &= \{i \in \{1, \dots, k_{\text{prod}} - 1\}\colon a_{i+1} = 1,\ x_i = 1\}.\end{aligned}$$

If $I_1(w) = \emptyset$ then $x_i = 1$ for all $i \in \{1, \dots, k_{\text{prod}} - 1\}$ and (5) reads as

$$\sum_{i\in B} x_i + \sum_{i\in A}(i - k_0 + 1 - d_{k_0 i})a_{i+1} \ \ge \sum_{t\in B} x_t \ = \ |B|.$$

Since $|B| = |B - A| + |A| \ge \sum_{i\in B\setminus A} d_i + |A|$, the inequality holds.

Assume that $I_1(w) \ne \emptyset$, and define $\hat{w} = (\hat{x}, \hat{y}, \hat{a})$ to be a feasible solution with $\hat{x} = x$, $\hat{y} = y$, and

$$\hat{a}_{k+1} = \begin{cases} a_{k+1} & k \in I_1(w) \cup I_2(w),\\ 0 & k \in I_3(w).\end{cases}$$

Clearly, $\hat{w}$ is a feasible solution with $I_1(\hat{w}) = I_1(w)$ and $I_2(\hat{w}) = I_2(w) \cup I_3(w)$ (hence $I_3(\hat{w}) = \emptyset$). Define $m = \max\{i\colon i \in A \cap I_1(\hat{w})\}$ if $A \cap I_1(\hat{w}) \ne \emptyset$ and $m = k_0 - 1$ otherwise. Then

$$\begin{aligned} &\sum_{i\in A\cap I_1(\hat{w})} \big((1 - \hat{x}_i) - (i - k_0 + 1 - d_{k_0 i})\big)\\ &\quad\le \sum_{i\in A\cap I_1(\hat{w})} (1 - \hat{x}_i) - (m - k_0 + 1 - d_{k_0 m})\end{aligned}$$

$$= \sum_{i \in A \cap I_1(\hat{w})} (1 - \hat{x}_i) - \sum_{i=k_0}^{m} (1 - d_i)$$

$$= \underbrace{\sum_{i \in A \cap I_1(\hat{w})} (d_i - \hat{x}_i)}_{UD} - \underbrace{\sum_{\substack{i=k_0 \\ i \notin A \cap I_1(\hat{w})}}^{m} (1 - d_i)}_{MPS} \leq \underbrace{\sum_{i \in B \setminus A} \hat{x}_i - \sum_{i \in B \setminus A} d_i}_{SS} . \tag{6}$$

The first bounding holds due to the fact that $A \subseteq \{k_0, \ldots, k_{\text{prod}} - 1\}$. The last bounding holds since the expression UD represents the demand in $A \cap I_1(\hat{w})$ unmet with production from the same periods, the expression MPS gives the maximum production surplus in the remaining periods in $\{k_0, \ldots, m\}$, and SS represents the surplus stock after period $k_0 - 1$. We must have $UD \leq SS + MPS$ in order to meet the demands up to the period m. We can now show the validity of (5):

$$\begin{aligned}
&\sum_{i \in B} x_i + \sum_{i \in A} (i - k_0 + 1 - d_{k_0 i}) a_{i+1} \\
&\quad \geq \sum_{i \in B \setminus A} \hat{x}_i + \sum_{i \in A} \hat{x}_i + \sum_{i \in A} (i - k_0 + 1 - d_{k_0 i}) \hat{a}_{i+1} \\
&\quad = \sum_{i \in B \setminus A} \hat{x}_i + \sum_{i \in A \cap I_2(\hat{w})} \hat{x}_i + \sum_{i \in A \cap I_1(\hat{w})} \big(\hat{x}_i + i - k_0 + 1 - d_{k_0 i}\big) \\
&\quad = \sum_{i \in B \setminus A} \hat{x}_i + |A \cap I_2(\hat{w})| + |A \cap I_1(\hat{w})| \\
&\qquad + \sum_{i \in A \cap I_1(\hat{w})} \big((\hat{x}_i - 1) + i - k_0 + 1 - d_{k_0 i}\big) \\
&\quad \geq \sum_{i \in B \setminus A} d_i + |A|.
\end{aligned}$$

In the last inequality we made use of the bound in (6). This allows us to conclude that (5) is a valid inequality for P. □

Remark 1 We have already identified a sub-family of the (k_0, M)-inequalities whose members define facets of P. Namely, if we set $k_0 = 1$, $M = \{1, \ldots, k_{\text{prod}} - 1\} \setminus \{s\}$ for some $s \in \{1, \ldots, k_{\text{prod}} - 1\}$, and $A = B = \{s\}$, then the (k_0, M)-inequality becomes

$$x_s + (s - d_{1s}) a_{s+1} \geq 1.$$

Theorem 4 implies that this inequality is facet-inducing for P.

In Sect. 4 we shall consider the case where the additional assumptions (3) are satisfied and also that the demands force the production to be nonzero in every period. We shall verify that under these assumptions the (k_0, M)-inequalities, together with the model constraints and the relaxed bounds on the binary variables, provide a complete description of P.

In addition, also in Sect. 4, we will see that the constraints (5) are closely related to the well-known (ℓ, S)-inequality for the general case of the uncapacitated lot-sizing problem. In fact, in the proof of Theorem 10 we show how this relationship can be achieved for the special case of high demands.

3 The combinatorial structure of P

In this section we consider a particular graph associated with a polyhedron P, namely the graph whose vertices correspond to the extreme points (i.e., the zero-dimensional faces) of P and where two vertices are adjacent if they belong to the same one-dimensional face of the polyhedron. In Sect. 3.1 we characterize the feasible solutions that are extreme points of P, a fact that is not straightforward due to the presence of continuous variables. In Sect. 3.2 we present some results concerning properties that ensure that two extreme points are neighbors in the associated graph. Finally, in Sect. 3.3 we provide lower and upper bounds on the diameter of the polytope, i.e., the longest distance between any two vertices in the associated graph.

3.1 Basic properties

In Sect. 1 we have denoted by $\mathcal{S}$ the set of feasible solutions for the formulation (1a)–(1f). Now we focus on the extreme points of the polyhedron $P = \text{conv}(S)$. Recall that w is an extreme point of P if there do not exist $w^1, w^2 \in P$, $w^1 \neq w^2$, such that $w = \alpha w^1 + (1-\alpha)w^2$ for some $\alpha \in (0, 1)$. Clearly, if $w \in P \cap \{0, 1\}^{3p}$ then w is an extreme point of P. However, a feasible solution may contain fractional x-variables and still be an extreme point, as the following theorem shows. If $w \in P$ we define $F(w) = \{t \in T : 0 < x_t < 1\}$. In case $F(w) \neq \emptyset$ we consider $F(w) = \{t_1, \ldots, t_k\}$ and $t_{k+1} = p + 1$, where $t_i < t_{i+1}$ for $i \in \{1, \ldots, k\}$. According to this notation we can state the following result.

Theorem 6 *Let $w = (x, y, a) \in S$ with $F(w) \neq \emptyset$. Then, w is an extreme point of P if and only if for every $i \in \{1, \ldots, k\}$, there exists $r \in \{t_i, \ldots, t_{i+1} - 1\}$ such that $x_{1r} = d_{1r}$.*

Proof For the forward implication, let $w = (x, y, a)$ be an extreme point of P. Suppose there exists some $i \in \{1, \ldots, k\}$ such that $x_{1r} > d_{1r}$ for every $r \in \{t_i, \ldots, t_{i+1} - 1\}$. We show that this assumption leads us to a contradiction by dividing our analysis into two cases.

Let $i < k$. We define two points $w^1 = (x^1, y, a)$ and $w^2 = (x^2, y, a)$ such that x^1 and x^2 only differ from x in the periods t_i and t_{i+1}, where they are defined as follows:

$$\begin{aligned} x^1_{t_i} &= x_{t_i} - \varepsilon, & x^2_{t_i} &= x_{t_i} + \varepsilon, \\ x^1_{t_{i+1}} &= x_{t_{i+1}} + \varepsilon, & x^2_{t_{i+1}} &= x_{t_{i+1}} - \varepsilon. \end{aligned}$$

Since $x_{1r} > d_{1r}$ for $r \in \{t_i, \ldots, t_{i+1} - 1\}$, both points belong to P if ε is small enough. Furthermore, $w = (x^1 + x^2)/2$ which implies that w is not an extreme point.

Now, if $i = k$, we define $w^1 = (x^1, y, a)$ and $w^2 = (x^2, y, a)$ such that x^1 and x^2 only differ from x in the period t_i. More precisely,

$$\begin{aligned} x^1_{t_i} &= x_{t_i} - \varepsilon, \\ x^2_{t_i} &= x_{t_i} + \varepsilon. \end{aligned}$$

Again, if ε is small enough then both points belong to P and $w = (w^1 + w^2)/2$, implying that w is not an extreme point.

For the converse implication, if w is not an extreme point of P it can be written as the convex combination of some points $w^1, \ldots, w^n \in P$. Let $w^i = (x^i, y^i, a^i)$ for $i \in \{1, \ldots, n\}$. Since the y- and a-variables are integral, $y^i = y$ and $a^i = a$, for $i = 1, \ldots, n$. The same argument shows that $x^i_t = x_t$ for $t \notin F(w)$.

Consider now the variable x_{t_1}, corresponding to the first index of $F(w)$. By hypothesis, there exists some $r \in \{t_1, \ldots, t_2 - 1\}$ such that $x_{1r} = d_{1r}$. Let $i \in \{1, \ldots, n\}$. Since w and w^i may only differ in their x_{t_1}-variable up to period $t_2 - 1$,

$$x^i_{1r} = x_{1,t_1-1} + x^i_{t_1} + x_{t_1+1,r} = d_{1r} + x^i_{t_1} - x_{t_1}.$$

Moreover, since $d_{1r} + x^i_{t_1} - x_{t_1} \geq d_{1r}$ we have that $x^i_{t_1} \geq x_{t_1}$. Therefore, $x^i_{t_1} \geq x_{t_1}$ for every $i \in \{1, \dots, n\}$. However, since x is a convex combination of $x^1, \dots, x^n$ we conclude $x^i_{t_1} = x_{t_1}$ for every $i \in \{1, \dots, n\}$. By repeating this argument we get $x^i_j = x_j$ for every $j \in F(w)$ and every $i \in \{1, \dots, n\}$. This shows that $w^i = w$ for $i \in \{1, \dots, n\}$ and then w must be an extreme point, i.e., a contradiction. □

Throughout this section we make use of the following well-known characterization of extreme points of a given polyhedron.

Lemma 1 [Nemhauser and Wolsey (1988)] *A feasible solution $w \in P$ is an extreme point of the polytope P if and only if there exists a linear objective function G_w such that w is the unique optimal solution of* $\max\{G_w(x): x \in P\}$.

Let us now introduce a linear function associated with every point in $\mathcal{S}$.

Definition 2 If $\bar{w} = (\bar{x}, \bar{y}, \bar{a}) \in S$, we define the linear function $G_{\bar{w}}$ on P as follows:

$$G_{\bar{w}}(w) = \underbrace{\sum_{j:\bar{y}_j=1} y_j + \sum_{j:\bar{y}_j=0} (1 - y_j)}_{G^1_{\bar{w}}(y)} + \underbrace{\sum_{j:\bar{a}_j=1} a_j + \sum_{j:\bar{a}_j=0} (1 - a_j)}_{G^2_{\bar{w}}(a)}$$
$$+ \underbrace{\sum_{j:\bar{x}_j=1} x_j + \sum_{j:\bar{x}_j=0} (1 - x_j)}_{G^3_{\bar{w}}(x)} + \underbrace{\frac{1}{p^2} \sum_{j \in F(\bar{w})} (p - j + 1)\,(1 - x_j)}_{G^4_{\bar{w}}(x)}$$

for $w \in P$.

Remark 2 The definition of $G_{\bar{w}}$ implies that $G^1_{\bar{w}}(\bar{y}) = G^2_{\bar{w}}(\bar{a}) = p$ and $G^3_{\bar{w}}(\bar{x}) = p - |F(\bar{w})|$. It is clear that $G^1_{\bar{w}}(y) \leq p$, $G^2_{\bar{w}}(a) \leq p$, and $G^3_{\bar{w}}(x) \leq p - |F(\bar{w})|$ for $w \in P$. If $F(\bar{w}) = \emptyset$, then $G_{\bar{w}}(\bar{w}) = 3p$ and $G_{\bar{w}}(w) \leq 3p$ for $w \in P$. If $F(\bar{w}) \neq \emptyset$, it holds that $G_{\bar{w}}(\bar{w}) > 3p - |F(\bar{w})|$. Let us now show that $G^4_{\bar{w}}(x) \leq 1$ for $w = (x, y, a) \in P$. Actually,

$$\begin{aligned}
G^4_{\bar{w}}(x) &= \frac{1}{p^2} \sum_{j \in F(\bar{w})} (p - j + 1)\,(1 - x_j) \\
&\leq \frac{1}{p^2} \sum_{j \in F(\bar{w})} (p - j + 1) \\
&\leq \frac{1}{p^2} \sum_{j \in T} (p - j + 1) = \frac{1}{p^2} \frac{(p+1)p}{2} \leq 1.
\end{aligned}$$

Therefore we have that $G_{\bar{w}}(w) \leq 3p - |F(\bar{w})| + 1$ for any $w \in P$.

Using this linear function we show the following necessary condition for $\bar{w}$ to be an extreme point.

Proposition 2 *If $\bar{w} = (\bar{x}, \bar{y}, \bar{a}) \in S$ is an extreme point of P then $\bar{w}$ is the only optimal solution of* $\max\{G_{\bar{w}}(w): w \in P\}$.

Proof For any $w = (x, y, a) \in \mathcal{S}$ such that $y \neq \bar{y}$ we have that $G^1_{\bar{w}}(y) < p$ and then $G_{\bar{w}}(w) \leq 3p - |F(\bar{w})|$ by Remark 2. This shows that this point cannot be optimal for $G_{\bar{w}}$. A similar situation holds for the a-variable. Therefore, any optimal solution of $\max\{G_{\bar{w}}(w) : w \in P\}$ can only differ from $\bar{w}$ in the x-variables.

Let $w' = (x', \bar{y}, \bar{a}) \in \mathcal{S}$ such that $x' \neq \bar{x}$. If $F(\bar{w}) = \emptyset$ and $G_{\bar{w}}(w') = 3p$ clearly implies that $w' = \bar{w}$. Now, if $F(\bar{w}) \neq \emptyset$ we show that if $A = G_{\bar{w}}(w') - G_{\bar{w}}(\bar{w})$ then $A < 0$. Let $F(\bar{w}) \cup \{p+1\} = \{t_1, \dots, t_k\} \cup \{t_{k+1}\}$. After Theorem 6, for every $i \in \{1, \dots, k\}$ there exists $r_i \in \{t_i, \dots, t_{i+1} - 1\}$ such that $\bar{x}_{1r_i} = d_{1r_i}$. Let

$$r_0 = 0,\ r_{k+1} = p,\ x'_{t_{k+1}} = \bar{x}_{t_{k+1}} = 0,$$

and, for $i \in \{1, \dots, k+1\}$, define $T^0(\bar{w}, r_i) = \{j \in \{r_{i-1}+1, \dots, r_i\} : \bar{x}_j = 0\}$ and $T^1(\bar{w}, r_i) = \{j \in \{r_{i-1}+1, \dots, r_i\} : \bar{x}_j = 1\}$. In addition, consider

$$\begin{aligned}
\varepsilon_1^{(i)} &= \sum_{j \in T^1(\bar{w}, r_i)} \left(x'_j - \bar{x}_j\right), \\
\varepsilon_2^{(i)} &= \sum_{j \in T^0(\bar{w}, r_i)} \left(x'_j - \bar{x}_j\right), \\
\varepsilon_3^{(i)} &= x'_{t_i} - \bar{x}_{t_i}, \\
\varepsilon_{r_i} &= x'_{1r_i} - x_{1r_i} = x'_{1r_i} - d_{1r_i}, \\
\varepsilon_{r_i} - \varepsilon_{r_{i-1}} &= \varepsilon_1^{(i)} + \varepsilon_2^{(i)} + \varepsilon_3^{(i)}.
\end{aligned}$$

By definition we have that if

$$A = \sum_{j:\bar{x}_j=1} \left(x'_j - \bar{x}_j\right) + \sum_{j:\bar{x}_j=0} \left(\bar{x}_j - x'_j\right) + \frac{1}{p^2} \sum_{j \in F(\bar{w})} (p - j + 1) \left(\bar{x}_j - x'_j\right),$$

then

$$\begin{aligned}
A =& \sum_{i=1}^{k+1} \varepsilon_1^{(i)} - \sum_{i=1}^{k+1} \varepsilon_2^{(i)} - \frac{1}{p^2} \sum_{i=1}^{k} (p - t_i + 1)\, \varepsilon_3^{(i)} \\
=& \sum_{i=1}^{k+1} \varepsilon_1^{(i)} - \sum_{i=1}^{k+1} \varepsilon_2^{(i)} - \frac{1}{p^2} (p - t_1 + 1) \left(\varepsilon_{r_1} - \varepsilon_1^{(1)} - \varepsilon_2^{(1)}\right) \\
& - \frac{1}{p^2} \sum_{i=2}^{k} (p - t_i + 1) \left(\varepsilon_{r_i} - \varepsilon_{r_{i-1}} - \varepsilon_1^{(i)} - \varepsilon_2^{(i)}\right).
\end{aligned}$$

Or, equivalently,

$$\begin{aligned}
A =& \sum_{i=1}^{k} \left(1 - \frac{p - t_i + 1}{p^2}\right) \varepsilon_1^{(i)} + \varepsilon_1^{(k+1)} - \sum_{i=1}^{k} \left(1 - \frac{p - t_i + 1}{p^2}\right) \varepsilon_2^{(i)} - \varepsilon_2^{(k+1)} \\
& - \frac{1}{p^2} \sum_{i=1}^{k-1} (t_{i+1} - t_i)\, \varepsilon_{r_i} - \frac{1}{p^2} (p - t_k + 1)\, \varepsilon_{r_k}
\end{aligned}$$

$$=\sum_{i=1}^{k+1}\left(1-\frac{p-t_i+1}{p^2}\right)\left(\varepsilon_1^{(i)}-\varepsilon_2^{(i)}\right)-\frac{1}{p^2}\sum_{i=1}^{k-1}(t_{i+1}-t_i)\,\varepsilon_{r_i}$$
$$-\frac{1}{p^2}(p-t_k+1)\,\varepsilon_{r_k}. \tag{7}$$

Now, by definition we have that $\varepsilon_1^{(i)} \leq 0$ and $\varepsilon_2^{(i)} \geq 0$ for $i \in \{1,\dots,k+1\}$. This implies that $(\varepsilon_1^{(i)}-\varepsilon_2^{(i)}) \leq 0$. By combining these observations with the fact that $(1-\frac{p-t_i+1}{p^2}) > 0$ for $i \in \{1,\dots,k+1\}$, we can conclude that $A \leq 0$.

Moreover, $A = 0$ if and only if $\varepsilon_1^{(i)} = \varepsilon_2^{(i)} = 0$ for $i \in \{1,\dots,k+1\}$ and $\varepsilon_{r_i} = 0$ for $i \in \{1,\dots,k\}$, that is, if $x' = \bar{x}$. This proves that $G_{\bar{w}}(w') < G_{\bar{w}}(\bar{w})$ for $w' \neq \bar{w}$.

If $F(\bar{w}) = \emptyset$ then $G_{\bar{w}}(\bar{w}) = 3p$ and, furthermore, $G_{\bar{w}}(w) \leq 3p$ for every feasible solution w. Since $\bar{w} \in \{0,1\}^{3p}$, any $w' \in P$ attaining $G_{\bar{w}}(w') = 3p$ must coincide with $\bar{w}$ thus implying that $w' = \bar{w}$. □

3.2 Neighboring extreme points

Proposition 2 provides the starting point for exploring the edges of the polytope P. Recall that e_i stands for the i-th unit vector of appropriate dimension.

Definition 3 Two extreme points of a polyhedron P are *neighbors* if they belong to the same one-dimensional face of P.

Remark 3 Two extreme points $\bar{w}$ and $\bar{w}'$ of P are neighbors if and only if there exists a linear function G such that they are the unique extreme points that are optimal solutions of $\max\{G(w): w \in P\}$. That is, the set $\{w \in P: G(w) = G(\bar{w})\}$ is a one-dimensional face of the polytope P.

In the following three propositions we consider particular pairs of extreme points of P which turn out to be neighbors.

Proposition 3 *Let $w^i = (x^i, y^i, a^i)$ be an extreme point of P with $y_i = 0$ and $i \in \{2,\dots,p\}$. Then $\bar{w}^i = (x^i, y^i + e_i, a^i)$ is an extreme point of P and w^i and $\bar{w}^i$ are neighbors.*

Proof It is clear that $\bar{w}^i$ is an extreme point of P. For any $w = (x, y, a) \in P$, let us consider the function $G_i(w) = G_{\bar{w}^i}(w) - y_i$, i.e., the functions G_i and $G_{\bar{w}^i}$ coincide except for the coefficient of the y_i-variable, which is not present in G_i. In order to show that w^i and $\bar{w}^i$ are neighbors, we will prove that the points w^i and $\bar{w}^i$ are the unique extreme points that are optimal solutions of $\max\{G_i(w): w \in P\}$.

For $w \in P$, $G_i(w)$ takes the following form

$$\begin{aligned}G_i(w) &= G_{\bar{w}^i}(w) - y_i\\ &= G^1_{\bar{w}^i}(y) - y_i + G^2_{\bar{w}^i}(a) + G^3_{\bar{w}^i}(x) + G^4_{\bar{w}^i}(x)\\ &= \sum_{\substack{j\neq i\\ \bar{y}^i_j=1}} y_j + \sum_{j:\bar{y}^i_j=0}(1-y_j) + G^2_{\bar{w}^i}(a) + G^3_{\bar{w}^i}(x) + G^4_{\bar{w}^i}(x).\end{aligned}$$

It is easy to prove that

$$G_{\bar{w}^i}(w) - G_{w^i}(w) = 1 \tag{8}$$

and

$$G_i(w^i) = G_i(\bar{w}^i). \tag{9}$$

We define the operator $H^i : \mathcal{S} \to \mathbb{R}^{3p}$ to be $H^i(w) = \widehat{w}^i = (\widehat{x}^i, \widehat{y}^i, \widehat{a}^i)$, with

$$\begin{aligned} \hat{x}^i &= x, \quad \hat{a}^i = a, \\ \hat{y}^i_i &= 1, \quad \hat{y}^i_j = y_j \quad j \neq i. \end{aligned}$$

Note that if $y_i = 1$ then $H^i(w) = w$.

If $w \in \mathcal{S}$ and $\widehat{w}^i = H^i(w)$ then

$$G_i(w) = G_{\bar{w}^i}(\widehat{w}^i) - 1. \tag{10}$$

Since $\bar{w}^i$ is optimal for $G_{\bar{w}^i}$, we have $G_{\bar{w}^i}(\bar{w}^i) \geq G_{\bar{w}^i}(w)$ for all $w \in P$, implying

$$G_i(w) \leq \ G_{\bar{w}^i}(\bar{w}^i) - 1 = \ G_i(\bar{w}^i). \tag{11}$$

By (9) we have that w^i and $\bar{w}^i$ satisfy (11) at equality, hence G_i achieves its maximum value at these points. If there were another extreme point $w^* \in P$ verifying $G_i(w^*) = G_i(\bar{w}^i)$ then $y^*_i = 0$ or $y^*_i = 1$. On the one hand, if $y^*_i = 1$, then $G_i(w) = G_{\bar{w}^i}(\widehat{w}^i) - 1$ and Proposition 2 imply that

$$G_i(w^*) = G_{\bar{w}^i}(w^*) - 1 < G_{\bar{w}^i}(\bar{w}^i) - 1 = G_i(\bar{w}^i),$$

a contradiction. On the other hand, if $y^*_i = 0$ then by combining (8), (10) and Proposition 2 we get

$$G_i(w^*) = G_{\bar{w}^i}(w^*) = G_{\bar{w}}(w^*) + 1 < G_{\bar{w}}(\bar{w}) + 1 = G_{\bar{w}^i}(\bar{w}) = G_i(\bar{w}),$$

which is also a contradiction. This shows that $\bar{w}^i$ and $\bar{w}^i$ are neighbors. □

The following results are proved in much the same way as Proposition 3, hence the proofs are omitted.

Proposition 4 *Let $w^i = (x^i, y^i, a^i)$ be an extreme point of P with $a^i_i = 0$ and $i \in \{2, \ldots, p\}$. Then $\bar{w}^i = (\bar{x}^i, \bar{y}^i, \bar{a} + e_i)$ is an extreme point of P and w^i and $\bar{w}^i$ are neighbors.*

Proposition 5 *Let $w^i = (x^i, y^i, a^i) \in \{0, 1\}^{3p}$ be an extreme point of P with $x^i_i = 0$ and $i \in \{k_{\text{prod}} + 1, \ldots, p\}$. We define $\bar{w}^i = (x^i + e_i, \bar{y}^i, a^i)$ such that $\bar{y}^i_i = 1$ and $\bar{y}^i_j = y^i_j$ for every $j \neq i$. Then $\bar{w}^i$ is also an extreme point of P and w^i and $\bar{w}^i$ are neighbors.*

Proposition 6 *Let $w^i = (x^i, y^i, a^i)$ be an extreme point of P with $F(w^i) = \{t_1, \ldots, t_k\}$ where $i \in \{r_k + 1, \ldots, p\}$ such that $x^i_i = 0$ for $r_k \in \{t_k, \ldots, p\}$ such that $x^i_{1r_k} = d_{1r_k}$. We define $\bar{w}^i = (x^i + e_i, \bar{y}^i, a^i)$ such that $\bar{y}^i_i = 1$ and $\bar{y}^i_s = y^i_s$ for $s \in T \setminus \{i\}$. Then $\bar{w}^i$ is also an extreme point and w^i and $\bar{w}^i$ are neighbors.*

In what follows, if $i = p$ we assume that $\{i + 1, \ldots, p\} = \emptyset$.

The following result shows under which conditions we can move from an extreme point to a neighboring extreme point by setting a non-zero and non-saturated x-variable to 1.

Lemma 2 *Let $w^i = (x^i, y^i, a^i)$ be an extreme point of P satisfying $0 < x^i_i < 1$ and $x^i_j \in \{0, 1\}$ for $j \in \{i + 1, \ldots, p\}$. Consider $\bar{w}^i = (\bar{x}^i, \bar{y}^i, \bar{a}^i)$ such that $\bar{y}^i = y^i$, $\bar{a}^i = a^i$, $\bar{x}^i_k = x^i_k$ for every k except for $\bar{x}^i_i = 1$. Then $\bar{w}^i$ is also an extreme point of P.*

Proof Let $F(w^i) = \{t \in T : 0 < x_t^i < 1\}$ and $F(w^i) \cup \{p+1\} = \{t_1, \ldots, t_{k+1}\}$ be as in Theorem 6, where $t_s < t_{s+1}$ for $s \in \{1, \ldots, k\}$ and $t_{k+1} = t_{p+1}$. By hypothesis, $t_k = i$. Since w^i is an extreme point of P, by Theorem 6, we have $x_{1r_s}^i = d_{1r_s}$ for every $s \in \{1, \ldots, k\}$ and for some $r_s \in \{t_s, \ldots, t_{s+1} - 1\}$. Since $w^i \in P$, the fact that $0 < x_i^i < 1$ implies that $\bar{y}_i = 1$ shows $\bar{w}^i \in P$.

Observe that $F(\bar{w}^i) = \{t \in T : 0 < \bar{x}_t^i < 1\} = F(\bar{w}) \setminus \{t_k\}$ and $F(\bar{w}^i) \cup \{p+1\} = \{t_1, \ldots, t_{k-1}\} \cup \{t_{k+1}\}$. Since $\bar{x}_{1j}^i = \bar{x}_{1j}$ for $j \in \{1, \ldots, t_k - 1\}$, we have that $\bar{x}_{1r_s}^i = d_{1r_s}$ for every $s \in \{1, \ldots, k-2\}$ and for some $r_s \in \{t_s, \ldots, t_{s+1} - 1\}$. Moreover, there exists $r_{k-1} \in \{t_{k-1}, \ldots, t_{k+1} - 1\}$ such that $\bar{x}_{1r_{k-1}}^i = d_{1r_{k-1}}$. By Theorem 6, we conclude that $\bar{w}^i$ is an extreme point of P. □

Using this last result we exhibit another pair of neighboring extreme points.

Proposition 7 *Let $w^i = (x^i, y^i, a^i)$ be an extreme point of P with $0 < x_i^i < 1$ and $x_j^i \in \{0, 1\}$ for $j \in \{i+1, \ldots, p\}$. Consider $\bar{w}^i = (\bar{x}^i, \bar{y}^i, \bar{a}^i)$ such that $\bar{y}^i = y^i$, $\bar{a}^i = a^i$, $\bar{x}_k^i = x_k^i$ for every k except for $\bar{x}_i^i = 1$. Then $\bar{w}^i$ is also an extreme point of P. Moreover, if $a_i^i = 1$ then w^i and $\bar{w}^i$ are neighbors.*

Proof By Lemma 2, $\bar{w}^i$ is an extreme point of P. In order to show that w^i and $\bar{w}^i$ are neighbors we define a particular objective function G_i over P and prove that the points w^i and $\bar{w}^i$ are the unique extreme points that are optimal solutions of $\max\{G_i(w) : w \in P\}$.

For any $w = (x, y, a) \in P$ we consider the function $G_i(w) = G_{\bar{w}^i}(w) - : x_i$, i.e., the functions G_i and $G_{\bar{w}^i}$ coincide except for the coefficient corresponding to the x_i-variable, which is not present in G_i.

For any $w \in P$, we have

$$\begin{aligned} G_i(w) &= G_{\bar{w}^i}(w) - x_i \\ &= G^1_{\bar{w}^i}(y) + G^2_{\bar{w}^i}(a) + G^3_{\bar{w}^i}(x) + G^4_{\bar{w}^i}(x) - x_i \\ &= G^1_{\bar{w}^i}(y) + G^2_{\bar{w}^i}(a) + \sum_{\substack{j \neq i \\ \bar{x}_j^i = 1}} x_j + \sum_{j : \bar{x}_j^i = 0} (1 - x_j) + G^4_{\bar{w}^i}(x). \end{aligned} \tag{12}$$

It is easy to prove that

$$G_i(w^i) = G_i(\bar{w}^i). \tag{13}$$

We define the operator $H^i : \mathcal{S} \to \mathbb{R}^{3p}$ to be $H^i(w) = \widehat{w}^i = (\widehat{x}^i, \widehat{y}^i, \widehat{a}^i)$ with

$$\begin{aligned} &\widehat{x}_j^i = x_j, \quad \widehat{y}_j^i = y_j, \quad \widehat{a}_j^i = a_j \quad j \neq i \quad \text{and} \\ &\widehat{x}_i^i = \widehat{y}_i^i = \widehat{a}_i^i = 1. \end{aligned} \tag{14}$$

Note that if $x_i = y_i = a_i = 1$ then $H^i(w) = w$.

If $w \in \mathcal{S}$ and $\widehat{w}^i = H^i(w)$ then

$$\begin{aligned} G_i(w) \leq\ & G^1_{\bar{w}^i}(y) + (1 - y_i) + G^2_{\bar{w}^i}(a) + (1 - a_i) \\ &+ \sum_{\substack{j \neq i \\ \bar{x}_j^i = 1}} x_j + (1 - \bar{x}_i^i) + \sum_{j : \bar{x}_j^i = 0} (1 - x_j) + G^4_{\bar{w}^i}(x) \\ =\ & G_{\bar{w}^i}(\widehat{w}^i) - \bar{x}_i^i \end{aligned} \tag{15}$$

Since $\bar{w}^i$ is optimal for $G_{\bar{w}^i}$, $G_{\bar{w}^i}(\bar{w}^i) \geq G_{\bar{w}^i}(w)$ for all $w \in P$ and then

$$G_i(w) \leq \ G_{\bar{w}^i}(\bar{w}^i) - 1 = \ G_i(\bar{w}^i) \tag{16}$$

If we combine the inequality in (16) for $w = w^i$ with (13), we have that G_i achieves it maximum value at w^i and $\bar{w}^i$.

Assume that there is another extreme point of P, namely $w^* = (x^*, y^*, a^*)$, verifying $G_i(w^*) = G_i(\bar{w}^i)$. We divide our analysis into the following cases:

Case 1 $x_i^* = y_i^* = a_i^* = 1$.
By using the operator in (14) we obtain that $H^i(w^*) = w^*$ is an extreme point different from $\bar{w}^i$. By Proposition 2, $G_{\bar{w}^i}(w^*) < G_{\bar{w}^i}(\bar{w}^i)$. From (15) we have

$$G_i(w^*) \leq G_{\bar{w}^i}(w^*) - \bar{x}_i^i < G_{\bar{w}^i}(\bar{w}^i) - \bar{x}_i^i = G_i(\bar{w}^i),$$

that is, a contradiction.
Case 2 $y_i^* = 0$ or $a_i^* = 0$.
By using the operator in (14) we have

$$\begin{aligned} G_i(w^*) < \ & G^1_{\bar{w}^i}(y^*) \ + (1 - y_i^*) + \ G^2_{\bar{w}i}(a^*) + (1 - a_i^*) + \sum_{\substack{j \neq i \\ \bar{x}_j^i = 1}} x_j^* \\ & + (1 - \bar{x}_i^i) + \sum_{j : \bar{x}_j^i = 0} (1 - x_j^*) + G^4_{\bar{w}^i}(x^*) \\ = \ & G_{\bar{w}^i}(H^i(w^*)) - \bar{x}_i^i. \end{aligned} \tag{17}$$

After (17) and Proposition 2 we have

$$G_i(w^*) < G_{\bar{w}^i}(H^i(w^*)) - \bar{x}_i^i \leq G_{\bar{w}^i}(\bar{w}^i) - \bar{x}_i^i = G_i(\bar{w}^i).$$

Again, a contradiction.
Case 3 $y_i^* = a_i^* = 1$ and $x_i^* < 1$.
Since w^* is different from w^i and $\bar{w}^i$ there must exist $k \neq i$ such that $y_k^* \neq y_k^i$ or $a_k^* \neq a_k^i$ or $x_k^* \neq x_k^i$. Indeed, if we assume that $y_k^* = y_k^i$, $a_k^* = a_k^i$ and $x_k^* = x_k^i$ for every $k \neq i$, then we can either have (a) $0 \leq x_i^* < x_i^i$ or (b) $\bar{x}_i < x_i^* < 1$.
If (a) holds then w^i would be a convex combination of w^* and $\bar{w}^i$. On the other hand, if (b) is satisfied, w^* would be a convex combination of w^i and $\bar{w}^i$. In both cases, we arrive at a contradiction.
Using Lemma 2 and the definition of H^i in (14), we have that $H^i(w^*)$ is an extreme point of P. Therefore, $H^i(w^*)$ and w^* coincide except for the x_i-variable. After Proposition 2,

$$G_i(w^*) = G_i(H^i(w^*)) < G_{\bar{w}^i}(\bar{w}^i) - \bar{x}_i^i = G_i(\bar{w}^i),$$

again, a contradiction.

This completes the proof that w^i and $\bar{w}^i$ are neighbors. □

3.3 Bounding the diameter of P

The results in the previous section allow us to bound the diameter of the polytope P. Studying this parameter is of theoretical interest, since the diameter provides a lower bound on the maximum number of iterations performed by any implementation of the simplex method on the convex hull of feasible solutions. Such an execution is theoretically beyond reach unless $P = NP$, so studying this parameter has—in principle—no practical consequences. Nevertheless, the diameter of a polytope is a measure of the connectivity properties of the polytope, and some polytopes arising in the context of integer programming formulations for combinatorial optimization problems have been shown to have low diameters. The results in this section show that this is also the case for the lot-sizing polytope studied in this work.

Given a polyhedron Q, we define an associated graph $G(Q)$ whose vertices are the extreme points of Q, and whose edges correspond to vertex pairs that belong to the same one-dimensional face of Q. That is, two extreme points are joined by an edge in $G(Q)$ if they are neighbors in the polyhedron. Taking this definition into account we use the notions of path, distance and diameter from graph theory.

Definition 4 Given v and w two extreme points of a polyhedron Q, a *path* (of length k) from v to w is a sequence $v = v_0, v_1, \ldots, v_k = w$ of extreme points such that v_{i-1} and v_i are neighbors for every $i \in \{1, \ldots, k\}$. The *distance from v to w* is the length of a shortest such path in the polyhedron Q and it is denoted by $\operatorname{dist}(v, w)$. If Q is a bounded polyhedron (i.e., a polytope), the *diameter of Q*, $\operatorname{diam}(Q)$, is the maximum distance between any two extreme points of Q.

Recall that throughout this paper we are assuming that $d_1 > 0$ and then $y_1 = a_1 = 1$.

Theorem 7 *If $P = \operatorname{conv}(\mathcal{S})$ where $\mathcal{S}$ is described by the inequalities (1a)–(1f) then*

$$\operatorname{diam}(P) \leq 4p - 2.$$

Proof Let $w = (x, y, a)$ and $w' = (x', y', a')$ be two extreme points of P. By applying Proposition 4 we get that the distance between w and $\hat{w} = (x, y, \mathbf{1})$ is at most $p - 1$ (since $\bar{a}_1 = 1$). If $F(\hat{w}) = \{t_1, \ldots, t_k\}$, we apply successively Proposition 7 starting at $\hat{w}$ for $i = t_k$ and over all the indices in $F(\hat{w})$. In this way we build a path of neighboring extreme points between $\hat{w}$ and $\tilde{w} = (\tilde{x}, y, \mathbf{1})$, where

$$\tilde{x}_j = \begin{cases} 1 & j \in F(\hat{w}), \\ \bar{x}_j & \text{elsewhere.} \end{cases}$$

This shows that $\operatorname{dist}(\hat{w}, \tilde{w}) \leq |F(\hat{w})|$ and then $\operatorname{dist}(w, \tilde{w}) \leq p + |F(\hat{w})| - 1$. Finally, we build a path of neighboring extreme points between $\tilde{w}$ and $\mathbf{1}$ by applying Proposition 5 to the null $x-$variables in $\tilde{w}$ and we get that the distance between $\tilde{w}$ and $\mathbf{1}$ is at most $p - |F(\hat{w})|$. Then

$$\operatorname{dist}(w, \mathbf{1}) \leq p + |F(\hat{w})| - 1 + p - |F(\hat{w})| = 2p - 1.$$

The same reasoning shows that $\operatorname{dist}(w', \mathbf{1}) \leq 2p - 1$. Therefore, $\operatorname{dist}(w, w') \leq 4p - 2$ for any pair w, w' of extreme points of P. Then, $\operatorname{diam}(P) \leq 4p - 2$. □

Now we focus on finding lower bounds for the diameter of P. To this end, we first present preliminary results studying conditions that ensure that two extreme points are not neighbors. We first show that this is the case if two points differ by two y-variables.

Proposition 8 *Let $w^{ij} = (x^{ij}, y^{ij}, a^{ij})$ be an extreme point of P such that $y_i^{ij} = y_j^{ij} = 0$ and $i \neq j$ in T. We define $\bar{w}^{ij} = (x^{ij}, \bar{y}^{ij}, a^{ij})$ such that $\bar{y}_i^{ij} = \bar{y}_j^{ij} = 1$ and $\bar{y}_s^{ij} = y_s^{ij}$ for $s \in T \setminus \{i, j\}$. Then $\bar{w}^{ij}$ is also an extreme point and w^{ij} and $\bar{w}^{ij}$ are not neighbors.*

Proof Clearly, $\bar{w}^{ij}$ is an extreme point. If w^{ij} and $\bar{w}^{ij}$ were neighbors, there would be a linear function f of the form $f(w) = cw$ for which they are the only optimal solutions in P. In particular, $f(w^{ij}) = f(\bar{w}^{ij}) = \max\{f(w)\colon w \in P\}$. If c_{y_s} is the coefficient of the y_s-variable then $c_{y_i} + c_{y_j} = 0$. If $c_{y_i} > 0$ then $\hat{w} = (x^{ij}, y^{ij} + e_i, a^{ij})$ would satisfy $f(\hat{w}) > f(w^{ij})$ contradicting the optimality of w^{ij}. Analogously, we have $c_{y_j} \leq 0$. Then, $c_{y_i} = c_{y_j} = 0$.

However, this fact implies that $\hat{w}$ is also an optimal solution, contradicting the fact that w^{ij} and $\bar{w}^{ij}$ are the only two optimal solutions for f in P. □

A similar result holds for the a-variables, and we omit the proof since it goes along the same arguments as Proposition 8.

Proposition 9 *Let $w^{ij} = (x^{ij}, y^{ij}, a^{ij})$ be an extreme point of P such that $a_i^{ij} = a_j^{ij} = 0$ and $i \neq j$ in T. We define $\bar{w}^{ij} = (x^{ij}, y^{ij}, \bar{a}^{ij})$ such that $\bar{a}_i^{ij} = \bar{a}_j^{ij} = 1$ and $\bar{a}_s^{ij} = a_s^{ij}$ for $s \in T \setminus \{i, j\}$. Then $\bar{w}^{ij}$ is also an extreme point and w^{ij} and $\bar{w}^{ij}$ are not neighbors.*

The following result provides conditions ensuring non-neighborhood and applies to the x-variables. Recall that for $w \in P$, $F(w) = \{s \in T : 0 < x_s < 1\}$. If it is nonempty we assume that $F(w) = \{t_1, \ldots, t_k\}$ and $t_{k+1} = p + 1$.

Proposition 10 *Let $w^{ij} = (x^{ij}, y^{ij}, a^{ij})$ be an extreme point of P with $F(w^{ij}) = \{t_1, \ldots, t_k\}$ and $i \neq j$ in the set $\{r_k + 1, \ldots, p\}$ such that $x_i^{ij} = x_j^{ij} = 0$ and $r_k \in \{t_k, \ldots, p\}$ satisfies $x_{1r_k}^{ij} = d_{1r_k}$. We define $\bar{w}^{ij} = (x^{ij} + e_i + e_j, \bar{y}^{ij}, a^{ij})$ such that $\bar{y}_i^{ij} = \bar{y}_j^{ij} = 1$ and $\bar{y}_s^{ij} = y_s^{ij}$ for $s \in T \setminus \{i, j\}$. Then $\bar{w}^{ij}$ is also an extreme point of P and w^{ij} and $\bar{w}^{ij}$ are not neighbors.*

Proof Let $i, j > r_k$. Since $F(\bar{w}^{ij}) = F(w^{ij})$ and $\bar{x}_{1r_s}^{ij} = x_{1r_s}^{ij} = d_{1r_s}$ for every $s \in F(\bar{w}^{ij})$, then, by Theorem 6, $\bar{w}^{ij}$ is an extreme point.

If w^{ij} and $\bar{w}^{ij}$ were neighbors, there would be a linear function f of the form $f(w) = cw$ for which they are the only optimal solutions in P. In particular, $f(w^{ij}) = f(\bar{w}^{ij}) = \max\{f(w) : w \in P\}$. Then if c_{x_s} is the coefficient of the x_s-variable and c_{y_s} is the coefficient of the y_s-variable, we have that $c_{x_i} + c_{x_j} + c_{y_i}(1 - y_i^{ij}) + c_{y_j}(1 - y_j^{ij}) = 0$.

With a similar reasoning as in the proof of Proposition 8 we obtain $c_{y_i} = c_{y_j} = 0$.

If $c_{x_i} > 0$ then $\hat{w} = (\hat{x}, \hat{y}, a^{ij})$ with $\hat{x}_i = \hat{y}_i = 1$, $\hat{x}_s = x_s^{ij}$ for $s \in T \setminus \{i\}$ and $\hat{y}_s = y_s^{ij}$ for $s \in T \setminus \{i\}$ would satisfy $f(\hat{w}) > f(w^{ij})$ contradicting the optimality of w^{ij}. Similarly, $c_{x_j} \leq 0$. Then, $c_{x_i} = c_{x_j} = 0$.

However, under these conditions $\hat{w}$ would be an optimal solution, contradicting the fact that w^{ij} and $\bar{w}^{ij}$ are the only two optimal solutions for f in P. □

The following result can be proved using similar techniques, hence its proof is omitted.

Proposition 11 *Let $w^{ij} = (x^{ij}, y^{ij}, a^{ij})$ be an extreme point of P with $F(w^{ij}) = \{t_1, \ldots, t_k\}$. Assume that $i = t_k$, $j \in \{r_k + 1, \ldots, p\}$ are such that $x_i^{ij} < 1$ and $x_j^{ij} = 0$, where $r_k \in \{t_k, \ldots, p\}$ satisfies $x_{1r_k}^{ij} = d_{1r_k}$. We define $\bar{w}^{ij} = (\bar{x}^{ij}, \bar{y}^{ij}, a^{ij})$ such that $\bar{x}_i^{ij} = \bar{x}_j^{ij} = 1$, $\bar{x}_s^{ij} = x_s^{ij}$ for $s \in T \setminus \{i, j\}$, $\bar{y}_i^{ij} = \bar{y}_j^{ij} = 1$ and $\bar{y}_s^{ij} = y_s^{ij}$ for $s \in T \setminus \{i, j\}$. Then $\bar{w}^{ij}$ is also an extreme point of P and w^{ij} and $\bar{w}^{ij}$ are not neighbors.*

We present a final preliminary proposition in order to prove the main result in this section.

Proposition 12 *Let $w^{ij} = (x^{ij}, y^{ij}, a^{ij})$ be an extreme point of P with $F(w^{ij}) = \{t_1, \ldots, t_k\}$. Assume $i = t_{k-1}$ and $j = t_k$. We define $\bar{w}^{ij} = (\bar{x}^{ij}, y^{ij}, a^{ij})$ such that $\bar{x}_i^{ij} = \bar{x}_j^{ij} = 1$ and $\bar{x}_s^{ij} = x_s^{ij}$ for $s \in T \setminus \{i, j\}$. Then $\bar{w}^{ij}$ is also an extreme point of P and w^{ij} and $\bar{w}^{ij}$ are not neighbors.*

Proof Let $i = t_{k-1}$ and $j = t_k$. Since $F(\bar{w}^{ij}) = F(w^{ij}) \setminus \{t_{k-1}, t_k\}$ and $\bar{x}_{1r_s}^{ij} = x_{1r_s}^{ij} = d_{1r_s}$ for every $s \in F(\bar{w}^{ij})$, then, by Theorem 6, $\bar{w}^{ij}$ is an extreme point.

If w^{ij} and $\bar{w}^{ij}$ were neighbors, there would be a linear function f of the form $f(w) = cw$ for which they are the only optimal solutions in P. In particular, $f(w^{ij}) = f(\bar{w}^{ij}) = \max\{f(w) : w \in P\}$. Then if c_{x_s} is the coefficient of the x_s-variable we have that $c_{x_i}(1 - x_i^{ij}) + c_{x_j}(1 - x_j^{ij}) = 0$.

If $c_{x_j} > 0$ then $\hat{w} = (\hat{x}^{ij}, y^{ij}, a^{ij})$ with $\bar{x}_j^{ij} = 1$ and $\bar{x}_s^{ij} = x_s^{ij}$ for $s \in T \setminus \{j\}$ would satisfy $f(\hat{w}) = f(w^{ij}) + c_{x_j}(1 - x_j^{ij}) > f(w^{ij})$ contradicting the optimality of w^{ij}. This implies that $c_{x_j} \leq 0$.

Now we assume that $c_{x_i} > 0$. Let $\varepsilon = 1 - x^{ij}$ and we define $\tilde{w} = (\tilde{x}^{ij}, y^{ij}, a^{ij})$ with $\tilde{x}_i^{ij} = 1$, $\tilde{x}_j^{ij} = (x_j^{ij} - \varepsilon)^+$ and $\tilde{x}_s^{ij} = x_s^{ij}$ for $s \in T \setminus \{i, j\}$. After Theorem 6, it is clear that $\tilde{w}$ is an extreme point of P. Then $\tilde{w}$ would satisfy

$$f(\tilde{w}) = f(w^{ij}) + c_{x_i}(1 - x_i^{ij}) + c_{x_j}((x_j^{ij} - \varepsilon)^+ - x_j^{ij}) > f(w^{ij})$$

contradicting the optimality of w^{ij}.

Then, $c_{x_i} = c_{x_j} = 0$. However, this implies that $\hat{w}$ is also an optimal solution, contradicting the fact that w^{ij} and $\bar{w}^{ij}$ are the only two optimal solutions for f in P. □

We are now in position of presenting the main result of this section, which provides both lower and upper bounds on the diameter of the polytope P. Let $\mathbb{P}$ be the family of the lot-sizing polytopes with continuous start-up costs and p periods.

Theorem 8 *If $\operatorname{diam}(\mathbb{P}) = \max_{Q \in \mathbb{P}} \operatorname{diam}(Q)$, then $p \leq \operatorname{diam}(\mathbb{P}) \leq 4p - 2$.*

Proof After Theorem 7 it only remains to show the lower bound. To this end, we present an instance for which P satisfies $\operatorname{diam}(P) \geq p$.

Consider $0 < \varepsilon < \frac{1}{p}$ and $d_i = \varepsilon$ for every $i \in \{1, \ldots, p\}$. Let $\bar{w} = (\varepsilon \mathbf{1}, \mathbf{1}, \mathbf{1})$. Note that in this case $F(\bar{w}) = T$ and $\bar{x}_{1i} = d_{1i}$ for every $i \in T$. Under the notation in Theorem 6, $r_i = i$ for every $i \in T$ and then, $\bar{w}$ is a extreme point of P.

We can apply Proposition 7 successively starting from $\bar{w}$ for $i = p$ over all the indices in T backwards and obtain a path of neighboring extreme points connecting $\bar{w}$ with the extreme point $\mathbf{1} \in P$. Then, $\operatorname{dist}(\bar{w}, \mathbf{1}) \leq p$. By Proposition 12 we get that the distance between $\bar{w}$ and $\mathbf{1}$ is exactly equal to p. Hence, we conclude that $\operatorname{diam}(P) \geq p$, hence $p \leq \operatorname{diam}(\mathbb{P})$. □

4 The polytope P for high demands

In this section we analyze the particular case of high demands implying nonzero production in every period. More precisely, we assume that $d_t \leq 1$ for every $t \in T$ but such that $d_1 < 1$ and $k_{\text{prod}} = p$. This ensures that $y_t = 1$ for every $t \in T$, and we show that the resulting polyhedral structure is much simplified in this case. We provide a complete characterization of P in terms of linear inequalities, and we show that the inequalities in such a characterization are separable in polynomial time, implying the existence of a polynomial-time algorithm for this particular lot-sizing situation.

The key addition to the model constraints is given by the family of valid inequalities described in Definition 1 with $k_{\text{prod}} = p$. The main result of this section asserts that the model constraints reinforced with these inequalities provides a complete characterization of P in this case. It is important to note that not all the inequalities (5) define facets of P, hence the characterization given by Theorem 10 below is—in this sense—redundant.

We now present two results by Pochet and Wolsey (2006), which are used to prove the main result of this section. Let us denote by X^{LS-U} [see Pochet and Wolsey (2006)] the standard uncapacitated lot-sizing model, i.e., the set of feasible solutions to:

$$s_{t-1} + x_t = d_t + s_t \qquad t \in \{1, \dots, p\}, \tag{18a}$$

$$x_t \leq My_t \qquad t \in \{1, \dots, p\}, \tag{18b}$$

$$s \in \mathbb{R}_+^{p+1}, \qquad x \in \mathbb{R}^p, \ y \in [0,1]^p, \tag{18c}$$

$$y \in \mathbb{Z}^p, \tag{18d}$$

$$s_0 = s_0^*, s_p = s_p^*, \tag{18e}$$

where

x_t: the amount produced in period t,
s_t: the amount in stock at the end of period t,
y_t: the 0-1 set-up variable which must have the value 1 if $x_t > 0$,
M: a large positive number and
s_0^*, s_p^*: nonnegative real numbers (fixed).

Proposition 13 [Pochet and Wolsey (2006)] *Let $\ell \in \{1, \dots, p\}$, $L = \{1, \dots, \ell\}$ and $S \subseteq L$, then the (ℓ, S)-inequality*

$$\sum_{j \in S} x_j \leq \sum_{j \in S} d_{j\ell} y_j + s_\ell \tag{19}$$

is valid for X^{LS-U}.

Theorem 9 [Pochet and Wolsey (2006)] *When $s_0 = s_p = 0$, the original constraints* (18a)–(18c) *plus the (ℓ, S)-inequalities* (19) *give a complete linear inequality description of* $\mathrm{conv}(X^{LS-U})$.

These results help us to provide a complete description of the polytope P for a particular situation described previously.

Theorem 10 *Let $0 < d_i \leq 1$ for $i \in T$ such that $d_1 < 1$ and $k_{\text{prod}} = p$. Then, the polytope P is completely described by the constraints* (1a), (1d), *$x, y, a \in [0,1]$ and the (k_0, M)-inequalities* (5) *for every $k_0 \geq 1$ and $M \subseteq \{k_0, \dots, p-1\}$.*

Proof Consider the linear system defined by the constraints (1a)–(1f). In our case we have $y_t = 1$ for $t \in T$ and then the system becomes

$$\begin{aligned} 0 < x_t &\leq 1 & t &\in T, \\ a_1 &= 1, & & \\ a_{t+1} &\geq 1 - x_t & t &\in T \setminus \{p\}, \\ d_{1t} &\leq \textstyle\sum_{k=1}^t x_k & t &\in T, \\ a_t &\in \{0,1\} & t &\in T. \end{aligned}$$

By applying the transformation $x'_t = 1 - x_t$ for $t \in T$ to the above system we obtain:

$$0 \le x'_t < 1 \qquad t \in T, \tag{20a}$$
$$a_1 = 1, \tag{20b}$$
$$a_{t+1} \ge x'_t \qquad t \in T \setminus \{p\}, \tag{20c}$$
$$\textstyle\sum_{k=1}^{t} x'_k \le t - d_{1t} = d'_t \qquad t \in T, \tag{20d}$$
$$a_t \in \{0, 1\} \qquad t \in T \setminus \{1\}. \tag{20e}$$

If $a'_t = a_{t+1}$, the constraints (20b), (20c) and (20e) become

$$\begin{aligned} a'_0 &= 1, \\ a'_t &\ge x'_t \qquad t \in T \setminus \{p\}, \\ a'_t &\in \{0, 1\} \quad t \in T \setminus \{p\}. \end{aligned}$$

Define $a'_p = 1$. For $t \in T$, let $s_t \in \mathbb{R}_+$ be such that

$$\sum_{k=1}^{t} x'_k + s_t = d'_t. \tag{21}$$

Then, by subtracting $\sum_{k=1}^{t-1} x'_k + s_{t-1} = d'_{t-1}$ from $\sum_{k=1}^{t} x'_k + s_t = d'_t$ we obtain

$$x'_t + s_t = d'_t - d'_{t-1} + s_{t-1}$$

or, equivalently, by the equality in (20d),

$$x'_t + s_t = 1 - d_t + s_{t-1}. \tag{22}$$

If

$$\begin{aligned} x''_j &= x'_{p-j+1}, \quad s''_j = s_{p-j}, \quad s''_0 = s_p, \\ a''_j &= a'_{p-j+1}, \quad d''_j = 1 - d_{p-j+1} \end{aligned} \tag{23}$$

then (22) implies that

$$x''_j + s''_{j-1} = d''_j + s''_j$$

for $j \in \{1, \dots, p\}$. According to this, the variable s''_j represents the stock at the end of period j for this new model.

By using this last transformation and by denoting $d''_{j\ell} = \sum_{k=j}^{\ell} d''_k$ we rewrite the linear system (20a)–(20e) as follows:

$$x''_j + s''_{j-1} = d''_j + s''_j \qquad j \in T, \tag{24a}$$
$$x''_j \le a''_j \qquad j \in T, \tag{24b}$$
$$a''_j \in \{0, 1\} \qquad j \in T, \tag{24c}$$
$$s'' \in \mathbb{R}_+^{p+1}, \; x'' \in [0, 1)^p, \tag{24d}$$
$$s''_0 = 0, \; s''_p = s^*_p. \tag{24e}$$

Observe that the a''-variables behave as if they were set-up variables for periods 1 to p. Moreover, the inequalities (24a)–(24e) give the feasible solution set of a lot-sizing problem of the form (18a)–(18e). Then, in order to describe the polytope $P = \mathrm{conv}(\mathcal{S})$, we only need to add the corresponding (ℓ, S)-inequalities (19) of Proposition 13. More precisely, let $\ell \in T$, $L = \{1, \dots, \ell\}$ and $S \subseteq L$. The (ℓ, S)-inequality becomes

$$\sum_{j\in S} x''_j \le \sum_{j\in S} d''_{j\ell} a''_j + s''_\ell. \tag{25}$$

After Theorem 9 for $s''_0 = s''_p = 0$ we have the complete description of P.

By using definition of $d''_{j\ell}$ in (23) we have that

$$d''_{j\ell} = \sum_{k=j}^{l} (1 - d_{p-k+1}) = \ell - j + 1 - d_{(p-\ell+1)(p-j+1)}$$

and

$$d''_{1\ell} = \ell - d_{(p-\ell+1)p}.$$

By denoting $j_p = p - j + 1$ for each $j \in T$ and by using the definitions in (23), the inequality (25) becomes

$$\sum_{j_p\in S_p} x'_{j_p} \le \sum_{j_p\in S_p} (j_p - \ell_p + 1 - d_{\ell_p j_p}) a'_{j_p} + s_{\ell_p - 1} \tag{26}$$

where $S_p = \{k \in T : p - k + 1 \in S\} \subseteq L_p = \{p - \ell + 1, \dots, p\}$.

By using (21) and the definition of d' in (20d) we have

$$\sum_{k=1}^{\ell_p - 1} x'_k + s_{\ell_p - 1} = d'_{\ell_p - 1} = \ell_p - 1 - d_{1(\ell_p - 1)}$$

and

$$\sum_{j_p\in S_p} x'_{j_p} \le \sum_{j_p\in S_p} (j_p - \ell_p + 1 - d_{\ell_p j_p}) a'_{j_p} + \ell_p - 1 - d_{1(\ell_p-1)} - \sum_{j_p=1}^{\ell_p - 1} x'_{j_p}.$$

This implies that

$$\sum_{j_p=1}^{\ell_p - 1} x'_{j_p} + \sum_{j_p\in S_p} x'_{j_p} \le \sum_{j_p\in S_p} (j_p - \ell_p + 1 - d_{\ell_p j_p}) a'_{j_p} + \ell_p - 1 - d_{1(\ell_p-1)}.$$

Finally, returning to the original variables we have that the (ℓ, S)-inequality in (25) corresponds to

$$\sum_{j_p=1}^{\ell_p - 1} x_{j_p} + \sum_{j_p\in S_p} x_{j_p} + \sum_{j_p\in S_p} (j_p - \ell_p + 1 - d_{\ell_p j_p}) a_{j_p+1} \ge d_{1(\ell_p-1)} + |S_p|. \tag{27}$$

By denoting $k_0 = \ell_p$, $M = L_p \setminus S_p$, $A = S_p$ and $B = \{1, \dots, \ell_p - 1\} \cup S_p$ we obtain that the inequality (27) is a (k_0, M)-inequality (see Definition 1). After Theorem 9 the result is proved. □

To conclude this section, we study the separation problem associated with the (k_0, M)-inequalities. If $\mathcal{C}$ is a family of valid inequalities, the *separation problem associated with* $\mathcal{C}$ takes as input a point (x, y, a) within the linear relaxation P_{LR} of P, and consists of deciding whether (x, y, a) violates some inequality from $\mathcal{C}$ or not. The separation problem for the (k_0, M)-inequalities is relevant in this context, since by Theorem 10 these are the only nontrivial inequalities in the description of P. If the separation problem for the class of facet-inducing inequalities of a polytope Q is solvable in polynomial time, then the results in Grötschel et al. (1988) imply that the optimization problem over Q is also solvable in polynomial time. The following theorem shows that this is indeed the case for P when $d_1 < 1$ and $k_{\text{prod}} = p$.

Theorem 11 *The separation problem for the* (k_0, M)*-inequalities* (5) *can be solved in* $O((p-1)^2)$ *time.*

Proof Consider the linear relaxation P_{LR} of P. Given $\hat{w} = (\hat{x}, \hat{y}, \hat{a}) \in P_{LR}$ we want to decide if there exists a (k_0, M)-inequality which is violated by this point. The (k_0, M)-inequality (5) can be written as

$$\sum_{i=1}^{k_0-1} (x_i - d_i) + \sum_{i \in A} \left(x_i + (i - k_0 + 1 - d_{k_0 i})a_{i+1} - 1\right) \geq 0, \tag{28}$$

where $k_0 \geq 1$, $M \subseteq \{k_0, \dots, p-1\}$, $A = \{i \in \{k_0, \dots, p-1\}\colon i \notin M\}$ and $B = \{i \in \{1, \dots, p-1\}\colon i \notin M\}$.

Given a possibly fractional solution $\hat{w} = (\hat{x}, \hat{y}, \hat{a})$, consider the following algorithm. For each $k_0 \in \{1, \dots, p-1\}$, define J to be the set of periods $j \in \{k_0, \dots, p-1\}$ such that $\hat{v}_{j,k_0} = \hat{x}_j + (j - k_0 + 1 - d_{k_0 i})\, \hat{a}_{j+1} - 1 < 0$. Some inequality (28) associated with k_0 is violated if and only if $\sum_{j \in J} \hat{v}_{j,k_0} < \sum_{i=1}^{k_0-1} \left(d_i - \hat{x}_i\right)$. For this inequality we take $A = J$. This algorithm provides an $O((p-1)^2)$ separation procedure. □

Theorem 11 implies that the problem associated with this special case can be solved in polynomial time for any linear objective function.

4.1 Shortest path algorithm for the special case of high demands

Recall that $F(w) = \{t \in T\colon 0 < x_t < 1\} := \{t_1, \dots, t_k\}$ for every $w \in P$. In this special case of high demands null production in a period is not feasible, therefore, $T \setminus F(w) = \{t \in T\colon x_t = 1\}$ for every $w \in P$. In Sect. 3 we have characterized the extreme points of P. Under the assumptions of high demands we can provide a combinatorial algorithm for solving the lot-sizing problem.

Corollary 1 *Let* $w = (x, y, a) \in \mathcal{S}$ *with* $F(w) \neq \emptyset$. *Then* w *is an extreme point if and only if*

$$\begin{aligned} x_{1t_1} &= d_{1t_1}, && \\ x_{(t_i+1)t_{i+1}} &= d_{(t_i+1)t_{i+1}} && \textit{for} \quad i \in \{1, \dots, k-1\}, \\ x_t &= 1 && \textit{for} \quad t \in T \setminus F(w). \end{aligned}$$

Proof It is immediate to check that if w satisfies the conditions as above, then it is an extreme point by Theorem 6. Now, let w be an extreme point of P. From Theorem 6 we know that for every $i \in \{1, \dots, k\}$ there exists $r_i \in \{t_i, \dots, t_{i+1} - 1\}$ such that $x_{1r_i} = d_{1r_i}$.

Moreover, we now prove that $r_i = t_i$ for $i \in \{1, \dots, k\}$. To this end, assume that this is not the case for some $j \in \{1, \dots, k\}$. This implies that there exists $r_j \in \{t_j + 1, \dots, t_{j+1} - 1\}(\neq \emptyset)$ such that $x_{1r_j} = d_{1r_j}$. The assumption $t_j \neq r_j$ implies $x_{1t_j} > d_{1t_j}$. However, by definition of the set $F(w)$ and the hypothesis of high demands, we have $x_t = 1$ for $t = t_j + 1, \dots, t_{j+1} - 1$ which, together with $d_t < 1$ for any $t \in T$, implies $x_{1r_j} > d_{1r_j}$, a contradiction. Therefore, the result follows. □

Using this characterization of the extreme points of P for this case of high demands we can follow the same reasoning used for the uncapacitated lot-sizing problem. Define a directed graph $D = (N, A)$ with $N = \{0, 1, \ldots, p\}$ and $A = \{(i, j) : i, j \in V, i < j\}$. The arc (i, j) will correspond to an interval $[i, j - 1]$ such that there is fractional production in i and j, with demand satisfaction at equality in these two periods, and saturated production in every period between them (this corresponds to the concept of *regeneration intervals* for the uncapacitated lot-sizing problems). Clearly, any path from 1 to p provides a feasible solution to our lot-sizing problem. Furthermore, this solution satisfies the structure of the extreme points in Corollary 1.

We now define costs associated to the arcs of D. Let G be the objective function for a point $w \in P$, i.e., $G(x, y, a) = \sum_{i\in T} c_i x_i + \sum_{i\in T} f_i y_i + \sum_{i\in T} g_i a_i$. Since every feasible solution has $y_i = 1$ for every $i \in T$ then the problem is equivalent to optimizing $\sum_{i\in T} c_i x_i + \sum_{i\in T} g_i a_i$. Hence, we define arc (i, j) to have cost

$$\bar{c}(i, j-1) = c_j(d_{(i+1)j} - (j - i - 1)) + g_{i+1} + \sum_{\substack{k=i+2 \\ g_k<0}}^{j} g_k.$$

The length of a shortest path in this graph plus a constant value $\sum_{i\in T} f_i$, corresponding to the y-variables, gives the value of the optimum solution of the capacitated lot-sizing problem with start-up costs (and high demands). This provides a polynomial-time combinatorial algorithm for this case.

5 Conclusions and open problems

In this work we have addressed a lot-sizing problem including a particular definition of start-up costs, which we call *continuous start-up* and gives rise to an interesting structure. We have presented a first polyhedral study of a natural integer programming formulation for this problem. Besides general results on the polytope, we have introduced an exponentially-sized family of valid inequalities and we have provided bounds on its diameter. Finally, we presented a complete characterization of the polytope for a special sub-problem, and we showed by polyhedral methods that in this case the lot-sizing problem can be solved in polynomial time.

Many open questions remain. It would be interesting to find further families of valid inequalities and to study their facetness properties. Many families of valid inequalities composed by $O(p)$ inequalities are known, but the (k_0, M)-inequalities presented in this work are the only known exponentially-sized family of valid inequalities particular to this polytope. Further families would provide more information on the structure of this polytope. It would also be interesting to know whether the known $O(p)$-families can be generalized into a common super-family.

The characterization in Sect. 4 of the polytope for a special case and the polynomial-time separation of the inequalities in the complete description are an example of a polynomiality result provided by a polyhedral approach. This result holds for any objective function expressed as a linear function of the model variables, so it may include production costs and inventory costs (associated with the x-variables), fixed production costs (associated with the y-variables), and continuous start-up costs (associated with the a-variables). The low-complexity combinatorial algorithm in Sect. 4.1 is the algorithmic counterpart of these polyhedral results. It may be worthwhile to explore whether the approach in that section can be generalized for other special cases of this lot-sizing problem.

Acknowledgments We would like to thank the anonymous reviewers of a first version of this manuscript, whose comments helped us to improve the final version and presentation of the results. In particular, we are greatly indebted to one of them for pointing us to the elegant proof of Theorem 10. We are also greatly indebted to the anonymous reviewer of this submission for his/her thorough comments and suggestions, and very specially for the suggestions that led to the results in Sect. 4.1. This work was partially supported by Grants PIP-ANPCyT 0361 and PID-UNR 415 (Argentina).

Appendix: An example

In this last section we summarize the results presented in Sect. 4 in a particular example. We consider the problem with a planning horizon consisting of periods $T = \{1, \dots, 5\}$. For periods in T, we have demands $d = (0.8, 0.7, 0.8, 0.9, 0.9)$. It verifies the conditions $d_i < 1$ for $1 \in T$ and $k_{\text{prod}} = 5$. Using the PORTA package (see http://www.zib.de/Optimization/Software/Porta/) we obtain the following description of P by equations and linear inequalities.

```
( 1a)              +y5-a1       == 0
( 2a)           +y4-y5          == 0
( 3a)        +y3-y4             == 0
( 4a)     +y2-y3                == 0
( 5a)  +y1-y2                   == 0
( 6a)                 +a1       == 1
( 1b)                  -10x1-10x2-10x3-10x4-10x5  <= -41
( 2b)     -3a3-5a4-6a5 -10x1-10x2-10x3-10x4       <= -38
( 3b)         -2a4-3a5 -10x1-10x2-10x3-10x4       <= -35
( 4b)             - a5 -10x1-10x2-10x3-10x4       <= -33
( 5b)     -3a3    -6a5 -10x1-10x2     -10x4       <= -28
( 6b)         -5a4-6a5 -10x1     -10x3-10x4       <= -28
( 7b)     -3a3-5a4     -10x1-10x2-10x3            <= -28
( 8b)         -2a4     -10x1-10x2-10x3            <= -25
( 9b)             -3a5 -10x1-10x2     -10x4       <= -25
(10b)     -3a3         -10x1-10x2                 <= -18
(11b)         -5a4     -10x1     -10x3            <= -18
(12b)         -7a4               -10x3            <= -10
(13b)             -3a5 - 5x1          - 5x4       <=  -9
(14b)  -a2             - 5x1                      <=  -5
(15b)             -4a5                - 5x4       <=  -5
(16b)     - a3              - 2x2                 <=  -2
(17b)                                    +  x5    <=   1
(18b)                               +  x4         <=   1
(19b)                          +  x3              <=   1
(20b)                     +  x2                   <=   1
(21b)                +  x1                        <=   1
(22b)            + a5                             <=   1
(23b)        + a4                                 <=   1
(24b)    + a3                                     <=   1
(25b)  +a2                                        <=   1
```

The following remarks categorize these 25 facets within the model constraints and the known families of valid inequalities.

- The equalities (1a)–(6a) correspond to the minimal equation system for P.
- The constraint (1b) corresponds to the total demand satisfaction (see Theorem 3(i)).
- The inequalities (17b)–(21b) induce facets as we prove in Theorem 3(ii).
- The constraints (22b)–(25b) correspond to the inequalities inducing facets in Theorem 3(v).
- The constraints (12b), (14b), (15b) and (16b) belong to the family of (k_0, M) -inequalities (see Remark 1) which define facets for P (see Theorem 4).
- The constraints (1b)–(11b) and (13b) are (k_0, M)-inequalities (see Definition 1 and Theorem 5).

References

Christof, T., & Löbel, A. PORTA—Polyhedron representation transformation algorithm. http://www.zib.de/Optimization/Software/Porta/.

Constantino, M. (1996). A cutting plane approach to capacitated lot-sizing with start-up costs. *Mathematical Programming*, *75*, 353–376.

Escalante, M., Marenco, J., & Varaldo, M. C. (2011). A polyhedral study of the single-item lot-sizing problem with continuous start-up costs. *Electronic Notes in Discrete Mathematics*, *37*, 261–266.

Grötschel, M., Lovász, L., & Schrijver, A. (1988). *Geometric algorithms and combinatorial optimization*. New York: Springer.

Nemhauser, G., & Wolsey, L. (1988). *Integer and combinatorial optimization*. New York: Wiley.

Pochet, Y., & Wolsey, L. (2006). Production planning by mixed integer programming. In T. V. Mikosch, S. M. Robinson & S. I. Resnick (Eds.), *Springer series in operations research and financial engineering*. New York: Springer.

Pulleyblank, W. R. (1989). Polyhedral combinatorics. In G. L. Nemhauser, A. H. G. Rinooy Kan & M. J. Todd (Eds.), *Handbooks in operations research and management science*, Chapter V, *Vol. 1* (pp. 371–446). Elsevier.

Toledo, F., dos Santos, M., Arenales, M., & Seleghim, P. (2008). Logistica de distribuicão de água em redes urbanas—Racionalizacão energética. *Pesquisa Operacional*, *88–1*, 75–91.

van Hoesel, G., Wagelmans, A., & Wolsey, L. (1994). Polyhedral characterization of the economic lot-sizing problem with start-up costs. *SIAM Journal of Discrete Mathematics*, *7–1*, 141–151.