On a class of mixed-integer sets with a single integer variable

Marco Di Summa*

Abstract

We consider mixed-integer sets defined by a linear system $Ax \geq b$ plus an integrality requirement on one variable, where A is a totally unimodular matrix with at most two nonzero entries per row. We give a complete linear-inequality description for the convex hull of any set of this type.

Keywords: mixed-integer programming, convex hull descriptions, network-flow matrices.

1 Introduction

In a recent paper, Conforti et al. [1] investigated the class $\mathcal{X}$ of mixed-integer sets of the type $X := \{x \in \mathbb{R}^n : Ax \geq b,\ x_i \text{ integer for } i \in I\}$, where A is a totally unimodular matrix with at most two nonzero entries per row, b is an arbitrary vector, and I is a subset of $\{1, \ldots, n\}$. They gave a linear system of inequalities that describes $\text{conv}(X)$ (the convex hull of X) in a higher-dimensional space by means of additional variables. As remarked in [1], the family $\mathcal{X}$ includes and generalizes several mixed-integer sets that had been studied previously [2, 3, 8, 11, 13, 14], most of them arising as relaxations of lot-sizing problems.

Despite the result of [1], a linear-inequality description for $\text{conv}(X)$ in the original variables x is unknown in the general case. However, the convex hull in the original space was found for the sets in some subfamilies of $\mathcal{X}$ [3, 4, 6, 8, 13]. In particular, a linear-inequality description in the original space is known for any set in $\mathcal{X}$ having only one continuous variable [4, 6]. The object of this study is, in a sense, the symmetric case: we deal with mixed-integer sets in $\mathcal{X}$ with a single integer variable.

The rest of this note is organized as follows. In Section 2 we state our main result, i.e., for any set $X \in \mathcal{X}$ with a single integer variable, we give a system of linear inequalities that describes $\text{conv}(X)$ in the original space. The result is then proven in Section 3.

Standard terminology and basic results of polyhedral theory will be used throughout the paper. We refer the reader to [10] or [12].

*Dipartimento di Informatica, Università degli Studi di Torino. Corso Svizzera 185, I-10149 Torino, Italy (`disumma@di.unito.it`).

2 The convex hull

Let $X := \{x \in \mathbb{R}^n : Ax \geq b,\ x_i \text{ integer for } i \in I\}$ be a mixed-integer set, where A is an $m \times n$ totally unimodular matrix with at most two nonzero entries per row and $I \subseteq \{1, \ldots, n\}$ with $|I| = 1$. We assume w.l.o.g. that x_n is the integer variable, i.e., $I = \{n\}$.

As shown in [1], a result in [9] implies that it is possible to multiply by -1 a subset of columns of A so that the resulting matrix has the following property: if a row contains two nonzero entries, then one of them is $+1$ and the other is -1. Therefore we can assume w.l.o.g. that A satisfies this property, and thus the linear system $Ax \geq b$ has the form

$$x_i - x_j \geq d_{ij}, \quad (i, j) \in D, \tag{1}$$

$$x_i \geq l_i, \quad i \in L, \tag{2}$$

$$x_i \leq u_i, \quad i \in U, \tag{3}$$

for some subsets $D \subseteq \{1, \ldots, n\} \times \{1, \ldots, n\}$ and $L, U \subseteq \{1, \ldots, n\}$, and rational numbers d_{ij}, l_i, u_i. If the above system does not include an explicit lower (resp., upper) bound on the integer variable x_n, we set $l_n := -\infty$ (resp., $u_n := +\infty$). Otherwise, if a lower (resp., upper) bound on x_n is given, we assume w.l.o.g. that l_n (resp., u_n) is an integer number.

Let $\mathcal{G} = (V, E)$ be the directed graph whose arc-node incidence matrix is constituted by the first $n-1$ columns of A, with a dummy node 0. That is, the vertex set of $\mathcal{G}$ is $V := \{0, \ldots, n-1\}$ and the arc set E is defined as follows:

(a) for each pair $(i, j) \in D$, where $i, j \neq n$, E contains arc (i, j);

(b) for each pair $(i, n) \in D$, E contains arc $(i, 0)$; symmetrically, for each pair $(n, j) \in D$, E contains arc $(0, j)$;

(c) for each index $i \in L$ with $i \neq n$, E contains arc $(i, 0)$;

(d) for each index $i \in U$ with $i \neq n$, E contains arc $(0, i)$.

Note that $\mathcal{G}$ may contain several pairs of parallel or opposite arcs.

Thus every inequality of the system $Ax \geq b$ (i.e., system (1)–(3)), except for the inequalities $l_n \leq x_n \leq u_n$, gives rise to an arc of $\mathcal{G}$. We assign weights to the arcs of $\mathcal{G}$ as follows: every arc of type (a) or (b), arising from a pair $(i, j) \in D$, receives weight d_{ij}; arcs of type (c) receive weight l_i; arcs of type (d) receive weight $-u_i$. The weight of an arc $e \in E$ is denoted by b_e. In other words, b_e is the right-hand side of the inequality of (1)–(3) (written in the "$\geq$" form) that corresponds to arc e.

Let $\mathcal{C}$ denote a sequence of arcs $e_0, \ldots, e_k$ forming an undirected cycle in $\mathcal{G}$ ($k \geq 2$). Let $i_0 = i_{k+1}$ be the node shared by arcs e_0 and e_k, and for $1 \leq t \leq k$, let i_t be the node shared by arcs e_{t-1} and e_t. Thus for $0 \leq t \leq k$, either $e_t = (i_t, i_{t+1})$ or $e_t = (i_{t+1}, i_t)$. Let $E^+_{\mathcal{C}} := \{e_t : e_t = (i_t, i_{t+1})\}$ and $E^-_{\mathcal{C}} := \{e_t : e_t = (i_{t+1}, i_t)\}$. Also, let $T^+_{\mathcal{C}}$ (resp., $T^-_{\mathcal{C}}$) be the set of indices t such that e_t is in $E^+_{\mathcal{C}}$ (resp., $E^-_{\mathcal{C}}$). We define

$$b^+_{\mathcal{C}} := \sum_{e \in E^+_{\mathcal{C}}} b_e, \quad b^-_{\mathcal{C}} := \sum_{e \in E^-_{\mathcal{C}}} b_e, \quad \beta_{\mathcal{C}} := b^+_{\mathcal{C}} - b^-_{\mathcal{C}}. \tag{4}$$

We now present the main result of this note, that is, a linear-inequality description for the convex hull of X, denoted $\mathrm{conv}(X)$. In the following, $f(\alpha)$ stands for the fractional part of a number α, i.e., $f(\alpha) := \alpha - \lfloor \alpha \rfloor$.

Theorem 1 *The polyhedron* $\mathrm{conv}(X)$ *is described by the original system* (1)–(3) *plus all inequalities of the form*

$$\sum_{t \in T^+_{\mathcal{C}}} (x_{i_t} - x_{i_{t+1}}) + \big(\varepsilon_{\mathcal{C}} + f(\beta_{\mathcal{C}})\big) x_n \geq b^-_{\mathcal{C}} + f(\beta_{\mathcal{C}}) \lfloor \beta_{\mathcal{C}} + 1 \rfloor, \tag{5}$$

where

- $\mathcal{C}$ *is a sequence of arcs* $e_0, \dots, e_k$ *forming an undirected cycle in* $\mathcal{G}$ $(k \geq 2)$*;* $i_0 = i_{k+1} = 0$ *is the node shared by arcs* e_0 *and* e_k*, and for* $1 \leq t \leq k$*,* i_t *is the node shared by arcs* e_{t-1} *and* e_t*;* e_0 *is an arc of type* (b)*, while* e_k *is either of type* (c) *or of type* (d)*;*
- $x_0 := 0$*,* $\beta_{\mathcal{C}}$ *is defined by* (4)*, and* $\varepsilon_{\mathcal{C}} := \begin{cases} 0 & \text{if } 0 \in T^+_{\mathcal{C}}, \\ -1 & \text{otherwise.} \end{cases}$

The next section is devoted to proving the above theorem.

3 Proof of the result

The following notation will be used. We denote by P the set $\{x \in \mathbb{R}^n : Ax \geq b\}$, i.e., the polyhedron defined by (1)–(3). Furthermore, we write $A = [M \mid a_n]$, where M is the column submatrix of A constituted by the first $n-1$ columns of A and a_n is the n-th column of A. Similarly, we decompose a vector $x \in \mathbb{R}^n$ as $x = (x_M, x_n)$.

Since the set X has a single integer variable, its convex hull is completely described by split cuts [5]. We recall that an inequality $cx \geq \delta$ is a split cut for P with respect to variable x_n if there exists an integer α such that the inequality $cx \geq \delta$ is valid for the two polyhedra $\{x \in P : x_n \leq \alpha\}$ and $\{x \in P : x_n \geq \alpha + 1\}$.

Let $cx \geq \delta$ be a non-redundant split cut associated with the disjunction $(x_n \leq \alpha) \vee (x_n \geq \alpha + 1)$ for some integer α, where we can assume w.l.o.g. that $l_n \leq \alpha < u_n$. Since the inequality $cx \geq \delta$ is valid for both polyhedra $\{x \in P : x_n \leq \alpha\}$ and $\{x \in P : x_n \geq \alpha + 1\}$, by Farkas' Lemma (see, e.g., [12]) there exist multipliers $(v, \lambda), (w, \mu) \in \mathbb{R}^m \times \mathbb{R}$ satisfying

$$vM = c_M = wM, \tag{6}$$

$$va_n - \lambda = c_n = wa_n + \mu, \tag{7}$$

$$vb - \lambda\alpha = \delta = wb + \mu(\alpha + 1), \tag{8}$$

$$v, w \geq \mathbf{0},\ \lambda, \mu \geq 0. \tag{9}$$

Thus the vector (v, λ, w, μ) belongs to the pointed cone defined by

$$vM = wM, \tag{10}$$

$$va_n - \lambda = wa_n + \mu, \tag{11}$$

$$vb - \lambda\alpha = wb + \mu(\alpha + 1), \tag{12}$$

$$v, w \geq \mathbf{0},\ \lambda, \mu \geq 0. \tag{13}$$

As $cx \geq \delta$ is non-redundant, (v, λ, w, μ) is an extreme ray of (10)–(13). Furthermore, as we are only interested in extreme rays of (10)–(13) with $\lambda, \mu > 0$ (otherwise inequality $cx \geq \delta$ would be implied by the original system $Ax \geq b$), we observe that (v, w) is an extreme ray of the cone defined by

$$vM = wM, \tag{14}$$
$$v, w \geq \mathbf{0}. \tag{15}$$

Now, if for an integer k we define

$$b^k := b - ka_n,$$

equations (11)–(12) give

$$\lambda = (w - v)b^{\alpha+1}, \quad \mu = (v - w)b^{\alpha}. \tag{16}$$

Using (6)–(8) and (16), inequality $cx \geq \delta$ can be written as

$$vMx_M + \left(vb^{\alpha} - wb^{\alpha+1}\right) x_n \geq vb^{\alpha} + \left(vb^{\alpha} - wb^{\alpha+1}\right) \alpha. \tag{17}$$

Therefore the polyhedron $\text{conv}(X)$ is described by the original system (1)–(3) plus the inequalities (17) for all extreme rays (v, w) of (14)–(15) and all α such that the corresponding values of λ and μ defined by (16) are positive. We now show that all these inequalities are of the form (5).

Lemma 2 *Let α be an integer such that $l_n \leq \alpha < u_n$ and let (v, w) be an extreme ray of the cone defined by* (14)–(15) *such that the values λ and μ defined by* (16) *are positive. Then the corresponding inequality* (17) *is of the form* (5).

Proof. We rewrite system (14)–(15) as

$$(v, w)\mathcal{M} = \mathbf{0}, \tag{18}$$
$$v, w \geq \mathbf{0}, \tag{19}$$

where $\mathcal{M} = \begin{bmatrix} M \\ -M \end{bmatrix}$.

Note that M may have some all-zero rows, namely the rows corresponding to inequalities $x_n \geq l_n$ and $x_n \leq u_n$ (if they appear in system $Ax \geq b$). Suppose, for instance, that inequality $x_n \geq l_n$ is the t-th row of system $Ax \geq b$. Then the vector (v, w) defined by setting $v_t = 1$ and all other entries equal to zero is the only extreme ray of (18)–(19) with $v_t \neq 0$. Similarly, the vector (v, w) defined by setting $w_t = 1$ and all other entries equal to zero is the only extreme ray of (18)–(19) with $w_t \neq 0$. However, since $\mu = l_n - \alpha \leq 0$ in the former case and $\lambda = l_n - \alpha - 1 \leq 0$ in the latter case, these rays need not be considered. A similar argument applies to inequality $x_n \leq u_n$. This shows that we can ignore the all-zero rows of system (18).

Now, system (18) describes flow-conservation constraints on a directed graph $\mathcal{H}$ with vertex set $\{0, \ldots, n-1\}$, where 0 is a dummy node, and arc set defined as follows: for every row of $\mathcal{M}$ containing a $+1$ in column i and a -1 in column j, there is an arc (i, j); for every row with a $+1$ (resp., -1) in column i and all other entries equal to 0, there is an arc $(i, 0)$ (resp., $(0, i)$). Note that $\mathcal{H}$ and the graph $\mathcal{G}$ introduced in Section 2 are defined on the same vertex set. Furthermore, every row of M generates an arc of $\mathcal{G}$ and a pair of opposite arcs of $\mathcal{H}$. If an arc e of $\mathcal{G}$ corresponds to the pair of opposite arcs e', e'' of $\mathcal{H}$, we say that e is the arc *underlying* e' and e''. Given any subset of arcs of $\mathcal{H}$, the underlying subset of arcs of $\mathcal{G}$ is defined similarly.

It is well-known [7] that the extreme rays of (18)–(19) are the 0-1 vectors (up to multiplication by a positive scalar) whose supports define directed cycles in $\mathcal{H}$.

Let $\mathcal{D}$ be a directed cycle in $\mathcal{H}$ defined by an extreme ray (v, w) of (18)–(19). If $\mathcal{D}$ consists of a pair of opposite arcs that correspond to the same arc of $\mathcal{G}$, then $v = w$. This implies that $\lambda = 0$ and thus we can ignore this ray.

Therefore from now on we assume that $\mathcal{D}$ is a directed cycle of $\mathcal{H}$ with at least three arcs. Let $\mathcal{C}$ be the underlying undirected cycle in $\mathcal{G}$. We denote the sequence of arcs of $\mathcal{C}$ by $e_0, \ldots, e_k$, where $k \geq 2$; furthermore, $i_0 = i_{k+1}$ is the node shared by arcs e_0 and e_k, and for $1 \leq t \leq k$, i_t is the node shared by arcs e_{t-1} and e_t.

Since the support of v (resp., w) corresponds to the arcs of $\mathcal{D}$ for which the underlying arcs of $\mathcal{C}$ are in $E^+_{\mathcal{C}}$ (resp., $E^-_{\mathcal{C}}$), we have

$$vb = b^+_{\mathcal{C}}, \qquad wb = b^-_{\mathcal{C}}. \tag{20}$$

Define $\delta := va_n$ and $\varepsilon := wa_n$. Since the support of column a_n corresponds to arcs of $\mathcal{G}$ of type (b) (see Section 2 for the definition of types (a)–(d)), we have the following:

> The value δ is the difference between the number of arcs of type (b) in $E^+_{\mathcal{C}}$ leaving node 0 and the number of arcs of type (b) in $E^+_{\mathcal{C}}$ entering node 0. Similarly, the value ε is the difference between the number of arcs of type (b) in $E^-_{\mathcal{C}}$ leaving node 0 and the number of arcs of type (b) in $E^-_{\mathcal{C}}$ entering node 0.

In particular, it follows that $\delta, \varepsilon \in \{0, \pm 1\}$, and if they are both nonzero, they are either both 1 or both -1. Thus $|\delta - \varepsilon| \leq 1$. Also note that

$$vb^\alpha = vb - \delta\alpha, \quad wb^{\alpha+1} = wb - \varepsilon(\alpha + 1). \tag{21}$$

Define $\rho := vb^\alpha - wb^{\alpha+1}$. Since $\rho = -\lambda + \delta = \mu + \varepsilon$, the condition $\lambda, \mu > 0$ is equivalent to $\varepsilon < \rho < \delta$, which is possible only if $\delta \geq \varepsilon + 1$. Now, because $|\delta - \varepsilon| \leq 1$, we necessarily have $\delta = \varepsilon + 1$. Then, using (21) and (20),

$$\rho = vb^\alpha - wb^{\alpha+1} = vb - wb - \alpha + \varepsilon = b^+_{\mathcal{C}} - b^-_{\mathcal{C}} - \alpha + \varepsilon = \beta_{\mathcal{C}} - \alpha + \varepsilon,$$

thus $\rho - \varepsilon = \beta_{\mathcal{C}} - \alpha$. Now, since $0 < \rho - \varepsilon < \delta - \varepsilon = 1$ and since ε and α are integer numbers, we have $\alpha = \lfloor \beta_{\mathcal{C}} \rfloor$ and $\rho = \varepsilon + f(\beta_{\mathcal{C}})$.

We now show that $\mathcal{C}$ satisfies the conditions of Theorem 1 and inequality (17) coincides with inequality (5).

As $\delta, \varepsilon \in \{0, \pm 1\}$ and $\delta = \varepsilon + 1$, either $\delta = 1$ and $\varepsilon = 0$, or $\delta = 0$ and $\varepsilon = -1$. Recalling the meaning of the values δ and ε pointed out above, we see that in both cases node 0 is part of the cycle and thus we can assume that $i_0 = 0$. Also, in both cases arc e_0 is of type (b), while e_k is either of type (c) or of type (d). Furthermore if $\delta = 1$ and $\varepsilon = 0$ then $e_0 \in E^+_{\mathcal{C}}$, while if $\delta = 0$ and $\varepsilon = -1$ then $e_0 \in E^-_{\mathcal{C}}$. Thus $\mathcal{C}$ satisfies the conditions of Theorem 1 and $\varepsilon = \varepsilon_{\mathcal{C}}$.

Since $\alpha = \lfloor \beta_{\mathcal{C}} \rfloor$ and $\delta = \varepsilon + 1$, the first equation in (21) gives

$$vb^\alpha = vb - (\varepsilon + 1)\alpha = b^+_{\mathcal{C}} - (\varepsilon + 1)\lfloor \beta_{\mathcal{C}} \rfloor = b^-_{\mathcal{C}} + \beta_{\mathcal{C}} - (\varepsilon + 1)\lfloor \beta_{\mathcal{C}} \rfloor.$$

Then, recalling that $\rho = \varepsilon + f(\beta_{\mathcal{C}})$, the right-hand side of (17) is

$$vb^\alpha + \rho\alpha = b^-_{\mathcal{C}} + \beta_{\mathcal{C}} - (\varepsilon + 1)\lfloor \beta_{\mathcal{C}} \rfloor + (\varepsilon + f(\beta_{\mathcal{C}}))\lfloor \beta_{\mathcal{C}} \rfloor = b^-_{\mathcal{C}} + f(\beta_{\mathcal{C}})\lfloor \beta_{\mathcal{C}} + 1 \rfloor,$$

which is exactly the right-hand side of inequality (5).

One can also verify that $vMx_M = \sum_{t \in T^+_{\mathcal{C}}} (x_{i_t} - x_{i_{t+1}})$, with the convention that $x_0 = 0$. Finally, the coefficient of x_n in inequality (17) is $vb^\alpha - wb^{\alpha+1} = \rho = \varepsilon + f(\beta_{\mathcal{C}}) = \varepsilon_{\mathcal{C}} + f(\beta_{\mathcal{C}})$. Thus (17) and (5) coincide. □

We can now prove the Theorem 1.

Proof of Theorem 1. By Lemma 2, all split cuts for P are inequalities of the form (17). To conclude, one needs just to observe that every inequality of the form (5) is valid for $\mathrm{conv}(X)$. This can be done as follows. Starting from the sequence of arcs $\mathcal{C}$, construct the unique directed cycle $\mathcal{D}$ in $\mathcal{H}$ whose underlying undirected cycle in $\mathcal{G}$ is $\mathcal{C}$. Let (v, w) be the characteristic vector of $\mathcal{D}$ and set $\alpha = \lfloor \beta_{\mathcal{C}} \rfloor$. By using the same arguments as in the proof of Lemma 2, one shows that $\lambda, \mu > 0$ and inequalities (17) and (5) coincide. Therefore inequality (5) is a split cut for P. □

Acknowledgement

The author is grateful to an anonymous referee, whose suggestions helped improve the presentation of the result.

References

[1] M. Conforti, M. Di Summa, F. Eisenbrand, and L. A. Wolsey. Network formulations of mixed-integer programs. *Mathematics of Operations Research*, 34(1):194–209, 2009.

[2] M. Conforti, M. Di Summa, and L. A. Wolsey. The intersection of continuous mixing polyhedra and the continuous mixing polyhedron with flows. In M. Fischetti and D. P. Williamson, editors, *Integer Programming and Combinatorial Optimization*, volume 4513 of *Lecture Notes in Computer Science*, pages 352–366. Springer, 2007.

[3] M. Conforti, M. Di Summa, and L. A. Wolsey. The mixing set with flows. *SIAM Journal on Discrete Mathematics*, 21(2):396–407, 2007.

[4] M. Conforti, L. A. Wolsey, and G. Zambelli. Projecting an extended formulation for mixed-integer covers on bipartite graphs. *Mathematics of Operations Research*, 2010, doi: 10.1287/moor.1100.0454

[5] W. Cook, R. Kannan, and A. Schrijver. Chvátal closures for mixed integer programming problems. *Mathematical Programming*, 47(1–3):155–174, 1990.

[6] M. Di Summa. *Formulations of Mixed-Integer Sets with Totally Unimodular Constraint Matrices.* PhD thesis, Università degli Studi di Padova, Italy, 2008.

[7] L. R. Ford and D. R. Fulkerson. *Flows in Networks.* Princeton University Press, 1962.

[8] O. Günlük and Y. Pochet. Mixing mixed-integer inequalities. *Mathematical Programming*, 90(3):429–457, 2001.

[9] I. Heller and C. B. Tompkins. An extension of a theorem of Dantzig. *Linear Inequalities and Related Systems*, pages 247–252, 1956.

[10] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization.* Wiley-Interscience New York, NY, USA, 1988.

[11] Y. Pochet and L. A. Wolsey. Polyhedra for lot-sizing with Wagner-Whitin costs. *Mathematical Programming*, 67(1):297–323, 1994.

[12] A. Schrijver. *Theory of Linear and Integer Programming.* John Wiley & Sons, Inc. New York, NY, USA, 1986.

[13] M. Van Vyve. The continuous mixing polyhedron. *Mathematics of Operations Research*, 30(2):441–452, 2005.

[14] M. Van Vyve. Linear programming extended formulations for the single-item lot-sizing problem with backlogging and constant capacity. *Mathematical Programming*, 108(1):53–78, 2006.