The Intersection of Continuous Mixing Polyhedra and the Continuous Mixing Polyhedron with Flows*

Michele Conforti[1], Marco Di Summa[1], and Laurence A. Wolsey[2]

[1] Dipartimento di Matematica Pura ed Applicata, Università degli Studi di Padova. Via Trieste 63, 35121 Padova, Italy. {conforti,mdsumma}@math.unipd.it
[2] Center for Operations Research and Econometrics (CORE), Université catholique de Louvain. 34, Voie du Roman Pays, 1348 Louvain-la-Neuve, Belgium. wolsey@core.ucl.ac.be

Abstract. In this paper we investigate two generalizations of the continuous mixing set studied by Miller and Wolsey [5] and Van Vyve [7]: the intersection set

$$X^{\mathrm{I}} = \{(\sigma, r, y) \in \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \mathbb{Z}^n_+ : \sigma_k + r_t + y_t \geq b_{kt},\ 1 \leq k, t \leq n\}$$

and the continuous mixing set with flows

$$X^{\mathrm{CMF}} = \{(s, r, x, y) \in \mathbb{R}_+ \times \mathbb{R}^n_+ \times \mathbb{R}^n_+ \times \mathbb{Z}^n_+ : \\ s + r_t + x_t \geq b_t,\ x_t \leq y_t,\ 1 \leq t \leq n\}\ ,$$

which appears as a strong relaxation of some single-item lot-sizing problems. We give two extended formulations for the convex hull of each of these sets. In particular, for X^{CMF} the sizes of the extended formulations are polynomial in the size of the original description of the set, thus proving that the corresponding linear optimization problem can be solved in polynomial time.

Keywords: integer programming.

1 Introduction

In the last 5-10 years several mixed-integer sets have been studied that are interesting in their own right as well as providing strong relaxations of single-item lot-sizing sets. One in particular is the *continuous mixing set* X^{CM}:

$$s + r_t + y_t \geq b_t,\ 1 \leq t \leq n$$
$$s \in \mathbb{R}_+,\ r \in \mathbb{R}^n_+,\ y \in \mathbb{Z}^n_+\ .$$

The *continuous mixing polyhedron* $\mathrm{conv}(X^{\mathrm{CM}})$, which is the convex hull of the above set, was introduced and studied by Miller and Wolsey in [5], where

* This work was partly carried out within the framework of ADONET, a European network in Algorithmic Discrete Optimization, contract no. MRTN-CT-2003-504438.

an extended formulation of conv(X^{CM}) with $\mathcal{O}(n^2)$ variables and $\mathcal{O}(n^2)$ constraints was given. Van Vyve [7] gave a more compact extended formulation of conv(X^{CM}) with $\mathcal{O}(n)$ variables and $\mathcal{O}(n^2)$ constraints and a formulation of conv(X^{CM}) in its original space.

We study here two generalizations of the continuous mixing set. First we consider the *intersection set* X^{I}, the intersection of several continuous mixing sets with distinct σ_k variables and common r and y variables:

$$\sigma_k + r_t + y_t \geq b_{kt},\ \ 1 \leq k, t \leq n \tag{1}$$

$$\sigma \in \mathbb{R}^n_+,\ r \in \mathbb{R}^n_+,\ y \in \mathbb{Z}^n_+\ . \tag{2}$$

Then we consider X^{CMF}, the "flow version" of the continuous mixing set:

$$s + r_t + x_t \geq b_t,\ \ 1 \leq t \leq n \tag{3}$$

$$x_t \leq y_t,\ \ 1 \leq t \leq n \tag{4}$$

$$s \in \mathbb{R}_+,\ r \in \mathbb{R}^n_+,\ x \in \mathbb{R}^n_+,\ y \in \mathbb{Z}^n_+\ . \tag{5}$$

We now show two links between the continuous mixing set with flows X^{CMF} and lot-sizing. The first is to the single-item constant capacity lot-sizing problems with backlogging over n periods, which can be formulated (including redundant equations) as:

$$s_{k-1} + \textstyle\sum_{u=k}^{t} w_u + r_t = \sum_{u=k}^{t} d_u + s_t + r_{k-1},\ \ 1 \leq k \leq t \leq n$$
$$w_u \leq C z_u,\ 1 \leq u \leq n;\ \ s \in \mathbb{R}^{n+1}_+,\ r \in \mathbb{R}^{n+1}_+,\ w \in \mathbb{R}^n_+,\ z \in \{0,1\}^n\ .$$

Here d_u is the demand in period u, s_u and r_u are the stock and backlog at the end of period u, z_u takes value 1 if there is a set-up in period u allowing production to take place, w_u is the production in period u and C is the capacity (i.e. the maximum production). To see that this set has a relaxation as the intersection of n continuous mixing sets with flows, take $C = 1$ wlog, fix k, set $s = s_{k-1}$, $x_t = \sum_{u=k}^{t} w_u$, $y_t = \sum_{u=k}^{t} z_u$ and $b_t = \sum_{u=k}^{t} d_u$, giving a first relaxation:

$$s + x_t + r_t \geq b_t,\ \ k \leq t \leq n \tag{6}$$

$$0 \leq x_u - x_{u-1} \leq y_u - y_{u-1} \leq 1,\ \ k \leq u \leq n \tag{7}$$

$$s \in \mathbb{R}_+,\ r \in \mathbb{R}^{n-k+1}_+,\ x \in \mathbb{R}^{n-k+1}_+,\ y \in \mathbb{Z}^{n-k+1}\ . \tag{8}$$

Now summing (7) over $k \leq u \leq t$ (for each fixed $t = k, \ldots, n$) and dropping the upper bound on y_t, one obtains precisely the continuous mixing set with flows X^{CMF}.

The set X^{CMF} also provides an exact model for the two stage stochastic lot-sizing problem with constant capacities and backlogging. Specifically, at time 0 one must choose to produce a quantity s at a per unit cost of h. Then in period 1, n different outcomes are possible. For $1 \leq t \leq n$, the probability of event t is ϕ_t, the demand is b_t and the unit production cost is p_t, with production in batches of size up to C; there are also a fixed cost of q_t per batch and a possible bound k_t on the number of batches. As an alternative to production there is a linear backlog (recovery) cost e_t. Finally the goal is to satisfy all demands and minimize the total expected cost. The resulting problem is

$$\begin{aligned} \min \quad & hs + \textstyle\sum_{t=1}^{n} \phi_t(p_t x_t + q_t y_t + e_t r_t) \\ \text{s.t.} \quad & s + r_t + x_t \geq b_t,\ 1 \leq t \leq n && (9) \\ & x_t \leq C y_t,\ y_t \leq k_t,\ 1 \leq t \leq n && (10) \\ & s \in \mathbb{R}_+,\ r \in \mathbb{R}_+^n,\ x \in \mathbb{R}_+^n,\ y \in \mathbb{Z}_+^n\ . && (11) \end{aligned}$$

When $k_t = 1$ for all t, this is a standard lot-sizing problem, and in general (assuming $C = 1$ wlog) this is the set $X^{\text{CMF}} \cap \{(s, r, x, y) : y_t \leq k_t,\ 1 \leq t \leq n\}$.

Now we describe the contents of this paper. Note that throughout, a *formulation* of a polyhedron $P \subseteq \mathbb{R}^n$ is an external description of P in its original space. It consists of a finite set of inequalities $Ax \leq d$ such that $P = \{x \in \mathbb{R}^n : Ax \leq d\}$. A formulation of P is *extended* whenever it gives an external description of P in a space $\mathbb{R}^{n+m}$ that includes the original space, so that, given $Q = \{(x, w) \in \mathbb{R}^{n+m} : A'x + B'w \leq d'\}$, P is the projection of Q onto the x-space. Given a mixed-integer set X, an extended formulation of conv(X) is *compact* if the size of the matrix $(A' \mid B' \mid d')$ is polynomial in the size of the original description of X.

In Sect. 2 we give two extended formulations for the polyhedron conv(X^{I}). In the first one, we split X^{I} into smaller sets, where the fractional parts of the σ variables are fixed. We then find an extended formulation for each of these sets and we use Balas' extended formulation for the convex hull of the union of polyhedra [1] to obtain an extended formulation of conv(X^{I}).

To construct the second extended formulation, we introduce extra variables to represent all possible fractional parts taken by the continuous variables at a vertex of conv(X^{I}). We then strengthen the original inequalities and show that the system thus obtained yields an extended formulation of conv(X^{I}).

When $b_{kt} = b_t - b_k$, $1 \leq t, k \leq n$, the intersection set is called a *difference set*, denoted X^{DIF}. For conv(X^{DIF}), we prove in Sect. 3 that our two extended formulations are compact. On the other hand, we show in Sect. 4 that the extended formulations of conv(X^{I}) are not compact when the values b_{kt} are arbitrary.

We then study the polyhedron conv(X^{CMF}). We show in Sect. 5 that there is an affine transformation which maps the polyhedron conv(X^{CMF}) into the intersection of a polyhedron conv(X^{DIF}) with a polyhedron that admits an easy external description. This yields two compact extended formulations for conv(X^{CMF}), showing in particular that one can optimize over X^{CMF} in polynomial time.

2 Two Extended Formulations for the Intersection Set

The *intersection set* X^{I} is the mixed-integer set defined by (1)–(2). Note that X^{I} is the intersection of n continuous mixing sets X_k^{CM}, each one associated with a distinct variable σ_k and having common variables r, y.

In order to obtain extended formulations for conv(X^{I}), we introduce two versions of the intersection set in which the fractional parts of the continuous variables σ_k, r_t are restricted in value.

In the following we call *fractional part* any number in $[0,1)$. Also, for a number $a \in \mathbb{R}$, $f(a) = a - \lfloor a \rfloor$ denotes the fractional part of a, and for a vector $v = (v_1, \ldots, v_q)$, $f(v)$ is the vector $(f(v_1), \ldots, f(v_q))$.

In the first case, we consider a list $\mathcal{L}_\sigma = \{f^1, \ldots, f^\ell\}$ of n-vectors whose components are fractional parts and a list $\mathcal{L}_r = \{g_1, \ldots, g_m\}$ of fractional parts and define the set

$$X_1^{\mathrm{I}} = \{(\sigma, r, y) \in X^{\mathrm{I}} : f(\sigma) \in \mathcal{L}_\sigma,\ f(r_t) \in \mathcal{L}_r,\ 1 \le t \le n\} \ .$$

We say that the lists $\mathcal{L}_\sigma, \mathcal{L}_r$ are *complete* for X^{I} if for every vertex $(\bar\sigma, \bar r, \bar y)$ of $\mathrm{conv}(X^{\mathrm{I}})$, $f(\bar\sigma) \in \mathcal{L}_\sigma$ and $f(\bar r_t) \in \mathcal{L}_r$, $1 \le t \le n$.

Remark 1. If $\mathcal{L}_\sigma, \mathcal{L}_r$ are complete lists for X^{I} then $\mathrm{conv}(X_1^{\mathrm{I}}) = \mathrm{conv}(X^{\mathrm{I}})$.

In the second case, we consider a single list $\mathcal{L} = \{f_1, \ldots, f_\ell\}$ of fractional parts and define the set

$$X_2^{\mathrm{I}} = \{(\sigma, r, y) \in X^{\mathrm{I}} : f(\sigma_k) \in \mathcal{L},\ f(r_t) \in \mathcal{L},\ 1 \le k, t \le n\} \ .$$

We say that the list $\mathcal{L}$ is *complete* for X^{I} if for every vertex $(\bar\sigma, \bar r, \bar y)$ of $\mathrm{conv}(X^{\mathrm{I}})$ and for every $1 \le k, t \le n$, $f(\bar\sigma_k) \in \mathcal{L}$ and $f(\bar r_t) \in \mathcal{L}$.

Remark 2. If $\mathcal{L}$ is a complete list for X^{I} then $\mathrm{conv}(X_2^{\mathrm{I}}) = \mathrm{conv}(X^{\mathrm{I}})$.

2.1 An Extended Formulation for conv(X_1^{I})

We give an extended formulation of $\mathrm{conv}(X_1^{\mathrm{I}})$ with $\mathcal{O}(\ell m n)$ variables and $\mathcal{O}(\ell n^2)$ constraints.

For each fixed vector $f^i \in \mathcal{L}_\sigma$, let $X_{1,i}^{\mathrm{I}} = \{(\sigma, r, y) \in X_1^{\mathrm{I}} : f(\sigma) = f^i\}$. Notice that $X_1^{\mathrm{I}} = \bigcup_{i=1}^{\ell} X_{1,i}^{\mathrm{I}}$. First we find an extended formulation for each of the sets $\mathrm{conv}(X_{1,i}^{\mathrm{I}})$, $1 \le i \le \ell$, and then, since $\mathrm{conv}(X_1^{\mathrm{I}}) = \mathrm{conv}\left(\bigcup_{i=1}^{\ell} \mathrm{conv}(X_{1,i}^{\mathrm{I}})\right)$, we use Balas' extended formulation for the convex hull of the union of polyhedra [1], in the fashion introduced in [3].

In the following we assume wlog $g_1 > g_2 > \cdots > g_m$. The set $X_{1,i}^{\mathrm{I}}$ can be modeled as the following mixed-integer set:

$$\begin{aligned}
&\sigma_k = \mu_k + f_k^i,\ 1 \le k \le n\\
&r_t = \nu_t + \textstyle\sum_{j=1}^m g_j \delta_{tj},\ 1 \le t \le n\\
&\mu_k + \nu_t + \textstyle\sum_{j=1}^m g_j \delta_{tj} + y_t \ge b_{kt} - f_k^i,\ 1 \le k, t \le n\\
&\textstyle\sum_{j=1}^m \delta_{tj} = 1,\ 1 \le t \le n\\
&\mu_k, \nu_t, y_t, \delta_{tj} \ge 0,\ 1 \le t, k \le n,\ 1 \le j \le m\\
&\mu_k, \nu_t, y_t, \delta_{tj} \text{ integer},\ 1 \le t, k \le n,\ 1 \le j \le m \ .
\end{aligned}$$

Using Chvátal-Gomory rounding, the above system can be tightened to

$$\sigma_k = \mu_k + f_k^i,\ 1 \le k \le n \tag{12}$$
$$r_t = \nu_t + \sum_{j=1}^m g_j \delta_{tj},\ 1 \le t \le n \tag{13}$$
$$\mu_k + \nu_t + \sum_{j:g_j \ge f(b_{kt}-f_k^i)} \delta_{tj} + y_t \ge \lfloor b_{kt} - f_k^i \rfloor + 1,\ 1 \le k, t \le n \tag{14}$$
$$\sum_{j=1}^m \delta_{tj} = 1,\ 1 \le t \le n \tag{15}$$
$$\mu_k, \nu_t, y_t, \delta_{tj} \ge 0,\ 1 \le t, k \le n, 1 \le j \le m \tag{16}$$
$$\mu_k, \nu_t, y_t, \delta_{tj} \text{ integer},\ 1 \le t, k \le n, 1 \le j \le m\ . \tag{17}$$

Let A be the constraint matrix of (14)–(15). We show that A is a totally unimodular (TU) matrix.

Order the columns of A according to the following ordering of the variables:

$$\mu_1, \dots, \mu_n;\ y_1, \nu_1, \delta_{11}, \dots, \delta_{1m};\ y_2, \nu_2, \delta_{21}, \dots, \delta_{2m};\ \dots\ ;\ y_n, \nu_n, \delta_{n1}, \dots, \delta_{nm}\ .$$

For each row of A, the 1's that appear in a block $[y_t, \nu_t, \delta_{t1}, \dots, \delta_{tm}]$ are consecutive and start from the first position. Furthermore, for each row of A only one of these blocks contains nonzero elements.

Consider an arbitrary column submatrix of A. We give color red to all the μ_i (if any) and then, for each of the blocks $[y_t, \nu_t, \delta_{t1}, \dots, \delta_{tm}]$, we give alternating colors, always starting with blue, to the columns of this block which appear in the submatrix. Since this is an equitable bicoloring, the theorem of Ghouila-Houri [4] shows that A is TU. Since the right-hand side of the constraints is integer, the theorem of Hoffman and Kruskal implies that (14)–(15) (along with the nonnegativity conditions) define an integral polyhedron.

Since (12)–(13) just define variables σ_k, r_t, we can remove the integrality constraints from (12)–(17), thus obtaining an extended formulation for $\text{conv}(X_{1,i}^{\text{I}})$:

$$\text{conv}(X_{1,i}^{\text{I}}) = \{(\sigma, r, y) \text{ such that there exist } \delta, \mu \text{ satisfying (12)–(16)}\}\ .$$

Note that this formulation involves $\mathcal{O}(mn)$ variables and $\mathcal{O}(n^2)$ constraints.

Using Balas' description for the union of polyhedra [1], we obtain:

Theorem 3. *The following linear system is an extended formulation of the polyhedron* $\text{conv}(X_1^{\text{I}})$ *with* $\mathcal{O}(\ell mn)$ *variables and* $\mathcal{O}(\ell n^2)$ *constraints:*

$$\sigma_k = \sum_{i=1}^\ell \sigma_k^i,\ 1 \le k \le n$$
$$r_t = \sum_{i=1}^\ell r_t^i,\ 1 \le t \le n$$
$$y_t = \sum_{i=1}^\ell y_t^i,\ 1 \le t \le n$$
$$\sum_{i=1}^\ell \lambda^i = 1$$
$$\sigma_k^i = \mu_k^i + f_k^i \lambda^i,\ 1 \le k \le n, 1 \le i \le \ell$$
$$r_t^i = \nu_t^i + \sum_{j=1}^m g_j \delta_{tj}^i,\ 1 \le t \le n, 1 \le i \le \ell$$
$$\mu_k^i + \nu_t^i + \sum_{j:g_j \ge f(b_{kt}-f_k^i)} \delta_{tj}^i + y_t^i \ge (\lfloor b_{kt} - f_k^i \rfloor + 1)\lambda^i,\ 1 \le k, t \le n, 1 \le i \le \ell$$
$$\sum_{j=1}^m \delta_{tj}^i = \lambda^i,\ 1 \le t \le n, 1 \le i \le \ell$$
$$\mu_k^i, \nu_t^i, y_t^i, \delta_{tj}^i, \lambda^i \ge 0,\ 1 \le k, t \le n, 1 \le j \le m, 1 \le i \le \ell\ .$$

By Remark 1 we then obtain:

Corollary 4. *If the lists $\mathcal{L}_\sigma, \mathcal{L}_r$ are complete for X^{I} then the linear system given in Theorem 3 is an extended formulation of* $\mathrm{conv}(X^{\mathrm{I}})$.

2.2 An Extended Formulation for $\mathrm{conv}(X_2^{\mathrm{I}})$

We give an extended formulation for $\mathrm{conv}(X_2^{\mathrm{I}})$ with $\mathcal{O}(\ell n)$ variables and $\mathcal{O}(\ell n^2)$ constraints. We include zero in the list $\mathcal{L}$. Also, for technical reasons we define $f_0 = 1$. Wlog we assume $1 = f_0 > f_1 > \cdots > f_\ell = 0$.

The set X_2^{I} can be modeled as the following mixed-integer set:

$$\sigma_k = \mu^k + \textstyle\sum_{j=1}^{\ell} f_j \delta_j^k, \; 1 \le k \le n \tag{18}$$

$$r_t = \nu^t + \textstyle\sum_{j=1}^{\ell} f_j \beta_j^t, \; 1 \le t \le n \tag{19}$$

$$\sigma_k + r_t + y_t \ge b_{kt}, \; 1 \le k, t \le n \tag{20}$$

$$\textstyle\sum_{j=1}^{\ell} \delta_j^k = 1, \; 1 \le k \le n \tag{21}$$

$$\textstyle\sum_{j=1}^{\ell} \beta_j^t = 1, \; 1 \le t \le n \tag{22}$$

$$\sigma_k \ge 0, \; r_t \ge 0, \; y_t \ge 0, \; 1 \le k, t \le n$$

$$\delta_j^k, \beta_j^t \ge 0, \; 1 \le k, t \le n, 1 \le j \le \ell$$

$$\mu^k, \nu^t, y_t, \delta_j^k, \beta_j^t \text{ integer}, \; 1 \le k, t \le n, 1 \le j \le \ell \; .$$

Now define the unimodular transformation

$$\mu_0^k = \mu^k, \; \mu_j^k = \mu^k + \textstyle\sum_{h=1}^{j} \delta_h^k, \; 1 \le k \le n, 1 \le j \le \ell$$

$$\nu_0^t = \nu^t + y_t, \; \nu_j^t = \nu^t + y_t + \textstyle\sum_{h=1}^{j} \beta_h^t, \; 1 \le t \le n, 1 \le j \le \ell \; .$$

Then (18) and (19) become

$$\sigma_k = \textstyle\sum_{j=0}^{\ell-1} (f_j - f_{j+1}) \mu_j^k, \; 1 \le k \le n$$

$$r_t = -y_t + \textstyle\sum_{j=0}^{\ell-1} (f_j - f_{j+1}) \nu_j^t, \; 1 \le t \le n \; ,$$

while (21)–(22) become $\mu_\ell^k - \mu_0^k = 1$, $1 \le k \le n$, and $\nu_\ell^t - \nu_0^t = 1$, $1 \le t \le n$.

Constraints $\delta_j^k \ge 0$, $1 \le k \le n, 1 \le j \le \ell$, can be modeled as $\mu_j^k - \mu_{j-1}^k \ge 0$. Similarly $\beta_j^t \ge 0$, $1 \le t \le n, 1 \le j \le \ell$, can be modeled as $\nu_j^t - \nu_{j-1}^t \ge 0$.

Inequalities $\sigma_k \ge 0$, $1 \le k \le n$, become $\mu_0^k \ge 0$, while $r_t \ge 0$, $1 \le t \le n$, become $\nu_0^t - y_t \ge 0$.

We now model (20). Define $\ell_{kt} = \max\{\tau : f_\tau \ge f(b_{kt})\}$. Also, for an index $0 \le j \le \ell_{kt} - 1$, define $h_{kt}^j = \max\{\tau : f_\tau \ge 1 + f(b_{kt}) - f_{j+1}\}$ and for an index $\ell_{kt} \le j \le \ell - 1$, define $h_{kt}^j = \max\{\tau : f_\tau \ge f(b_{kt}) - f_{j+1}\}$.

Lemma 5. *Assume that a point (σ, r, y) satisfies* (18), (19), (21) *and* (22). *Then (σ, r, y) satisfies* (20) *if and only if the following inequalities are valid for (σ, r, y):*

$$\mu_{h_{kt}^j}^k + \nu_j^t \ge \lfloor b_{kt} \rfloor, \; 0 \le j \le \ell_{kt} - 1 \tag{23}$$

$$\mu_{h_{kt}^j}^k + \nu_j^t \ge \lfloor b_{kt} \rfloor + 1, \; \ell_{kt} \le j \le \ell - 1 \; . \tag{24}$$

Proof. We first assume that (σ, r, y) satisfies (18)–(22). Suppose $0 \le j \le \ell_{kt} - 1$. Constraint (20) can be written as $\mu^k + \nu^t + y_t + \sum_{i=1}^{\ell} f_i \delta_i^k + \sum_{i=1}^{\ell} f_i \beta_i^t \ge (\lfloor b_{kt} \rfloor - 1) + 1 + f(b_{kt})$. Since the δ_i^k's (resp. β_i^t's) are binary variables such that $\sum_{i=1}^{\ell} \delta_i^k = 1$ (resp. $\sum_{i=1}^{\ell} \beta_i^t = 1$), this implies $\mu^k + \nu^t + y_t + \sum_{i=1}^{h_{kt}^j} f_i \delta_i^k + f_{h_{kt}^j+1} + \sum_{i=1}^{j} f_i \beta_i^t + f_{j+1} \ge (\lfloor b_{kt} \rfloor - 1) + 1 + f(b_{kt})$, thus $\mu_{h_{kt}^j}^k + \nu_j^t \ge (\lfloor b_{kt} \rfloor - 1) + 1 + f(b_{kt}) - f_{h_{kt}^j+1} - f_{j+1}$. As $1 + f(b_{kt}) - f_{h_{kt}^j+1} - f_{j+1} > 0$ for $0 \le j \le \ell_{kt} - 1$ and as $\mu_{h_{kt}^j}^k + \nu_j^t$ is an integer, (23) is valid.

Suppose now $\ell_{kt} \le j \le \ell - 1$. Constraint (20) can be written as $\mu^k + \nu^t + y_t + \sum_{i=1}^{\ell} f_i \delta_i^k + \sum_{i=1}^{\ell} f_i \beta_i^t \ge \lfloor b_{kt} \rfloor + f(b_{kt})$. Similarly as before, this implies $\mu_{h_{kt}^j}^k + \nu_j^t \ge \lfloor b_{kt} \rfloor + f(b_{kt}) - f_{h_{kt}^j+1} - f_{j+1}$. As $f(b_{kt}) - f_{h_{kt}^j+1} - f_{j+1} > 0$ for $\ell_{kt} \le j \le \ell - 1$ and as $\mu_{h_{kt}^j}^k + \nu_j^t$ is an integer, (24) is valid.

Now assume that (σ, r, y) satisfies (18), (19), (21) and (22), along with (23)–(24). Specifically, assume $\sigma_k = \mu^k + f_i$ and $r_t = \nu^t + f_l$.

Suppose $l \le \ell_{kt}$. Inequality (23) for $j = l - 1$ is $\mu_{h_{kt}^{l-1}}^k + \nu^t + y^t \ge \lfloor b_{kt} \rfloor$. If $i \le h_{kt}^{l-1}$, the inequality is $\mu^k + \nu^t + y_t \ge \lfloor b_{kt} \rfloor - 1$, thus $\sigma_k + r_t + y_t \ge \lfloor b_{kt} \rfloor - 1 + f_i + f_l \ge \lfloor b_{kt} \rfloor + f(b_{kt}) = b_{kt}$. And if $i > h_{kt}^{l-1}$, the inequality is $\mu^k + \nu^t + y_t \ge \lfloor b_{kt} \rfloor$, thus $\sigma_k + r_t + y_t \ge \lfloor b_{kt} \rfloor + f_l \ge \lfloor b_{kt} \rfloor + f(b_{kt}) = b_{kt}$. Thus (20) is satisfied when $l \le \ell_{kt}$. The case $l > \ell_{kt}$ is similar. □

Thus we obtain the following result.

Theorem 6. *The following linear system is an extended formulation of the polyhedron* $\mathrm{conv}(X_2^{\mathrm{I}})$ *with* $\mathcal{O}(\ell n)$ *variables and* $\mathcal{O}(\ell n^2)$ *constraints:*

$$\sigma_k = \sum_{j=0}^{\ell-1} (f_j - f_{j+1}) \mu_j^k, \ 1 \le k \le n \tag{25}$$

$$r_t = -y_t + \sum_{j=0}^{\ell-1} (f_j - f_{j+1}) \nu_j^t, \ 1 \le t \le n \tag{26}$$

$$\mu_{h_{kt}^j}^k + \nu_j^t \ge \lfloor b_{kt} \rfloor, \ 1 \le k,t \le n, \ 0 \le j \le \ell_{kt} - 1 \tag{27}$$

$$\mu_{h_{kt}^j}^k + \nu_j^t \ge \lfloor b_{kt} \rfloor + 1, \ 1 \le k,t \le n, \ \ell_{kt} \le j \le \ell - 1 \tag{28}$$

$$\mu_\ell^k - \mu_0^k = 1, \ \nu_\ell^t - \nu_0^t = 1, \ 1 \le k,t \le n \tag{29}$$

$$\mu_j^k - \mu_{j-1}^k \ge 0, \ \nu_j^t - \nu_{j-1}^t \ge 0, \ 1 \le k,t \le n, \ 1 \le j \le \ell \tag{30}$$

$$\mu_0^k \ge 0, \ \nu_0^t - y_t \ge 0, \ y_t \ge 0, \ 1 \le k,t \le n \ . \tag{31}$$

Proof. X_2^{I} is the set of points (σ, r, y) such that there exist *integral* vectors δ, μ satisfying (25)–(31). Changing the sign of the ν_j^t and y_t variables, the constraint matrix of (27)–(31) is a *dual network matrix* (that is, the transpose of a network flow matrix), in particular it is TU. Since the right-hand side is an integer vector and since (25)–(26) just define variables σ_k, r_t,

$$\mathrm{conv}(X_2^{\mathrm{I}}) = \{(\sigma, r, y) \text{ such that there exist } \delta, \mu \text{ satisfying (25)–(31)}\} \ .$$

□

By Remark 2 we then obtain:

Corollary 7. *If the list $\mathcal{L}$ is complete for X^{I} then the linear system given in Theorem 6 is an extended formulation of* $\mathrm{conv}(X^{\mathrm{I}})$.

3 The Difference Set

The following set is the *difference set* X^{DIF}:

$$\begin{aligned}&\sigma_k + r_t + y_t \ge b_t - b_k,\ \ 0 \le k < t \le n\\ &\sigma \in \mathbb{R}^{n+1}_+,\ r \in \mathbb{R}^n_+,\ y \in \mathbb{Z}^n_+\ ,\end{aligned}$$

where $0 = b_0 \le b_1 \le \ldots \le b_n$. Note that X^{DIF} is an intersection set where $b_{kt} = b_t - b_k$, as for $k \ge t$ the constraint $\sigma_k + r_t + y_t \ge b_t - b_k$ is redundant.

Here we prove that the extended formulations given in Sect. 2 are compact for a set of the type X^{DIF}. This will be useful in Sect. 5, where we study X^{CMF}.

Theorem 8. *Let (σ^*, r^*, y^*) be a vertex of* $\mathrm{conv}(X^{\mathrm{DIF}})$. *Then there exists an index $h \in \{0, \ldots, n\}$ such that $\sigma^*_k > 0$ for $k < h$ and $\sigma^*_k = 0$ for $k \ge h$. Furthermore there is an index $\ell \ge h$ such that $f(\sigma^*_k) = f(b_\ell - b_k)$ for $0 \le k < h$.*

Proof. Let (σ^*, r^*, y^*) be a vertex of $\mathrm{conv}(X^{\mathrm{DIF}})$, let $\alpha = \max_{1 \le t \le n}\{b_t - r^*_t - y^*_t\}$ and let $T_\alpha \subseteq \{1, \ldots, n\}$ be the subset of indices for which this maximum is achieved.
CLAIM 1: *For each $1 \le k \le n$, $\sigma^*_k = \max\{0, \alpha - b_k\}$.*
PROOF. The inequalities that define X^{DIF} show that $\sigma^*_k \ge \max\{0, \alpha - b_k\}$. If $\sigma^*_k > \max\{0, \alpha - b_k\}$, then there is an $\varepsilon > 0$ such that $(\sigma^*, r^*, y^*) \pm \varepsilon(e_k, \mathbf{0}, \mathbf{0})$ are both in $\mathrm{conv}(X^{\mathrm{DIF}})$, a contradiction to the fact that (σ^*, r^*, y^*) is a vertex. This concludes the proof of the claim.

Let $h = \min\{k : \alpha - b_k \le 0\}$. (This minimum is well defined: since the only inequality involving σ_n is $\sigma_n \ge 0$, certainly $\sigma^*_n = 0$; then, by Claim 1, $\alpha - b_n \le 0$.) Since $0 = b_0 \le b_1 \le \cdots \le b_n$, Claim 1 shows that $\sigma^*_k > 0$ for $k < h$ and $\sigma^*_k = 0$ for $k \ge h$ and this proves the first part of the theorem. Furthermore $\sigma^*_k + r^*_t + y^*_t = b_t - b_k$ for all $k < h$ and $t \in T_\alpha$.
CLAIM 2: *Either $r^*_t = 0$ for some $t \in T_\alpha$ or $f(r_t) = f(b_t - b_h)$ for every $t \in T_\alpha$.*
PROOF. We use the fact that (σ^*, r^*) is a vertex of the polyhedron:

$$Q = \{(\sigma, r) \in \mathbb{R}^{n+1}_+ \times \mathbb{R}^n_+ : \sigma_k + r_t \ge b_t - b_k - y^*_t,\ 0 \le k < t \le n\}\ .$$

We now consider the following two cases:
CASE 1: $\alpha - b_h < 0$.
For $k \ge h$, the only inequality that is tight for (σ^*, r^*) and contains σ_k in its support is $\sigma_k \ge 0$. For $k < h$, the only inequalities that are tight for (σ^*, r^*) and contain σ_k in their support are $\sigma_k + r_t \ge b_t - b_k - y^*_t$, $t \in T_\alpha$.

Let e_H be the $(n+1)$-vector having the first h components equal to 1 and the others to 0, let e_{T_α} be the incidence vector of T_α and assume that $r^*_t > 0$ for

all $t \in T_\alpha$. Then the vectors $(\sigma^*, r^*) \pm \varepsilon(e_H, -e_{T_\alpha})$ for some $\varepsilon > 0$ are both in Q, contradicting the fact that (σ^*, r^*) is a vertex of Q. So $r_t^* = 0$ for some $t \in T_\alpha$.
CASE 2: $\alpha - b_h = 0$.
Then (σ^*, r^*, y^*) satisfies $\sigma_h^* + r_t^* + y_t^* = b_t - b_h$ for all $t \in T_\alpha$. Since $\sigma_h^* = 0$ and y_t^* is integer, then $f(r_t^*) = f(b_t - b_h)$ for all $t \in T_\alpha$ and this completes the proof of Claim 2.

Assume $r_t^* = 0$ for some $t \in T_\alpha$. Since $\sigma_k^* + r_t^* + y_t^* = b_t - b_k$ for all $k < h$ and y_t^* is an integer, then $f(\sigma_k^*) = f(b_t - b_k)$ for all $k < h$.

If $f(r_t^*) = f(b_t - b_h)$ for all $t \in T_\alpha$, since $\sigma_k^* + r_t^* + y_t^* = b_t - b_k$ for all $t \in T_\alpha$ and for all $k < h$ and since y^* is an integer vector, then $f(\sigma_k^*) = f(b_h - b_k)$ for all $k < h$. □

Corollary 9. *If (σ^*, r^*, y^*) is a vertex of $\text{conv}(X^{\text{DIF}})$, then $f(r_t^*) \in \{f(b_t - b_k), 1 \le k \le n\}$ for $1 \le t \le n$.*

Proof. The result follows from Theorem 8 and the observation that at a vertex of $\text{conv}(X^{\text{DIF}})$ either $r_t^* = 0$ or $\sigma_k^* + r_t^* + y_t^* = b_t - b_k$ for some k. □

We then obtain the following result.

Theorem 10. *The polyhedron $\text{conv}(X^{\text{DIF}})$ admits an extended formulation of the type given in Theorem 3 with $\mathcal{O}(n^5)$ variables and $\mathcal{O}(n^4)$ constraints and an extended formulation of the type given in Theorem 6 with $\mathcal{O}(n^3)$ variables and $\mathcal{O}(n^4)$ constraints.*

Proof. Recall that X^{DIF} is an intersection set. Define $\mathcal{L}_\sigma$ as the set of all possible $(n+1)$-vectors of fractional parts taken by σ at a vertex of $\text{conv}(X^{\text{DIF}})$ and $\mathcal{L}_r$ as the set of all possible fractional parts taken by the variables r_t at a vertex of $\text{conv}(X^{\text{DIF}})$. Since these lists are complete for X^{DIF}, Corollary 4 implies that the linear system given in Theorem 3 is an extended formulation of $\text{conv}(X^{\text{DIF}})$. By Theorem 8, $\ell = |\mathcal{L}_\sigma| = \mathcal{O}(n^2)$ and by Corollary 9, $m = |\mathcal{L}_r| = \mathcal{O}(n^2)$, therefore this formulation has $\mathcal{O}(n^5)$ variables and $\mathcal{O}(n^4)$ constraints.

Now define $\mathcal{L}$ as the set of all possible fractional parts taken by the variables σ_k, r_t at a vertex of $\text{conv}(X^{\text{DIF}})$. Since this list is complete for X^{DIF}, by Corollary 7 the system given in Theorem 6 is an extended formulation of $\text{conv}(X^{\text{DIF}})$. Since $\ell = |\mathcal{L}| = \mathcal{O}(n^2)$ (see Theorem 8 and Corollary 9), this formulation has $\mathcal{O}(n^3)$ variables and $\mathcal{O}(n^4)$ constraints. □

We point out that the result of the above theorem can be improved as follows.

Consider the first formulation. If for each set $X_{1,i}^{\text{I}}$ we define a different list of fractional parts for the variables r_t, say $\mathcal{L}_r^i$, then we can easily choose such lists so that $|\mathcal{L}_r^i| = \mathcal{O}(n)$. In this case the first extended formulation for $\text{conv}(X^{\text{DIF}})$ involves $\mathcal{O}(n^4)$ variables.

Consider now the second formulation. Instead of defining a unique list for all variables, we can define a list for each variable, say $\mathcal{L}_{\sigma_k}$ and $\mathcal{L}_{r_t}$, $1 \le k, t \le n$. It is not difficult to verify that the construction of the extended formulation can be carried out with straightforward modifications. Since in this case $|\mathcal{L}_{\sigma_k}| = \mathcal{O}(n)$ (by Theorem 8) and $|\mathcal{L}_{r_t}| = \mathcal{O}(n)$ (by Corollary 9), the second extended formulation involves $\mathcal{O}(n^2)$ variables and $\mathcal{O}(n^3)$ constraints.

Theorem 11. *The polyhedron* $\mathrm{conv}(X^{\mathrm{CMF}})$ *admits an extended formulation with* $\mathcal{O}(n^2)$ *variables and* $\mathcal{O}(n^3)$ *constraints.*

4 Intersection Sets with an Exponential Number of Fractional Parts

In this section we show that the extended formulations derived in Sect. 2 are not compact in general. Specifically, we prove here the following result:

Theorem 12. *In the set of vertices of the polyhedron defined by*

$$\sigma_k + r_t \geq \frac{3^{(t-1)n+k}}{3^{n^2+1}}, \ 1 \leq k, t \leq n \tag{32}$$

$$\sigma \in \mathbb{R}^n_+, \ r \in \mathbb{R}^n_+ \tag{33}$$

the number of distinct fractional parts taken by variable σ_n *is exponential in* n.

Remark 13. Since the vertices of the above polyhedron are the vertices on the face defined by $y = \mathbf{0}$ of the polyhedron $\mathrm{conv}(X^{\mathrm{I}})$ with the same right-hand side, Theorem 12 shows that any extended formulation that explicitly takes into account a list of all possible fractional parts taken at a vertex by the continuous variables (such as those introduced to model $\mathrm{conv}(X_1^{\mathrm{I}})$ and $\mathrm{conv}(X_2^{\mathrm{I}})$) will not be compact in general.

Now let b_{kt} be as in the theorem, i.e. $b_{kt} = \frac{3^{(t-1)n+k}}{3^{n^2+1}}$, $1 \leq k, t \leq n$.

Remark 14. $b_{kt} < b_{k't'}$ if and only if $(t, k) \prec (t', k')$, where $\prec$ denotes the lexicographic order. Thus $b_{11} < b_{21} < \cdots < b_{n1} < b_{12} < \cdots < b_{nn}$.

Lemma 15. *The following properties hold.*

1. *Suppose that* $\alpha \in \mathbb{Z}^q_+$ *with* $\alpha_j < \alpha_{j+1}$ *for* $1 \leq j \leq q-1$, *and define* $\Phi(\alpha) = \sum_{j=1}^{q} (-1)^{q-j} 3^{\alpha_j}$. *Then* $\frac{1}{2} 3^{\alpha_q} < \Phi(\alpha) < \frac{3}{2} 3^{\alpha_q}$.
2. *Suppose that* α *is as above and* $\beta \in \mathbb{Z}^{q'}_+$ *is defined similarly. Then* $\Phi(\alpha) = \Phi(\beta)$ *if and only if* $\alpha = \beta$.

Proof. 1. $\sum_{j=0}^{\alpha_q - 1} 3^j = \frac{3^{\alpha_q} - 1}{3-1} < \frac{1}{2} 3^{\alpha_q}$. Now $\Phi(\alpha) \geq 3^{\alpha_q} - \sum_{j=1}^{\alpha_q - 1} 3^j > 3^{\alpha_q} - \frac{1}{2} 3^{\alpha_q} = \frac{1}{2} 3^{\alpha_q}$, and $\Phi(\alpha) \leq 3^{\alpha_q} + \sum_{j=1}^{\alpha_q - 1} 3^j < 3^{\alpha_q} + \frac{1}{2} 3^{\alpha_q} = \frac{3}{2} 3^{\alpha_q}$.

2. Suppose $\alpha \neq \beta$. Wlog we assume $q \geq q'$. Assume first $(\alpha_{q-q'+1}, \ldots, \alpha_q) = \beta$. Then $q > q'$ (otherwise $\alpha = \beta$) and, after defining $\bar{\alpha} = (\alpha_1, \ldots, \alpha_{q-q'})$, we have $\Phi(\alpha) - \Phi(\beta) = \Phi(\bar{\alpha}) > 0$ by 1. Now assume $(\alpha_{q-q'+1}, \ldots, \alpha_q) \neq \beta$. Define $h = \min\{\tau : \alpha_{q-\tau} \neq \beta_{q'-\tau}\}$ and suppose $\alpha_{q-h} > \beta_{q'-h}$ (the other case is similar). If we define the vectors $\bar{\alpha} = (\alpha_1, \ldots, \alpha_{q-h})$ and $\bar{\beta} = (\beta_1, \ldots, \beta_{q'-h})$, 1. gives $\Phi(\alpha) - \Phi(\beta) = \Phi(\bar{\alpha}) - \Phi(\bar{\beta}) > \frac{1}{2} 3^{\alpha_{q-h}} - \frac{3}{2} 3^{\beta_{q'-h}} \geq 0$, as $\alpha_{q-h} > \beta_{q'-h}$. □

We now give a construction of an exponential family of vertices of (32)–(33) such that at each vertex variable σ_n takes a distinct fractional part. Therefore this construction proves Theorem 12.

Let $(k_1, \ldots, k_m)$ and $(t_1, \ldots, t_{m-1})$ be two increasing subsets of $\{1, \ldots, n\}$ with $k_1 = 1$ and $k_m = n$. For $1 \le k, t \le n$, let $p(k) = \max\{j : k_j \le k\}$ and $q(t) = \max\{j : t_j \le t\}$, with $q(t) = 0$ if $t < t_1$.

Consider the following system of equations:

$$\begin{aligned} \sigma_{k_1} &= 0 \\ \sigma_{k_j} + r_{t_j} &= b_{k_j t_j}, \quad & 1 \le j \le m-1 \\ \sigma_{k_{j+1}} + r_{t_j} &= b_{k_{j+1} t_j}, \quad & 1 \le j \le m-1 \\ \sigma_{k_{q(t)+1}} + r_t &= b_{k_{q(t)+1} t}, \quad & t \notin \{t_1, \ldots, t_{m-1}\} \\ \sigma_k + r_{t_{p(k)}} &= b_{k t_{p(k)}}, \quad & k \notin \{k_1, \ldots, k_m\} \ . \end{aligned}$$

The unique solution of this system is:

$$\begin{aligned} \sigma_{k_1} &= 0 \\ \sigma_{k_j} &= \textstyle\sum_{\ell=1}^{j-1} b_{k_{\ell+1} t_\ell} - \sum_{\ell=1}^{j-1} b_{k_\ell t_\ell}, \quad & 2 \le j \le m \\ r_{t_j} &= \textstyle\sum_{\ell=1}^{j} b_{k_\ell t_\ell} - \sum_{\ell=1}^{j-1} b_{k_{\ell+1} t_\ell}, \quad & 1 \le j \le m-1 \\ \sigma_k &= b_{k t_{p(k)}} - r_{t_{p(k)}}, & k \notin \{k_1, \ldots, k_m\} \\ r_t &= b_{k_{q(t)+1} t} - \sigma_{k_{q(t)+1}}, & t \notin \{t_1, \ldots, t_{m-1}\} \ . \end{aligned}$$

As each of these variables σ_k, r_t takes a value of the form $\Phi(\alpha)/3^{n^2+1}$, by Lemma 15 (i) we have that $\sigma_{k_j} > \frac{1}{2} b_{k_j t_{j-1}} > 0$ for $2 \le j \le m$, $r_{t_j} > \frac{1}{2} b_{k_j t_j} > 0$ for $1 \le j \le m-1$, $\sigma_k > \frac{1}{2} b_{k t_{p(k)}} > 0$ for $k \notin \{k_1, \ldots, k_m\}$ and $r_t > \frac{1}{2} b_{k_{q(t)+1} t} > 0$ for $t \notin \{t_1, \ldots, t_{m-1}\}$. Therefore the nonnegativity constraints are satisfied.

Now we show that the other constraints are satisfied. Consider the k, t constraint with $t \notin \{t_1, \ldots, t_{m-1}\}$. We distinguish some cases.

1. $p(k) \le q(t)$. Then $\sigma_k + r_t \ge r_t > \frac{1}{2} b_{k_{q(t)+1} t} \ge \frac{1}{2} b_{k_{p(k)+1} t} \ge \frac{3}{2} b_{kt} > b_{kt}$.
2. $p(k) > q(t)$ and $k \notin \{k_1, \ldots, k_m\}$. Then $\sigma_k + r_t \ge \sigma_k > \frac{1}{2} b_{k t_{p(k)}} \ge \frac{1}{2} b_{k t_{q(t)+1}} \ge \frac{3^n}{2} b_{kt} > b_{kt}$.
3. $p(k) = q(t) + 1$ and $k = k_j$ for some $1 \le j \le m$ (thus $p(k) = j = q(t) + 1$). In this case the k, t constraints is satisfied at equality by construction.
4. $p(k) > q(t) + 1$ and $k = k_j$ for some $1 \le j \le m$ (thus $p(k) = j > q(t) + 1$). Then $\sigma_k + r_t \ge \sigma_k > \frac{1}{2} b_{k t_{j-1}} \ge \frac{1}{2} b_{k t_{q(t)+1}} \ge \frac{3^n}{2} b_{kt} > b_{kt}$.

The argument with $k \notin \{k_1, \ldots, k_m\}$ is similar.

Finally suppose that $k = k_j$ and $t = t_h$ with $h \notin \{j-1, j\}$. If $h > j$, $\sigma_k + r_t \ge r_t > \frac{1}{2} b_{k_h t_h} \ge \frac{3}{2} b_{k_j t_h} > b_{kt}$. If $h < j-1$, $\sigma_k + r_t \ge \sigma_k > \frac{1}{2} b_{k_j t_{j-1}} \ge \frac{3^n}{2} b_{k_j t_h} > b_{kt}$. This shows that the solution is feasible and as it is unique, it defines a vertex of (32)–(33).

Now let $a_{kt} = (t-1)n + k$, so that $b_{kt} = 3^{a_{kt}}/3^{n^2+1}$ and take

$$\alpha = (a_{k_1 t_1}, a_{k_2 t_1}, a_{k_2 t_2}, a_{k_3 t_2}, \ldots, a_{k_m t_{m-1}}) \ .$$

As $\sigma_n = \Phi(\alpha)/3^{n^2+1}$, Lemma 15 (ii) implies that in any two vertices constructed as above by different sequences $(k_1,\dots,k_m)$, $(t_1,\dots,t_{m-1})$ and $(k'_1,\dots,k'_{m'})$, $(t'_1,\dots,t'_{m'-1})$, the values of σ_n are distinct numbers in the interval $(0,1)$. As the number of such sequences is exponential in n, this proves Theorem 12.

5 An Extended Formulation for conv(X^{CMF})

Now we address the question of showing that the linear optimization problem over the continuous mixing set with flows (3)–(5) is solvable in polynomial time. Specifically we derive compact extended formulations for $\mathrm{conv}(X^{\mathrm{CMF}})$.

We assume that $0 < b_1 \le \cdots \le b_n$. Consider the set Z:

$$s + r_t + y_t \ge b_t,\ \ 1 \le t \le n \tag{34}$$
$$s + r_k + x_k + r_t + y_t \ge b_t,\ \ 1 \le k < t \le n \tag{35}$$
$$s + r_t + x_t \ge b_t,\ \ 1 \le t \le n \tag{36}$$
$$s \in \mathbb{R}_+,\ r \in \mathbb{R}^n_+,\ x \in \mathbb{R}^n,\ y \in \mathbb{Z}^n_+\ . \tag{37}$$

Note that x is unrestricted in Z.

Proposition 16. *Let X^{CMF} and Z be defined on the same vector b. Then $X^{\mathrm{CMF}} \subseteq Z$ and $X^{\mathrm{CMF}} = Z \cap \{(s,r,x,y) :\ \mathbf{0} \le x \le y\}$.*

Proof. Clearly (34)–(37) are valid for the points in X^{CMF}. The only inequalities that define X^{CMF} but do not appear in the definition of Z are $\mathbf{0} \le x \le y$. □

Lemma 17. *The $3n+1$ extreme rays of $\mathrm{conv}(X^{\mathrm{CMF}})$ are the vectors $(1,\mathbf{0},\mathbf{0},\mathbf{0})$, $(0,e_i,\mathbf{0},\mathbf{0})$, $(0,\mathbf{0},\mathbf{0},e_i)$, $(0,\mathbf{0},e_i,e_i)$. The $3n+1$ extreme rays of $\mathrm{conv}(Z)$ are the vectors $(1,\mathbf{0},-\mathbf{1},\mathbf{0})$, $(0,e_i,-e_i,\mathbf{0})$, $(0,\mathbf{0},e_i,\mathbf{0})$, $(0,\mathbf{0},\mathbf{0},e_i)$. Therefore both recession cones of $\mathrm{conv}(X^{\mathrm{CMF}})$ and $\mathrm{conv}(Z)$ are full-dimensional simplicial cones, thus showing that $\mathrm{conv}(X^{\mathrm{CMF}})$ and $\mathrm{conv}(Z)$ are full-dimensional polyhedra.*

Proof. The first part is obvious. We characterize the extreme rays of $\mathrm{conv}(Z)$. The recession cone C of $\mathrm{conv}(Z)$ is defined by

$$s + r_k + x_k + r_t + y_t \ge 0,\ \ 1 \le k < t \le n$$
$$s + r_t + x_t \ge 0,\ \ 1 \le t \le n$$
$$s \in \mathbb{R}_+,\ r \in \mathbb{R}^n_+,\ x \in \mathbb{R}^n,\ y \in \mathbb{R}^n_+\ .$$

One can verify that the vectors $\rho = (1,\mathbf{0},-\mathbf{1},\mathbf{0})$, $u_i = (0,e_i,-e_i,\mathbf{0})$, $v_i = (0,\mathbf{0},e_i,\mathbf{0})$, $z_i = (0,\mathbf{0},\mathbf{0},e_i)$ are extreme rays of $\mathrm{conv}(Z)$ by checking that each of them satisfies at equality $3n$ linearly independent inequalities defining C (including nonnegativity constraints).

Thus we only have to show that every vector in C can be expressed as conic combination of the above rays. Let $(\bar{s},\bar{r},\bar{x},\bar{y})$ be in C. Notice that $(\bar{s},\bar{r},\bar{x},\bar{y}) = \bar{s}\rho + \sum_{i=1}^n \bar{r}_i u_i + \sum_{i=1}^n (\bar{s} + \bar{r}_i + \bar{x}_i) v_i + \sum_{i=1}^n \bar{y}_i w_i$. Since $(\bar{s},\bar{r},\bar{x},\bar{y}) \in C$, all the coefficients appearing in the above combination are nonnegative.

It can also be checked that the above rays are linearly independent. □

Lemma 18. *Let* (s^*, r^*, x^*, y^*) *be a vertex of* $\mathrm{conv}(Z)$*. Then*

$$s^* = \max\{0;\ b_t - r^*_t - y^*_t,\ 1 \le t \le n\}\ ,$$
$$x^*_k = \max\{b_k - s^* - r^*_k;\ b_t - s^* - r^*_k - r^*_t - y^*_t,\ 1 \le k < t \le n\}\ .$$

Proof. Assume $s^* > 0$ and $s^* + r^*_t + y^*_t > b_t$, $1 \le t \le n$. Then, there is an $\varepsilon \ne 0$ such that $(s^*, r^*, x^*, y^*) \pm \varepsilon(1, \mathbf{0}, -\mathbf{1}, \mathbf{0})$ belong to $\mathrm{conv}(Z)$, a contradiction. This proves the first statement. The second one is obvious. □

Proposition 19. *Let* (s^*, r^*, x^*, y^*) *be a vertex of* $\mathrm{conv}(Z)$*. Then* $\mathbf{0} \le x^* \le y^*$*.*

Proof. Assume that $\{t : x^*_t < 0\} \ne \emptyset$ and let $h = \min\{t : x^*_t < 0\}$. Then $s^* + r^*_h > b_h > 0$ and together with $y^*_h \ge 0$, this implies $s^* + r^*_h + y^*_h > b_h$.
Claim: $r^*_h > 0$.
Proof. Assume $r^*_h = 0$. Then $s^* > b_h > 0$. By Lemma 18, $s^* + r^*_t + y^*_t = b_t$ for some index t. It follows that $s^* \le b_t$, thus $t > h$ (as $b_h < s^* \le b_t$). Equation $s^* + r^*_t + y^*_t = b_t$, together with $s^* + r^*_h + x^*_h + r^*_t + y^*_t \ge b_t$, gives $r^*_h + x^*_h \ge 0$, thus $r^*_h > 0$, as $x^*_h < 0$, and this concludes the proof of the claim.

The inequalities $s^* + r^*_h + y^*_h > b_h$ and $r^*_k + x^*_k \ge 0$, $1 \le k < h$, imply $s^* + r^*_k + x^*_k + r^*_h + y^*_h > b_h$, $1 \le k < h$.

All these observations show the existence of an $\varepsilon \ne 0$ such that both points $(s^*, r^*, x^*, y^*) \pm \varepsilon(0, e_h, -e_h, \mathbf{0})$ belong to $\mathrm{conv}(Z)$, a contradiction to the fact that the point (s^*, r^*, x^*, y^*) is a vertex of $\mathrm{conv}(Z)$. Thus $x^* \ge \mathbf{0}$.

Suppose now that there exists h such that $x^*_h > y^*_h$. Then constraint $s + r_h + y_h \ge b_h$ gives $s^* + r^*_h + x^*_h > b_h$. Lemma 18 then implies that $s^* + r^*_h + x^*_h + r^*_t + y^*_t = b_t$ for some $t > h$. This is not possible, as inequalities $x^*_h > y^*_h \ge 0$, $r^*_h \ge 0$ and $s^* + r^*_t + y^*_t \ge b_t$ imply $s^* + r^*_h + x^*_h + r^*_t + y^*_t > b_t$. Thus $x^* \le y^*$. □

For the main theorem of this section we present a lemma whose proof is given in [2].

For a polyhedron P in $\mathbb{R}^n$ and a vector $a \in \mathbb{R}^n$, let $\mu_P(a)$ be the value $\min\{ax,\ x \in P\}$ and $M_P(a)$ be the face $\{x \in P : ax = \mu_P(a)\}$, where $M_P(a) = \emptyset$ whenever $\mu_P(a) = -\infty$.

Lemma 20. *Let* $P \subseteq Q$ *be two pointed polyhedra in* $\mathbb{R}^n$*, with the property that every vertex of* Q *belongs to* P*. Let* $Cx \ge d$ *be a system of inequalities that are valid for* P *such that for every inequality* $cx \ge \delta$ *of the system,* $P \not\subseteq \{x \in \mathbb{R}^n : cx = \delta\}$*. If for every* $a \in \mathbb{R}^n$ *such that* $\mu_P(a)$ *is finite but* $\mu_Q(a) = -\infty$*,* $Cx \ge d$ *contains an inequality* $cx \ge \delta$ *such that* $M_P(a) \subseteq \{x \in \mathbb{R}^n : cx = \delta\}$*, then* $P = Q \cap \{x \in \mathbb{R}^n : Cx \ge d\}$*.*

Proof. See [2].

Theorem 21. *Let* X^{CMF} *and* Z *be defined on the the same vector* b*. Then* $\mathrm{conv}(X^{\mathrm{CMF}}) = \mathrm{conv}(Z) \cap \{(s, r, x, y) :\ \mathbf{0} \le x \le y\}$*.*

Proof. By Proposition 16, $\mathrm{conv}(X^{\mathrm{CMF}}) \subseteq \mathrm{conv}(Z)$. By Propositions 19 and 16, every vertex of $\mathrm{conv}(Z)$ belongs to $\mathrm{conv}(X^{\mathrm{CMF}})$.

Let $a = (h, d, p, q)$, $h \in \mathbb{R}^1$, $d \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, $q \in \mathbb{R}^n$, be such that $\mu_{\mathrm{conv}(X^{\mathrm{CMF}})}(a)$ is finite and $\mu_{\mathrm{conv}(Z)}(a) = -\infty$. Since by Lemma 17, the extreme rays of $\mathrm{conv}(Z)$ that are not rays of $\mathrm{conv}(X^{\mathrm{CMF}})$ are the vectors $(0, \mathbf{0}, e_i, \mathbf{0})$, $(0, e_i, -e_i, \mathbf{0})$ and $(1, \mathbf{0}, -\mathbf{1}, \mathbf{0})$, then either $p_i < 0$ for some index i or $d_i < p_i$ for some index i or $h < \sum_{t=1}^n p_t$.

If $p_i < 0$, then $M_{\mathrm{conv}(X^{\mathrm{CMF}})}(a) \subseteq \{(s, r, x, y) : x_i = y_i\}$.

If $d_i < p_i$, then $M_{\mathrm{conv}(X^{\mathrm{CMF}})}(a) \subseteq \{(s, r, x, y) : x_i = 0\}$, otherwise, given an optimal solution with $x_i > 0$, we could increase r_i by a small $\varepsilon > 0$ and decrease x_i by ε, thus obtaining a feasible point with lower objective value.

If $h < \sum_{t=1}^n p_t$, let $N^+ = \{j : p_j > 0\}$ and $k = \min\{j : j \in N^+\}$: we show that $M_{\mathrm{conv}(X^{\mathrm{CMF}})}(a) \subseteq \{(s, r, x, y) : x_k = 0\}$. Suppose that $x_k > 0$ in some optimal solution. As the solution is optimal and $p_k > 0$, we cannot just decrease x_k and remain feasible. Thus $s + r_k + x_k = b_k$, which implies that $s < b_k$. Then for all $j \in N^+$ we have $r_j + x_j \geq b_j - s > b_j - b_k \geq 0$, as $j \geq k$. Since we can assume $d_t \geq p_t$ for every t (otherwise we are in the previous case), $r_t = 0$ for every t: if not, chosen an index t such that $r_t > 0$, one can decrease r_t by a small $\varepsilon > 0$ and increase x_t by ε, thus obtaining a feasible point with lower objective value, a contradiction. So $r_t = 0$ for every t and thus, since $r_j + x_j > 0$ for all $j \in N^+$, we have $x_j > 0$ for all $j \in N^+$. Then we can increase s by a small $\varepsilon > 0$ and decrease x_j by ε for all $j \in N^+$. The new point is feasible in X^{CMF} and has lower objective value, a contradiction.

We have shown that for every vector a such that $\mu_{\mathrm{conv}(X^{\mathrm{CMF}})}(a)$ is finite and $\mu_{\mathrm{conv}(Z)}(a) = -\infty$, the system $\mathbf{0} \leq x \leq y$ contains an inequality which is tight for the points in $M_{\mathrm{conv}(X^{\mathrm{CMF}})}(a)$. To complete the proof, since $\mathrm{conv}(X^{\mathrm{CMF}})$ is full-dimensional (Lemma 17), the system $\mathbf{0} \leq x \leq y$ does not contain an improper face of $\mathrm{conv}(X^{\mathrm{CMF}})$. So we can now apply Lemma 20 to $\mathrm{conv}(X^{\mathrm{CMF}})$, $\mathrm{conv}(Z)$ and the system $\mathbf{0} \leq x \leq y$. □

Therefore, if we have a compact extended formulation of $\mathrm{conv}(Z)$, then this will immediately yield a compact extended formulation of $\mathrm{conv}(X^{\mathrm{CMF}})$. Such a formulation exists, as Z is equivalent to a difference set:

Theorem 22. *Let X^{DIF} be a difference set and X^{CMF} be defined on the same vector b. The affine transformation $\sigma_0 = s$, $\sigma_t = s + r_t + x_t - b_t$, $1 \leq t \leq n$, maps* $\mathrm{conv}(X^{\mathrm{CMF}})$ *into* $\mathrm{conv}(X^{\mathrm{DIF}}) \cap \{(\sigma, r, y) : 0 \leq \sigma_k - \sigma_0 - r_k + b_k \leq y_k,\ 1 \leq k \leq n\}$.

Proof. Let Z be defined on the same vector b. It is straightforward to check that the affine transformation $\sigma_0 = s$, $\sigma_t = s + r_t + x_t - b_t$, $1 \leq t \leq n$, maps $\mathrm{conv}(Z)$ into $\mathrm{conv}(X^{\mathrm{DIF}})$. By Theorem 21, $\mathrm{conv}(X^{\mathrm{CMF}}) = \mathrm{conv}(Z) \cap \{(s, r, x, y) : \mathbf{0} \leq x \leq y\}$ and the result follows. □

Then the extended formulations of $\mathrm{conv}(X^{\mathrm{DIF}})$ described in Sects. 2–3 give extended formulations of $\mathrm{conv}(X^{\mathrm{CMF}})$ which are compact. By Theorem 11 we have:

Theorem 23. *The polyhedron* $\text{conv}(X^{\text{CMF}})$ *admits an extended formulation with* $\mathcal{O}(n^2)$ *variables and* $\mathcal{O}(n^3)$ *constraints. It follows that the linear optimization problem over* X^{CMF} *can be solved in polynomial time.*

5.1 An Extended Formulation for the Two Stage Stochastic Lot-sizing Problem with Constant Capacities and Backlogging

We briefly consider the set $X^{\text{CMF}} \cap W$, where

$$W = \{(s, r, x, y) : l_j \le y_j \le u_j,\ l_{jk} \le y_j - y_k \le u_{jk},\ 1 \le j, k \le n\}\ ,$$

with $l_j, u_j, l_{jk}, u_{jk} \in \mathbb{Z} \cup \{+\infty, -\infty\}$, $1 \le j, k \le n$. We assume that for each $1 \le i \le n$, W contains a point satisfying $y_i \ge 1$.

In the following we show that an extended formulation of $\text{conv}(X^{\text{CMF}} \cap W)$ is obtained by adding the inequalities defining W to one of the extended formulations of $\text{conv}(X^{\text{CMF}})$ derived above. The proof uses the same technique as in Sect. 5, where Z (resp. X^{CMF}) has to be replaced with $Z \cap W$ (resp. $X^{\text{CMF}} \cap W$). We only point out the main differences.

To see that the proof of Theorem 21 is still valid, note that the extreme rays of $\text{conv}(Z \cap W)$ are of the following types:

1. $(1, \mathbf{0}, -\mathbf{1}, \mathbf{0})$, $(0, e_i, -e_i, \mathbf{0})$, $(0, \mathbf{0}, e_i, \mathbf{0})$;
2. $(0, \mathbf{0}, \mathbf{0}, y)$ for suitable vectors $y \in \mathbb{Z}^n$.

However, the rays of type 2 are also rays of $\text{conv}(X^{\text{CMF}} \cap W)$. Also, the condition that for every index i, W contains a vector with $y_i > 0$, shows that none of the inequalities $0 \le x_i \le y_i$ defines an improper face of $\text{conv}(X^{\text{CMF}} \cap W)$ and Lemma 20 can still be applied. Thus the proof of Theorem 21 is still valid.

The rest of the proof is a straightforward adaptation of Theorem 22.

Since (9)–(11) define a set of the type $X^{\text{CMF}} \cap W$ (assuming $C = 1$ wlog), the above result yields an extended formulation for the feasible region of the two stage stochastic lot-sizing problem with constant capacities and backlogging.

References

1. Balas, E.: Disjunctive programming: properties of the convex hull of feasible points. Invited paper, Discrete Appl. Math. **89** (1988) 1–44
2. Conforti, M., Di Summa, M., Wolsey, L. A.: The mixing set with flows. CORE DP 2005/92, Université catholique de Louvain. SIAM J. Discret. Math. (to appear)
3. Conforti, M., Wolsey, L. A.: Compact formulations as a union of polyhedra. CORE DP 2005/62, Université catholique de Louvain, submitted to Math. Program.
4. Ghouila-Houri, A.: Caractérisations des matrices totalement unimodulaires. C. R. Acad. Sci. Paris **254** (1968) 155–163
5. Miller, A. J., Wolsey, L. A.: Tight formulations for some simple mixed integer programs and convex objective integer programs. Math. Program. B **98** (2003) 73–88
6. Van Vyve, M.: A solution approach of production planning problems based on compact formulations for single-item lot-sizing models. Ph. D. thesis, Faculté des Sciences Appliquées, Université catholique de Louvain (2003)
7. Van Vyve, M.: The continuous mixing polyhedron. Math. Oper. Res. **30** (2005) 441–452