Math. Program., Ser. A (2010) 121:337–352
DOI 10.1007/s10107-008-0236-7

FULL LENGTH PAPER

Lattice based extended formulations for integer linear equality systems

Karen Aardal · Laurence A. Wolsey

Received: 14 February 2007 / Accepted: 20 June 2008 / Published online: 23 July 2008

Abstract We study different extended formulations for the set $X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$ with $\boldsymbol{A} \in \mathbb{Z}^{m\times n}$ in order to tackle the feasibility problem for the set $X \cap \mathbb{Z}^n_+$. Pursuing the work of Aardal, Lenstra et al. using the reformulation $X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{x} - \boldsymbol{x}^0 = \boldsymbol{Q}\boldsymbol{\lambda},\ \boldsymbol{\lambda} \in \mathbb{Z}^{n-m}\}$, our aim is to derive reformulations of the form $\{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{P}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{T}\boldsymbol{\mu},\ \boldsymbol{\mu} \in \mathbb{Z}^s\}$ with $0 \le s \le n - m$ where preferably all the coefficients of $\boldsymbol{P}$ are small compared to the coefficients of $\boldsymbol{A}$ and $\boldsymbol{T}$. In such cases the new variables $\boldsymbol{\mu}$ appear to be good branching directions, and in certain circumstances permit one to deduce rapidly that the instance is infeasible. We give a polynomial time algorithm for identifying such $\boldsymbol{P}$, $\boldsymbol{T}$ if possible, and for the case that $\boldsymbol{A}$ has one row $\boldsymbol{a}$ we analyze the reformulation when $s = 1$, that is, one μ-variable is introduced. In particular, we determine the integer width of the extended formulations in the direction of the μ-variable, and derive a lower bound on the Frobenius number of $\boldsymbol{a}$. We conclude with some preliminary tests to see if the reformulations are effective when the number s of additional constraints and variables is limited.

This work was partly carried out within the framework of ADONET, a European network in Algorithmic Discrete Optimization, contract no. MRTN-CT-2003-504438. The first author is financed in part by the Dutch BSIK/BRICKS project. The research was carried out in part while the second author visited CWI, Amsterdam with the support of the NWO visitor grant number B 61-556.

K. Aardal (✉)
Centrum voor Wiskunde en Informatica, Postbus 94079, 1090 GB Amsterdam, The Netherlands
e-mail: karen.aardal@cwi.nl

K. Aardal
Faculteit Wiskunde en Informatica, Technische Universiteit Eindhoven,
Postbus 513, 5600 MB Eindhoven, The Netherlands

L. A. Wolsey
CORE and INMA Université Catholique de Louvain, 34 Voie du Roman Pays,
1348 Louvain-la-Neuve, Belgium
e-mail: laurence.wolsey@uclouvain.be

Keywords Integer programming · Lattice basis reduction · Integer width · Frobenius number

Mathematics Subject Classification (2000) 90C10 · 45A05 · 11Y50

1 Introduction

Over the years a variety of approaches, other than cutting planes, have been proposed for the reformulation of part or all of the feasible region of a linear integer program $\{\boldsymbol{x} \in \mathbb{Z}^n_+ \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$. These include the conversion to a knapsack constraint [5,20], coefficient reduction of knapsack constraints [4,7,19] that can be used in preprocessing, and lattice reformulations in which the set $X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$ is rewritten as $\{\boldsymbol{x} \mid \boldsymbol{x} = \boldsymbol{x}^0 + \boldsymbol{Q}\boldsymbol{\lambda},\ \boldsymbol{\lambda} \in \mathbb{Z}^{n-m}\}$ involving $n - m$ new variables, see among others [6]. In a recent series of papers Aardal, Lenstra and others [2,3] have demonstrated the interest of the latter reformulation when $\boldsymbol{Q}$ is a reduced basis, and in particular when the reduced basis $\boldsymbol{Q}$ has a partition $\boldsymbol{Q} = (\boldsymbol{R}, \boldsymbol{S})$ in which the coefficients of $\boldsymbol{R}$ are small and those of $\boldsymbol{S}$ large. For the special case when $m = 1$ and $\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2$ with $\boldsymbol{p}^1$, $\boldsymbol{p}^2$ small and M_1, M_2 large, they show that $\boldsymbol{S}$ contains a single column and the corresponding variable λ_{n-m} is an important variable for branching. They have also derived strong lower bounds on the Frobenius number when $M_2 = 1$ and $\boldsymbol{p}^1 \in \mathbb{Z}^n_{>0}$. Computationally it has been shown in several papers [1,3,18] that the reformulation is much more effective than the original formulation for certain hard integer programs when using a commercial branch-and-bound system or a specialized enumeration algorithm. Another viewpoint is that branching on one of the new variables is equivalent to branching on an easily calculated hyperplane in the original space. Interesting tests of branching on hyperplanes arising from Gomory mixed integer cuts have been carried out recently by Karamanov and Cornuéjols [13].

Here we pursue the viewpoint of reformulating with reduced bases, but from a somewhat different angle. To motivate our approach, consider the set

$$\{\boldsymbol{x} \in \mathbb{Z}^4_+ \mid 1003x_1 - 7004x_2 + 998x_3 - 1999x_4 = 27001\}.$$

Given the coefficients, a natural idea is to separate the multiples of 1000 from the multiples of 1 leading to the equivalent set

$$\begin{array}{rcrcrcrcl} x_1 & - & 7x_2 & + & x_3 & - & 2x_4 & = & 27 + \mu \\ 3x_1 & - & 4x_2 & - & 2x_3 & + & x_4 & = & 1 - 1000\mu \end{array}$$

$$\boldsymbol{x} \in \mathbb{Z}^4_+,\ \mu \in \mathbb{Z}^1$$

Here the new integer variable μ is a natural candidate for branching as any nonzero value of μ immediately forces some of the $\boldsymbol{x}$-variables to take large values.

This suggests three questions:

1. Given the set $X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}^0\}$, how can we find reformulations of the form

$$\{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x}^0 + \boldsymbol{T}\boldsymbol{\mu},\ \boldsymbol{\mu} \in \mathbb{Z}^s\}$$

involving just a small number s of new variables?
2. Can we detect when the matrix $\boldsymbol{A}$ has "special" structure as in the example above, so that we can produce a reformulation in which the coefficients of $\boldsymbol{P}$ are small relative to those of $\boldsymbol{T}$ and one can see directly that branching on the $\boldsymbol{\mu}$-variables has a significant impact?
3. Can we measure this impact explicitly?

We now describe the contents of this paper. In the rest of this section, we present some links between lattices in $\mathbb{Z}^n$ and the feasible regions of integer programs.

In Sect. 2 we show that questions 1 and 2 have simple answers, and that the existence of special structure corresponds almost exactly to the structure observed by Aardal and Lenstra [3], i.e., $\boldsymbol{Q} = (\boldsymbol{R}, \boldsymbol{S})$ with $\boldsymbol{R}$ consisting of short vectors and $\boldsymbol{S}$ of long vectors. One consequence is that for knapsack constraints when $m = 1$ we can detect when $\boldsymbol{a}$ can be written in the form $\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2$ with $\boldsymbol{p}^1$, $\boldsymbol{p}^2$ short relative to $\boldsymbol{a}$.

In Sect. 3 we consider this special case with $m = 1$ explicitly. We reformulate it with two constraints and one additional variable μ, and we then calculate the width of the underlying polyhedron in the direction of μ. This allows us to simplify and generalize a result of Aardal and Lenstra and leads also to lower bounds on the Frobenius number of $\boldsymbol{a}$.

For lattice reformulations to be useful as a tool in preprocessing, it is an open question whether the reformulations should only involve a small number of new variables. In Sect. 4 we carry out some preliminary tests to see whether reformulations with a small number of new variables, as opposed to the $n-m$ appearing in the reformulations of Aardal et al. [2], can have a significant effect in reducing the size of enumeration trees.

1.1 Preliminaries

The set of nonnegative (positive) integers in $\mathbb{R}^n$ is denoted by $\mathbb{Z}^n_+$ ($\mathbb{Z}^n_{>0}$). Similar notation is used for the nonnegative real numbers.

The *Hermite Normal Form* of a matrix $\boldsymbol{A} \in \mathbb{Z}^{m\times n}$ of full row rank, HNF($\boldsymbol{A}$), is obtained by multiplying $\boldsymbol{A}$ by an $n \times n$ unimodular matrix $\boldsymbol{U}$ to obtain the form $(\boldsymbol{D},\ \boldsymbol{0})$, where $\boldsymbol{D} \in \mathbb{Z}^{m\times m}$ is a nonsingular, nonnegative lower triangular matrix with the unique row maximum along the diagonal. Note that when $\boldsymbol{D} \neq \boldsymbol{I}$, the matrix $\boldsymbol{D}^{-1}\boldsymbol{A}$ is integral and HNF($\boldsymbol{D}^{-1}\boldsymbol{A}$) = $(\boldsymbol{I},\ \boldsymbol{0})$.

Let $\boldsymbol{b}^1, \ldots, \boldsymbol{b}^l$ be linearly independent vectors in $\mathbb{R}^n$. The set

$$L(\boldsymbol{B}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \boldsymbol{B}\boldsymbol{\lambda},\ \boldsymbol{\lambda} \in \mathbb{Z}^l\}$$

is called a *lattice*. The set of vectors $\{\boldsymbol{b}^1, \ldots, \boldsymbol{b}^l\}$ is called a *lattice basis*. $\boldsymbol{B}'$ is an alternative basis of $L(\boldsymbol{B})$ if and only if $\boldsymbol{B}' = \boldsymbol{BU}$ where $\boldsymbol{U}$ is an $l \times l$ unimodular matrix. The *rank* of a given lattice L, rk L, is equal to the dimension of the Euclidean vector space generated by a basis of L.

From now on, we assume that $\boldsymbol{B}$ is an integer matrix. If the lattice contains every integer point in the subspace generated by $\boldsymbol{B}$, i.e., $L(\boldsymbol{B}) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \boldsymbol{Bz},\ \boldsymbol{z} \in \mathbb{R}^l\} \cap \mathbb{Z}^n$, then the lattice is called a *pure sublattice of* $\mathbb{Z}^n$.

The following observations all have a short proof, or follow immediately from the above definitions.

Observation 1 *$L(\boldsymbol{B})$ is a pure sublattice of $\mathbb{Z}^n$ if and only if $HNF(\boldsymbol{B}^T) = (\boldsymbol{I}, \boldsymbol{0})$.*

Suppose $\boldsymbol{A}$ is an $m \times n$ integer matrix of full row rank. We use the notation $\ker_{\mathbb{Z}} \boldsymbol{A}$ to denote the lattice $\{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{Ax} = \boldsymbol{0}\}$. If $\boldsymbol{Q} \in \mathbb{Z}^{n \times (n-m)}$ is a basis of $\ker_{\mathbb{Z}} \boldsymbol{A}$, then we can write

$$\ker_{\mathbb{Z}} \boldsymbol{A} = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{Ax} = \boldsymbol{0}\} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \boldsymbol{Q\lambda},\ \boldsymbol{\lambda} \in \mathbb{Z}^{n-m}\} = L(\boldsymbol{Q}).$$

Observation 2 *$ker_{\mathbb{Z}} \boldsymbol{A}$ is a pure sublattice of $\mathbb{Z}^n$ and hence $HNF(\boldsymbol{Q}^T) = (\boldsymbol{I}, \boldsymbol{0})$.*

Now we consider when one can go from the basis formulation $L(\boldsymbol{Q})$ to an IP formulation $\ker_{\mathbb{Z}} \boldsymbol{A}$.

Observation 3 *If $\boldsymbol{Q} \in \mathbb{Z}^{n \times r}$ with $rk(\boldsymbol{Q}) = r$, there exists an integer $(n - r) \times n$ matrix $\boldsymbol{A}$ such that $L(\boldsymbol{Q}) = ker_{\mathbb{Z}} \boldsymbol{A}$ if and only if $HNF(\boldsymbol{Q}^T) = (\boldsymbol{I}, \boldsymbol{0})$ ($L(\boldsymbol{Q})$ is a pure sublattice of $\mathbb{Z}^n$). In addition, there exists an $\boldsymbol{A}$ with $HNF(\boldsymbol{A}) = (\boldsymbol{I}, \boldsymbol{0})$.*

Observation 4 *If $L(\boldsymbol{Q}) = ker_{\mathbb{Z}} \boldsymbol{A}$ with $HNF(\boldsymbol{A}) = (\boldsymbol{I}, \boldsymbol{0})$ and $HNF(\boldsymbol{Q}^T) = (\boldsymbol{I}, \boldsymbol{0})$, then $L(\boldsymbol{A}^T) = ker_{\mathbb{Z}} \boldsymbol{Q}^T$.*

Observation 5 *If $\boldsymbol{Q}$ is a lattice basis, $HNF(\boldsymbol{Q}^T) = (\boldsymbol{I}, \boldsymbol{0})$ and $\boldsymbol{Q} = (\boldsymbol{R}, \boldsymbol{S})$, then $HNF(\boldsymbol{R}^T) = (\boldsymbol{I}, \boldsymbol{0})$.*

A *reduced basis* $\boldsymbol{B}'$ of $L(\boldsymbol{B})$ is a lattice basis in which the columns are relatively short and orthogonal. See Lenstra, Lenstra, and Lovász [14] for precise definitions. For a more detailed exposition of lattices and the importance of reduced bases, see for instance Cassels [8], Lovász [17], Kannan [11] and Lenstra [15, 16].

An algorithm [14], known as the LLL algorithm, calculates a reduced basis in polynomial time. Using this algorithm it is also possible to find a reduced basis $\boldsymbol{Q}$ of $\ker_{\mathbb{Z}} \boldsymbol{A}$ in polynomial time, see [2]. The Hermite Normal Form can also be calculated in polynomial time using for instance the LLL algorithm [21], or the algorithm of Kannan and Bachem [12].

2 Alternative formulations and hidden structure

The basic set that we consider is $X = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{Ax} = \boldsymbol{b}\}$ where $\boldsymbol{A} \in \mathbb{Z}^{m \times n}$ with full row rank, and $\boldsymbol{b} \in \mathbb{Z}^m$. We assume that $X \neq \emptyset$. For any vector $\boldsymbol{x}^0 \in X$ we can write $\boldsymbol{b} = \boldsymbol{Ax}^0$ and

$$X(\boldsymbol{x}^0) = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{0}\} = \{\boldsymbol{x}^0\} + \ker_{\mathbb{Z}} \boldsymbol{A}.$$

Later, the reformulations of $\ker_{\mathbb{Z}} A$ that we obtain below will be used in testing feasibility or optimizing over the set

$$X_+(x^0) = X(x^0) \cap \{x \mid x \geq 0\}. \tag{1}$$

Since A has full row rank, rk $\ker_{\mathbb{Z}} A = n - m$. In Sect. 2.1 we derive a family of reformulations of $\ker_{\mathbb{Z}} A$ of the form $\{x \in \mathbb{Z}^n \mid Px = T\mu,\ \mu \in \mathbb{Z}^s\}$, where $P \in \mathbb{Z}^{(m+s)\times n}$, $T \in \mathbb{Z}^{(m+s)\times s}$ for $0 \leq s \leq n - m$, and in Sect. 2.2 we give a polynomial time algorithm to compute P and T for any value of s. Notice that $s = 0$ corresponds to the case $Ax = 0$, and $s = n - m$ to the case $Ix = Q\mu$ where Q is a basis for $\ker_{\mathbb{Z}} A$.

2.1 A family of extended formulations

Given A and Q, a basis of $\ker_{\mathbb{Z}} A$, our goal in this subsection is to derive alternatives to the reformulation $\{x \in \mathbb{R}^n \mid x = Q\lambda,\ \lambda \in \mathbb{Z}^{n-m}\}$ of $\ker_{\mathbb{Z}} A$.

Theorem 1 *Let Q be a basis of $ker_{\mathbb{Z}} A$ and (R, S) a partition of Q with $R \in \mathbb{Z}^{n\times r}$, $S \in \mathbb{Z}^{n\times s}$ and $r + s = n - m$. Moreover, let $P^T \in \mathbb{Z}^{n\times(m+s)}$ be a basis for the lattice $ker_{\mathbb{Z}} R^T$. Then,*

$$ker_{\mathbb{Z}} A = \{x \in \mathbb{Z}^n \mid Px = PS\mu,\ \mu \in \mathbb{Z}^s\}.$$

Proof From Observation 2 it follows that HNF$(P) = (I, 0)$ and from Observations 2 and 5 that HNF$(R^T) = (I, 0)$. As $\ker_{\mathbb{Z}} R^T = L(P^T)$, we obtain from Observation 4 that $\ker_{\mathbb{Z}} P = L(R)$.

Now $\{x \in \mathbb{Z}^n \mid Px = PS\mu,\ \mu \in \mathbb{Z}^s\}$

$= \{x \in \mathbb{Z}^n \mid P(x - S\mu) = 0,\ \mu \in \mathbb{Z}^s\}$

$= \{x \in \mathbb{Z}^n \mid x - S\mu = R\lambda,\ \mu \in \mathbb{Z}^s, \lambda \in \mathbb{Z}^{n-m-s}\}$

$= \{x \in \mathbb{Z}^n \mid x = R\lambda + S\mu,\ \mu \in \mathbb{Z}^s, \lambda \in \mathbb{Z}^{n-m-s}\}$

$= \ker_{\mathbb{Z}} A.$ □

We now show that if we have a reformulation of $\ker_{\mathbb{Z}} A$ of the form $\{x \in \mathbb{Z}^n \mid Px = T\mu,\ \mu \in \mathbb{Z}^s\}$ with P, T integer matrices and with HNP$(P) = (I, 0)$, it is necessarily of the above form.

Proposition 1 *If $ker_{\mathbb{Z}} A = \{x \in \mathbb{Z}^n \mid Px = Tu,\ u \in \mathbb{Z}^t\}$ with $P \in \mathbb{Z}^{(m+s)\times n}$, $T \in \mathbb{Z}^{(m+s)\times t}$, $rk(P) = m + s$, $rk(T) = t$, and $HNF(P) = (I, 0)$, then $s = t$ and P is generated as in Theorem* 1.

Proof Let $U \in \mathbb{Z}^{n\times n}$ be the unimodular matrix used to bring P in Hermite Normal Form, i.e., $P(U_1, U_2) = (I, 0)$. Then

$$\{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{P}\boldsymbol{x} = \boldsymbol{T}\boldsymbol{u},\ \boldsymbol{u} \in \mathbb{Z}^t\}$$
$$= \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{x} = \boldsymbol{U}\boldsymbol{w},\ (\boldsymbol{I}, \boldsymbol{0})\begin{pmatrix}\boldsymbol{w}^1\\ \boldsymbol{w}^2\end{pmatrix} = \boldsymbol{T}\boldsymbol{u},\ \boldsymbol{u} \in \mathbb{Z}^t,\ \boldsymbol{w}^1 \in \mathbb{Z}^{m+s},\ \boldsymbol{w}^2 \in \mathbb{Z}^{n-m-s}\}$$
$$= \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{x} = \boldsymbol{U}^1\boldsymbol{T}\boldsymbol{u} + \boldsymbol{U}^2\boldsymbol{w}^2,\ \boldsymbol{u} \in \mathbb{Z}^t,\ \boldsymbol{w}^2 \in \mathbb{Z}^{n-m-s}\}.$$

Now $\boldsymbol{Q} = (\boldsymbol{U}^2, \boldsymbol{U}^1\boldsymbol{T})$ must be a basis of $\ker_{\mathbb{Z}}\boldsymbol{A}$ and thus $s = t$. Taking $\boldsymbol{R} = \boldsymbol{U}^2$ and $\boldsymbol{S} = \boldsymbol{U}^1\boldsymbol{T}$, $\boldsymbol{P}\boldsymbol{U}^2 = 0$ and $\boldsymbol{P}\boldsymbol{S} = \boldsymbol{P}\boldsymbol{U}^1\boldsymbol{T} = \boldsymbol{T}$ and the claim follows. □

2.2 Using reduced bases to find structure

Here we present an algorithm for determining a matrix $\boldsymbol{P}$ as described in the previous subsection for varying values of s. Notice that since $\boldsymbol{A} \in \ker_{\mathbb{Z}}\boldsymbol{Q}^T$, then $\boldsymbol{A} \in \ker_{\mathbb{Z}}\boldsymbol{R}^T$. Then, as $\ker_{\mathbb{Z}}\boldsymbol{R}^T = L(\boldsymbol{P}^T)$, we can write $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{P}$ for some $m \times (m+s)$ integer matrix $\boldsymbol{M}$.

(i) Find a reduced basis $\boldsymbol{Q}$ of $\ker_{\mathbb{Z}}\boldsymbol{A}$.
(ii) Suppose $\boldsymbol{Q}$ consists of s long vectors and $r = n-m-s$ short ones. How to define "long" and "short" is up to the user. (If all vectors of $\boldsymbol{Q}$ are of approximately the same length we set $s = n - m$.) We define $\boldsymbol{R}$ to be the set of short vectors of $\boldsymbol{Q}$ and $\boldsymbol{S}$ to be the set of long ones.
(iii) Find a (reduced) basis $\boldsymbol{P}^T$ of $\ker_{\mathbb{Z}}\boldsymbol{R}^T$.
(iv) Solve the system of equations $\boldsymbol{M}\boldsymbol{P} = \boldsymbol{A}$, $\boldsymbol{M} \in \mathbb{Z}^{m\times(m+s)}$ to find the matrix of multipliers $\boldsymbol{M}$ if an explicit decomposition of $\boldsymbol{A}$ is desired.

We can establish the following relationship between $\boldsymbol{M}$ and the matrix $\boldsymbol{P}\boldsymbol{S}$.

Proposition 2 *Given $\boldsymbol{A}$, suppose $\boldsymbol{P}$ is obtained as in Theorem 1 and $\boldsymbol{M}$ as in step iv) of the above algorithm. Then $\boldsymbol{P}\boldsymbol{S}$ is a basis of $\ker_{\mathbb{Z}}\boldsymbol{M}$.*

Proof Let $\boldsymbol{T} = \boldsymbol{P}\boldsymbol{S}$. We have $\boldsymbol{0} = \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{M}\boldsymbol{P}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{M}\boldsymbol{T}\boldsymbol{\mu}$ for $\boldsymbol{\mu} \in \mathbb{Z}^s$. Hence, the columns of $\boldsymbol{T}$ lie in $\ker_{\mathbb{Z}}\boldsymbol{M}$. Suppose that they do not form a lattice basis. Then there exists an element $\boldsymbol{t}^* \in \ker_{\mathbb{Z}}\boldsymbol{M}$ that is not in $L(\boldsymbol{T})$. By Observation 2 we have $\mathrm{HNF}(\boldsymbol{P}) = (\boldsymbol{I}, \boldsymbol{0})$, and hence there exists a vector $\boldsymbol{x}^*$ such that $\boldsymbol{P}\boldsymbol{x}^* = \boldsymbol{t}^*$. Now the vector $\boldsymbol{y}^* = \boldsymbol{x}^0 + \boldsymbol{x}^*$ lies in $X(\boldsymbol{x}^0)$ as $\boldsymbol{A}\boldsymbol{y}^* - \boldsymbol{A}\boldsymbol{x}^0 = \boldsymbol{A}\boldsymbol{x}^* = \boldsymbol{M}\boldsymbol{P}\boldsymbol{x}^* = \boldsymbol{M}\boldsymbol{t}^* = \boldsymbol{0}$. Moreover, $\boldsymbol{P}(\boldsymbol{y}^* - \boldsymbol{x}^0) = \boldsymbol{P}\boldsymbol{x}^* = \boldsymbol{t}^*$. By the assumption that $\boldsymbol{t}^* \notin L(\boldsymbol{T})$ we have $\boldsymbol{P}(\boldsymbol{y}^* - \boldsymbol{x}^0) \neq \boldsymbol{T}\boldsymbol{\mu}$ for any $\boldsymbol{\mu} \in \mathbb{Z}^s$, contradicting Theorem 1. □

The choice of a reduced basis $\boldsymbol{Q}$ at the beginning of our algorithm, and the choice of $\boldsymbol{R}$ as the set of "short" basis vectors of $\boldsymbol{Q}$ determines the set of interesting branching variables $\boldsymbol{\mu}$. Then any basis $\boldsymbol{P}^T$ of $\ker_{\mathbb{Z}}\boldsymbol{R}^T$ will lead to a reformulation with $\boldsymbol{\mu}$ as additional variables. However if $\boldsymbol{R}$ consists of short vectors, then there exists a basis $\boldsymbol{P}^T$ consisting of relatively short vectors. Therefore choosing a reduced basis $\boldsymbol{P}^T$ brings out the real structure of $\boldsymbol{A}$, namely $\boldsymbol{A} = \boldsymbol{M}\boldsymbol{P}$ with $\boldsymbol{P}$ having small, and $\boldsymbol{M}$ potentially large elements. This is illustrated in the following example.

Example 1 We consider a matrix $\boldsymbol{A}$ consisting of one row $\boldsymbol{a}$ only:

$$\boldsymbol{a} = (12223\ 12224\ 36674\ 61119\ 85569).$$

and $b = 89643481$. This is instance `cuww1` from Cornuéjols et al. [10]. Here we derive a reformulation using $m+s = 2$ constraints and $n+1 = 6$ variables, and show how its hidden structure can be uncovered.

The vector $(\boldsymbol{x}^0)^T = (-635, 30, 1428, -511, 887)$ satisfies $\boldsymbol{a}\boldsymbol{x}^0 = b$. A reduced basis of $\ker_{\mathbb{Z}}\boldsymbol{a}$ is equal to

$$\boldsymbol{Q} = \begin{pmatrix} 0 & -3 & -1 & 2059 \\ 1 & 1 & -3 & 157 \\ -1 & -1 & -1 & -3336 \\ -1 & 1 & 0 & 2687 \\ 1 & 0 & 1 & -806 \end{pmatrix}.$$

Here we observe that the last column of the reduced basis $\boldsymbol{Q}$ is much longer than the other columns. Taking $r = 3$ and $s = 1$, $\boldsymbol{R}$ will consist of the first three columns of $\boldsymbol{Q}$, and $\boldsymbol{S}$ will consist of the last column of $\boldsymbol{Q}$. A reduced basis $\boldsymbol{P}^T$ the of $\ker_{\mathbb{Z}}\boldsymbol{R}^T$ is

$$\boldsymbol{P}^T = \begin{pmatrix} -1 & 2 \\ 0 & 1 \\ 2 & 1 \\ -1 & 6 \\ 1 & 6 \end{pmatrix}.$$

A reformulation of $X_+(\boldsymbol{x}^0)$ is given by the set of nonnegative integer $\boldsymbol{x}$ satisfying

$$\begin{aligned} -x_1 \quad\quad + \quad 2x_3 \quad - \quad x_4 \quad + \quad x_5 \quad &= \quad 4889 \quad - \quad 12224\mu \\ 2x_1 \quad + \quad x_2 \quad + \quad x_3 \quad + \quad 6x_4 \quad + \quad 6x_5 \quad &= \quad 2444 \quad + \quad 12225\mu \end{aligned}$$

for some integer μ.

Notice also that the vector $(M_1, M_2) = (12225, 12224)$ solves $\boldsymbol{M}\boldsymbol{P} = \boldsymbol{a}$, so we can write $\boldsymbol{a} = 12225\boldsymbol{p}^1 + 12224\boldsymbol{p}^2$, with $\boldsymbol{p}^1$ being the first row of $\boldsymbol{P}$, and $\boldsymbol{p}^2$ being the second row of $\boldsymbol{P}$. □

3 Knapsack sets replaced by two equations

Recall the set (1)

$$X_+(\boldsymbol{x}^0) = \{\boldsymbol{x} \in \mathbb{Z}^n \mid \boldsymbol{A}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{0}\} \cap \{\boldsymbol{x} \mid \boldsymbol{x} \geq \boldsymbol{0}\},$$

where $\boldsymbol{x}^0$ is an integer vector satisfying $\boldsymbol{A}\boldsymbol{x}^0 = \boldsymbol{b}$ for given $\boldsymbol{b} \in \mathbb{Z}^m$. Using our reformulation of $\ker_{\mathbb{Z}}\boldsymbol{A}$ for a given choice of $\boldsymbol{P}$ we obtain the following formulation of the set (1):

$$X_+(\boldsymbol{x}^0, \boldsymbol{P}) = \{\boldsymbol{x} \in \mathbb{Z}^n_+ \mid \boldsymbol{P}(\boldsymbol{x} - \boldsymbol{x}^0) = \boldsymbol{P}\boldsymbol{S}\boldsymbol{\mu},\ \boldsymbol{\mu} \in \mathbb{Z}^s\}.$$

Here we analyze the case $m = 1$ and $s = 1$ in more detail. We suppose that $a_j > 0$ for all $1 \leq j \leq n$ and that $\gcd(a_1, \ldots, a_n) = 1$. We generate $\boldsymbol{P} \in \mathbb{Z}^{2\times n}$ and $\boldsymbol{M} = (M_1, M_2)$ using the algorithm in Sect. 2.2, with a possible adjustment of

the signs so that $M_1, M_2 > 0$. Notice that HNF$(\boldsymbol{P}) = (\boldsymbol{I}, \boldsymbol{0})$ from Observation 2, and that $\gcd(M_1, M_2) = 1$, which implies the existence of integers $\boldsymbol{q} \in \mathbb{Z}^2$ with $M_1 q_1 + M_2 q_2 = 1$. Suppose $\boldsymbol{q}' \in \mathbb{Z}^2$ satisfies $M_1 q_1' + M_2 q_2' = 1$. The set of all valid $(q_1, q_2)^T$ can be written as

$$\begin{pmatrix} q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} q_1' \\ q_2' \end{pmatrix} + \lambda \begin{pmatrix} -M_2 \\ M_1 \end{pmatrix}. \tag{2}$$

Now we can derive specific values for $\boldsymbol{p}^i \boldsymbol{x}^0$, $\boldsymbol{p}^i \boldsymbol{S}$, $i = 1, 2$. Choose $\boldsymbol{q} \in \mathbb{Z}^2$ such that $M_1 q_1 + M_2 q_2 = 1$. Because $\boldsymbol{a}\boldsymbol{x}^0 - \boldsymbol{b} = M_1(\boldsymbol{p}^1 \boldsymbol{x}^0 - q_1 b) + M_2(\boldsymbol{p}^2 \boldsymbol{x}^0 - q_2 b) = 0$ we can take $\boldsymbol{p}^i \boldsymbol{x}^0 = q_i b$ for $i = 1, 2$. By Proposition 2, $\boldsymbol{P}\boldsymbol{S}$ forms a basis of $\ker_{\mathbb{Z}} \boldsymbol{M}$. Hence, we can set $\boldsymbol{p}^1 \boldsymbol{S} = M_2$ and $\boldsymbol{p}^2 \boldsymbol{S} = -M_1$.

Using these specific values of $\boldsymbol{p}^i \boldsymbol{x}^0$, $\boldsymbol{p}^i \boldsymbol{S}$, $i = 1, 2$, and rearranging the terms, yields the following reformulation of $X_+(\boldsymbol{x}^0)$.

$$X_+(\boldsymbol{x}^0, \boldsymbol{P}) = \{\boldsymbol{x} \in \mathbb{Z}_+^n, \mu \in \mathbb{Z}^1 \mid \tag{3}$$

$$\boldsymbol{p}^1 \boldsymbol{x} - M_2 \mu = q_1 b \tag{4}$$

$$\boldsymbol{p}^2 \boldsymbol{x} + M_1 \mu = q_2 b\}. \tag{5}$$

The linear relaxation of $X_+(\boldsymbol{x}^0, \boldsymbol{P})$ is denoted by $Y_+(\boldsymbol{x}^0, \boldsymbol{P})$.

We derive two results. The first concerns the width of the polyhedron $Y_+(\boldsymbol{x}^0, \boldsymbol{P})$ in the direction μ. The second uses the width to derive a lower bound on the Frobenius number, simplifying and generalizing a result of Aardal and Lenstra [3] that is valid under the assumptions that $\boldsymbol{p}^1 \in \mathbb{Z}_{>0}$ and $M_2 = 1$.

3.1 The integer width

The integer width of a rational polytope P in the integer direction $\boldsymbol{d}$, $w_I(P, \boldsymbol{d})$, is defined as

$$w_I(P, \boldsymbol{d}) = \left\lfloor \max\{\boldsymbol{d}^T \boldsymbol{x} \mid \boldsymbol{x} \in P\} \right\rfloor - \left\lceil \min\{\boldsymbol{d}^T \boldsymbol{x} \mid \boldsymbol{x} \in P\} \right\rceil + 1,$$

and is equal to the number of parallel lattice hyperplanes in direction $\boldsymbol{d}$ that are intersecting P.

Our goal is to calculate the integer width of

$$Y_+(\boldsymbol{x}^0, \boldsymbol{P}) = \{(\boldsymbol{x}, \mu) \in \mathbb{R}_+^n \times \mathbb{R} \mid \boldsymbol{p}^1 \boldsymbol{x} - M_2 \mu = q_1 b,\ \boldsymbol{p}^2 \boldsymbol{x} + M_1 \mu = q_2 b\}$$

in the direction of the variable μ. To do this, let

$$\bar{Y}_+(\boldsymbol{x}^0, \boldsymbol{P}) = \{(\boldsymbol{x}, \mu) \in \mathbb{R}_+^n \times \mathbb{R} \mid \boldsymbol{p}^1 \boldsymbol{x} - M_2 \mu = q_1,\ \boldsymbol{p}^2 \boldsymbol{x} + M_1 \mu = q_2\} \tag{6}$$

be the scaled down version of this polyhedron with $b = 1$.

Below we derive the values $\bar{z} = \max\{\mu \mid (\boldsymbol{x}, \mu) \in \bar{Y}_+(\boldsymbol{x}^0, \boldsymbol{P})\}$ and $\underline{z} = \min\{\mu \mid (\boldsymbol{x}, \mu) \in \bar{Y}_+(\boldsymbol{x}^0, \boldsymbol{P})\}$. Once we have the values $\bar{z}$ and $\underline{z}$ it will be straightforward to compute the width $w_I(Y_+(\boldsymbol{x}^0, \boldsymbol{P}), \boldsymbol{e}^{n+1})$ as $\lfloor b\bar{z} \rfloor - \lceil b\underline{z} \rceil + 1$.

Lemma 1 *Given the polyhedron* $\bar{Y}_+(\boldsymbol{x}^0, \boldsymbol{P})$, *let*

$$j = \arg\min\{i \mid p_i^1 / a_i\}$$
$$k = \arg\max\{i \mid p_i^1 / a_i\}.$$

Then,

$$\bar{z} = \frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2} \text{ and } \underline{z} = \frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2}.$$

Proof Consider the linear program $\bar{z} = \max\{\mu \mid (\boldsymbol{x}, \mu) \in \bar{Y}_+(\boldsymbol{x}^0, \boldsymbol{P})\}$, and let $\boldsymbol{\gamma}$ be the corresponding LP-dual variables.

The dual problem is:

$$\bar{z} = \min \quad q_1\gamma_1 + q_2\gamma_2 \tag{7}$$
$$\text{s.t.} \quad p_i^1\gamma_1 + p_i^2\gamma_2 \geq 0, \quad 1 \leq i \leq n, \tag{8}$$
$$-M_2\gamma_1 + M_1\gamma_2 = 1, \tag{9}$$
$$\boldsymbol{\gamma} \in \mathbb{R}^2.$$

From constraint (9) we obtain $\gamma_1 = \frac{M_1\gamma_2 - 1}{M_2}$. Substituting for γ_1 in the dual objective function (7) gives

$$\bar{z} = \min \frac{\gamma_2}{M_2}(q_1M_1 + q_2M_2) - \frac{q_1}{M_2} = \min \frac{\gamma_2}{M_2} - \frac{q_1}{M_2}, \tag{10}$$

where the last term is a constant. Since $M_2 > 0$ we want to find the minimum value of γ_2. Constraints (8) now yield $p_i^1(\frac{M_1\gamma_2 - 1}{M_2}) + p_i^2\gamma_2 \geq 0, \quad 1 \leq i \leq n$.

Rewriting gives $\gamma_2(p_i^2 + p_i^1(M_1/M_2)) - p_i^1/M_2 = 1/M_2[\gamma_2(p_i^1M_1 + p_i^2M_2) - p_i^1] \geq 0$, or $\gamma_2 \geq \frac{p_i^1}{a_i}$, $1 \leq i \leq n$. From the definition of the index k we obtain $\gamma_2 = \frac{p_k^1}{a_k}$. Finally, we substitute for $\boldsymbol{\gamma}_2$ in the rewritten dual objective function (10) yielding the optimal dual objective value

$$\bar{z} = \frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2}. \tag{11}$$

The calculation of $\underline{z}$ is almost identical. □

Immediately we obtain the integer width.

Theorem 2

$$w_I(Y_+(\boldsymbol{x}^0, \boldsymbol{P}), \boldsymbol{e}^{n+1}) = \left\lfloor \frac{bp_k^1}{M_2 a_k} - \frac{bq_1}{M_2} \right\rfloor - \left\lceil \frac{bp_j^1}{M_2 a_j} - \frac{bq_1}{M_2} \right\rceil + 1,$$

where the indices j and k are defined as in Lemma 1.

Notice that $\bar{z}$ and $\underline{z}$, and hence $w_I(Y_+(\boldsymbol{x}^0, \boldsymbol{P}), \boldsymbol{e}^{n+1})$, can be expressed in several ways by using various equations. An expression for $\bar{z}$ (and similarly for $\underline{z}$) that does not contain M_1 or M_2 is obtained by using $M_1 q_1 + M_2 q_2 = 1$ and $a_k = M_1 p_k^1 + M_2 p_k^2$ in (11). We then obtain $\bar{z} = (p_k^1 q_2 - p_k^2 q_1)/a_k$.

Example 1, *cont.* Consider the instance of Example 1 and its decomposition. Now we apply Theorem 2 to this example. We have $j = 1$, $k = 3$, $q_1 = 1$, $q_2 = -1$. We obtain

$$\begin{aligned} w_I(Y_+(\boldsymbol{x}^0, \boldsymbol{P}), \boldsymbol{e}^{n+1}) &= \left\lfloor \frac{bp_k^1}{M_2 a_k} - \frac{bq_1}{M_2} \right\rfloor - \left\lceil \frac{bp_j^1}{M_2 a_j} - \frac{bq_1}{M_2} \right\rceil + 1 \\ &= \left\lfloor \frac{bp_3^1}{M_2 a_3} - \frac{bq_1}{M_2} \right\rfloor - \left\lceil \frac{bp_1^1}{M_2 a_1} - \frac{bq_1}{M_2} \right\rceil + 1 \\ &= \left\lfloor \frac{89643481}{12224} \left(\frac{2}{36674} - 1 \right) \right\rfloor \\ &\quad - \left\lceil \frac{89643481}{12224} \left(\frac{-1}{12223} - 1 \right) \right\rceil + 1 \\ &= \lfloor -7333.00003 \rfloor - \lceil -7333.9999 \rceil \\ &\quad + 1 = -7334 + 7333 + 1 = 0. \end{aligned}$$

It follows that $X_+(\boldsymbol{x}^0, \boldsymbol{P}) = \emptyset$. Applying branch-and-bound, and branching first on the μ variable, this infeasibility would immediately be apparent. This is not the case using branch-and-bound starting from the original formulation $\{\boldsymbol{x} \in \mathbb{Z}_+^n \mid \boldsymbol{ax} = b\}$. In particular Cplex fails to prove infeasibility within 500 million nodes. □

A natural question is whether the integer width differs if we use a different member of the family of extended formulations. Consider the two sets

$$Y_+(\boldsymbol{x}^0, \boldsymbol{I}) = \{(\boldsymbol{x}, \boldsymbol{\mu}) \in \mathbb{R}_+^n \times \mathbb{R}^{s+r} \mid \boldsymbol{Ix} = \boldsymbol{Ix}^0 + \boldsymbol{S\mu}^S + \boldsymbol{R\mu}^R\}, \text{ and}$$

$$Y_+(\boldsymbol{x}^0, \boldsymbol{P}) = \{(\boldsymbol{x}, \boldsymbol{\mu}) \in \mathbb{R}_+^n \times \mathbb{R}^s \mid \boldsymbol{Px} = \boldsymbol{Px}^0 + \boldsymbol{PS\mu}^S\}.$$

Observation 6

(i) *For $s = 1$, the width of $Y_+(\boldsymbol{x}^0, \boldsymbol{P})$ in direction μ is independent of the choice of $\boldsymbol{x}^0$ (or $\boldsymbol{q}$).*

(ii) *For fixed $\boldsymbol{R}$, the width of $Y_+(\boldsymbol{x}^0, \boldsymbol{P})$ in the direction of any of the $\boldsymbol{\mu}$-variables is independent of which basis $\boldsymbol{P}^T$ of $ker_{\mathbb{Z}} \boldsymbol{R}^T$ is chosen.*

(iii) *$proj_{\boldsymbol{x},\boldsymbol{\mu}^S} Y_+(\boldsymbol{x}^0, \boldsymbol{I}) = Y_+(\boldsymbol{x}^0, \boldsymbol{P})$, and thus the width of the polytopes $Y_+(\boldsymbol{x}^0, \boldsymbol{I})$ and $Y_+(\boldsymbol{x}^0, \boldsymbol{P})$ in any direction $\boldsymbol{d}$ over the $(\boldsymbol{x}, \boldsymbol{\mu}^S)$ variables is identical.*

Note however that the multipliers M_1, M_2 as well as the values q_1, q_2 do change with different choices of $\boldsymbol{P}$.

3.2 A lower bound on the Frobenius number

The Frobenius number of $\boldsymbol{a}$, $F(\boldsymbol{a})$, is the largest integer value of b such that $\boldsymbol{ax} = b$ does not have a nonnegative integer solution. Without loss of generality we choose $\boldsymbol{q}$ such that $|q_1| \leq M_2/2$. So, if $|q_1'| > M_2/2$, we can determine new valid values of q_1, q_2 such that $|q_1'| \leq M_2/2$ by identifying an appropriate value of λ in Expression (2).

Theorem 3 *Let $\boldsymbol{a} = M_1\boldsymbol{p}^1 + M_2\boldsymbol{p}^2$ with $\boldsymbol{a}, M_1, M_2, \boldsymbol{p}^1, \boldsymbol{p}^2$ satisfying the assumptions given in the beginning of Sect.* (3). *Moreover, let $\underline{z}$, $\bar{z}$ and the indices j and k be as defined in Lemma* 1.

If $(-M_2/2) \leq q_1 \leq 0$ and

(1a) $\frac{p_j^1}{a_j} > q_1$

(2a) $\frac{p_k^1}{a_k} < M_2 + q_1$

(3a) $(\frac{1-\bar{z}}{\bar{z}-\underline{z}})\underline{z} \notin \mathbb{Z}$

then

$$F(\boldsymbol{a}) \geq \frac{a_j a_k(M_2 + q_1) - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{M_2}{\frac{p_j^1}{a_j} - q_1}. \tag{12}$$

or if $0 < q_1 \leq M_2/2$ and

1b) $\frac{p_j^1}{a_j} > -M_2 + q_1$

(2b) $\frac{p_k^1}{a_k} < q_1$

(3b) $(\frac{1+\underline{z}}{\bar{z}-\underline{z}})\bar{z} \notin \mathbb{Z}$,

then

$$F(\boldsymbol{a}) \geq \frac{a_j a_k(M_2 - q_1) + p_j^1 a_k}{p_k^1 a_j - p_j^1 a_k} + \frac{M_2}{\frac{p_k^1}{a_k} - q_1}.$$

Proof We have already determined the width of $\bar{Y}_+(\boldsymbol{x}^0, \boldsymbol{P})$ in the direction of μ corresponding to $b = 1$ in the proof of Lemma 1. Specifically we have shown that μ lies in the interval $[I_j, I_k]$, where

$$I_j := \underline{z} = \frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2} \text{ and } I_k := \bar{z} = \frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2},$$

whose width is

$$D := I_k - I_j = \frac{a_j p_k^1 - a_k p_j^1}{M_2 a_j a_k} > 0.$$

Any integer right-hand side value $b = t$ for which the corresponding interval $[tI_j,\ tI_k]$ does not contain an integer is a lower bound on the Frobenius number $F(\boldsymbol{a})$. Below we will show that

$$t \geq \frac{a_j a_k (M_2 + q_1) - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{M_2}{\frac{p_j^1}{a_j} - q_1}$$

is such a lower bound in the case that $-M_2/2 \leq q_1 < 0$. A sketch of the proof for the case $0 < q_1 \leq M_2/2$ is given in Appendix 1.

If $q_1 \leq 0$, Assumptions 1a and 2a imply that $0 < I_j < I_k < 1$. Moreover, since $q_1 \geq -M_2/2$ we obtain $I_k \leq p_k^1/(M_2 a_2) + 1/2$. Let $s := \frac{1-I_k}{D}$. Notice that $1 - I_k > 0$ since $I_k < 1$. The interval $[sI_j,\ sI_k]$ has length $1 - I_k$. Notice that $sI_j \notin \mathbb{Z}$ due to Assumption 3 of the theorem. Define $\ell := \lfloor sI_j \rfloor$ and $s' := \ell / I_j$. The number s' satisfies $s - \frac{1}{I_j} < s' < s$, and yields the interval $[I'_j, I'_k] := [s'I_j,\ s'I_k]$, with I'_j integral. The length of $[I'_j, I'_k]$ is less than the length $1 - I_k$ of $[sI_j,\ sI_k]$. Therefore, $[I'_j, I'_k + I_k]$ has length less than 1, and since I'_j is integral it follows that $(I'_j, I'_k + I_k]$ does not contain an integer.

Now, define $s^* := \lfloor s' \rfloor + 1$ and the interval $[I_j^*, I_k^*] := [s^* I_j,\ s^* I_k]$. We have $I'_j < I_j^* \leq I'_j + I_j$ and $I'_k < I_k^* \leq I'_k + I_k$. The result that $[I_j^*,\ I_k^*]$ does not contain an integer follows from the observation that $(I'_j, I'_k + I_k]$ does not contain an integer.

We finally observe that

$$s^* = \lfloor s' \rfloor + 1 > \lfloor s - \frac{1}{I_j} \rfloor + 1 \geq s - \frac{1}{I_j} - 1 + 1 = s - \frac{1}{I_j},$$

so we can conclude that $s - \frac{1}{I_j} = \frac{1-I_k}{D} - \frac{1}{I_j}$ yields a lower bound on the Frobenius number $F(\boldsymbol{a})$. Rewriting $\frac{1-I_k}{D} - \frac{1}{I_j}$ results in the expression

$$\begin{aligned}\frac{1 - I_k}{D} - \frac{1}{I_j} &= \frac{1 - \frac{p_k^1}{M_2 a_k} + \frac{q_1}{M_2}}{\frac{a_j p_k^1 - a_k p_j^1}{M_2 a_j a_k}} - \frac{1}{\frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2}} \\ &= \frac{a_j a_k (M_2 + q_1) - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{M_2}{\frac{p_j^1}{a_j} - q_1}.\end{aligned}$$

□

We notice the similarity with the expression for the lower bound on the Frobenius number derived by Aardal and Lenstra [3] for the case that $M_2 = 1$ and $\boldsymbol{p}^1 \in \mathbb{Z}^n_{>0}$. If

we set $M_2 = 1$ and $q = 0$ in Expression (12) we obtain

$$\frac{a_j a_k - p_k^1 a_j}{p_k^1 a_j - p_j^1 a_k} - \frac{a_j}{p_j^1}.$$

The only difference in the two expressions is in the numerator of the first term, where we have $p_k^1 a_j$ instead of $2p_j^1 a_k$ in [3]. This is a result of a different choice of the number s in the proof. In [3] s was chosen as $s = (1 - 2I_j)/D$ under a constraint on the relationship between I_j and I_k.

4 Computation: feasibility testing and quality of the Frobenius bound

We tested the quality of the extended formulations $\{\boldsymbol{x} \in \mathbb{Z}_+^n \mid \boldsymbol{P}(\boldsymbol{x}-\boldsymbol{x}^0) = \boldsymbol{P}\boldsymbol{S}\boldsymbol{\mu},\ \boldsymbol{\mu} \in \mathbb{Z}^s\}$ for different choices of s on some instances of integer equality knapsacks and the Cornuéjols-Dawande market split problem. We use our algorithm presented in Sect. 2.2 to determine $\boldsymbol{P}$ and $\boldsymbol{S}$ for each chosen s. In particular, the reduced bases $\boldsymbol{Q}$ and $\boldsymbol{P}$ in Steps i and iii of the algorithm, and the vector $\boldsymbol{x}^0$, are computed using the algorithm of Aardal et al. [2].

The integer knapsack instances were taken from Aardal and Lenstra [3]. Instances prob1–4 are such that the vector $\boldsymbol{a}$ decomposes with short $\boldsymbol{p}^1$, $\boldsymbol{p}^2$, whereas for the instances prob11–14 the $\boldsymbol{a}$-coefficients, randomly generated from $U[10,000, 150,000]$, are of the same size on average as in prob1–4. Instances prob11–14 have no apparent structure, and the columns of a reduced basis $\boldsymbol{Q}$ of $\ker_{\mathbb{Z}} \boldsymbol{A}$ are of approximately the same length. We use the Frobenius number of the vector $\boldsymbol{a}$ as right-hand side coefficient for all knapsack instances. Instances prob1–4 have eight variables and prob10–14 have ten variables. For details of the instances, see [3].

The market split instances [9] are multiple row equality knapsack problems in $\{0, 1\}$-variables with m rows and $n = 10(m-1)$ variables. The elements of $\boldsymbol{a}^i$ for each row i are generated randomly from $U[0, 99]$, and the right-hand side coefficients are calculated as $b_i = \lfloor (\sum_{j=1}^n a_j^i)/2 \rfloor$. We generated two sets of market split instances with 4 constraints and 30 variables, and 5 constraints and 40 variables respectively.

In Tables 1–2 we report on the number of nodes used by the integer programming solver Xpress Version 16.01.01 [22] to solve the various reformulations. Column "orig" refers to the original formulation in $\boldsymbol{x}$-variables. Column "AHL" refers to the Aardal-Hurkens-Lenstra lattice reformulation in which the $\boldsymbol{x}$-variables have been removed from the formulation, i.e., the formulation $\{\boldsymbol{\mu} \in \mathbb{Z}^{n-1} \mid \boldsymbol{Q}\boldsymbol{\mu} \geq -\boldsymbol{x}^0\}$ in the knapsack case and the formulation $\{\boldsymbol{\mu} \in \mathbb{Z}^{n-m} \mid -\boldsymbol{x}^0 \leq \boldsymbol{Q}\boldsymbol{\mu} \leq 1 - \boldsymbol{x}^0\}$ in the market split case. For formulations $X_+(\boldsymbol{x}^0, \boldsymbol{P})$ we report on results for different values of s. Notice that the formulations AHL and $X_+(\boldsymbol{x}^0, \boldsymbol{P})$ for $s = n-m$ are mathematically equivalent, but the $X_+(\boldsymbol{x}^0, \boldsymbol{P})$-formulations contain the $\boldsymbol{x}$-variables with the identity matrix as coefficients. Since the solver reacts differently to the presence of the redundant $\boldsymbol{x}$-variables, this leads to slight deviations in the number of enumeration nodes needed.

Instances prob1–4, which decompose in short $\boldsymbol{p}^1$, $\boldsymbol{p}^2$, are very difficult to tackle with branch-and-bound applied to the original formulation. The Frobenius numbers for these instances are also large, see Table 3. None of the instances could be solved

Table 1 The number of branch-and-bound nodes: knapsack instances

Instance	Orig	AHL	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$	$s=7$	$s=9$
prob1	$>10^8$	1	59	15	3	3	1	1	1	–
prob2	$>10^8$	3	23	7	3	1	1	1	1	–
prob3	$>10^8$	13	37	29	5	7	11	9	5	–
prob4	$>10^8$	3	13	5	1	1	1	1	1	–
prob11	100,943	61	2,237	7,683	317	89	51	69	49	61
prob12	160,783	93	10,981	1,105	967	523	179	105	117	71
prob13	188,595	91	10,205	12,261	239	321	35	57	39	59
prob14	140,301	87	2,443	627	689	389	283	115	105	87

Table 2 The number of branch-and-bound nodes: CD-instances 4×30 and 5×40

Instance	Orig	AHL	$s=1$	$s=5$	$s=10$	$s=15$	$s=20$	$s=26$
$4 \times 30\_1$	157,569	281	124,695	71,641	8,033	1,397	1,021	607
$4 \times 30\_2$	169,455	167	154,505	51,989	3,794	1,487	610	535
$4 \times 30\_3$	209,741	325	178,697	181,373	32,367	1,831	1,025	845
$4 \times 30\_4$	202,513	199	156,047	4,685	3,583	829	493	9,527
$4 \times 30\_5$	115,173	311	73,151	17,201	1,197	391	353	3,135

Instance	orig	AHL	$s=5$	$s=10$	$s=20$	$s=30$	$s=35$
$5 \times 40\_1$	$>10^7$	5,873	$>10^7$	3,144,737	160,701	32,507	32,099
$5 \times 40\_2$	$>10^7$	1,643	$>10^7$	2,821,042	128,707	30,302	12,734
$5 \times 40\_3$	$>10^7$	7,349	$>10^7$	8,264,955	86,483	28,491	25,541
$5 \times 40\_4$	$>10^7$	6,870	$>10^7$	1,854,280	70,949	19,616	16,557
$5 \times 40\_5$	$>10^7$	6,651	$>10^7$	7,805,023	1,107,713	35,989	36,897

Table 3 The value of the lower bound of the Frobenius number.

Instance	$F(a)$	lower bound on $F(a)$
prob1	33,367,335	26,061,675
prob2	14,215,206	10,894,273
prob3	58,424,799	31,510,625
prob4	60,575,665	56,668,034
prob11	577,134	98,774
prob12	944,183	113,114
prob13	765,260	67,752
prob14	680,230	60,476

within 100 million nodes. As could be expected, the $X_+(\boldsymbol{x}^0, \boldsymbol{P})$-formulation with $s = 1$, which is a formulation with the $\boldsymbol{x}$-variables and one variable μ, is easy to solve and comparable to the AHL-formulation. In contrast, instances prob11–14 are

solvable using the original formulation, mainly due to the smaller value of the right-hand side coefficients. Here, one could expect that we would need to set $s = n - m$ to see a noticeable improvement compared to the original formulation, but in fact even taking $s = 1$ reduces the number of enumeration nodes by at least an order of magnitude, and with s around 5 we obtain results comparable to those obtained with the AHL-formulation.

For the market split instances, which have no clear structure of the $\boldsymbol{Q}$-matrix, we notice similar results to those obtained for the knapsack instances prob11–14. The algorithm of Sect. 2.2 prescribes $s = n - m$ for these types of instances. The computational results suggest that smaller values of s already yield significant computational improvement.

In Table 3 we report on the value of the Frobenius number as well as the value produced by the lower bound given in Theorem 3. For instances prob1–4, the lower bound is of the same order of magnitude as the Frobenius number, whereas for instances prob11–14 the bound is off by an order of magnitude. The bound might be improved by a different choice of the value s in the proof of the theorem.

Appendix 1

Proof of Theorem 6 for the case $0 < q_1 \leq (M_2/2)$. If $0 < q_1 \leq (M_2/2)$, Assumptions 1b and 2b imply that $-1 < I_j < I_k < 0$, so the interval $[I_j,\ I_k]$ does not contain an integer. In addition, $I_j \geq p_j^1/(M_2 a_j) - 1/2$.

Let $s := \frac{1+I_j}{D}$. The length of the interval $[sI_j,\ sI_k]$ is equal to $1 + I_j$, and since $-1 < I_j < 0$ we have that $0 < 1 + I_j < 1$.

Notice that $sI_k \notin \mathbb{Z}$ due to Assumption 3b of the theorem. Define $\ell := \lceil sI_k \rceil$ and $s' := \ell/I_k$. The number s' satisfies $s + \frac{1}{I_k} < s' < s$, and yields the interval $[I'_j, I'_k] := [s'I_j,\ s'I_k]$, with I'_k integral. The length of $[I'_j, I'_k]$ is less than the length $1 + I_j$ of $[sI_j,\ sI_k]$. Therefore, $[I'_j + I_j, I'_k]$ has length less than 1, and since I'_k is integral it follows that $[I'_j + I_j, I'_k)$ does not contain an integer.

Now, define $s^* := \lfloor s' \rfloor + 1$ and the interval $[I^*_j, I^*_k] := [s^*I_j,\ s^*I_k]$. We have $I'_j + I_j \leq I^*_j < I'_j$ and $I'_k + I_k \leq I^*_k < I'_k$. The result that $[I^*_j,\ I^*_k]$ does not contain an integer follows from the observation that $[I'_j + I_j, I'_k)$ does not contain an integer.

We finally observe that

$$s^* = \lfloor s' \rfloor + 1 > \lfloor s + \frac{1}{I_k} \rfloor + 1 \geq s + \frac{1}{I_k} - 1 + 1 = s + \frac{1}{I_k},$$

so we can conclude that $s + \frac{1}{I_k} = \frac{1+I_j}{D} + \frac{1}{I_k}$ yields a lower bound on the Frobenius number $F(\boldsymbol{a})$. Rewriting $\frac{1+I_j}{D} + \frac{1}{I_k}$ results in the expression

$$\frac{1 + I_k}{D} + \frac{1}{I_k} = \frac{1 + \frac{p_j^1}{M_2 a_j} - \frac{q_1}{M_2}}{\frac{a_j p_k^1 - a_k p_j^1}{M_2 a_j a_k}} + \frac{1}{\frac{p_k^1}{M_2 a_k} - \frac{q_1}{M_2}} = \frac{a_j a_k (M_2 - q_1) + p_j^1 a_k}{p_k^1 a_j - p_j^1 a_k} + \frac{M_2}{\frac{p_k^1}{a_k} - q_1}.$$

References

1. Aardal, K., Bixby, R.E., Hurkens, C.A.J., Lenstra, A.K., Smeltink, J.W.: Market split and basis reduction: Towards a solution of the Cornuéjols-Dawande instances. INFORMS J. Comput. **12**, 192–202 (2000)
2. Aardal, K., Hurkens, C.A.J., Lenstra, A.K.: Solving a system of diophantine equations with lower and upper bounds on the variables. Math. Oper. Res. **25**, 427–442 (2000)
3. Aardal, K., Lenstra, A.K.: Hard equality constrained integer knapsacks. Mathematics of Operations Research, **29**(3), 724–738 (2004). Erratum: Mathematics of Operations Research, **31**(4), p. 846 (2006)
4. Andersen, K., Pochet, Y.: Coefficient strengthening: a tool for formulating mixed integer programs. CORE DP 2007/24, Université catholique de Louvain (2007)
5. Bradley, G.H.: Transformation of integer programs to knapsack problems. Discrete Math. **1**, 29–45 (1971)
6. Bradley, G.H.: Equivalent integer programs and canonical problems. Manage. Sci. **17**, 354–366 (1971)
7. Bradley, G.H., Hammer, P.L., Wolsey, L.A.: Coefficient reduction for inequalities in 0-1 variables. Math. Program. **7**, 263–282 (1974)
8. Cassels, J.W.S.: An Introduction to the Geometry of Numbers. Classics in Mathematics. Springer, Berlin (1997). Second Printing, Corrected, Reprint of the 1971 ed.
9. Cornuéjols, G., Dawande, M.: A class of hard small 0-1 programs. INFORMS J. Comput. **11**, 205–210 (1999)
10. Cornuéjols, G., Urbaniak, R., Weismantel, R., Wolsey, L.A.: Decomposition of integer programs and of generating sets. In: Burkard, R.E., Woeginger, G.J. (eds.) Algorithms–ESA '97. Lecture Notes in Computer Science, vol. 1284. pp. 92–103. Springer, Berlin (1997)
11. Kannan, R.: Algorithmic geometry of numbers. Annu. Rev. Comput. Sci. **2**, 231–267 (1987)
12. Kannan, R., Bachem, A.: Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix. SIAM J. Comput. **8**, 499–507 (1979)
13. Karamanov, M., Cornuéjols, G.: Branching on general disjunctions. Working paper, Tepper School of Business, Carnegie Mellon University (2005). Revised September 2007. To appear in: Chvátal, V., Sbihi, N. (eds.) Proceedings of the Montreal 2006 NATO Conference
14. Lenstra, A.K., Lenstra, H.W. Jr., Lovász, L.: Factoring polynomials with rational coefficients. Math. Ann. **261**, 515–534 (1982)
15. Lenstra, H.W. Jr.: Flags and lattice basis reduction. In: Casacuberta, C., Miró-Roig, R.M., Verdera, J., Xambó-Descamps, S. (eds.) Proceedings of the third European Congress of Mathematics, vol. I, pp. 37–51. Birkhäuser Verlag, Basel (2000)
16. Lenstra, H.W. Jr.: Lattices. To appear in: Surveys in Algorithmic Number Theory, Mathematical Sciences Research Institute Publications, Cambridge University Press, Cambridge (2005)
17. Lovász, L.: An Algorithmic Theory of Numbers, Graphs and Convexity. CBMS-NSF Regional Conference Series in applied mathematics, vol. 50. SIAM, Philadelphia (1986)
18. Louveaux, Q., Wolsey, L.A.: Combining problem structure with basis reduction to solve a class of hard integer programs. Math. Oper. Res. **27**(3), 470–484 (2002)
19. Martin, R.K., Schrage, L.: Subset coefficient reduction cuts for 0-1 mixed integer programming. Oper. Res. **33**, 505–526 (1985)
20. Padberg, M.W.: Equivalent knapsack-type formulations of bounded integer programs: an alternative approach. Naval Res. Log. Q. **19**, 699–708 (1972)
21. Schrijver, A.: Theory of Linear and Integer Programming. Wiley, Chichester (1986)
22. Xpress-MP Optimization Software. Dash optimization. http://www.dashoptimization.com/home/index.html