FULL LENGTH PAPER

Graver basis and proximity techniques for block-structured separable convex integer minimization problems

**Raymond Hemmecke · Matthias Köppe ·
Robert Weismantel**

Abstract We consider N-fold 4-block decomposable integer programs, which simultaneously generalize N-fold integer programs and two-stage stochastic integer programs with N scenarios. In previous work (Hemmecke et al. in Integer programming and combinatorial optimization. Springer, Berlin, 2010), it was proved that for fixed blocks but variable N, these integer programs are polynomial-time solvable for any linear objective. We extend this result to the minimization of separable convex objective functions. Our algorithm combines Graver basis techniques with a proximity result (Hochbaum and Shanthikumar in J. ACM 37:843–862, 1990), which allows us to use convex continuous optimization as a subroutine.

R. Hemmecke
Zentrum Mathematik, M9, Technische Universität München,
Boltzmannstr. 3, 85747 Garching, Germany
e-mail: hemmecke@ma.tum.de

M. Köppe (✉)
Department of Mathematics, University of California,
Davis, One Shields Avenue, Davis, CA 95616, USA
e-mail: mkoeppe@math.ucdavis.edu

R. Weismantel
Institute for Operations Research, ETH Zürich, Rämistrasse 101,
8092 Zurich, Switzerland
e-mail: robert.weismantel@ifor.math.ethz.ch

1 Introduction

We consider a family of nonlinear integer minimization problems over block-
structured linear constraint systems in variable dimension. The objective is to minimize
a separable convex objective function $f: \mathbb{R}^n \to \mathbb{R}$, defined as

$$f(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(x_i),$$

with convex functions $f_i: \mathbb{R} \to \mathbb{R}$ of one variable each.

Hochbaum and Shanthikumar [13] present a general technique for transforming
algorithms for linear integer minimization to algorithms for separable convex integer
minimization. The key ingredients of this transformation technique are scaling tech-
niques and proximity results between optimal integer solutions and optimal solutions
of the continous relaxation. This technique leads directly to polynomial time algo-
rithms if all the subdeterminants of the constraint matrix are bounded polynomially.

Of course, this is quite a restrictive hypothesis, but an important corollary of this
work is a polynomial time algorithm for minimizing a separable convex function
over systems of inequalities associated with a unimodular matrix. This generalizes, in
particular, earlier work of Minoux [14] on minimum cost flows with separable convex
cost functions.

An impossibility result on the existence of a strongly polynomial algorithm for
minimizing a general separable convex function over network flow constraints has
been shown in [12].

In the present paper, we study a certain family of block-structured separable convex
integer minimization problems over polyhedra, which does not satisfy the hypothesis
of polynomially bounded subdeterminants. The constraint matrix of these problems
is *N-fold* 4-*block decomposable* as follows:

$$\begin{pmatrix} C & D \\ B & A \end{pmatrix}^{(N)} := \begin{pmatrix} C & D & D & \cdots & D \\ B & A & O & & O \\ B & O & A & & O \\ \vdots & & & \ddots & \\ B & O & O & & A \end{pmatrix}$$

for some given $N \in \mathbb{Z}_+$ and N copies of A, B, and D. This problem type was studied
recently in [8].

N-fold 4-block decomposable matrices arise in many contexts and have been stud-
ied in various special cases, three of which are particularly relevant. We denote by O
a zero matrix of compatible dimensions and by · a matrix with no columns or no rows.

(i) For $C = \cdot$ and $D = \cdot$ we recover the problem matrix $\left(\begin{smallmatrix} \cdot & \cdot \\ B & A \end{smallmatrix}\right)^{(N)}$ of a two-stage stochastic integer optimization problem. Then, B is the matrix associated with the first stage decision variables and A is associated with the decision on stage 2. The number of occurrences of blocks of the matrix A reflect all the possible scenarios that pop up once a first stage decision has been made. We refer to [11] for a survey on state of the art techniques to solve this problem.

(ii) For $B = \cdot$ and $C = \cdot$ we recover the problem matrix $\left(\begin{smallmatrix} \cdot & D \\ \cdot & A \end{smallmatrix}\right)^{(N)}$ of a so-called N-fold integer problem. Here, if we let A be the node-edge incidence matrix of the given network and set D to be the identity, then the resulting N-fold IP is a multicommodity network flow problem. Separable convex N-fold IPs can be solved in polynomial time, provided that the matrices A and D are fixed [2,10].

(iii) For totally unimodular matrices C, A their so-called 1-sum $\left(\begin{smallmatrix} C & O \\ O & A \end{smallmatrix}\right)$ is totally unimodular. Similarly, total unimodularity is preserved under the so-called 2-sum and 3-sum composition [21,24]. For example, for matrices C and A, column vector $\mathbf{a}$ and row vector $\mathbf{b}^{\mathsf{T}}$ of appropriate dimensions, the 2-sum of $(\,C\ \mathbf{a}\,)$ and $\left(\begin{smallmatrix} \mathbf{b}^{\mathsf{T}} \\ A \end{smallmatrix}\right)$ gives $\left(\begin{smallmatrix} C & \mathbf{ab}^{\mathsf{T}} \\ O & A \end{smallmatrix}\right)$. The 2-sum of $\left(\begin{smallmatrix} C & \mathbf{ab}^{\mathsf{T}} & \mathbf{a} \\ O & A & O \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} \mathbf{b}^{\mathsf{T}} \\ B \end{smallmatrix}\right)$ creates the matrix

$$\begin{pmatrix} C & \mathbf{ab}^{\mathsf{T}} & \mathbf{ab}^{\mathsf{T}} \\ O & A & O \\ O & O & A \end{pmatrix},$$

which is the 2-fold 4-block decomposable matrix $\left(\begin{smallmatrix} C & \mathbf{ab}^{\mathsf{T}} \\ O & A \end{smallmatrix}\right)^{(2)}$. Repeated application of certain 1-sum, 2-sum and 3-sum compositions leads to a particular family of N-fold 4-block decomposable matrices with special structure regarding the matrices B and D.

(iv) The general case appears in stochastic integer programs with second order dominance relations [4] and stochastic integer multi-commodity flows. See [8] for further details of the model as an N-fold 4-block decomposable problem. To give one example consider a stochastic integer multi-commodity flow problem, introduced in [15,20]. Let M integer (in contrast to continuous) commodities to be transported over a given network. While we assume that supply and demands are deterministic, we assume that the upper bounds for the capacities per edge are uncertain and given initially only via some probability distribution. In a first stage we have to decide how to transport the M commodities over the given network without knowing the true capacities per edge. Then, after observing the true capacities per edge, penalties have to be paid if the capacity is exceeded. Assuming that we have knowledge about the probability distributions of the uncertain upper bounds, we wish to minimize the costs for the integer multi-commodity flow plus the expected penalties to be paid for exceeding capacities. To solve this problem, we discretize as usual the probability distribution for the uncertain upper bounds into N scenarios. Doing so, we obtain a (typically large-scale) (two-stage stochastic) integer programming problem as shown in Fig. 1. Herein, A is the node-edge incidence matrix of the given network, I is an identity matrix of appropriate size, and the columns containing $-I$ correspond to the penalty variables.
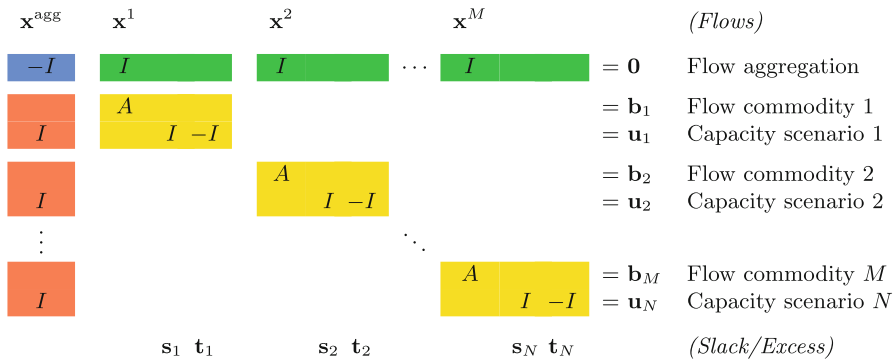
Fig. 1 Modeling a two-stage stochastic integer multi-commodity flow problem as an N-fold 4-block decomposable problem. Without loss of generality, the number of commodities and the number of scenarios are assumed to be equal

2 Main results and proof outline

In [8], the authors proved the following result.

Theorem 2.1 *Let $A \in \mathbb{Z}^{d_A \times n_A}$, $B \in \mathbb{Z}^{d_A \times n_B}$, $C \in \mathbb{Z}^{d_C \times n_B}$, $D \in \mathbb{Z}^{d_C \times n_A}$ be fixed matrices. For given $N \in \mathbb{Z}_+$ let $\mathbf{l} \in (\mathbb{Z} \cup \{-\infty\})^{n_B + N n_A}$, $\mathbf{u} \in (\mathbb{Z} \cup \{+\infty\})^{n_B + N n_A}$, $\mathbf{b} \in \mathbb{Z}^{d_C + N d_A}$, and let $f : \mathbb{R}^{n_B + N n_A} \to \mathbb{R}$ be a separable convex function that takes integer values on $\mathbb{Z}^{n_B + N n_A}$ and denote by $\hat{f}$ an upper bound on the maximum of $|f|$ over the feasible region of the N-fold 4-block decomposable convex integer minimization problem*

$$(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}: \qquad \min \left\{ f(\mathbf{z}) : \begin{pmatrix} C & D \\ B & A \end{pmatrix}^{(N)} \mathbf{z} = \mathbf{b},\ \mathbf{l} \leq \mathbf{z} \leq \mathbf{u},\ \mathbf{z} \in \mathbb{Z}^{n_B + N n_A} \right\}.$$

We assume that f is given only by a comparison oracle that, when queried on $\mathbf{z}$ and $\mathbf{z}'$ decides whether $f(\mathbf{z}) < f(\mathbf{z}')$, $f(\mathbf{z}) = f(\mathbf{z}')$ or $f(\mathbf{z}) > f(\mathbf{z}')$. Then the following holds:

(a) *There exists an algorithm with input N, $\mathbf{l}$, $\mathbf{u}$, $\mathbf{b}$ that computes a feasible solution to $(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ or decides that no such solution exists and that runs in time polynomial in N and in the binary encoding lengths $\langle \mathbf{l}, \mathbf{u}, \mathbf{b} \rangle$.*

(b) *There exists an algorithm with input N, $\mathbf{l}$, $\mathbf{u}$, $\mathbf{b}$ and a feasible solution $\mathbf{z}_0$ to $(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ that decides whether $\mathbf{z}_0$ is optimal or finds a better feasible solution $\mathbf{z}_1$ to the problem $(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ with $f(\mathbf{z}_1) < f(\mathbf{z}_0)$ and that runs in time polynomial in N and in the binary encoding lengths $\langle \mathbf{l}, \mathbf{u}, \mathbf{b}, \hat{f} \rangle$.*

(c) *For the restricted problem where f is linear, there exists an algorithm with input N, $\mathbf{l}$, $\mathbf{u}$, $\mathbf{b}$ that finds an optimal solution to the problem $(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ or decides that $(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ is infeasible or unbounded and that runs in time polynomial in N and in the binary encoding lengths $\langle \mathbf{l}, \mathbf{u}, \mathbf{b}, \hat{f} \rangle$.*

This theorem generalizes a similar statement for N-fold integer programming and for two-stage stochastic integer programming. In these two special cases, one can even

prove claim (c) of Theorem 2.1 for all separable convex functions and for a certain class of separable convex functions, respectively. In [8], it was posed as an open question whether Theorem 2.1 can be extended, for the full class of N-fold 4-block decomposable problems, from linear f to general separable convex functions f.

In the present paper, we settle this question, proving the following result for separable convex functions f, for which we assume that the following approximate continuous convex optimization oracle is available:

Problem 2.2 (*Approximate continuous convex optimization*) Given the data A, B, C, D, N, $\mathbf{l}$, $\mathbf{u}$, $\mathbf{b}$ and a number $\epsilon \in \mathbb{Q}_{>0}$, find a feasible solution $\mathbf{r}_\epsilon \in \mathbb{Q}^{n_B + N n_A}$ for the continuous relaxation

$$(\mathrm{CP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}: \qquad \min \left\{ f(\mathbf{r}) : \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \mathbf{r} = \mathbf{b}, \ \mathbf{l} \le \mathbf{r} \le \mathbf{u}, \ \mathbf{r} \in \mathbb{R}^{n_B + N n_A} \right\}.$$

such that there exists an optimal solution $\hat{\mathbf{r}}$ to $(\mathrm{CP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ with

$$\|\hat{\mathbf{r}} - \mathbf{r}_\epsilon\|_\infty \le \epsilon,$$

or report INFEASIBLE or UNBOUNDED.

Under suitable assumptions on the function f, this oracle may be implemented using the ellipsoid method or an interior point method.

Theorem 2.3 *For the problem of Theorem 2.1, we assume that the objective function f is given by an evaluation oracle and an approximate continuous convex optimization oracle for $(\mathrm{CP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$.*

Then there exists an algorithm that finds an optimal solution to $(\mathrm{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ or decides that $(\mathrm{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ is infeasible or unbounded and that runs in time polynomial in N and in the binary encoding lengths $\langle \mathbf{l}, \mathbf{u}, \mathbf{b}, \hat{f} \rangle$.

The main new technical contribution of the present paper is to combine Graver basis techniques with a proximity result developed by Hochbaum and Shanthikumar [13] in the context of their so-called proximity-scaling technique.

This allows us to first use the approximate continuous convex optimization oracle to find a point, in whose proximity the optimal integer solution has to lie. The integer problem restricted to this neighborhood is then efficiently solvable with primal (augmentation) algorithms using Graver bases, which will find the optimal integer solution in a polynomial number of steps.

We now briefly explain the Graver basis techniques; we refer the reader to the survey paper [18] or the monograph [17] for more details. Let $E \in \mathbb{Z}^{d \times n}$ be a matrix. We associate with E a finite set $\mathcal{G}(E)$ of vectors with remarkable properties, the *Graver basis* of E. The set $\mathcal{G}(E)$ consists of all nonzero vectors $\mathbf{v} \in \ker(E) \cap \mathbb{Z}^n$ that cannot be written as a sum $\mathbf{v} = \mathbf{v}' + \mathbf{v}''$ of nonzero vectors $\mathbf{v}', \mathbf{v}'' \in \ker(E) \cap \mathbb{Z}^n$ that lie in the same orthant (or equivalently, have the same sign pattern in $\{\ge 0, \le 0\}^n$) as $\mathbf{v}$. The set $\mathcal{G}(E)$ was introduced by Graver [5], who proved that it constitutes an *optimality certificate* (*test set*) for the family of integer linear programs that share the same problem matrix, E. By this we mean that $\mathcal{G}(E)$ provides an augmenting vector for any

non-optimal feasible solution and hence allows the design of a simple augmentation algorithm to solve the integer linear program in a finite number of augmentations.

The augmentation technique can also be used to efficiently construct a feasible solution in the first place, in a procedure similar to phase I of the simplex algorithm [7].

More recently, it has been shown in [16] that $\mathcal{G}(E)$ constitutes an optimality certificate for a wider class of integer minimization problems, namely for those minimizing a separable convex objective function over a feasible region of the form

$$\{ \mathbf{z} : E\mathbf{z} = \mathbf{b}, \ \mathbf{l} \leq \mathbf{z} \leq \mathbf{u}, \ \mathbf{z} \in \mathbb{Z}^n \}.$$

Moreover, several techniques have been found to turn the augmentation algorithm into an efficient algorithm, bounding the number of augmentation steps polynomially. Three such speed-up techniques are known in the literature: For 0/1 integer linear problems, a simple *bit-scaling technique* suffices [23]. For general integer linear problems, one can use the *directed augmentation technique* [22], in which one uses Graver basis elements $\mathbf{v} \in \mathcal{G}(E)$ that are improving directions for the nonlinear functions $\mathbf{c}^\mathsf{T}\mathbf{v}^+ + \mathbf{d}^\mathsf{T}\mathbf{v}^-$, which are adjusted during the augmentation algorithm. For separable convex integer problems, one can use the *Graver-best augmentation technique* [10], where one uses an augmentation vector $\mathbf{v}$ that is at least as good as the best augmentation step of the form $\gamma\mathbf{g}$ with $\gamma \in \mathbb{Z}_+$ and $\mathbf{g} \in \mathcal{G}(E)$.

In [8], the authors found a way to implement the directed augmentation technique efficiently for N-fold 4-block decomposable integer programs, despite the exponential size of the Graver basis. This gives an efficient optimization algorithm for the case of linear objective functions, proving Theorem 2.1. It is still an open question whether the Graver-best augmentation technique can be implemented efficiently. This would give an alternative proof of Theorem 2.3.

The paper [8] and the present paper crucially rely on the following structural result about $\mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\right)$, which was proved in [8].

Theorem 2.4 *If* $A \in \mathbb{Z}^{d_A \times n_A}$, $B \in \mathbb{Z}^{d_A \times n_B}$, $C \in \mathbb{Z}^{d_C \times n_B}$, $D \in \mathbb{Z}^{d_C \times n_A}$ *are fixed matrices, then* $\max\left\{ \|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\right) \right\}$ *is bounded by a polynomial in* N.

We note that in the special case of N-fold IPs, the ℓ_1-norm is bounded by a constant (depending only on the fixed problem matrices and not on N), and in the special case of two-stage stochastic IPs, the ℓ_1-norm is bounded linearly in N. This fact suggests that N-fold 4-block IPs are much richer and more difficult to solve than the two special cases.

3 Proof of the results

3.1 Aggregation technique

We will use an aggregation/disaggregation technique, which is based on the following folklore fact on Graver bases (see, for example, Corollary 3.2 in [6]):

Lemma 3.1 (Aggregation) *Let $G = (F \mathbf{f} \mathbf{f})$ be a matrix with two identical columns. Then the Graver bases of $(F \mathbf{f})$ and G are related as follows:*

$$\mathcal{G}(G) = \{(\mathbf{u}, v, w) : vw \geq 0, (\mathbf{u}, v + w) \in \mathcal{G}((F \mathbf{f}))\} \cup \{\pm(\mathbf{0}, 1, -1)\}.$$

Thus, the maximum ℓ_1-norm of Graver basis elements does not change if we repeat columns.

Corollary 3.2 *Let G be a matrix obtained from a matrix F by repeating columns. Then*

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(G)\} = \max\{2, \max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(F)\}\}.$$

3.2 Bounds for Graver basis elements

Let us start by bounding the ℓ_1-norm of Graver basis elements of matrices. The following result can be found, for instance, in [17, Lemma 3.20].

Lemma 3.3 (Determinant bound) *Let $A \in \mathbb{Z}^{m \times n}$ be a matrix of rank r and let $\Delta(A)$ denote the maximum absolute value of subdeterminants of A. Then $\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(A)\} \leq (n - r)(r + 1)\Delta(A)$. Moreover, $\Delta(A) \leq (\sqrt{m}M)^m$, where M is the maximum absolute value of an entry of A.*

As a corollary of Lemma 3.3 and the aggregation technique (Corollary 3.2), we obtain the following result:

Corollary 3.4 (Determinant bound, aggregated) *Let $A \in \mathbb{Z}^{m \times n}$ be a matrix of rank r and let d be the number of different columns in A and M the maximum absolute value of an entry of A. Then*

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(A)\} \leq (d - r)(r + 1)\left(\sqrt{m}M\right)^m.$$

For matrices with only one row ($m = r = 1$), there are only $2M + 1$ different columns, and so this bound simplifies to $4M^2$. However, a tighter bound is known for this special case. The following lemma is a straight-forward consequence of Theorem 2 in [3].

Lemma 3.5 (PPI bound) *Let $A \in \mathbb{Z}^{1 \times n}$ be a matrix consisting of only one row and let M be an upper bound on the absolute values of the entries of A. Then we have $\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(A)\} \leq 2M - 1$.*

Let us now prove some more general degree bounds on Graver bases that we will use in the proof of the main theorem below.

Lemma 3.6 (Graver basis length bound for stacked matrices) *Let $L \in \mathbb{Z}^{d \times n}$ and let $F \in \mathbb{Z}^{m \times n}$. Moreover, put $E := \binom{F}{L}$. Then we have*

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(E)\} \leq \max\{\|\boldsymbol{\lambda}\|_1 : \boldsymbol{\lambda} \in \mathcal{G}(F \cdot \mathcal{G}(L))\} \cdot \max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(L)\}.$$

Proof Let $\mathbf{v} \in \mathcal{G}(E)$. Then $\mathbf{v} \in \ker(L)$ implies that $\mathbf{v}$ can be written as a nonnegative integer linear sign-compatible sum $\mathbf{v} = \sum \lambda_i \mathbf{g}_i$ using Graver basis vectors $\mathbf{g}_i \in \mathcal{G}(L)$. Adding zero components if necessary, we can write $\mathbf{v} = \mathcal{G}(L)\lambda$. We now claim that $\mathbf{v} \in \mathcal{G}(E)$ implies $\lambda \in \mathcal{G}(F \cdot \mathcal{G}(L))$.

First, observe that $\mathbf{v} \in \ker(F)$ implies $F\mathbf{v} = F \cdot (\mathcal{G}(L)\lambda) = (F \cdot \mathcal{G}(L))\lambda = \mathbf{0}$ and thus, $\lambda \in \ker(F \cdot \mathcal{G}(L))$. If $\lambda \notin \mathcal{G}(F \cdot \mathcal{G}(L))$, then it can be written as a sign-compatible sum $\lambda = \mu + \nu$ with $\mu, \nu \in \ker(F \cdot \mathcal{G}(L))$. But then

$$\mathbf{v} = (\mathcal{G}(L)\mu) + (\mathcal{G}(L)\nu)$$

gives a sign-compatible decomposition of $\mathbf{v}$ into vectors $\mathcal{G}(L)\mu, \mathcal{G}(L)\nu \in \ker(E)$, contradicting the minimality property of $\mathbf{v} \in \mathcal{G}(E)$. Hence, $\lambda \in \mathcal{G}(F \cdot \mathcal{G}(L))$.

From $\mathbf{v} = \sum \lambda_i \mathbf{g}_i$ with $\mathbf{g}_i \in \mathcal{G}(L)$ and $\lambda \in \mathcal{G}(F \cdot \mathcal{G}(L))$, the desired estimate follows. $\qquad\square$

We will employ the following simple corollary.

Corollary 3.7 *Let $L \in \mathbb{Z}^{d \times n}$ and let $\mathbf{a}^\mathsf{T} \in \mathbb{Z}^n$ be a row vector. Moreover, put $E := \begin{pmatrix} \mathbf{a}^\mathsf{T} \\ L \end{pmatrix}$. Then we have*

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(E)\} \leq \left(2 \cdot \max\left\{|\mathbf{a}^\mathsf{T}\mathbf{v}| : \mathbf{v} \in \mathcal{G}(L)\right\} - 1\right) \cdot \max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(L)\}.$$

In particular, if $M := \max\{|a^{(i)}| : i = 1, \ldots, n\}$ then

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(E)\} \leq 2nM \left(\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(L)\}\right)^2.$$

Proof By Lemma 3.6, we already get

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(E)\} \leq \max\{\|\lambda\|_1 : \lambda \in \mathcal{G}(\mathbf{a}^\mathsf{T} \cdot \mathcal{G}(L))\} \cdot \max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(L)\}.$$

Now, observe that $\mathbf{a}^\mathsf{T} \cdot \mathcal{G}(L)$ is a $1 \times |\mathcal{G}(L)|$-matrix. Thus, the degree bound of primitive partition identities, Lemma 3.5, applies, which gives

$$\max\{\|\lambda\|_1 : \lambda \in \mathcal{G}(\mathbf{a}^\mathsf{T} \cdot \mathcal{G}(L))\} \leq 2 \cdot \max\left\{|\mathbf{a}^\mathsf{T}\mathbf{v}| : \mathbf{v} \in \mathcal{G}(L)\right\} - 1,$$

and thus, the first claim is proved. The second claim is a trivial consequence of the first. $\qquad\square$

Let us now extend this corollary to a form that we need to prove Theorem 2.4.

Corollary 3.8 *Let $L \in \mathbb{Z}^{d \times n}$ and let $F \in \mathbb{Z}^{m \times n}$. Let the entries of F be bounded by M in absolute value. Moreover, put $E := \begin{pmatrix} F \\ L \end{pmatrix}$. Then we have*

$$\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(E)\} \leq (2nM)^{2^m - 1} \left(\max\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}(L)\}\right)^{2^m}.$$

Proof This claim follows by simple induction, adding one row of F at a time, and by using the second inequality of Corollary 3.7 to bound the sizes of the intermediate Graver bases in comparison to the Graver basis of the matrix with one row of F fewer. □

In order to give a proof of Theorem 2.4, let us consider the submatrix $\left(\begin{smallmatrix} \cdot & \cdot \\ B & A \end{smallmatrix}\right)^{(N)}$. A main result from [11] is the following.

Theorem 3.9 (Graver basis for stochastic IPs) *Let $A \in \mathbb{Z}^{d_A \times n_A}$ and $B \in \mathbb{Z}^{d_A \times n_B}$, and let $\mathcal{G} = \mathcal{G}\left(\left(\begin{smallmatrix} \cdot & \cdot \\ B & A \end{smallmatrix}\right)^{(N)}\right)$. There exist numbers $g, \xi, \eta \in \mathbb{Z}_+$ depending only on A and B but not on N such that the following holds:*

(a) *For every $N \in \mathbb{Z}_+$ and for every $\mathbf{v} \in \mathcal{G}$, we have $\|\mathbf{v}\|_\infty \le g$, i.e., the components of $\mathbf{v}$ are bounded by g in absolute value.*
(b) *As a corollary, $\|\mathbf{v}\|_1 \le (n_B + Nn_A)g$ for all $\mathbf{v} \in \mathcal{G}$.*
(c) *More precisely, there exists a finite set $X \subseteq \mathbb{Z}^{n_B}$ of cardinality $|X| \le \xi$ and for each $\mathbf{x} \in X$ a finite set $Y_\mathbf{x} \subseteq \mathbb{Z}^{n_A}$ of cardinality $|Y_\mathbf{x}| \le \eta$ such that the elements $\mathbf{v} \in \mathcal{G}$ take the form $\mathbf{v} = (\mathbf{x}, \mathbf{y}^1, \ldots, \mathbf{y}^n)$, with $\mathbf{x} \in X$ and $\mathbf{y}^1, \ldots, \mathbf{y}^n \in Y_\mathbf{x}$.*

Remark 3.10 The finiteness of the numbers g, ξ, η comes from a saturation result in commutative algebra. Concrete bounds on these numbers are unfortunately not available. However, for given matrices A and B, the finite sets X and $Y_\mathbf{x}$ for $\mathbf{x} \in X$ can be computed using the Buchberger-type completion algorithm in [11, Section 3.3]. Thus, the numbers g, ξ, η are effectively computable.

Combining this result with Corollary 3.8, we get a bound for the ℓ_1-norms of the Graver basis elements of $\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}$.

Proposition 3.11 (Graver basis length bound for 4-block IPs) *Let $A \in \mathbb{Z}^{d_A \times n_A}$, $B \in \mathbb{Z}^{d_A \times n_B}$, $C \in \mathbb{Z}^{d_C \times n_B}$, $D \in \mathbb{Z}^{d_C \times n_A}$ be given matrices. Moreover, let M be a bound on the absolute values of the entries in C and D, and let $g \in \mathbb{Z}_+$ be the number from Theorem 3.9. Then for any $N \in \mathbb{Z}_+$ we have*

$$\max\left\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\right)\right\}$$

$$\le (2(n_B + Nn_A)M)^{2^{d_C}-1}\left(\max\left\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} \cdot & \cdot \\ B & A \end{smallmatrix}\right)^{(N)}\right)\right\}\right)^{2^{d_C}}$$

$$\le (2(n_B + Nn_A)M)^{2^{d_C}-1}((n_B + Nn_A)g)^{2^{d_C}}.$$

If A, B, C, D are fixed matrices, then $\max\left\{\|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\right)\right\}$ is bounded by $O(N^{2^{d_C}+1})$, a polynomial in N.

Proof The first claim is a direct consequence of Theorem 3.9 and Corollary 3.8 with $L = \left(\begin{smallmatrix} \cdot & \cdot \\ B & A \end{smallmatrix}\right)^{(N)}$, $F = \left(\begin{smallmatrix} C & D \\ \cdot & \cdot \end{smallmatrix}\right)^{(N)}$, and $E = \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}$. The polynomial bound for fixed matrices A, B, C, D and varying N follows by observing that n_A, n_B, d_C, M, g are constants as they depend only on the fixed matrices A, B, C, D. □

The above result has appeared before in [8]; we included the proof to make the present paper more self-contained. We now complement it with a useful alternative bound, which is given by the following new result.

Proposition 3.12 (Alternative length bound for 4-block IPs) *Let* $A \in \mathbb{Z}^{d_A \times n_A}$, $B \in \mathbb{Z}^{d_A \times n_B}$, $C \in \mathbb{Z}^{d_C \times n_B}$, $D \in \mathbb{Z}^{d_C \times n_A}$ *be given matrices. Moreover, let* M *be a bound on the absolute values of the entries in* C *and* D, *and let* $g, \xi, \eta \in \mathbb{Z}_+$ *be the numbers, depending on* A *and* B, *from Theorem* 3.9. *Then for any* $N \in \mathbb{Z}_+$ *we have*

$$\max \left\{ \|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \right) \right\}$$
$$\leq \xi \cdot (N + \eta)^\eta \cdot d_C \cdot \left(\sqrt{d_C}(n_B + N n_A) M g \right)^{d_C} \cdot (n_B + N n_A) g.$$

If A, B, C, D *are fixed matrices, then* $\max \left\{ \|\mathbf{v}\|_1 : \mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \right) \right\}$ *is bounded by* $\mathrm{O}(N^{d_C + \eta})$, *a polynomial in* N.

Either of the two results implies Theorem 2.4.

Remark 3.13 Comparing the two results is difficult because bounds for the finite numbers g, ξ, and, in particular, η (which appears in the exponent) are unknown. However, one should expect that the bound of Proposition 3.12 is better for matrices with large upper blocks $\left(\begin{smallmatrix} C & D \end{smallmatrix} \right)$, whereas the bound of Proposition 3.11 is better for matrices with large lower blocks $\left(\begin{smallmatrix} B & A \end{smallmatrix} \right)$.

Proof of Proposition 3.12 Let $L = \left(\begin{smallmatrix} B & A \end{smallmatrix} \right)^{(N)}$ and $F = \left(\begin{smallmatrix} C & D \end{smallmatrix} \right)^{(N)} = (C, D, \ldots, D)$. First of all, Theorem 3.9(b) gives the bound

$$\|\mathbf{v}\|_1 \leq (n_B + N n_A) g \quad \text{for} \quad \mathbf{v} \in \mathcal{G}(L), \tag{3.1}$$

where g is a constant that only depends on A and B.

We now consider the matrix $F \cdot \mathcal{G}(L)$. Each column of it is given by

$$F\mathbf{v} = C\mathbf{x} + D \sum_{i=1}^{N} \mathbf{y}^i \quad \text{with} \quad \mathbf{v} = \left(\mathbf{x}, \mathbf{y}^1, \ldots, \mathbf{y}^N \right) \in \mathcal{G}(L).$$

By Theorem 3.9(c), there are at most $\xi = \mathrm{O}(1)$ different vectors $\mathbf{x}$ and for each $\mathbf{x}$ at most $\eta = \mathrm{O}(1)$ different vectors $\mathbf{y}^i$. We now determine the number σ of different sums $\mathbf{s} = \sum_{i=1}^{N} \mathbf{y}^i$ that can arise from these choices. This number is bounded by the number of weak compositions of N into η non-negative integer parts: $\sigma \leq \binom{N+\eta-1}{\eta-1} \leq (N + \eta)^\eta = \mathrm{O}(N^\eta)$. Thus $F \cdot \mathcal{G}(L)$ has at most $d := \xi \cdot \sigma \leq \xi \cdot (N + \eta)^\eta = \mathrm{O}(N^\eta)$ different columns.

Using the bound on the entries of C and D, we find from (3.1) that the maximum absolute value of the entries of $F \cdot \mathcal{G}(L)$ is bounded by $(n_B + N n_A) M g$.

We now determine a length bound for the elements $\boldsymbol{\lambda}$ of $\mathcal{G}(F \cdot \mathcal{G}(L))$. By Corollary 3.4, we find that

$$\|\boldsymbol{\lambda}\|_1 \leq d \cdot d_C \cdot \left(\sqrt{d_C}(n_B + Nn_A)Mg\right)^{d_C}$$
$$\leq \xi \cdot (N + \eta)^\eta \cdot d_C \cdot \left(\sqrt{d_C}(n_B + Nn_A)Mg\right)^{d_C}. \tag{3.2}$$

Combining the two bounds (3.1) and (3.2) using Lemma 3.6 then gives the result.
$\square$

3.3 Constructing a feasible solution

For constructing a feasible solution to the problem $(\text{IP})_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$, we will use the algorithm of Theorem 2.1 (a), first introduced in [8]. For sake of completeness, we describe the algorithm here and thus give the proof of Theorem 2.1(a).

Proof of Theorem 2.1(a) Let $N \in \mathbb{Z}_+$, $\mathbf{l}, \mathbf{u} \in \mathbb{Z}^{n_B + Nn_A}$, $\mathbf{b} \in \mathbb{Z}^{d_C + Nd_A}$. First, construct an integer solution to the system $\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\mathbf{z} = \mathbf{b}$. This can be done in polynomial time using the Hermite normal form of $\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}$. Then we turn it into a feasible solution (satisfying $\mathbf{l} \leq \mathbf{z} \leq \mathbf{u}$) by a sequence of at most $\text{O}(Nd_A)$ many integer linear programs (with the same problem matrix $\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}$, but with bounds $\tilde{\mathbf{l}}, \tilde{\mathbf{u}}$ adjusted so that the current solution is feasible) with auxiliary objective functions that move the components of $\mathbf{z}$ into the direction of the given original bounds $\mathbf{l}, \mathbf{u}$, see [7]. This step is similar to phase I of the simplex method in linear programming.

In order to solve these auxiliary integer linear programs with polynomially many augmentation steps, we use the speed-up provided by the directed augmentation procedure [22]. This procedure requires us to repeatedly find, for certain vectors $\mathbf{c}$ and $\mathbf{d}$ that it constructs, an augmentation vector $\mathbf{v}$ with respect to the (separable convex) piecewise linear function $h(\mathbf{v}) = \mathbf{c}^\mathsf{T}\mathbf{v}^+ + \mathbf{d}^\mathsf{T}\mathbf{v}^-$.

Consequently, we only need to show how to find, for a given solution $\mathbf{z}_0$ that is feasible for $(\text{IP})_{N,\mathbf{b},\tilde{\mathbf{l}},\tilde{\mathbf{u}},h}$, an augmenting Graver basis element $\mathbf{v} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\right)$ for a separable convex piecewise linear function $h(\mathbf{v})$ in polynomial time in N and in the binary encoding lengths of $\mathbf{z}_0$ and of $\mathbf{c}, \mathbf{d}$.

Let us now assume that we are given a solution $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0^1, \ldots, \mathbf{y}_0^N)$ that is feasible for $(\text{IP})_{N,\mathbf{b},\tilde{\mathbf{l}},\tilde{\mathbf{u}},h}$ and that we wish to decide whether there exists another feasible solution $\mathbf{z}_1$ with $h(\mathbf{z}_1 - \mathbf{z}_0) < 0$. By [5,16], it suffices to decide whether there exists some vector $\mathbf{v} = (\bar{\mathbf{x}}, \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N)$ in the Graver basis of $\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}$ such that $\mathbf{z}_0 + \mathbf{v}$ is feasible and $h(\mathbf{v}) < 0$. By Propositions 3.11 or 3.12, the ℓ_1-norm of $\mathbf{v}$ is bounded polynomially in N. Thus, since n_B is constant, there is only a polynomial number of candidates for the $\bar{\mathbf{x}}$-part of $\mathbf{v}$. Since the bounds given by Propositions 3.11 and 3.12 are effectively computable (cf. Remark 3.10), we can actually list all possible vectors $\bar{\mathbf{x}}$ that satisfy these bounds.

For each such candidate $\bar{\mathbf{x}}$, we can find a best possible choice for $\bar{\mathbf{y}}_1, \ldots, \bar{\mathbf{y}}_N$ by solving the following N-fold IP:

$$\min \left\{ h\left(\mathbf{v}\right) : \begin{array}{c} \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} (\mathbf{z}_0 + \mathbf{v}) = \mathbf{b}, \\ \tilde{\mathbf{l}} \le (\mathbf{z}_0 + \mathbf{v}) \le \tilde{\mathbf{u}}, \\ \mathbf{v} = \left(\bar{\mathbf{x}}, \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N \right) \in \mathbb{Z}^{n_B + N n_A} \end{array} \right\}$$

$$= \min \left\{ h \begin{pmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{y}}^1 \\ \vdots \\ \bar{\mathbf{y}}^N \end{pmatrix} : \begin{array}{c} \left(\begin{smallmatrix} \cdot & D \\ \cdot & A \end{smallmatrix} \right)^{(N)} \begin{pmatrix} \bar{\mathbf{y}}^1 \\ \vdots \\ \bar{\mathbf{y}}^N \end{pmatrix} = \mathbf{b} - \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \mathbf{z}_0 - \left(\begin{smallmatrix} C & \cdot \\ B & \cdot \end{smallmatrix} \right)^{(N)} \bar{\mathbf{x}}, \\ \tilde{\mathbf{l}} - \mathbf{z}_0 \le \begin{pmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{y}}^1 \\ \vdots \\ \bar{\mathbf{y}}^N \end{pmatrix} \le \tilde{\mathbf{u}} - \mathbf{z}_0, \\ \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N \in \mathbb{Z}^{n_A} \end{array} \right\}$$

for given $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0^1, \ldots, \mathbf{y}_0^N)$ and $\bar{\mathbf{x}}$. As shown in the second line, this problem does indeed simplify to a separable convex N-fold IP with problem matrix $\left(\begin{smallmatrix} \cdot & D \\ \cdot & A \end{smallmatrix} \right)^{(N)}$ because $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0^1, \ldots, \mathbf{y}_0^N)$ and $\bar{\mathbf{x}}$ are fixed. Since the matrices A and D are fixed, each such N-fold IP is solvable in polynomial time [10]. In fact, as shown in [9], because the function h is "2-piecewise affine", this problem can be solved in time $O(N^3 L)$ by Graver-based dynamic programming, where $L = \langle \mathbf{c}, \mathbf{d}, \tilde{\mathbf{l}}, \tilde{\mathbf{u}}, \mathbf{z}_0, \bar{\mathbf{x}} \rangle$.

If the N-fold IP is infeasible, there does not exist an augmenting vector using the particular choice of $\bar{\mathbf{x}}$. If it is feasible, let $\mathbf{v} = (\bar{\mathbf{x}}, \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N)$ be the optimal solution. Now if we have $h(\mathbf{v}) \ge 0$, then no augmenting vector can be constructed using this particular choice of $\bar{\mathbf{x}}$. If, on the other hand, we have $h(\mathbf{v}) < 0$, then $\mathbf{v}$ is a desired augmenting vector for $\mathbf{z}_0$ and we can stop.

As we solve polynomially many polynomially solvable N-fold IPs, one for each choice of $\bar{\mathbf{x}}$, an optimality certificate or a desired augmentation step can be computed in polynomial time and the claim follows.   □

3.4 Using Hochbaum–Shanthikumar's proximity results

Hochbaum and Shanthikumar [13] present an algorithm for nonlinear separable convex integer minimization problems for matrices with small subdeterminants. The algorithm is based on the so-called proximity-scaling technique. It is pseudo-polynomial in the sense that the running time depends polynomially on the absolute value of the largest subdeterminant of the problem matrix. The results of the paper [13] cannot be directly applied to our situation, since the subdeterminants of N-fold 4-block decomposable matrices typically grow exponentially in N. In the following we adapt a lemma from [13] that establishes proximity of optimal solutions of the integer problem and its continuous relaxation; we do not use the scaling technique, however.

We consider the separable convex integer minimization problem

$$\min \left\{ f(\mathbf{z}) : E\mathbf{z} = \mathbf{b}, \ \mathbf{l} \le \mathbf{z} \le \mathbf{u}, \ \mathbf{z} \in \mathbb{Z}^n \right\}. \tag{3.3}$$

Theorem 3.14 (Proximity) *Let $\hat{\mathbf{r}}$ be an optimal solution of the continuous relaxation of* (3.3),

$$\min \left\{ f(\mathbf{r}) : E\mathbf{r} = \mathbf{b}, \ \mathbf{l} \le \mathbf{r} \le \mathbf{u}, \ \mathbf{r} \in \mathbb{R}^n \right\}. \tag{3.4}$$

Then there exists an optimal solution $\mathbf{z}^*$ of the integer optimization problem (3.3) *with*

$$\|\hat{\mathbf{r}} - \mathbf{z}^*\|_\infty \le n \cdot \max \left\{ \|\mathbf{v}\|_\infty : \mathbf{v} \in \mathcal{G}(E) \right\}.$$

We remark that we actually just need a bound on the circuits of E. A vector in $\ker(E)$ is called a *circuit* of E if its support is inclusion minimal among all elements in $\ker(E)$ and its components are integer and relatively prime. The set of circuits forms a subset of the Graver basis of E.

Hochbaum and Shanthikumar [13] prove a version of this result where the maximum of the absolute values of the subdeterminants of E appears on the right-hand side. Our proof is almost identical, but we include it here for completeness.

Proof Let $\hat{\mathbf{z}}$ be an optimal solution of the integer optimization problem (3.3). Since $\hat{\mathbf{z}}$ is a feasible solution to the continuous relaxation, there exists a conformal (orthant-compatible) decomposition of $\hat{\mathbf{r}} - \hat{\mathbf{z}}$ into rational multiples of the circuits of E,

$$\hat{\mathbf{r}} - \hat{\mathbf{z}} = \sum_{i=1}^{n} \alpha_i \mathbf{u}^i, \quad \alpha_i \ge 0, \ \mathbf{u}^i \in \mathcal{C}(E),$$

where, due to Carathéodory's theorem, at most n circuits are needed. Then

$$\hat{\mathbf{r}} - \hat{\mathbf{z}} = \sum_{i=1}^{n} \lfloor \alpha_i \rfloor \mathbf{u}^i + \sum_{i=1}^{n} \beta_i \mathbf{u}^i,$$

setting $\beta_i = \alpha_i - \lfloor \alpha_i \rfloor$. Now we define

$$\mathbf{r}^* = \hat{\mathbf{z}} + \sum_{i=1}^{n} \beta_i \mathbf{u}^i, \quad \text{and} \quad \mathbf{z}^* = \hat{\mathbf{z}} + \sum_{i=1}^{n} \lfloor \alpha_i \rfloor \mathbf{u}^i.$$

Since the vectors $\mathbf{u}^i$ lie in the kernel of matrix E, both $\mathbf{z} = \mathbf{z}^*$ and $\mathbf{z} = \mathbf{r}^*$ satisfy the equation $E\mathbf{z} = \mathbf{b}$. Moreover, since both $\hat{\mathbf{r}}$ and $\hat{\mathbf{z}}$ lie within the lower and upper bounds and the vectors $\mathbf{u}^i$ lie in the same orthant as $\hat{\mathbf{r}} - \hat{\mathbf{z}}$, also $\mathbf{z}^*$ and $\mathbf{r}^*$ lie within the lower and upper bounds. Thus, $\mathbf{r}^*$ is a feasible solution to the continuous relaxation of (3.3). Since $\mathbf{z}^*$ is also an integer vector, it is a feasible solution to the integer optimization problem (3.3).

We can write

$$\hat{\mathbf{r}} - \hat{\mathbf{z}} = [\mathbf{r}^* - \hat{\mathbf{z}}] + [\mathbf{z}^* - \hat{\mathbf{z}}].$$

Then we use an important superadditivity property of separable convex functions (see [13, Lemma 3.1] and [16]), which gives

$$f(\hat{\mathbf{r}}) - f(\hat{\mathbf{z}}) \ge [f(\mathbf{r}^*) - f(\hat{\mathbf{z}})] + [f(\mathbf{z}^*) - f(\hat{\mathbf{z}})], \tag{3.5}$$

or, equivalently,

$$f(\hat{\mathbf{r}}) - f(\mathbf{r}^*) \geq f(\mathbf{z}^*) - f(\hat{\mathbf{z}}). \tag{3.6}$$

Since $\hat{\mathbf{r}}$ is an optimal solution to the continuous relaxation and $\mathbf{r}^*$ is a feasible solution to it, the left-hand side is nonpositive, and so $f(\mathbf{z}^*) \leq f(\hat{\mathbf{z}})$. Thus, since $\mathbf{z}^*$ is a feasible solution to (3.3), it is, in fact, another optimal solution of the integer optimization problem and $f(\mathbf{z}^*) = f(\hat{\mathbf{z}})$.

We now verify the proximity of $\mathbf{z}^*$ to $\hat{\mathbf{r}}$. From the definition of $\mathbf{z}^*$, we immediately get

$$
\begin{aligned}
\left\| \hat{\mathbf{r}} - \mathbf{z}^* \right\|_\infty &= \left\| [\hat{\mathbf{r}} - \hat{\mathbf{z}}] + [\hat{\mathbf{z}} - \mathbf{z}^*] \right\|_\infty \\
&= \left\| \sum_{i=1}^n \alpha_i \mathbf{u}^i - \sum_{i=1}^n \lfloor \alpha_i \rfloor \mathbf{u}^i \right\|_\infty \\
&= \left\| \sum_{i=1}^n \beta_i \mathbf{u}^i \right\|_\infty \\
&\leq n \cdot \max\{ \| \mathbf{u}^j \|_\infty : j = 1, \ldots, n \} \\
&\leq n \cdot \max\{ \| \mathbf{v} \|_\infty : \mathbf{v} \in \mathcal{G}(E) \}.
\end{aligned}
$$

This concludes the proof.   □

As an immediate corollary, we obtain the following result.

Corollary 3.15 *Let $\epsilon \geq 0$ and let $\hat{\mathbf{r}}$ be an optimal solution to the continuous relaxation* (3.4). *Setting*

$$
\begin{aligned}
\mathbf{l}' &= \max\{\mathbf{l}, \lfloor \hat{\mathbf{r}} - (n \cdot \ell)\mathbf{1} \rfloor\}, \\
\mathbf{u}' &= \min\{\mathbf{u}, \lceil \hat{\mathbf{r}} + (n \cdot \ell)\mathbf{1} \rceil\},
\end{aligned}
$$

where $\ell = \max \{\| \mathbf{v} \|_\infty : \mathbf{v} \in \mathcal{G}(E)\}$, we have

$$
\begin{aligned}
&\min\{f(\mathbf{z}) : E\mathbf{z} = \mathbf{b}, \, \mathbf{l} \leq \mathbf{z} \leq \mathbf{u}, \, \mathbf{z} \in \mathbb{Z}^n\} \\
&= \min\{f(\mathbf{z}) : E\mathbf{z} = \mathbf{b}, \, \mathbf{l}' \leq \mathbf{z} \leq \mathbf{u}', \, \mathbf{z} \in \mathbb{Z}^n\}. \tag{3.7}
\end{aligned}
$$

Later we will use a simple modification of Corollary 3.15, using an ϵ-approximate optimal solution to the continuous relaxation (3.4).

For $E = \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)}$, we can control the size of ℓ using Propositions 3.11 or 3.12 and thus obtain an equivalent IP with small (polynomial-sized) bounds.

We note that though the bounds are small, the dimension is still variable, and so the problem cannot be solved efficiently with elementary techniques such as dynamic programming. In the following subsections, we show how to solve this IP with Graver basis techniques.

3.5 Graver-best augmentation for the restricted problem

In the restricted problem, no long augmentation steps are possible, and therefore it is possible to efficiently construct a Graver-best augmentation vector. Using this observation, we prove the following theorem.

Theorem 3.16 *Let $A \in \mathbb{Z}^{d_A \times n_A}$, $B \in \mathbb{Z}^{d_A \times n_B}$, $C \in \mathbb{Z}^{d_C \times n_B}$, $D \in \mathbb{Z}^{d_C \times n_A}$ be fixed matrices. Then there exists an algorithm that, given the input data $N \in \mathbb{Z}_+$, $\mathbf{b} \in \mathbb{Z}^{d_C + N d_A}$, $\mathbf{l}', \mathbf{u}' \in \mathbb{Z}^{n_B + N n_A}$, a feasible solution $\mathbf{z}_0$, and a comparison oracle for the function $f : \mathbb{R}^{n_B + N n_A} \to \mathbb{R}$, finds an optimal solution to*

$$\min \left\{ f(\mathbf{z}) : \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \mathbf{z} = \mathbf{b}, \ \mathbf{l}' \leq \mathbf{z} \leq \mathbf{u}', \ \mathbf{z} \in \mathbb{Z}^{n_B + N n_A} \right\}$$

and that runs in time that is polynomially bounded in N, in $k := \|\mathbf{u}' - \mathbf{l}'\|_\infty$, and in the binary encoding lengths $\langle \mathbf{b}, \hat{f} \rangle$.

Proof By the Graver-best speed-up technique [10], it suffices to show that for a given feasible solution $\mathbf{z}_0$, we can construct a vector $\gamma \mathbf{g}$, where $\gamma \in \mathbb{Z}_+$ and $\mathbf{g} \in \mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \right)$, such that $\mathbf{z}_0 + \gamma \mathbf{g}$ is feasible, and γ and $\mathbf{g}$ minimize $f(\mathbf{z}_0 + \gamma \mathbf{g})$ among all possible choices. It actually suffices to construct any vector $\mathbf{v}$ such that $\mathbf{z}_0 + \mathbf{v}$ is feasible and $f(\mathbf{z}_0 + \mathbf{v}) \leq f(\mathbf{z}_0 + \gamma \mathbf{g})$.

Write $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0^1, \ldots, \mathbf{y}_0^N)$ and let $\mathbf{v} = (\bar{\mathbf{x}}, \ldots)$ be any vector in the Graver basis of $\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)}$. By Propositions 3.11 or 3.12, the ℓ_1-norm of $\mathbf{v}$ is bounded polynomially in N. Thus, since n_B is constant, there is only a polynomial number of candidates for the $\bar{\mathbf{x}}$-part of $\mathbf{v}$. Since the bounds given by Propositions 3.11 and 3.12 are effectively computable (cf. Remark 3.10), we can actually list all possible vectors $\bar{\mathbf{x}}$ that satisfy these bounds.

For each such vector $\bar{\mathbf{x}}$, we now consider all vectors of the form $(\gamma \bar{\mathbf{x}}, \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N)$ as candidate augmentation vectors, not just multiples $\gamma \mathbf{v}$ of Graver basis elements.

In the special case $\bar{\mathbf{x}} = \mathbf{0}$, this is equivalent to the construction of a Graver-best augmentation vector for the N-fold IP with the problem matrix $\left(\begin{smallmatrix} \cdot & D \\ \cdot & A \end{smallmatrix} \right)^{(N)}$, which can be done in polynomial time [10].

Otherwise, if $\bar{\mathbf{x}} \neq \mathbf{0}$, we determine the largest step length $\hat{\gamma} \in \mathbb{Z}_+$ such that $\mathbf{x}_0 + \hat{\gamma} \bar{\mathbf{x}}$ lies within the bounds $\mathbf{l}', \mathbf{u}'$. Certainly $\hat{\gamma} \leq k$. We now check each possible step length $\gamma = 1, 2, \ldots, \hat{\gamma}$ separately. To find a best possible choice for $\bar{\mathbf{y}}_1, \ldots, \bar{\mathbf{y}}_N$, we solve the following N-fold IP:

$$\min \left\{ f(\mathbf{v}) : \begin{array}{c} \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} (\mathbf{z}_0 + \mathbf{v}) = \mathbf{b}, \\ \mathbf{l}' \leq (\mathbf{z}_0 + \mathbf{v}) \leq \mathbf{u}', \\ \mathbf{v} = \left(\gamma \bar{\mathbf{x}}, \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N \right) \in \mathbb{Z}^{n_B + N n_A} \end{array} \right\}.$$

Since the matrices A and D are fixed, each such N-fold IP is solvable in polynomial time [10].

If the N-fold IP is infeasible, there does not exist an augmenting vector using the particular choice of $\bar{\mathbf{x}}$ and γ. If it is feasible, let $\mathbf{v} = (\gamma \bar{\mathbf{x}}, \bar{\mathbf{y}}^1, \ldots, \bar{\mathbf{y}}^N)$ be an optimal solution. Now if we have $f(\mathbf{v}) \geq 0$, then no augmenting vector can be constructed

using this particular choice of $\bar{\mathbf{x}}$ and γ. If, on the other hand, we have $f(\mathbf{v}) < 0$, then $\mathbf{v}$ is a candidate for the Graver-best augmentation vector.

By iterating over all $\bar{\mathbf{x}}$ and all γ, we efficiently construct a Graver-best augmentation vector.   □

Remark 3.17 A more precise complexity analysis is as follows.

(a) For the construction in the special case $\bar{\mathbf{x}} = \mathbf{0}$: In fact, by [9, Lemma 3.4 and proof of Theorem 4.2], for any of the possible step lengths $\gamma = 1, 2, \ldots, k$, we can find in linear time $O(N)$ an augmenting vector $\gamma \mathbf{v}$ that is at least as good as the best Graver step $\gamma \mathbf{g}$ with $\mathbf{g} \in \mathcal{G}\left(\begin{smallmatrix} \cdot & D \\ \cdot & A \end{smallmatrix} \right)^{(N)}$. Checking all step lengths, we get a complexity of $O(kN)$.

(b) For the solution of the N-fold subproblem in the general case $\bar{\mathbf{x}} \neq \mathbf{0}$: This optimization, in turn, uses another Graver-best augmentation technique. In Phase I, the possible step lengths are large, but the auxiliary objective functions are linear, and so the running time is $O(N^3 L)$ by Graver-based dynamic programming [9, Theorem 3.9], where $L = \langle \mathbf{l}', \mathbf{u}', \mathbf{z}_0, \bar{\mathbf{x}} \rangle$. In Phase II, there are few possible step lengths, $\gamma = 1, 2, \ldots, k$, so we can try them all. By [9, Lemma 3.4 and proof of Theorem 4.2], we can find for a fixed γ in linear time $O(N)$ an augmenting vector $\gamma \mathbf{v}$ that is at least as good as the best Graver step $\gamma \mathbf{g}$ with $\mathbf{g} \in \mathcal{G}\left(\begin{smallmatrix} \cdot & D \\ \cdot & A \end{smallmatrix} \right)^{(N)}$. Checking all step lengths, we get a complexity of $O(kN)$. Using the results of [10] (modified with the optimality criterion of [16]), the number of Graver-best augmentations is bounded by $O(N \langle \hat{f} \rangle)$. Thus the complexity of this subproblem is $O(N^2 k \langle \hat{f} \rangle + N^3 L)$.

(c) The number of steps in the overall Graver-best augmentation algorithm for the restricted 4-block decomposable problem is again bounded by $O(N \langle \hat{f} \rangle)$.

Remark 3.18 Other augmentation techniques can be used to prove Theorem 3.16. For example, following [13, Section 2], we can reformulate a separable convex integer minimization problem with small bounds as a 0/1 linear integer minimization problem in the straightforward way. Then we can apply the bit-scaling speed-up technique, for instance; see [23].

3.6 Putting all together

For each set of fixed matrices A, B, C, D and for any function $\epsilon(N)$ that is bounded polynomially in N, we consider the following algorithm.

Algorithm 3.19 (Graver proximity algorithm)

1: **input** $N \in \mathbb{Z}_+$, bounds $\mathbf{l}, \mathbf{u} \in \mathbb{Z}^{n_B + N n_A}$, right-hand side $\mathbf{b} \in \mathbb{Z}^{d_C + N d_A}$, evaluation oracle for a separable convex function $f : \mathbb{R}^{n_B + N n_A} \to \mathbb{R}$, approximate continuous convex optimization oracle.

2: **output** an optimal solution $\mathbf{z}^*$ to $(\text{IP})_{N, \mathbf{b}, \mathbf{l}, \mathbf{u}, f}$ or INFEASIBLE or UNBOUNDED.

3: Let $n = n_B + N n_A$ denote the dimension of the problem.

4: Call the approximate continuous convex optimization oracle with $\epsilon = \epsilon(N)$ to find an approximate solution $\mathbf{r}_\epsilon \in \mathbb{Q}^{n_B + N n_A}$ to the continuous relaxation

$$\min \left\{ f(\mathbf{r}) : \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix} \right)^{(N)} \mathbf{r} = \mathbf{b}, \ \mathbf{l} \le \mathbf{r} \le \mathbf{u}, \ \mathbf{r} \in \mathbb{R}^{n_B + N n_A} \right\}.$$

5:   **if** oracle returns INFEASIBLE **then**

6:     **return** INFEASIBLE.

7:   **else if** oracle returns UNBOUNDED **then**

8:     **return** UNBOUNDED.

9:   **else**

10:   Compute an upper bound ℓ on the maximum ℓ_1-norm of the vectors in $\mathcal{G}\left(\left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\right)$, using Proposition 3.11 or Proposition 3.12.

11:   Let $\mathbf{l}' = \max\{\mathbf{l}, \lfloor \mathbf{r}_\epsilon - (n \cdot \ell + \epsilon)\mathbf{1} \rfloor\}$ and $\mathbf{u}' = \min\{\mathbf{u}, \lceil \mathbf{r}_\epsilon + (n \cdot \ell + \epsilon)\mathbf{1} \rceil\}$.

12:   Let $k = \|\mathbf{u}' - \mathbf{l}'\|_\infty$.

13:   Using the algorithm of Theorem 2.1 (a), find a feasible solution $\mathbf{z}_0$ for the restricted convex integer minimization problem

$$\min\left\{ f(\mathbf{z}) : \left(\begin{smallmatrix} C & D \\ B & A \end{smallmatrix}\right)^{(N)}\mathbf{z} = \mathbf{b},\ \mathbf{l}' \le \mathbf{z} \le \mathbf{u}',\ \mathbf{z} \in \mathbb{Z}^{n_B + N n_A} \right\}.$$

14:   Solve the problem to optimality using the algorithm of Theorem 3.16.

By analyzing this algorithm, we now prove the main theorem of this paper.

Proof of Theorem 2.3 We first show that Algorithm 3.19 is correct. If the continuous relaxation $(CP)_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ is infeasible or unbounded, then so is the problem $(IP)_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$. In the following, assume that $(CP)_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ has an optimal solution. Then there exists an optimal solution $\hat{\mathbf{r}}$ to $(CP)_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ with $\|\hat{\mathbf{r}} - \mathbf{r}_\epsilon\|_\infty \le \epsilon$. By Theorem 3.14, there exists an optimal solution $\mathbf{z}^*$ of the integer optimization problem $(IP)_{N,\mathbf{b},\mathbf{l},\mathbf{u},f}$ with $\|\hat{\mathbf{r}} - \mathbf{z}^*\|_\infty \le n \cdot \ell$. By the triangle inequality, this solution then satisfies $\|\mathbf{z}^* - \mathbf{r}_\epsilon\|_\infty \le n \cdot \ell + \epsilon$ and is therefore a feasible solution to the restricted IP with variable bounds $\mathbf{l}'$ and $\mathbf{u}'$. Thus it suffices to solve the restricted IP to optimality, which is done with the algorithm of Theorem 3.16.

The algorithm has the claimed complexity because

$$k \le 2((n_B + N n_A) \cdot \ell + \epsilon)$$

is bounded polynomially in N by Propositions 3.11 or 3.12. The complexity then follows from Theorem 3.16.   □

References

1. De Loera, J.A., Hemmecke, R., Onn, S., Rothblum, U.G., Weismantel, R.: Convex integer maximization via Graver bases. J. Pure Appl. Algebra **213**, 1569–1577 (2009)

2. De Loera, J.A., Hemmecke, R., Onn, S., Weismantel, R.: N-fold integer programming. Discret. Optim. **5**(2), 231–241 (2008) doi:10.1016/j.disopt.2006.06.006 (In Memory of George B. Dantzig)

3. Diaconis, P., Graham, R. L., Sturmfels, B.: Primitive partition identities. In: Miklós, D., Sós, V.T., Szőnyi, D. (eds.), Combinatorics, Paul Erdős is Eighty, Volume 2. Bolyai Society Mathematical Studies, vol. 2, pp. 173–192 (1996)

4. Gollmer, R., Gotzes, U., Schultz, R.: A note on second-order stochastic dominance constraints induced by mixed-integer linear recourse. Math. Program. **126**, 179–190 (2011)

5. Graver, J.E.: On the foundations of linear and integer linear programming I. Math. Program. **8**, 207–226 (1975)

6. Hemmecke, R.: Test sets for integer programs with $\mathbb{Z}$-convex objective. eprint arXiv:math/0309154 (2003)

7. Hemmecke, R.: On the positive sum property and the computation of Graver test sets. Math. Program. Ser. B **96**, 247–269 (2003)

8. Hemmecke, R., Köppe, M., Weismantel, R.: A polynomial-time algorithm for optimizing over N-fold 4-block decomposable integer programs, In: Eisenbrand, F., Shepherd, F.B. (eds.), Integer Programming and Combinatorial Optimization. Lecture Notes in Computer Science, vol. 6080, Springer, Berlin, pp. 219–229 (2010). doi:10.1007/978-3-642-13036-6_17

9. Hemmecke, R., Onn, S., Romanchuk, L.: N-fold integer programming in cubic time. Math. Program. 1–17 (2011). doi:10.1007/s10107-011-0490-y

10. Hemmecke, R., Onn, S., Weismantel, R.: A polynomial oracle-time algorithm for convex integer minimization. Math. Program. **126**, 97–117 (2011). doi:10.1007/s10107-009-0276-7

11. Hemmecke, R., Schultz, R.: Decomposition of test sets in stochastic integer programming. Math. Program. **94**(2–3), 323–341 (2003)

12. Hochbaum, D.S.: Lower and upper bounds for allocation problems. Math. Oper. Res. **19**, 390–409 (1994)

13. Hochbaum, D.S., Shanthikumar, J.G.: Convex separable optimization is not much harder than linear optimization, J. ACM **37**, 843–862 (1990). doi:10.1145/96559.96597

14. Minoux, M.: Solving integer minimum cost flows with separable convex cost objective polynomially. Math. Prog. Study **26**, 237–239 (1986)

15. Mirchandani, P., Soroush, H.: The stochastic multicommodity flow problem. Networks **20**, 121–155 (1990)

16. Murota, K., Saito, H., Weismantel, R.: Optimality criterion for a class of nonlinear integer programs. Oper. Res. Lett. **32**, 468–472 (2004)

17. Onn, S.: Nonlinear Discrete Optimization. Zurich Lectures in Advanced Mathematics, European Mathematical Society (2010)

18. Onn, S.: Theory and applications of N-fold integer programming. In: The IMA Volumes in Mathematics and its Applications, Mixed Integer Nonlinear Programming, pp. 559–593. Springer, Berlin (2012)

19. Onn, S., Rothblum, U.G.: Convex combinatorial optimization. Disc. Comp. Geom. **32**, 549–566 (2004)

20. Powell, W., Topaloglu, H.: Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems. INFORMS J. Comput. **18**, 31–42 (2006)

21. Schrijver, A.: Theory of Linear and Integer Programming. Wiley, New York (1986)

22. Schulz, A.S., Weismantel, R.: An oracle-polynomial time augmentation algorithm for integer programming. In: Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 967–968 (1999)

23. Schulz, A.S., Weismantel, R., Ziegler, G.M.: 0/1 integer programming: optimization and augmentation are equivalent. In: Spirakis, P. (ed.), Algorithms-ESA 95, Lecture Notes in Computer Science, vol. 979, Springer, Berlin, pp. 473–483 (1995)

24. Seymour, P.D.: Decomposition of regular matroids. J. Comb. Theory **28**, 305–359 (1980)