A polynomial oracle-time algorithm for convex integer minimization

Raymond Hemmecke
University of Magdeburg, Germany

Shmuel Onn
Technion Haifa, Israel

Robert Weismantel
University of Magdeburg, Germany

Abstract

In this paper we consider the solution of certain convex integer minimization problems via greedy augmentation procedures. We show that a greedy augmentation procedure that employs only directions from certain Graver bases needs only polynomially many augmentation steps to solve the given problem. We extend these results to convex N-fold integer minimization problems and to convex 2-stage stochastic integer minimization problems. Finally, we present some applications of convex N-fold integer minimization problems for which our approach provides polynomial time solution algorithms.

1 Introduction

For an integer matrix $A \in \mathbb{Z}^{d\times n}$, we define the circuits $\mathcal{C}(A)$ and the Graver basis $\mathcal{G}(A)$ as follows. Herein, an integer vector $v \in \mathbb{Z}^n$ is called *primitive* if all its components are coprime, that is, $\gcd(v_1, \ldots, v_n) = 1$.

Definition 1 *Let $A \in \mathbb{Z}^{d\times n}$ and let $\mathbb{O}_j$, $j = 1, \ldots, 2^n$ denote the 2^n orthants of $\mathbb{R}^n$. Then the cones*

$$C_j := \ker(A) \cap \mathbb{O}_j = \{z \in \mathbb{O}_j : Az = 0\}$$

are pointed rational polyhedral cones. Let R_j and H_j denote the (unique) minimal sets of primitive integer vectors generating C_j over $\mathbb{R}_+$ and $C_j \cap \mathbb{Z}^n$ over $\mathbb{Z}_+$, respectively. Then we define

$$\mathcal{C}(A) := \bigcup_{j=1}^{2^n} R_j \setminus \{0\} \quad \textit{and} \quad \mathcal{G}(A) := \bigcup_{j=1}^{2^n} H_j \setminus \{0\}$$

to be the set $\mathcal{C}(A)$ of circuits of A and the Graver basis $\mathcal{G}(A)$ of A.

Remark 2 It is not hard to show that $\mathcal{C}(A)$ corresponds indeed to all primitive support-minimal vectors in $\ker(A)$ [7].

Already in 1975, Graver showed that $\mathcal{C}(A)$ and $\mathcal{G}(A)$ provide optimality certificates for a large class of continuous and integer linear programs, namely for

$$(\mathrm{LP})_{A,u,b,f}: \qquad \min\{f(z) : Az = b, 0 \leq z \leq u, z \in \mathbb{R}_+^n\},$$

and

$$(\mathrm{IP})_{A,u,b,f}: \qquad \min\{f(z) : Az = b, 0 \leq z \leq u, z \in \mathbb{Z}_+^n\},$$

where the linear objective function $f(x) = c^\intercal x$, the upper bounds vector u, and the right-hand side vector b are allowed to be changed [7]. A solution z^0 to $(\mathrm{LP})_{A,u,b,f}$ is optimal if and only if there are no $g \in \mathcal{C}(A)$ and $\alpha \in \mathbb{R}_+$ such that $z^0 + \alpha g$ is a feasible solution to $(\mathrm{LP})_{A,u,b,f}$ that has a smaller objective function value $f(z^0 + \alpha g) < f(z^0)$. Analogously, an integer solution z^0 to $(\mathrm{IP})_{A,u,b,f}$ is optimal if and only if there are no $g \in \mathcal{G}(A)$ and $\alpha \in \mathbb{Z}_+$ such that $z^0 + \alpha g$ is a feasible solution to $(\mathrm{IP})_{A,u,b,f}$ that has a smaller objective function value $f(z^0 + \alpha g) < f(z^0)$.

Thus, the directions from $\mathcal{C}(A)$ and $\mathcal{G}(A)$ allow a simple augmentation procedure that iteratively improves a given feasible solution to optimality. While this augmentation process has to terminate for bounded IPs, it may show some zig-zagging behaviour, even to non-optimal solutions for LPs [8]:

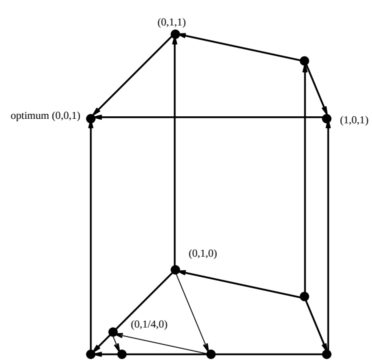
Example 3 Consider the problem

$$\min\{z_1 + z_2 - z_3 : 2z_1 + z_2 \leq 2, z_1 + 2z_2 \leq 2, z_3 \leq 1, (z_1, z_2, z_3) \in \mathbb{R}^3_{\geq 0}\}$$

with optimal solution $(0, 0, 1)$. Introducing slack variables z_4, z_5, z_6 we obtain the problem $\min\{c^\intercal z : Az = (2, 2, 1)^\intercal, z \in \mathbb{R}^6_{\geq 0}\}$ with $c^\intercal = (1, 1, -1, 0, 0, 0)$ and

$$A = \begin{pmatrix} 2 & 1 & 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

The vectors $(1, 0, 0, -2, -1, 0), (0, 1, 0, -1, -2, 0), (1, -2, 0, 0, 3, 0), (2, -1, 0, -3, 0, 0), (0, 0, 1, 0, 0, -1)$ together with their negatives are the circuits of A. The improving directions are given by all circuits v for which $c^\intercal v > 0$.



Now start with the feasible solution $z_0 = (0,1,0,1,0,1)$. Following the directions $(0,1,0,-1,-2,0)$ and $(0,0,-1,0,0,1)$ as far as possible, we immediately arrive at $(0,0,1,2,2,0)$ which corresponds to the desired optimal solution $(0,0,1)$ of our problem. However, alternatively choosing only the vectors $(-1,2,0,0,-3,0)$ and $(2,-1,0,-3,0,0)$ as improving directions, the augmentation process does not terminate. In our original space $\mathbb{R}^3$, this corresponds to the sequence of movements

$$(0,1,0) \to \left(\frac{1}{2},0,0\right) \to \left(0,\frac{1}{4},0\right) \to \left(\frac{1}{8},0,0\right) \to \left(0,\frac{1}{16},0\right) \to \dots$$

clearly shows the zig-zagging behaviour to the non-optimal point $(0,0,0)$. $\square$

Indeed, in order to avoid zig-zagging, certain conditions on the selection of the potential augmenting circuits must be imposed. As suggested in [8], one can avoid such an undesired convergence

- by first choosing an augmenting circuit direction freely, and
- by then moving only along such circuit directions that do not increase the objective value, that is $c^\intercal g \leq 0$, and which introduce an additional zero component in the current feasible solution, that is $\operatorname{supp}(z^0 + \alpha g) \subsetneq \operatorname{supp}(z^0)$. After $O(n)$ such steps, we have again reached a vertex and may perform a free augmentation step if possible.

A natural question that arises is, whether there are strategies to choose a direction from $\mathcal{C}(A)$ and $\mathcal{G}(A)$, respectively, to augment any given feasible solution of $(\mathrm{LP})_{A,u,b,f}$ or $(\mathrm{IP})_{A,u,b,f}$ to optimality in only polynomially many augmentation steps. In this paper, we answer this question affirmatively. For this let us introduce the notion of a *greedy augmentation vector*.

Definition 4 *Let $\mathcal{F} \subseteq \mathbb{R}^n$ be a set of feasible solutions, $z_0 \in \mathcal{F}$, $f : \mathbb{R}^n \to \mathbb{R}$ any objective function, and let $S \subseteq \mathbb{R}^n$ be a (finite) set of directions. Then we call any optimal solution to*

$$\min\{f(z_0 + \alpha g) : \alpha \in \mathbb{R}_+, g \in S, z_0 + \alpha g \in \mathcal{F}\},$$

a **greedy augmentation vector** *(from S for z_0).*

Theorem 5 *Let $A \in \mathbb{Z}^{d\times n}$, $u \in \mathbb{Q}^n$, $b \in \mathbb{Z}^d$ and $c \in \mathbb{Q}^n$ be given. Moreover, let $f(z) = c^\intercal z$. Then the following two statements hold.*

(a) Any feasible solution z^0 to $(LP)_{A,u,b,f}$ can be augmented to an optimal solution of $(LP)_{A,u,b,f}$ by iteratively applying the following greedy procedure:

1. *Choose a greedy direction αg from $\mathcal{C}(A)$ and set $z^0 := z^0 + \alpha g$.*

   If $\alpha g = 0$, return z_0 as optimal solution.

2. *As long as it is possible, find a circuit direction $g \in \mathcal{C}(A)$ and $\alpha > 0$ such that $z^0 + \alpha g$ is feasible, $c^\intercal(z^0 + \alpha g) \leq c^\intercal z^0$, and $\operatorname{supp}(z^0 + \alpha g) \subsetneq \operatorname{supp}(z^0)$, and set $z^0 := z^0 + \alpha g$.*

   Go back to Step 1.

The number of augmentation steps in this augmentation procedure is polynomially bounded in the encoding lengths of A, u, b, c, and z^0.

(b) Any feasible solution z^0 to $(IP)_{A,u,b,f}$ can be augmented to an optimal solution of $(IP)_{A,u,b,f}$ by iteratively applying the following greedy procedure:

Choose a greedy direction αg from $\mathcal{G}(A)$ and set $z^0 := z^0 + \alpha g$.

If $\alpha g = 0$, return z_0 as optimal solution.

The number of augmentation steps in this augmentation procedure is polynomially bounded in the encoding lengths of A, u, b, c, and z^0.

For our proof of Theorem 5 we refer to Section 5.1. Note that in [4] it was shown that the Graver basis $\mathcal{G}(A)$ allows to design a polynomial time augmentation procedure. This procedure makes use of the oracle equivalence of so-called oriented augmentation and linear optimization established in [15]. However, the choice of the Graver basis element that has to be used as a next augmenting vector using the machanism of [15] is far more technical than our simple greedy strategy suggested by Theorem 5, Part (b).

In this paper, we generalize Part (b) of Theorem 5 to certain $\mathbb{Z}$-convex objective functions. We say that a function $g : \mathbb{Z} \to \mathbb{Z}$ is $\mathbb{Z}$-convex, if for all $x, y \in \mathbb{Z}$ and for all $0 \le \lambda \le 1$ with $\lambda x + (1-\lambda)y \in \mathbb{Z}$, the inequality $g(\lambda x + (1-\lambda)y) \ge \lambda g(x) + (1-\lambda)g(y)$ holds. With this notion of $\mathbb{Z}$-convexity, we generalize Part (b) of Theorem 5 to *nonlinear convex* objectives of the form $f(c^\intercal z, c_1^\intercal z, \ldots, c_s^\intercal z)$, where

$$f(y_0, y_1, \ldots, y_s) = \sum_{i=1}^{s} f_i(y_i) + y_0 \tag{1}$$

is a separable $\mathbb{Z}$-convex function and where $c_0, \ldots, c_s \in \mathbb{Z}^n$ are given fixed vectors. In particular, each function $f_i : \mathbb{Z} \to \mathbb{Z}$ is $\mathbb{Z}$-convex. When all $f_i \equiv 0$, we recover linear integer optimization as a special case. To state our result, let C denote the $s \times n$ matrix with rows $c_1, \ldots, c_s$ and let $\mathcal{G}(A, C)$ denote the Graver basis of $\left(\begin{smallmatrix} A & 0 \\ C & I_s \end{smallmatrix}\right)$ projected onto the first n variables. As was shown in [9, 13], this finite set provides an improving direction for any non-optimal solution z^0 of $(\mathrm{IP})_{A,u,b,f}$.

Theorem 6 *Let $A \in \mathbb{Z}^{d\times n}$, $u \in \mathbb{Z}^n$, $b \in \mathbb{Z}^d$, $c \in \mathbb{Q}^n$, $c_1, \ldots, c_s \in \mathbb{Q}^n$. Moreover, let $\bar{f}(z) := f(c^\intercal z, c_1^\intercal z, \ldots, c_s^\intercal z)$, where f denotes a separable $\mathbb{Z}$-convex function as in (1) given by a polynomial time comparison oracle which, when queried on $x, y \in \mathbb{Z}^{s+1}$, decides whether $f(x) < f(y)$, $f(x) = f(y)$, or $f(x) > f(y)$ holds in time polynomial in the encoding lengths of x and y. Moreover, let H be an upper bound for the difference of maximum and minimum value of $\bar{f}$ over the feasible set $\{z : Az = b, 0 \le z \le u, z \in \mathbb{Z}_+^n\}$ and assume that the encoding length of H is of polynomial size in the encoding lengths of $A, u, b, c, c_1, \ldots, c_s$. Then the following statement holds.*

Any feasible solution z^0 to $(IP)_{A,u,b,\bar{f}}$ can be augmented to an optimal solution of $(IP)_{A,u,b,\bar{f}}$ by iteratively applying the following greedy procedure:

Choose a greedy direction αg from $\mathcal{G}(A, C)$ and set $z^0 := z^0 + \alpha g$.

If $\alpha g = 0$, return z_0 as optimal solution.

The number of augmentation steps in this augmentation procedure is polynomially bounded in the encoding lengths of A, u, b, $c, c_1, \ldots, c_s$, and z^0.

For our proof of Theorem 6 we refer to Section 5.2. As a consequence to Theorem 6, we construct in Sections 2 and 3 polynomial time algorithms to solve convex N-fold integer minimization problems and convex 2-stage stochastic integer minimization problems. In the first case, the Graver basis under consideration is of polynomial size in the input data and hence the greedy augmentation vector αg can be found in polynomial time. In the second case, the Graver basis is usually of exponential size in the input data. Despite this fact, the desired greedy augmentation vector αg can be constructed in polynomial time, if the f_i are convex *polynomial* functions. Finally, we present some applications of convex N-fold integer minimization problems for which our approach provides a polynomial time solution algorithm. We conclude the paper with our proofs of Theorems 5 and 6.

2 N-fold convex integer minimization

Let $A \in \mathbb{Z}^{d_a \times n}$, $B \in \mathbb{Z}^{d_b \times n}$, and $c_1, \ldots, c_s \in \mathbb{Z}^n$ be fixed and consider the problem

$$\min \left\{ \sum_{i=1}^{N} f^{(i)}\left(x^{(i)}\right) : \sum_{i=1}^{N} Bx^{(i)} = b^{(0)}, Ax^{(i)} = b^{(i)}, 0 \leq x^{(i)} \leq u^{(i)}, x^{(i)} \in \mathbb{Z}^n, i = 1, \ldots, N \right\},$$

where we have

$$f^{(i)}(z) := \sum_{j=1}^{s} f_j^{(i)}\left(c_j^\intercal z\right) + c^{(i)\intercal} z$$

for given convex functions $f_j^{(i)}$ and vectors $c^{(i)} \in \mathbb{Z}^n$, $i = 1, \ldots, N$, $j = 1, \ldots, s$. If we dropped the coupling constraint $\sum_{i=1}^{N} Bx^{(i)} = b^{(0)}$, this optimization problem would decompose into N simpler convex problems

$$\min \left\{ f^{(i)}\left(x^{(i)}\right) : Ax^{(i)} = b^{(i)}, 0 \leq x^{(i)} \leq u^{(i)}, x^{(i)} \in \mathbb{Z}^n \right\}, i = 1, \ldots, N,$$

which could be solved independently. Hence the name "N-fold convex integer program".

Definition 7 *The* N**-fold matrix of the ordered pair** A, B *is the following* $(d_b + Nd_a) \times Nn$ *matrix,*

$$[A, B]^{(N)} := \begin{pmatrix} B & B & B & \cdots & B \\ A & 0 & 0 & \cdots & 0 \\ 0 & A & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A \end{pmatrix}.$$

For any vector $x = (x^1, \ldots, x^N)$ with $x^i \in \mathbb{Z}^n$ for $i = 1, \ldots, N$, we call the number $|\{i : x^i \neq 0\}|$ of nonzero building blocks $x^i \in \mathbb{Z}^n$ of x the **type** *of x.*

In [11], it was shown that there exists a constant $g(A,B)$ such that for all N the types of the Graver basis elements in $\mathcal{G}([A,B]^{(N)})$ are bounded by $g(A,B)$. In [4], this was exploited to solve *linear* N-fold IP in polynomial time.

Lemma 8 (Results from [4])

- *For fixed matrices A and B the sizes of the Graver bases $\mathcal{G}([A,B]^{(N)})$ increase only polynomially in N.*
- *For any choice of the right-hand side vector b, an initial feasible solution z_0 can be constructed in time polynomial in N and in the encoding length of b.*
- *For any linear objective function $c^\intercal z$, this solution z_0 can be augmented to optimality in time polynomial in N and in the encoding lengths of b, c, u, and z_0.*

Using Theorem 6, we can now generalize this polynomial time algorithm to convex objectives of the form above. Let us prepare the main result of this section by showing that the encoding lengths of Graver bases from [9, 13] increase only polynomially in N. For this, let C denotes the $s \times n$ matrix with rows $c_1, \ldots, c_s$.

Lemma 9 *Let the matrices $A \in \mathbb{Z}^{d_a \times n}$, $B \in \mathbb{Z}^{d_b \times n}$, and $C \in \mathbb{Z}^{s \times n}$ be fixed. Then the encoding lengths of the Graver bases of*

$$([A,B],C)^{(N)} := \left(\begin{array}{cccc|cccc} B & B & \cdots & B & & & & \\ A & & & & & & & \\ & A & & & & & & \\ & & \ddots & & & & & \\ & & & A & & & & \\ \hline C & & & & I_s & & & \\ & C & & & & I_s & & \\ & & \ddots & & & & \vdots & \\ & & & C & & & & I_s \end{array}\right).$$

increase only polynomially in N.

Proof. The claim follows from the results in [4] by rearranging the rows and columns as follows

$$([A,B],C)^{(N)} := \left(\begin{array}{cccccccc} B & 0 & B & 0 & \cdots & & B & 0 \\ A & 0 & & & & & & \\ C & I_s & & & & & & \\ & & A & 0 & & & & \\ & & C & I_s & & & & \\ & & & \ddots & \ddots & & & \\ & & & & & & A & 0 \\ & & & & & & C & I_s \end{array}\right).$$

This is the matrix of an N-fold IP with $\bar{A} = \left(\begin{smallmatrix} A & 0 \\ C & I_s \end{smallmatrix}\right)$ and with $\bar{B} = \left(\begin{smallmatrix} B & 0 \end{smallmatrix}\right)$. Hence, the sizes and the encoding lengths of the Graver bases increase only polynomially in N. □

Now that we have shown that the Graver basis is of polynomial size, we can consider each Graver basis element g independently and search for the best $\alpha \in \mathbb{Z}_+$ such that $z_0 + \alpha g$ is feasible and has a smallest objective value. This can be done in polynomial time as the following lemma shows.

Lemma 10 *Let $f : \mathbb{R} \to \mathbb{R}$ be a convex function given by a comparision oracle. Then for any given numbers $l, u \in \mathbb{Z}$, the one-dimensional minimization problem $\min\{f(\alpha) : l \leq \alpha \leq u\}$ can be solved by polynomially many calls to the comparision oracle.*

Proof. If the interval $[l, u]$ contains at most 2 integers, return l or u as the minimum, depending on the values of $f(l)$ and $f(u)$. If the interval $[l, u]$ contains at least 3 integers, consider the integers $\lfloor (l+u)/2 \rfloor - 1$, $\lfloor (l+u)/2 \rfloor$, $\lfloor (l+u)/2 \rfloor + 1 \in [l, u]$ and exploit convexity of f to bisect the interval $[l, u]$ as follows:

If $f(\lfloor (l+u)/2 \rfloor - 1) < f(\lfloor (l+u)/2 \rfloor)$ holds, then the minimum of f must be attained in the interval $[l, \lfloor (l+u)/2 \rfloor]$. If, on the other hand, $f(\lfloor (l+u)/2 \rfloor) > f(\lfloor (l+u)/2 \rfloor + 1)$, then the minimum of f must be attained in the interval $[\lfloor (l+u)/2 \rfloor + 1, u]$. If none of the two holds, the minimum of f is attained in the point $\alpha = \lfloor (l+u)/2 \rfloor$.

Clearly, after $O(\log(u - l))$ bisection steps, the minimization problem is solved. □

The results in [4] together with the previous two lemmas now immediately imply the main result of this section.

Theorem 11 *Let A, B, C be fixed integer matrices of appropriate dimensions. Then the following holds. Moreover, let $f_j^{(i)} : \mathbb{R} \to \mathbb{R}$ be convex functions mapping $\mathbb{Z}$ to $\mathbb{Z}$ given by polynomial time evaluation oracles. Then the problem*

$$\min \left\{ \sum_{i=1}^{N} f^{(i)} \left(x^{(i)} \right) : \sum_{i=1}^{N} B x^{(i)} = b^{(0)}, A x^{(i)} = b^{(i)}, 0 \leq x^{(i)} \leq u^{(i)}, x^{(i)} \in \mathbb{Z}^n, i = 1, \ldots, N \right\},$$

can be solved in time polynomial in the encoding length of the input data.

Proof. Polynomial time construction of an initial feasible solution from which we can start our augmentation process follows immediately from the results in [4].

To show that this feasible solution can be augmented to optimality in polynomial time, we note that by Theorem 6 that only polynomially many greedy augmentation steps are needed. By Lemma 9, we only need to check polynomially many directions g to search for a greedy augmentation vector. But this can be done in polynomial time by Lemma 10. □

3 Convex 2-stage stochastic integer minimization

Multistage stochastic integer programming has become an important field of optimization, see [3, 12, 14] for details. From a mathematical point of view, the data describing a 2-stage stochastic integer program is as follows. Let $T \in \mathbb{Z}^{d\times m}$, $W \in \mathbb{Z}^{d\times n}$, $c_1, \ldots, c_s \in \mathbb{Z}^m$, $d_1, \ldots, d_s \in \mathbb{Z}^n$ be fixed, and consider the problem

$$\min\{\mathbb{E}_\omega(f^\omega(x,y)) : Tx + Wy = b^\omega, 0 \leq x \leq u_x, 0 \leq y \leq u_y, x \in \mathbb{Z}^m, y \in \mathbb{Z}^n\},$$

where ω is some probability distribution in a suitable probability space and where f is a convex function of the form

$$f^\omega(x,y) := \sum_{j=1}^{s} f_j^\omega\left(c_j^\intercal x + d_j^\intercal y\right)$$

in which each $f_j^\omega : \mathbb{R} \to \mathbb{R}$ is a convex function.

Discretizing the probability distribution using N scenarios, we obtain the following convex integer minimization problem

$$\min\left\{\sum_{i=1}^{N} f^{(i)}\left(x, y^{(i)}\right) : Tx + Wy^{(i)} = b^{(i)}, 0 \leq x \leq u_x, 0 \leq y^{(i)} \leq u_y^{(i)}, x \in \mathbb{Z}^m, y^{(i)} \in \mathbb{Z}^n, i = 1, \ldots, N\right\},$$

where we have

$$f^{(i)}(x,y) := \sum_{j=1}^{s} f_j^{(i)}\left(c_j^\intercal x + d_j^\intercal y\right)$$

for given convex functions $f_j^{(i)}$. Note that fixing the first-stage decision x would decompose the optimization problem into N simpler convex problems

$$\min\left\{f^{(i)}\left(x, y^{(i)}\right) : Ax^{(i)} = b^{(i)}, 0 \leq y^{(i)} \leq u_y^{(i)}, y^{(i)} \in \mathbb{Z}^n\right\}, i = 1, \ldots, N,$$

which could be solved independently. However, the problem of finding a first-stage decision x with smallest overall costs would still remain to be solved.

Lemma 12 (Results from [10])

- *A vector $(v, w_1, \ldots, w_N)$ is in the kernel of the matrix*

$$\overline{[T,W]}^{(N)} := \begin{pmatrix} T & W & 0 & \cdots & 0 \\ T & 0 & W & \cdots & 0 \\ \vdots & & & \ddots & \\ T & 0 & 0 & \cdots & W \end{pmatrix}$$

  if and only if $(v, w_i) \in \ker\left(\overline{[T,W]}^{(1)}\right)$ for all i, that is, if $Tv + Ww_i = 0$ for all i.

- *The Graver bases for the matrices $\overline{[T,W]}^{(N)}$ decompose into a finite number of first-stage and second-stage building blocks that are independent on N.*

- *For any given* linear *objective, any given right-hand side vector and any non-optimal feasible solution z_0, an improving vector to z_0 can be reconstructed from the building blocks in time linear in the number N of scenarios.*

Note that this finiteness result from [10] does not imply that the Graver basis of $\overline{[T,W]}^{(N)}$ is of polynomial size in N. In fact, one can easily construct an exponential size counter-example. Before we present the main result of this section, let show that there exists a polynomial time optimality certificate also for convex 2-stage stochastic integer minimization problems of the type above, if the matrices T and W are kept fix. For this, let C denote the $s\times m$ matrix with rows $c_1,\ldots,c_s$, and let D denote the $s\times n$ matrix with rows $d_1,\ldots,d_s$.

Lemma 13 *The Graver bases of the matrices*

$$\overline{[T,W,C,D]}^{(N)} := \begin{pmatrix} T & W & & & & & & \\ T & & W & & & & & \\ \vdots & & & \ddots & & & & \\ T & & & & W & & & \\ C & D & & & & I_s & & \\ C & & D & & & & I_s & \\ \vdots & & & \ddots & & & & \ddots \\ C & & & & D & & & & I_s \end{pmatrix}$$

decompose into a finite number of first-stage and second-stage building blocks that are independent on N.

For any given convex *objective, any given right-hand side vector and any non-optimal feasible solution z_0, an improving vector to z_0 can be reconstructed from the building blocks in time linear in the number N of scenarios.*

Proof. To prove our first claim, we rearrange blocks within the matrix $\overline{[T,W,C,D]}^{(N)}$ as follows:

$$\begin{pmatrix} T & W & 0 & & & & & \\ C & D & I_s & & & & & \\ T & & & W & 0 & & & \\ C & & & D & I_s & & & \\ \vdots & & & & & \ddots & & \\ T & & & & & & W & 0 \\ C & & & & & & D & I_s \end{pmatrix} = \overline{\left[\begin{pmatrix} T \\ C \end{pmatrix}, \begin{pmatrix} W & 0 \\ D & I_s \end{pmatrix}\right]}^{(N)},$$

which is the matrix of a 2-stage stochastic integer program with N scenarios and fixed matrices $\left(\begin{smallmatrix} T \\ C \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} W & 0 \\ D & I_s \end{smallmatrix}\right)$. Hence, its Graver basis consists out of a constant number of building blocks independent on N. This proves the first claim.

To prove the second claim, note that the results from [9, 13] show that the Graver basis of $\overline{[T,W,C,D]}^{(N)}$ projected down onto the variables corresponding to T and W columns gives improving directions for non-optimal solutions z_0 to

$$\min\left\{\sum_{i=1}^{N} f^{(i)}\left(x, y^{(i)}\right) : Tx + Wy^{(i)} = b^{(i)}, 0 \leq x \leq u_x, 0 \leq y \leq u_y, x \in \mathbb{Z}^m, y \in \mathbb{Z}^n, i = 1, \ldots, N\right\}.$$

Thus, these directions consist out of only a constant number of building blocks independent on N. Let $z = (x, y^{(1)}, \ldots, y^{(N)})$ be a feasible solution and let $g = (v, w^{(1)}, \ldots, w^{(N)})$ be an augmenting vector formed out of the constant number of first-stage and second-stage building blocks. To be an improving direction, g must satisfy the following constraints:

- $T(x + v) + W(y^{(i)} + w^{(i)}) = b^{(i)}$, $i = 1, \ldots, N$,
- $0 \leq x + v \leq u_x$,
- $0 \leq y^{(i)} + w^{(i)} \leq u_y^{(i)}$, $i = 1, \ldots, N$,
- $\sum_{i=1}^{N} f^{(i)}\left(x + y, y^{(i)} + w^{(i)}\right) < \sum_{i=1}^{N} f^{(i)}\left(x, y^{(i)}\right)$.

For each of the finitely many first-stage building blocks perform the following test: If $0 \leq x + v \leq u_x$, try to find suitable second-stage building blocks satisfying the remaining constraints, which for fixed v simplify to

- $Tv + Ww^{(i)} = 0$, $i = 1, \ldots, N$,
- $0 \leq y^{(i)} + w^{(i)} \leq u_y^{(i)}$, $i = 1, \ldots, N$,
- $\sum_{i=1}^{N} f^{(i)}\left(x + v, y^{(i)} + w^{(i)}\right) < \sum_{i=1}^{N} f^{(i)}\left(x, y^{(i)}\right)$.

For fixed v, this problem decomposes into N independent minimization problems:

$$\min\left\{f^{(i)}\left(x + v, y^{(i)} + w^{(i)}\right) : Tv + Ww^{(i)} = 0, 0 \leq y^{(i)} + w^{(i)} \leq u_y^{(i)}\right\}, i = 1, \ldots, N.$$

If for those optimal values $\sum_{i=1}^{N} f^{(i)}\left(x + v, y^{(i)} + w^{(i)}\right) < \sum_{i=1}^{N} f^{(i)}\left(x, y^{(i)}\right)$ holds, we have found an improving vector $g = (v, w^{(1)}, \ldots, w^{(N)})$ for z_0. If one of these minimization problems is infeasible or if $\sum_{i=1}^{N} f^{(i)}\left(x + v, y^{(i)} + w^{(i)}\right) \geq \sum_{i=1}^{N} f^{(i)}\left(x, y^{(i)}\right)$, then no augmenting vector for z_0 can be constructed using the first-stage building block v. If for no first-stage building block v, an augmenting vector can be constructed z_0 must be optimal. If there was an augmenting vector for z_0 with some first-stage building block v, this vector or even a better augmenting vector would have been constructed by the procedure above when the first-stage building block v was considered. $\square$

Note that the augmenting vector constructed in the proof of the previous lemma need not be a Graver basis element (it may not be minimal), but every Graver basis element could be constructed, guaranteeing the optimality certificate. It remains to show how to construct a *greedy* augmentation vector from the building blocks from the Graver basis. Note that the procedure in the previous proof constructs an augmenting vector also for a *fixed* step length α. To compute a greedy augmentation vector, however, one has to allow α to vary. But then, the minimization problem does not decompose into N independent simpler problems. It is this difficulty that enforces us to restrict the set of possible convex functions.

Definition 14 *We call a convex function $f : \mathbb{R}^{m+n} \to \mathbb{R}$ that maps $\mathbb{Z}^{m+n}$ to $\mathbb{Z}$* **splittable**, *if for all fixed vectors $x \in \mathbb{Z}^m$, $y, g_1, g_2 \in \mathbb{Z}^n$, and for all finite intervals $[l, u] \subseteq \mathbb{R}$, there exists polynomially many (in the encoding length of the problem data) intervals $I_1, \ldots, I_r$ such that*

- $[l, u] = \bigcup\limits_{i=1}^{r} I_r$,
- $I_i \cap I_j \cap \mathbb{Z} = \emptyset$ *for all* $1 \leq i < j \leq r$, *and*
- *for each $j = 1, \ldots, r$, either $f(x, y + \alpha g_1) \leq f(x, y + \alpha g_2)$ or $f(x, y + \alpha g_1) \geq f(x, y + \alpha g_2)$ holds for all $\alpha \in I_j$.*

Note that convex polynomials of fixed maximal degree k are splittable, as $f(x, y+\alpha g_1) - f(x, y+\alpha g_2)$ switches its sign at most k times. Hence each interval $[l, u]$ can be split into at most $k+1$ intervals with the desired property. With the notion of splittable convex functions, we can now state and prove the main theorem of this section.

Theorem 15 *Let T, W, C, D be fixed integer matrices of appropriate dimensions. Then the following holds.*

(a) For any choice of the right-hand side vector b, an initial feasible solution z_0 to

$$\min \left\{ \sum_{i=1}^{N} f^{(i)} \left(x, y^{(i)} \right) : Tx + Wy^{(i)} = b^{(i)}, 0 \leq x \leq u_x, 0 \leq y^{(i)} \leq u_y^{(i)}, x \in \mathbb{Z}^m, y^{(i)} \in \mathbb{Z}^n, i = 1, \ldots, N \right\},$$

can be constructed in time polynomial in N and in the encoding length of the input data.

(b) Then, for any choice of splittable convex functions $f^{(i)}$, this solution z_0 can be augmented to optimality in time polynomial in the encoding length of the input data.

Proof. Let us prove Part (b) first. This proof follows the main idea behind the proof of Lemma 13. Let $z = (x, y^{(1)}, \ldots, y^{(N)})$ be a feasible solution and let $g = (v, w^{(1)}, \ldots, w^{(N)})$ be an augmenting vector formed out of the constant number of first-stage and second-stage building blocks. Again, for fixed v, we wish to consider each scenarios independently. For this, note that the possible step length $\alpha \in \mathbb{Z}_+$ is bounded from above by some polynomial size bound u_α, since our feasible region is a polytope. Since the convex functions $f^{(i)}$ are splittable, we can for each scenario partition the interval $[0, u_\alpha]$ into polynomially subintervals $I_{i,1}, \ldots, I_{i,r_i}$ such that for each interval $I_{i,j}$ there is either no building block leading to a feasible solution or a well-defined building block $w_{i,j}$ with $Tv + Ww_{i,j} = 0$ and $0 \leq y^{(i)} + \alpha w_{i,j} \leq u^{(i)}$ that minimizes $f^{(i)}(x + v, y^{(i)} + \alpha w_{i,j})$ for all $\alpha \in I_{i,j}$.

Taking the common refinement of all intervals $I_{i,j}$, $i = 1, \ldots, N$, $j = 1, \ldots, r_i$, one obtains polynomially many intervals $J_1, \ldots, J_t$, such that for each interval J_i and for all $\alpha \in J_i$, there is a well-defined building block for each scenario minimizing the function value. For this fixed vector $g = (v, w^{(1)}, \ldots, w^{(N)})$ we then compute the best $\alpha \in J_i$, and then compare these values $\sum_{i=1}^{N} f^{(i)} \left(x + \alpha v, y^{(i)} + \alpha w^{(i)} \right)$ to find the desired greedy augmentation vector. Applying Theorem 6, this proves Part (b).

Finally, let us prove Part (a). For this, introduce nonnegative integer slack-variables into the second-stages to obtain a linear IP with problem matrix

$$\overline{[T,(W,I_d,-I_d)]}^{(N)} := \begin{pmatrix} T & W & I_d & -I_d & 0 & 0 & 0 & \cdots & 0 & & \\ T & 0 & 0 & 0 & W & I_d & -I_d & \cdots & 0 & & \\ \vdots & & & & & & & \ddots & & & \\ T & 0 & 0 & 0 & 0 & 0 & 0 & \cdots & W & I_d & -I_d \end{pmatrix}$$

whose associated Graver basis is formed out of only constantly many first- and second-stage building blocks. Using this extended formulation, we may immediately write down a feasible solution. Using only greedy directions from the Graver basis of $\overline{[T,(W,I_d,-I_d)]}^{(N)}$, we can minimize the sum of all slack-variables in polynomially many augmentation steps. Part (b) now implies that an optimal solution to this extended problem can be found in polynomial time. If all slack-variables are 0, we have found a feasible solution to our intial problem, otherwise the initial problem is infeasible. $\square$

Let us conclude with the remark that these polynomiality results for convex 2-stage stochastic integer minimization can be extended to the multi-stage situation by applying the finiteness results from [2].

4 Some Applications

Consider the following general nonlinear problems over an arbitrary set $\mathcal{F} \subseteq \mathbb{Z}^n$ of feasible solutions:

Separable convex minimization: Find a feasible point $x \in \mathcal{F}$ minimizing a *separable* convex cost function $f(x) := \sum_{i=1}^n f_i(x_i)$ with each f_i a univariate convex function. It generalizes standard linear optimization with cost $f(x) = \sum_{i=1}^n c_i^\intercal x_i$ recovered with $f_i(x_i) := c_i^\intercal x_i$ for some costs c_i.

Minimum l_p-distance: Find a feasible point $x \in \mathcal{F}$ minimizing the l_p-distance to a partially specified "goal" point $\bar{x} \in \mathbb{Z}^n$. More precisely, given $1 \leq p \leq \infty$ and the restriction $\bar{x}_I := (\bar{x}_i : i \in I)$ of $\bar{x}$ to a subset $I \subseteq \{1,\ldots,n\}$ of the coordinates, find $x \in \mathcal{F}$ minimizing the l_p-distance $\|x_I - \bar{x}_I\|_p := (\sum_{i\in I} |x_i - \bar{x}_i|^p)^{\frac{1}{p}}$ for $1 \leq p < \infty$ and $\|x_I - \bar{x}_I\|_\infty := \max_{i\in I} |x_i - \bar{x}_i|$ for $p = \infty$.

Note that a common special case of the above is the natural problem of l_p-norm minimization over $\mathcal{F}$, $\min\{\|x\|_p : x \in \mathcal{F}\}$; in particular, the l_∞-norm minimization problem is the min-max problem $\min\{\max_{i=1}^n |x_i| : x \in \mathcal{F}\}$.

In our discussion of N-fold systems below it will be convenient to index the variable vector as $x = (x^1,\ldots,x^N)$ with each block indexed as $x^i = (x_{i,1},\ldots,x_{i,n})$, $i = 1,\ldots,N$.

We have the following corollary of Theorem 11, which will be used in the applications to follow.

Corollary 16 *Let A and B be fixed integer matrices of compatible sizes. Then there is an algorithm that, given any positive integer N, right-hand sides b^i, and upper bound vectors u^i, of suitable dimensions, solves the above problems over the following set of integer points in an N-fold program*

$$\mathcal{F} \;=\; \{x = (x^1, \dots, x^N) \in \mathbb{Z}^{N\times n} \;:\; \sum_{i=1}^{N} Bx^i = b^0,\; Ax^i = b^i,\; 0 \le x^i \le u^i,\; i = 1, \dots, N\} \qquad (2)$$

in time which is polynomial in N and in the binary encoding length of the rest of the input, as follows:

1. *For $i = 1, \dots, N$ and $j = 1, \dots, n$, let $f_{i,j}$ denote convex univariate functions. Moreover, let $f(x) := \sum_{i=1}^{N}\sum_{j=1}^{n} f_{i,j}(x_{i,j})$ be given by a comparision oracle. Then the algorithm solves the separable convex minimization problem*

$$\min\left\{\sum_{i=1}^{N}\sum_{j=1}^{n} f_{i,j}(x_{i,j}) \;:\; x \in \mathcal{F}\right\} .$$

2. *Given any $I \subseteq \{1, \dots, N\} \times \{1, \dots, n\}$, any partially specified integer point $\bar{x}_I := (x_{i,j} : (i,j) \in I)$, and any integer $1 \le p < \infty$ or $p = \infty$, the algorithm solves the minimum l_p-distance problem*

$$\min\{\|x_I - \bar{x}_I\|_p \;:\; x \in \mathcal{F}\} .$$

In particular, the algorithm solves the l_p-norm minimization problem $\min\{\|x\|_p : x \in \mathcal{F}\}$.

Proof. Consider first the separable convex minimization problem. Then this is just the special case of Theorem 11 with $c_j := \mathbf{1}_j$ the standard j-th unit vector in $\mathbb{Z}^n$ for $j = 1, \dots, n$ and $c^i := 0$ in $\mathbb{Z}^n$ for $i = 1, \dots, N$. The objective function in Theorem 11 then becomes the desired objective,

$$\sum_{i=1}^{N} f^i(x^i) \;=\; \sum_{i=1}^{N}\sum_{j=1}^{n} f_{i,j}(c_j^{\intercal} x^i) + c^i x^i \;=\; \sum_{i=1}^{N}\sum_{j=1}^{n} f_{i,j}(x_{i,j}) .$$

Next consider the minimum l_P-distance problem. Consider first an integer $1 \le p < \infty$. Then we can minimize the integer-valued p-th power $\|x\|_p^p$ instead of the l_p-norm itself. Define

$$f_{i,j}(x_{i,j}) \;:=\; \begin{cases} |x_{i,j} - \bar{x}_{i,j}|^p, & \text{if } (i,j) \in I; \\ 0, & \text{otherwsie.} \end{cases}$$

With these $f_{i,j}$, the objective in the separable convex minimization becomes the desired objective,

$$\sum_{i=1}^{N}\sum_{j=1}^{n} f_{i,j}(x_{i,j}) \;=\; \sum_{(i,j)\in I} |x_{i,j} - \bar{x}_{i,j}|^p \;=\; |x_I - \bar{x}_I|_p^p .$$

Next, consider the case $p = \infty$. Let $w := \max\{|u_{i,j}| : i = 1, \dots, N, j = 1, \dots, n\}$ be the maximum upper bound on any variable. We may assume $w > 0$ else $\mathcal{F} \subseteq \{0\}$ and the integer program is trivial. Choose a positive integer q satisfying $q\log(1+(2w)^{-1}) > \log(Nn)$. Now solve the minimum

l_q-distance problem and let $x^* \in \mathcal{F}$ be an optimal solution. We claim that x^* also minimizes the l_∞-distance to $\bar{x}$. Consider any $x \in \mathcal{F}$. By standard inequalities between the l_∞ and l_q norms,

$$\|x_I^* - \bar{x}_I\|_\infty \ \leq \ \|x_I^* - \bar{x}_I\|_q \ \leq \ \|x_I - \bar{x}_I\|_q \ \leq \ (Nn)^{\frac{1}{q}} \|x_I - \bar{x}_I\|_\infty \ .$$

Therefore

$$\|x_I^* - \bar{x}_I\|_\infty - \|x_I - \bar{x}_I\|_\infty \ \leq \ ((Nn)^{\frac{1}{q}} - 1)\|x_I - \bar{x}_I\|_\infty \ \leq \ ((Nn)^{\frac{1}{q}} - 1)2w \ < \ 1$$

where the last inequality holds by the choice of q. Since $\|x_I^* - \bar{x}_I\|_\infty$ and $\|x_I - \bar{x}_I\|_\infty$ are integers we find that indeed $\|x_I^* - \bar{x}_I\|_\infty \leq \|x_I - \bar{x}_I\|_\infty$ holds for all $x \in \mathcal{F}$ and the claim follows. □

4.1 Congestion-avoiding (multi-way) transportation and routing

The classical (discrete) transportation problem is the following. We wish to transport commodities (in containers or bins) on a traffic network (by land, sea or air), or route information (in packets) on a communication network, from n suppliers to N customers. The demand by customer i is d_i units and the supply from supplier j is s_j units. We need to determine the number $x_{i,j}$ of units to transport to customer i from supplier j on channel $i \leftarrow j$ subject to supply-demand requirements and upper bounds $x_{i,j} \leq u_{i,j}$ on channel capacity so as to minimize total delay or cost. The classical approach assumes a channel cost $c_{i,j}$ per unit flow, resulting in linear total cost $\sum_{i=1}^N \sum_{j=1}^n c_{i,j} x_{i,j}$. But due to channel congestion when subject to heavy traffic or heavy communication load, the transportation delay or cost on a channel are actually a nonlinear convex function of the flow over it, such as $f_{i,j}(x_{i,j}) = c_{i,j}|x_{i,j}|^{\alpha_{i,j}}$ for suitable $\alpha_{i,j} > 1$, resulting in nonlinear total cost $\sum_{i,j} f_{i,j}(x_{i,j})$, which is much harder to minimize.

It is often natural that the number of suppliers is small and fixed while the number of customers is very large. Then the transportation problem is an N-fold integer programming problem. To see this, index the variable vector as $x = (x^1, \ldots, x^N)$ with $x^i = (x_{i,1}, \ldots, x_{i,n})$ and likewise for the upper bound vector. Let $b^i := d_i$ for $i = 1, \ldots, N$ and let $b^0 := (s_1, \ldots, s_n)$. Finally, let $A = (1, \ldots, 1)$ be the $1 \times n$ matrix with all entries equal to 1 and let B be the $n \times n$ identity matrix. Then the N-fold constraints $Ax^i = b^i$, $i = 1, \ldots, N$ and $B(\sum_{i=1}^N x^i) = b^0$ represent, respectively the demand and supply constraints. The feasible set in (2) then consists of the feasible transportations and the solution of the congestion-avoiding transportation problem is provided by Corollary 16 part 1. So we have:

Corollary 17 *Fix the number of suppliers and let $f_{i,j}$, $i = 1, \ldots, N$, $j = 1, \ldots, n$, denote convex univariate functions. Moreover, let $f(x) := \sum_{i=1}^N \sum_{j=1}^n f_{i,j}(x_{i,j})$ be given by a comparision oracle. Then the congestion-avoiding transportation problem can be solved in polynomial time.*

This result can be extended to multi-way (high-dimensional) transportation problems as well. In the 3-*way line-sum transportation problem*, the set of feasible solutions consists of all nonnegative integer $L \times M \times N$ arrays with specified line-sums and upper bound (capacity) constraints,

$$\mathcal{F} \ := \ \{x \in \mathbb{Z}^{L \times M \times N} \ : \ \sum_i x_{i,j,k} = r_{j,k} \, , \ \sum_j x_{i,j,k} = s_{i,k} \, , \ \sum_k x_{i,j,k} = t_{i,j} \, , \ 0 \leq x_{i,j,k} \leq u_{i,j,k} \} \ . \tag{3}$$

If at least two of the array-size parameters L, M, N are variable then even the classical linear optimization problem over $\mathcal{F}$ is NP-hard [5]. In fact, remarkably, *every* integer program *is* a $3 \times M \times N$ transportation program for some M and N [6]. But when both L and M are relatively small and fixed, the resulting problem over "long" arrays, with a large and variable number N of layers, is again an N-fold program. To see this, index the variable array as $x = (x^1, \ldots, x^N)$ with $x^i = (x_{1,1,i}, \ldots, x_{L,M,i})$ and likewise for the upper bound vector. Let A be the $(L+M) \times LM$ incidence matrix of the complete bipartite graph $K_{L,M}$ and let B be the $LM \times LM$ identity matrix. Finally, suitably define the right-hand side vectors b^h, $h = 0, \ldots, N$ in terms of the given line sums $r_{j,k}$, $s_{i,k}$, and $t_{i,j}$. Then the n-fold constraint $B(\sum_{h=1}^N x^h) = b^0$ represents the line-sum constraints where summation over layers occurs, whereas $Ax^h = b^h$, $h = 1, \ldots, N$, represent the line-sum constraints where summations are within a single layer at a time. Then we can minimize in polynomial time any separable convex cost function $\sum_{i=1}^L \sum_{j=1}^M \sum_{k=1}^N f_{i,j,k}(x_{i,j,k})$ over the set of feasible transportations $\mathcal{F}$ in (3). So we have:

Corollary 18 *Fix any L and M and let $f_{i,j,k}$, $i = 1, \ldots, L$, $j = 1, \ldots, M$, $k = 1, \ldots, N$, denote convex univariate functions. Moreover, let $f(x) := \sum_{i=1}^L \sum_{j=1}^M \sum_{k=1}^N f_{i,j,k}(x_{i,j,k})$ be given by a comparision oracle. Then the congestion-avoiding 3-way transportation problem can be solved in polynomial time.*

Even more generally, this result holds for "long" d-way transportations of any fixed dimension d and for any *hierarchical* sum constraints, see Section 4.3 below for the precise definitions.

4.2 Error-correcting codes

Linear-algebraic error correcting codes generalize the "check-sum" idea as follows: a message to be communicated on a noisy channel is arranged in a vector x. To allow for error correction, several sums of subsets of entries of x are communicated as well. Multi-way tables provide an appealing way of organizing the check-sum protocol. The sender arranges the message in a multi-way $M_1 \times \cdots \times M_d$ array x and sends it along with the sums of some of its lower dimensional sub-arrays (margins). The receiver obtains an array $\bar{x}$ with some entries distorted on the way; it then finds an array $\hat{x}$ having the specified check-sums (margins), that is l_p-closest to the received distorted array $\bar{x}$, and declares it as the retrieved message. For instance, when working over the $\{0, 1\}$ alphabet, the useful *Hamming distance* is precisely the l_1-distance. Note that the check-sums might be distorted as well; to overcome this difficulty, we determine ahead of time an upper bound U on all possible check-sums, and make it a fixed part of the communication protocol; then we blow each array to size $(M_1+1) \times \cdots \times (M_d+1)$, and fill in the new "slack" entries so as to sum up with the original entries to U.

To illustrate, consider 3-way arrays of format $L \times M \times N$ (already augmented with slack variables). Working over alphabet $\{0, \ldots, u\}$, define upper bounds $u_{i,j,k} := u$ for original message variables and $u_{i,j,k} := U$ for slack variable. Then the set of possible messages that the receiver has to choose from is

$$\mathcal{F} \ := \ \{x \in \mathbb{Z}^{L \times M \times N} \ : \ \sum_i x_{i,j,k} = \sum_j x_{i,j,k} = \sum_k x_{i,j,k} = U \ , \ 0 \leq x_{i,j,k} \leq u_{i,j,k} \} \ . \qquad (4)$$

Choosing L and M to be relatively small and fixed, $\mathcal{F}$ is again the set of integer points in an N-fold system. Corollary 16 part 2 now enables the efficient solution of the *error-correcting decoding* problem

$$\min\{\|\hat{x}-\bar{x}\|_p \ : \ \hat{x}\in\mathcal{F}\} \ .$$

Corollary 19 *Fix L, M. Then 3-way l_p error-correcting decoding can be done in polynomial time.*

4.3 Hierarchically-constrained multi-way arrays

The transportation and routing problem, as well as the error-correction problem, have very broad and useful generalizations, to arrays of any dimension and to any *hierarchical* sum constraints. We proceed to define such systems of arrays.

Consider d-way arrays $x=(x_{i_1,\dots,i_d})$ of size $M_1\times\cdots\times M_d$. For any d-tuple $(i_1,\dots,i_d)$ with $i_j\in\{1,\dots,M_j\}\cup\{+\}$, the corresponding *margin* $x_{i_1,\dots,i_d}$ is the sum of entries of x over all coordinates j with $i_j=+$. The *support* of $(i_1,\dots,i_d)$ and of $x_{i_1,\dots,i_d}$ is the set $\mathrm{supp}(i_1,\dots,i_d):=\{j : i_j\neq +\}$ of non-summed coordinates. For instance, if x is a $4\times 5\times 3\times 2$ array then it has 12 margins with support $H=\{1,3\}$ such as $x_{3,+,2,+}=\sum_{i_2=1}^{5}\sum_{i_4=1}^{2}x_{3,i_2,2,i_4}$. Given a family $\mathcal{H}$ of subsets of $\{1,\dots,d\}$ and margin values $v_{i_1,\dots,i_d}$ for all tuples with support in $\mathcal{H}$, consider the set of integer nonnegative and suitably upper-bounded arrays with these margins,

$$\mathcal{F}_{\mathcal{H}} \ := \ \left\{ x\in\mathbb{Z}^{M_1\times\cdots\times M_d} \ : \ x_{i_1,\dots,i_d} \ = \ v_{i_1,\dots,i_d}\,, \ \ \mathrm{supp}(i_1,\dots,i_d)\in\mathcal{H}\,, \ \ 0\le x_{i_1,\dots,i_d}\le u_{i_1,\dots,i_d}\right\} \ .$$

The congestion-avoiding transportation problem over $\mathcal{F}_{\mathcal{H}}$ is to find $x\in\mathcal{F}_{\mathcal{H}}$ minimizing a given separable convex cost $\sum_{i_1,\dots,i_d}f_{i_1,\dots,i_d}(x_{i_1,\dots,i_d})$. The error-correcting decoding problem over $\mathcal{F}_{\mathcal{H}}$ is to estimate an original message as $\hat{x}\in\mathcal{F}_{\mathcal{H}}$ minimizing a suitable l_p-distance $\|\hat{x}-\bar{x}\|_p$ to a received message $\bar{x}$.

Again, for long arrays, that is, of format $M_1\times\cdots\times M_{d-1}\times N$ with d and $M_1,\dots,M_{d-1}$ fixed and only the length (number of layers) N variable, the set $\mathcal{F}_{\mathcal{H}}$ is the set of feasible points in an N-fold systems and, as a consequence of Corollary 16, we can solve both problems in polynomial time.

Corollary 20 *Fix any $d, M_1,\dots,M_{d-1}$ and family $\mathcal{H}$ of subsets of $\{1,\dots,d\}$. Then congestion-avoiding transportation and error-correcting decoding over $\mathcal{F}_{\mathcal{H}}$ can be solved in polynomial time for any array length $M_d:=N$ and any margin values $v_{i_1,\dots,i_d}$ for all tuples $(i_1,\dots,i_d)$ with support in $\mathcal{H}$.*

5 Proofs of Theorems 5 and 6

In this section we finally prove Theorems 5 and 6. For this, we employ the following fact.

Lemma 21 (Theorem 3.1 in Ahuja et al. [1]) *Let H be the difference between maximum and minimum objective function values of an (integer valued) optimization problem.*

Suppose that f^k is the objective function value of some solution of a minimization problem at the k-th interation of an algorithm and f^* is the minimum objective function value. Furthermore, suppose that the algorithm guarantees that for every iteration k,

$$(f^k - f^{k+1}) \geq \beta(f^k - f^*)$$

(i.e., the improvement at iteration $k+1$ is at least β times the total possible improvement) for some constant $0 < \beta < 1$ (which is independent of the problem data). Then the algorithm terminates in $O((\log H)/\beta)$.

5.1 Proof of Theorem 5

Let Δ denote the least common multiple of all non-vanishing maximal subdeterminants of A. Note that the encoding length $\log \Delta$ is polynomially bounded in the encoding lengths of the input data A, u, b and c. Hence, the objective function values of two vertices are either the same or differ by at least $1/\Delta$.

Let $f^0 = \Delta \cdot c^\intercal z^0$ denote the normalized objective value of the initially given feasible solution and by $f^1, f^2, \ldots$ denote the normalized objective values of the vertices $z^1, z^2, \ldots$ that we reach at the end of the second steps of the augmentation procedure. Note that the difference H between maximum and minimum normalized objective function values of $(\mathrm{LP})_{A,u,b,f}$ has an encoding length $\log H$ that is polynomially bounded in the encoding lengths of the input data A, u, b and c. We now show that

$$(f^k - f^{k+1}) \geq \beta(f^k - f^*)$$

holds for $0 < \beta = 1/n < 1$ and conclude by Lemma 21, that we only have to enumerate $O((\log H)n)$, that is polynomially many, vertices.

Cosider the vector $z^* - z^k \in \ker(A)$. There is some orthant $\mathbb{O}_j$ such that $z^* - z^k \in \ker(A) \cap \mathbb{O}_j$. Hence, we can write

$$z^* - z^k = \sum_{i=1}^{n} \alpha_i g_i$$

for some $\alpha_i \in \mathbb{R}_+$ and $g_i \in \mathcal{C}(A) \cap \mathbb{O}_j$, $i = 1, \ldots, n$. As $\alpha_i g_i$ has the same sign pattern as $z^k - z^*$, one can easily check that the components of $z^k + \alpha_i g_i$ lie between the components of z^k and of z^*. Hence they are nonnegative. As $g_i \in \ker(A)$, we have $Ag_i = 0$ and thus $A(z^k + \alpha_i g_i) = Az^k = b$ for any choice of $i = 1, \ldots, n$. Consequently, $z^k + \alpha_i g_i$ is a feasible solution for any choice of $i = 1, \ldots, n$. Finally, we have

$$\Delta \cdot c^\intercal (z^k - z^*) = \sum_{i=1}^{n} \Delta \cdot c^\intercal(-\alpha_i g_i) = \sum_{i=1}^{n} \Delta \cdot c^\intercal(z^k - (z^k + \alpha_i g_i))$$

from which we conclude that there is some index i_0 such that

$$\Delta \cdot c^\intercal(z^k - (z^k + \alpha_{i_0} g_{i_0})) = \Delta \cdot c^\intercal(-\alpha_{i_0} g_{i_0}) \geq \frac{1}{n} \sum_{i=1}^{n} \Delta \cdot c^\intercal(-\alpha_i g_i) = \frac{1}{n} \Delta \cdot c^\intercal(z^k - z^*) = \frac{1}{n}(f^k - f^*).$$

Note that a greedy choice for an augmentating vector cannot make a smaller augmentation step than the vector $\alpha_{i_0} g_{i_0}$. Thus,

$$f^k - f^{k+1} \geq \Delta \cdot c^\intercal (z^k - (z^k + \alpha_{i_0} g_{i_0})) \geq \frac{1}{n}(f^k - f^*).$$

This proves Part (a).

The proof of Part (b) is is nearly literally the same. Clearly, in the integer situation, we may choose $\Delta = 1$. If $z^1, z^2, \ldots$ denote the vectors that we reach from our initial feasible solution z^0 via greedy augmentation steps, we only have to be careful about the choice of β. In the integer situation, we need to choose $\beta = 1/(2n-2)$, since for the integer vector $z^* - z^k \in \ker(A) \cap \mathbb{O}_j$ at most $2n-2$ vectors from the Hilbert basis of $C_j = \ker(A) \cap \mathbb{O}_j$ are needed to represent each lattice point in $C_j \cap \mathbb{Z}^n$ as a nonnegative *integer* linear combination of elements in $\mathcal{G}(A) \cap \mathbb{O}_j$ [16]. Thus, we need to apply $O((\log H)(2n-2)) = O((\log H)n)$ augmentation steps, a number being polynomial in the encoding length. □

5.2 Proof of Theorem 6

In [9, 13], it was shown that $\mathcal{G}(A, C)$ allows a representation

$$(z^* - z^k, -C(z^* - z^k)) = \sum_{i=1}^{2(n+s)-2} \alpha_i (g_i, -Cg_i),$$

where each $\alpha_i \in \mathbb{Z}_+$ and where each $(g_i, -Cg_i)$ lies in the same orthant as $(z^* - z^k, -C(z^* - z^k))$. It follows again from the results in [16] that at most $2(n+s)-2$ summands are needed. Similarly to the proof of Theorem 5, we can already conclude from this representation that $z^k + \alpha_i g_i$ is feasible for all $i = 1, \ldots, 2(n+s)-2$.

Moreover, in [13] it was shown that for such a representation superadditivity holds, that is,

$$\bar{f}(z^*) - \bar{f}(z^k) \geq \sum_{i=1}^{2(n+s)-2} [\bar{f}(z^k + \alpha_i g_i) - \bar{f}(z^k)]$$

and thus, rewritten,

$$\bar{f}^k - \bar{f}^* = \bar{f}(z^k) - f(z^*) \leq \sum_{i=1}^{2(n+s)-2} [\bar{f}(z^k) - \bar{f}(z^k + \alpha_i g_i)].$$

Therefore, there is some index i_0 such that

$$\bar{f}^k - \bar{f}^{k+1} = \bar{f}(z^k) - \bar{f}(z^k + \alpha_{i_0} g_{i_0}) \geq \frac{1}{2(n+s)-2}[\bar{f}(z^k) - \bar{f}(z^*)] = \frac{1}{2(n+s)-2}(\bar{f}^k - \bar{f}^*),$$

and the result follows from Lemma 21. □

References

[1] R. K. Ahuja, T. L. Magnanti, and J. B. Orlin. *Network flows: theory, algorithms, and applications*. Prentice-Hall, Inc., New Jersey, 1993.

[2] M. Aschenbrenner and R. Hemmecke. Finiteness theorems in stochastic integer programming. *Foundations of Computational Mathematics* **7** (2007), 183–227.

[3] J. R. Birge and F. V. Louveaux. Intorduction to Stochastic Programming. Springer, New York, 1997.

[4] J. A. De Loera, R. Hemmecke, S. Onn, and R. Weismantel. N-fold integer programming. *Discrete Optimization*, to appear.

[5] J. A. De Loera, and S. Onn: The complexity of three-way statistical tables. *SIAM Journal on Computing* **33** (2004), 819–836.

[6] J. A. De Loera, and S. Onn: All linear and integer programs are slim 3-way transportation programs. *SIAM Journal on Optimization* **17** (2006), 806–821.

[7] J. E. Graver. On the foundation of linear and integer programming I. *Mathematical Programming* **9** (1975), 207–226.

[8] R. Hemmecke. On the positive sum property and the computation of Graver test sets. *Mathematical Programming* **96** (2003), 247–269.

[9] R. Hemmecke. Test Sets for Integer Programs with Z-convex Objective Function. e-print available from `http://front.math.ucdavis.edu/math.CO/0309154`, 2003.

[10] R. Hemmecke and R. Schultz. Decomposition of test sets in stochastic integer programming. *Mathematical Programming*, **94** (2003), 323–341.

[11] S. Hoşten and S. Sullivant. Finiteness theorems for Markov bases of hierarchical models. *Journal of Combinatorial Theory: Series A* **114** (2007), 311–321.

[12] F. V. Louveaux and R. Schultz. Stochastic Integer Programming. In: *Handbooks in Operations Research and Management Science, 10: Stochastic Programming*, A. Ruszczynski, A. Shapiro (eds.), Elsevier Science, Amsterdam, 2003, 213–266.

[13] K. Murota, H. Saito, and R. Weismantel. Optimality criterion for a class of nonlinear integer programs. *Operations Research Letters* **32** (2004), 468–472.

[14] W. Römisch, R. Schultz. Multistage stochastic integer programming: an introduction. In: *Online Optimization of Large Scale Systems*, M. Grötschel, S. O. Krumke, J. Rambau (eds.), Springer-Verlag Berlin, 2001, 581–600.

[15] A. Schulz and R. Weismantel. An oracle-polynomial time augmentation algorithm for integer programming. In: *Proc. of the 10th ACM-SIAM Symposium on Discrete Algorithms*, Baltimore, 1999, 967–968.

[16] A. Sebö. Hilbert bases, Caratheodory's Theorem and combinatorial optimization. In: *Proc. of the IPCO conference*, Waterloo, Canada, 1990, 431–455.