New Bounds on Augmenting Steps of Block-Structured Integer Programs

Lin Chen
Department of Computer Science, Texas Tech University, Lubbock, TX, US
chenlin198662@gmail.com

Martin Koutecký
Computer Science Institute, Charles University, Prague, Czech Republic
koutecky@iuuk.mff.cuni.cz

Lei Xu
Department of Computer Science, University of Texas Rio Grande Valley, TX, US
xuleimath@gmail.com

Weidong Shi
Department of Computer Science, University of Houston, TX, US
larryshi@ymail.com

Abstract

Iterative augmentation has recently emerged as an overarching method for solving Integer Programs (IP) in variable dimension, in stark contrast with the volume and flatness techniques of IP in fixed dimension. Here we consider *4-block n-fold integer programs*, which are the most general class considered so far. A 4-block n-fold IP has a constraint matrix which consists of n copies of small matrices A, B, and D, and one copy of C, in a specific block structure. Iterative augmentation methods rely on the so-called *Graver basis* of the constraint matrix, which constitutes a set of fundamental augmenting steps. All existing algorithms rely on bounding the ℓ_1- or ℓ_∞-norm of elements of the Graver basis. Hemmecke et al. [Math. Prog. 2014] showed that 4-block n-fold IP has Graver elements of ℓ_∞-norm at most $\mathcal{O}_{FPT}(n^{2^{s_D}})$, leading to an algorithm with a similar runtime; here, s_D is the number of rows of matrix D and $\mathcal{O}_{FPT}$ hides a multiplicative factor that is only dependent on the small matrices A, B, C, D, However, it remained open whether their bounds are tight, in particular, whether they could be improved to $\mathcal{O}_{FPT}(1)$, perhaps at least in some restricted cases.

We prove that the ℓ_∞-norm of the Graver elements of 4-block n-fold IP is upper bounded by $\mathcal{O}_{FPT}(n^{s_D})$, improving significantly over the previous bound $\mathcal{O}_{FPT}(n^{2^{s_D}})$. We also provide a matching lower bound of $\Omega(n^{s_D})$ which even holds for arbitrary non-zero lattice elements, ruling out augmenting algorithm relying on even more restricted notions of augmentation than the Graver basis. We then consider a special case of 4-block n-fold in which C is a zero matrix, called 3-block n-fold IP. We show that while the ℓ_∞-norm of its Graver elements is $\Omega(n^{s_D})$, there exists a different decomposition into lattice elements whose ℓ_∞-norm is bounded by $\mathcal{O}_{FPT}(1)$, which allows us to provide improved upper bounds on the ℓ_∞-norm of Graver elements for 3-block n-fold IP. The key difference between the respective decompositions is that a Graver basis guarantees a *sign-compatible* decomposition; this property is critical in applications because it guarantees each step of the decomposition to be feasible. Consequently, our improved upper bounds let us establish faster algorithms for 3-block n-fold IP and 4-block IP, and our lower bounds strongly hint at parameterized hardness of 4-block and even 3-block n-fold IP. Furthermore, we show that 3-block n-fold IP is without loss of generality in the sense that 4-block n-fold IP can be solved in FPT oracle time by taking an algorithm for 3-block n-fold IP as an oracle.

2012 ACM Subject Classification Theory of computation → Parameterized complexity and exact algorithms; Mathematics of computing → Combinatorial algorithms

Keywords and phrases Integer Programming, Graver basis, Fixed parameter tractable

Digital Object Identifier 10.4230/LIPIcs.ESA.2020.33

28th Annual European Symposium on Algorithms (ESA 2020).
Editors: Fabrizio Grandoni, Grzegorz Herman, and Peter Sanders; Article No. 33; pp. 33:1–33:19
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Related Version A full version of this paper is available at `https://arxiv.org/abs/1805.03741`.

Funding *Lin Chen*: Research was supported in part by NSF 1756014.
Martin Koutecký: Partially supported by Israel Science Foundation grant 308/18, Charles University project UNCE/SCI/004 and by the project 19-27871X of GA ČR.

1 Introduction

A powerful mathematical tool for modeling of various optimization problems is Integer Programming:

$$\min\{\mathbf{w}\cdot\mathbf{x} \,:\, \mathcal{A}\mathbf{x}=\mathbf{b},\, \mathbf{l}\le\mathbf{x}\le\mathbf{u},\, \mathbf{x}\in\mathbb{Z}^N\}, \tag{IP}$$

where $\mathbf{w}, \mathbf{b}, \mathbf{l}, \mathbf{u}$ are integer vectors of the *objective function*, *right hand side*, and *lower and upper bounds*, respectively, $\mathcal{A}$ is an integer *constraint matrix*, and $\mathbf{x}$ is a vector of *variables.* It plays a key role in theory as a component in the design of approximation and parameterized algorithms, as well as in practice, with current solvers being routinely utilized in industry and capable of handling models with thousands of variables.

In general, Integer Programming is NP-hard, as was shown already by Karp [23], which motivates the search for tractable special cases. Famous polynomially solvable cases are IPs with few rows and small coefficients as shown by Papadimitriou in 1981 [30], and IPs with few variables as shown by Lenstra in 1983. Arguably the most significant development in the last 20 years has been the introduction of *iterative augmentation* methods which led to the development of fast algorithms for wide classes of IPs whose constraint matrix has a special block structure, and to subsequent breakthrough applications in parameterized and approximation algorithms [5, 21, 27]. In fact, essentially *all* known tractable classes of IP in variable dimension are of this kind, except total unimodular IPs from the '60s.

An iterative augmentation algorithm starts with an initial feasible solution $\mathbf{x}$ and iteratively finds *augmenting steps* $\mathbf{g}\in\mathbb{Z}^N$, i.e., $\mathbf{x}+\mathbf{g}$ is feasible and $\mathbf{w}(\mathbf{x}+\mathbf{g}) < \mathbf{w}\mathbf{x}$. A major question is *where* to obtain "good" augmenting steps. The *Graver basis of* $\mathcal{A}$, $\mathcal{G}(\mathcal{A})$, has emerged as an excellent choice, with good guarantees on convergence to optimal solutions while still being algorithmically "tame". Specifically, at the heart of iterative augmentation techniques are bounds on the ℓ_1- and ℓ_∞-norm of elements of the Graver basis, which enable dynamic programming to be used to find Graver elements.

We stress the role of bounds on the elements of $\mathcal{G}(\mathcal{A})$. Historically, all tractable classes of IP were discovered by proving new norm bounds and subsequently designing a dynamic program around them, with the former typically being much harder than the latter. Moreover, recent runtime improvements have followed from improving existing bounds [8, 28], and the most challenging questions in the field are tightly connected to norm bounds. Our focus here is the currently least understood class of IPs, *4-block n-fold IP*:

$$(\mathrm{IP})_{n,\mathbf{b},\mathbf{l},\mathbf{u},\mathbf{w}}: \quad \min\{\mathbf{w}\cdot\mathbf{x} : H\mathbf{x}=\mathbf{b},\, \mathbf{l}\le\mathbf{x}\le\mathbf{u},\, \mathbf{x}\in\mathbb{Z}^{t_B+nt_A}\}, \tag{1}$$

where H (called a 4-*block* n-*fold matrix*) is build from smaller blocks A, B, C and D:

$$H=\begin{pmatrix} C & D \\ B & A\end{pmatrix}^{(n)} := \begin{pmatrix} C & D & D & \cdots & D \\ B & A & 0 & & 0 \\ B & 0 & A & & 0 \\ \vdots & & & \ddots & \\ B & 0 & 0 & & A \end{pmatrix} .$$

Here, A, B, C, D are $s_i \times t_i$ matrices, $i = A, B, C, D$, respectively, and H consists of n copies of A, B, D and one copy of C. Notice that by plugging A, B, C, D into the above block structure we require that $s_C = s_D$, $s_A = s_B$, $t_B = t_C$ and $t_A = t_D$. Let Δ be the largest absolute value among all the entries of A, B, C, D. Let H_0 be a matrix obtained from H by setting $C = \mathbf{0}$. We also study *3-block n-fold IP*, obtained by replacing H with H_0.

For ease of presentation, we introduce the submatrices E and F such that

$$E := \begin{pmatrix} D & D & \cdots & D \\ A & 0 & & 0 \\ 0 & A & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & A \end{pmatrix} \qquad F := \begin{pmatrix} B & A & 0 & & 0 \\ B & 0 & A & & 0 \\ \vdots & & & \ddots & \\ B & 0 & 0 & & A \end{pmatrix}, \tag{2}$$

4-block n-fold IP remains the simplest case of block-structured IPs for which an algorithm of runtime $f(s_A, t_A, \ldots, s_D, t_D, \Delta)n^{O(1)}$ (i.e., an FPT algorithm; see below) remains unknown. From another perspective, Koutecký et al. [28] has recently resolved the complexity of IP with respect to the structural parameters *primal* and *dual treedepth* td_P and td_D, respectively, by showing that IPs with small td_P and td_D are efficiently solvable. IPs with small *incidence treedepth* td_I subsume both of the aforementioned classes as well as 4-block n-fold IP, and 4-block n-fold IP remains the simplest open case with respect to td_I.

Our Contribution. Because we are interested in efficient algorithms, we wish to confine the exponential dependence on the input into the small numbers s_i, t_i, $i = A, B, C, D$, and Δ. Thus we take the perspective of *parameterized complexity*: for a problem instance I with a *parameter* k, we call an algorithm with runtime $f(k)|I|^{O(1)}$ a *fixed-parameter tractable (FPT)* algorithm, and an algorithm with runtime $|I|^{f(k)}$ an *XP algorithm* (for *slice-wise polynomial*). If such algorithms exist, we say that the problem is *FPT* or *XP parameterized by* k, respectively.

In this paper, we provide new and improved upper bounds and resulting algorithms for 4-block and 3-block n-fold IP, as well as the very first lower bounds for these classes which we believe to hint at the parameterized hardness of these problems. We denote by $\ker_{\mathbb{Z}}(H) = \{\mathbf{x} \in \mathbb{Z}^{t_B + nt_A} \mid H\mathbf{x} = \mathbf{0}\}$ the *integer kernel of* H, also called the *lattice of* H, and by $g_\infty(H) = \max_{\mathbf{g} \in \mathcal{G}(H)} \|\mathbf{g}\|_\infty$ the largest ℓ_∞-norm of an element of the Graver basis $\mathcal{G}(H)$ (a precise definition is given below in Section 2); analogously for H_0. First, we show an upper bound on $g_\infty(H)$.

▶ **Theorem 1.** *For any 4-block n-fold matrix H, $g_\infty(H) \le \mathcal{O}_{FPT}(n^{s_D})$.*

This improves on the previous bound of $\mathcal{O}_{FPT}(n^{2^{s_D}})$ [16]. We also establish the first explicit lower bound matching our upper bound, making it tight up to an FPT factor. Importantly, our lower bound even applies to the first t_B coordinates (denoted $\mathbf{x}^0$ for a vector $\mathbf{x} \in \mathbb{Z}^{t_B + nt_A}$) which play a special role in algorithms for 4-block n-fold IP. What is more, our lower bound even applies to *any* non-zero element of $\ker_{\mathbb{Z}}(H)$:

▶ **Theorem 2.** *For arbitrary integer $t \in \mathbb{N}$, there exists a 4-block n-fold matrix H such that $s_i, t_i \in O(t)$ for $i = A, B, C, D$, and for* any *$\mathbf{g} \in \ker_{\mathbb{Z}}(H)$ we have $\|\mathbf{g}^0\|_\infty = \Omega(n^t)$.*

Therefore, even augmenting via a different set of steps may have to deal with steps that are unbounded by $\mathcal{O}_{FPT}(1)$. Combining Theorem 1 with the original idea of Hemmecke et al. [16] and a strongly polynomial framework of Koutecký et al. [28], we obtain the currently fastest algorithm for 4-block n-fold IP:

▶ **Theorem 3.** *4-block n-fold IP can be solved in time* $\mathcal{O}_{FPT}(n^{O(s_D t_B)})$.

Second, we restrict our attention to 3-block n-fold IP. The motivation is that 3-block n-fold IP is essentially no less general than 4-block n-fold IP. Indeed, for any 4-block n-fold IP, there exists an equivalent 3-block n-fold IP where the largest coefficient, number of rows and columns of the submatrices only increase by $\mathcal{O}(1)$ times (see Theorem 19 and Definition 17 in Appendix 5 for a formal statement).

Interestingly, the lattice elements (i.e., augmenting step candidates) of 3-block n-fold IP admit a decomposition with ℓ_∞-norm bounded by $\mathcal{O}_{FPT}(1)$:

▶ **Theorem 4.** *Any* $\mathbf{g} \in \ker_{\mathbb{Z}}(H_0)$ *decomposes to* $\sum_{i=1}^N \mathbf{e}_i$ *for some* $N \in \mathbb{Z}_{\geq 0}$ *with* $\mathbf{e}_i \in \ker_{\mathbb{Z}}(H_0)$ *and* $\|\mathbf{e}_i\|_\infty \leq \mathcal{O}_{FPT}(1)$ *for each* i.

However, this decomposition is not "sign-compatible", meaning possibly none of its elements is a feasible step on its own, which makes its immediate algorithmic use complicated. Nevertheless, we are able to use it to establish an upper bound of $\min\{\mathcal{O}_{FPT}(n^{s_D}), \mathcal{O}_{FPT}(n^{t_A^2+1})\}$ (below, and Theorem 1):

▶ **Theorem 5.** *For any 3-block n-fold matrix* H_0, $g_\infty(H_0) \leq \mathcal{O}_{FPT}(n^{t_A^2+1})$.

This upper bound of $\mathcal{O}_{FPT}(n^{t_A^2+1})$, which is singly exponential in t_A, is much more involved compared with the upper bound of Theorem 1. This coincides with the existing results for 4-block n-fold IP [16], where an upper bound depending on A, B (instead of C, D) is much more complicated. Our proof relies on a completely new approach, which first establishes the decomposition of Theorem 4 and then modifies it into a sign-compatible decomposition through merging summands. This may be of separate interest for deriving upper bounds on $g_\infty(\mathcal{A})$ for other classes of matrices $\mathcal{A}$, particularly for deriving an upper bound on $g_\infty(H)$ which has an explicit dependency on s_A, s_B, t_A, t_B in the exponent of n. Moreover, we show that any 4-block n-fold IP can be embedded in a 3-block n-fold IP (Theorem 19) in a particular way, which allows us to transfer the 4-block n-fold lower bound (now restricted to *feasible* lattice elements):

▶ **Theorem 6.** *For arbitrary integer* $t \in \mathbb{N}$*, there exists a 3-block n-fold IP with a matrix H such that* $s_i, t_i \in O(t)$ *for* $i = A, B, C, D$*, and for any* feasible *nonzero* $\mathbf{g} \in \ker_{\mathbb{Z}}(H_0)$ *we have* $\|\mathbf{g}^0\|_\infty = \Omega(n^t)$.

Finally, using our new upper bound of Theorem 5, we get that:

▶ **Theorem 7.** *3-block n-fold IP can be solved in time* $\min\{\mathcal{O}_{FPT}(n^{O(s_D t_B)}, \mathcal{O}_{FPT}(n^{O(t_A^2 t_B)})\}$.

Related Work

4-block n-fold IP originated as a generalization of two previously studied classes of IP, the *n-fold* and *2-stage stochastic* IP, which are obtained by substituting the constraint matrix H with E and F we defined before. We also call E the n-fold matrix and F the 2-stage stochastic matrix, respectively. The origins of iterative augmentation methods for 2-stage stochastic IP reach the work of Hemmecke and Schultz in 2001 [19]. De Loera et al. [7] first studied n-fold IP in 2008. Later, Hemmecke et al. [17] showed an FPT algorithm for n-fold IP based on dynamic programming, which led to a breakthrough in computational social choice [26] and was also applied in the context of scheduling by Knop and Koutecký [25]. Later, this FPT algorithm inspired a better algorithm for a special case of *combinatorial*

n-fold IP developed by Knop et al. [27], who also apply it to problems in stringology and graph algorithms. Finally, this algorithm was lifted to the general n-fold IP by Koutecký et al. [28] and Eisenbrand et al. [8].

An extension of n-fold IP to tree-structured matrices called *tree-fold IP* was developed by Chen and Marx [5] and applied to scheduling problems. Jansen et al. [21] have used n-fold IP to obtain efficient PTASes for scheduling problems. An extension of 2-stage stochastic IP analogous to tree-folds is called multi-stage stochastic and was studied by Aschenbrenner and Hemmecke [4]. Ganian and Ordyniak [12] studied the structural parameters primal treedepth and treewidth, and later Ganian et al. [13] studied dual and incidence treedepth and treewidth. Koutecký et al. [28] discovered that tree-fold and multi-stage stochastic IPs are essentially equivalent to IPs with small dual and primal treedepth, settling the parameterized complexity with respect to these parameters. The work of Koutecký et al. [28] subsumes essentially all current knowledge about the solvability of IP in variable dimension with the exception of totally unimodular constraint matrices and two related classes [2, 3], with the main remaining open problem being the complexity of 4-block n-fold IP and, more generally, IP with respect to incidence treedepth.

Bounds on $g_\infty(\mathcal{A})$ and $g_1(\mathcal{A}) = \max_{\mathbf{g}\in\mathcal{G}(\mathcal{A})} \|\mathbf{g}\|_1$ play a central role in the recent developments. For example, Chen and Marx [5] showed that tree-fold IP is FPT, but a naïve analysis yields a tower-of-exponentials dependence on the parameters. Eisenbrand et al. [8] lower this to double-exponential by improving the bounds on $g_1(\mathcal{A})$, and, at least with the current approach, the only way to obtain a single-exponential algorithm is by obtaining single-exponential bounds on $g_1(\mathcal{A})$. It has been known for a long time that 2-stage stochastic IP is FPT [19], however, there are no known bounds at all for this algorithm except for the computability of the parameter dependence f due to no bounds being available for $g_\infty(F)$. Very recently, Klein [24] is able to obtain such a bound for $g_\infty(F)$, which yields an FPT algorithm with a concrete running time. Lower bounds on $g_\infty(\mathcal{A})$ have been rare so far. Finhold and Hemmecke [11] study them in the context of n-fold IP. Koutecký et al. [28] show lower bounds (only using elementary techniques) for IPs in terms of their primal and dual treewidth.

We use the Steinitz Lemma, which has recently gained renewed attention [10, 8, 22].

2 Preliminaries

Notations

We write vectors in boldface, e.g. $\mathbf{x}, \mathbf{y}$, and their entries in normal font, e.g. x_i, y_i. Any (t_B+nt_A)-dimensional vector $\mathbf{x}$ can be divided into $n+1$ *bricks*, such that $\mathbf{x} = (\mathbf{x}^0, \mathbf{x}^1, \cdots, \mathbf{x}^n)$ where $\mathbf{x}^0 \in \mathbb{Z}^{t_B}$ and each $\mathbf{x}^i \in \mathbb{Z}^{t_A}$, $1 \le i \le n$. We call $\mathbf{x}^i$ the *i-th brick* for $0 \le i \le n$. We write $0_{s\times t}$ for an $s \times t$ matrix consisting of 0, and I_t for an $t \times t$ identity matrix. For a vector or a matrix, we write $\|\cdot\|_\infty$ to denote the maximal absolute value of its elements. For two vectors $\mathbf{x}, \mathbf{y}$ of the same dimension, $\mathbf{x} \cdot \mathbf{y}$ denotes their inner product.

Throughout this paper, we write $\mathcal{O}_{FPT}(1)$ to represent a parameter that is only dependent on $\Delta, s_A, s_B, s_C, s_D, t_A, t_B, t_C, t_D$ where Δ is the maximal absolute value among all the entries of A, B, C, D, that is, $\mathcal{O}_{FPT}(1)$ is only dependent on the small matrices A, B, C, D and is independent of n. For any computable function $f(x)$, we write $\mathcal{O}_{FPT}(f)$ to represent a computable function $f'(x)$ such that $|f'(x)| \le \mathcal{O}_{FPT}(1) \cdot |f(x)|$, and $\Omega_{FPT}(f)$ to represent a function f'' such that $|f''(x)| \ge \Omega(1) \cdot |f(x)|$. If no FPT-term is hidden, we will use $\mathcal{O}$ in its standard meaning (e.g., in Theorem 6).

Two vectors $\mathbf{x}$ and $\mathbf{y}$ are called *sign-compatible* if $x_i \cdot y_i \geq 0$ holds for every pair of coordinates (x_i, y_i). Furthermore, we call a summation $\sum_i \mathbf{x}_i$ *sign-compatible* if the summands are pair-wise sign-compatible.

Graver basis

Consider the general integer linear programming in the standard form (IP). Let $\sqsubseteq$ be the *conformal order* in $\mathbb{R}^m$ defined such that $\mathbf{x} \sqsubseteq \mathbf{y}$ if $\mathbf{x}$ and $\mathbf{y}$ lie in the same orthant, i.e., $x_i \cdot y_i \geq 0$ for each $i = 1, \ldots, m$, and $|x_i| \leq |y_i|$ for each $i = 1, \ldots, m$. Given any subset $X \subseteq \mathbb{R}^n$, we say $\mathbf{x}$ is an $\sqsubseteq$-*minimal* element of X if $\mathbf{x} \in X$ and there does not exist $\mathbf{y} \in X$, $\mathbf{y} \neq \mathbf{x}$ such that $\mathbf{y} \sqsubseteq \mathbf{x}$. It is known that every subset of $\mathbb{Z}^m$ has finitely many $\sqsubseteq$-minimal elements. We study the Graver basis:

▶ **Definition 8** (Graver basis [14]). *The* Graver basis *of an integer matrix E is the finite set $\mathcal{G}(E) \subseteq \ker_{\mathbb{Z}}(E)$ of all $\sqsubseteq$-minimal elements of $\ker_{\mathbb{Z}}(E) \setminus \{\mathbf{0}\}$.*

For clarity, we sometimes emphasize that $\mathbf{g}$ comes from $\mathcal{G}(H)$ by writing it as $\mathbf{g}(H)$, and similarly for other vectors. We use the fact that any $\mathbf{x} \in \ker_{\mathbb{Z}}(H)$, $\mathbf{x} \neq 0$ can be written as $\mathbf{x} = \sum_i \alpha_i \mathbf{g}_i(H)$, where $\alpha_i \in \mathbb{Z}_+$, $\mathbf{g}_i(H) \in \mathcal{G}(H)$ and $\mathbf{g}_i(H) \sqsubseteq \mathbf{x}$ [29, Lemma 3.4].

The Graver basis $\mathcal{G}(H)$ is only dependent on H. Let $\|B\|_\infty$ be the largest absolute value over all entries. For any $\mathbf{g} \in \mathcal{G}(A)$, we have the following rough estimation for some constant c_1, c_2 [29]:

$$|\mathcal{G}(H)| \leq (c_1 \|H\|_\infty)^{mn} \quad \text{and } \|g\|_\infty \leq (c_2 \|A\|_\infty)^{mn}.$$

Augmentation algorithms for IP and Graver-best oracle

There is a general framework for solving (IP) by utilizing $\mathcal{G}(\mathcal{A})$, which was developed in a series of papers [5, 17, 21, 27]. A recent paper by Koutecký et al. [28] formalizes this framework and extends it to also obtaining *strongly polynomial* algorithms (algorithms whose number of arithmetic operations does not depend on the length of the numbers on input).

We say that $\mathbf{x}$ is *feasible* for (IP) if $\mathcal{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{l} \leq \mathbf{x} \leq \mathbf{u}$. Let $\mathbf{x}$ be a feasible solution for (IP). We call $\mathbf{g}$ a *feasible step* if $\mathbf{x} + \mathbf{g}$ is feasible for (IP). Further, call a feasible step $\mathbf{g}$ *augmenting* if $\mathbf{w}(\mathbf{x} + \mathbf{g}) < \mathbf{w}(\mathbf{x})$. An augmenting step $\mathbf{g}$ and a *step length* $\rho \in \mathbb{Z}$ form an $\mathbf{x}$-*feasible step pair* with respect to a feasible solution $\mathbf{x}$ if $\mathbf{l} \leq \mathbf{x} + \rho\mathbf{g} \leq \mathbf{u}$. An augmenting step $\mathbf{h}$ is a *Graver-best step* for $\mathbf{x}$ if $\mathbf{w}(\mathbf{x} + \mathbf{h}) \leq \mathbf{w}(\mathbf{x} + \rho\mathbf{g})$ for all $\mathbf{x}$-feasible step pairs $(\mathbf{g}, \rho) \in \mathcal{G}(\mathcal{A}) \times \mathbb{Z}$. The next definition and theorem show that it is sufficient to focus all our attention on finding Graver-best steps. This takes care of matters such as finding an initial feasible solution, using a proximity theorem to shrink $\mathbf{w}, \mathbf{b}, \mathbf{l}, \mathbf{u}$ and so on.

▶ **Definition 9** (Graver-best oracle). *A* Graver-best oracle *for an integer matrix $\mathcal{A}$ is one that, queried on $\mathbf{w}, \mathbf{b}, \mathbf{l}, \mathbf{u}$ and $\mathbf{x}$ feasible to* (IP)*, returns a Graver-best step $\mathbf{h}$ for $\mathbf{x}$.*

▶ **Theorem 10** ([28]). *Given a Graver-best oracle for E, (IP) can be solved in strongly polynomial oracle time.*

We remark that the polynomial dependence on the dimension N and in particular the number of bricks n when it comes to 4-block n-fold IP, can be reduced using an *approximate Graver-best oracle* introduced by Altmanová et al. [1] and implicitly by Eisenbrand et al. [8].

Finiteness theorems for n-fold and 2-stage stochastic matrices

Consider an n-fold matrix E that consists of A and D (i.e., $B = C = 0$ in a 4-block n-fold matrix). It is shown that $g_\infty(E)$ is $\mathcal{O}_{FPT}(1)$. More precisely, we have the following lemma.

▶ **Lemma 11** ([9, Lemma 28]). *Let E be an n-fold matrix. Then $g_1(E) \leq (\|E\|_\infty s_D s_A)^{O(s_D s_A)}$.*

▶ **Lemma 12** ([9, Lemma 26]). *Let F be a two-stage stochastic matrix. Then $g_\infty(F) \leq f(t_B, t_A, \|A, B\|_\infty)$ for a double-exponential function f.*

Both lemmas hold for more general classes of tree-fold and multi-stage stochastic matrices.

The Steinitz lemma

The Steinitz lemma has been utilized in several recent papers [8, 10, 22] to establish better algorithms for IP. We use it as well.

▶ **Lemma 13** ([15]). *Let an arbitrary norm be given in $\mathbb{R}^\kappa$ and assume that $\|\mathbf{x}_i\| \leq \zeta$ for $1 \leq i \leq m$ and $\sum_{i=1}^m \mathbf{x}_i = \mathbf{x}$. Then there exists a permutation π such that for all positive integers $\ell \leq m$, $\|\sum_{i=1}^{\ell} \mathbf{x}_{\pi(i)} - \frac{\ell-\kappa}{m}\mathbf{x}\| \leq \kappa\zeta$.*

3 4-block n-fold IP

In this section we consider IP (1) for arbitrary H and derive matching upper and lower bounds on the ℓ_∞-norm of its Graver basis depending on the parameter $s_{\mathrm{C}} = s_{\mathrm{D}}$.

We first establish the following upper bound that improves significantly the current result.

▶ **Theorem 1.** *For any 4-block n-fold matrix H, $g_\infty(H) \leq \mathcal{O}_{FPT}(n^{s_{\mathrm{D}}})$.*

Proof. Let $\mathbf{g} \in \mathcal{G}(H)$. Recall the definition of F in Eq (2). As $F \cdot \mathbf{g} = 0$, there exist $\alpha_j \in \mathbb{Z}_+$, $\mathbf{g}_j(F) \in \mathcal{G}(F)$ and $\mathbf{g}_j(F) \sqsubseteq \mathbf{g}$ such that $\mathbf{g} = \sum_{j=1}^m \alpha_j \mathbf{g}_j(F)$. Furthermore, $\|\mathbf{g}_j(F)\|_\infty = \mathcal{O}_{FPT}(1)$ according to Lemma 12. Let $\mathbf{h}_j = C \cdot \mathbf{g}_j^0(F) + \sum_{i=1}^n D\mathbf{g}_j^i(F)$, which is an s_{D}-dimensional vector such that $\|\mathbf{h}_j\|_\infty = \mathcal{O}_{FPT}(n)$. As $H\mathbf{g} = 0$, it follows that

$$\sum_{j=1}^m \alpha_j \mathbf{h}_j = \underbrace{\mathbf{h}_1 + \mathbf{h}_1 + \cdots + \mathbf{h}_1}_{\alpha_1} + \underbrace{\mathbf{h}_2 + \mathbf{h}_2 + \cdots + \mathbf{h}_2}_{\alpha_2} + \cdots + \underbrace{\mathbf{h}_m + \mathbf{h}_m + \cdots + \mathbf{h}_m}_{\alpha_m} = 0,$$

i.e., the sequence of $\mathbf{h}_i$'s sum up to 0. According to Lemma 13, there exists a permutation of the sequence such that $\|\sum_{i=1}^{\ell} \mathbf{z}_i\|_\infty \leq s_{\mathrm{D}} \cdot \mathcal{O}_{FPT}(n) = \mathcal{O}_{FPT}(n)$ for all $\ell \leq m'$, where $m' = \sum_{i=1}^m \alpha_i$ and $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{m'}$ is a permutation of the sequence $\underbrace{\mathbf{h}_1, \mathbf{h}_1, \cdots, \mathbf{h}_1}_{\alpha_1}, \underbrace{\mathbf{h}_2, \mathbf{h}_2, \cdots, \mathbf{h}_2}_{\alpha_2}, \cdots, \underbrace{\mathbf{h}_m, \mathbf{h}_m, \cdots, \mathbf{h}_m}_{\alpha_m}$. Let $\tau = \mathcal{O}_{FPT}(n)$ be the upper bound on $\|\sum_{i=1}^{\ell} \mathbf{z}_i\|_\infty$, then we know that $\sum_{i=1}^{\ell} \mathbf{z}_i \in \{-\tau, -\tau+1, \cdots, \tau\}^{s_{\mathrm{D}}}$. Consequently, if $m' > (2\tau+1)^{s_{\mathrm{D}}} + 1$, there exists $\ell_1 < \ell_2$ such that $\sum_{i=1}^{\ell_1} \mathbf{z}_i = \sum_{i=1}^{\ell_2} \mathbf{z}_i$, i.e., $\sum_{i=1}^{\ell_2-\ell_1} \mathbf{z}_i = 0$. Recall that every $\mathbf{z}_i$ corresponds to some $\mathbf{h}_{i'}$. Suppose $\sum_{i=1}^{\ell_2-\ell_1} \mathbf{z}_i = \sum_{j=1}^m \alpha'_j \mathbf{h}_j$ for $\alpha'_j \leq \alpha_j$, then by the definition of $\mathbf{h}_j$ it follows that

$$C\left(\sum_{j=1}^m \alpha'_j \mathbf{g}_j^0(F)\right) + \sum_{i=1}^n D\left(\sum_{j=1}^m \alpha'_j \mathbf{g}_j^i(F)\right) = 0.$$

Hence, $H\sum_{j=1}^m \alpha'_j \mathbf{g}_j(F) = 0$. That is, if $m' = \sum_{j=1}^m \alpha_j > (2\tau+1)^{s_{\mathrm{D}}} + 1$, then there exists some $\mathbf{g}' = \alpha'_j \mathbf{g}_j(F)$ such that $H\mathbf{g}' = 0$, $\mathbf{g}' \sqsubset \mathbf{g}$ and $\mathbf{g}' \neq \mathbf{g}$, contradicting the fact that $\mathbf{g} \in \mathcal{G}(H)$. Thus, $\sum_{j=1}^m \alpha_j \leq (2\tau+1)^{s_{\mathrm{D}}} + 1$, implying that $\|\mathbf{g}\|_\infty = \mathcal{O}_{FPT}(n^{s_{\mathrm{D}}})$. ◀

We complement our upper bound by establishing a matching lower bound. We remark that lower bound from Theorem 2 not only holds for the ℓ_∞-norm of Graver basis elements, but even holds for any non-zero lattice element. This gives a sharp contrast to 3-block n-fold IP. As we will show later in Theorem 6 and Theorem 4, a similar lower bound also exists for the ℓ_∞-norm of Graver basis elements of $\ker_{\mathbb{Z}}(H_0)$, however, $\ker_{\mathbb{Z}}(H_0)$ does admit a decomposition into lattice elements whose ℓ_∞-norm is bounded by $\mathcal{O}_{FPT}(1)$.

▶ **Theorem 2.** *For arbitrary integer $t \in \mathbb{N}$, there exists a 4-block n-fold matrix H such that $s_i, t_i \in O(t)$ for $i = A, B, C, D$, and for* any *$\mathbf{g} \in \ker_{\mathbb{Z}}(H)$ we have $\|\mathbf{g}^0\|_\infty = \Omega(n^t)$.*

Proof. We let $A = I_{t\times t}$, $B = -I_{t\times t}$. We define $(t-1)\times t$ matrices D and C such that

$$D = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix} \qquad C = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & \cdots & -1 & 0 \end{pmatrix}$$

Consider any nonzero $\mathbf{y} \in \ker_{\mathbb{Z}}(H)$. According to $A\mathbf{y}^0 - B\mathbf{y}^i = 0$, we know that $\mathbf{y}^0 = \mathbf{y}^i$ for every $1 \le i \le n$. According to $C\mathbf{y}^0 + \sum_{i=1}^n D\mathbf{y}^i = 0$, we have $(C + nD)\mathbf{y}^0 = 0$, i.e.,

$$\begin{pmatrix} n-1 & -n & 0 & \cdots & 0 & 0 \\ 0 & n-1 & -n & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \\ 0 & 0 & 0 & \cdots & n-1 & -n \end{pmatrix} \cdot \mathbf{y} = 0$$

Let $\mathbf{y}^0 = (y_1, y_2, \cdots, y_t)$, the following is true:

$$(n-1)y_i = ny_{i+1}, \quad 1 \le i \le t-1 \tag{3}$$

It is easy to see that as long as $\mathbf{y} \neq 0$, we have $\mathbf{y}^0 \neq 0$ and consequently $y_i \neq 0$ for every $1 \le i \le t$. According to $(n-1)y_{t-1} = ny_t$, y_{t-1} is dividable by n. Let $y_{t-1} = nz_{t-1}$ for some $z_{t-1} \in \mathbb{Z}_{\neq 0}$. According to $(n-1)y_{t-2} = ny_{t-1} = n^2 z_{t-1}$, we know that y_{t-2} is dividable by n^2. Let $y_{t-2} = n^2 z_{t-2}$ and we plug it into $(n-1)y_{t-3} = ny_{t-2}$. In general, suppose we have shown that $y_{t-k} = n^k z_{t-k}$ for all $k \le k_0$. Now for $k = k_0+1$, we have $(n-1)y_{t-k_0-1} = ny_{t-k_0} = n^{k_0+1} z_{n-k_0}$, then y_{t-k_0-1} is dividable by n^{k_0+1}. Hence, we conclude that y_1 is dividable by n^{t-1}, i.e., $\|\mathbf{y}\|_\infty = \Omega(n^{t-1})$ and Theorem 2 is proved. ◀

4 3-block n-fold IP

In this section we focus on 3-block n-fold IP where $H = H_0$, i.e., $C = \mathbf{0}$. As we will show in this section, 3-block n-fold IP admits several properties that make it a particularly interesting and important special case. First, 3-block n-fold IP is without loss of generality – any 4-block n-fold IP reduces to 3-block n-fold IP with a constant increase in the parameters. Second, any element of $\ker_{\mathbb{Z}}(H_0)$ admits a decomposition into lattice elements with bounded ℓ_∞-norm, which is in certain contrast to Theorem 2. Unfortunately, a strong lower bound of $\Omega(n^t)$ for *feasible* lattice elements still exists for $s_i = t_i = \mathcal{O}(t)$. Nevertheless, we establish an alternative upper bound of $\mathcal{O}_{FPT}(n^{t_A^2+1})$ on the ℓ_∞-norm of the Graver basis elements for 3-block n-fold IP which only depends on parameters of A.

4.1 Decomposition into lattice elements with bounded ℓ_∞-norm

The goal of this subsection is to prove the following theorem.

▶ **Theorem 4.** *Any $\mathbf{g} \in \ker_{\mathbb{Z}}(H_0)$ decomposes to $\sum_{i=1}^{N} \mathbf{e}_i$ for some $N \in \mathbb{Z}_{\geq 0}$ with $\mathbf{e}_i \in \ker_{\mathbb{Z}}(H_0)$ and $\|\mathbf{e}_i\|_\infty \leq \mathcal{O}_{FPT}(1)$ for each i.*

Proof. Since $H_0\mathbf{g} = 0$, we know that $F \cdot \mathbf{g} = 0$. Therefore, there exist $\alpha_j \in \mathbb{Z}_+$ and $\mathbf{g}_j(F) \sqsubseteq \mathbf{g}$ such that $\mathbf{g} = \sum_j \alpha_j \mathbf{g}_j(F)$, where $\mathbf{g}_j(F) \in \mathcal{G}(F)$. Consider each $\mathbf{g}_j(F)$. As F is a two-stage stochastic matrix (recall its definition in Eq (2)), by Lemma 12 it holds for every j that $\|\mathbf{g}_j(F)\|_\infty = \mathcal{O}_{FPT}(1)$. Note that each $\mathbf{g}_j(F)$ can be written into $n+1$ bricks such that $\mathbf{g}_j(F) = \big(\mathbf{g}_j^0(F), \mathbf{g}_j^1(F), \cdots, \mathbf{g}_j^n(F)\big)$ where $\mathbf{g}_j^0(F)$ is a t_B-dimensional vector, and $\mathbf{g}_j^i(F)$ is a t_A-dimensional vector for every $1 \leq i \leq n$. It is obvious that $\|\mathbf{g}_j^i(F)\|_\infty = \mathcal{O}_{FPT}(1)$ for every $0 \leq i \leq n$, and it holds that

$$B\mathbf{g}_j^0(F) + A\mathbf{g}_j^i(F) = 0, \quad \forall 1 \leq i \leq n.$$

The claim below follows from picking a suitable $\mathbf{v}_j^*$ such that $\mathbf{g}_j^i(F) - \mathbf{v}_j^*$ has "balanced" coefficients.

▷ Claim 14. For every $\mathbf{g}_j(F)$ and $1 \leq \ell \leq |\mathcal{G}(A)|$, there exist some $\mathbf{v}_j^*$ and $\alpha_{j,\ell}^i \in \mathbb{Z}_{\geq 0}$ with

- $\mathbf{g}_j^i(F) - \mathbf{v}_j^* = \sum_{\ell=1}^{|\mathcal{G}(A)|} \alpha_{j,\ell}^i \mathbf{g}_\ell(A), \quad \forall 1 \leq i \leq n.$
- For every $1 \leq \ell \leq |\mathcal{G}(A)|$, either $|\{i : \alpha_{j,\ell}^i > 0\}| = 0$, or $|\{i : \alpha_{j,\ell}^i > 0\}| \geq n/2$.
- $\max_{i,j,\ell} |\alpha_{j,\ell}^i| \leq \alpha_{\max}$, where $\alpha_{\max} = 2g_\infty(F) = \mathcal{O}_{FPT}(1)$
- $\|\mathbf{v}_j^*\|_\infty = \mathcal{O}_{FPT}(1)$.

Proof. Consider an arbitrary $\mathbf{v}_j$ such that $\begin{pmatrix} \mathbf{g}_j^0(F) \\ \mathbf{v}_j \end{pmatrix} \in \ker_{\mathbb{Z}}([B, A])$ and $\| \begin{pmatrix} \mathbf{g}_j^0(F) \\ \mathbf{v}_j \end{pmatrix} \|_\infty \leq g_\infty(F) = \alpha_{\max}/2$. Such $\mathbf{v}_j$ exists since $\mathbf{g}_j(F)$ satisfies that $B\mathbf{g}_j^0(F) + A\mathbf{g}_j^i(F) = 0$ for any $1 \leq i \leq n$. We have $A(\mathbf{g}_j^i(F) - \mathbf{v}_j) = 0$ for every $1 \leq i \leq n$, hence there exist $\bar{\alpha}_{j,\ell}^i \in \mathbb{Z}_+$, $\mathbf{g}_\ell(A) \in \mathcal{G}(A)$ and $\mathbf{g}_\ell(A) \sqsubseteq \mathbf{g}_j^i(F) - \mathbf{v}_j$ such that for some integer m,

$$\mathbf{g}_j^i(F) - \mathbf{v}_j = \sum_{\ell=1}^{m} \bar{\alpha}_{j,\ell}^i \mathbf{g}_\ell(A), \quad \forall 1 \leq i \leq n.$$

Since $\| \begin{pmatrix} \mathbf{g}_j^0(F) \\ \mathbf{v}_j \end{pmatrix} \|_\infty \leq \alpha_{\max}/2$, consequently $\|\mathbf{g}_j^i(F) - \mathbf{v}_j\|_\infty \leq \alpha_{\max}$, and $\bar{\alpha}_{j,\ell}^i \leq \alpha_{\max}$. Consider the cardinality of the set $\{i : \bar{\alpha}_{j,\ell}^i > 0\}$. If $1 \leq |\{i : \bar{\alpha}_{j,\ell}^i > 0\}| \leq \lfloor n/2 \rfloor$, we say ℓ is *unbalanced* for $\mathbf{g}_j(F)$. Let $\bar{\alpha}_{j,\max}^i = \max_{1 \leq \ell \leq m} \bar{\alpha}_{j,\ell}^i$ and UB_j be the set of all unbalanced indices ℓ, we define

$$\mathbf{v}_j^* := \mathbf{v}_j + \sum_{\ell \in UB_j} \bar{\alpha}_{j,\max}^i \mathbf{g}_\ell(A),$$

$$\text{then, } \mathbf{g}_j^i(F) - \mathbf{v}_j^* = \sum_{\ell \in \{1,2,\cdots,m\} \setminus UB_j} \bar{\alpha}_{j,\ell}^i \mathbf{g}_\ell(A) + \sum_{\ell \in UB_j} (\bar{\alpha}_{j,\max}^i - \bar{\alpha}_{j,\ell}^i) \cdot (-\mathbf{g}_\ell(A)), \quad \forall 1 \leq i \leq n.$$

Note that $-\mathbf{g}_\ell(A) \in \mathcal{G}(A)$. For all the $\mathbf{g}_\ell(A)$'s in $\mathcal{G}(A)$ that do not appear in the above equation, we take their coefficients as 0. Furthermore, we have $|\bar{\alpha}_{j,\ell}^i| \leq \alpha_{\max}$ and $|\bar{\alpha}_{j,\max}^i - \bar{\alpha}_{j,\ell}^i| \leq \alpha_{\max}$ for all i, ℓ. As $\|\mathbf{v}_j\|_\infty = \mathcal{O}_{FPT}(1)$, $\|\mathbf{g}_\ell(A)\|_\infty = \mathcal{O}_{FPT}(1)$, we know that $\|\mathbf{v}_j^*\|_\infty = \mathcal{O}_{FPT}(1)$. Thus, the claim is proved. ◁

We call $(\mathbf{g}_j^0(F), \mathbf{v}_j^*, \mathbf{v}_j^*, \cdots, \mathbf{v}_j^*)$ as a canonical vector (of $\mathbf{g}_j(F)$). It is easy to see that $F(\mathbf{g}_j^0(F), \mathbf{v}_j^*, \mathbf{v}_j^*, \cdots, \mathbf{v}_j^*) = 0$. Since $\|\mathbf{v}_j^*\|_\infty = \mathcal{O}_{FPT}(1)$ and $\|\mathbf{g}_j^0(F)\|_\infty = \mathcal{O}_{FPT}(1)$, there are at most $\tau = \mathcal{O}_{FPT}(1)$ different kinds of canonical vectors. This means, there may be different $\mathbf{g}_k(F)$'s with the same canonical vector. We list all the τ possible canonical vectors and let $\mathbf{r}_j := (\mathbf{p}_j^*, \mathbf{v}_j^*, \mathbf{v}_j^*, \cdots, \mathbf{v}_j^*)$ be the j-th one. Let CA_j be the set of indices of all $\mathbf{g}_k(F)$'s whose canonical vector is $\mathbf{r}_j$, then we have

$$\mathbf{g} = \sum_{j=1}^{\tau}(\sum_{k\in CA_j} \alpha_k)\mathbf{r}_j + \sum_{j=1}^{\tau}\sum_{k\in CA_j} \alpha_k\left(\mathbf{g}_k(F) - \mathbf{r}_j\right). \tag{4}$$

We say an n-dimensional vector $\boldsymbol{\alpha} = (\alpha^1, \alpha^2, \cdots, \alpha^n) \in \mathbb{Z}^n_{\geq 0}$ is *balanced*, if $\boldsymbol{\alpha} = 0$, or $\|\boldsymbol{\alpha}\|_\infty \leq \alpha_{\max} = \mathcal{O}_{FPT}(1)$ and $|\{i : \alpha^i > 0\}| \geq n/2$. Then the following observation is true.

▶ **Observation 15.** *For any nonzero balanced vector $\boldsymbol{\alpha}$ it holds that $\|\boldsymbol{\alpha}\|_1 \geq n/2 \cdot \alpha^i/\alpha_{\max}$ for every $1 \leq i \leq n$.*

Using the concept of a balanced vector, Claim 14 indicates that if $\mathbf{r}_j$ is a canonical vector of $\mathbf{g}_k(F)$, then $\mathbf{g}_k^i(F) - \mathbf{v}_j^* = \sum_{\ell=1}^{|\mathcal{G}(A)|} \alpha_{k,\ell}^i \mathbf{g}_\ell(A)$ such that the vector $(\alpha_{k,\ell}^1, \alpha_{k,\ell}^2, \cdots, \alpha_{k,\ell}^n)$ is a balanced vector. The nice thing about balanced vectors is that we can have the following claim, which will be used several times later.

▷ Claim 16. Let $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_k$ be a sequence of balanced vectors in $\mathbb{Z}^n_{\geq 0}$ such that $\|\sum_{h=1}^k \mathbf{y}_h\|_1 \leq n\Lambda$ where $\Lambda = \mathcal{O}_{FPT}(1)$, then $\|\sum_{h=1}^k \mathbf{y}_h\|_\infty \leq 2\alpha_{\max}\Lambda = \mathcal{O}_{FPT}(1)$.

Proof of Claim 16. We prove by contradiction. Suppose on the contrary that $\|\sum_{h=1}^k \mathbf{y}_h\|_\infty > 2\alpha_{\max}\Lambda$, then there exists some i^* such that $\sum_{h=1}^k \mathbf{y}_h^{i*} > 2\alpha_{\max}\Lambda$. Since $\mathbf{y}_h$'s are balanced vectors, according to Observation 15, we have

$$\|\sum_{h=1}^{k}\mathbf{y}_h\|_1 = \sum_{h=1}^{k}\|\mathbf{y}_h\|_1 \geq n \cdot \frac{\sum_{h=1}^k \mathbf{y}_h^{i*}}{2\alpha_{\max}} > n\Lambda,$$

which contradicts the fact that $\|\sum_{h=1}^k \mathbf{y}_h\|_1 \leq n\Lambda$. Hence, the claim is true. ◁

Since $\mathbf{r}_j$ is a canonical vector of $\mathbf{g}_k(F)$, by Claim 14, there exist balanced vectors $\boldsymbol{\beta}_{k,\ell}$ such that Eq (4) can be rewritten as (ignoring $\mathbf{g}^0$):

$$\mathbf{g}^i = \sum_{j=1}^{\tau}(\sum_{k\in CA_j} \alpha_k)\mathbf{v}_j^* + \sum_{j=1}^{\tau}\sum_{k\in CA_j} \alpha_k \left(\sum_{\ell=1}^{|\mathcal{G}(A)|} \beta_{k,\ell}^i \mathbf{g}_\ell(A)\right), \quad \forall 1 \leq i \leq n,$$

$$\text{or equivalently,} \quad \mathbf{g}^i = \sum_{j=1}^{\tau} \alpha_j' \mathbf{v}_j^* + \sum_{\ell=1}^{|\mathcal{G}(A)|} \beta_\ell^i \mathbf{g}_\ell(A), \quad \forall 1 \leq i \leq n, \tag{5}$$

where $\alpha_j' = \sum_{k\in CA_j} \alpha_k$ and each $\beta_\ell = (\beta_\ell^1, \cdots, \beta_\ell^n)$ is the summation of balanced vectors.

$$\text{As } [0, D, D, \cdots, D]\mathbf{g} = 0, \text{ we have } \sum_{j=1}^{\tau} n\alpha_j' D\mathbf{v}_j^* + \sum_{\ell=1}^{|\mathcal{G}(A)|}(\sum_{i=1}^{n}\beta_\ell^i) D\mathbf{g}_\ell(A) = 0. \tag{6}$$

Note that $|\mathcal{G}(A)| = \mathcal{O}_{FPT}(1)$, the equation above can be rewritten as

$$[D\mathbf{v}_1^*, \cdots, D\mathbf{v}_\tau^*, D\mathbf{g}_1(A), \cdots, D\mathbf{g}_{|\mathcal{G}(A)|}(A)] \cdot (n\alpha_1', \cdots, n\alpha_\tau', \sum_{i=1}^{n}\beta_1^i, \cdots, \sum_{i=1}^{n}\beta_{|\mathcal{G}(A)|}^i) = 0. \tag{7}$$

Let $V = [D\mathbf{v}_1^*, D\mathbf{v}_2^*, \cdots, D\mathbf{v}_\tau^*, D\mathbf{g}_1(A), D\mathbf{g}_2(A), \cdots, D\mathbf{g}_{|\mathcal{G}(A)|}(A)]$, which is an $\mathcal{O}_{FPT}(1) \times \mathcal{O}_{FPT}(1)$-matrix with $\|V\|_\infty = \mathcal{O}_{FPT}(1)$, then there exist $\lambda_k \in \mathbb{Z}_+$ and $\mathbf{g}_k(V) \in \mathcal{G}(V)$, such that $(n\alpha_1', n\alpha_2', \cdots, n\alpha_\tau', \sum_{i=1}^n \beta_1^i, \cdots, \sum_{i=1}^n \beta_{|\mathcal{G}(A)|}^i) = \sum_k \lambda_k \mathbf{g}_k(V)$. Note that since $\alpha_j', \beta_\ell^i \geq 0$, we can restrict that every $\mathbf{g}_k(V) \in \mathbb{Z}_{\geq 0}^{\tau+|\mathcal{G}(A)|}$.

For ease of description, from now on we take the viewpoint of a packing problem. We view each canonical vector $\mathbf{r}_j^*$ and $\mathbf{g}_\ell(A)$ as an item, whereas there are $\tau + |\mathcal{G}(A)|$ different kinds of items. There are $n+1$ different bins. Bin 0 can only be used to pack items $\mathbf{r}_j^*$, $1 \leq j \leq \tau$, and bin i ($1 \leq i \leq n$) can only be used to pack items $\mathbf{g}_\ell(A)$, $1 \leq \ell \leq |\mathcal{G}(A)|$. Currently there are α_j' copies of item $\mathbf{r}_j^*$ in bin 0, and β_ℓ^i copies of item $\mathbf{g}_\ell(A)$ in bin i. This is called a packing profile. Now we want to split this packing profile into several "sub-profiles", i.e., we want to determine integers $\mu_j^h, \sigma_\ell^{i,h} \in \mathbb{Z}_{\geq 0}$ such that the followings are true:

(i) $\mu_j^h, \sigma_\ell^{i,h} = \mathcal{O}_{FPT}(1)$ and $\mu_j^h + \sigma_\ell^{i,h} > 0$.

(ii) $\sum_h \mu_j^h = \alpha_j'$, $\sum_h \sigma_\ell^{i,h} = \beta_\ell^i$;

(iii) $[D\mathbf{v}_1^*, \cdots, D\mathbf{v}_\tau^*, D\mathbf{g}_1(A), \cdots, D\mathbf{g}_{|\mathcal{G}(A)|}(A)]$
$\cdot (n\mu_1^h, \cdots, n\mu_\tau^h, \sum_{i=1}^n \sigma_\ell^{i,h}, \cdots, \sum_{i=1}^n \sigma_{|\mathcal{G}(A)|}^{i,h}) = 0$ for every h.

A packing with μ_j^h copies of $\mathbf{r}_j^*$ in bin 0 and $\sigma_\ell^{i,h}$ copies of $\mathbf{g}_\ell(A)$ in bin i is called a sub-profile. Any sub-profile corresponds to a $(t_A + nt_B)$-dimensional vector $\mathbf{e}_h = (\mathbf{e}_h^0, \mathbf{e}_h^1, \cdots, \mathbf{e}_h^n)$ where

$$\mathbf{e}_h^0 = \sum_{j=1}^{\tau} \mu_j^h \mathbf{p}_j^*$$

$$\mathbf{e}_h^i = \sum_{j=1}^{\tau} \mu_j^h \mathbf{v}_j^* + \sum_{\ell=1}^{|\mathcal{G}(A)|} \sigma_\ell^{i,h} \mathbf{g}_\ell(A), \quad \forall 1 \leq i \leq n$$

If all the three conditions on sub-profiles hold, then we know that $\|\mathbf{e}_h\|_\infty = \mathcal{O}_{FPT}(1)$, $\mathbf{g} = \sum_h \mathbf{e}_h$ and $H_0\mathbf{e}_h = 0$ (to see why $H_0\mathbf{e}_h = 0$ holds, simply recall that $F\mathbf{r}_j^* = 0$ and condition (iii) implies that $[0, D, D, \cdots, D]\mathbf{e}_h = 0$), and furthermore, there are at most $\sum_j \alpha_j' + \sum_{i,\ell} \beta_\ell^i$ sub-profiles, which is finite. Hence, $\mathbf{g} = \sum_h \mathbf{e}_h$ and the theorem is proved.

We will construct $\mathbf{e}_h$'s iteratively. Once $\mathbf{e}_h$ is constructed, we continue our decomposition procedure on $\mathbf{g} - \sum_{k=1}^h \mathbf{e}_k$.

Suppose we have constructed $\mathbf{e}_1$ to $\mathbf{e}_{h_0-1}$ where conditions (i) and (iii) are satisfied for each $\mathbf{e}_h$, $\alpha_j' - \sum_{h=1}^{h_0-1} \mu_j^h \geq 0$, $\bar{\beta}_\ell^i := \beta_\ell^i - \sum_{h=1}^{h_0-1} \sigma_\ell^{i,h} \geq 0$ and furthermore, each vector $\bar{\boldsymbol{\beta}}_\ell = (\bar{\beta}_\ell^1, \cdots, \bar{\beta}_\ell^n)$ can be expressed as a summation of all but one balanced vectors, more precisely, there exist balanced vectors $\phi_{\ell,k} \in \mathbb{Z}_{\geq 0}^n$, $1 \leq k \leq k_{\max}$ such that

$$\bar{\beta}_\ell = \sum_{k=1}^{k_{\max}-1} \phi_{\ell,k} + \bar{\phi}_{\ell,k_{\max}}, \quad \text{where } \bar{\phi}_{\ell,k_{\max}} \sqsubseteq \phi_{\ell,k_{\max}}.$$

We show how to construct $\mathbf{e}_{h_0}$. Let $\bar{\alpha}_j' = \alpha_j' - \sum_{h=1}^{h_0-1} \mu_j^h$. According to condition (iii) of each $\mathbf{e}_h$, we know that

$$[D\mathbf{v}_1^*, \cdots, D\mathbf{v}_\tau^*, D\mathbf{g}_1(A), \cdots, D\mathbf{g}_{|\mathcal{G}(A)|}(A)] \cdot (n\bar{\alpha}_1', \cdots, n\bar{\alpha}_\tau', \sum_{i=1}^n \bar{\beta}_1^i, \cdots, \sum_{i=1}^n \bar{\beta}_{|\mathcal{G}(A)|}^i) = 0$$

Consequently, there exist $\lambda_k' \in \mathbb{Z}_{\geq 0}$ and $\mathbf{g}_k \in \mathbb{Z}_{\geq 0}^{\tau+|\mathcal{G}(A)|} \cap \mathcal{G}(V)$ such that

$$(n\bar{\alpha}_1', n\bar{\alpha}_2', \cdots, n\bar{\alpha}_\tau', \sum_{i=1}^n \bar{\beta}_1^i, \cdots, \sum_{i=1}^n \bar{\beta}_{|\mathcal{G}(A)|}^i) = \sum_k \lambda_k' \mathbf{g}_k(V).$$

There are two possibilities.

Case 1. If there exists some $\lambda'_k \geq n$, we consider the vector-summand $n\mathbf{g}_k(V)$ out of $\lambda'_k\mathbf{g}_k(V)$. Let $n\mathbf{g}_k(V) = (n\zeta_1, n\zeta_2, \cdots, n\zeta_{\tau+|\mathcal{G}(A)|})$. We set $\mu_j^{h_0} = \zeta_j = \mathcal{O}_{FPT}(1)$ for $1 \leq j \leq \tau$. We set the values of σ_ℓ^{i,h_0} such that $\sum_{i=1}^n \sigma_\ell^{i,h_0} = n\zeta_{\tau+\ell}$. Consequently, condition (iii) is satisfied for $\mathbf{e}_{h_0}$. Now it suffices to set the values of each σ_ℓ^{i,h_0} such that they are bounded by $\mathcal{O}_{FPT}(1)$. Equivalently, this means out of the $\bar{\beta}_\ell^i$ copies of $\mathbf{g}_\ell(A)$, our goal is to take σ_ℓ^{i,h_0} copies such that in total we take $n\zeta_{\tau+\ell}$ copies and $\sigma_\ell^{i,h_0} = \mathcal{O}_{FPT}(1)$. We achieve this in a simple greedy way. Let k^* be the index such that

$$\sum_{k=k^*+1}^{k_{\max}-1} \|\phi_{\ell,k}\|_1 + \|\bar{\phi}_{\ell,k_{\max}}\|_1 < n\zeta_{\tau+\ell} \leq \sum_{k=k^*}^{k_{\max}-1} \|\phi_{\ell,k}\|_1 + \|\bar{\phi}_{\ell,k_{\max}}\|_1$$

Let $\bar{\phi}_{\ell,k^*} \sqsubseteq \phi_{\ell,k^*}$ be an arbitrary vector such that

$$\|\bar{\phi}_{\ell,k^*}\|_1 + \sum_{k=k^*+1}^{k_{\max}-1} \|\phi_{\ell,k}\|_1 + \|\bar{\phi}_{\ell,k_{\max}}\|_1 = n\zeta_{\tau+\ell}.$$

We set $\sigma_\ell^{i,h_0} = \bar{\phi}_{\ell,k^*}^i + \sum_{k=k^*+1}^{k_{\max}-1} \phi_{\ell,k}^i + \bar{\phi}_{\ell,k_{\max}}^i$. It is obvious that in total we have taken $n\zeta_{\tau+\ell}$ copies of $\mathbf{g}_\ell(A)$. Now it remains to show that $\|\sigma_\ell^{h_0}\|_\infty = \|\bar{\phi}_{\ell,k^*} + \sum_{k=k^*+1}^{k_{\max}-1} \phi_{\ell,k} + \bar{\phi}_{\ell,k_{\max}}\|_\infty = \mathcal{O}_{FPT}(1)$. To see this, notice that each $\phi_{\ell,k}$ is a balanced vector, hence

$$\|\phi_{\ell,k^*}\|_1 + \sum_{k=k^*+1}^{k_{\max}-1} \|\phi_{\ell,k}\|_1 + \|\phi_{\ell,k_{\max}}\|_1 \leq n\zeta_{\tau+\ell} + 2n\alpha_{\max} = \mathcal{O}_{FPT}(n).$$

According to Claim 16, $\|\phi_{\ell,k^*} + \sum_{k=k^*+1}^{k_{\max}-1} \phi_{\ell,k} + \phi_{\ell,k_{\max}}\|_\infty = \mathcal{O}_{FPT}(1)$. Consequently, $\|\sigma_\ell^{h_0}\|_\infty = \mathcal{O}_{FPT}(1)$.

Also notice that after we take σ_ℓ^{i,h_0} copies of $\mathbf{g}_\ell(A)$, $\bar{\boldsymbol{\beta}}_\ell - \sigma_\ell^{h_0} = \sum_{k=1}^{k^*-1} \phi_{\ell,k} + (\phi_{\ell,k^*} - \bar{\phi}_{\ell,k^*})$, which is still the summation of all but one balanced vector. Hence we can continue to decompose $\mathbf{g} - \sum_{h=1}^{h_0} \mathbf{e}_h$.

Case 2. $\lambda'_k < n$ for every k. We claim that $\|\mathbf{g} - \sum_{h=1}^{h_0-1} \mathbf{e}_h\|_\infty = \mathcal{O}_{FPT}(1)$. If this claim is true, then $\mathbf{g} = \sum_{h=1}^{h_0-1} \mathbf{e}_h + (\mathbf{g} - \sum_{h=1}^{h_0-1} \mathbf{e}_h)$, and Theorem 4 is proved. To show the claim, we use a similar argument as that of case 1. First, $n\bar{\alpha}'_j \leq (\sum_k \lambda_k) \cdot \max_k \|\mathbf{g}_k(V)\|_\infty = \mathcal{O}_{FPT}(n)$, hence $\bar{\alpha}'_j = \mathcal{O}_{FPT}(1)$. Second, we consider the n-dimensional vector $\boldsymbol{\beta} = \sum_{\ell=1}^{|\mathcal{G}(A)|} \boldsymbol{\beta}_\ell$. Recall that $\bar{\beta}_\ell^i := \beta_\ell^i - \sum_{h=1}^{h_0-1} \sigma_\ell^{i,h} \geq 0$ and each vector $\bar{\boldsymbol{\beta}}_\ell$ satisfy that

$$\bar{\beta}_\ell = \sum_{k=1}^{k_{\max}-1} \phi_{\ell,k} + \bar{\phi}_{\ell,k_{\max}}$$

where $\bar{\phi}_{\ell,k_{\max}} \sqsubseteq \phi_{\ell,k_{\max}}$. Let $\bar{\beta}'_\ell = \sum_{k=1}^{k_{\max}} \phi_{\ell,k}$ and $\boldsymbol{\beta}' = \sum_{\ell=1}^{|\mathcal{G}(A)|} \boldsymbol{\beta}'_\ell$. Given that $\bar{\phi}_{\ell,k_{\max}} \sqsubseteq \phi_{\ell,k_{\max}}$ and $\phi_{\ell,k_{\max}}$ is a balanced vector, $\|\bar{\beta}'_\ell\|_1 \leq \|\bar{\beta}_\ell\|_1 + n\alpha_{\max}$. Consequently

$$\begin{aligned}\|\beta'\|_1 \leq \sum_{\ell=1}^{|\mathcal{G}(A)|} \|\bar{\beta}'_\ell\|_1 \quad &\leq \sum_{\ell=1}^{|\mathcal{G}(A)|} \|\bar{\beta}_\ell\|_1 + n\alpha_{\max} \cdot |\mathcal{G}(A)| \\ &\leq \sum_k \lambda'_k \cdot \max_k \|\mathbf{g}_k(V)\|_1 + n\alpha_{\max} \cdot |\mathcal{G}(A)| = \mathcal{O}_{FPT}(n).\end{aligned}$$

Note that $\boldsymbol{\beta}'$ is the summation of balanced vectors. According to Claim 16, $\|\boldsymbol{\beta}'\|_\infty = \mathcal{O}_{FPT}(1)$, consequently $\|\boldsymbol{\beta}\|_\infty \leq \|\boldsymbol{\beta}'\|_\infty = \mathcal{O}_{FPT}(1)$. Combining the fact that $\|\mathbf{p}_j^*\|_\infty = \mathcal{O}_{FPT}(1)$, $\|\mathbf{v}_j^*\|_\infty = \mathcal{O}_{FPT}(1)$ and $\|\mathbf{g}_\ell(A)\|_\infty = \mathcal{O}_{FPT}(1)$, we have $\|\mathbf{g} - \sum_{i=1}^{h_0-1} \mathbf{e}_h\|_\infty = \mathcal{O}_{FPT}(1)$. ◀

Theorem 4 indicates that, there exists some "basis" for 3-block n-fold IP with FPT-bounded ℓ_∞-norms. Unfortunately, this basis need not be Graver basis; indeed, we will show later that the Graver basis of 3-block n-fold IP does not have an FPT-bounded ℓ_∞-norm. However, Theorem 4 provides a new perspective on the structure of the kernel space, which can be utilized to bound the ℓ_∞-norm of the Graver basis through a "merging" technique for the proof of Theorem 5 as we illustrate in the following subsection.

4.2 A sign-compatible decomposition

We have shown in the previous subsection that any element of $\ker_{\mathbb{Z}}(H_0)$ admits a decomposition into lattice elements whose ℓ_∞-norm is bounded by $\mathcal{O}_{FPT}(1)$. However, this decomposition is not necessarily "sign-compatible", meaning that possibly none of its elements is a feasible step on its own, which makes its immediate algorithmic use complicated. Towards the algorithm for 3-block n-fold IP, we resort to Graver basis. The goal of this subsection is to prove the following theorem.

▶ **Theorem 5.** *For any* 3*-block* n*-fold matrix* H_0*,* $g_\infty(H_0) \le \mathcal{O}_{FPT}(n^{t_A^2+1})$*.*

Following the line of arguments in previous papers [4, 16, 18, 20], it seems very difficult to derive an upper bound singly exponential in t_A. To prove Theorem 5, we use a completely different approach. We give a brief overview of the proof idea. The reader is referred to the full version of this paper [6] for details.

Proof idea. The basic idea is to show that if $\|\mathbf{g}(H_0)\|_\infty$ is too large for some $\mathbf{g}(H_0) \in \mathcal{G}(H_0)$, then we are able to find some $\mathbf{z} \sqsubset \mathbf{g}(H_0)$ and $H_0\mathbf{z} = 0$, contradicting the fact that $\mathbf{g}(H_0)$ is a Graver basis element. Suppose $\mathbf{y} = \mathbf{g}(H_0)$ and $\|\mathbf{y}\|_\infty$ is very large. The crucial idea is that we do not search directly for $\mathbf{z} \sqsubset \mathbf{y}$, but rather search for $\mathbf{z} \sqsubset \tilde{\mathbf{y}}$ where $\tilde{\mathbf{y}}$ is an "equalization", of $\mathbf{y}$, and then prove that such a $\mathbf{z}$ also satisfies that $\mathbf{z} \sqsubset \mathbf{y}$.

Roughly speaking, we will divide the n bricks of $\mathbf{y}$, i.e., $\mathbf{y}^i$ for $i = 1, 2, \cdots, n$, into $\sigma = \mathcal{O}_{FPT}(1)$ groups $N_1, N_2, \cdots, N_\sigma$ such that for any $k \in N_j$, $\tilde{\mathbf{y}}^k \approx \frac{1}{|N_j|}\sum_{i \in N_j} \mathbf{y}^i$. Why do we need to take such a detour in the proof? The problem is that by directly arguing on $\mathbf{y}$ we run into a bound that is similar as [16]. Therefore, we use a completely different approach – we adopt the decomposition of Theorem 4, and then modify such a decomposition into a sign-compatible one by "merging" summands. Towards this, we first prove a merging lemma (see Lemma 5 of the full version) which states that given a summation of a sequence of vectors, we can always divide them into disjoint subsets such the summation of vectors in each subset becomes sign-compatible. The merging lemma can turn an arbitrary decomposition into a sign-compatible one, despite the fact that the cardinality of each subset is exponential in the dimension of the vectors (which means the ℓ_∞-norm of the summands will explode by a factor that is exponential in the dimension). Consequently, if we directly merge the $\mathcal{O}_{FPT}(n)$-dimensional vectors in the decomposition of Theorem 4, we get a very weak bound. To handle this, we consider an alternative sum $\tilde{\mathbf{y}}$, which is derived by averaging multiple bricks of $\mathbf{y}$ as we mentioned above.

By altering the decomposition of $\mathbf{y}$, we get a decomposition of $\tilde{\mathbf{y}}$ such that the following is true: all the $n+1$ bricks of each vector-summand can be divided into $\mathcal{O}_{FPT}(1)$ subsets where in each subset the bricks are identical. This indicates that, although we are summing up $\mathcal{O}_{FPT}(n)$-dimensional vectors to $\tilde{\mathbf{y}}$, it is essentially the same as summing up $\mathcal{O}_{FPT}(1)$-dimensional vectors. Such a transformation comes at a cost – summands summing up to $\tilde{\mathbf{y}}$ do not have $\mathcal{O}_{FPT}(1)$-bounded ℓ_∞-norms, indeed, each vector-summand consists of n bricks whose ℓ_∞-norm is $\mathcal{O}_{FPT}(1)$, and at most 1 brick (which is a t_A-dimensional vector) whose ℓ_∞-norm is $\mathcal{O}_{FPT}(n)$. Applying our merging lemma, we derive a sign-compatible decomposition of $\tilde{\mathbf{y}}$ where the summands have an ℓ_∞-norm bounded by $\mathcal{O}_{FPT}(n^{t_A^2+1})$.

It remains to show that at least one summand $\mathbf{z}$ in the sign-compatible decomposition of $\tilde{\mathbf{y}}$ also satisfies that $\mathbf{z} \sqsubset \mathbf{y}$. To show this we need to go back to the definition of $\tilde{\mathbf{y}}$. We are averaging bricks of $\mathbf{y}$, but which bricks shall we average? Each brick is a t_A-dimensional vector and we consider each coordinate. We set up a threshold Γ. If the absolute value of a coordinate is larger than Γ, we say it is (positive or negative) large. Otherwise it is small. Therefore, each brick can be characterized by identifying its coordinates being positive large, negative large or small (which is defined as the *quantity type* of a brick). We only average the bricks of the same quantity type. By doing so, we can ensure that $\tilde{\mathbf{y}}^i$ is roughly sign-compatible with $\mathbf{y}^i$ – if the j-th coordinate of $\mathbf{y}^i$ is positive or negative large, then this coordinate of $\tilde{\mathbf{y}}^i$ is also positive or negative. Hence, any $\mathbf{z} \sqsubset \tilde{\mathbf{y}}^i$ is almost sign-compatible with $\mathbf{y}$ – indeed, if we can ensure additionally that the j-th coordinate of $\mathbf{z}^i$ is 0 as long as the j-th coordinate of $\mathbf{y}^i$ is small, then we can conclude that $\mathbf{z} \sqsubset \mathbf{y}$. This "if" can be proved using a counting argument, and we get Theorem 5. ◀

5 4-block n-fold IP reduces to 3-block n-fold IP

In this section, we will show that for any 4-block n-fold IP, there exists an equivalent 3-block n-fold IP which is *kernel preserving*, as we define in the following.

▶ **Definition 17** (Extended formulation). *Let $n' \geq n$, $m' \in \mathbb{N}$, $\mathcal{A} \in \mathbb{Z}^{m\times n}$, $\mathbf{b} \in \mathbb{Z}^m$, $\mathbf{l}, \mathbf{u} \in (\mathbb{Z} \cup \{\pm\infty\})^n$ and $\mathcal{A}' \in \mathbb{Z}^{m'\times n'}$, $\mathbf{b}' \in \mathbb{Z}^{m'}$, $\mathbf{l}', \mathbf{u}' \in (\mathbb{Z} \cup \{\pm\infty\})^{n'}$. We say that*

$$\mathcal{A}'(\mathbf{x}, \mathbf{y}) = \mathbf{b}',\ \mathbf{l}' \leq (\mathbf{x}, \mathbf{y}) \leq \mathbf{u}' \tag{EF}$$

is an extended formulation *of*

$$\mathcal{A}\mathbf{x} = \mathbf{b},\ \mathbf{l} \leq \mathbf{x} \leq \mathbf{u} \tag{OrigF}$$

if $\{\mathbf{x} \mid \mathcal{A}\mathbf{x} = \mathbf{b}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}\} = \{\mathbf{x} \mid \exists \mathbf{y} : \mathcal{A}'(\mathbf{x}, \mathbf{y}) = \mathbf{b}', \mathbf{l}' \leq (\mathbf{x}, \mathbf{y}) \leq \mathbf{u}'\}$.

▶ **Definition 18** (Feasibly kernel-preserving extended formulation). *We say that* (EF) *is a* feasibly kernel preserving *extended formulation of* (OrigF) *if for each* $(\mathbf{x}, \mathbf{y})$ *feasible in* (EF)*,*

$$\mathcal{A}'(\mathbf{g}, \mathbf{h}) = \mathbf{0},\ \mathbf{l}' \leq (\mathbf{x}, \mathbf{y}) + (\mathbf{g}, \mathbf{h}) \leq \mathbf{u}' \qquad \Longrightarrow \qquad \mathcal{A}\mathbf{g} = \mathbf{0},\ \mathbf{l} \leq \mathbf{x} + \mathbf{g} \leq \mathbf{u},$$

that is, each element $(\mathbf{g}, \mathbf{h})$ *of* $\ker(\mathcal{A}')$ *which is feasible with respect to* $(\mathbf{x}, \mathbf{y})$ *corresponds to an element* $\mathbf{g} \in \ker(\mathcal{A})$ *which is feasible with respect to* $\mathbf{x}$.

Extended formulations are commonly used to show how a set of solutions can be embedded in an *extended space*, perhaps using less inequalities or obeying some extra structural requirements. The basic observation is that if we take an objective function f over the original formulation (OrigF) and optimize $f'(\mathbf{x}, \mathbf{y}) = f(\mathbf{x})$ over (EF), the optimal solution $(\mathbf{x}, \mathbf{y})$ over (EF) is such that $\mathbf{x}$ is an optimum over (OrigF). In the subsequent theorem we will use it to show that any 4-block n-fold IP can be embedded in a 3-block n-fold IP without blowing up the block sizes too much. The specific notion of a feasibly kernel-preserving extended formulation is useful to show that also our lower bounds on lattice elements are transferred, as we will show subsequently in Theorem 6.

Now we come to the main result of this subsection.

▶ **Theorem 19.** *Any 4-block n-fold IP with parameters $\Delta, s_A, s_B, s_C, s_D, t_A, t_B, t_C, t_D$ has a feasibly kernel-preserving extended formulation whose constraint matrix is a 3-block n-fold matrix with parameters $\hat{\Delta}, \hat{s}_A, \hat{s}_B, \hat{s}_D, \hat{t}_A, \hat{t}_B, \hat{t}_D$ satisfying*

$$\hat{\Delta} = \Delta \quad \hat{t}_A = \hat{t}_D = 2t_C + t_D + s_A \quad \hat{t}_B = t_B \quad \hat{s}_A = \hat{s}_B = s_B + t_C \quad \hat{s}_D = s_D = s_C\ .$$

Proof. Let us construct a 3-block n-fold IP instance which models the given 4-block IP instance. It's matrix $\hat{H}_0$ is a 3-block n-fold matrix composed of blocks $\hat{A}, \hat{B}$ and $\hat{D}$, and the remaining data is $\hat{\mathbf{b}}, \hat{\mathbf{l}}, \hat{\mathbf{u}}$ and $\hat{\mathbf{w}}$. Let the blocks be defined as follows.

$$\hat{D} = (C \ D \ \mathbf{0} \ \mathbf{0}) \qquad \hat{A} = \left(\begin{smallmatrix} -I & \mathbf{0} & I & \mathbf{0} \\ \mathbf{0} & A & \mathbf{0} & I \end{smallmatrix} \right) \qquad \hat{B} = \left(\begin{smallmatrix} I \\ B \end{smallmatrix} \right)$$

We call the four block columns of $\hat{A}$ and $\hat{D}$ *subbricks* and index them by greek letters α, β, γ and δ, i.e., $\mathbf{x}^{1\alpha}$ is the α-subbrick of the first brick.

Now, we add an extra brick which we call an *aggregation* brick, denoted $\mathbf{x}^d$ where $d = n+1$. The idea is that the α subbrick is non-zero only at the aggregation brick and corresponds to the first-stage variables of the original 4-block n-fold IP. We shall ensure that this is true using lower and upper bounds. However, to subsequently "assign" the aggregated values to the first stage variables, we also need to modify the B block, which, in turn, forces us to introduce new slack variables. This is the meaning of the γ subbrick (slack variables for bricks $i \neq d$) and δ subbrick (slack variables for the dth brick).

The right hand side $\hat{\mathbf{b}}$ is simply $\hat{\mathbf{b}}^0 = \mathbf{b}^0$ and $\hat{\mathbf{b}}^i = (\mathbf{0} \ \mathbf{b}^i)$ for $i \neq d$ and $\hat{\mathbf{b}}^d = (\mathbf{0} \ \mathbf{0})$. We set the new lower and upper bounds $\hat{\mathbf{l}}, \hat{\mathbf{u}}$ as follows:

α subbrick $\hat{\mathbf{l}}^{i\alpha} = \hat{\mathbf{u}}^{i\alpha} = \mathbf{0}$ for all $i \neq d$, and $\hat{\mathbf{l}}^{d\alpha} = -\infty$, $\hat{\mathbf{u}}^{d\alpha} = +\infty$. This ensures the α subbrick to be only possibly non-zero in brick d.

β subbrick $\hat{\mathbf{l}}^{i\beta} = \mathbf{l}^i$ and $\hat{\mathbf{u}}^{i\beta} = \mathbf{u}^i$ for all $i \neq d$ and $\hat{\mathbf{l}}^{d\beta} = \hat{\mathbf{u}}^{d\beta} = \mathbf{0}$. This ensures that the β subbrick has the meaning of the original variables $\mathbf{x}^i$ for all bricks except brick d, where we enforce $\hat{\mathbf{x}}^{d\beta} = \mathbf{0}$.

γ subbrick $\hat{\mathbf{l}}^{d\gamma} = \hat{\mathbf{u}}^{d\gamma} = \mathbf{0}$ and $\hat{\mathbf{l}}^{i\gamma} = -\infty$, $\hat{\mathbf{u}}^{i\gamma} = +\infty$ for $i \neq d$. Without these variables and due to the structure of $\hat{A}$ and $\hat{B}$, we would be enforcing for *each* brick $i \neq d$ that $\hat{\mathbf{x}}^0 = \hat{\mathbf{x}}^{i\alpha}$, and since $\hat{\mathbf{x}}^{i\alpha} = \mathbf{0}$ this would mean $\hat{\mathbf{x}}^0 = \mathbf{0}$. The γ subbrick relaxes this to $\hat{\mathbf{x}}^0 = \hat{\mathbf{x}}^{i\alpha} + \hat{\mathbf{x}}^{i\gamma} = \mathbf{0} + \hat{\mathbf{x}}^{i\gamma}$ which is trivially satisfiable considering our setting of the bounds $\hat{\mathbf{l}}^{d\gamma}$ and $\hat{\mathbf{u}}^{d\gamma}$.

δ subbrick $\hat{\mathbf{l}}^{i\delta} = \hat{\mathbf{u}}^{i\delta} = \mathbf{0}$ for all $i \neq d$, and $\hat{\mathbf{l}}^{d\delta} = -\infty$, $\hat{\mathbf{u}}^{d\delta} = +\infty$, i.e., the same as for the α subbrick. Similarly to the γ subbrick, without the δ subbrick we would be enforcing $B\hat{\mathbf{x}}^0 + A\hat{\mathbf{x}}^{d\beta} = \mathbf{0}$, however $\hat{\mathbf{x}}^{d\beta} = \mathbf{0}$ so we would be forcing $B\hat{\mathbf{x}}^0 = \mathbf{0}$, which is undesired. Thus we relax it to $B\hat{\mathbf{x}}^0 + A\hat{\mathbf{x}}^{d\beta} + \hat{\mathbf{x}}^{d\delta} = B\hat{\mathbf{x}}^0 + \hat{\mathbf{x}}^{d\delta} = \mathbf{0}$ which is trivially satisfiable.

To show that the constructed system

$$H^0\hat{\mathbf{x}} = \hat{\mathbf{b}}, \, \hat{\mathbf{l}} \leq \hat{\mathbf{x}} \leq \hat{\mathbf{u}}, \, \hat{\mathbf{x}} \in \mathbb{Z}^{\hat{t}_B + (n+1)\hat{t}_A} \tag{8}$$

is truly an extended formulation of $H\mathbf{x} = \mathbf{b}, \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \mathbf{x} \in \mathbb{Z}^{t_C + nt_A}$, let us define a projection $\pi : \mathbb{Z}^{\hat{t}_B + (n+1)\hat{t}_A} \to \mathbb{Z}^{t_C + nt_A}$ which defines the mapping from the extended formulation to the original instance. Specifically, we let

$$\pi((\hat{x}^0, \hat{x}^{1\alpha}, \hat{x}^{1\beta}, \hat{x}^{1\gamma}, \hat{x}^{1\delta}, \hat{x}^{2\alpha}, \ldots, \hat{x}^{n\delta}, \hat{x}^{d\alpha}, \ldots, \hat{x}^{d\delta}) = (\hat{x}^0, \hat{x}^{1\beta}, \hat{x}^{2\beta}, \ldots, \hat{x}^{n\beta}) \ .$$

By the arguments above we see that $\hat{\mathbf{x}}^0$ has precisely the meaning of $\mathbf{x}^0$ and $\hat{\mathbf{x}}^{i\beta}$ for $i \neq d$ has the meaning of $\mathbf{x}^i$.

Finally, let us argue that this extended formulation is also feasibly kernel-preserving. Consider now a feasible solution $\hat{\mathbf{x}}$ of (8), and consider any $\hat{\mathbf{g}}$ in $\ker(H_0)$ such that $\hat{\mathbf{x}} + \hat{\mathbf{g}}$ is again feasible. We have to show that $H\pi(\hat{\mathbf{x}}) = \mathbf{0}$ and $\mathbf{l} \leq \mathbf{x} + \pi(\hat{\mathbf{x}}) \leq \mathbf{u}$. The latter follows easily from the fact that $\hat{\mathbf{l}} \leq \hat{\mathbf{x}} + \hat{\mathbf{g}} \leq \hat{\mathbf{u}}$ and that $\pi(\hat{\mathbf{l}}) = \mathbf{l}$ and $\pi(\hat{\mathbf{u}}) = \mathbf{u}$. To see the former, consider separately first the upper row $(C \ D \ \cdots \ D)$ of H and after that the remaining rows.

We have that

$$C\hat{\mathbf{x}}^{d\alpha} + D\hat{\mathbf{x}}^{d\beta} + \mathbf{0}\hat{\mathbf{x}}^{d\gamma} + \mathbf{0}\hat{\mathbf{x}}^{d\delta} + \sum_{i=1}^{n} C\hat{\mathbf{x}}^{i\alpha} + D\hat{\mathbf{x}}^{i\beta} + \mathbf{0}\hat{\mathbf{x}}^{i\gamma} + \mathbf{0}\hat{\mathbf{x}}^{i\delta} = \mathbf{0} \ .$$

Omitting the zero blocks, we obtain

$$C\hat{\mathbf{x}}^{a\alpha} + D\hat{\mathbf{x}}^{a\beta} + \sum_{i=1}^{n} C\hat{\mathbf{x}}^{i\alpha} + D\hat{\mathbf{x}}^{i\beta} = \mathbf{0} \ .$$

Recall that our bounds enforce $\hat{\mathbf{x}}^{d\beta} = \mathbf{0}$ and $\hat{\mathbf{x}}^{i\alpha} = \mathbf{0}$ for $i \neq d$, and finally $\hat{\mathbf{x}}^{0} = \hat{\mathbf{x}}^{d\alpha}$, so plugging these in we obtain

$$C\hat{\mathbf{x}}^{0} + \sum_{i=1}^{n} D\hat{\mathbf{x}}^{i\beta} = \mathbf{0},$$

which by the definition of π implies that $C\pi(\mathbf{x})^0 + \sum_{i=1}^{n} D\pi(\mathbf{x})^i = \mathbf{0}$ as desired. Now it is left to show that, for each $i \neq d$, $B\pi(\mathbf{x})^0 + A\pi(\mathbf{x})^i = \mathbf{0}$. We have that

$$B\hat{\mathbf{x}}^{0} + \mathbf{0}\hat{\mathbf{x}}^{i\alpha} + A\hat{\mathbf{x}}^{i\beta} + \mathbf{0}\hat{\mathbf{x}}^{i\gamma} + I\hat{\mathbf{x}}^{i\delta} = \mathbf{0} \ .$$

Omitting the zero blocks and recalling that our bounds enforce $\hat{\mathbf{x}}^{i\delta} = \mathbf{0}$ for each $i \neq d$, we have

$$A\hat{\mathbf{x}}^{0} + A\hat{\mathbf{x}}^{i\beta} = \mathbf{0},$$

which, by definition of π, is what we wanted to show. ◀

▶ Remark 20. Theorem 19 has several consequences. One is that 4-block n-fold IP is in FPT if and only if 3-block n-fold IP is in FPT. Furthermore, as the reduction is kernel preserving, we can also utilize Theorem 19 to transfer the Graver basis elements between 4-block n-fold IP and 3-block n-fold IP, as is implied by Theorem 6.

▶ **Theorem 6.** *For arbitrary integer $t \in \mathbb{N}$, there exists a 3-block n-fold IP with a matrix H such that $s_i, t_i \in O(t)$ for $i = A, B, C, D$, and for any* feasible *nonzero $\mathbf{g} \in \ker_{\mathbb{Z}}(H_0)$ we have $\|\mathbf{g}^0\|_\infty = \Omega(n^t)$.*

Proof. Consider the instance constructed in Theorem 2 with H being the 4-block n-fold matrix from the proof. Apply Theorem 19 to this instance to obtain its feasible kernel-preserving extended formulation, which is a 3-block n-fold IP, and consider any $\hat{\mathbf{x}}$ which is a feasible solution for it. Denote by π the projection from the proof of Theorem 19.

Now let $\hat{\mathbf{g}} \in \ker_{\mathbb{Z}}(H_0) \subseteq \ker(H_0)$ be feasible with respect to $\hat{\mathbf{x}}$. By Definition 18, we have $\mathbf{g} = \pi(\hat{\mathbf{g}}) \in \ker_{\mathbb{Z}}(H)$, and by Theorem 2 we have $\|\mathbf{g}\|_\infty = \Omega(n^{t-1})$ and in particular $\|\mathbf{g}^0\|_\infty = \Omega(n^{t-1})$. By the definition of π these lower bounds transfer to $\hat{\mathbf{g}}$ and $\hat{\mathbf{g}}^0$. ◀

▶ Remark 21. The reader may wonder what if we take a "fat" kernel element of a 4-block n-fold IP, which cannot be decomposed into "thin" kernel element whose infinity norm bounded by $\mathcal{O}_{FPT}(1)$, then use Theorem 19 to construct an equivalent 3-block n-fold IP, and apply Theorem 4 to decompose the kernel element of the 3-block n-fold IP into thin elements and transform them back to the original 4-block n-fold IP. This seems to suggest that Theorem 19 is contradicting Theorem 2 and Theorem 4. We emphasize that such a contradiction does not exist. Indeed, our definition of a feasibly-preserving extended formulation is such that it takes kernel elements from 3-block n-fold IP to 4-block n-fold IP *only* if they are feasible in the 3-block n-fold IP. Note that the construction of Theorem 19 requires specific lower and upper bounds on the extended variables $\mathbf{y}$. This is where the non-conformality of the decomposition of Thm 4 comes into play: what happens is that we decompose a kernel element of 3-block into “thin” kernel elements, however, we cannot take them back to the 4-block because they do not satisfy the bounds on the extended variables, and thus the definition of feasibly kernel-preserving extended formulation does not guarantee anything for them.

6 Algorithms

Using the upper bound on the Graver basis elements, we can derive algorithms for 3-block and 4-block n-fold IP by combining the idea from [16] and the recent progress in [28, 8], as indicated by Theorem 7 and Theorem 3.

In the following we prove Theorem 7. Theorem 3 can be proved by plugging in the upper bound of 4-block n-fold IP and proceed with the same argument.

▶ **Theorem 7.** *3-block n-fold IP can be solved in time* $\min\{\mathcal{O}_{FPT}(n^{O(s_D t_B)}, \mathcal{O}_{FPT}(n^{O(t_A^2 t_B)})\}$.

Proof. Using the idea of approximate Graver-best oracle introduced by Altmanová et al. [1] and implicitly by Eisenbrand et al. [8], it suffices for us to solve the following IP for each fixed value $\rho_0 = 2^0, 2^1, 2^2, \cdots$:

$$\min\{\mathbf{w}\mathbf{x} : H_0\mathbf{x} = 0, \mathbf{l} \le \mathbf{x}_0 + \rho_0\mathbf{x} \le \mathbf{u}, \mathbf{x} \in \mathbb{Z}^m, \|\mathbf{x}\|_\infty \le \min\{\mathcal{O}_{FPT}(n^{s_c}), \mathcal{O}_{FPT}(n^{t_A^2+1})\}\}$$

Let $\mathbf{x}_*$ be the optimal solution. Given that $\|\mathbf{x}_*\|_\infty \le \mathcal{O}_{FPT}(n^{t_A^2+1})$, we can guess $\mathbf{x}_*^0$ and there are $\mathcal{O}_{FPT}(n^{(t_A^2+1)t_B})$ different possibilities. For each guess, say, $\mathbf{x}_*^0 = \mathbf{v}$, we solve the following problem:

$$\min\{\mathbf{w}\cdot\mathbf{x} : H_0\mathbf{x} = 0, \mathbf{l} \le \mathbf{x}_0 + \rho_0\mathbf{x} \le \mathbf{u}, \mathbf{x} \in \mathbb{Z}^m, \mathbf{x}^0 = \mathbf{v}\}$$

By fixing $\mathbf{x}^0$, the above problem becomes exactly an n-fold IP, which can be solved efficiently in $\mathcal{O}_{FPT}(n^2 \log n^2)$ time [8]. Notice that ρ_0 may take $\mathcal{O}_{FPT}(n \log n)$ distinct values, the overall running time is $\min\{\mathcal{O}_{FPT}(n^{s_D t_B+3}) \log^3 n, \mathcal{O}_{FPT}(n^{(t_A^2+1)t_B+3} \log^3 n)\}$. ◀

7 Conclusion

We consider 4-block n-fold IP and its important special case 3-block n-fold IP, both generalizing the two-stage stochastic IP and n-fold IP. We show that lattice elements of 3-block n-fold IP admit a decomposition whose ℓ_∞-norm is bounded in $\mathcal{O}_{FPT}(1)$, while any non-zero integral element in the kernel of 4-block n-fold IP may have an ℓ_∞-norm at least $\Omega(n^{s_c})$. We provide a matching upper bound on the ℓ_∞-norm of the Graver basis for 4-block n-fold IP, which gives an exponential improvement upon the best known result. We also establish an upper bound of $\min\{\mathcal{O}_{FPT}(n^{s_c}), \mathcal{O}_{FPT}(n^{t_A^2} + 1)\}$ on the ℓ_∞-norm of the Graver basis for 3-block n-fold IP. A remaining important open problem is whether 4-block n-fold IP, or equivalently, 3-block n-fold IP, is in FPT. Our lower bounds give some indication that this is unlikely.

References

1 Katerina Altmanová, Dusan Knop, and Martin Koutecký. Evaluating and tuning n-fold integer programming. *ACM Journal of Experimental Algorithmics*, 24(1):2.2:1–2.2:22, 2019. doi:10.1145/3330137.

2 Gautam Appa, Balázs Kotnyek, Konstantinos Papalamprou, and Leonidas Pitsoulis. Optimization with binet matrices. *Operations research letters*, 35(3):345–352, 2007.

3 Stephan Artmann, Robert Weismantel, and Rico Zenklusen. A strongly polynomial algorithm for bimodular integer linear programming. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017*, pages 1206–1219. ACM, 2017.

4 Matthias Aschenbrenner and Raymond Hemmecke. Finiteness theorems in stochastic integer programming. *Foundations of Computational Mathematics*, 7(2):183–227, 2007.

5 Lin Chen and Daniel Marx. Covering a tree with rooted subtrees–parameterized and approximation algorithms. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2801–2820. SIAM, 2018.

6 Lin Chen, Lei Xu, Weidong Shi, and Martin Koutecký. New bounds on augmenting steps of block-structured integer programs. *arXiv preprint*, 2018. arXiv:1805.03741.

7 Jesús A De Loera, Raymond Hemmecke, Shmuel Onn, and Robert Weismantel. N-fold integer programming. *Discrete Optimization*, 5(2):231–241, 2008.

8 Friedrich Eisenbrand, Christoph Hunkenschröder, and Kim-Manuel Klein. Faster algorithms for integer programs with block structure. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 49:1–49:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ICALP.2018.49.

9 Friedrich Eisenbrand, Christoph Hunkenschröder, Kim-Manuel Klein, Martin Koutecký, Asaf Levin, and Shmuel Onn. An algorithmic theory of integer programming. *arXiv preprint*, 2019. arXiv:1904.01361.

10 Friedrich Eisenbrand and Robert Weismantel. Proximity results and faster algorithms for integer programming using the steinitz lemma. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 808–816. SIAM, 2018.

11 Elisabeth Finhold and Raymond Hemmecke. Lower bounds on the graver complexity of m-fold matrices. *Annals of Combinatorics*, 20(1):73–85, 2016.

12 Robert Ganian and Sebastian Ordyniak. The complexity landscape of decompositional parameters for ILP. *Artificial Intelligence*, 2018.

13 Robert Ganian, Sebastian Ordyniak, and MS Ramanujan. Going beyond primal treewidth for (m) ilp. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

14 Jack E Graver. On the foundations of linear and integer linear programming i. *Mathematical Programming*, 9(1):207–226, 1975.

15 Victor S Grinberg and Sergey V Sevast'yanov. Value of the steinitz constant. *Functional Analysis and Its Applications*, 14(2):125–126, 1980.

16 Raymond Hemmecke, Matthias Köppe, and Robert Weismantel. Graver basis and proximity techniques for block-structured separable convex integer minimization problems. *Mathematical Programming*, 145(1-2):1–18, 2014.

17 Raymond Hemmecke, Shmuel Onn, and Lyubov Romanchuk. N-fold integer programming in cubic time. *Mathematical Programming*, 137(1-2):325–341, 2013.

18 Raymond Hemmecke, Shmuel Onn, and Robert Weismantel. A polynomial oracle-time algorithm for convex integer minimization. *Mathematical Programming*, 126(1):97–117, 2011.

19 Raymond Hemmecke and Rüdiger Schultz. Decomposition methods for two-stage stochastic integer programs. In *Online optimization of large scale systems*, pages 601–622. Springer, 2001.

20 Serkan Hoşten and Seth Sullivant. A finiteness theorem for markov bases of hierarchical models. *Journal of Combinatorial Theory, Series A*, 114(2):311–321, 2007.

21 Klaus Jansen, Kim-Manuel Klein, Marten Maack, and Malin Rau. Empowering the configuration-ip - new PTAS results for scheduling with setups times. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 44:1–44:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ITCS.2019.44.

22 Klaus Jansen and Lars Rohwedder. On integer programming and convolution. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, volume 124 of *LIPIcs*, pages 43:1–43:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.ITCS.2019.43.

23 Richard M Karp. Reducibility among combinatorial problems. In *Complexity of computer computations*, pages 85–103. Springer, 1972.

24 Kim-Manuel Klein. About the complexity of two-stage stochastic ips. In Daniel Bienstock and Giacomo Zambelli, editors, *Integer Programming and Combinatorial Optimization - 21st International Conference, IPCO 2020, London, UK, June 8-10, 2020, Proceedings*, volume 12125 of *Lecture Notes in Computer Science*, pages 252–265. Springer, 2020. doi:10.1007/978-3-030-45771-6_20.

25 Dusan Knop and Martin Koutecký. Scheduling meets n-fold integer programming. *J. Sched.*, 21(5):493–503, 2018. doi:10.1007/s10951-017-0550-0.

26 Dusan Knop, Martin Kouteckỳ, and Matthias Mnich. Voting and bribing in single-exponential time. In *LIPIcs-Leibniz International Proceedings in Informatics*, volume 66. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

27 Dušan Knop, Martin Kouteckỳ, and Matthias Mnich. Combinatorial n-fold integer programming and applications. *Mathematical Programming*, pages 1–34, 2019.

28 Martin Koutecký, Asaf Levin, and Shmuel Onn. A parameterized strongly polynomial algorithm for block structured integer programs. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, volume 107 of *LIPIcs*, pages 85:1–85:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.ICALP.2018.85.

29 Shmuel Onn. Nonlinear discrete optimization. *Zurich Lectures in Advanced Mathematics, European Mathematical Society*, 2010.

30 Christos H Papadimitriou. On the complexity of integer programming. *Journal of the ACM (JACM)*, 28(4):765–768, 1981.