

AN ENHANCED PRIMAL-SIMPLEX BASED TARDOS' ALGORITHM FOR LINEAR OPTIMIZATION

Shinji Mizuno
Tokyo Institute of Technology

Noriyoshi Sukegawa
Chuo University

Antoine Deza
McMaster University / Université Paris Sud

(Received March 28, 2017; Revised October 6, 2017)

Abstract While the algorithmic complexity is in general worse than the one of Tardos' original algorithms, the authors, motivated by the practicality of such methods, recently proposed a simplex-based variant that is strongly polynomial if the coefficient matrix is totally unimodular and the auxiliary problems are non-degenerate. In this paper, we introduce a slight modification that circumvents the determination of the largest sub-determinant while keeping the same theoretical performance. Assuming that the coefficient matrix is integer-valued and the auxiliary problems are non-degenerate, the proposed algorithm is polynomial in the dimension of the input data and the largest absolute value of a sub-determinant of the coefficient matrix.

Keywords: Linear programming, Tardos' algorithm, simplex method, strong polynomiality

1. Introduction, Main Result, and Related Work

1.1. Introduction

Linear optimization deals with the minimization problem: $\min\{\mathbf{c}^\top \mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ where the coefficient matrix $A \in \mathbb{R}^{m \times n}$, the right hand side vector $\mathbf{b} \in \mathbb{R}^m$, and the objective function vector $\mathbf{c} \in \mathbb{R}^n$ are given data. The celebrated Tardos' algorithm [24, 25] for linear optimization runs in polynomial time in m , n , and the size L_A of A . We recall that the size of a matrix $A = [a_{ij}]$ is defined for integer A as $L_A = mn + \sum_{i,j} (1 + \lceil \log_2(|a_{ij}| + 1) \rceil)$. Thus, Tardos' algorithm is strongly-polynomial if L_A is polynomial in m and n which is the case for combinatorial problems such as minimum cost flow, bipartite matching, multicommodity flow, and vertex packing in chordal graphs.

A key element of Tardos' algorithm is to identify the coordinates equal to zero at optimality by solving several auxiliary dual problems via an ellipsoid or interior-point method. Considering only the primal problem, Orlin [17] proposed a modification of Tardos' algorithm which specifically identifies the coordinates strictly positive at optimality. Mizuno [15] modified Tardos' algorithm by using a dual simplex method to solve the auxiliary problems. Assuming that A is integer-valued and the auxiliary problems are non-degenerate, Mizuno's algorithm runs in polynomial time in m , n , and Δ_A . We recall that Δ_A denotes the largest absolute value of a sub-determinant of A and that the non-degeneracy assumption holds if the basic variables are strictly positive for every basic feasible solution. In particular, Mizuno's algorithm is strongly polynomial if A is totally unimodular and non-degeneracy holds. Note that the complexity analysis uses Kitahara and Mizuno's bounds [11, 12] which depend on the values of the entries rather than on the data length. Thus, the complexity of Mizuno's algorithm depends on Δ_A rather than on L_A . Combining Orlin's and Mizuno's approaches, the authors introduced a primal-simplex based Tardos' algorithm with the same

theoretical complexity as Mizuno's algorithm, see [16]. Tardos' algorithm and the mentioned modifications by Orlin, Mizuno, and Mizuno et al. are of rather theoretical interest. In particular, the determination of Δ_A might be challenging as the naive upper bound of $m!A_{\max}^m$ is typically impractically large; we recall that $A_{\max} = \max |a_{i,j}|$. In addition, the coefficients of the auxiliary problems might be impractically large too. For instance, the size of the coefficients in Orlin's algorithm can be m times larger than those in Tardos' or Mizuno's algorithm. The complexity analysis of the Mizuno et al. algorithm requires total unimodularity for A . Zadeh [29] showed that the original network simplex method may require an exponential number of pivots for some specific min-cost flow problems; that is, the simplex method may be inefficient even if A is totally unimodular. This issue was addressed by Orlin [18] who proposed a variant of the network simplex method which is strongly polynomial for min-cost flow problems.

1.2. Main result

We propose an enhanced primal-simplex based Tardos' algorithm circumventing the computation of Δ_A while achieving a strong-polynomial complexity for a slightly wider class of problems. The enhanced algorithm is obtained by modifying the auxiliary problem used in the Mizuno et al. algorithm. Assuming A is integer-valued and the auxiliary problems are non-degenerate, the enhanced Mizuno et al. algorithm is polynomial in m , n , and Δ_A . Thus the strong polynomiality holds for a slightly larger class than totally unimodular matrices, e.g. a coefficient matrix resulting from the addition to a totally unimodular matrix of a fixed number of rows (constraints) with entries polynomially bounded in m and n . The determination of Δ_A is circumvented via a simple search procedure and the practicality of the algorithm improves as the coefficients of the auxiliary problems are typically substantially smaller in the enhanced Mizuno et al. algorithm. Note that the results hold under dual degeneracy via lexicographic treatment of the basis exchange in Dantzig's rule. After recalling some related work in Section 1, the pre-processing and reformulations into auxiliary problems are presented in Section 2. Sections 3 and 4 outline the proposed algorithm: the main procedure – which requires the determination of Δ_A – and, then, the enhanced algorithm which circumvents the determination of Δ_A . The correctness and the complexity analysis of the algorithm are proven in Sections 5 and 6.

1.3. Related work

We recall a few results dealing with the geometry, combinatorics and computational aspects of linear optimization. Finding a good bound on the maximal diameter $\Delta(n, m)$ of the vertex-edge graph of a polytope in terms of its dimension n and the number of its facets m is one of the basic open questions in polytope theory. Although some bounds are known, the behaviour of the function $\Delta(n, m)$ is largely unknown. The Hirsch conjecture, formulated in 1957 and reported in [7], states that $\Delta(n, m)$ is linear in m and n : $\Delta(n, m) \leq m - n$. The conjecture was recently disproved by Santos [20]. However, the asymptotic behaviour of $\Delta(n, m)$ is not well understood: the best upper bounds — due to Kalai and Kleitman [10], Todd [27], and the subsequent improvements [22, 23] — are quasi-polynomial. The behaviour of $\Delta(n, m)$ is historically closely connected with the theory of the simplex method as $\Delta(n, m)$ is a lower bound for the worst complexity of simplex methods. Bonifas et al. [3] showed that the diameter is an $O(n^4 \Delta_A^2 \log(n \Delta_A))$ extending the previous bound of $O(m^{16} n^3 (\log mn)^3)$ by Dyer and Frieze [8] for totally unimodular instances. Dadush and Hähnle [6] used another parameter associated to the coefficient matrix A , called the *curvature* δ_A , to analyze the behaviour of the shadow simplex method. They showed that the expected number of pivots of the shadow simplex method is an $O(\frac{n^3}{\delta_A} \log \frac{n}{\delta_A})$. Both δ_A and Δ_A can be regarded as a

measure of how well-conditioned A is. Note that $1/\delta_A \leq n\Delta_A^2$. Under a non-degeneracy assumption weaker than the one assumed for the complexity analysis of our algorithm, Bilen, Csizmadia, and Illés [2] proposed an algorithm for linear feasibility which runs in polynomial time in m and Δ where Δ involves A and \mathbf{b} , and hence differs from Δ_A . Bilen, Csizmadia, and Illés' algorithm is a version of the simplex method with a special pivot rule for dealing with degenerate bases, see Csizmadia [5] for further details. Vavasis and Ye [28] proposed a primal-dual path-following interior-point algorithm with an $O(n^{3.5} \log(n\bar{\chi}_A))$ iteration complexity bound where $\bar{\chi}_A$ can be regarded as condition number associated with A . Megiddo et al. [13] proposed a modification that circumvents the computation of $\bar{\chi}_A$ to enhance the implementability. Another variant of Vavasis and Ye algorithm was proposed by Monteiro and Tsuchiya [14]

In a similar fashion, we circumvent the determination of Δ_A while Megiddo et al. circumvent the determination of $\bar{\chi}_A$. While we assume non-degeneracy, Dadush and Hähnle algorithm is non-deterministic. In practice, degenerate pivots are typically rare and our algorithm may exhibit reasonable performance under moderate degeneracy. The proposed algorithm may visit infeasible points as, for example, Anstreicher and Terlaky's monotonic build-up simplex and Paparizos exterior point simplex, or Fukuda and Terlaky's criss-cross method; see [1, 9, 19, 26] and references therein.

2. Pre-processing and Reformulation via Auxiliary Problems

We consider the following linear optimization formulation:

$$\begin{aligned} & \text{minimize} && \mathbf{c}^\top \mathbf{x} \\ & \text{subject to} && A\mathbf{x} = \mathbf{b}, \\ & && \mathbf{x} \geq \mathbf{0} \end{aligned} \tag{2.1}$$

where $A \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, and $\mathbf{c} \in \mathbb{R}^n$ are given. In addition, A is assumed to have full row rank m and be integer. We recall that A_{\max} and Δ_A , respectively, denote the largest absolute value of an entry and a sub-determinant of A . Note that $A_{\max} \leq \Delta_A$ and $\Delta_A \leq m!A_{\max}^m$.

2.1. Pre-processing and problem reformulations: reduction and scaling

An optimal solution of (2.1), if any, is assumed without loss of generality to be unique. Otherwise \mathbf{c} could be perturbed by $(\epsilon, \epsilon^2, \dots, \epsilon^n)$ for a sufficiently small $\epsilon > 0$. Such perturbations have no impact on the analysis of the proposed algorithm as it is based on the results of Kitahara and Mizuno [11, 12] which are independent of \mathbf{c} . From an algorithmic viewpoint, perturbations are not required as one can instead consider a lexicographical order if a tie occurs when choosing the entering variable via the simplex method with Dantzig's rule.

Let $K^* \subseteq N = \{1, 2, \dots, n\}$ be the set of indices i such that $x_i^* > 0$ for the optimal solution \mathbf{x}^* of (2.1) – which is assumed to exist. The proposed algorithm inductively builds a subset $\bar{K} \subseteq K^*$ through solving auxiliary problems. If $\bar{K} = K^*$, we obtained the optimal solution. Otherwise, we obtain a smaller, yet equivalent, problem by deleting the variables corresponding to \bar{K} . We first observe that (2.1) is equivalent to:

$$\begin{aligned} & \text{minimize} && \mathbf{c}_{\bar{K}}^\top \mathbf{x}_{\bar{K}} + \mathbf{c}_K^\top \mathbf{x}_K \\ & \text{subject to} && A_{\bar{K}} \mathbf{x}_{\bar{K}} + A_K \mathbf{x}_K = \mathbf{b}, \\ & && \mathbf{x}_{\bar{K}} \text{ free, } \mathbf{x}_K \geq \mathbf{0} \end{aligned} \tag{2.2}$$

where $K = N \setminus \bar{K}$ and \bar{K} is an arbitrary subset of K^* , and X_M denotes the restriction of X to the coordinates belonging to M .

The *reduced problem* (2.3) is obtained by eliminating free variables in $\mathbf{x}_{\bar{K}}$ as follows. Let G be a $m \times m$ sub-matrix of A such that the first $|\bar{K}|$ columns form $A_{\bar{K}}$, and $H = G^{-1}$. The Gaussian elimination for $A\mathbf{x} = \mathbf{b}$ of the variables x_i for $i \in \bar{K}$ is performed via $HA\mathbf{x} = H\mathbf{b}$. Let H_1 consist of the first $|\bar{K}|$ rows of H , and H_2 denote the remainder. The equality $HA\mathbf{x} = H\mathbf{b}$ yields:

$$\begin{aligned} H_1 A_{\bar{K}} \mathbf{x}_{\bar{K}} + H_1 A_K \mathbf{x}_K &= H_1 \mathbf{b}, \\ H_2 A_{\bar{K}} \mathbf{x}_{\bar{K}} + H_2 A_K \mathbf{x}_K &= H_2 \mathbf{b} \end{aligned}$$

where $H_1 A_{\bar{K}} = I$ and $H_2 A_{\bar{K}} = 0$ by the definition of H . Hence, the reduced problem is:

$$\begin{aligned} \text{minimize} \quad & (\mathbf{c}_K^\top - \mathbf{c}_{\bar{K}}^\top H_1 A_K) \mathbf{x}_K + \mathbf{c}_{\bar{K}}^\top H_1 \mathbf{b} \\ \text{subject to} \quad & H_2 A_K \mathbf{x}_K = H_2 \mathbf{b}, \\ & \mathbf{x}_K \geq \mathbf{0}. \end{aligned} \quad (2.3)$$

Let \mathbf{x}_K be an optimal solution of (2.3). Then $\mathbf{x} = (\mathbf{x}_{\bar{K}}, \mathbf{x}_K)$ with $\mathbf{x}_{\bar{K}} = H_1 \mathbf{b} - H_1 A_K \mathbf{x}_K$ is an optimal solution of (2.2). Setting $A' = H_2 A_K$, $\mathbf{b}' = H_2 \mathbf{b}$, $\mathbf{c}' = \mathbf{c}_K - A_K^\top H_1^\top \mathbf{c}_{\bar{K}}$, and $\mathbf{x}' = \mathbf{x}_K$, the reduced problem (2.3) is reformulated as the standard linear optimization problem (2.4) where the constant in the objective function is omitted:

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}'^\top \mathbf{x}' \\ \text{subject to} \quad & A' \mathbf{x}' = \mathbf{b}', \\ & \mathbf{x}' \geq \mathbf{0}. \end{aligned} \quad (2.4)$$

Observe that HA has full row rank and thus A' too. Problems (2.1) and (2.4) are equivalent: If L^* is an optimal basis of (2.4), then $\bar{K} \cup L^*$ is an optimal basis of (2.1). Let $m' = m - |\bar{K}|$, respectively $n' = n - |\bar{K}|$, denote the number of equality constraints and variables of (2.4). To obtain the desired auxiliary problem, we rewrite (2.4) and get a simplex *tableau* with respect to some basis $L \subseteq K$ of A' and set $\bar{L} = K \setminus L$ as follows:

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}'^\top \mathbf{x}' \\ \text{subject to} \quad & \mathbf{x}'_L + (A'_L)^{-1} A'_L \mathbf{x}'_{\bar{L}} = (A'_L)^{-1} \mathbf{b}', \\ & \mathbf{x}' \geq \mathbf{0}. \end{aligned} \quad (2.5)$$

Considering a scaling factor $\kappa = \|A'^\top (A' A'^\top)^{-1} \mathbf{b}'\|_2 / (m n n' A_{\max} \Delta + m')$ for some Δ , yields the following scaled problem:

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}'^\top \mathbf{x}' \\ \text{subject to} \quad & \mathbf{x}'_L + (A'_L)^{-1} A'_L \mathbf{x}'_{\bar{L}} = (A'_L)^{-1} \mathbf{b}' / \kappa, \\ & \mathbf{x}' \geq \mathbf{0}. \end{aligned} \quad (2.6)$$

The scaling factor κ is always strictly positive in our algorithm and, thus, a basis is optimal for (2.6) if and only if it is optimal for (2.4) and for the simplex tableau (2.5). Note that since A' has full row rank, $A' A'^\top$ is positive definite and thus $(A' A'^\top)^{-1}$ is well defined.

2.2. Auxiliary problem

The auxiliary problem is obtained from (2.6) by rounding up the right hand side vector where $\lceil \mathbf{a} \rceil$ denotes the vector whose i -th coordinate is the smallest integer not less than the i -th coordinate of \mathbf{a} :

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}'^\top \mathbf{x}' \\ \text{subject to} \quad & \mathbf{x}'_L + (A'_L)^{-1} A'_L \mathbf{x}'_{\bar{L}} = \lceil (A'_L)^{-1} \mathbf{b}' / \kappa \rceil, \\ & \mathbf{x}' \geq \mathbf{0}, \end{aligned} \quad (2.7)$$

Note that a feasible basis of (2.6) is feasible for (2.7) as (2.7) is a relaxation of (2.6). The key feature of (2.7) is that the coordinates of the right hand side vectors are scaled integers, see Lemma 6. Thus, (2.7) can be solved efficiently by the simplex method, yielding the strong polynomiality analysis.

3. Main Procedure

The main procedure of the algorithm is frequently called to solve (2.1) and involves, as subroutine, a two-phase simplex method to solve (2.7), see Section 3.1. While $\Delta \geq \Delta_A$ guarantees that (2.1) is solved by the main procedure, (2.1) may be solved even if $\Delta < \Delta_A$. The enhanced primal-simplex based Tardos' algorithm is presented in Section 4.

3.1. Two-phase simplex method TWOS((2.1); \mathbf{F} , \bar{K}^*)

Input: Problem (2.1).

Output: \mathbf{F} which is either INFEASIBLE OR UNBOUNDED or FEASIBLE AND FINITE, and an optimal basis \bar{K}^* of (2.1) if $\mathbf{F} = \text{FEASIBLE AND FINITE}$.

Phase I: Solve the following auxiliary problem via the simplex method with Dantzig's rule: a non-negative slack variable is added for each constraint and the sum of the slacks is minimized. The optimal value σ of this auxiliary problem is zero if and only if (2.1) is feasible. Output INFEASIBLE OR UNBOUNDED for \mathbf{F} if $\sigma > 0$. If $\sigma = 0$, the associated optimal basis yields a feasible basis \bar{K}^0 for (2.1) used to initialize Phase II.

Phase II: Starting from \bar{K}^0 , solve (2.1) via the simplex method with Dantzig's rule. Output INFEASIBLE OR UNBOUNDED for \mathbf{F} if (2.1) is unbounded; otherwise output FEASIBLE AND FINITE and an optimal basis \bar{K}^* for (2.1) is obtained.

3.2. Main procedure PROC((2.1), Δ ; \mathbf{F} , \bar{K}^*)

Input: Problem (2.1) and $\Delta > 0$.

Output: \mathbf{F} which is either INFEASIBLE OR UNBOUNDED, FEASIBLE AND FINITE, DEGENERATE, or UNIDENTIFIED, and an optimal basis \bar{K}^* for (2.1) if $\mathbf{F} = \text{FEASIBLE AND FINITE}$.

Initialization $\bar{K} := \emptyset$.

Step 1: If $\bar{K} \neq \emptyset$, remove the variables x_i in (2.1) for all $i \in \bar{K}$ to obtain the reduced problem (2.4). If $\bar{K} = \emptyset$, set $A' = A$, $\mathbf{b}' = \mathbf{b}$, $\mathbf{c}' = \mathbf{c}$, and $\mathbf{x}' = \mathbf{x}$. Go to Step 2.

Step 2: Consider the simplex tableau (2.5) associated to a basis L of the reduced problem (2.3). If $(A'_L)^{-1}\mathbf{b}' = \mathbf{0}$, output $\mathbf{F} = \text{DEGENERATE}$. Otherwise, determine $\kappa = \|A'^\top (A'A'^\top)^{-1}\mathbf{b}'\|_2 / (m'nn'A_{\max}\Delta + m')$ and obtain the auxiliary problem (2.7). Go to Step 3.

Step 3: Perform TWOS((2.7); \mathbf{F} , L^*). Output \mathbf{F} if $\mathbf{F} = \text{INFEASIBLE OR UNBOUNDED}$. Otherwise, let \mathbf{x}'' be the optimal solution of (2.7) associated to the optimal basis L^* . Output $\mathbf{F} = \text{FEASIBLE AND FINITE}$ and set $\bar{K}^* = \bar{K} \cup L^*$ if $\bar{K} \cup L^*$ is an optimal basis for (2.1). Otherwise, go to Step 4.

Step 4: Update $\bar{K} := \bar{K} \cup J$ with $J = \{i \in L^* \mid x''_i \geq m'nA_{\max}\Delta\}$. If $|\bar{K}| = m$, output $\mathbf{F} = \text{UNIDENTIFIED}$. Otherwise, go to Step 1.

We show in Section 4 that PROC((2.1), Δ ; \mathbf{F} , \bar{K}^*) can be used to solve (2.1) if $\Delta \geq \Delta_A$, and thus to extend the Mizuno et al. [16] primal-simplex based Tardos' algorithm.

3.2.1. Annotations for PROC((2.1), Δ ; \mathbf{F} , \bar{K}^*)

We outline the stopping criteria before providing additional details about the main procedure.

- (i) If $(A'_L)^{-1}\mathbf{b}' = \mathbf{0}$, the original problem (2.1) is degenerate, and (2.3) is either unbounded or admits zero as an optimal solution.

(ii) If (2.7) is infeasible then (2.1) is infeasible, and if (2.7) is unbounded then (2.1) is unbounded or infeasible. Indeed, (2.7) being a relaxation of (2.6) for any Δ , the infeasibility of (2.7) implies the infeasibility of (2.6) and of (2.3), and thus of (2.1). If (2.7) is unbounded, then (2.6) is unbounded or infeasible, and thus (2.1) is unbounded or infeasible. The set J defined in Step 4 satisfies $J \neq \emptyset$ (see Lemma 1), and thus the main procedure is finite as at most m auxiliary problems are solved. As shown in Corollary 1, if $\Delta \geq \Delta_A$, then $x_i^* > 0$ for $i \in J$ with x^* the optimal solution of (2.1). Thus, Corollary 1 shows that $J \subset K^*$ and validates Step 4, and hence, the correctness of the main procedure for $\Delta \geq \Delta_A$. As the main procedure is guaranteed to solve (2.1) only if $\Delta \geq \Delta_A$, F is set UNIDENTIFIED if $\Delta < \Delta_A$. However, the correct solution may be obtained even if $\Delta < \Delta_A$. For example, if $\bar{K} = \emptyset$; i.e. no reduction is performed in Step 4, and an optimal basis for (2.7) turns out to be feasible for (2.1) in Step 3, then this basis is optimal for (2.1) as (2.7) is a relaxation of (2.1).

3.2.2. Warm start for PROC((2.1), Δ ; F, \bar{K}^*)

Although the main procedure builds the simplex tableau (2.5) and the reduced problem (2.3) from scratch at each iteration, it is only to clarify the exposition. Indeed, in practice, we observe that $L^* \setminus J$ can serve as the basis L for (3) at the next iteration, thus enabling a warm start – as already noticed in the Mizuno et al. algorithm [16].

4. An Enhanced Primal-simplex Based Tardos' Algorithm

The proposed algorithm circumvents the determination of Δ_A via a simple search procedure in the following algorithm ALG((2.1), Δ^0 , λ ; F, \bar{K}^*) where, typically, one can use $\Delta^0 = 1$ and $\lambda = mA_{\max}$. Assuming non-degeneracy and Δ_A being polynomially bounded in m and n , the proposed algorithm is strongly polynomial – as shown in Theorem 2.

Input: Problem (2.1), $\Delta^0 > 0$, and $\lambda > 1$.

Output: F which is either INFEASIBLE OR UNBOUNDED, DEGENERATE, or FEASIBLE AND FINITE and an optimal basis \bar{K}^* for (2.1) if F=FEASIBLE AND FINITE.

Initialization $\Delta := \Delta^0$.

Step 1: Perform PROC((2.1), Δ ; F, \bar{K}^*). Output F if F=INFEASIBLE OR UNBOUNDED or F=DEGENERATE. Output F and \bar{K}^* if F=FEASIBLE AND FINITE. Otherwise, go to Step 2.

Step 2: Update $\Delta := \lambda\Delta$. Go to Step 1.

Theorem 1. *The enhanced primal-simplex based Tardos' algorithm ALG((2.1), Δ^0 , λ ; F, \bar{K}^*) solves (2.1).*

Theorem 1 is a consequence of (i) Δ eventually satisfies $\Delta \geq m!A_{\max}^m \geq \Delta_A$, and (ii) $J \subset K^*$ for $\Delta \geq \Delta_A$ as shown in Corollary 1; that is, Step 4 is valid.

Theorem 2. *ALG((2.1), $\Delta^0 = 1, \lambda = mA_{\max}; F, \bar{K}^*$) performs PROC((2.1), Δ ; F, \bar{K}^*) at most $m + 1$ times. PROC((2.1), Δ ; F, \bar{K}^*) performs TWOS((2.7); F, \bar{K}^*) at most m times. If all the auxiliary problems are non-degenerate, the number of arithmetic operations used by ALG((2.1), $\Delta^0 = 1, \lambda = mA_{\max}; F, \bar{K}^*$) to solve (2.1) is polynomial in m, n , and Δ_A .*

The first statement of Theorem 2 is implied by the stopping criterion $\Delta \geq m!A_{\max}^m$ and the setting $\Delta^0 = 1$ and $\lambda = mA_{\max}$. As mentioned in Section 3.2.1, the second statement of Theorem 2 is implied by Lemma 1. Thus, to complete the proof of Theorem 2 one has to show that TWOS((2.7); F, \bar{K}^*) is polynomial in m, n , and Δ_A – as proved in Section 6.

Instances of coefficient matrices with Δ_A polynomial in m and n include the one associated to capacitated network flow problems with additional linear constraints considered by Chen and Saigal [4]. The coefficient matrix they consider consists in the incidence matrix

of a directed network, and thus totally unimodular, to which a fixed number of arbitrary linear constraints on arc flow are added – assuming the entries are polynomial in m and n .

5. Proof of Theorem 1

Lemma 1 implies that, for any Δ , PROC((2.1), Δ ; F, \bar{K}^*) performs TWOS((2.7); F, \bar{K}^*) at most m times.

Lemma 1. *For any $\Delta > 0$, a basic solution \mathbf{x}'' of the auxiliary problem (2.7) satisfies $\|\mathbf{x}''\|_\infty \geq m'nA_{\max}\Delta$. Hence, the set J defined in Step 4 satisfies $J \neq \emptyset$.*

Proof. Let \mathbf{x}'' be a solution of (2.7). We have $A'\mathbf{x}'' = A'_L[(A'_L)^{-1}\mathbf{b}'/\kappa]$, $A'A'^\top$ is positive definite, and, for any \mathbf{g} , $A'^\top(A'A'^\top)^{-1}\mathbf{g}$ is the minimal l_2 -norm point satisfying $A'\mathbf{x} = \mathbf{g}$. Thus,

$$\begin{aligned} \|\mathbf{x}''\|_2 &\geq \|A'^\top(A'A'^\top)^{-1}A'_L[(A'_L)^{-1}\mathbf{b}'/\kappa]\|_2 \\ &\geq \|A'^\top(A'A'^\top)^{-1}\mathbf{b}'/\kappa\|_2 - \|A'^\top(A'A'^\top)^{-1}A'_L\mathbf{d}\|_2 \\ &= (m'nn'A_{\max}\Delta + m') - \left\| A'^\top(A'A'^\top)^{-1}A' \begin{pmatrix} \mathbf{d} \\ \mathbf{0}_{\bar{L}} \end{pmatrix} \right\|_2 \\ &\geq m'nn'A_{\max}\Delta + m' - \left\| \begin{pmatrix} \mathbf{d} \\ \mathbf{0}_{\bar{L}} \end{pmatrix} \right\|_2 \\ &= m'nn'A_{\max}\Delta + m' - \|\mathbf{d}\|_2 \end{aligned}$$

where $\kappa = \|A'^\top(A'A'^\top)^{-1}\mathbf{b}'\|_2/(m'nn'A_{\max}\Delta + m')$ and $\mathbf{d} = [(A'_L)^{-1}\mathbf{b}'/\kappa] - (A'_L)^{-1}\mathbf{b}'/\kappa$. Since $\|\mathbf{d}\|_\infty < 1$ and $\|\mathbf{d}\|_2 \leq m'\|\mathbf{d}\|_\infty$, we obtain: $\|\mathbf{x}''\|_\infty \geq \|\mathbf{x}''\|_2/n' > (m'nn'A_{\max}\Delta + m' - m')/n' = m'nA_{\max}\Delta$. \square

Applying a key result of Schrijver, recalled in Lemma 2, to (2.6) and (2.7) yields Lemma 3, whose direct consequence, Corollary 1, guarantees $J \subset K^*$, i.e. Step 4 is valid, for $\Delta \geq \Delta_A$.

Lemma 2 ([21], Theorem 10.5). *Let A be an $m \times n$ -matrix, and let Δ^* be such that for each nonsingular submatrix B of A all entries of B^{-1} are at most Δ^* in absolute value. Let \mathbf{c} be a column n -vector, and let \mathbf{b}'' and \mathbf{b}^* be column m -vectors such that $P'' : \max\{\mathbf{c}^\top\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}''\}$ and $P^* : \max\{\mathbf{c}^\top\mathbf{x} \mid A\mathbf{x} \leq \mathbf{b}^*\}$ are finite. Then, for each optimal solution \mathbf{x}'' of P'' , there exists an optimal solution \mathbf{x}^* of P^* with $\|\mathbf{x}'' - \mathbf{x}^*\|_\infty \leq n\Delta^*\|\mathbf{b}'' - \mathbf{b}^*\|_\infty$.*

Lemma 3. *Assume that the scaled problem (2.6) and the auxiliary problem (2.7) are both feasible and finite. Then, for an optimal solution \mathbf{x}'' of (2.7), there exists an optimal solution \mathbf{x}^* of (2.6) such that $\|\mathbf{x}'' - \mathbf{x}^*\|_\infty \leq n\Delta_A\|A_L\mathbf{d}\|_\infty$ with $\mathbf{d} = [(A'_L)^{-1}\mathbf{b}'/\kappa] - (A'_L)^{-1}\mathbf{b}'/\kappa$.*

Proof. Let \mathbf{x}'' be an optimal solution of (2.7). Then, $\tilde{\mathbf{x}}'' = (\tilde{\mathbf{x}}''_{\bar{K}}, \tilde{\mathbf{x}}''_K)$, with $\tilde{\mathbf{x}}''_K = \mathbf{x}''$ and $\tilde{\mathbf{x}}''_{\bar{K}} = H_1(\mathbf{b}/\kappa + A_L\mathbf{d}) - H_1A_K\mathbf{x}''$, is an optimal solution of:

$$\begin{aligned} &\text{minimize} && \mathbf{c}'^\top \mathbf{x}_K \\ &\text{subject to} && \mathbf{x}_{\bar{K}} + H_1A_K\mathbf{x}_K = H_1(\mathbf{b}/\kappa + A_L\mathbf{d}), \\ &&& H_2A_K\mathbf{x}_K = H_2(\mathbf{b}/\kappa + A_L\mathbf{d}), \\ &&& \mathbf{x}_K \geq \mathbf{0}. \end{aligned}$$

Multiplying both sides of the equations from the left by $G = H^{-1}$, and recalling the definitions of H_1 and H_2 given in Section 2.1, yields:

$$\begin{aligned} &\text{minimize} && \mathbf{c}'^\top \mathbf{x}_K \\ &\text{subject to} && A_{\bar{K}}\mathbf{x}_{\bar{K}} + A_K\mathbf{x}_K = \mathbf{b}/\kappa + A_L\mathbf{d}, \\ &&& \mathbf{x}_K \geq \mathbf{0}. \end{aligned}$$

By Lemma 2, there exists an optimal solution $\tilde{\mathbf{x}}^* = (\tilde{\mathbf{x}}_{\bar{K}}^*, \tilde{\mathbf{x}}_K^*)$ of:

$$\begin{aligned} & \text{minimize} && \mathbf{c}'^\top \mathbf{x}_K \\ & \text{subject to} && A_{\bar{K}} \mathbf{x}_{\bar{K}} + A_K \mathbf{x}_K = \mathbf{b}/\kappa, \\ & && \mathbf{x}_K \geq \mathbf{0} \end{aligned}$$

such that $\|\tilde{\mathbf{x}}'' - \tilde{\mathbf{x}}^*\|_\infty \leq n\Delta_A \|(\mathbf{b}/\kappa + A_L \mathbf{d}) - \mathbf{b}/\kappa\|_\infty$, and thus $\|\tilde{\mathbf{x}}'' - \tilde{\mathbf{x}}^*\|_\infty \leq n\Delta_A \|A_L \mathbf{d}\|_\infty$. In addition, $\|\mathbf{x}'' - \tilde{\mathbf{x}}_K^*\|_\infty \leq \|\tilde{\mathbf{x}}'' - \tilde{\mathbf{x}}^*\|_\infty$ since \mathbf{x}'' and $\tilde{\mathbf{x}}_K^*$ are sub-vectors of, respectively, $\tilde{\mathbf{x}}''$ and $\tilde{\mathbf{x}}^*$. Note that $\tilde{\mathbf{x}}_K^*$ is an optimal solution of the scaled problem (2.6), and equal to \mathbf{x}^* . \square

Note that in Lemma 3, $\|\mathbf{x}'' - \mathbf{x}_K^*\|_\infty \leq n\Delta_A \|A_L \mathbf{d}\|_\infty < m'nA_{\max}\Delta_A$. Hence, if $i \in J$ and $\Delta \geq \Delta_A$, then x_i is an optimal basic variable of both (2.6) and (2.1). Recalling that (2.1) has a unique optimal solution yields Corollary 1.

Corollary 1. *For $i \in J$ and $\Delta \geq \Delta_A$, $x_i^* > 0$ with \mathbf{x}^* the optimal solution of (2.1).*

6. Proof of Theorem 2

As mentioned in Section 4, we need to show that TwoS((2.7); F, \bar{K}^*) is polynomial in m , n , and Δ_A which is achieved via the following result of Kitahara and Mizuno and Lemma 5.

Lemma 4 ([12], Corollary 3). *If the problem is nondegenerate, the simplex method with the most negative pivoting rule, i.e. Dantzig's rule, or the best improvement pivoting rule finds an optimal solution in at most $n \lceil m \frac{\gamma}{\delta} \log(m \frac{\gamma}{\delta}) \rceil$ iterations where m is the number of constraints, n is the number of variables, and δ and γ are, respectively, the minimum and the maximum values of all the positive elements of the primal basic feasible solutions.*

Lemma 5. *Let L be a basis of A . Then, each coordinate of $A_L^{-1}(A, I)$ is a rational number whose denominator is $\det A_L$ and the absolute value of the numerator is bounded above by Δ_A .*

Proof. For $j = 1, 2, \dots, m+n$, let \mathbf{y}_j be j -th column vector of $A_L^{-1}(A, I)$. Then, $A_L \mathbf{y}_j = \mathbf{a}_j$ where \mathbf{a}_j is j -th column vector of (A, I) . By Cramer's rule, the i -th coordinate of \mathbf{y}_j is $y_{ji} = \det A_L(i, j) / \det A_L$ with $A_L(i, j)$ being the matrix where the i -th column vector of A_L is replaced by \mathbf{a}_j . \square

In order to apply Lemma 4, the quantities γ and δ associated the auxiliary problem (2.7) are estimated in Lemma 6 and yields $\gamma/\delta \leq m^2(m'nn'A_{\max}\Delta + m')\Delta_A^3 + m\Delta_A^2$. This bound for γ/δ combined with Lemma 4 completes the proof of Theorem 2.

Lemma 6. *Each positive element of a basic feasible solution \mathbf{x}'' of (2.7) is bounded above by $m^2(m'nn'A_{\max}\Delta + m')\Delta_A^2 + m\Delta_A$ and below by $1/\Delta_A$.*

Proof. Let \mathbf{x}'' be a basic feasible solution of (2.7). Then, $\tilde{\mathbf{x}} = (\tilde{\mathbf{x}}_{\bar{K}}, \tilde{\mathbf{x}}_K)$, with $\tilde{\mathbf{x}}_K = \mathbf{x}''$ and $\tilde{\mathbf{x}}_{\bar{K}} = H_1 A_L \mathbf{f} - H_1 A_K \mathbf{x}''$, is a basic solution of:

$$\begin{aligned} \mathbf{x}_{\bar{K}} + H_1 A_K \mathbf{x}_K &= H_1 A_L \mathbf{f}, \\ H_2 A_K \mathbf{x}_K &= H_2 A_L \mathbf{f} \end{aligned}$$

where $\mathbf{f} = [(A'_L)^{-1} \mathbf{b}'/\kappa]$. Multiplying both sides of the equations by $G = H^{-1}$ from the left yields $A\mathbf{x} = A_L \mathbf{f}$. Since $\tilde{\mathbf{x}}$ is a basic feasible solution of $A\mathbf{x} = A_L \mathbf{f}$, any positive coordinate \tilde{x}_i of $\tilde{\mathbf{x}}$ is a rational number whose denominator is equal to the determinant of the basis matrix (see Lemma 5) and numerator is bounded below by 1 by the integrality of $A_L \mathbf{f}$. Hence the denominator of the coordinate of \mathbf{x}'' is bounded by Δ_A .

Similarly, $\hat{\mathbf{x}} := (\hat{\mathbf{x}}_{\bar{K}}, \hat{\mathbf{x}}_K)$, with $\hat{\mathbf{x}}_K := \mathbf{x}''$ and $\hat{\mathbf{x}}_{\bar{K}} := H_1 A_L \mathbf{d} - H_1 A_K \mathbf{x}''$, is a basic solution of:

$$\begin{aligned} \mathbf{x}_{\bar{K}} + H_1 A_K \mathbf{x}_K &= H_1 A_L \mathbf{d}, \\ H_2 A_K \mathbf{x}_K &= \mathbf{b}'/\kappa + H_2 A_L \mathbf{d}. \end{aligned}$$

Multiplying both sides of the equations by $G = H^{-1}$ from the left yields:

$$A\mathbf{x} = G \begin{pmatrix} \mathbf{0}_{\bar{K}} \\ \mathbf{b}'/\kappa \end{pmatrix} + A_L \mathbf{d}.$$

Since G and A_L are submatrices of (A, I) and $\hat{\mathbf{x}}$ is a basic solution of this system, from Lemma 5 and the integrality of A , we have

$$\|\hat{\mathbf{x}}\|_\infty \leq m\Delta_A \|\mathbf{b}'/\kappa\|_\infty + m\Delta_A \|\mathbf{d}\|_\infty.$$

Let $\mathbf{v} = A'^\top (A'A'^\top)^{-1} \mathbf{b}'$; that is, $A'\mathbf{v} = \mathbf{b}'$. Since the absolute value of any entry of A' is bounded by Δ_A by Cramer's rule, we have

$$\|\mathbf{b}'\|_\infty \leq m\Delta_A \|\mathbf{v}\|_\infty = m(m'nn'A_{\max}\Delta + m')\Delta_A \kappa$$

and thus $\|\mathbf{x}''\|_\infty \leq \|\hat{\mathbf{x}}\|_\infty \leq m^2(m'nn'A_{\max}\Delta + m')\Delta_A^2 + m\Delta_A$. \square

Acknowledgment

Research supported in part by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Science Research (A) 26242027 program and Grant-in-Aid for Research Activity Start-up 15H06617 program, by the Natural Sciences and Engineering Research Council of Canada Discovery Grant program (RGPIN-2015-06163), and by the Digiteo Chair C&O program. Part of this research was done while the authors were at the Laboratoire de Recherches en Informatique (LRI), Université Paris-Sud, Orsay, France, within the Digiteo Chair C&O program.

References

- [1] K. Anstreicher and T. Terlaky: A monotonic build-up simplex algorithm for linear programming. *Operations Research*, **42** (1994), 556–561.
- [2] F. Bilen, Z. Csizmadia, and T. Illés: Anstreicher-Terlaky type monotonic simplex algorithms for linear feasibility problems. *Optimization Methods and Software*, **22** (2007), 679–695.
- [3] N. Bonifas, M.D. Summa, F. Eisenbrand, N. Hähnle, and M. Niemeier: On sub-determinants and the diameter of polyhedra. *Discrete and Computational Geometry*, **52** (2014), 102–115.
- [4] S. Chen and R. Saigal: A primal algorithm for solving a capacitated network flow problem with additional linear constraints. *Networks*, **7** (1977), 59–79.
- [5] Z. Csizmadia: New pivot based methods in linear optimization, and an application in petroleum industry. Ph.D. Thesis, Eötvös Loránd University of Sciences, Budapest, Hungary (2007).
- [6] D. Dadush and N. Hähnle: On the shadow simplex method for curved polyhedra. *Discrete and Computational Geometry*, **56** (2016), 882–909.
- [7] G.B. Dantzig: *Linear Programming and Extensions* (Princeton University Press, Princeton, 1963).

- [8] M. Dyer and A. Frieze: Random walks, totally unimodular matrices, and a randomized dual simplex algorithm. *Mathematical Programming*, **64** (1994), 1–16.
- [9] K. Fukuda and T. Terlaky: Criss-cross methods: A fresh view on pivot algorithms. *Mathematical Programming*, **79** (1997), 369–395.
- [10] G. Kalai and D. Kleitman: A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bulletin of the American Mathematical Society*, **26** (1992), 315–316.
- [11] T. Kitahara and S. Mizuno: On the number of solutions generated by the dual simplex method. *Operations Research Letters*, **40** (2012), 172–174.
- [12] T. Kitahara and S. Mizuno: A bound for the number of different basic solutions generated by the simplex method. *Mathematical Programming*, **137** (2013), 579–586.
- [13] N. Megiddo, S. Mizuno, and T. Tsuchiya: A modified layered-step interior-point algorithm for linear programming. *Mathematical Programming*, **82** (1998), 339–355.
- [14] R.D. Monteiro and T. Tsuchiya: A variant of the Vavasis–Ye layered-step interior-point algorithm for linear programming. *SIAM Journal on Optimization*, **13** (2003), 1054–1079.
- [15] S. Mizuno: The simplex method using Tardos' basic algorithm is strongly polynomial for totally unimodular LP under nondegeneracy assumption. *Optimization Methods and Software*, **31** (2016), 1298–1304.
- [16] S. Mizuno, N. Sukegawa, and A. Deza: A primal-simplex based Tardos' algorithm. *Operations Research Letters*, **43** (2015), 625–628.
- [17] J.B. Orlin: A dual version of Tardos's algorithm for linear programming. *Operations Research Letters*, **5** (1986), 221–226.
- [18] J.B. Orlin: A polynomial time primal network simplex algorithm for minimum cost flows. *Mathematical Programming*, **78** (1997), 109–129.
- [19] K. Paparrizos: An infeasible (exterior) simplex algorithms for assignment problems. *Mathematical Programming*, **51** (1991), 45–54.
- [20] F. Santos: A counterexample to the Hirsch conjecture. *Annals of Mathematics*, **176** (2012), 383–412.
- [21] A. Schrijver: *Theory of Linear and Integer Programming* (Wiley-Interscience, New York, 1986).
- [22] N. Sukegawa: Improving bounds on the diameter of a polyhedron in high dimensions. *Discrete Mathematics*, **340** (2017), 2134–2142.
- [23] N. Sukegawa and T. Kitahara: A refinement of Todd's bound for the diameter of a polyhedron. *Operations Research Letters*, **43** (2015), 534–536.
- [24] E. Tardos: A strongly polynomial minimum cost circulation algorithm. *Combinatorica*, **5** (1985), 247–255.
- [25] E. Tardos: A strongly polynomial algorithm to solve combinatorial linear programs. *Operations Research*, **34** (1986), 250–256.
- [26] T. Terlaky: A convergent criss-cross method. *Optimization*, **16** (1985), 683–690.
- [27] M. Todd: An improved Kalai–Kleitman bound for the diameter of a polyhedron. *SIAM Journal on Discrete Mathematics*, **28** (2014), 1944–1947.
- [28] S. Vavasis and Y. Ye: A primal-dual accelerated interior point method whose running time depends only on the constraint matrix. *Mathematical Programming*, **74** (1996), 79–120.
- [29] N. Zadeh: A bad network problem for the simplex method and other minimum cost flow algorithms. *Mathematical Programming*, **5** (1973), 255–266.

Noriyoshi Sukegawa
Department of Information and System Engineering
Faculty of Science and Engineering, Chuo University
3602, 1-13-27 Kasuga
Bunkyo-ku, Tokyo 112-8551, JAPAN
E-mail: sukegawa@ise.chuo-u.ac.jp