Les Cahiers du GERAD ISSN: 0711–2440

The minimum mean cycle-canceling algorithm for linear programs

J. B. Gauthier, J. Desrosiers

G–2021–05

February 2021

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2021
– Bibliothèque et Archives Canada, 2021

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2021
– Library and Archives Canada, 2021

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

The minimum mean cycle-canceling algorithm for linear programs

Jean Bertrand Gauthier [a] [b]

Jacques Desrosiers [a] [c]

[a] *GERAD, Montréal (Québec), Canada, H3T 2A7*

[b] *Johannes Gutenberg University Mainz, 55122 Mainz, Germany*

[c] *Department of Decision Sciences, HEC Montréal, Montréal (Québec), Canada, H3T 2A7*

jean-bertrand.gauthier@hec.ca
jacques.desrosiers@hec.ca

February 2021
Les Cahiers du GERAD
G–2021–05

Abstract : This paper presents the properties of the minimum mean cycle-canceling algorithm for solving linear programming models. Originally designed by Goldberg and Tarjan (1989) for solving network flow problems for which it runs in strongly polynomial time, most of its properties are preserved. This is at the price of adapting the fundamental decomposition theorem of a network flow solution together with various definitions: that of a cycle and the way to calculate its cost, the residual problem, and the improvement factor at the end of a phase. We also use the primal and dual necessary and sufficient optimality conditions stated on the residual problem (Gauthier et al., 2014) for establishing the pricing step giving its name to the algorithm. It turns out that the successive solutions need not be basic, there are no degenerate pivots, and the improving directions are potentially interior in addition to those on edges. For solving an $m \times n$ linear program, it requires $O(n\Delta)$ so-called phases, where Δ depends on the number of rows and the coefficient matrix. Since each phase comprises at most n iterations solvable in polynomial time by an interior point algorithm, the overall complexity is pseudo-polynomial.

Keywords: Linear program, network flow problem, residual problem, cycle cancellation, complexity analysis, interior direction

Résumé : Cet article présente les propriétés de l'algorithme MMCC (*minimum mean cycle-canceling*) pour la résolution de programmes linéaires. Initialement conçu par Goldberg and Tarjan (1989) pour résoudre les problèmes de réseau pour lesquels il s'exécute en temps fortement polynomial, la plupart de ses propriétés sont préservées. Ceci au prix de l'adaptation du théorème de décomposition d'une solution ainsi de certaines définitions: celle d'un cycle et la manière de calculer son coût, le problème résiduel et le facteur d'amélioration en fin de phase. Nous utilisons également les conditions d'optimalité primales et duales énoncées sur le problème résiduel (Gauthier et al., 2014). Il s'avère que les solutions successives n'ont pas besoin d'être de base, qu'il n'y a pas de pivots dégénérés, et que les directions d'amélioration sont potentiellement intérieures en plus de celles suivant les arêtes. Pour résoudre un programme linéaire de dimension $m \times n$, il faut $O(n\Delta)$ phases, où Δ dépend du nombre de contraintes m et de la matrice de coefficients. Puisque chaque phase comprend au plus n itérations que l'on peut résoudre en temps polynomial par un algorithme de points intérieurs, la complexité globale est pseudo-polynomiale.

Mots clés : Programme linéaire, problème de flots, problème résiduel, annulation de cycle, analyse de complexité, direction intérieure

Acknowledgements: Jacques Desrosiers acknowledges the National Sciences and Engineering Research Council of Canada and the HEC Montréal Foundation for their financial support. Jean Bertrand Gauthier acknowledges the GERAD — Group for Research in Decision Analysis — for its financial support.

Introduction

In this paper, we analyze an adaptation to linear programs of the *minimum mean cycle-canceling algorithm* (MMCC) originally designed for solving capacitated network flow problems (Goldberg and Tarjan, 1989). We largely make use of the mathematical results presented in Gauthier et al. (2015) as well as the description of MMCC in Ahuja et al. (1993, Section 10.3). By adequately redefining the notions of residual problem, cycle and cycle cost, improvement factors, and exploiting the primal and dual optimality conditions expressed on the residual problem (Gauthier et al., 2014), we show that the fundamental properties of MMCC well apply to linear programs. Although almost identical to those given in earlier papers, for completeness of the presentation, we provide the adapted proofs for all the properties. Finally note that the linear programming version of the algorithm was mentioned first in Gauthier et al. (2014) but later denoted MWCC (minimum weighted cycle-canceling algorithm) in Gauthier et al. (2018).

The paper is organized as follows. We start with a description of the algorithm in Section 1. This is followed by illustrations in Section 2 on a network flow problem and a linear program: these show that improving interior-directions are possible with MMCC. Section 3 provides the complexity analysis that turns out to be pseudo-polynomial. In Section 4, we then compare MMCC to three other primal algorithms. Concluding remarks appear in Section 5.

1 Algorithm

We consider the following linear program LP with non-negative upper bounded variables:

$$\begin{aligned} z^\star = \min \quad & \mathbf{c}^\intercal \mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \quad [\boldsymbol{\pi}] \\ & \mathbf{0} \le \mathbf{x} \le \mathbf{u}, \end{aligned} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n_+$ whereas the data vectors are rational, i.e., $\mathbf{c}$, $\mathbf{u} \in \mathbb{Q}^n$, $\mathbf{b} \in \mathbb{Q}^m$, and $\mathbf{A} \in \mathbb{Q}^{m \times n}$ with $m < n$. We assume that the matrix $\mathbf{A}$ is of full row rank, LP feasible, and $z^\star$ finite. The dual vector $\boldsymbol{\pi} \in \mathbb{R}^m$ associated with the equality constraints appears within brackets. Vectors and matrices are written in bold face by respectively using lower and upper cases and we denote by $\mathbf{0}$ (resp. $\mathbf{1}$) a vector with all zeros (resp. ones) entries of contextually appropriate dimension.

1.1 Residual problem

The iterative solution process is indexed by $k \ge 0$. We define the *residual problem* $RP(\mathbf{x}^k)$ with respect to a primal feasible solution $\mathbf{x}^k$ for (1) with cost $z^k = \mathbf{c}^\intercal \mathbf{x}^k$ as follows. Every variable x_j, $j \in \{1, \ldots, n\}$, is replaced by $x_j = x_j^k + y_j - y_{j+n}$, where relatively to x_j^k:

- the forward variable y_j of cost $d_j = c_j$ represents the possible increase of x_j, with residual upper bound $r_j^k = u_j - x_j^k$;
- the backward variable y_{j+n} of cost $d_{j+n} = -c_j$ represents its possible decrease, with residual upper bound $r_{j+n}^k = x_j^k$.
- Moreover, the condition $y_j\, y_{j+n} = 0$ holds for all pairs, that is, at most one direction is used.

Let $\mathbf{k}_j$, $j \in \{1, \ldots, 2n\}$, be equal to $\mathbf{a}_j$ or $-\mathbf{a}_j$ depending on whether y_j is forward or backward, respectively. Define the set $J^k = \{j \in \{1, \ldots, 2n\} \mid r_j^k > 0\}$ which contains only the indexes of the *residual* variables, that is, the y-variables with positive residual upper bounds. A formulation for $RP(\mathbf{x}^k)$ is given by

$$\begin{aligned} z^\star = z^k + \min \quad & \sum_{j \in J^k} d_j\, y_j \\ \text{s.t.} \quad & \sum_{j \in J^k} \mathbf{k}_j\, y_j = \mathbf{0} \quad [\boldsymbol{\pi}] \\ & 0 \le y_j \le r_j^k \qquad \forall j \in J^k. \end{aligned} \tag{2}$$

Formulations (1) and (2) are equivalent in the sense that there is a one-to-one correspondence between feasible solutions that preserves the cost value.

Definition 1 *A **cycle** $\mathbf{y} = [y_j]_{j \in J^k}$ in (2) is an extreme ray of the homogeneous system $\sum_{j \in J^k} \mathbf{k}_j y_j = \mathbf{0}$, $y_j \geq 0$, $\forall j \in J^k$. We interchangeably speak of cycle W which denotes the positive variable support, i.e., $W = \{j \in J^k \mid y_j > 0\}$. Let $\mathcal{C}(\mathbf{x}^k)$ be the set of cycles in $RP(\mathbf{x}^k)$.*

Definition 2 *The **cost** of cycle W is defined as $d_W = \frac{|W|}{\sum_{j \in W} y_j} \sum_{j \in W} d_j y_j$, where $\sum_{j \in W} y_j < \infty$ normalizes the objective function with respect to the intuitive normalization $\sum_{j \in W} y_j = |W|$ established in network flows. The **mean cost** is defined with respect to the number of positive variables, i.e., $\frac{d_W}{|W|} = \frac{\sum_{j \in W} d_j y_j}{\sum_{j \in W} y_j}$. A **negative cycle** is a cycle with a negative cost (or negative mean cost).*

For any given $\boldsymbol{\pi}$, let the reduced cost of x_j, $j \in \{1, \dots, n\}$ be computed as $\bar{c}_j = c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j$. Similarly, we have $\bar{d}_j = d_j - \boldsymbol{\pi}^\intercal \mathbf{k}_j$ for y_j, $j \in \{1, \dots, 2n\}$, that is, for $j \in \{1, \dots, n\}$, $\bar{d}_j = \bar{c}_j$ and $\bar{d}_{j+n} = -\bar{c}_j$. The complementary slackness optimality conditions are well known to verify whether or not a primal-dual pair of feasible solutions is optimal for LP (1). Two equivalent necessary and sufficient optimality conditions are expressed on the residual problem.

Theorem 1 (Gauthier et al., 2014, Theorem 4) *A primal feasible solution $\mathbf{x}^k$ to LP (1) is optimal if and only if the following equivalent conditions are satisfied:*

Primal: *$RP(\mathbf{x}^k)$ contains no negative cycle, i.e., $d_W \geq 0$, $\forall W \in \mathcal{C}(\mathbf{x}^k)$.*
Dual: *$\exists \boldsymbol{\pi}$ such that $\bar{d}_j \geq 0$, $\forall j \in J^k$.*

Negative cycles are center pieces of Theorem 2 and describe the essence of a cycle-canceling algorithm. That is, we repeatedly move from one solution to the next, by increasing flow on a negative cycle until a residual upper bound is saturated, as long as optimality is not reached.

Theorem 2 (Gauthier et al., 2014, Theorem 3) *Let $\mathbf{x}^k$ and $\mathbf{x}^\ell$ be any two solutions to LP (1). Then $\mathbf{x}^\ell$ can be written as $\mathbf{x}^k$ plus the non-negative combination of at most n cycles of $RP(\mathbf{x}^k)$. Furthermore, the cost of $\mathbf{x}^\ell$ equals $\mathbf{c}^\intercal \mathbf{x}^k$ plus the cost of that combination of cycles.*

This is an existence theorem later used in the proof of Proposition 5. Note that the converse is also true, that is, $\mathbf{x}^k$ can be written as $\mathbf{x}^\ell$ plus the non-negative combination of at most n cycles of $RP(\mathbf{x}^\ell)$. Such a cycle in $RP(\mathbf{x}^\ell)$, say W_-, is the reverse of W_+ identified in $RP(\mathbf{x}^k)$ to move from $\mathbf{x}^k$ to $\mathbf{x}^\ell$, obviously with a reverse cost $c_{W_-} = -c_{W_+}$.

1.2 Pricing problem

Dantzig and Thapa (2003, Section 10.2) suggest to identify such negative cycles by imposing the arbitrary normalization constraint $\sum_{j \in J^k} y_j = 1$ to scale each extreme ray. Alternatively, the pricing problem, denoted $PP(\mathbf{x}^k)$, captures the rationale of the primal and dual optimality conditions of Theorem 1 to identify negative cycles until none remain.

Let us first consider the dual point of view: $\mathbf{x}^k$ is optimal if and only if there exists $\boldsymbol{\pi}$ such that $d_j - \boldsymbol{\pi}^\intercal \mathbf{k}_j \geq 0$, $\forall j \in J^k$. This condition can be verified by optimizing $\boldsymbol{\pi}$ so as to maximize the smallest reduced cost i.e.,

$$\mu^k = \max_{\boldsymbol{\pi}} \min_{j \in J^k} \bar{d}_j = \max_{\boldsymbol{\pi}} \min_{j \in J^k} d_j - \boldsymbol{\pi}^\intercal \mathbf{k}_j. \tag{3}$$

Optimality of $\mathbf{x}^k$ is confirmed if $\mu^k = 0$. We can linearize (3) as

$$\begin{aligned} \mu^k = \max_{\boldsymbol{\pi}} \quad & \mu \\ \text{s.t.} \quad & \mu + \boldsymbol{\pi}^\intercal \mathbf{k}_j \leq d_j \quad [y_j] \quad \forall j \in J^k, \end{aligned} \tag{4}$$

for which the dual is expressed in terms of the y-variables of $RP(\mathbf{x}^k)$:

$$\mu^k = \min \quad \sum_{j \in J^k} d_j \, y_j \tag{5a}$$

$$\text{s.t.} \quad \sum_{j \in J^k} \mathbf{k}_j \, y_j = \mathbf{0} \qquad [\boldsymbol{\pi}] \tag{5b}$$

$$\sum_{j \in J^k} y_j = 1 \qquad [\mu] \tag{5c}$$

$$y_j \geq 0 \qquad \forall j \in J^k. \tag{5d}$$

While formulation (4) of $PP(\mathbf{x}^k)$ verifies whether the dual optimality condition on $RP(\mathbf{x}^k)$ can be fulfilled, formulation (5) does it for the primal condition. Interestingly, the normalization constraint (5c) appears naturally from deriving an optimization program to verify the primal optimality condition. By Definition 2 the optimum corresponds to a minimum mean cost. Given positive weights $\mathbf{w} = [w_j]_{j \in \{1 \ldots, 2n\}}$, a straightforward generalization of the normalization follows by starting with $\frac{\bar{d}_j}{w_j}$ in (3), see Section 3.3. Finally, observe that (5) is a linear program that matches the canonical form of Karmarkar's original algorithm (Karmarkar, 1984).

1.3 Solution process

The MMCC algorithm is initialized with $\mathbf{x}^0$ of cost z^0. We improve it using negative cycles.

- At iteration $k \geq 0$, $PP(\mathbf{x}^k)$ (5) identifies a cycle W^k of minimum mean cost μ^k.
- Step size ρ^k is computed so as to saturate a residual variable (cycle-canceling).
- Improved solution $\mathbf{x}^{k+1}$ of cost $z^{k+1} < z^k$ is obtained; $RP(\mathbf{x}^{k+1})$ and $PP(\mathbf{x}^{k+1})$ are updated.
- Repeat until $\mu = 0$, the residual problem contains no negative cycle.

The n-dimensional solution $\mathbf{x}^{k+1} = [x_j^{k+1}]_{j \in \{1, \ldots, n\}}$ is obtained as

$$x_j^{k+1} = \begin{cases} x_j^k + \rho^k \, y_j^k \,, & \text{if } \; j \in W^k \\ x_j^k - \rho^k \, y_j^k \,, & \text{if } \; n + j \in W^k \\ x_j^k, & \text{otherwise,} \end{cases} \tag{6}$$

where ρ^k is computed with respect to the residual upper bounds of the variables forming cycle W^k divided by their respective contribution: $\rho^k = \min_{j \in W^k} \; r_j^k / y_j^k$.

2 Illustrations

MMCC (Goldberg and Tarjan, 1989) is a strongly polynomial algorithm for capacitated network flow problems. Applying it to the max-flow problem indeed corresponds to the algorithm of Edmonds and Karp (1972). We illustrate next that the pricing problem can identify interior-directions.

Let us start with a network problem. Figure 1 (left) shows the original network $G = (N, A)$, where N is the set of nodes and A is the set of arcs. Four units are available at the central node 0 and one unit is requested at all four other nodes. The cost and flow bounds appear next to the arcs. We assume $c < 0$, hence the optimal solution sends one unit of flow on the *radial arcs* $(0,1), (0,2), (0,3), (0,4)$ and 100 units on the counterclockwise *ring arcs* $(1,2), (2,3), (3,4), (4,1)$. Therefore $z^\star = 400c$ at $\mathbf{x}^\star = (1,1,1,1,100,100,100,100)^\intercal$.

Initial solution $\mathbf{x}^0 = (1,1,1,1,0,0,0,0)^\intercal$ of cost $z^0 = 0$ sends one unit on radial arcs and nothing on ring arcs; $\mathbf{x}^0$ is basic, where the basic variables are $x_{0j}^0 = 1$, $j = 1, \ldots, 4$, strictly within the interval $[0,2]$ on G. Figure 1 (right) depicts the residual network $G(\mathbf{x}^0)$, a representation of $RP(\mathbf{x}^0)$. For $j = 1, \ldots, 4$, y_{0j} and y_{j0} have a cost of zero while their residual upper bounds are equal to 1; for the ring variables y_{12}, y_{23}, y_{34}, and y_{41}, the cost is c and the residual upper bounds are equal to 100. A convex combination of variables in (5c) is a *directed cycle* in $G(\mathbf{x}^0)$. Let W denote such a cycle with a flow of $1/|W|$ on every arc. Finding μ^0 reduces to evaluating the minimum *mean cost* of cycles.

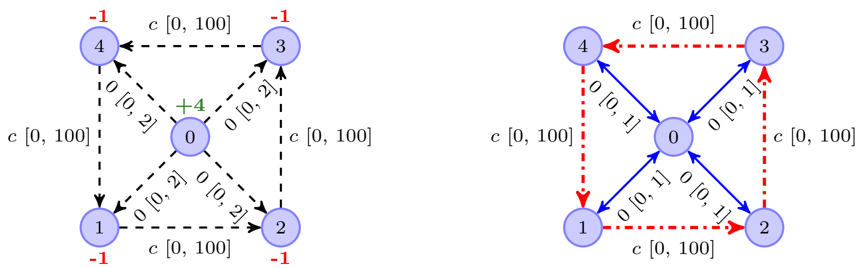

Figure 1: Original network G (left) and residual network $G(\mathbf{x}^0)$ (right).

$G(\mathbf{x}^0)$ comprises 13 negative cycles:

- for $j = 1, \ldots, 4$: 3-node cycles denoted C_3^j of mean cost $c/3$ represented by the node sequences (0120), (0230), (0340) and (0410);
- for $j = 1, \ldots, 4$: 4-node cycles C_4^j of mean cost $2c/4$ given by (01230), (02340), (03410) and (04120);
- for $j = 1, \ldots, 4$: 5-node cycles C_5^j of mean cost $3c/5$ given by (012340), (023410), (034120) and (041230);
- *outer* cycle $C_o = (12341)$ of mean cost $4c/4 = c$, indeed the minimum value for μ^0.

The pricing problem selects the 4-arc cycle C_o and the step size must satisfy $\rho(1/4) \leq 100$ on the ring arcs, hence $\rho = 400$, $z^1 = z^0 + 400\,c$. MMCC reaches the optimal vertex $\mathbf{x}^\star$ in a single iteration. At this point, $G(\mathbf{x}^\star)$ contains no negative cost cycles, hence satisfies the primal optimality condition of Theorem 1. Cycle C_o is in fact the combination of the four 3-node cycles $C_3^j, j = 1, \ldots, 4$, each one being an edge-direction. As such, C_o induces an interior-direction.

We continue with the 3D-linear maximization program (7) comprising four inequality constraints (Gauthier et al., 2018):

$$\begin{array}{rrrrcl}
\max\limits_{\mathbf{x}\geq\mathbf{0}} & 130x_1 & +80x_2 & +60x_3 & & \\
\text{s.t.} & 2x_1 & -\,x_2 & +\,2x_3 & \leq & 21 \\
 & -\,x_1 & +\,x_2 & -\,x_3 & \leq & 8 \\
 & 2x_1 & -\,x_2 & -\,x_3 & \leq & 15 \\
 & -\,x_1 & -\,x_2 & +\,2x_3 & \leq & 32
\end{array} \tag{7}$$

The initial zero-cost basic solution $\mathbf{x}^0 \in \mathbb{R}_+^7$ uses the slack variables at values $x_4^0 = 21$, $x_5^0 = 8$, $x_6^0 = 15$ and $x_7^0 = 32$. For this non-degenerate solution, there are three edge-directions according to the possible entering variables x_1, x_2, and x_3. While the primal simplex algorithm selects x_1 with a maximum reduced cost of 130, MMCC with $\mu^0 = 140/3$ rather combines the three edge-directions to move with the *interior-direction* $(1/6, 1/4, 1/12)$. As a consequence, the successive primal solutions are not necessarily (*and need not be*) basic in MMCC.

3 Complexity analysis

As for network flow problems, the dual values are used to relocate the cost information along the cycles: finding a minimum mean cost cycle is equivalent to finding a minimum *mean reduced cost* cycle.

Proposition 1 *For any $\boldsymbol{\pi}$ and cycle W in (5), $\sum_{j\in W} d_j y_j = \sum_{j\in W} \bar{d}_j y_j$.*

Proof. We have $\sum_{j\in W} \mathbf{k}_j y_j = 0$ by (5b) for any cycle W. Therefore

$$\sum_{j\in W} \bar{d}_j y_j = \sum_{j\in W} (d_j - \boldsymbol{\pi}^\intercal \mathbf{k}_j) y_j = \sum_{j\in W} d_j\, y_j - \boldsymbol{\pi}^\intercal \sum_{j\in W} \mathbf{k}_j y_j = \sum_{j\in W} d_j y_j.$$

□

3.1 Optimality parameter

The mechanics of MMCC do not require the computation of any reduced costs. The algorithm simply relies on the *primal* condition of Theorem 1. However, the complexity analysis exploits the dual point of view of that condition. The idea is to study the convergence towards zero of the *optimality parameter* μ, the current most negative reduced cost. In the following, the superscript k in $\bar{d}_j^k$ is understood to signify that the computation of the reduced cost is done with $\boldsymbol{\pi}^k$, $k \ge 0$.

Proposition 2 *Given a non-optimal solution* $\mathbf{x}^k, k \ge 0$, *solutions to* (4)–(5) *are such that*

$$\text{(a) } \mu^k = \min_{j\in J^k} \bar{d}_j^k; \qquad \text{(b) } \bar{d}_j^k = \mu^k,\ \forall j \in W^k; \qquad \text{(c) } \mu^{k+1} \ge \mu^k.$$

Proof.

(a) At $k \ge 0$, the constraints in (4) can be written as $\mu \le d_j - \boldsymbol{\pi}^\intercal \mathbf{k}_j$, $\forall j \in J^k$. At optimality, $\mu^k \le \bar{d}_j^k$, $\forall j \in J^k$, for some optimized vector $\boldsymbol{\pi}^k$ and the maximization in the objective function pushes μ^k to the smallest reduced cost.

(b) The complementary slackness conditions guarantee that the equality holds in (4) for all $y_j^k > 0$, that is, $\mu^k = \bar{d}_j^k$, $\forall j \in W^k$.

(c) In $RP(\mathbf{x}^{k+1})$, the saturated variables in W^k are removed and new variables appear in the reverse direction with a reduced cost equal to $-\mu^k > 0$. Therefore, by construction, every variable of $RP(\mathbf{x}^{k+1})$ has a reduced cost $\bar{d}_j^k \ge \mu^k$ computed with respect to $\boldsymbol{\pi}^k$. Since the reduced cost of a cycle is at least as great as the smallest reduced cost of any of its terms, $\mu^{k+1} \ge \mu^k$.

□

Proposition 2 alone is not sufficient to ensure convergence, we also need μ to strictly increase.

Definition 3 *Given* $\boldsymbol{\pi}^k$, W^k *is a* ***cycle of Type 1*** *if it contains only* y*-variables of negative reduced costs, i.e.,* $\bar{d}_j^k < 0$, $\forall j \in W^k$. *Otherwise, it is a* ***cycle of Type 2*** *and it contains at least one* y*-variable with a non-negative reduced cost, i.e.,* $\exists s \in W^k$ *such that* $\bar{d}_s^k \ge 0$.

On network flow problems, the flow is the same on all the arcs of a cycle W^k, i.e., $y_j^k = 1/|W^k|$, $\forall j \in W^k$. This is not the case in general for linear programs. For any $k \ge 0$, let us compute the ratio $R^k > 0$ of the smallest to the largest y-values in W^k and denote by R the smallest data-dependent ratio that can occur in any cycle:

$$R^k = \frac{\min_{j\in W^k} y_j^k}{\max_{j\in W^k} y_j^k}, \qquad R = \min_{\mathbf{x}, W\in\mathcal{C}(\mathbf{x})} \frac{\min_{j\in W} y_j}{\max_{j\in W} y_j}. \tag{8}$$

Then, because $\max_{j\in W^k} y_j^k \ge 1/|W^k|$,

$$\forall j \in W^k: \quad y_j^k \ge \min_{j\in W^k} y_j^k \ge \frac{R^k}{|W^k|} \ge \frac{R}{m+1} \ge \frac{1}{\Delta}, \quad \text{where} \quad \Delta = \left\lceil \frac{m+1}{R} \right\rceil. \tag{9}$$

We show next that μ strictly increases within n iterations, otherwise MMCC finds an optimal solution.

Proposition 3 *Let $\mathbf{x}^k$ be a non-optimal solution, $k \geq 0$. After at most n iterations, MMCC finds an optimal solution to LP (1), otherwise $\mu^{k+\ell} > \mu^k$, for some $1 \leq \ell \leq n$.*

Proof. By Proposition 1, $\sum_{j\in J^k} d_j\, y_j$ in (5a) can be replaced by $\sum_{j\in J^k} \bar{d}_j\, y_j$, that is, any reduced cost vector can be used instead of the cost vector. For every iteration $k+\ell$, where $1 \leq \ell \leq n$, let us find $\mu^{k+\ell}$ using $\bar{d}_j^k$, $\forall j \in J^{k+\ell}$, that is, the reduced costs are computed using $\boldsymbol{\pi}^k$. There are n variables in *LP* (1), hence in any pricing problem, there are at most n variables with negative reduced costs, the others being reverse y-variables with opposite signs for the reduced costs.

With respect to $\boldsymbol{\pi}^k$, an optimal cycle $W^{k+\ell}$ is of Type 1 or Type 2. Each time a Type 1 cycle is found, at least one variable reaches its residual upper bound. Such a variable y_j with a negative reduced cost $\bar{d}_j^k$ is eliminated from the next residual problem and replaced by the reverse variable for which $-\bar{d}_j^k$ is positive. Consequently, if within n iterations, the algorithm finds no Type 2 cycle, then all the variables in the pricing problem have non-negative reduced costs, the residual problem contains no negative cycle, and the algorithm terminates with an optimal solution by Theorem 1.

Now consider that at $\ell \leq n$, the pricing problem finds $W^{k+\ell}$, a Type 2 cycle. By Definition 3, there must exist some index $s \in W^{k+\ell}$ such that $\bar{d}_s^k \geq 0$, where $y_s^{k+\ell} \geq \frac{1}{\Delta}$. At this point, all variables that have not been reversed during the iterative process still have their reduced costs greater than or equal to μ^k by Proposition 2(a) while the added variables have positive reduced costs. Therefore

$$\begin{aligned}\mu^{k+\ell} &= \bar{d}_s^k\, y_s^{k+\ell} + \textstyle\sum_{j\in W^{k+\ell}\setminus\{s\}} \bar{d}_j^k\, y_j^{k+\ell} \;\geq \textstyle\sum_{j\in W^{k+\ell}\setminus\{s\}} \bar{d}_j^k\, y_j^{k+\ell} \\ &\geq \mu^k \textstyle\sum_{j\in W^{k+\ell}\setminus\{s\}} y_j^{k+\ell} \;= \mu^k(1-y_s^{k+\ell}) \geq \mu^k(1-\frac{1}{\Delta}) > \mu^k.\end{aligned}$$

□

A Type 2 cycle at the end of iteration $k+\ell$, $1 \leq \ell \leq n$, is a marker for the *improvement factor* on μ^k expressed by $\left(1-\frac{1}{\Delta}\right)$. This leads to the following definitions.

Definition 4 *A **phase** is a sequence of iterations terminated by a Type 2 cycle. A **phase solution** $\mathbf{x}^h, h \geq 0$, is the solution at the beginning of phase h.*

While the two phase numbers h and $h+1$ are consecutive, the number of cycles canceled within phase h is at most n. For the remainder of this paper, the notations k and h are reserved for iteration and phase numbers, respectively.

Proposition 4 *Let $\mathbf{x}^h$, $h \geq 0$, be a non-optimal phase solution. Then $\mu^{h+\Delta} > \frac{\mu^h}{2}$.*

Proof. A Type 2 cycle implies that $\mu^{h+1} \geq \mu^h\left(1-\frac{1}{\Delta}\right)$. Likewise, $\mu^{h+\Delta} \geq \mu^h\left(1-\frac{1}{\Delta}\right)^\Delta > \frac{\mu^h}{2}$ as $\left(1-\frac{1}{\Delta}\right)^\Delta$ converges to $e^{-1} \simeq 0.368$ from below. □

3.2 Pseudo-polynomial number of phases

In this section, we present a theorem on the overall complexity of MMCC for linear programs.

Proposition 5 *Let $\mathbf{x}^k$, $k \geq 0$, be a non-optimal solution. If $\bar{d}_f^k \geq -2\Delta\mu^k$, for $f \in \{1,\ldots,2n\}$, then y_f is implicitly set to zero in all subsequent pricing problems, that is, $y_f^{k+\ell} = 0$, $\forall \ell \geq 1$.*

Proof. For $f \in \{1,\ldots,n\}$, let $\bar{d}_f^k \geq -2\Delta\mu^k > 0$. The forward variable y_f exists in $RP(\mathbf{x}^k)$ because y_{n+f} cannot as it would have a negative reduced cost $\bar{d}_{f+n}^k = -\bar{d}_f^k \leq 2\Delta\mu^k < \mu^k$, a contradiction with Proposition 2(a). Hence $x_f^k = 0$. We show next that we then implicitly have $x_f^\star = 0$, i.e., $x_f^{k+\ell}$ remains at zero for all $\ell \geq 1$. For $f \in \{n+1,\ldots,2n\}$ and $\bar{d}_f^k \geq -2\Delta\mu^k > 0$, similar arguments yield $x_f^\star = u_f$.

Assume on the contrary that $x_f^{k+\ell} > 0$ for some $\ell \geq 1$. Using Proposition 2(c), we have $\mu^k \leq \mu^{k+\ell}$. By Theorem 2, $\mathbf{x}^{k+\ell}$ can be written as $\mathbf{x}^k$ plus the non-negative combination of at most n cycles of $RP(\mathbf{x}^k)$. One of these, denoted W_+^s at iteration $s \in [k+1, k+\ell]$, is such that the forward variable y_f^s is positive, hence $y_f^s \geq 1/\Delta$ by (9). Given $\boldsymbol{\pi}^k$ and $\mu^k < 0$ for the reference iteration k, the mean reduced cost of cycle W_+^s is

$$\begin{aligned} \frac{\bar{d}_{W_+^s}}{|W_+^s|} &= \sum_{j \in W_+^s \setminus \{f\}} \bar{d}_j^k \, y_j^s \; + \bar{d}_f^k \, y_f^s \quad \geq \quad \mu^k(1 - y_f^s) - (2\Delta\mu^k)\, y_f^s \\ &\geq \mu^k(1 - \frac{1}{\Delta}) \quad - (2\Delta\mu^k)\frac{1}{\Delta} \quad = \quad -\mu^k \; - \; \mu^k/\Delta \; > \; 0. \end{aligned} \tag{10}$$

Now consider the reverse cycle W_-^s that appears in $RP(\mathbf{x}^{k+\ell})$ (Theorem 2). It has a negative mean reduced cost $\frac{\bar{d}_{W_-^s}}{|W_+^s|} = -\frac{\bar{d}_{W_+^s}}{|W_+^s|}$ at most equal to $\mu^k + \mu^k/\Delta < \mu^k \leq \mu^{k+\ell}$, a contradiction on the optimality of $\mu^{k+\ell}$ by Proposition 2(a). Hence $x_f^{k+\ell} = 0$ for all $\ell \geq 1$, i.e., $x_f^\star = 0$. □

Proposition 6 *MMCC solves LP (1) in $O(n\Delta \log \Delta)$ phases.*

Proof. Consider a block of $L = \Delta(\lceil \log \Delta \rceil + 1)$ phases and phase solutions $\mathbf{x}^h$ and $\mathbf{x}^{h+L}$. By Proposition 4, μ^h increases by a factor of at least $1/2$ every Δ phases so that the increase in a block is

$$\mu^{h+L} \geq \left(\frac{1}{2^{\lceil \log \Delta \rceil + 1}} \right) \mu^h \geq \frac{1}{2\Delta}\, \mu^h, \qquad \text{or equivalently,} \quad \mu^h \leq 2\Delta\, \mu^{h+L}.$$

By Proposition 1, the mean reduced cost μ^h for W^h is independent of the dual vector used, that is,

$$\mu^h = \sum_{j \in W^h} d_j \, y_j^h = \sum_{j \in W^h} \bar{d}_j^{h+L} \, y_j^h.$$

The values $\{\bar{d}_j^{h+L}\}_{j \in W^h}$ cannot *all* be strictly greater than μ^h, hence, there exists $f \in W^h$ such that $\bar{d}_f^{h+L} \leq \mu^h \leq 2\Delta\, \mu^{h+L}$. The step size ρ^h is computed so as to cancel W^h which makes the reverse cycle appear in the next residual problem: a y-variable with a reduced cost equal to $-\bar{d}_f^{h+L} \geq -2\Delta\, \mu^{h+L}$ appears, either for y_{f+n} if $f \in \{1, \ldots, n\}$ or y_{f-n} if $f \in \{n+1, \ldots, 2n\}$. In any case, this y-variable is fixed to zero by Proposition 5, that is, the corresponding x-variable is fixed either at zero or its upper bound. All in all, each block fixes a different variable and since there are n variables, $O(n\Delta \log \Delta)$ phases are sufficient to solve LP (1). □

Following the lines of Radzik and Goldberg (1994) or perhaps more easily that of Gauthier et al. (2015), we can reduce the total number of phases to $O(n\Delta)$. Since this is a tight complexity bound based on implicitly fixing values of variables from optimal multipliers, the complexity of solving at most n pricing problems in a phase becomes a moot point although we could rest on the polynomial complexity of interior-point algorithms.

Theorem 3 *MMCC solves LP (1) in pseudo-polynomial time.*

For the capacitated minimum cost network flow problem with $|N|$ nodes and $|A|$ arcs, a cycle W comprises no more than $|N|$ nodes. The positive y-variables in (5) are all equal to $1/|W|$, the ratio $R = 1$, $\Delta = |N|$ in the worst-case (one constraint being redundant), and the improving factor at the end of a phase is at least $(1 - 1/|N|)$. As expected, this simplifies to the tight bound $O(|A||N|)$ phases and known complexity result.

Theorem 4 (Radzik and Goldberg, 1994) *Given arbitrary real-valued arc costs for a network flow problem defined on $G = (N, A)$, MMCC is strongly polynomial as it performs $O(|A||N|)$ phases, each one comprising at most $|A|$ cycle cancellations, each pricing problem being solved by dynamic programming (Karp, 1978) in $O(|A||N|)$. Therefore, MMCC runs in $O(|A|^3|N|^2)$ time.*

3.3 Normalization constraint

For the ease of the presentation, the sum of the variables $\sum_{j\in J^k} y_j = 1$ appears in (5). However, it can be replaced by the more general weighted version

$$\sum_{j\in J^k} w_j y_j = 1, \qquad y_j \geq 0, w_j > 0, \ \forall j \in \{1,\ldots,2n\}. \tag{11}$$

The weight w_j can be one of the norms of $\mathbf{k}_j$, i.e., $w_j = \|\mathbf{k}_j\|$, see for example Rosat et al. (2017). The correspondence between the non-zero extreme points and extreme rays remains the same, although the order in which the cycles are canceled differs. To derive (11), the dual condition of Theorem 1 is verified by optimizing $\boldsymbol{\pi}$ so as to maximize the smallest *weighted reduced cost*, i.e., $PP(\mathbf{x}^k)$ (3) is rather formulated as

$$\mu^k = \max_{\boldsymbol{\pi}} \ \min_{j\in J^k} \ \frac{\bar{d}_j}{w_j} = \max_{\boldsymbol{\pi}} \ \min_{j\in J^k} \ \frac{d_j - \boldsymbol{\pi}^\intercal \mathbf{k}_j}{w_j}, \tag{12}$$

and the optimality of $\mathbf{x}^k$ is again confirmed if $\mu^k = 0$. Two equivalent primal-dual pairs of linear programs are presented below, depending on the position of the weights:

$$\begin{array}{ll|ll}
\mu^k = \max\limits_{\boldsymbol{\pi}} & \mu & \min & \sum_{j\in J^k} d_j y_j \\
\text{s.t.} & w_j\mu + \boldsymbol{\pi}^\intercal \mathbf{k}_j \leq d_j \quad [y_j] \ \ \forall j \in J^k & \text{s.t.} & \sum_{j\in J^k} \mathbf{k}_j y_j = \mathbf{0} \quad [\boldsymbol{\pi}] \\
 & & & \sum_{j\in J^k} w_j y_j = 1 \quad [\mu] \\
 & & & y_j \geq 0 \quad \forall j \in J^k,
\end{array} \tag{13}$$

$$\begin{array}{ll|ll}
\mu^k = \max\limits_{\boldsymbol{\pi}} & \mu & \min & \sum_{j\in J^k} \left(\frac{d_j}{w_j}\right) y_j \\
\text{s.t.} & \mu + \frac{\boldsymbol{\pi}^\intercal \mathbf{k}_j}{w_j} \leq \frac{d_j}{w_j} \quad [y_j] \ \ \forall j \in J^k & \text{s.t.} & \sum_{j\in J^k} \left(\frac{\mathbf{k}_j}{w_j}\right) y_j = \mathbf{0} \quad [\boldsymbol{\pi}] \\
 & & & \sum_{j\in J^k} y_j = 1 \quad [\mu] \\
 & & & y_j \geq 0 \quad \forall j \in J^k.
\end{array} \tag{14}$$

Using the first pair (13), the pseudo-polynomial number of phases remains valid, it suffices to redefine the ratios R^k and R in (8) as

$$R^k = \frac{\min\limits_{j\in W^k} \ w_j y_j^k}{\max\limits_{j\in W^k} \ w_j y_j^k}, \qquad R = \min_{\mathbf{x}, W\in\mathcal{C}(\mathbf{x})} \ \frac{\min\limits_{j\in W} \ w_j y_j}{\max\limits_{j\in W} \ w_j y_j}. \tag{15}$$

4 Comparisons and implementation issues

Let us complete the analysis by comparing MMCC with three other algorithms, namely the classical primal simplex PS (Dantzig, 1947), the improved primal simplex IPS (El Hallaoui et al., 2011), and the linear fractional approximation LFA (Bouarab et al., 2017). The comparisons presented in Table 1 are based on the formulation of the pricing problems from a dual point of view in (16).

At iteration $k \geq 0$, let us assume that $\mathbf{x}^k$ is a basic solution. First recall that a degenerate solution comprises basic variables at their bounds. Second, none of the three equivalent optimality conditions of Theorem 1 is related to a basic representation of a solution. Indeed, there is no need to associate a basis with an extreme point solution. Third, for each algorithm, the dual constraints are divided in two subsets to highlight their similarities and differences. Let us recall and define various subsets of variables.

Table 1: Comparisons between MMCC (1989), PS (1947), IPS (2011), and LFA (2017).

Algorithm	Cycle	Direction	Optimality parameter	Objective	Sum
MMCC	feasible	edge, face, interior	$\mu^k \le \mu^{k+1}$, $\mu^k < \mu^{k+n}$	$z^k > z^{k+1}$	oscillation
PS	not always feasible	edge	oscillation	$z^k \ge z^{k+1}$	oscillation
IPS	feasible	edge	oscillation	$z^k > z^{k+1}$	oscillation
LFA	combination of cycles	new $\mathbf{x}$-solution	$\mu^k < \mu^{k+1}$	$z^k > z^{k+1}$	$s^k < s^{k+1}$

- $J = \{1, \ldots, n\}$ for the original variables;
- $B^k = \{j \in J \mid x_j^k \text{ is basic}\}$ and $N^k = J \setminus B^k$ for the non-basic variables;
- $L^k = \{j \in J \mid x_j^k = 0\}$, $U^k = \{j \in J \mid x_j^k = u_j\}$ and $F^k = \{j \in J \mid 0 < x_j^k < u_j\}$;
- $J^k = \{j \in \{1, \ldots, 2n\} \mid r_j^k > 0\}$ for the residual variables.

We also refer to Figure 2, where we recognize the basic variables in *free* ones $F^k \subseteq B^k$ and *degenerate* ones in $L^k \cup U^k$. Note that J^k comprises the original variables (J) and those in reverse within F^k such that none of the reverse degenerate variables are present. The θ-variable arises only within LFA.

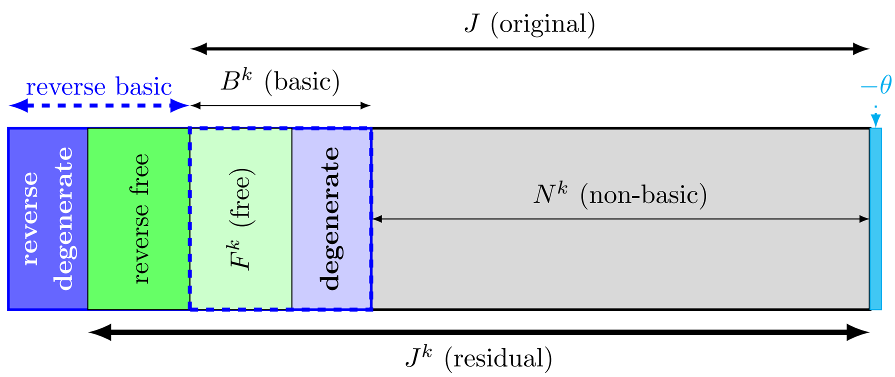

Figure 2: Primal variables at iteration $k \ge 0$.

The formulation of the dual pricing problems for MMCC, PS, IPS, and LFA appear below. The first is a rewriting of (4) with $\bar{d}_j = \bar{c}_j, \forall j \in J$, and $\bar{d}_{n+j} = -\bar{c}_j, \forall j \in F^k$.

$$
\begin{array}{rllll}
\mu^k = \max\limits_{\boldsymbol{\pi}} & \mu \quad \text{s.t.} & & & \\
\textbf{MMCC} & \mu \le \quad c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j & [y_j \ge 0] & \forall j \in J & \text{(original)} \qquad (16a) \\
& \mu \le -(c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j) & [y_j \ge 0] & \forall j \in F^k & \text{(reverse free)} \qquad (16b) \\
\textbf{PS} & 0 = c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j & [y_j \in \mathbb{R}] & \forall j \in B^k & \text{(basic and reverse basic)} \qquad (16c) \\
& \mu \le c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j & [y_j \ge 0] & \forall j \in N^k & \text{(non-basic)} \qquad (16d) \\
\textbf{IPS} & 0 = c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j & [y_j \in \mathbb{R}] & \forall j \in F^k & \text{(free and reverse free)} \qquad (16e) \\
& \mu \le c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j & [y_j \ge 0] & \forall j \in L^k & \text{(lower bound)} \qquad (16f) \\
\textbf{LFA} & \mu \le c_j - \boldsymbol{\pi}^\intercal \mathbf{a}_j & [y_j \ge 0] & \forall j \in J & \text{(original)} \qquad (16g) \\
& \boldsymbol{\pi}^\intercal \mathbf{b} = z^k & [\theta > 0] & & \text{(dual obj.} = \text{primal obj.)} \qquad (16h)
\end{array}
$$

The inequalities for MMCC imply that the y-variables in the primal formulation are greater-than-or-equal to zero, as in (5). This is not the case for PS. Imposing a zero reduced cost for the *basic* variables implies that $y_j \in \mathbb{R}$, $\forall j \in B^k$, or equivalently, $y_j, y_{n+j} \ge 0$, $\forall j \in B^k$. Unfortunately, this is only allowed for the free variables $F^k \subseteq B^k$, not for the degenerate ones: a variable at its lower bound cannot decrease while one at its upper bound cannot increase. Finding an optimal cycle which wrongly uses a basic variable with a residual bound at zero results in a degenerate pivot.

Regarding IPS, a zero reduced cost is imposed for the *free* variables (F^k), in accordance with the complementary slackness conditions of Theorem 1. This remains valid for any subset $\tilde{F}^k \subseteq F^k$, MMCC being at one end ($\tilde{F}^k = \emptyset$), IPS at the other ($\tilde{F}^k = F^k$), see Gauthier et al. (2018). IPS always identifies improving *edge*-directions but has an exponential time complexity on the Klee?Minty polytope (MMCC takes one iteration).

Finally, LFA imposes the same constraints as MMCC on the original variables (J) but restricts $\boldsymbol{\pi}$ to satisfy $\boldsymbol{\pi}^\intercal \mathbf{b} = z^k$, a condition that holds at optimality (and is satisfied in PS with $\boldsymbol{\pi}^\intercal = \mathbf{c}_\mathbf{B}^\intercal \mathbf{B}^{-1}$, where $\mathbf{B}$ is the current basis). LFA requires that the sum s^k of the x-variables be positive, $\forall k \geq 0$. The main properties come from the rewriting of the primal pricing problem as a linear fractional program using the transformation of Charnes and Cooper (1962), where $\theta^k = 1/s^k$: at every iteration until optimality is reached, z^{k+1} decreases while both μ^{k+1} and s^{k+1} increase. It also has a super-geometric growth rate on μ given by $\mu^{k+1} = \mu^0 \left(1 - \sqrt[k+1]{s^0/s_{max}}\right)^{k+1}$, where s_{max} is an upper bound on the sum of the x-variables.

Polynomial complexity is not a warrant of efficiency. A landmark in linear programming history that underscores this is the first polynomial time ellipsoid algorithm of Khachiyan (1979). Although PS may have degenerate pivots and not converge, the number of iterations is approximately $3m$ in practice. The adaptation of IPS to solve by column generation the linear relaxation of set partitioning problems, the so-called *dynamic constraint aggregation* algorithm, also turns out to be very efficient for various crew scheduling type problems (El Hallaoui et al., 2005). Little is known on LFA except that it should be used for problems whose objective value is not correlated to the sum of the variables, otherwise the approximation is almost the original problem.

Despite the fact that MMCC runs in strongly polynomial time on network flow problems, it is not at all efficient in a naive implementation. It sometimes takes more time to solve a single pricing step than directly solving the problem at hand (Gauthier et al., 2017). It is therefore not surprising that the same is observed for the solution of linear programs. However, as all other algorithms, MMCC requires adequate heuristic strategies, notably a partial pricing scheme. By Theorem 2, all the negative cycles to move from $\mathbf{x}^h$ to $\mathbf{x}^\star$ are in the residual problem $RP(\mathbf{x}^h)$. We can therefore identify several cycles, combine them in a restricted master problem (by Dantzig-Wolfe decomposition) to determine the best (likely interior) direction and step size, modify the dual variables, and *again using the same residual problem*, identify additional negative cycles. This is very similar to the *Cancel-and-Tighten* strategy (CT) proposed by Goldberg and Tarjan (1989) for network problems. As CT exploits both the primal and dual properties of the cycle-canceling algorithm, the heuristic selection as well as the order in which the cycles of Type I are canceled become irrelevant. Gauthier et al. (2015) show that the complexity of the bottleneck operation, i.e., solving a phase, is reduced from $O(|A|^2|N|)$ to only $O(|A| \log |N|)$, hence MMCC runs in $O(|A|^2|N| \log |N|)$ time. A similar strategy has been used with IPS in Metrane et al. (2010) while identifying and combining edge directions. Our hope is that this can be successfully adapted for solving large-scale linear programs with MMCC, specially in the context of the column generation algorithm where, too often, very time consuming subproblems select variables that lead to degenerate pivots. In fact, any heuristic solution to the pricing problem (5) results in a strict decrease of the objective function.

5 Conclusion

This paper presents the properties on linear programs of the minimum mean cycle-canceling algorithm originally designed for solving capacitated minimum cost network flow problems. The adaptation needs a similar decomposition theorem of a solution together with various definitions: that of a cycle and the way to calculate its cost, the residual problem, and the improvement factor at the end of a phase. From the primal point of view, the objective function decreases at every iteration, interior-directions are possible in addition to edge ones, and the successive solutions need not be basic. Every pricing problem can be solved in polynomial running time by interior point algorithms. From the dual point of view, the smallest reduced cost is increasing from one phase to the next and converges to zero in a pseudo-polynomial number of these. The overall complexity is hence a pseudo-polynomial running time. A naive implementation is not efficient, neither for network flow problems nor linear programs. Further research should concentrate on accelerating and heuristic strategies, especially in the context of column generation in order to avoid time-consuming degenerate pivots.

References

Ravindra Kumar Ahuja, Thomas Lee Magnanti, and James Berger Orlin. Network Flows: Theory, Algorithms, and Applications. Prentice Hall, Upper Saddle River, NJ, USA, 1993.

Hocine Bouarab, Guy Desaulniers, Jacques Desrosiers, and Jean Bertrand Gauthier. Linear fractional approximations for master problems in column generation. Operations Research Letters, 45(5):503–507, 2017. doi: 10.1016/j.orl.2017.08.004.

Abraham Charnes and William Wager Cooper. Programming with linear fractional functionals. Naval Research Logistics Quarterly, 9(3–4):181–186, 1962. doi: 10.1002/nav.3800090303.

George Bernard Dantzig and Mukund Narain Thapa. Linear Programming 2: Theory and Extensions. Springer Series in Operations Research and Financial Engineering (Book 2). Springer, New York, NY, USA, 2003. doi: 10.1007/b97283.

Jack Edmonds and Richard Manning Karp. Theoretical improvements in algorithmic efficiency for network flow problems. Journal of the ACM, 19(2):248–264, 1972. doi: 10.1145/321694.321699.

Issmail El Hallaoui, Daniel Villeneuve, François Soumis, and Guy Desaulniers. Dynamic aggregation of set-partitioning constraints in column generation. Operations Research, 53(4):632–645, 2005. doi: 10.1287/opre.1050.0222.

Issmail El Hallaoui, Abdelmoutalib Metrane, Guy Desaulniers, and François Soumis. An improved primal simplex algorithm for degenerate linear programs. INFORMS Journal on Computing, 23(4):569–577, 2011. doi: 10.1287/ijoc.1100.0425.

Jean Bertrand Gauthier, Jacques Desrosiers, and Marco E. Lübbecke. Decomposition theorems for linear programs. Operations Research Letters, 42(8):553–557, 2014. doi: 10.1016/j.orl.2014.10.001.

Jean Bertrand Gauthier, Jacques Desrosiers, and Marco E. Lübbecke. About the minimum mean cycle-canceling algorithm. Discrete Applied Mathematics, 196:115–134, 2015. doi: 10.1016/j.dam.2014.07.005.

Jean Bertrand Gauthier, Jacques Desrosiers, and Marco E. Lübbecke. A strongly polynomial contraction-expansion algorithm for network flow problems. Computers & Operations Research, 84:16–32, August 2017. doi: 10.1016/j.cor.2017.02.019.

Jean Bertrand Gauthier, Jacques Desrosiers, and Marco E. Lübbecke. Vector space decomposition for solving large-scale linear programs. Operations Research, 66(5):1376–1389, 2018. doi: 10.1287/opre.2018.1728.

Andrew Vladislav Goldberg and Robert Endre Tarjan. Finding minimum-cost circulations by canceling negative cycles. Journal of the ACM, 36(4):873–886, 1989. doi: 10.1145/76359.76368.

Narendra Karmarkar. A new polynomial-time algorithm for linear programming. Combinatorica, 4(4):373–395, 1984. doi: 10.1007/BF02579150.

Richard Manning Karp. A characterization of the minimum cycle mean in a digraph. Discrete Mathematics, 23(3):309–311, 1978. doi: 10.1016/0012-365X(78)90011-0.

Leonid Genrikhovich Khachiyan. A polynomial algorithm in linear programming. Doklady Akademii Nauk SSSR, 244(5):1093–1096, 1979. (english translation in Soviet Mathematics Doklady, 20:191–194, 1979).

Abdelmoutalib Metrane, François Soumis, and Issmail El Hallaoui. Column generation decomposition with the degenerate constraints in the subproblem. European Journal of Operational Research, 207(1):37–44, 2010. doi: 10.1016/j.ejor.2010.05.002.

Tomasz Radzik and Andrew Vladislav Goldberg. Tight bounds on the number of minimum-mean cycle cancellations and related results. Algorithmica, 11(3):226–242, 1994. doi: 10.1007/BF01240734.

Samuel Rosat, Issmail El Hallaoui, François Soumis, and Driss Chakour. Influence of the normalization constraint on the integral simplex using decomposition. Discrete Applied Mathematics, 217(Part 1):53–70, 2017. doi: 10.1016/j.dam.2015.12.015.