Mathematical Programming manuscript No.
(will be inserted by the editor)

Donald Goldfarb · Yiqing Lin

Combinatorial Interior Point Methods for Generalized Network Flow Problems ⋆

Received: June 1, 2001 / Revised version: Revised May 23, 2002

Abstract. We present combinatorial interior point methods for the generalized minimum cost flow and the generalized circulation problems based on Wallacher and Zimmermann's combinatorial interior point method for the minimum cost network flow problem. The algorithms have features of both a combinatorial algorithm and an interior point method. They work towards optimality by iteratively reducing the value of a potential function while maintaining interior point solutions. At each iteration, flow is augmented along a generalized circulation, which is computed by solving a TVPI (Two Variables Per Inequality) system. The algorithms run in $O(m^2n^2(\log m + \log^2 n)L)$ time, where m and n are, respectively, the number of arcs and nodes in the graph, and L is the length of the input data.

1. Introduction

We present combinatorial interior point algorithms for the generalized minimum cost flow and the generalized circulation problems based on Wallacher and Zimmermann's ([11]) combinatorial interior point method for the minimum cost network flow (MCNF) problem. Compared to a "pure" network flow problem, in a generalized network flow problem, each arc has a gain/loss factor associated with it. Since the "pure" and generalized MCNF problems are linear programs, any linear programming algorithm can be used to solve these problems, and in particular, interior point methods can be used to solve them in polynomial time. However, if these methods are applied directly, they are not combinatorial.

Wallacher and Zimmermann use Gonzaga's ([5], [6]) interior point method and exploit the special structure of network problems to derive a combinatorial interior point method for the MCNF problem. We extend Wallacher and Zimmermann's work to generalized network flow problems. Specifically we show how to exploit the special structure of a generalized network to obtain combinatorial interior point methods for the generalized MCNF and the generalized circulation problems.

Donald Goldfarb: Department of Industrial Engineering and Operations Research, Columbia University, New York, New York 10027, Email: goldfarb@columbia.edu

Yiqing Lin: United Technologies Research Center, East Hartford, Connecticut 06108, Email: liny@utrc.utc.com

Mathematics Subject Classification (1991): 20E28, 20G40, 20C20

⋆ This research was supported in part by NSF Grants DMS 94-14438, DMS 95-27124, CDA 97-26385 and DMS 01-04282, and DOE Grant DE-FG02-92ER25126

At each iteration of Gonzaga's interior point method, a descent direction with respect to a potential function is computed. For the MCNF problem, Wallacher and Zimmermann's algorithm reduces the problem of finding such a direction to finding a nonpositive cost cycle with respect to some specified cost, thus making the algorithm combinatorial. For the generalized network flow problems, we can also reduce the problem of finding such directions to finding certain types of flows in the generalized network, thus making the algorithm combinatorial. These flows can be found by solving a system of linear inequalities, each of which has two variables. To do this, we use Cohen and Megiddo's ([3]) algorithm to solve these TVPI (two variables per inequality) systems.

In the next section, we briefly review Wallacher and Zimmermann's potential reduction interior point method on which our algorithms are based. In Section 3, we present our algorithm for the generalized MCNF problem and theorems related to it, and show that the problem of finding a suitable descent direction for the potential function can be solved by finding a nonpositive cost generalized flow. In Section 4, we show how to find a nonpositive cost generalized flow based upon combinatorial algorithms for solving TVPI systems. This section gives a fairly detailed explanation of such algorithms in terms of generalized flows in networks. We briefly outline our algorithm for the generalized circulation problem in Section 5. We discuss how to find an initial feasible interior point solution in Section 6. In Section 7, we compare our algorithms with a related algorithm developed by Wayne ([12]). We also discuss there the extension of our results to the more general problem of optimizing a linear function subject to a TVPI system and an alternative approach that achieves a polylogarithmic improvement in the overall complexity of our algorithms.

2. A potential reduction interior point method

Our algorithms are based on Wallacher and Zimmermann's algorithm for the MCNF problem, which in turn is based on Gonzaga's potential reduction interior point method for the standard form linear programming problem:

$$\begin{aligned} \min\ & c^\top z \\ s.t.\ & Az = b \\ & z \geq 0, \end{aligned} \tag{1}$$

where $c, z \in R^n$, $A \in R^{m\times n}$ and $b \in R^m$. Let α_* be the optimal value of the problem and $\alpha \leq \alpha_*$ be a lower bound. Gonzaga's algorithm reduces the potential function $\phi_\alpha(z) : R^n \to R$:

$$\phi_\alpha(z) = q \cdot \log(c^\top z - \alpha) - \sum_{i=1}^{n} \log z_i$$

at each iteration, where q is a fixed number with $q > n$. It is well known that if L denotes the length of the input data, $\phi_{\alpha^1}(z^1) = O(nL)$ and for a series of interior points $\{z^i\}$ and lower bounds $\{\alpha^i\}$ with $\alpha^i \le \alpha^{i+1}$, $\phi_{\alpha^{i+1}}(z^{i+1}) \le \phi_{\alpha^i}(z^i) - \delta$, for $i = 1, 2, \ldots$ for a fixed number $\delta > 0$, then an interior point z with $c^\top z - \alpha_* \le 2^{-L}$ can be found in $O((q/\delta)L)$ iterations ([11]). The point z can either be accepted as a 2^{-L}-approximate optimal solution or it can be rounded to an optimal solution in polynomial time.

Let $S^0 = \{z \in R^n : Az = b, z > 0\}$ be the set of feasible interior points. For an interior point $\bar{z} = (\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_n)^\top \in S^0$, let $D_{\bar{z}} = diag(\bar{z}_1, \ldots, \bar{z}_n)$ and $y = D_{\bar{z}}^{-1} z$ for any $z \in S^0$. We define the scaled potential function for any $y > 0$ as:

$$\psi_\alpha(y) = q \cdot \log(c^\top D_{\bar{z}} y - \alpha) - \sum_{i=1}^{n} \log y_i.$$

It is trivial to verify that $\phi_\alpha(z^{k+1}) - \phi_\alpha(z^k) = \psi_\alpha(y^{k+1}) - \psi_\alpha(y^k)$ for any two points z^k and $z^{k+1} \in S^0$. Clearly,

$$\nabla \psi_\alpha(y) = \frac{q}{c^\top D_{\bar{z}} y - \alpha} D_{\bar{z}} c - D_y^{-1} \mathbf{1},$$

where $D_y := diag(y_1, \ldots, y_n)$ and $\mathbf{1} = (1, \ldots, 1)^\top$. To simplify the notation, in the remaining of the paper, we will denote $D_{\bar{z}}$ by D.

Let

$$a_\alpha := \nabla \psi_\alpha(\mathbf{1}) = \frac{q}{c^\top D \mathbf{1} - \alpha} Dc - \mathbf{1}.$$

The following theorem is from Wallacher and Zimmermann ([11]):

Theorem 1: (Theorem 2.1 in [11])
For any $\alpha \le \alpha_*$, let h be a search direction (i.e., $ADh = 0$) with $a_\alpha^\top h < 0$. Then

$$\psi_\alpha(\mathbf{1} + \bar{\lambda} h) - \psi_\alpha(\mathbf{1}) \le -\frac{1}{8} \min \left\{ \left(\frac{a_\alpha^\top h}{\|h\|_1} \right)^2, \ 4 \right\},$$

where $\bar{\lambda} = arg\min\{\rho(\lambda) : 0 \le \lambda ||h||_\infty \le \frac{1}{2}\}$ and $\rho(\lambda) = \lambda a_\alpha^\top h + 2\lambda^2 ||h||_2^2$.

At each iteration of Wallacher and Zimmermann's algorithm, a search direction $\hat{h}$ corresponding to a cycle flow that satisfies $\frac{a_\alpha^\top \hat{h}}{\|\hat{h}\|_1} \le -1/2$ is sought. They prove that either such a cycle exists and suffcient improvement can be made by augmenting flow in it, or the lower bound α can be improved. In both cases, the value of the potential function can be reduced by an amount that is greater than a positive constant δ, thus making the algorithm polynomial.

3. Generalized minimum cost flow problem

Let $G = (V, E)$ be a generalized network with $|V| = n$ nodes and $|E| = m$ arcs. Each arc $i \in E$ has integral capacity u_i and gain γ_i, which can be represented as the ratio of two integers, and the flow in it is denoted by x_i. For arc i, we denote its reverse arc by $m + i$ and set $\gamma_{m+i} = 1/\gamma_i$.

Let $F(v)$, $B(v)$ denote the set of arcs leaving and entering node v, i.e., $F(v) = \{(v,w) \in E : w \in V\}$ and $B(v) = \{(w,v) \in E : w \in V\}$. The generalized minimum cost flow problem can be formulated as:

$$\min \sum_{i \in E} \bar{c}_i x_i$$

$$s.t. \sum_{i \in B(v)} \gamma_i x_i - \sum_{i \in F(v)} x_i = \bar{b}_v, \quad v \in V$$

$$0 \leq x_i \leq u_i, \quad i \in E,$$

where $\bar{c}_i$ is the integral cost for each arc $i \in E$, and $\bar{b}_v$ is the integral demand/supply for each node $v \in V$.

The gain of a directed path $P = v_0 e_1 v_1 \dots e_k v_k$ is defined as $\gamma(P) \equiv \prod_{e \in P} \gamma(e)$. The unit cost of this path, i.e., the cost incurred in sending one unit of flow along it, is defined as $c(P) \equiv \sum_{i=1}^{k} c(e_i) \prod_{1 \leq j \leq i-1} \gamma(e_j)$. The gain and unit cost of a directed cycle $W = v_0 e_1 v_1 \dots e_k v_k$ with starting node $v_0 = v_k$ are $\gamma(W) = \prod_{e \in W} \gamma(e)$, and $c(W) = \sum_{i=1}^{k} c(e_i) \prod_{1 \leq j \leq i-1} \gamma(e_j)$, respectively. If $\gamma(W) > 1$ ($\gamma(W) < 1$), then the cycle is called a *flow-generating* (*flow-absorbing*) cycle; if $\gamma(W) = 1$, then the cycle is called a *unit gain cycle*. Sending one unit of flow along a flow-generating (flow-absorbing) cycle W with gain $\gamma(W)$ creates $\gamma(W) - 1$ ($1 - \gamma(W)$) units of excess (deficit) at its starting node; sending flow along a unit gain cycle does not create any excess or deficit.

Any generalized flow can be decomposed into five types of elementary generalized flows. For each type below, the flow x_i is positive only on the arcs in the designated subnetwork.
type 1: a path from a node with deficit to a node with excess
type 2: a flow-generating cycle plus a path connecting the cycle to a node with excess (in the generalized circulation if the node with excess is the sink, then it is called a generalized augmenting path (GAP))
type 3: a path connecting a node with deficit to a flow-absorbing cycle plus the cycle
type 4 (unit gain cycle): a cycle with gain 1
type 5 (bicycle): a path connecting a flow-generating cycle to a flow-absorbing cycle plus those cycles

If the path in a bicycle (GAP) consists of a single node, we call the bicycle (GAP) *degenerate*. Also, the arcs in the two cycles of a bicycle need not be disjoint. We have the following theorem from [4]:

Theorem 2: (Theorem 2.6 in [4])
A generalized flow g can be decomposed into components $g_1, \dots, g_k$ with $k \leq m$ such that

$$g = \sum_i g_i,$$

where each component g_i is an elementary generalized flow of one of the above five types that is positive only on arcs (v, w) with $g(v, w) > 0$. This decomposition can be found in $O(mn)$ time.

We call a generalized flow a *generalized circulation* if the flow satisfies the flow conservation constraints at all nodes. We call unit gain cycles and bicycles *elementary generalized circulations*, since they do not create excess or deficit at any node. It is easy to see that a generalized circulation can be decomposed into elementary generalized circulations.

If we introduce slack variables $y_i = u_i - x_i$, $i \in E$, the generalized MCNF problem can be written as a standard form linear program (1), where $z^\top = (x^\top, y^\top)$, $b^\top = (\bar{b}^\top, u^\top)$, $c^\top = (\bar{c}^\top, 0)$,

$$A = \begin{pmatrix} \bar{A} & 0 \\ I & I \end{pmatrix}, \text{ and } \bar{A} = \{\bar{a}_{vi}\},\ \bar{a}_{vi} = \begin{cases} \gamma_i, & i \in B(v) \\ -1, & i \in F(v) \\ 0, & \text{otherwise.} \end{cases}$$

Let $z^\top = (x^\top, y^\top)$ be a feasible interior point. Then x is a feasible generalized flow in G with flow value x_i $(0 < x_i < u_i)$ on arc i. Let $G_x = (V, E_x)$ be the residual graph of G corresponding to x. Clearly, every arc in E and its reverse arc are in E_x.

Let $\Delta z \equiv (\Delta x, \Delta y)^\top$ satisfy $A\Delta z = 0$, then $\Delta x_i = \Delta z_i = -\Delta z_{m+i} = -\Delta y_i$. Δz corresponds to a generalized flow (circulation) $g = g(\Delta z)$ in G_x, where for all $i \in E_x$, $g_i = \max(\Delta x_i, 0)$, and $g_{m+i} = \max(\Delta y_i, 0) \cdot \gamma_i$. Hence g satisfies $g_i \cdot g_{m+i} = 0$ for all $i \in E_x$. Conversely, if g is a generalized circulation in G_x, then it corresponds to some Δz with $A\Delta z = 0$, where

$$-\Delta z_{m+i} = \Delta z_i = \begin{cases} g_i & \text{if } g_i > 0 \\ -g_{m+i}/\gamma_i & \text{otherwise.} \end{cases}$$

In our algorithm, any search direction h at the current point z must satisfy $ADh = 0$ where $D := diag(z_1, \dots, z_{2m})$; i.e., $g(Dh)$ is a generalized circulation. If z_{opt} is an optimal solution, there exists a search direction h such that $z_{opt} = z + Dh$ and $g(Dh)$ can be decomposed into elementary generalized circulations.

In the remainder of this section we present a number of theorems and formally state our basic algorithm. The theorems, their proofs and the basic algorithm are essentially identical to those in ([11]) except that the set of cycleflows used by Wallacher and Zimmermann is replaced by the set of elementary generalized circulations. We provide these proofs for the sake of completeness.

Let h be a search direction and $\tilde{C}_x$ denote the set of elementary generalized circulations in G_x. From Theorem 2, the corresponding nonnegative flow $g(Dh)$ can be decomposed into elementary generalized circulations with $g^i(\Delta z^i) \in \tilde{C}_x$, $1 \le i \le r$, i.e.,

$$g(Dh) = \mu_1 g^1(\Delta z^1) + \cdots + \mu_r g^r(\Delta z^r),$$

with $r \leq 2m$ and $\mu_i > 0$, $i = 1, 2, \ldots, r$, and where $g_i^\nu \cdot g_{m+i}^\rho = 0$, $1 \leq i \leq m, 1 \leq \nu, \rho \leq r$. Moreover, $Dh = \mu_1 \Delta z^1 + \cdots + \mu_r \Delta z^r$ satisfies

$$(Dh)_i \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} 0 \Longrightarrow \Delta z_i^\nu \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} 0,\ 1 \leq i \leq 2m, 1 \leq \nu \leq r.$$

Let $h^\nu := D^{-1}\Delta z^\nu$, $\nu = 1, 2, \ldots, r$. Then $h = \mu_1 h^1 + \cdots + \mu_r h^r$ and

$$h_i \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} 0 \Longrightarrow h_i^\nu \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} 0,\ 1 \leq i \leq 2m, 1 \leq \nu \leq r.$$

This implies that $\|h\|_1 = \|\mu_1 h^1\|_1 + \cdots + \|\mu_r h^r\|_1$.

Lemma 1: (Lemma 3.2 in [11])
Let $a_1, \ldots, a_s \in R$ and $b_1, \ldots, b_s$ be positive real numbers. Then $\min_{1 \leq \nu \leq s} \frac{a_\nu}{b_\nu} \leq \frac{a_1 + \cdots + a_s}{b_1 + \cdots + b_s}$.

Based on this lemma, we have the following theorem:

Theorem 3: (analogue of Theorem 3.1 in [11])
If $a_\alpha = \nabla\psi_\alpha(\mathbf{1}) = \frac{q}{c^\top D\mathbf{1} - \alpha} Dc - \mathbf{1}$ and h is a search direction with $a_\alpha^\top h < 0$, then

$$0 > \frac{a_\alpha^\top h}{\|h\|_1} \geq \frac{a_\alpha^\top \tilde{h}}{\|\tilde{h}\|_1},\ \text{for some } \tilde{h} \text{ with } g(D\tilde{h}) \in \tilde{C}_x.$$

Proof.

$$0 > \frac{a_\alpha^\top h}{\|h\|_1} = \frac{\mu_1 a_\alpha^\top h^1 + \cdots + \mu_r a_\alpha^\top h^r}{\|\mu_1 h^1\|_1 + \cdots + \|\mu_r h^r\|_1} \geq \min\{\frac{a_\alpha^\top h^\nu}{\|h^\nu\|_1} : 1 \leq \nu \leq r\} = \frac{a_\alpha^\top \tilde{h}}{\|\tilde{h}\|_1}.$$

□

Theorem 4: (analogue of Theorem 3.2 in [11])
Let $q \geq 4m, \alpha = \alpha_*$, and $a_\alpha := \nabla\psi_\alpha(\mathbf{1})$. Then $\frac{a_\alpha^\top \tilde{h}}{\|\tilde{h}\|_1} \leq -1$ for some $\tilde{h}$ with $g(D\tilde{h}) \in \tilde{C}_x$.

Proof. Let h be the step from z to some optimal point z_{opt}; i.e., $z_{opt} = D(\mathbf{1}+h)$. Since $c^\top D(\mathbf{1} + h) = \alpha_*$,

$$\frac{a_\alpha^\top h}{\|h\|_1} = \frac{1}{\|h\|_1}\left(\frac{q}{c^\top D\mathbf{1} - c^\top D\mathbf{1} - c^\top Dh} c^\top Dh - \mathbf{1}^\top h\right) = -\frac{q + \mathbf{1}^\top h}{\|h\|_1}.$$

Define h^+ and $h^- \in R^{2m}$ by

$$h_i^+ = \begin{cases} h_i & \text{if } h_i \geq 0, \\ 0 & \text{otherwise,} \end{cases} \qquad h_i^- = \begin{cases} h_i & \text{if } h_i \leq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$\mathbf{1}^\top h = \mathbf{1}^\top h^+ + \mathbf{1}^\top h^- = \|h^+\|_1 - \|h^-\|_1,$$
$$\|h\|_1 = \mathbf{1}^\top h^+ - \mathbf{1}^\top h^- = \|h^+\|_1 + \|h^-\|_1,$$

and $0 \leq \|h^-\|_1 \leq 2m,$ since $\mathbf{1} + h \geq 0$ (h is a feasible step). Therefore

$$\frac{a_\alpha^\top h}{\|h\|_1} = -\frac{q + \mathbf{1}^\top h}{\|h\|_1} = -\frac{q + \|h^+\|_1 - \|h^-\|_1}{\|h^+\|_1 + \|h^-\|_1} \leq -\frac{2m + \|h^+\|_1}{\|h^+\|_1 + 2m} = -1,$$

and from Theorem 3

$$-1 \geq \frac{a_\alpha^\top h}{\|h\|_1} \geq \frac{a_\alpha^\top \tilde{h}}{\|\tilde{h}\|_1}$$

for some $\tilde{h}$ with $g(D\tilde{h})$ in the decomposition of $g(Dh)$. But from the previous discussion, $g(D\tilde{h}) \in \tilde{C}_x$.

□

The next theorem provides criteria for when to improve the lower bound:

Theorem 5: (analogue of Theorem 3.3 in [11])
Let $q \geq 4m$, $\alpha \leq \alpha_*$, $a_\alpha := \nabla\psi_\alpha(\mathbf{1})$, and $0 < s < 1$. If $\frac{a_\alpha^\top \tilde{h}}{\|\tilde{h}\|_1} > -s$ for all $\tilde{h}$ with $g(D\tilde{h}) \in \tilde{C}_x$, then

$$\tilde{\alpha} = \frac{1-s}{2} c^\top D\mathbf{1} + \frac{1+s}{2}\alpha$$

is a lower bound for α_* with

$$\frac{c^\top D\mathbf{1} - \tilde{\alpha}}{c^\top D\mathbf{1} - \alpha} = \frac{1+s}{2}.$$

Proof. Let h be a step such that $z_{opt} = D(\mathbf{1}+h)$. Then $c^\top D\mathbf{1} + c^\top Dh = \alpha_*$ and from the previous theorem

$$\frac{a_{\alpha_*}^\top h}{\|h\|_1} = -\frac{q + \mathbf{1}^\top h}{\|h\|_1} = -\theta \leq -1.$$

Hence,

$$\|h\|_1 = \frac{q + \mathbf{1}^\top h}{\theta} \text{ with } \theta \geq 1.$$

Since $a_\alpha^\top \tilde{h}/\|\tilde{h}\|_1 > -s$ for all $\tilde{h}$ with $g(D\tilde{h}) \in \tilde{C}_x$, and h is a direction such that $g(Dh)$ can be decomposed into flows in $\tilde{C}_x$,

$$\frac{a_\alpha^\top h}{\|h\|_1} = \frac{\theta}{q + \mathbf{1}^\top h}\left(\frac{q(\alpha_* - c^\top D\mathbf{1})}{c^\top D\mathbf{1} - \alpha} - \mathbf{1}^\top h\right) > -s.$$

Since $\mathbf{1}^\top h \geq -\|h^-\|_1 \geq -2m$, $q + \mathbf{1}^\top h \geq 2m$ and we have

$$\frac{q(\alpha_* - c^\top D\mathbf{1})}{c^\top D\mathbf{1} - \alpha} > \frac{-s(q + \mathbf{1}^\top h)}{\theta} + \mathbf{1}^\top h \geq -s(q + \mathbf{1}^\top h) + \mathbf{1}^\top h$$

$$= (1-s)\mathbf{1}^\top h - sq \geq -2m(1-s) - sq.$$

From $c^\top D\mathbf{1} - \alpha > 0$, ($\alpha$ is a lower bound for the current objective value $c^\top D\mathbf{1}$),

$$\begin{aligned}
\alpha_* &> -\frac{2m(1-s)}{q}(c^\top D\mathbf{1} - \alpha) - s(c^\top D\mathbf{1} - \alpha) + c^\top D\mathbf{1} \\
&\geq -\frac{1-s}{2}(c^\top D\mathbf{1} - \alpha) - s(c^\top D\mathbf{1} - \alpha) + c^\top D\mathbf{1} \\
&= \frac{1-s}{2}c^\top D\mathbf{1} + \frac{1+s}{2}\alpha = \tilde{\alpha}.
\end{aligned}$$

Thus $\tilde{\alpha}$ is a lower bound for α_*. Furthermore, it is easy to verify that $\frac{c^\top D\mathbf{1}-\tilde{\alpha}}{c^\top D\mathbf{1}-\alpha} = \frac{1+s}{2}$.

□

Let $q = 4m$ and $s = \frac{1}{2}$. If there exists an elementary generalized circulation $g(D\tilde{h}) \in \tilde{C}_x$ such that $a_\alpha^\top \tilde{h}/\|\tilde{h}\|_1 \leq -\frac{1}{2}$, then from Theorem 1, the value of the scaled potential function can be reduced by at least $\frac{1}{32}$ by taking a step in the direction $\tilde{h}$. Otherwise, from Theorem 5, the lower bound α can be improved, and the decrease of the potential function is at least

$$\psi_{\tilde{\alpha}}(\mathbf{1}) - \psi_\alpha(\mathbf{1}) = q \cdot \log \frac{c^\top D\mathbf{1} - \tilde{\alpha}}{c^\top D\mathbf{1} - \alpha} = 4m \log \frac{3}{4} < -1 \text{ for } m \geq 1.$$

Therefore, an algorithm based on the above takes at most $O(mL)$ iterations. To complete the description of such an algorithm we must show how to compute the step direction $\tilde{h}$ combinatorially.

In Wallacher and Zimmermann's algorithm, since the set $\tilde{C}_x$ is the set of cycleflows, the problem of finding a step direction $\tilde{h}$ corresponds to finding a cycle $g(D\tilde{h}) \in \tilde{C}_x$ with ratio $a_\alpha^\top \tilde{h}/\|\tilde{h}\|_1 \leq -1/2$. Wallacher and Zimmermann show that there is a way to assign a cost to each arc in the residual graph so that this problem can be further reduced to finding a nonpositive cost cycle. Such cycles can be found by using a shortest-path algorithm.

In our algorithm, the set $\tilde{C}_x$ is more complicated. However, we can still assign a cost to each arc in the residual graph so that the step direction subproblem is reduced to finding a nonpositive cost generalized circulation. We use Cohen and Megiddo's algorithm for solving TVPI systems to find such a generalized circulation. This is described in Section 4.

Now each direction h with $g(Dh) \in \tilde{C}_x$ satisfies

$$-D_{m+i}h_{m+i} = D_i h_i = \begin{cases} \Gamma_i & \text{if } i \in C \\ -\Gamma_{m+i}/\gamma_i & \text{if } m+i \in C \\ 0 & \text{otherwise,} \end{cases}$$

where C is an elementary generalized circulation in $\tilde{C}_x$, and Γ_i is the flow value on arc i in C so that flow conservation is maintained at each node. We have

$$1^\top h = \sum_{i=1}^{2m} h_i = \sum_{i\in C, 1\leq i\leq m} (\frac{1}{x_i} - \frac{1}{y_i})\Gamma_i + \sum_{m+i\in C, 1\leq i\leq m} (-\frac{1}{x_i} + \frac{1}{y_i})\frac{\Gamma_i}{\gamma_i},$$

$$\|h\|_1 = \sum_{i=1}^{2m} |h_i| = \sum_{i\in C, 1\le i\le m} (\frac{1}{x_i} + \frac{1}{y_i})\Gamma_i + \sum_{m+i\in C, 1\le i\le m} (\frac{1}{x_i} + \frac{1}{y_i})\frac{\Gamma_i}{\gamma_i},$$

and, recalling that $c_i = \bar{c}_i$, for $1 \le i \le m$ and $c_i = 0$, for $m < i \le 2m$,

$$c^\top Dh = \sum_{i=1}^{2m} c_i D_i h_i = \sum_{i\in C, 1\le i\le m} c_i \Gamma_i + \sum_{m+i\in C, 1\le i\le m} (-c_i \frac{\Gamma_i}{\gamma_i}).$$

If we set

$$-\gamma_i \sigma_{m+i} \equiv \sigma_i \equiv \frac{q}{c^\top D1 - \alpha} c_i - (\frac{1}{x_i} - \frac{1}{y_i}) \quad \text{and} \quad \gamma_i \tau_{m+i} \equiv \tau_i \equiv \frac{1}{x_i} + \frac{1}{y_i},\ 1 \le i \le m,$$

then

$$a_\alpha^\top h = \frac{q}{c^\top D1 - \alpha} c^\top Dh - 1^\top h = \sum_{i\in C, 1\le i\le 2m} \sigma_i \Gamma_i = \sigma(C)$$

and

$$\|h\|_1 = \sum_{i\in C, 1\le i\le 2m} \tau_i \Gamma_i = \tau(C).$$

Since $\tau(C) > 0$,

$$\frac{\sigma(C)}{\tau(C)} \le -\frac{1}{2} \Longleftrightarrow \sigma(C) + \frac{1}{2}\tau(C) \le 0 \Longleftrightarrow \sum_{i\in C, 1\le i\le 2m} (\sigma_i + \frac{1}{2}\tau_i)\Gamma_i \le 0.$$

If we define $\hat{c}_i = \sigma_i + \frac{1}{2}\tau_i$ for $1 \le i \le 2m$, we change the problem of finding h with $g(Dh) \in \tilde{C}_x$ and $a_\alpha^\top h/\|h\|_1 \le -1/2$ into finding a nonpositive cost elementary generalized circulation in G_x.

Combining the results of Theorems 1 and 5 gives the following algorithm:

Algorithm GN(*generalized network*):(analogue of Algorithm 3.1 in [11])
Given $\alpha \le \alpha_*$, $q = 4m$ and an initial interior solution z:
while $c^\top z - \alpha > 2^{-L}$ **do**
  if there is a nonpositive cost generalized circulation $g(D\hat{h})$ with respect to cost $\hat{c}$, **then** find such a circulation and the corresponding step direction $\hat{h}$;
  Compute the step size λ^* by doing an approximate line search such that $\psi_\alpha(\mathbf{1} + \lambda^*\hat{h}) \le \psi_\alpha(\mathbf{1}) + \rho(\bar{\lambda})$, where $\rho(\bar{\lambda}) = \min\{\lambda a_\alpha^\top \hat{h} + 2\lambda^2\|\hat{h}\|_2^2 :$ $0 \le \lambda\|\hat{h}\|_\infty \le 1/2\}$ and
  set $z := D(\mathbf{1} + \lambda^*\hat{h})$;
    else (update lower bound) $\alpha \leftarrow \frac{1}{4}c^\top D\mathbf{1} + \frac{3}{4}\alpha$;
  end if;
end while.

In Section 4, we will show how to compute a suitable step direction or determine that none exists in $O(mn^2(\log m + \log^2 n))$ time. Consequently, the worst case running time of Algorithm GN is $O(m^2n^2(\log m + \log^2 n)L)$.

4. Two variables per inequality systems

In this section we describe how to solve the step direction problem in Algorithm GN. With this in mind, consider the following problem in G_x:

$$\begin{aligned} z^* = \min \sum_{i \in E_x} \hat{c}_i x_i \\ s.t. \sum_{i \in B(v)} \gamma_i x_i - \sum_{i \in F(v)} x_i = 0, \quad v \in V \\ x_i \geq 0, \quad i \in E_x, \end{aligned} \tag{2}$$

and its dual

$$\begin{aligned} \max \ \sum_{i \in V} 0 \cdot \pi_i \\ s.t.\ \pi_k - \gamma_{kl} \pi_l \leq \hat{c}_{kl} \quad \forall (k,l) \in E_x. \end{aligned} \tag{3}$$

Since each solution z is an interior point, all arcs in the original graph and their reverse arcs are in E_x. When solving problem (2), the following cases are possible.

(i): Problem (2) is unbounded. Then there exists a negative cost elementary generalized circulation that is a suitable direction $\hat{h}$ for Algorithm GN.
(ii): Problem (2) has an optimal solution $z^* = 0$.
 (a) If there exists an elementary generalized circulation with zero cost, then it is a suitable direction $\hat{h}$;
 (b) Otherwise, we update the lower bound.

The system of linear inequalities in problem (3) is a TVPI system. There are algorithms for TVPI systems that can either find a feasible solution to (3) or identify a certificate of infeasibility that can be used to derive a negative cost elementary generalized circulation. We show that by using some intermediate results from those algorithms, we can further check if there exists an elementary generalized circulation with zero cost. Although our step direction algorithm is based on algorithms for TVPI systems, we will describe it in the context of generalized network flows. As far as we can tell, this description is new. Moreover, it may be useful in developing improved algorithms for solving or determining a certificate of infeasibility of a TVPI system.

Let S denote the feasible region of problem (3), i.e., $S = \{\pi | \pi_k - \gamma_{kl}\pi_l \leq \hat{c}_{kl},\ \forall (k,l) \in E_x\}$, and $\mathcal{X}_k(S)$ be the projection of the region S on the π_k-axis. We can view $\mathcal{X}_k(S)$ as the set of values of π_k for which a solution to the full TVPI system exists, i.e., $\mathcal{X}_k(S) = \{\hat{\pi}_k : S \cap \{\pi | \pi_k = \hat{\pi}_k\} \neq \emptyset\}$. Since S is a closed convex polyhedron, $\mathcal{X}_k(S)$ is a closed interval for each k, which we denote by $[\pi_{k(min)}, \pi_{k(max)}]$. Shostak([9]), Aspvall and Shiloach([2]), Megiddo ([8])and Cohen and Megiddo ([3]) developed algorithms that find a feasible solution to a TVPI system by first finding $\pi_{k(min)}$ and $\pi_{k(max)}$ for each variable. We can use any of these algorithms to find a suitable direction for our Algorithm GN or determine that none exists. All of these algorithms work as follows:

Consider a directed cycle $W = v_1 e_1 v_2 e_2 \ldots e_{l-1} v_l e_l v_1$ in G_x that starts and ends at node v_1. Each arc e_j on this cycle corresponds to an inequality of the form $\pi_j - \gamma_j \pi_{j+1} \leq \hat{c}_j$. Consequently,

$$\begin{aligned} \pi_1 &\leq \hat{c}_1 + \gamma_1 \pi_2 \leq \hat{c}_1 + \gamma_1(\hat{c}_2 + \gamma_2 \pi_3) = \hat{c}_1 + \gamma_1 \hat{c}_2 + \gamma_1 \gamma_2 \pi_3 \leq \cdots \\ &\leq \hat{c}_1 + \gamma_1 \hat{c}_2 + \cdots + \gamma_1 \gamma_2 \cdots \gamma_{l-1} \hat{c}_l + \gamma_1 \gamma_2 \cdots \gamma_l \pi_1, \end{aligned}$$

i.e.,

$$\pi_1 \leq \hat{c}(W) + \gamma(W) \pi_1. \tag{4}$$

If $\gamma(W) > 1$, then for any feasible solution π, $\pi_1 \geq \hat{c}(W)/(1-\gamma(W))$; if $\gamma(W) < 1$, then $\pi_1 \leq \hat{c}(W)/(1-\gamma(W))$; if $\gamma(W) = 1$ and $\hat{c}(W) \geq 0$, the cycle W does not impose any constraint on π_1; if $\gamma(W) = 1$ and $\hat{c}(W) < 0$, inequality (4) is not satisfiable, and thus the TVPI system does not have any feasible solution. We call a cycle with $\gamma(W) = 1$ and $\hat{c}(W) < 0$ an *infeasible directed cycle*.

From the above discussion, we see that each directed cycle that starts and ends at node v_k and has gain not equal to 1 (there may be an exponential number of them) gives an inequality of the form $\pi_k \geq h$ or $\pi_k \leq h$. Let $\pi_k \geq \pi_k low$ and $\pi_k \leq \pi_k high$ be the two tightest such inequalities. Aspvall and Shiloach ([2]) present a routine that either detects an infeasible directed cycle or finds $\pi_k low$ and $\pi_k high$ using binary search for each variable along with the directed cycles that give these bounds. In each iteration of their algorithm, a guessed value of π_k is "pushed" through the graph in a breadth-first manner. By analyzing the outcome, it is possible to reduce the interval for either $\pi_k low$ or $\pi_k high$ by at least half. Their routine runs in $O(mn^3 L)$ time. Then for each variable, given $\pi_k low$ and $\pi_k high$, $\pi_{k(min)}$ and $\pi_{k(max)}$ can be found in $O(mn)$ time by a Bellman-Ford like algorithm. Therefore, their complete algorithm for finding $\pi_{k(min)}$ and $\pi_{k(max)}$ for a TVPI system runs in $O(mn^3 L)$ time.

Megiddo's ([8]) strongly polynomial time algorithm further improves upon Aspvall and Shiloach's algorithm. Megiddo's algorithm for finding $\pi_k low$ and $\pi_k high$, as well as finding a solution to a TVPI system runs in $O(mn^3 \log m)$ time. Cohen and Megiddo ([3]) give another algorithm that solves the TVPI system even faster – i.e., in $O(mn^2(\log m + \log^2 n))$ time, which we use to evaluate the complexity of our algorithm. Though these algorithms use different methods from Aspvall and Shiloach's, they are still combinatorial.

Now, let us discuss how to find the step direction for Algorithm GN or decide that none exists. First we use any algorithm that finds $\pi_k low$ and $\pi_k high$ for each variable or gives an infeasible directed cycle, i.e., a unit gain cycle with negative cost, if the TVPI system is infeasible. In the latter case, problem (3) is infeasible (and Problem (2) is unbounded). This cycle provides a suitable direction and we are done. In the former case we proceed as follows.

Let A denote the flow-generating cycle corresponding to $\pi_k \geq \pi_k low$, i.e., $\pi_k low = \hat{c}(A)/(1-\gamma(A))$. From the fact that $\pi_k low$ is the tightest bound, we have $\pi_k low \geq \hat{c}(W)/(1-\gamma(W))$, for any cycle W starting from node v_k with $\gamma(W) > 1$. Thus

$$\frac{\hat{c}(W)}{\gamma(W)-1} \geq -\pi_k low = \frac{\hat{c}(A)}{\gamma(A)-1}.$$

Since $(\gamma(W)-1)$ and $\hat{c}(W)$ are respectively the units of flow generated and the cost incurred by sending one unit flow along the flow-generating cycle W starting from node v_k, we can interpret $(-\pi_k low)$ as the minimum cost of generating one unit of flow at node v_k through flow-generating cycles starting and ending at node v_k.

Similarly, let B denote the flow-absorbing cycle corresponding to the inequality $\pi_k \leq \pi_k high$, i.e., $\pi_k high = \hat{c}(B)/(1-\gamma(B))$. We have

$$\pi_k high \leq \hat{c}(W)/(1-\gamma(W)),$$

for any cycle W starting from node v_k with $\gamma(W) < 1$. Therefore, $\pi_k high$ represents the minimum cost of absorbing one unit of flow at node v_k through flow-absorbing cycles starting and ending at node v_k.

If $\pi_k low = -\infty$ ($\pi_k high = \infty$), there is no flow-generating (flow-absorbing) cycle passing through node v_k. If $\pi_k low = -\infty$ and $\pi_k high = \infty$ for every variable, there are no flow-generating or flow-absorbing cycles in G_x, and hence in G, since $E \subseteq E_x$. In this case, since every cycle has unit gain and nonnegative cost, we can check if there is a zero-cost unit gain cycle using a shortest path routine that handles networks with negative cost arcs and no negative cycles. If such a cycle is found, it provides a suitable step direction. In the following discussion, we assume $\pi_k low \neq -\infty$ or $\pi_k high \neq \infty$ for at least one variable.

If for some variable π_k, $\pi_k low \geq \pi_k high$, then we can find a nonpositive cost degenerate bicycle with node v_k as the common node on both cycles. Again, let A, B denote the cycles achieving $\pi_k low$ and $\pi_k high$, respectively. We have

$$\pi_k high - \pi_k low = \frac{\hat{c}(B)}{1-\gamma(B)} + \frac{\hat{c}(A)}{\gamma(A)-1} \leq 0,$$

or equivalently, since $\gamma(A) > 1$

$$\hat{c}(A,B) \equiv \hat{c}(A) + \frac{\gamma(A)-1}{1-\gamma(B)}\hat{c}(B) \leq 0.$$

But $\hat{c}(A,B)$ is just the cost of sending one unit flow along the bicycle starting from node v_k. Sending one unit of flow through cycle A creates $\gamma(A)-1$ units of excess at node v_k and incurs a cost of $\hat{c}(A)$. In order to cancel this excess, we have to send $(\gamma(A)-1)/(1-\gamma(B))$ units of flow along cycle B starting from node v_k, incurring a cost of $(\gamma(A)-1)/(1-\gamma(B)) \cdot \hat{c}(B)$. Conversely, if there exists a nonpositive cost degenerate bicycle that has node v_k on both the flow-generating cycle A and the flow-absorbing cycle B, then $\pi_k low \geq \pi_k high$. Thus we have the following:

Theorem 6: There is a nonpositive cost degenerate bicycle in the residual graph G_x with node v_k as a common node on both cycles if and only if $\pi_k low \geq \pi_k high$.

If, on the other hand, $\pi_k low < \pi_k high$ for all nodes v_k, we can either find a nonpositive cost generalized circulation or determine that none exists using the following Bellman-Ford like algorithm:

Algorithm BF(*Bellman-Ford*):
for each node v_k, set $\pi_k^l = \pi_k low$, $\pi_k^h = \pi_k high$;
for $p = 1$ to n **do**
  scan all arcs in E_x: for each arc $(u, v) \in G_x$ scanned, set
  $\pi_v^l = \max\{\pi_v^l, (\pi_u^l - \hat{c}_{uv})/\gamma_{uv}\}$ and $\pi_u^h = \min\{\pi_u^h, \gamma_{uv}\pi_v^h + \hat{c}_{uv}\}$;
  if $\pi_v^l \geq \pi_v^h$, **then** STOP (we have found a nonpositive cost bicycle in G_x).
  end if;
end for.

If the cost of generating one unit flow at node u is $-\pi_u^l$, then $(-\pi_u^l + \hat{c}_{uv})/\gamma_{uv}$ is the cost of generating one unit flow at node v through arc (u, v). Similarly, if the cost of absorbing one unit flow at node v is π_v^h, then $\gamma_{uv}\pi_v^h + \hat{c}_{uv}$ denotes the cost of absorbing one unit flow at node u through arc (u, v). Initially, $-\pi_u^l = -\pi_u low$ ($\pi_u^h = \pi_u high$) is the minimum costs of generating (absorbing) one unit flow along any simple flow-generating (flow-absorbing) cycle starting from node u. When we update π_v^l through $\pi_v^l = \max\{\pi_v^l, (\pi_u^l - \hat{c}_{uv})/\gamma_{uv}\}$, we are using the criterion of finding the minimum cost of generating one unit of flow at node v through cycles and paths. Thus at any phase of Algorithm BF, $-\pi_u^l$ denotes our *current* estimate of this minimum cost. Similarly, π_v^h denotes the *current* estimate of the minimum cost of absorbing one unit of flow at node v through cycles and paths.

Theorem 7: If $\pi_u low < \pi_u high$ for each node u, then there is a nonpositive cost bicycle in G_x if and only if $\pi_v^l \geq \pi_v^h$ for some node v.

Proof. Since $\pi_v low < \pi_v high$, if $\pi_v^l \geq \pi_v^h$ in Algorithm BF, at least one of π_v^l and π_v^h has been updated. Suppose $\pi_v^l = (\pi_u low - \hat{c}(P))/\gamma(P)$ for some node u, and some path P from u to v, and $\pi_v^h = \gamma(P')\pi_{u'} high + \hat{c}(P')$ for some node u' and path P' from v to u'. If $P = \emptyset$, then $\pi_v^l = \pi_v low$ and if $P' = \emptyset$, then $\pi_v^h = \pi_v high$. From $\pi_v^l \geq \pi_v^h$, we have

$$\pi_v^h - \pi_v^l = \gamma(P')\pi_{u'} high + \hat{c}(P') - \frac{\pi_u low - \hat{c}(P)}{\gamma(P)} \leq 0$$

or equivalently,

$$\begin{aligned}\hat{c}(A, P'', B) &\equiv [-\pi_u low + \hat{c}(P'') + \gamma(P'')\pi_{u'} high](\gamma(A) - 1) \\ &= \hat{c}(A) + [\gamma(A) - 1]\hat{c}(P'') + \frac{(\gamma(A) - 1)\gamma(P'')}{1 - \gamma(B)}\hat{c}(B) \leq 0,\end{aligned}$$

where P'' is the path from node u to node u' obtained by concatenating the paths P and P' (note that P'' may not be a simple path), $\gamma(P'') = \gamma(P)\gamma(P')$, $\hat{c}(P'') = \hat{c}(P) + \gamma(P)\hat{c}(P')$, and A and B are the cycles corresponding to $\pi_u low$ and $\pi_{u'} high$. $\hat{c}(A, P'', B)$ is the cost of sending one unit of flow along cycle A starting from node u, $(\gamma(A) - 1)$ units of flow along path P'' from node u to node u' and $(\gamma(A)-1)\gamma(P'')/(1-\gamma(B))$ units of flow along cycle B starting from node u'. It is easy to see that this flow does not create excess or deficit at any nodes.

We now prove that if there is a nonpositive cost bicycle, then $\pi_v^l \geq \pi_v^h$ for some variable. Suppose there is a bicycle that contains a flow-generating cycle A, a flow-absorbing cycle B and a simple path P'' that connects cycles A and B at nodes u and u'. Then the cost of sending one unit flow starting from node u along the bicycle is $\hat{c}(A, P'', B)$.

From the definition of $\pi_u low$ and $\pi_{u'} high$ and the fact that $\gamma(A) > 1$, $\hat{c}(A, P'', B) \leq 0$ is equivalent to

$$\pi_u low \geq -\frac{\hat{c}(A)}{\gamma(A)-1} \geq \hat{c}(P'') + \gamma(P'')\frac{\hat{c}(B)}{1-\gamma(B)} \geq \hat{c}(P'') + \gamma(P'')\pi_{u'} high.$$

Since path P'' has length less than n, after n iterations of Algorithm BF,

$$\pi_{u'}^l \geq \frac{\pi_u low - \hat{c}(P)}{\gamma(P)}.$$

Therefore

$$\hat{c}(P) + \gamma(P)\pi_{u'} high \leq \pi_u low \leq \hat{c}(P) + \gamma(P)\pi_{u'}^l,$$

which implies that $\pi_{u'}^h \leq \pi_{u'} high \leq \pi_{u'}^l$ for node u'.

□

Aspvall and Shiloach ([2]) use Algorithm BF to find $\pi_{u(min)}$ and $\pi_{u(max)}$. Therefore, $-\pi_{u(min)}$ and $\pi_{u(max)}$ can be interpreted as the minimum cost of generating and absorbing one unit flow at node u, using paths and cycles.

Note that a bicycle $\bar{G}$ in the residual graph G_x may contain both arc i and its reverse arc $m+i$; i.e., for a given flow g' in $\bar{G}$, both $g'_i > 0$ and $g'_{m+i} > 0$. When we add the generalized circulation g' to the current generalized flow in the original graph G, we must first transform (reduce) g' so that for every arc i in $\bar{G}$ either it or its reverse arc has zero flow, i.e., in the transformed flow $\bar{g}$, $\bar{g}_i \cdot \bar{g}_{m+i} = 0$. For each arc $i \in \bar{G}$, if $g'_i \geq g'_{m+i}/\gamma_i$, then we set $\bar{g}_i = g'_i - g'_{m+i}/\gamma_i \geq 0$ and $\bar{g}_{m+i} = 0$; if $g'_i < g'_{m+i}/\gamma_i$, then we set $\bar{g}_i = 0$ and $\bar{g}_{m+i} = g'_{m+i} - \gamma_i g'_i$. It is easy to see that the flow $\bar{g}$ does not create excess or deficit and $\hat{c}(\bar{g}) = \sum_{i \in E_x} \hat{c}_i \bar{g}_i \leq \sum_{i \in E_x} \hat{c}_i g'_i = \hat{c}(g') \leq 0$. The latter follows from the facts that

$$\hat{c}_i = \alpha_i + \frac{1}{2}\beta_i,\ \hat{c}_{m+i} = -\frac{\alpha_i}{\gamma_i} + \frac{1}{2}\cdot\frac{\beta_i}{\gamma_i} \text{ and } \hat{c}_i + \gamma_i \hat{c}_{m+i} = \beta_i > 0, \quad 1 \leq i \leq m,$$

and we are decreasing flows in arcs i and $m+i$, by some positive amounts θ and $\gamma_i\theta$, respectively, which decreases the total cost by $\theta\beta_i$. Hence, $\bar{g}$ is a nonpositive cost generalized circulation and is a suitable step direction for Algorithm GN. Figure 1 illustrates this transformation for a degenerate bicycle.

In Figure 1 we are assuming that the flow in path $a{\to}c$ in bicycle I is set to zero to give bicycle II. If we assume in I that $\gamma(A) > 1$ and $\gamma(B) < 1$, where cycle $A = (a{\to}b{\to}c{\to}a)$ and cycle $B = (a{\to}c{\to}d{\to}a)$, then the cycles in II are A and $C = (a{\to}b{\to}c{\to}d{\to}a)$. Since $\gamma(c{\to}a)\gamma(a{\to}c) = 1$, $\gamma(C) = \gamma(A)\gamma(B)$. Furthermore, assume that $e = (a', a)$ is the last arc on path $c \rightarrow a$. Then if one unit of flow is sent through the bicycle in I starting at node a, $g(a', a) = \gamma(A)/\gamma(a', a)$ and $g(a, a') = (\gamma(A) - 1)/(1 - \gamma(B))$. From

$$g(a', a) - g(a, a')/\gamma(a', a) = \frac{1 - \gamma(A)\gamma(B)}{\gamma(a', a)(1 - \gamma(B))} \geq 0$$

we have $\gamma(C) = \gamma(A)\gamma(B) \leq 1$. If $\gamma(C) < 1$, then II is a valid bicycle. Note that the arcs in cycles A and C are not disjoint. If $\gamma(C) = 1$, then the flow on path $c \rightarrow a$ in II is zero and II reduces to the simple unit gain cycle C.

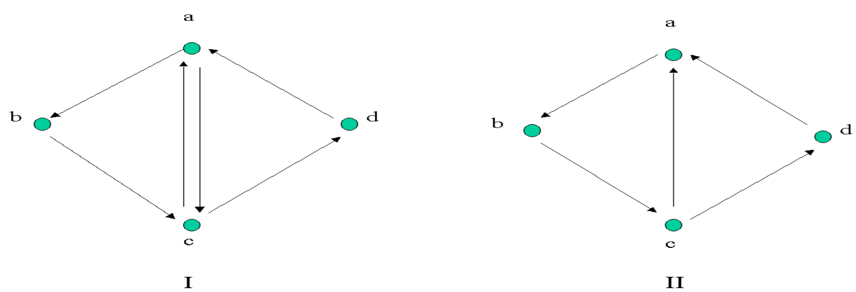

Figure 1. Flow reduction in a degenerate bicycle

Suppose now that $\pi_u^l < \pi_u^h$ for all nodes u. The only direction that we have not considered is a zero-cost unit gain cycle. Without loss of generality, we assume that G is connected. Since $0 < x < u$, there is a directed path from any node u to any other node v in G_x. Thus if there is a flow-generating cycle in G_x, then $\pi_u^l \neq -\infty$ for each node u since there is at least one path from that flow-generating cycle to every node. Similarly, if there is a flow-absorbing cycle in G_x, then $\pi_u^h \neq \infty$ for each node u, since there is at least one path from any node to that flow-absorbing cycle. At this point in our step direction search it is not possible for $\pi_u low = -\infty$ and $\pi_u high = \infty$ for all nodes u; thus there is at least one flow-generating or one flow-absorbing cycle in G_x.

Let us assume that there is at least one flow-generating cycle in G_x and hence that $\pi_u^l > -\infty$ for all u. (The case when there is at least one flow-absorbing cycle in G_x is similar.) Under this assumption, it is easy to see that there is a feasible solution to the system $Ax = b, x \geq 0$ for any $b > 0$, where A is the constraint matrix in problem (2). Thus matrix A is a Leontief matrix and it follows from results in Adler and Cosares ([1]) that there is a maximal solution to the system $A^\top y \leq \hat{c}$ if the system is feasible. Moreover, since $\pi = -y$ (the constraint matrix of problem (3) is $-A^\top$) and $\pi_u^l < \pi_u^h$ for each variable, this system is feasible and $\pi^l = (\pi_1^l, \ldots, \pi_n^l)^\top$ is a feasible solution to the TVPI system $S = \{\pi : \pi_u - \gamma_{uv}\pi_v \leq \hat{c}_{uv}, (u, v) \in E_x\}$ that we are considering.

Let $\bar{E} = \{(u, v) \in E_x : \hat{c}_{uv} - \pi_u^l + \gamma_{uv}\pi_v^l = 0\}$. Since π^l is a feasible solution to the TVPI system S, $\pi_u^l - \gamma_{uv}\pi_v^l < \hat{c}_{uv}$, for arcs $(u, v) \notin \bar{E}$. If there is a unit gain cycle W in $\bar{E}$ passing through node v_1, then

$$\pi_1^l = \hat{c}(W) + \gamma(W)\pi_1^l \tag{5}$$

where $\hat{c}(W)$ is the cost of sending one unit of flow around W starting from node v_1. Since $\gamma(W) = 1$, we must have $\hat{c}(W) = 0$. Thus any unit gain cycle in $\bar{E}$ has cost 0.

Conversely, if there is a unit gain cycle W whose cost $\hat{c}(W)$ starting from node v_1 equals 0, it must be in $\bar{E}$. If not, then for some arc (u, v) on W we have $\pi_u^l - \gamma_{uv}\pi_v^l < \hat{c}_{uv}$, which implies that $\pi_1^l < \hat{c}(W) + \gamma(W)\pi_1^l$, and hence, since $\gamma(W) = 1$, that $\hat{c}(W) > 0$ – a contradiction.

Therefore, in order to find a unit gain cycle with zero cost we only need to find a unit gain cycle in $\bar{E}$. Note that there is no flow-absorbing cycle in $\bar{E}$, since each directed cycle W in $\bar{E}$ starting from node u satisfies (5), and consequently, if $\gamma(W) < 1$, then $\pi_u^l = c(W)/(1 - \gamma(W)) \geq \pi_u high \geq \pi_u^h$. But this contradicts $\pi_u^l < \pi_u^h$. Thus any cycle in $\bar{E}$ has gain at least 1, and hence, a nonnegative length with respect to arc lengths $d(e) = \log \gamma(e)$. Therefore, we can use a shortest path algorithm to find a unit gain cycle in $\bar{E}$. Specifically, after computing a shortest path tree T and reduced costs $\bar{d}(e)$, i.e., $\bar{d}(e) = 0$ for all $e \in T$, if there exists any $e = (u, v) \notin T$ with $\bar{d}(e) = 0$ and v is on the path from the root of tree T to u, then we have a zero cost unit gain cycle. Since a shortest path algorithm can be run in $O(mn)$ time, we can check if there exists any zero-cost unit gain cycles in $O(mn)$ time. Such a cycle yields a suitable step direction for Algorithm GN.

Thus the overall algorithm for finding a suitable direction or deciding that none exists is:

Algorithm SD(*Step Direction*):
call a subroutine that computes $\pi_u low$ and $\pi_u high$ and gives their corresponding cycles for each node u;
if during the subroutine an infeasible cycle is found, **then** it is a suitable direction, STOP;
end if;
if $\pi_u low = -\infty$ and $\pi_u high = \infty$ for every node u, **then** check if there is a zero cost unit-gain cycle using a shortest path routine;
  if such a cycle exists, **then** it is a suitable direction, STOP;
  else no suitable direction exists, STOP;
  end if;
end if;
if $\pi_u low \geq \pi_u high$ for some node u, **then** the nonpositive cost degenerate bicycle formed from the cycles with common node u associated with $\pi_u low$ and $\pi_u high$ is a suitable direction, STOP;
else $\pi_u low < \pi_u high$ for each node u, call Algorithm BF;
  if Algorithm BF finds a nonpositive cost bicycle, **then** it is a suitable direction, STOP;
  else
    if $\pi_u^l > -\infty$ for all $u \in V$, **then**
      set $d(e) = \log \gamma(e)$ for all $e \in \bar{E} \equiv \{(u, v) \in E : \hat{c}_{uv} - \pi_u^l + \gamma_{uv}\pi_v^l = 0\}$;
      call a shortest path subroutine to check if there is a zero-length directed cycle in $\bar{E}$ with respect to the arc lengths $d(e)$;
      if there is such a zero cost unit-gain cycle, **then** it is a suitable

direction, STOP;
end if;
end if;
if $\pi_u^h < \infty$ for all $u \in V$, **then**
set $d(e) = -\log \gamma(e)$ for all $e \in \bar{E} \equiv \{(u,v) \in E : \hat{c}_{uv} - \pi_u^h + \gamma_{uv}\pi_v^h = 0\}$;
call a shortest path subroutine to check if there is a zero-length directed cycle in $\bar{E}$ with respect to the arc lengths;
if there is such a zero cost unit-gain cycle, **then** it is a suitable direction, STOP;
end if;
end if;
otherwise, no suitable direction exists, STOP;
end if;
end if.

Given $\pi_u low$ and $\pi_u high$ for each variable, the above algorithm finds a suitable direction or decides that none exists in $O(mn^2)$ time. The bottleneck is still finding πlow and $\pi high$, which can be done in $O(mn^2(\log m + \log^2 n))$ time using the algorithm of Cohen and Megiddo ([3]).

5. Generalized circulation problem

The generalized circulation problem is a special case of the generalized MCNF problem and hence can be solved by any algorithm for the latter problem. In this section, we show that a slightly modified version of our algorithm can solve the generalized circulation problem directly without first converting the problem into a generalized MCNF problem.

Let node s be the specified sink. The generalized circulation problem can be formulated as:

$$\max \sum_{i \in B(s)} \gamma_i x_i - \sum_{i \in F(s)} x_i$$

$$s.t. \sum_{i \in B(v)} \gamma_i x_i - \sum_{i \in F(v)} x_i = 0, v \in V \backslash \{s\}$$

$$0 \le x_i \le u_i, \quad i \in E.$$

As with the generalized MCNF problem, the generalized circulation problem can be written as a standard form linear program (1) in variables $z^\top = (x^\top, y^\top)$, where $A = \begin{pmatrix} \bar{A} & 0 \\ I & I \end{pmatrix}$, $b = \begin{pmatrix} 0 \\ u \end{pmatrix}$, $c = \begin{pmatrix} \bar{c} \\ 0 \end{pmatrix}$ and $\bar{c}_i = \begin{cases} -\gamma_i, & i \in B(s) \\ 1 & i \in F(s) \\ 0 & \text{otherwise.} \end{cases}$ Note that $\bar{A}$ and $\bar{c}$ are different from what these quantities are in the generalized MCNF problem.

Let h be a search direction that satisfies $ADh = 0$. Then flow $g(Dh)$ can be decomposed into type 2,3,4 and 5 flows. Since type 3 flows can only decrease the objective value, if z_{opt} is an optimal solution, there exists a search direction h such that $z_{opt} = z + Dh$ and $g(Dh)$ can be decomposed into type 2,4 and 5 flows. Furthermore, if $g(\Delta z^i)$, $1 \leq i \leq r$ are the elementary generalized flows in the decomposition, then $Dh = \sum_{i=1}^{r} \mu_i \Delta z^i$, $\mu_i \geq 0$, $1 \leq r \leq 2m$, and

$$(Dh)_i \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} 0 \Longrightarrow \Delta z_i^{\nu} \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} 0, \quad 1 \leq i \leq 2m, 1 \leq \nu \leq r.$$

Since this is the only property required in the proofs of the results in Section 3, all of these results can be extended to the generalized circulation case.

The differences between our algorithm for the generalized circulation problem and the one already described for the generalized MCNF problem are:

(1) A suitable direction for the generalized circulation problem is either a type 2,4 or 5 elementary flow. Each of the type 2 flow (GAP) ends in the sink node. Therefore, for this problem, the set $\tilde{C}_x$ is composed of type 2 flows as well as type 4 (unit-gain cycle) and type 5 (bicycle) flows.
(2) When trying to find a nonpositive cost generalized circulation in Algorithm GN, the arc costs defined may be different, since the original arc costs are different. However, this does not affect the complexity of the algorithm.
(3) For the generalized circulation problem, besides the termination condition mentioned in Algorithm SD, the algorithm can be stopped in the following three cases:
   (i) if $\pi_u low = -\infty$ and $\pi_u high = \infty$ for every node u, then flow $g = 0$ is optimal;
   (ii) If $\pi_s low \geq 0$ (s is the sink node), then the flow-generating cycle that gives the inequality $\pi_s \geq \pi_s low$ is a nonpositive cost degenerate GAP, which provides a suitable direction;
   (iii) In Algorithm BF, if $\pi_s^l \geq 0$, then the GAP that gives π_s^l provides a suitable direction.

6. An initial solution

In order to use Algorithm GN, we need to provide it with an initial interior feasible solution. We can find such a solution using Algorithm GN with the following modifications. This is similar to Wallacher and Zimmermann's approach of finding an initial solution.

For the generalized minimum cost network flow problem, let $x \in R^n$ satisfy $0 < x < u$ and let $e(i)$ denote the imbalance at node i with respect to x for $1 \leq i \leq m$. Extend graph $G(V, E)$ to $G^e = (V^e, E^e)$ with

$$V^e = V \cup \{v, w\}$$

and

$$E^e = E \cup \{(i, v) | e(i) - b_i > 0\} \cup \{(w, i) | e(i) - b_i < 0\} \cup \{(v, w)\}.$$

For each arc $i \in E^e \setminus (E \cup \{(v,w)\})$, set $l_i = 0, u_i = \sum_{j \in E} \gamma_j u_j$ and $\gamma_i = 1$. Let $M = \sum_{i:e(i)>b_i}(e(i) - b_i)$ and $N = \sum_{i:e(i)<b_i}(b_i - e(i))$. For arc (v,w), set $l_{vw} = 0$, $u_{vw} = \sum_{j \in E} \gamma_j u_j$ and $\gamma_{vw} = N/M$. Set $b_v = b_w = 0$. x can be extended to an interior feasible solution x^e with respect to the problem in G^e if we set

$$x^e_{iv} = e(i) - b_i, \quad x^e_{wi} = b_i - e(i) \text{ and } \quad x^e_{vw} = \sum_{i \in V} x^e_{iv}.$$

From this construction, we know that a feasible solution exists in G if and only if there exists a feasible flow x^e in G^e with $x^e_{vw} = 0$. Introducing costs $\bar{c}$, where $\bar{c}_{vw} = 1$ and $\bar{c}_{ij} = 0$ for $ij \neq vw$, there exists a feasible solution in G if and only if the optimal value of the generalized minimum cost network flow problem for the graph G^e with respect to costs $\bar{c}$ is zero.

Now we can apply Algorithm GN to graph G^e with respect to costs $\bar{c}$ using x^e as an initial interior feasible solution. We can use $\alpha = 0$ as a lower bound for the optimal value. The algorithm stops when either a feasible flow in G does not exist (i.e., when α needs to be updated during the algorithm), or with an initial solution to the original problem.

For the generalized circulation problem, the only difference is that G^e does not include the sink s, the arcs incident to it and $b_i = 0$ for all $i \in V \backslash \{s\}$, i.e.,

$$E^e = E \cup \{(i,v)|e(i) > 0, i \neq s\} \cup \{(w,i)|e(i) < 0, i \neq s\} \cup \{(v,w)\}.$$

All other arguments are the same.

7. Comparison with other algorithms

Our algorithms are combinatorial interior point methods. They are combinatorial algorithms since at each iteration, a suitable generalized circulation in the current residual graph is chosen and flow is augmented along it. They are interior point methods since all the solutions found in our algorithms are interior points, and the algorithms achieve optimality by iteratively reducing the value of a potential function.

We were able to extend the results of Wallacher and Zimmermann relating to their algorithm for the minimum cost flow problem to our algorithms for both the generalized minimum cost flow and generalized circulation problems by noting that all of their lemmas and theorems can be extended to the generalized flow case because of the analogy between the decomposition theorems for ordinary and generalized flows. The main difference between their algorithm and our algorithms is that at each iteration, their algorithm tries to find a nonpositive cost cycle, while in our algorithms, we try to find a nonpositive cost generalized circulation.

Wayne ([12]) developed a combinatorial algorithm for the generalized minimum cost flow problem that is based on Wallacher's ([10]) combinatorial algorithm for the ordinary minimum cost flow problem. As in our algorithms, the direction finding step in Wayne's algorithm seeks a negative cost generalized circulation in the current residual network. However, his algorithm is not an interior point method, and it always tries to find a descent direction with respect to the original objective function while our algorithms try to find a descent direction with respect to a potential function. Hence the arc costs defined for the direction finding problem in the two algorithms are different. Wayne's algorithm runs in $\tilde{O}(m^3 n^2 \log B)$ time, where B is the largest integer needed to represent the integral costs and capacities and rational gain factors. Our algorithms run in $O(m^2 n^2 (\log m + \log^2 n)L)$ time. Since $L = O(m \log B)$ and $\tilde{O}(f)$ represents $f \log^{O(1)} m$, the algorithms have the same worst-case running time.

Note that Theorem 5 still holds if we relax the strict inequality in the condition $\frac{a_\alpha^\top \tilde{h}}{\|\tilde{h}\|_1} > -s$ to be nonstrict. By doing so, the step direction problem is reduced to finding a negative cost generalized circulation in the current residual graph and can be solved directly by any algorithm that either finds a feasible solution to a TVPI system or gives a certificate of infeasibility. In this case, Hochbaum and Naor's algorithm ([7]) can be used instead of Cohen and Meggido's algorithm to achieve a slightly better overall complexity of $O(m^2 n^2 \log m \cdot L)$. Moreover, both of these algorithms can be combined with Wallacher and Zimmermann's algorithm to solve the general problem of optimizing a TVPI system. This problem is the dual of a "generalized" minimum cost flow problem that allows negative as well as positive gains. Such an approach uses the fact that the decomposition theorem (Theorem 2) can be extended to such "generalized" networks with negative gains. Specifically, if an additional type 6 elementary "generalized" flow is defined that consists of two type 2 elementary flows and a negative gain arc connecting the two nodes with excesses, then a "generalized" circulation that has negative gain arcs can be decomposed into types 4,5 and 6 flows. This is because removing negative gain arcs in such a "generalized" circulation creates an ordinary generalized flow (i.e., a flow with only positive gains) with nonnegative imbalance at each node. By Theorem 2, such a flow can be decomposed into types 2,4 and 5 flows and adding back negative gain arcs combines type 2 flows into type 6 "generalized" flows. Wayne ([12]) gives a detailed definition of residual networks for this case.

Wayne ([12]) also provides a combinatorial method for rounding a B^{-7m}-approximate optimal solution to an optimal solution in $\tilde{O}(m^2 n)$ time. This method, with obvious changes in the generalized circulation case, can be used to obtain an optimal solution once our algorithms terminate.

Acknowledgements. The authors are thankful to an anonymous referee for several helpful suggestions for improving the presentation of the paper and for pointing out the polylogarithmic improvement in complexity obtained by using Hochbaum and Naor's TVPI algorithm.

References

1. I. Adler and S. Cosares, A strongly polynomial algorithm for a special class of linear programs, *Operations Research* **39,** (1991) 955-960.

2. B. Aspvall and Y. Shiloach, A polynomial time algorithm for solving systems of linear inequalities with two variables per inequality, *SIAM J. Comput.* **9,** (1980) 827-845.
3. E. Cohen and N. Megiddo, Improved algorithms for linear inequalities with two variables per inequalities, *SIAM J. Comput.* **23,** (1994) 1313-1347.
4. A. Goldberg, S. Plotkin and E. Tardos, Combinatorial algorithms for the generalized circulation problem, *Mathematics of Operations Research*, **16,** (1991) 351-379.
5. C. Gonzaga, Polynomial affine algorithms for linear programming, *Mathematical Programming* **49,** (1990) 7-21.
6. C. Gonzaga, Large step path-following methods for linear programming II: Potential reduction method, *SIAM J. Optim.* **1,** (1991) 280-292.
7. D. Hochbaum and J. Naor, Simple and fast algorithms for linear and integer programs with two variables per inequality, *SIAM J. Comput.* **23,** (1994) 1179-1192.
8. N. Megiddo, Towards a genuinely polynomial algorithm for linear programming, *SIAM J. Comput.* **12,** (1983) 347-353.
9. R. Shostak, Deciding linear inequalities by computing loop residues, *Assoc. Comput. Mach.* **28,** (1981) 769-779.
10. C. Wallacher, A generalization of the minimum-mean cycle selection rule in cycle canceling algorithms, Report. Inst. für Ang. Math, Braunschweig, Germany, 1991.
11. C. Wallacher and U. Zimmermann, A combinatorial interior point method for network flow problems, *Mathematical Programming* **56,** (1992) 321-335.
12. K. Wayne, A polynomial combinatorial algorithm for generalized minimum cost flow, *ACM Symposium on Theory of Computing* (1999) 11-18.