

## SCALING METHODS FOR FINDING A MAXIMUM FREE MULTIFLOW OF MINIMUM COST

ANDREW V. GOLDBERG AND ALEXANDER V. KARZANOV

Suppose we are given an undirected graph with nonnegative integer-valued edge capacities and costs in which a subset of nodes is specified. We consider the problem of finding a collection of flows between arbitrary pairs of specified nodes such that the capacity constraints are satisfied and the sum of costs of flows is minimum, provided that the sum of values of flows is maximum. It is known that this problem has a *half-integer* optimal solution and such a solution can be found in strongly polynomial time using the ellipsoid method.

In this paper we give two "purely combinatorial" polynomial algorithms for finding a half-integer optimal solution. These are based on capacity and cost scaling techniques and use the double covering method earlier worked out for the problem.

**1. Introduction.** By a *graph* we mean a finite undirected graph with possible multiple edges;  $VG$  and  $EG$  are the sets of nodes and edges of a graph  $G$ , respectively. A *network* is a quadruple  $N = (G, T, c, a)$  consisting of a graph  $G$ , a subset  $T \subseteq VG$  of its nodes, called *terminals*, and nonnegative integer-valued functions  $c$  (of *capacities*) and  $a$  (of *costs*) on the edges of  $G$ .

A simple path in  $G$  between two distinct terminals is called a *T-path*; let  $\mathcal{P} = \mathcal{P}(G, T)$  denote the set of *T-paths*. A ( $c$ -*admissible*) *multicommodity flow*, or, briefly, a *multiflow*, in  $N$  is a nonnegative rational-valued function  $f: \mathcal{P} \rightarrow \mathbb{Q}_+$  satisfying the capacity constraint

$$\xi^f(e) := \sum (f(P): e \in P \in \mathcal{P}) \leq c(e) \quad \text{for all } e \in EG$$

(when writing  $e \in P$ , we consider a path as an edge-set). The *total value*,  $v_f$ , of  $f$  is  $\sum (f(P): P \in \mathcal{P})$ , and the *total cost*,  $a_f$ , of  $f$  is  $\sum (a(e)\xi^f(e): e \in EG)$ . A multiflow of the maximum total value is called *maximum*. We deal with the following *minimum cost maximum multiflow* problem:

(1.1) Find a maximum multiflow  $f$  in  $N$  whose total cost  $a_f$  is as small as possible.

The simplest case arises when  $|T| = 2$ . Then (1.1) turns into the well-known (undirected) minimum cost maximum *flow* problem. A fundamental fact is that the latter problem has an optimal solution which is integer-valued; see, e.g., Ford and Fulkerson (1962). This is, in general, not true for  $|T| \geq 3$ . Nevertheless, a weaker property of *half-integrality* has been proved, as follows.

**THEOREM 1 (KARZANOV 1979).** *There exists an optimal solution  $f$  to (1.1) such that  $2f$  is integer-valued.*

**REMARK 1.** Problem (1.1) is a special case of the general undirected minimum cost maximum multiflow problem in which pairs  $\{s_1, t_1\}, \dots, \{s_r, t_r\}$  of nodes are

specified, and the goal is to connect these pairs by flows so that the resulting multiflow is  $c$ -admissible and has the minimum cost provided that its total value is maximum. The special case we study allows flows between arbitrary pairs of terminals; in other words, the graph  $H$  induced by the pairs  $\{s_1, t_1\}, \dots, \{s_r, t_r\}$  as edges is a complete graph. Following Lomonosov (1985), we call a multiflow for a complete  $H$  a *free* multiflow. The relationship between the denominators of multiflow entries and the structure of  $H$  is studied in Karzanov (1987). It turned out that if  $H$  is not a complete multi-partite graph, then for each  $k \in \mathbb{Z}_+$ , there exists an instance of  $G, c, a$  such that  $kf$  is not integral for any optimal solution  $f$  for  $G, H, c, a$ . On the other hand, when  $H$  is a complete multi-partite graph, the problem is easily reduced to the free multiflow case, whence the problem has a half-integer optimal solution by Theorem 1.

REMARK 2. The existence of a half-integer maximum free multiflow was established, independently, by Lovász (1976) and Cherkassky (1977). Theorem 1 extends that result to the cost case.

Theorem 1 appeared as a consequence of a similar result concerning a more general *parametric problem*, namely:

(1.2) Given  $p \in \mathbb{Z}_+$ , find a multiflow  $f$  in  $N$  which maximizes  $pv_f - a_f$ .

THEOREM 2 (KARZANOV 1979). For any  $p \geq 0$ , (1.2) has a half-integer optimal solution.

Obviously, (1.2) becomes equivalent to (1.1) when  $p$  is large enough (from the existence of a half-integer solution to (1.2) it follows that  $p = 2a(EG)c(EG) + 1$  is sufficient). So Theorem 1 immediately follows from Theorem 2. Here and later on for a function  $g$  on a set  $S$  and a subset  $S' \subseteq S$ ,  $g(S')$  denotes  $\sum(g(e) : e \in S')$ .

Theorem 2 was originally proved by constructive means, relying on a pseudo-polynomial algorithm which solves (1.2) in time  $O(\min\{(c(EG) + 1)q, (a(EG) + 1)q', 2^{q''}\})$ , where  $q, q', q''$  are polynomials in  $|VG|, |EG|$ . Further results were obtained in Karzanov (1994) where a relatively simple, nonalgorithmic, proof of Theorem 2 was given and it was shown how to find a half-integer optimal solution to (1.2) (and, therefore, (1.1)) in strongly polynomial time by use of the ellipsoid method. However, no polynomial algorithm which uses solely "combinatorial means" has been known until the present.

Like the classic algorithm for the min-cost max-flow problem due to Ford and Fulkerson (1962), the algorithm in Karzanov (1979) works in frameworks of the primal-dual method of linear programming and iteratively reduces (1.2) to simpler, noncost, subproblems on multiflows in "feasible subgraphs"  $\Gamma$  of  $G$ , which are solved by use of combinatorial techniques. The key idea is that the subproblem allows a reduction to a certain single-commodity flow problem in a skew-symmetric digraph  $D$  such that each edge of  $\Gamma$  corresponds to a pair of "skew-symmetric arcs" in  $D$ . Such a  $D$  is called the *double covering* over  $\Gamma$ . This explains the existence of a half-integer optimal solution to (1.2): The flow problem in  $D$  has an integer solution  $g$  and the back reduction increases the solution "fractionality" twice, transforming  $g$  into a half-integer multiflow in  $\Gamma$ .

In this paper we design two "purely combinatorial" polynomial algorithms for finding a half-integer optimal solution to (1.2). The first algorithm works with an arbitrary  $p \in \mathbb{Z}_+$ , while the second algorithm works when  $p$  is large; in particular, both algorithms solve (1.1). At the high level, instead of the primal-dual approach, they apply variants of the scaling method: the first algorithm scales capacities (cf. Edmonds and Karp 1972) and runs in  $O(\log(c(EG) + 2))$  iterations, while the second one scales costs (cf. Röck 1980; Bland and Jensen 1992) and runs in  $O(\log(a(EG) +$

$2) + \log(c(Z) + 2))$  iterations. Here  $Z$  is the set of edges  $e$  with  $a(e) = 0$ . An iteration of each algorithm deals with a subproblem which is somewhat different from, but closely related to, that of Karzanov (1979). We show that each subproblem is reduced to a sequence of directed flow problems and that the total time required to perform an iteration is polynomial in  $|VG|, |EG|$ .

It should be noted that the original version of this paper (Goldberg and Karzanov 1993) presented polynomial time capacity and cost scaling algorithms based on similar ideas, but was technically more complicated, using the concept of so-called transitive fork environments—a formalism which allows us to extend standard augmenting and alternating path techniques to rather general situations. This formalism was, in fact, introduced in Karzanov (1993) for solving the integer strengthening of (1.1).

For technical reasons, throughout the paper we assume that the cost function  $a$  is *positive*, i.e.,  $a(e) > 0$  for all  $e \in EG$ . This assumption does not lead to loss of generality. Indeed, if the set  $Z = \{e \in EG: a(e) = 0\}$  is nonempty, we can replace  $p$  by  $p' = (2c(Z) + 1)p$ , and  $a$  by  $a'$  defined by

$$\begin{aligned} a'(e) &= (2c(Z) + 1)a(e) \quad \text{for } e \in EG - Z, \\ &= 1 \quad \text{for } e \in Z. \end{aligned}$$

Then for any two *half-integer* maximum multiflows  $f$  and  $f'$  with  $\Delta = (pv_f - a_f) - (pv_{f'} - a_{f'}) > 0$ , we have  $\Delta \geq 1/2$  and

$$\begin{aligned} &(p'v_f - a'_f) - (p'v_{f'} - a'_{f'}) \\ &= (2c(Z) + 1)((pv_f - a_f) - (pv_{f'} - a_{f'})) - \zeta^f(Z) + \zeta^{f'}(Z) \\ &\geq (2c(Z) + 1)\Delta - c(Z) \geq c(Z) + \frac{1}{2} - c(Z) > 0. \end{aligned}$$

Therefore, if  $f$  is an optimal solution for  $p', a'$ , then  $f$  is also an optimal solution for  $p, a$ . Note that a factor of  $c(Z) + 1$  in the definition of  $a'$  causes the rising of the term involving  $c$  in the complexity of the cost scaling algorithm declared above.

Although we allow multiple edges in  $G$ , when it is not confusing, we denote an edge with end nodes  $u$  and  $v$  by  $uv$ .

This paper is organized as follows. Section 2 contains background for the algorithms that we develop. It considers the dual of (1.2), describes the construction of double covering digraphs  $D$ , and briefly reviews the relationship between multiflows in  $G$  and flows in  $D$ . Sections 3–6 are devoted to the capacity scaling algorithm. Section 7 contains the cost scaling algorithm; its description is shorter because both algorithms utilize many common tools.

**2. Double covering.** The linear program dual of (1.2) is

$$(2.1) \quad \text{Minimize } c \cdot \gamma \text{ over all } \gamma: EG \rightarrow \mathbb{Q}_+ \text{ such that } \gamma(P) \geq p - a(P) \text{ for each } P \in \mathcal{P},$$

viewing  $P$  as an edge-set and denoting by  $g \cdot h$  the inner product  $\sum(g(e)h(e): e \in S)$  of functions  $g$  and  $h$  on  $S$ . For  $l \in \mathbb{Q}_+^{EG}$ , let  $\text{dist}_l(x, y)$  denote the  $l$ -distance between nodes  $x, y \in VG$ , i.e., the minimum  $l$ -length  $l(P)$  of a path  $P$  from  $x$  to  $y$ . The constraints in (2.1) can be expressed in a more compact form as

$$(2.2) \quad \text{dist}_{a+\gamma}(s, t) \geq p \quad \text{for any } s, t \in T, s \neq t.$$

The linear programming duality theorem applied to (1.2) and (2.1) implies that a multiflow  $f$  and a vector  $\gamma \in \mathbb{Q}_+^{EG}$  satisfying (2.2) are optimal solutions to these problems if and only if the following "complementary slackness" conditions hold:

$$(2.3) \text{ For } P \in \mathcal{P}, \text{ if } f(P) > 0 \text{ then } a(P) + \gamma(P) = p;$$

$$(2.4) \text{ For } e \in EG, \text{ if } \gamma(e) > 0 \text{ then } e \text{ is saturated by } f, \text{ i.e., } \zeta^f(e) = c(e).$$

In the rest of this section we consider a positive function  $l$  on  $EG$  which satisfies  $\text{dist}_l(s, t) \geq p$  for all distinct  $s, t \in T$  (usually  $l$  is equal to  $a + \gamma$ ). For brevity, we denote  $\text{dist}_l(\cdot, \cdot)$  by  $\text{dist}(\cdot, \cdot)$  and omit the prefix  $l$  in the terms  $l$ -length,  $l$ -distance and  $l$ -shortest. A  $T$ -path of the length exactly  $p$  is called a  $(T, l)$ -line, or, briefly, a  $T$ -line. A part of a  $T$ -line is called a line. By (2.2), every line  $P$  is a shortest path; so  $P$  is simple (as  $l$  is positive). The structure of lines has important properties described below; for detailed proofs, see Karzanov (1979, 1994).

We need some terminology and notations. The subgraph of  $G$  formed by  $T$  and the nodes and edges contained in lines is called the  $l$ -graph and denoted by  $\Gamma = \Gamma^l$ . For  $v \in V\Gamma$ , the potential,  $\pi(v) = \pi_l(v)$ , of  $v$  is the distance from  $v$  to  $T$  (i.e.,  $\min\{\text{dist}(v, s) : s \in T\}$ ), and  $T(v) = T_l(v)$  denotes the set of terminals  $s \in T$  closest to  $v$  (i.e.,  $\text{dist}(v, s) = \pi(v)$ ).

Obviously,  $\pi(v) \leq p/2$  for any  $v \in V\Gamma$ ; moreover,  $\pi(v) = p/2$  implies  $|T(v)| \geq 2$ , and  $\pi(v) < p/2$  implies  $|T(v)| = 1$ . Therefore, the nodes of  $\Gamma$  are naturally partitioned into the sets  $V_s = \{v \in V\Gamma : T(v) = \{s\}\}$  for  $s \in T$  and  $V^\bullet = \{v \in V\Gamma : |T(v)| \geq 2\}$ ; the nodes in  $V^\bullet$  are called central. One can obtain the following characterization of the lines (in particular, the edges of  $\Gamma$ ) in terms of potentials:

(2.5) A path  $P = (v_0, e_1, v_1, \dots, e_k, v_k)$  in  $G$  with both ends in  $\Gamma$  is a line if and only if either (i)  $v_0, v_k \in V_s \cup V^\bullet$  and  $|\pi(v_k) - \pi(v_0)| = l(P)$  for some  $s \in T$ , or (ii)  $v_0 \in V_s, v_k \in V_t$  and  $\pi(v_0) + \pi(v_k) + l(P) = p$  for some distinct  $s, t \in T$ .

In particular, in view of the positivity of  $l$ ,  $\pi(x) \neq \pi(y)$  for any edge  $xy \in E\Gamma$  with  $x, y \in V_s$ , and no edge of  $\Gamma$  connects two central nodes. Property (2.5) easily implies that

(2.6) A  $T$ -path  $P = (v_0 = s, e_1, v_1, \dots, e_k, v_k = t)$  in  $\Gamma$  is a  $T$ -line if and only if there is  $0 \leq i < k$  such that  $v_0, \dots, v_i \in V_s; v_{i+2}, \dots, v_k \in V_t; \pi(v_0) < \dots < \pi(v_i); \pi(v_{i+2}) > \dots > \pi(v_k)$ ; and either  $v_{i+1} \in V^\bullet$ , or  $v_{i+1} \in V_t$  and  $\pi(v_{i+1}) > \pi(v_{i+2})$ .

Using (2.5) for one-edge paths (edges), we now construct the digraph  $D^l = D = (VD, AD)$ , the double covering over  $\Gamma$ , mentioned in the Introduction, as follows. Split each  $v \in V\Gamma$  into  $2|T(v)|$  nodes  $v_s^1$  and  $v_s^2$  ( $s \in T(v)$ ). If  $v$  is noncentral, i.e.,  $T(v)$  consists of a single terminal  $s$ , then  $v_s^1$  and  $v_s^2$  are called the first and second copies of  $v$  in  $D$ , respectively. The arcs of  $D$  are defined as follows:

- (2.7) (i) Each edge  $e = uv \in E\Gamma$  with  $u \in V_s, v \in V_s \cup V^\bullet$  and  $\pi(u) < \pi(v)$  generates two arcs  $(u_s^1, v_s^1)$  and  $(v_s^2, u_s^2)$ ;  
(ii) Each edge  $e = uv \in E\Gamma$  with  $u \in V_s$  and  $v \in V_t$  ( $s \neq t$ ) generates two arcs  $(u_s^1, v_t^2)$  and  $(v_t^1, u_s^2)$ ;  
(iii) Each central node  $v$  generates arcs  $(v_s^1, v_t^2)$  for all distinct  $s, t \in T(v)$ .

An illustration is given in Figure 1 where  $T = \{s, t, q\}$ ,  $p = 4$ , the numbers on edges indicate values of  $l$ , and the arcs of  $D$  are directed upward. An arc as in (2.7)(iii) is

called *central*, or a *v-arc*, and has infinite capacity. An arc as in (2.7)(i) or (ii) has the same capacity as that of the edge which generates this arc. We keep the same notation  $c$  for the capacities in  $D$  and think of  $T^1 = \{s^1: s \in T\}$  and  $T^2 = \{s^2: s \in T\}$  as the sets of *sources* and *sinks* of  $D$ , respectively.

Since for each  $b = (u_s^i, v_t^j) \in AD$ , the pair  $b' = (v_t^{3-j}, u_s^{3-i})$  is also an arc of  $D$ , the mapping  $\sigma: v_s^i \rightarrow v_t^{3-i}$  gives a (skew) *symmetry* of  $D$ ; we extend  $\sigma$  to the arcs denoting  $b'$  by  $\sigma(b)$ . Next, the construction of  $D$  yields a natural mapping  $\omega$  of  $VD \cup AD$  to  $V\Gamma \cup E\Gamma$  which brings a node  $v_s^i$  to  $v$ , a noncentral arc  $(u_s^i, v_t^j)$  to the edge  $uv$ , and a central arc  $(v_s^1, v_t^2)$  to the node  $v$ . We extend  $\sigma$  and  $\omega$  in a natural way to the dipaths and other objects in  $D$ . For example, the dipath  $\sigma(P)$  symmetric to a dipath  $P = (x_0, b_1, x_1, \dots, b_k, x_k)$  is  $(\sigma(x_k), \sigma(b_k), \sigma(x_{k-1}), \dots, \sigma(b_1), \sigma(x_0))$ ; the path  $\omega(P)$  for this  $P$  is the sequence  $(\omega(x_0), \omega(b_1), \omega(x_1), \dots, \omega(b_k), \omega(x_k))$  with the repeated central nodes (if any) deleted. A function  $h$  on  $AD$  is *symmetric* if  $h(b) = h(\sigma(b))$  for each  $b \in AD$ .

Important properties of  $D$ , easily derived from (2.6) and (2.7), are as follows:

(2.8) Dipaths  $P$  and  $\sigma(P)$  are disjoint, and  $\omega(\sigma(P))$  is reverse to  $\omega(P)$ ;

(2.9)  $\omega$  yields a *one-to-one correspondence* between the set of  $T$ -lines and the set of  $T^1$  to  $T^2$  dipaths in  $D$ .

Such a correspondence is further extended to a relationship between certain multiflows in  $G$  and flows in  $D$ , as follows. A multiflow  $f$  is called *going along  $T$ -lines* if each  $T$ -path  $P$  with  $f(P) > 0$  is a  $T$ -line. A ( $c$ -admissible  $T^1$  to  $T^2$ ) *flow* in  $D$  is a function  $g: AD \rightarrow \mathbb{Q}_+$  satisfying the conservation condition

$$\text{div}_g(x) := \sum_{y: (x,y) \in AD} g(x,y) - \sum_{y: (y,x) \in AD} g(y,x) = 0$$

for all  $x \in VD - (T^1 \cup T^2)$ ,

and the capacity constraint  $g(b) \leq c(b)$  for all  $b \in AD$ . Note that  $D$  is acyclic, therefore, a flow  $g$  can be represented as the sum of elementary flows along dipaths. That is, there are  $T^1$  to  $T^2$  dipaths  $P_1, \dots, P_m$  and rationals  $\alpha_1, \dots, \alpha_m \geq 0$  such that  $\sum(\alpha_i; b \in P_i) = g(b)$  for each  $b \in AD$ ; we call  $\mathcal{D} = \{(P_i, \alpha_i)\}$  a *decomposition* of

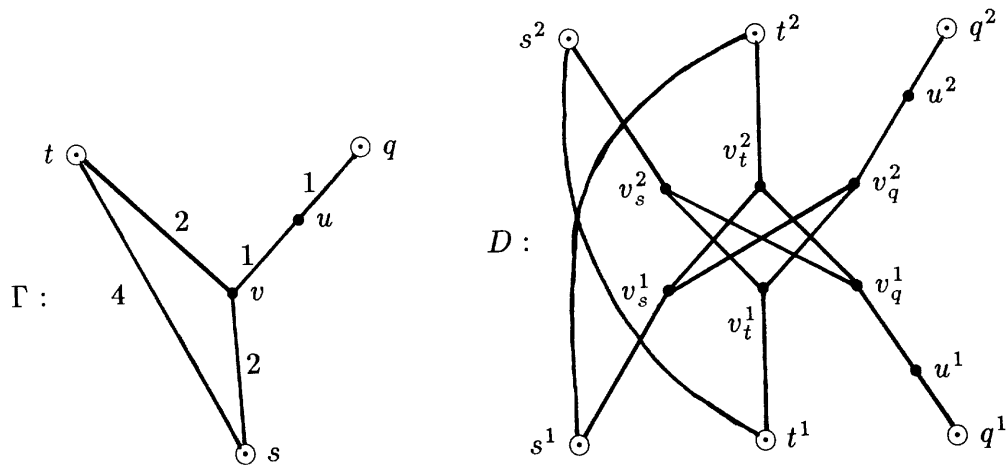


FIGURE 1.

$g$ . Such a  $\mathcal{D}$  induces the function  $f = f^{\mathcal{D}}$  on  $\mathcal{P}$  by setting  $f(\omega(P_i)) = \alpha_i/2$  for  $i = 1, \dots, m$  and  $f(P) = 0$  for the remaining  $T$ -paths  $P$ . For any noncentral arc  $b \in AD$ , we have

$$(2.10) \quad \zeta^f(\omega(b)) = \frac{1}{2}(g(b) + g(\sigma(b))) \leq \frac{1}{2}(c(b) + c(\sigma(b))) = c(\omega(b)).$$

Hence,  $f$  is  $c$ -admissible. Also  $f$  goes along  $T$ -lines (by (2.9)). Furthermore, if  $g$  is integer-valued, then it has an *integer* decomposition  $\mathcal{D}$  (i.e., with all  $\alpha_i$ 's integral), whence  $f^{\mathcal{D}}$  is *half-integral* (this is the crucial observation in the proof of Theorem 2 in Karzanov 1979, 1994).

Conversely, let  $f: \mathcal{P} \rightarrow \mathbb{Q}_+$  be a multiflow going along  $T$ -lines. Define the function  $g = g^f$  on  $AD$  so that for  $b \in AD$ ,  $g(b)$  is the sum of numbers  $f(\omega(P))$  over all  $T^1$  to  $T^2$  dipaths  $P$  in  $D$  that contain  $b$  or  $\sigma(b)$ . Then  $g$  is a symmetric flow, and for each noncentral arc  $b \in AD$ , we have

$$(2.11) \quad g(b) = g(\sigma(b)) = \zeta^f(\omega(b)).$$

Now return to consideration of a function  $\gamma$  satisfying (2.2) and let  $l = a + \gamma$ . It is convenient to extend  $\gamma$  to the arcs  $b$  of  $D^l$  by setting  $\gamma(b) = \gamma(\omega(b))$  if  $b$  is noncentral, and  $\gamma(b) = 0$  otherwise. We say that an arc  $b$  is *feasible* if  $\gamma(b) = 0$ , and a flow  $g$  in  $D$  is *feasible* if it saturates each infeasible arc  $b$ ,  $g(b) = c(b)$ . Summing up arguments above and using (2.10) and (2.11), we derive the following.

STATEMENT 2.1 (KARZANOV 1979). The flow  $g^f$  induced by an optimal solution  $f$  to (1.2) is feasible. Moreover, if, in addition,  $f$  is half-integral, so is  $g$ . Conversely, the multiflow  $f^{\mathcal{D}}$  induced by a feasible flow  $g$  and a decomposition  $\mathcal{D}$  of  $g$  is an optimal solution to (1.2). Moreover, if both  $g$  and  $\mathcal{D}$  are integral, then  $f$  is half-integral.

Our algorithms will also use a result on the fractionality of dual problem (2.1).

THEOREM 2.2 (KARZANOV 1994). (2.1) has a half-integer optimal solution.

**3. Capacity scaling algorithm.** As before, we assume that  $a$  is positive. Also one may assume that  $p > 0$ . At the high level, the algorithm for (1.2) that we now develop applies a standard capacity scaling approach and consists of *big* (or *scaling*) *iterations*. The numbers of these iterations is equal to the number of ones in the binary notation of the capacities.

More precisely, let  $I$  be the maximum of numbers  $\lceil \log_2(c(e) + 1) \rceil$  among all  $e \in EG$ . For  $i = 0, \dots, I$  and  $e \in EG$ , define the truncated capacity  $c^i(e) = \lfloor c(e)/2^{I-i} \rfloor$ . Then  $c^0 = 0$  and  $c^I = c$ . For  $i = 1, \dots, I$  define  $U^i$  to be the set of edges  $e$  with  $c^i(e) > 2c^{i-1}(e)$ . Assuming that the elements of  $U^i$  are arbitrarily ordered as  $z_1, \dots, z_{k(i)}$  ( $k(i) = |U^i|$ ), define the capacities  $c_j^i(e) = c^i(e) - 1$  ( $= 2c^{i-1}(e)$ ) for  $e = z_{j+1}, \dots, z_k$ , and  $c_j^i(e) = c^i(e)$  for the other edges  $e$  in  $G$ ,  $j = 0, \dots, k(i)$ . In particular,  $c_0^i = 2c^{i-1}$ .

The algorithm starts with obvious optimal primal and dual solutions  $f^0$  and  $\gamma^0$  to (1.2) for  $G, T, c^0, a, p$ . More precisely, we put  $f^0 = 0$ , and choose  $\gamma^0$  so as to provide  $\text{dist}_{a+\gamma^0}(s, t) \geq p$  for all distinct  $s, t \in T$  (e.g., for  $e = xy \in EG$ , one can take  $\gamma^0(e)$  to be  $\max(0, p - a(e))$  if  $x, y \in T$ , 0 if  $x, y \in VG - T$ , and  $\max(0, p/2 - a(e))$  otherwise).

The input of a current big iteration (with  $1 \leq i \leq I$  and  $1 \leq j \leq k(i)$ ) consists of optimal primal and dual solutions  $f_{j-1}^i$  and  $\gamma_{j-1}^i$  for  $G, T, c_{j-1}^i, a, p$ , and the goal is to transform them into optimal  $f_j^i$  and  $\gamma_j^i$  for  $c_j^i$  (letting  $f_0^i = 2f^{i-1}$  and  $\gamma_0^i = \gamma^{i-1}$ , where  $f^{i-1}$  and  $\gamma^{i-1}$  are optimal solutions found for  $c^{i-1}$ ). So the last big iteration finds optimal solutions for the original  $c$ .

Thus, the following auxiliary problem has to be solved at most  $|EG|$  times; for convenience we keep the same notation as for the original problem.

(3.1) Given optimal half-integral  $f$  and  $\gamma$  for  $G, T, \hat{c}, a, p$  and an edge  $z \in EG$ , find optimal half-integral  $f'$  and  $\gamma'$  for  $G, T, c, a, p$ , where  $c(z) = \hat{c}(z) + 1$  and  $c(e) = \hat{c}(e)$  for all  $e \in EG - \{z\}$ .

We describe an algorithm to solve (3.1) in strongly polynomial time, thus providing the desired result for (1.2). For reasons which will be clear later, we assume that each pair of distinct terminals is connected by an edge of a large capacity and zero cost; the set of these edges is denoted by  $W$ . The latter assumption does not lead to loss of generality because, obviously, each element of  $W$  must be saturated in any optimal solution for  $\hat{c}$  or  $c$  (in view of  $p > 0$ ). Also we may assume that  $c$  is positive; for if  $c(e) = 0$  for some  $e$ , we can simply delete  $e$  from  $G$ .

If  $\gamma(z) = 0$ , then  $f$  and  $\gamma$  remain optimal for  $c$  (since (2.4) holds for  $z$ ). So assume that  $\gamma(z) > 0$ . Then  $\zeta^f(z) = \hat{c}(z) = c(z) - 1$ , by (2.4). Our aim is either to make  $z$  saturated by increasing the total multiflow through  $z$  by one, or to reduce  $\gamma(z)$  to zero, in order to restore the violated complementary slackness condition at  $z$ . This task will require at most two *primal updates* (which increase  $\zeta(z)$ ) and  $O(|VG|)$  *dual updates* (which decrease  $\gamma(z)$ ).

We use the construction from the previous section. Let  $l = a + \gamma$  and construct the  $l$ -graph  $\Gamma = \Gamma^l$ , digraph  $D = D^l$  and flow  $g = g^f$ . As before,  $\gamma$  is extended to  $AD$ . We may assume that  $z$  is in  $\Gamma$  because if  $z$  belongs to no  $T$ -line, we can decrease  $\gamma(z)$  by the maximum amount allowed by (2.2). Then  $\gamma$  takes zero values on all edges in  $EG - E\Gamma$  (in view of (2.3), (2.4) (for  $\hat{c}$ ) and the positivity of  $c$ ). Let  $\bar{z} = (q, r)$  and  $\bar{z}' = (r', q')$  be the arcs forming  $\omega^{-1}(z)$ . We need to slightly modify  $D$  and  $g$  to  $H = H^l$  and  $h$ , the *extended double covering* and a *circulation* in it, as follows.

(i)  $H$  is formed by adding to  $D$  arcs of infinite capacity and zero cost from each sink  $s^2$  to each source  $t^1$ . The set of these arcs is denoted by  $\bar{W}$ ; we formally extend  $\gamma$  by zero to  $\bar{W}$ . Extend  $g$  to each arc in  $\bar{W}$  by a *rather large* positive half-integer in such a way that the resulting function  $g'$  is a circulation, i.e.,  $\text{div}_{g'}(v) = 0$  for all nodes  $v$ . This can be done since the edges from  $W$  generate arcs  $b$  from sources  $s^1$  to sinks  $t^2$ , for all distinct  $s, t \in T$ , and  $g(b) = c(b)$  is large.

(ii) Transform  $g'$  into an *integer* circulation  $h$  feasible for  $\hat{c}, \gamma$  by increasing or decreasing  $g'(b)$  by  $1/2$  at each arc  $b$  for which  $g'(b)$  is not integral. This can be done since these  $b$ 's are feasible (as, obviously,  $0 < g'(b) < \hat{c}(b)$ , yielding  $\gamma(b) = 0$ ), and the fact that each node  $v$  satisfies  $\text{div}_{g'}(v) = 0$  implies that  $v$  is incident to an even number of arcs  $b$  with  $g'(b)$  not integral.

Thus,  $h$  is an integer circulation in  $H$ , and  $\bar{z}, \bar{z}'$  are the only arcs which make  $h$  infeasible for  $c$ .

**4. Primal update.** We attempt to increase  $h$  at the arcs  $\bar{z}, \bar{z}'$  using techniques motivated by usual augmenting paths in flow problems. Let  $F$  be the set of feasible arcs in  $H$ , and  $H_F$  the digraph  $(VH, F \cup \{\bar{z}, \bar{z}'\})$ . For a digraph  $D'$  and a subset  $X \subseteq VD'$ , let  $\delta_{D'}^+(X)$  ( $\delta_{D'}^-(X)$ ) denote the set of arcs  $b = (x, y) \in AD'$  leaving  $X$ , i.e., with  $x \in X \not\equiv y$  (respectively, entering  $X$ , i.e., with  $x \notin X \ni y$ ), and let  $\delta_{D'}(X) = \delta_{D'}^+(X) \cup \delta_{D'}^-(X)$  (the *cut* induced by  $X$ ).

A path  $P = (v_0, b_1, v_1, \dots, b_k, v_k)$  in  $H_F$  is called *active* if  $h(b_i) < c(b_i)$  for all *forward* arcs  $b_i = (v_{i-1}, v_i)$ ,  $h(b_i) > 0$  for all *backward* arcs  $b_i = (v_i, v_{i-1})$  in  $P$ , and  $P$  uses neither  $\bar{z}$  nor  $\bar{z}'$  as a backward arc. If, in addition,  $P$  is a simple circuit and contains at least one of  $\bar{z}$  and  $\bar{z}'$ ,  $P$  is called an *augmenting circuit*. One can

efficiently find such a circuit or certify that it does not exist (cf. Ford and Fulkerson 1962). Three cases are possible.

*Case 1.* An augmenting circuit  $P$  containing both  $\bar{z}, \bar{z}'$  is found. In this case we push a unit of flow along  $P$ , i.e., increase  $h$  by one at each forward arc of  $P$ , and decrease it by one at each backward arc. Then the resulting function  $h'$  is a ( $c$ -admissible) integer circulation with  $h'(\bar{z}) = h'(\bar{z}') = c(z)$ . Therefore, the half-integer multiflow  $f^{\mathcal{D}}$  induced by an integer decomposition  $\mathcal{D}$  of  $h'_{|AD}$ , along with  $\gamma$ , solves (3.1).

*Case 2.* No augmenting circuit exists. Then the set  $X \subseteq VH$  reachable by active paths from  $r$  does not contain  $q$ , and the set  $Y$  reachable by active paths from  $q'$  does not contain  $r'$ . Note that if, say,  $Y$  contains  $q$ , then, obviously,  $X$  does not contain  $r'$ . So, without loss of generality, one may assume that  $q, r' \notin X$ .

*Case 3.* An augmenting circuit  $P$  containing exactly one of  $\bar{z}$  and  $\bar{z}'$ , say  $\bar{z}' \in P$ , is found. Push a unit of flow along  $P$ , obtaining a circulation  $h'$  with  $h'(\bar{z}') = c(z)$ . Then search for an augmenting circuit  $P'$  for  $h'$  and  $\bar{z}$ . (i) If such a  $P'$  is found, push unit flow along  $P'$ . The resulting circulation saturates (for  $c$ ) both  $\bar{z}, \bar{z}'$  and induces an optimal solution to (3.1). (ii) Now suppose that  $P'$  does not exist. Then the set  $X \subseteq VH$  reachable from  $r$  by active paths for  $h'$  does not contain  $q$ . Furthermore, the following is true (this will also be used in §§6 and 7).

STATEMENT 4.1. Let  $h'$  be a circulation in  $H$  with  $h'(\bar{z}) < h'(\bar{z}')$ , and let  $X$  be the set of nodes of  $H$  reachable from  $r$  by active paths for  $h'$ . Let  $q \notin X$ . Then  $r' \notin X$ .

PROOF. Let  $F_1$  and  $F_2$  (respectively,  $Q_1$  and  $Q_2$ ) be the sets of feasible (respectively, infeasible) arcs in  $\delta_H^+(X)$  and  $\delta_H^-(X)$ , respectively. Then  $\bar{z} \in Q_2, h'(b) = c(b)$  for each  $b \in (F_1 \cup Q_1 \cup Q_2) - \{\bar{z}, \bar{z}'\}$ , and  $h'(b) = 0$  for each  $b \in F_2$  (in view of the definition of  $X$ ).

Suppose that  $r' \in X$ . Let  $\alpha = c(\bar{z}) - h'(\bar{z})$  and  $\beta = c(\bar{z}') - h'(\bar{z}')$ ; then  $\alpha > \beta$ . If  $q' \in X$  (whence  $\bar{z}' \notin \delta_H(X)$ ), put  $\Delta = \alpha$ , and if  $q' \notin X$  (whence  $\bar{z}' \in Q_1$ ), put  $\Delta = \alpha - \beta$ . We have

$$(4.1) \quad \begin{aligned} 0 = \operatorname{div}_{h'}(X) &= h'(F_1) + h'(Q_1) - h'(F_2) - h'(Q_2) \\ &= c(F_1) + c(Q_1) - c(Q_2) + \Delta. \end{aligned}$$

Consider the symmetric set  $X' = \sigma(X)$ . By symmetry,  $r, r' \in X' \not\equiv q'$ . Also  $q \in X'$  if and only if  $q' \in X$ . Define the subsets  $F'_1, F'_2, Q'_1, Q'_2$  of  $\delta_H(X')$  similarly to those for  $\delta_H(X)$ . Then  $\bar{z}' \in Q'_1$ . If  $q \in X'$  (whence  $\bar{z} \notin \delta_H(X')$ ), put  $\Delta' = \beta$ , and if  $q \notin X'$  (whence  $\bar{z} \in Q'_2$ ), put  $\Delta' = \beta - \alpha$ . We have

$$(4.2) \quad \begin{aligned} 0 = -\operatorname{div}_{h'}(X') &= h'(F'_2) + h'(Q'_2) - h'(F'_1) - h'(Q'_1) \\ &\leq c(F'_2) + c(Q'_2) - c(Q'_1) + \Delta'. \end{aligned}$$

Note that  $F'_2, Q'_1, Q'_2$  are symmetric to  $F_1, Q_2, Q_1$ , respectively. Therefore,  $c(F'_2) + c(Q'_2) - c(Q'_1) = c(F_1) + c(Q_1) - c(Q_2)$ . Now (4.1) and (4.2) yield  $\Delta \leq \Delta'$ , which is impossible (in both cases) since  $\alpha > \beta$ .  $\square$

In Case 1 and in (i) of Case 3, the primal updates applied (that update  $h$ ) give an optimal solution to (3.1). In what follows we assume that no augmenting circuit exists, i.e., we are in Case 2 or in (ii) of Case 3. Then the algorithm continues by performing the dual update described in the next section. For convenience we keep the notation  $h$  for the current circulation in  $H$ , and the notation  $f$  for a corresponding multiflow in  $G$  (defined up to a decomposition of  $h_{|AD}$ ).



For  $v \in VH$ , let  $X_v$  denote the set of nodes reachable from  $v$  by active paths in  $H$ . The above analysis shows that, up to symmetry, we deal with one of the following situations:

(4.3)

- (i)  $h(\bar{z}) < h(\bar{z}') (= c(z))$ , and  $q, r' \notin X_r$ ;
- (ii)  $h(\bar{z}) = h(\bar{z}') < c(z)$ ,  $q, r' \notin X_r$ , and  $q \in X_{q'} \not\equiv r'$ ;
- (iii)  $h(\bar{z}) = h(\bar{z}') < c(z)$ ,  $q, r' \notin X_r \cup X_{q'}$ , and  $X_r \cap X_{q'} \neq \emptyset$ ;
- (iv)  $h(\bar{z}) = h(\bar{z}') < c(z)$ ,  $q, r' \notin X_r \cup X_{q'}$ , and  $X_r \cap X_{q'} = \emptyset$ .

We call (i), (ii), (iii), (iv) *Situations (of type) 1, 2, 3, 4*, respectively. In this hierarchy, type  $i$  is considered as more preferable than type  $j$  if  $i < j$ . We shall see in §6 that each dual update either solves (3.1), or leads to a primal update, or is followed by another dual update of the same or a more preferable type. Define  $R$  to be  $\{r, q'\}$  in Situation 3, and  $\{r\}$  in the other situations ( $R$  is the set of *roots* of active paths we are going to deal with). In what follows by the *reachability* we always mean the reachability by use of active paths (from or to a specified node or subset of nodes). The set of nodes reachable from  $R$  is denoted by  $X$ , and the symmetric set  $\sigma(X)$  by  $X'$ .

REMARK. At first glance, it may look artificial to distinguish between Situations 3 and 4 and define  $R$  (and, therefore,  $X$ ) in these situations differently. However, we shall see later that if  $R$  were defined in the same way (to be either  $\{r\}$  or  $\{r, q'\}$ ) in these situations, we could not guarantee that both the dual half-integrality and the property of  $z$  to be an edge of the current  $l$ -graph are simultaneously maintained.

In Situation 2, we also introduce the set

$$(4.4) \quad M = \{x \in VH: x \text{ is reachable from } q' \text{ and reachable to } q\}.$$

A subset  $Y \subseteq VH$  is called *tight* if  $h(b) = c(b)$  for each  $b \in \delta_H^+(Y) \cap F$ , and  $h(b) = 0$  for each  $b \in \delta_H^-(Y) \cap F$ ; e.g.,  $X$  is tight. The sets  $X$  and  $X'$  have important properties, which will often be used later on.

STATEMENT 4.2. (i) If  $Y \subseteq VH$  is tight, then  $VH - \sigma(Y)$  is also tight; in particular,  $VH - X'$  is tight. (ii)  $X \cap X' = \emptyset$ . (iii)  $X'$  coincides with the set  $X''$  of nodes reachable to  $R' = \sigma(R)$ .

PROOF. (i) is proved similarly to Statement 4.1. To see (ii), observe that  $q, r' \notin X$  implies  $R \cap X' = \emptyset$ . Therefore,  $R \subseteq X - X'$ . Next, it is easy to see that if  $Y$  and  $Z$  are tight, then both  $Y \cap Z$  and  $Y \cup Z$  are tight too. This implies that  $X - X'$  is tight (since both  $X$  and  $VH - X'$  are tight). From the fact that each element of  $X$  is reachable from  $R$  it easily follows that no proper subset of  $X$  which includes  $R$  can be tight. Hence,  $X - X' = X$ , yielding (ii). Finally, to see (iii), observe that  $VH - (X' \cap X'')$  is tight (since both  $VH - X'$  and  $VH - X''$  are tight). Also  $R' \subseteq X' \cap X''$ . From the definition of  $X''$  it follows that  $VH - X''$  is the maximum tight set disjoint from  $R'$ . Therefore,  $X' \cap X'' = X''$ , whence  $X'' \subseteq X'$ . The reverse inclusion  $X' \subseteq X''$  follows from the inclusion  $X \subseteq \sigma(X'')$  (since  $\sigma(X'')$  is tight, by (i)).  $\square$

The fact that  $H$  contains the set  $\bar{W}$  of arcs, each with nonzero flow, leads to the following important statement.

STATEMENT 4.3.  $X \cap (T^1 \cup T^2) = \emptyset$  (and, therefore,  $X' \cap (T^1 \cup T^2) = \emptyset$ ).

PROOF. Suppose that  $X$  contains some  $s^i \in T^1 \cup T^2$ . The existence of an arc  $b = (s^2, s^1)$  (in  $\bar{W}$ ) with  $0 < h(b) < c(b)$  implies that  $s^{3-i}$  is also reachable. Hence, both  $s^1, s^2$  are in  $X$ . By symmetry,  $s^1, s^2 \in X'$ , contrary to the fact that  $X$  and  $X'$  are disjoint (by Statement 4.2).  $\square$

Finally, in Situation 2, we observe the following:

- (4.5)
- (i)  $M$  is self-symmetric, i.e.,  $M = \sigma(M)$ ;
  - (ii)  $M$  and  $X \cup X'$  are disjoint.

PROOF.  $M$  is the intersection of  $X_q$  and the set  $Y$  of nodes reachable to  $q$ . Arguing as in the proof of (iii) in Statement 4.2, one shows that  $Y = \sigma(X_q)$ . This implies (i). Next, if (ii) is false, then  $M$  and  $X$  have a common node (as  $M = \sigma(M)$  and  $X' = \sigma(X)$ ). Hence, there is an active path from  $r$  to  $q$ ; a contradiction.  $\square$

**5. Dual update.** In the classic algorithm by Ford and Fulkerson (1962) for the min-cost max-flow problem, the current dual vector is transformed within a single cut. In contrast, the dual update described below deals simultaneously with two cuts, namely,  $\delta_H(X)$  and  $\delta_H(X')$ , where  $X$  and  $X'$  were defined in the previous section. Another feature, which makes the procedure more involved, is that the update of  $\gamma$  may change the  $l$ -graph  $\Gamma$  and, therefore, the double covering  $D$  over  $\Gamma$ . In fact, our method uses some basic ideas of the dual update from the algorithm in Karzanov (1979). However, compared with the latter, we mainly work with the  $l$ -graph  $\Gamma$  rather than the double covering  $D$ , which seems to be simpler and more enlightening. When it is not confusing, we omit  $H$  in  $\delta_H^+(\cdot)$ ,  $\delta_H^-(\cdot)$  and  $\delta_H(\cdot)$ .

First of all we have to explain how  $X$  can intersect the subgraph  $H(v) = \omega^{-1}(v)$  generated by a central vertex  $v$  ( $H(v)$  is induced by the  $v$ -arcs and called a *central subgraph*, or the  *$v$ -subgraph*). For  $i = 1, 2$ , let  $V^i(v)$  denote the set  $\{v_s^i: s \in T(v)\}$  (so  $\{V^1(v), V^2(v)\}$  is a partition of  $VH(v)$ ).

STATEMENT 5.1. Let  $Y = X \cap VH(v)$  be nonempty. Then: (i)  $Y = \{v_s^1\} \cup (V^2(v) - \{v_s^2\})$  for some  $s \in T(v)$ ; and (ii)  $h(b) = 0$  for each arc from  $V^1(v) - \{v_s^1\}$  to  $V^2(v) - \{v_s^2\}$ .

PROOF. Suppose that  $Y$  contains  $v_s^1$  for some  $s \in T(v)$ . Since  $X$  is tight and each arc  $(v_s^1, v_t^2)$  is not saturated (as it has infinite capacity),  $X$  contains  $v_t^2$  for all  $t \in T(v) - \{s\}$ . Then  $X'$  contains the symmetric nodes  $v_s^2$  and  $v_t^1$  ( $t \neq s$ ), and now the fact that  $X$  and  $X'$  are disjoint (by Statement 4.2) gives (i). To see (ii), observe that each arc  $b$  from  $V^1(v) - \{v_s^1\}$  to  $V^2(v) - \{v_s^2\}$  is feasible and entries  $X$ ; therefore,  $h(b) = 0$ .

Next suppose that  $X$  contains a node  $v_s^2$  and none of the nodes in  $V^1(v)$ . Then  $h(b) = 0$  for each arc  $b$  from  $V^1(v)$  to  $v_s^2$ . By the construction of  $D$ , no other arcs enter  $v_s^2$ . Therefore, the total flow entering  $v_s^2$  is zero, whence  $\text{div}_h(v_s^2) = 0$  implies that  $h(b) = 0$  for each arc  $b$  leaving  $v_s^2$ . But then  $v_s^2$  cannot be reachable from  $R$ ; a contradiction.  $\square$

In order to describe the update of  $\gamma$ , we associate with  $X$  certain subsets and functions in  $\Gamma$ . For  $s \in T$ , define the sets

$$(5.1) \quad A_s = \{v \in V_s: v_s^1 \in X\}; \quad A_s^\bullet = \{v \in V^\bullet: s \in T(v), v_s^1 \in X\};$$

$$B_s = \{v \in V_s: v_s^2 \in X\}; \quad \tilde{V}_s = V_s \cup A_s^\bullet.$$

From Statements 4.2(ii) and 5.1 it follows that the sets  $A_s, A_s^\bullet, B_s$  for all  $s \in T$  are pairwise disjoint, and their preimages in  $D$  give a partition of  $X \cup X'$ . Let  $\tilde{V}^\bullet = V^\bullet - \cup(A_s^\bullet: s \in T)$ . Define the function  $\rho$  on  $V\Gamma$  by

$$(5.2) \quad \begin{aligned} \rho(v) &= -1 && \text{if } v \in A_s \cup A_s^\bullet, s \in T, \\ &= 1 && \text{if } v \in B_s, s \in T, \\ &= 0 && \text{otherwise;} \end{aligned}$$

and the function  $\lambda$  on  $E\Gamma$  by

$$(5.3) \quad \begin{aligned} \lambda(e) &= \rho(v) - \rho(u) && \text{if } e = uv, u, v \in \tilde{V}_s \cup \tilde{V}^\bullet, \pi(u) < \pi(v), \\ &= -\rho(u) - \rho(v) && \text{if } e = uv, u \in \tilde{V}_s, v \in \tilde{V}_t, s \neq t. \end{aligned}$$

The algorithm chooses some rational number  $\epsilon \geq 0$  and makes the  $\epsilon$ -transformation  $\gamma \rightarrow \gamma^\epsilon$  which replaces  $\gamma(e)$ ,  $e \in E\Gamma$ , by

$$(5.4) \quad \gamma^\epsilon(e) = \max\{0, \gamma(e) + \epsilon\lambda(e)\},$$

and maintains  $\gamma^\epsilon(e) = 0$  for  $e \in EG - E\Gamma$ . Accordingly, define the length function  $l^\epsilon$  to be  $a + \gamma^\epsilon$ . Since  $\gamma^\epsilon$  is nonnegative,  $l^\epsilon$  is positive.

First we show that such an update of  $\gamma$  is correct for a sufficiently small  $\epsilon > 0$ . The proof falls into several parts. For  $j = -2, -1, 0, 1, 2$ , define  $E^j = \{e \in E\Gamma: \lambda(e) = j\}$ . For  $U, W \subseteq VH$ , let  $(U, W)$  denote the set of arcs from  $U$  to  $W$  in  $H$ . Let  $C$  denote the set of central arcs in  $H$ .

STATEMENT 5.2. Let  $Z = VH - (X \cup X')$ . The following are true:

$$(5.5) \quad \begin{aligned} E^{-2} &= \{\omega(b): b \in (X', X) - C\}, \\ E^{-1} &= \{\omega(b): b \in (Z, X)\} \quad (= \{\omega(b): b \in (X', Z)\}), \\ E^1 &= \{\omega(b): b \in (X, Z)\} \quad (= \{\omega(b): b \in (Z, X')\}), \\ E^2 &= \{\omega(b): b \in (X, X')\}. \end{aligned}$$

PROOF. Consider an arc  $b = (u_s^i, v_t^j) \in \delta(X)$ . Let  $b$  belong to  $(X', X)$ . Since  $i \leq j$  (in view of (2.7)) and  $\sigma(b) \in (X', X)$ , we may assume that  $i = 1$ . (2.7) shows that either (i)  $s = t$ ,  $j = 1$ ,  $u \in V_s$ ,  $v \in V_s \cup V^\bullet$  and  $\pi(u) < \pi(v)$ , or (ii)  $s \neq t$ ,  $j = 2$ ,  $u \in V_s$  and  $v \in V_t$ . In case (i),  $X$  contains  $u_s^2$  and  $v_s^1$ , and we have  $u \in B_s$  and  $v \in A_s \cup A_s^\bullet$  (by (5.1)), whence  $\rho(u) = 1$  and  $\rho(v) = -1$  (by (5.2)). In case (ii),  $X$  contains  $u_s^2$  and  $v_t^2$ , whence  $\rho(u) = \rho(v) = 1$ . In both cases,  $\lambda(\omega(b)) = -2$ , by (5.3).

The cases with  $b$  in  $(Z, X)$ ,  $(X, Z)$  or  $(X, X')$  are examined in a similar way (note that in these cases  $b$  is noncentral, by Statement 5.1). Also a similar analysis shows that if  $b$  is noncentral and not in  $\delta(X) \cup \delta(X')$ , then  $\lambda(\omega(b)) = 0$ . A careful examination of these remaining cases is left to the reader.  $\square$

For  $i = -1, -2$ , define  $J^i = \{e \in E\Gamma: \gamma(e) = 0, \lambda(e) = i\}$ . Expression (5.5) and the tightness of  $X$  imply that

$$(5.6) \quad \begin{aligned} (i) \quad & z \in E^{-1} \cup E^{-2}; \\ (ii) \quad & \zeta^f(e) = 0 \text{ for each } e \in J^{-1} \cup J^{-2}; \\ (iii) \quad & \zeta^f(e) = c(e) \text{ for each edge } e \neq z \text{ in } (E^{-2} - J^{-2}) \cup (E^{-1} - J^{-1}) \cup E^1 \cup E^2. \end{aligned}$$

Therefore, increasing  $\epsilon$  decreases  $\gamma(z)$ , and each edge  $e \neq z$  with  $\gamma^\epsilon(e) > 0$  remains saturated by  $f$ . The choice of  $\epsilon$  for the update of  $\gamma$  is restricted by several factors. The first upper bound on  $\epsilon$  is

$$(5.7) \quad \epsilon_1 = \min\{\min\{\gamma(e): e \in E^{-1} - J^{-1}\}, \min\{\gamma(e)/2: e \in E^{-2} - J^{-2}\}\}.$$

In particular,  $\epsilon_1 \leq \gamma(z)$ . Obviously,  $\epsilon_1 > 0$ . In other words,  $\epsilon_1$  is the minimum  $\epsilon$  for which the  $\epsilon$ -transformation reduces some positive  $\gamma(e)$  to zero. In what follows we

always assume that  $0 < \epsilon \leq \epsilon_1$ . Another important property provided by the choice of  $\epsilon_1$  is as follows.

STATEMENT 5.3. Let  $P$  be a  $(T, l)$ -line. Then  $l^\epsilon(P) \geq p$ . Moreover,  $l^\epsilon(P) = p$  if  $f(P) > 0$ .

PROOF. Consider the  $T^1$  to  $T^2$  dipath  $Q = \omega^{-1}(P)$ . By Statement 4.3, the ends of  $Q$  are outside both  $X$  and  $X'$ . Hence,  $Q$  enters  $X$  as many times as it leaves  $X$ , and similarly for  $X'$ . Also if  $\delta(X) \cup \delta(X')$  contains a central arc  $b$ , then  $b \in (X', X)$ . These arguments together with (5.5) show that  $\lambda(P) = 2\alpha$ , where  $\alpha$  is the number of times  $Q$  meets central arcs in  $\delta(X)$  (clearly  $\alpha$  is 0 or 1). Next, since  $\epsilon \leq \epsilon_1$ , we derive from (5.4) and (5.7) that  $\gamma^\epsilon(P) = \gamma(P) + \epsilon\lambda(P) + \epsilon\beta_1 + 2\epsilon\beta_2$ , where  $\beta_i$  is the number of edges of  $P$  in  $J^{-i}$ . Hence,

$$\begin{aligned} l^\epsilon(P) &= a(P) + \gamma^\epsilon(P) = a(P) + \gamma(P) + 2\epsilon\alpha + \epsilon\beta_1 + 2\epsilon\beta_2 \\ &\geq a(P) + \gamma(P) = l(P) = p. \end{aligned}$$

The second part of the statement is valid because if  $f(P) > 0$  then  $\beta_1 = \beta_2 = 0$  (by (5.6)(ii)) and  $\alpha = 0$  (by Statement 5.1 and the fact that  $h(b) > 0$  for all arcs  $b$  of  $Q$ ).  $\square$

A  $(T, l)$ -line  $P$  with  $l^\epsilon(P) = p$  is called a *nonbroken line*. By Statement 5.3, each edge  $e \in E\Gamma$  with  $\zeta^J(e) > 0$  belongs to a nonbroken line. Let us call an arc  $b$  of  $D$  *forbidden* if it is feasible (i.e.,  $\gamma(b) = 0$ ) and either enters  $X$  or leaves  $X'$  or both. From the above proof one can see that

(5.8) for a  $(T, l)$ -line  $P$ , the following are equivalent:

- (i)  $P$  is nonbroken;
- (ii)  $\omega^{-1}(P)$  has no forbidden arcs;
- (iii)  $P$  neither meets  $J^{-1} \cup J^{-2}$  nor contains three consecutive nodes  $u \in V_t$ ,  $v \in A_s^\bullet$ ,  $w \in V_{t'}$  for  $s \neq t, t'$ .

Statement 5.3 implies that for a sufficiently small  $\epsilon$ , the  $l^\epsilon$ -distance between distinct terminals is at least  $p$ , and  $f$  goes along  $(T, l^\epsilon)$ -lines (taking into account that  $l^\epsilon$  is continuous in  $\epsilon$ , that  $l(P) > p$  for each  $T$ -path  $P$  in  $G$  which is not a  $(T, l)$ -line, and that the number of these (simple) paths is finite). Our aim is to find a reasonable bound on  $\epsilon$  to guarantee these properties.

Let  $\Gamma'$  be the subgraph of  $\Gamma$  formed by  $T$  and the nodes and edges of nonbroken lines. By Statement 5.3,  $f$  goes along lines in  $\Gamma'$ . This implies that  $\gamma^\epsilon$  takes nonzero values only within  $E\Gamma' \cup \{z\}$  (since the edges in  $E^1 \cup E^2$  are saturated, by (5.6)(iii)). Using (5.8), one can give an efficient procedure to extract  $\Gamma'$  from  $\Gamma$ . There is a simple explicit description of  $\Gamma'$ , as follows.

Let  $J$  be the set of edges  $e$  of  $\Gamma$  such that any  $(T, l)$ -line passing through  $e$  meets  $J^{-1} \cup J^{-2}$ ; clearly  $J$  can be found efficiently. By (5.8), the elements of  $J$  cannot belong to nonbroken lines. The converse property also holds, as shown by the following statement which will be proved in the next section.

STATEMENT 5.4. Each of the following belongs to a nonbroken line: (i) the node  $\omega(x)$  for each  $x \in X$ ; (ii) each edge in  $E\Gamma - J$ ; (iii) the edge  $z$ .

Thus, assuming that this statement is true, we observe that  $E\Gamma' = E\Gamma - J$ , that  $\Gamma'$  includes  $A_s$ ,  $A_s^\bullet$  and  $B_s$  for all  $s \in T$ , and that  $\gamma^\epsilon(e) = 0$  for all  $e \in E\Gamma - E\Gamma'$ . For  $s \in T$ , let  $V'_s = \tilde{V}_s \cap V\Gamma'$ , and let  $V'^\bullet = \tilde{V}^\bullet \cap V\Gamma'$  (cf. (5.1)). Then  $A_s, A_s^\bullet, B_s \subseteq V'_s$ . For  $v \in V\Gamma'$ , define

$$(5.9) \quad \pi^\epsilon(v) = \pi(v) + \epsilon\rho(v).$$

Comparing (2.5) with (5.3) and (5.4), we observe that for  $e = uv \in E\Gamma$ ,

$$(5.10) \quad \begin{aligned} l^\epsilon(e) &= |\pi^\epsilon(v) - \pi^\epsilon(u)| && \text{if } u, v \in V'_s \cup V'^\bullet, s \in T, \\ &= p - \pi^\epsilon(u) - \pi^\epsilon(v) && \text{if } u \in V'_s, v \in V'_t, s \neq t. \end{aligned}$$

When  $\epsilon$  is growing from zero,  $\pi^\epsilon$  decreases by  $\epsilon$  on the elements of each  $A_s \cup A_s^\bullet$ , increases by  $\epsilon$  on the elements of each  $B_s$ , and remains unchanged on the other nodes of  $\Gamma$ . This motivates the next upper bound on  $\epsilon$  to be

$$(5.11) \quad \epsilon_2 = \min\{p/2 - \pi(v) : v \in B_s, s \in T\},$$

which says that the growth of  $\epsilon$  should be stopped as soon as some increasing number  $\pi^\epsilon(v)$  achieves  $p/2$ . From (5.10) it follows that  $\pi^\epsilon$  is the potential function for  $l^\epsilon$  in  $\Gamma'$  whenever  $\epsilon \leq \epsilon_2$  (i.e., for each  $v \in V\Gamma'$ ,  $\pi^\epsilon(v)$  is the minimum  $l^\epsilon$ -length of a path from  $T$  to  $v$  in  $\Gamma'$ ). Since each set  $B_s$  has no central nodes for  $l$ ,  $\epsilon_2 > 0$ .

Consider two nodes  $u$  and  $v$  in  $\Gamma'$ . Let  $d(u, v)$  be the minimum cost  $a(P)$  of a path  $P$  from  $u$  to  $v$  in  $G$  with all edges and inner nodes (if any) not in  $\Gamma'$ ; if such a path does not exist,  $d(u, v) = \infty$ . When  $\epsilon$  is growing, it may happen at some moment that such a minimum cost path  $P$  becomes a line for the current length function (and further growing  $\epsilon$  would make the distance between some terminals less than  $p$ ). To account for such a moment, we associate with the *ordered* pair  $(u, v)$  two numbers  $\bar{l}(u, v)$  and  $\bar{\lambda}(u, v)$ , where

$$(5.12) \quad \begin{aligned} \bar{l}(u, v) &= \pi(v) - \pi(u) && \text{and } \bar{\lambda}(u, v) = \rho(v) - \rho(u) \\ &&& \text{if } u, v \in V'_s \cup V'^\bullet, s \in T, \\ \bar{l}(u, v) &= p - \pi(u) - \pi(v) && \text{and } \bar{\lambda}(u, v) = -\rho(u) - \rho(v) \\ &&& \text{if } u \in V'_s, v \in V'_t, s \neq t. \end{aligned}$$

One can see that if  $u$  and  $v$  are connected by a line (e.g., an edge)  $P$  in  $\Gamma'$ , then  $l^\epsilon(P)$  equals  $\max\{\bar{l}(u, v) + \epsilon\bar{\lambda}(u, v), \bar{l}(v, u) + \epsilon\bar{\lambda}(v, u)\}$ . Our third upper bound on  $\epsilon$  is

$$(5.13) \quad \epsilon_3 = \min\{\epsilon(u, v) : u, v \in V\Gamma', \epsilon(u, v) \geq 0\},$$

where

$$(5.14) \quad \epsilon(u, v) = (d(u, v) - \bar{l}(u, v)) / \bar{\lambda}(u, v).$$

We observe that  $d(u, v) > \bar{l}(u, v)$ ; therefore,  $\epsilon_3 > 0$ .

STATEMENT 5.5. Let  $0 \leq \epsilon \leq \min\{\epsilon_1, \epsilon_2, \epsilon_3\}$ . Then  $\text{dist}_T^\epsilon(s, t) \geq p$  for all distinct  $s, t \in T$ . In particular,  $\Gamma'$  is a subgraph of  $\Gamma^{l^\epsilon}$ .

PROOF. We know that if  $\epsilon' \leq \epsilon_1, \epsilon_2$ , then  $\pi^{\epsilon'}$  is the potential function for  $l^{\epsilon'}$  within  $\Gamma'$ . Suppose that  $l^{\epsilon'}(P) < p$  for some  $T$ -path  $P$ . Using the potentials  $\pi^{\epsilon'}$  in  $\Gamma'$  and the fact that each node of  $\Gamma'$  belongs to a  $(T, l^{\epsilon'})$ -line in it, one can see that there is a part  $P'$  of  $P$ , say, from  $u$  to  $v$ , such that both  $u, v$  are in  $\Gamma'$  while the other nodes of  $P'$  are not, and either (i)  $u, v \in V'_s \cup V'^\bullet$  and  $l^{\epsilon'}(P') < |\pi^{\epsilon'}(v) - \pi^{\epsilon'}(u)|$ , or (ii)  $u \in V'_s, v \in V'_t$  ( $s \neq t$ ) and  $l^{\epsilon'}(P') < p - \pi^{\epsilon'}(u) - \pi^{\epsilon'}(v)$ .

In case (i), assume for definiteness  $\pi^\epsilon(v) \geq \pi^\epsilon(u)$ . We have  $l^\epsilon(P') = a(P') \geq d(u, v)$  (by the definition of  $d$ ). Also  $\pi^\epsilon(v) - \pi^\epsilon(u) = \pi(v) + \epsilon\rho(v) - \pi(u) - \epsilon\rho(u) = \bar{l}(u, v) + \epsilon\bar{\lambda}(u, v)$ . Therefore,  $l^\epsilon(P') < \pi^\epsilon(v) - \pi^\epsilon(u)$  implies  $d(u, v) < \bar{l}(u, v) + \epsilon\bar{\lambda}(u, v)$ . This is impossible when  $\bar{\lambda}(u, v) \leq 0$  (since  $d(u, v) > \bar{l}(u, v)$ ). And if  $\bar{\lambda}(u, v) \geq 0$ , then  $\epsilon(u, v) > 0$ , and now (5.14) and  $\epsilon \leq \epsilon_3 \leq \epsilon(u, v)$  imply  $d(u, v) \geq \bar{l}(u, v) + \epsilon\bar{\lambda}(u, v)$ ; a contradiction.

Case (ii) leads to a contradiction in a similar way.  $\square$

Thus, the choice of  $\epsilon$  to be  $\min\{\epsilon_1, \epsilon_2, \epsilon_3\}$  gives  $\gamma^\epsilon$  which, together with  $f$ , satisfies (2.2)–(2.4) for  $\hat{c}$ . In the next section we show that such a choice ensures fast convergence to optimal solutions for (3.1).

**6. Complexity of the algorithm and dual half-integrality.** Our goal is to show that  $O(|VG|)$  consecutive dual updates either reduces  $\gamma(z)$  to zero or make it possible to apply the primal update increasing the total multiflow at  $z$ . First of all we establish three facts (Statements 6.1–6.3).

**STATEMENT 6.1.** The edge  $z$  belongs to a nonbroken line, i.e., (iii) in Statement 5.4 is true.

**PROOF.** If  $\zeta^f(z) > 0$ , the result immediately follows from Statement 5.3. So assume that  $\zeta^f(z) = 0$ . Then  $h(\bar{z}) = h(\bar{z}') = 0$ , therefore, we are in Situation 2, 3 or 4 (see (4.3)). Choose a dipath  $Q_1$  from  $r$  to  $T^2$  in  $D$ , and let  $b$  be the first arc of  $Q_1$  leaving  $X$ . Since  $h(b) > 0$  (as  $X$  is tight), we may assume that  $Q_1$  is chosen so that all its arcs following  $b$  have nonzero flow. Then  $Q_1$  has no forbidden arcs.

Suppose that a situation of type 3 takes place. Then  $q \in X'$  (as  $R = \{r, q'\}$ ). The argument “symmetric” to that above shows that there is a  $T^1$  to  $q$  dipath  $Q_2$  without forbidden arcs. Now the concatenation of the path  $Q_2$ , the arc  $\bar{z}$  and the path  $Q_1$  gives a  $T^1$  to  $T^2$  dipath which contains  $\bar{z}$  and has no forbidden arcs.

In Situation 4, we argue in a similar way using the facts that  $q$  belongs to the set  $\sigma(X_{q'})$  which is disjoint from  $X$  and that  $VH - \sigma(X_{q'})$  is tight (as  $X_r \cap X_{q'} = \emptyset$ ,  $\sigma(X_r) \cap X_{q'} = \emptyset$  and  $X_{q'}$  is tight).

Now consider a situation of type 2. Let  $P = (x_0, b_1, x_1, \dots, b_k, x_k)$  be an active path from  $q'$  to  $q$ . Note that  $P$  cannot contain only forward arcs; for otherwise there would exist a  $T^1$  to  $T^2$  dipath  $L$  in  $D$  which passes through both  $\bar{z}$  and  $\bar{z}'$ , implying that the path  $\omega(L)$  is not a  $T$ -line, contrary to (2.9). Thus,  $P$  has a backward arc; let  $b_i$  be the last of such arcs. Since  $h(b_i) > 0$ , there is a dipath  $Q_2$  from  $T^1$  to  $x_i$  with nonzero flow at all its arcs. Concatenating the path  $Q_2$ , the part of  $P$  from  $x_i$  to  $q$ , the arc  $\bar{z}$  and the path  $Q_1$ , we obtain a  $T^1$  to  $T^2$  dipath which contains  $\bar{z}$  and avoids forbidden arcs.  $\square$

**STATEMENT 6.2.** Let  $b = (x, y)$  be an arc of  $D$  with  $x, y \in X$ . Then the element  $\omega(b)$  belongs to a nonbroken line.

**PROOF.** Choose a dipath  $Q_1$  from  $y$  to  $T^2$  and an active path  $P$  from  $R$  to  $x$ . By the argument as in the proof of Statement 6.1, we may assume that  $h(b) = 0$ , and  $Q_1$  has no forbidden arcs. Let  $w$  be the last node of  $P$  such that either  $w \in R$  or  $w$  is the tail of an arc with nonzero flow. Then there is a  $T^1$  to  $w$  dipath  $Q_2$  without forbidden arcs (if  $w \in R$ ,  $Q_2$  exists by Statement 6.1). Now the concatenation of the path  $Q_2$ , the part of  $P$  from  $w$  to  $x$ , the arc  $b$  and the path  $Q_1$  gives a  $T^1$  to  $T^2$  dipath which contains  $b$  and has no forbidden arcs.  $\square$

**STATEMENT 6.3.** Assuming that situation of type 2 occurs, let  $M$  be the set defined in (4.4). Let  $b = (x, y)$  be an arc of  $D$  with  $x, y \in M$ . Then the element  $\omega(b)$  belongs to a nonbroken line.