RIMS-1687

Half-integrality of node-capacitated multiflows and tree-shaped facility locations on trees

By

Hiroshi HIRAI

January 2010

京都大学　数理解析研究所

RESEARCH INSTITUTE FOR MATHEMATICAL SCIENCES

KYOTO UNIVERSITY, Kyoto, Japan

Half-integrality of node-capacitated multiflows and tree-shaped facility locations on trees

Hiroshi HIRAI
Research Institute for Mathematical Sciences,
Kyoto University, Kyoto 606-8502, Japan
hirai@kurims.kyoto-u.ac.jp

January, 2010

Abstract

In this paper, we establish a novel duality relationship between node-capacitated multiflows and tree-shaped facility locations. We prove that the maximum value of a tree-distance-weighted maximum node-capacitated multiflow problem is equal to the minimum value of the problem of locating subtrees in a tree, and the maximum is attained by a half-integral multiflow. Utilizing this duality, we show that a half-integral optimal multiflow and an optimal location can be found in strongly polynomial time. These extend previously known results in the maximum free multiflow problems. We also show that the set of tree-distance weights is the only class having bounded fractionality in maximum node-capacitated multiflow problems.

1 Introduction

A node-capacitated *network* (V, E, S, b, c) consists of an undirected graph (V, E), a specified node subset $S \subseteq V$, called a *terminal set*, a nonnegative integral node-capacity $b : V \to \mathbf{Z}_+$, and a nonnegative integral edge-capacity $c : E \to \mathbf{Z}_+$. A path connecting distinct terminals is called an *S-path*. A *multiflow* (*multicommodity flow*) is a pair $(\mathcal{P}, \lambda)$ of a set $\mathcal{P}$ of S-paths and a nonnegative flow-value function $\lambda : \mathcal{P} \to \mathbf{R}_+$ satisfying the capacity constraint:

$$\begin{aligned} \sum\{\lambda(P) \mid P \in \mathcal{P}, x \in VP\} &\leq b(x) \quad (x \in V), \\ \sum\{\lambda(P) \mid P \in \mathcal{P}, e \in EP\} &\leq c(e) \quad (e \in E), \end{aligned}$$

where VP and EP denote the sets of nodes and edges in P, respectively. Let $\mu : \binom{S}{2} \to \mathbf{Q}_+$ be a nonnegative rational-valued function defined on the set $\binom{S}{2}$ of terminal pairs. For a multiflow $f = (\mathcal{P}, \lambda)$, the total flow-value $\mathrm{val}(\mu, f)$ with respect to μ is defined by $\sum_{P \in \mathcal{P}} \mu(s_P, t_P)\lambda(P)$, where s_P and t_P denote the ends of P. We consider the following problem:

(1.1) Maximize $\mathrm{val}(\mu, f)$ over all multiflows f in (V, E, S, b, c).

Example 1. Consider the case $S = \{s, t\}$ and $\mu(s, t) = 1$. This is a maximum flow problem. By the max-flow min-cut theorem or a version of Menger's theorem, the maximum flow value is equal to the minimum cut value, and there exists an *integral* maximum flow in (1.1).

Example 2. Consider the case where $|S| \geq 3$ and $\mu(s,t) = 1$ for all terminal pairs s,t. Then (1.1) is the fractional S-path packing problem, or called the maximum *free multiflow* problem. In this case, the integrality theorem does not hold. In edge-only-capacitated case ($b \to +\infty$), Lovász [18] and Cherkassky [2] proved the existence of a *half-integral* optimal multiflow, and a combinatorial min-max formula. Also, in node-capacitated case, a similar property holds. Vazirani [30] proved the dual half-integrality. Pap [23, 24] proved the primal half-integrality and gave a strongly polynomial time algorithm to find a half-integral optimal multiflow; see also a further algorithmic development [1].

The main contribution of this paper is to extend this half-integrality result to more general weights μ related to *trees*, with establishing a combinatorial min-max relation by a *tree-shaped facility location* on the tree associated with μ. A rational terminal weight $\mu : \binom{S}{2} \to \mathbf{Q}_+$ is called a *tree distance* if there exist a positive rational γ, a tree $\Gamma = (V\Gamma, E\Gamma)$, and a family $\{R_s \mid s \in S\}$ of its subtrees indexed by S such that

$$\mu(s,t) = \gamma d_\Gamma(R_s, R_t) \quad (s,t \in S).$$

Here subtree R_s is regarded as a node subset, and $d_\Gamma(R_s, R_t)$ denotes the shortest path distance between R_s and R_t. Namely $\mu(s,t)$ is represented as distances among subtrees $\{R_s\}$ in a tree Γ with uniform edge-length γ. We also say that μ *is realized by* $(\Gamma, \{R_s\}_{s\in S}; \gamma)$. We particularly call μ a *tree metric* if each R_s is a single node. As far as we know, tree distances were first studied by [6] although tree metrics have been well studied in connection with phylogenetic trees [26]. Now suppose that μ is a tree distance realized by $(\Gamma, \{R_s\}_{s\in S}; \gamma)$. Let $\mathcal{F}\Gamma \subseteq 2^{V\Gamma}$ be a set of all subtrees in Γ. For a subtree F, the *diameter* $\mathrm{diam}F$ is defined by the maximum distance of nodes in F. Consider the following subtree location problem:

$$\begin{aligned} (1.2) \qquad \text{Min.} \quad & \gamma \sum_{y\in V} b(y)\mathrm{diam}F(y) + \gamma \sum_{xy\in E} c(xy) d_\Gamma(F(x), F(y)) \\ \text{s.t.} \quad & F : V \to \mathcal{F}\Gamma, \\ & F(s) \cap R_s \neq \emptyset \quad (s \in S). \end{aligned}$$

This problem can be read as follows. We associate each x with a tree-shaped facility $F(x)$ in Γ. Facilities $F(x)$ have the *communication cost* which is a nondecreasing function of distances $d_\Gamma(F(x), F(y))$ among subtrees. So if subtrees become large, then the communication cost becomes small. However this could be balanced by the *size cost* which is a nondecreasing function of diameters $\mathrm{diam}F(x)$. Then the problem is to find the subtree configuration of the smallest cost. Our main theorem establishes the duality between (1.1) and (1.2) together with half-integrality property:

Theorem 1.1. *Let μ be a tree distance on S realized by $(\Gamma, \{R_s\}_{s\in S}; \gamma)$. For any network (V, E, S, b, c), the maximum value of* (1.1) *is equal to the minimum value of* (1.2). *Moreover there exists a half-integral optimal solution in* (1.1).

We also show that a half-integral optimal multiflow and an optimal subtree map can be found in strongly polynomial time.

The problem of locating several tree-shaped facilities in a tree network was studied by Tamir and Lowe [27] and Hakimi, Schmeichel, and Labbe [4], extending single tree-shaped facility location by Minieka [20]. In their model, subtrees are imposed to be disjoint, and the size cost (constraint) is the number of edges, instead of the diameter in our model. Many problem formulations (center, median, and etc) were shown to be

NP-hard [4]. So the polynomial time solvability of (1.2) has own interest in the location theory.

Consider the case where the network is edge-only-capacitated ($b \to +\infty$). Then the diameter of each subtree $F(x)$ must be zero, i.e., $F(x)$ is a single point. Thus (1.2) reduces to a point-location problem on a tree:

$$\begin{aligned} (1.3) \qquad \text{Min.} \quad & \gamma \sum_{xy \in E} c(xy) d_{\Gamma}(\rho(x), \rho(y)) \\ \text{s.t.} \quad & \rho : V \to V\Gamma, \\ & \rho(s) \in R_s \quad (s \in S). \end{aligned}$$

The location problem of this type has been well studied in the literature [28]. In the multiflow theory (of edge-only-capacitated), the corresponding duality relationship has been discovered by Karzanov [16, 17], and further developed by the author [7, 8, 9], for more general weights beyond tree distances. So our main theorem is a node-capacitated variation of these results.

In examples below, networks are supposed to be node-only-capacitated; see the above-cited references for edge-only-capacitated examples. In this case, for each edge $xy \in E$, two facilities $F(x)$ and $F(y)$ must intersect, i.e., $d_{\Gamma}(F(x), F(y)) = 0$. Then (1.2) reduces to:

$$\begin{aligned} (1.4) \qquad \text{Min.} \quad & \gamma \sum_{y \in V} b(y) \mathrm{diam} F(y) \\ \text{s.t.} \quad & F : V \to \mathcal{F}\Gamma, \\ & F(x) \cap F(y) \neq \emptyset \quad (xy \in E), \\ & F(s) \cap R_s \neq \emptyset \quad (s \in S). \end{aligned}$$

This problem gives a unified interpretation of *combinatorial dual solutions* as follows. For simplicity, we assume that the node-capacity of each terminal s is sufficiently large ($b(s) \to +\infty$), and there is no edge joining terminals.

Example 3. Consider single flows; $S = \{s, t\}$ and $\mu(s, t) = 1$. Obviously μ is a tree metric realized by one edge $v_s v_t$ with $R_s = \{v_s\}$ and $R_t = \{v_t\}$ and $\gamma = 1$. Then (1.2) is given by

$$\begin{aligned} (1.5) \qquad \text{Min.} \quad & \sum_{y \in V \setminus \{s,t\}} b(y) \mathrm{diam} F(y) \\ \text{s.t.} \quad & F : V \to \{\{v_s\}, \{v_s, v_t\}, \{v_t\}\}, \\ & F(x) \cap F(y) \neq \emptyset \quad (xy \in E), \\ & (F(s), F(t)) = (\{v_s\}, \{v_t\}). \end{aligned}$$

For a feasible solution F in (1.5), the inverse image of $\{v_s, v_t\}$ is an (s, t)-node cut since there is no edge between $F^{-1}(\{v_s\})$ and $F^{-1}(\{v_t\})$. Moreover the objective value is the sum of the node-capacity over $F^{-1}(\{v_s, v_t\})$. Hence the maximum flow value is equal to the minimum capacity of (s, t)-node cuts.

Example 4. Consider the free multiflows; $S = \{s_1, s_2, \ldots, s_k\}$ and $\mu(s_i, s_j) = 1$ for each i, j. Then μ is a tree metric realized by a star Γ of k leaves $v_1, v_2, \ldots, v_k$ with center v_0,

edge-length $\gamma = 1/2$, and $R_{s_i} = \{v_i\}$ for $i = 1, 2, \ldots, k$. Then (1.2) is given by

$$
\begin{aligned}
(1.6) \qquad \text{Min.} \quad & \frac{1}{2} \sum_{y \in V \setminus S} b(y) \mathrm{diam} F(y) \\
\text{s.t.} \quad & F(x) = \{v_i\}, \{v_j, v_0\}, \text{or } \{v_0, v_1, v_2, \ldots, v_k\} \quad (x \in V), \\
& F(x) \cap F(y) \neq \emptyset \quad (xy \in E), \\
& F(s_j) = \{v_j\} \quad (j = 1, 2, \ldots, k).
\end{aligned}
$$

This problem can also be represented as

$$
\begin{aligned}
(1.7) \qquad \text{Min.} \quad & b(U_0) + \frac{1}{2} \sum_{i=1}^{k} b(\mathrm{Bd} U_i) \\
\text{s.t.} \quad & \text{disjoint node subsets } U_0, U_1, U_2, \ldots, U_k, \\
& s_i \in U_i \quad (i = 1, 2, \ldots, k),
\end{aligned}
$$

where $\mathrm{Bd}(U_i)$ is the set of nodes in U_i incident to $V \setminus (U_i \cup U_0)$. This min-max formula coincides with one given in [30, Section 19.3], and can also be derived from Mader's S-path theorem [19]; see [24]. The equivalence between (1.6) and (1.7) can be seen as follows. For $U_0, U_1, U_2, \ldots, U_k$ feasible to (1.7), define subtree map $F : V \to \mathcal{F}\Gamma$ by

$$
F(x) = \begin{cases} \{v_0, v_1, v_2, \ldots, v_k\} & \text{if } x \in U_0, \\ \{v_0, v_i\} & \text{if } x \in \mathrm{Bd} U_i, \\ \{v_i\} & \text{if } x \in U_i \setminus \mathrm{Bd} U_i, \\ \{v_0\} & \text{otherwise} \end{cases}
$$

Then F is feasible to (1.6) with the same objective value. Conversely, for a feasible subtree map F in (1.6), let $U_0 = \{x \in V \mid F(x) = \{v_0, v_1, v_2, \ldots, v_k\}\}$ and $U_i = \{x \in V \mid F(x) = \{v_i\} \text{ or } \{v_0, v_i\}\}$. Then $U_0, U_1, \ldots, U_k$ are feasible to (1.7) and does not increase the objective value by $\mathrm{Bd} U_i \subseteq F^{-1}(\{v_0, v_i\})$.

The rest of the paper is organized as follows. In Section 2, we give some preliminary results for linear programming dual to (1.1), fractional b-matchings, and tree distances. In Section 3, we prove the former part of Theorem 1.1, which follows from a key fact (Lemma 3.2) that every minimal dual solution of (1.1) can be represented as distances and diameters of subtrees in a *geometric realization* of tree Γ. Then the LP-dual of (1.1) reduces to a *continuous* tree-shaped facility location problem, and its extreme solutions correspond to subtree maps in the discrete one (1.2). In Section 4, we prove a stronger half-integral assertion: there exists a half-integral optimal multiflow of *minimum total cost.* This extends a classical result by Karzanov on minimum cost free multiflows [12], and its node-capacitated extension by Pap [23, 24]. Also in the minimum cost problem, its LP-dual reduces to a *convex-cost* continuous tree-shaped facility location. Our proof is a combination of this subtree representation and the ideas of Karzanov [14] (construction of an optimal multiflow from a dual optimum) and of Pap [24] (use of fractional b-matchings). By the optimality criterion with an optimal subtree map F^*, we can construct a half-integral fractional b-matching ζ^*, which coincides with the flow-support of some optimal multiflow. From ζ^*, we can recover a half-integral optimum f^*. These constructions can be done in a strongly polynomial time by using Tardos' method [29]. A design of a combinatorial strongly polynomial algorithm is still open. In Section 5, we prove that the set of tree distances is the only class admitting the half-integrality theorem, more strongly, if μ is not a tree distance, then there is no positive integer k such that (1.1) has a $1/k$-integral optimal solution for every network. This

completes the fractionality problem in maximum node-capacitated multiflows; recently the edge-capacitated version of this problem was solved in [9]. In Section 6, we give some concluding remarks.

Notation. Let $\mathbf{R}$, $\mathbf{Q}$, and $\mathbf{Z}$ denote the sets of reals, rationals, and integers, respectively and let $\mathbf{R}_+$, $\mathbf{Q}_+$, and $\mathbf{Z}_+$ denote the sets of nonnegative reals, rationals, and integers, respectively. For a subset $S \subseteq V$, the characteristic function (vector) of S is denoted by χ_S, i.e., $\chi_S(x) = 1$ if $x \in S$ and $\chi_S(x) = 0$ otherwise. For a function b on a set V and $S \subseteq V$, let $b(S)$ denote the sum of $b(s)$ over $s \in S$. Throughout the paper, a graph $G = (V, E)$ is an undirected graph without parallel edges and loops. For a node $x \in V$, the set of all edges incident to x is denoted by δx. By a path we mean a simple path, and denote it by a chain of nodes, such as $P = (x_1, x_2, \ldots, x_m)$. For a set V, let E_V denote the edge set of the complete graph on V, i.e., $E_V = \binom{V}{2}$. We often do not distinguish $\{x\}$ and x.

2 Preliminaries

Here we present some preliminary results.

LP formulations. A naive LP dual to (1.1) is given by

$$\begin{aligned} (2.1) \qquad \text{Min.} \quad & \sum_{y \in V} b(y)h(y) + \sum_{xy \in E} c(xy)d(xy) \\ \text{s.t.} \quad & h(VP) + d(EP) \geq \mu(s_P, t_P) \quad (P\text{: an } S\text{-path}), \\ & (h, d) \in \mathbf{R}_+^V \times \mathbf{R}_+^E. \end{aligned}$$

This LP has the following extended formulation of polynomial size:

$$\begin{aligned} (2.2) \qquad \text{Min.} \quad & \sum_{y \in V} b(y)h(y) + \sum_{xy \in E} c(xy)d(xy) \\ \text{s.t.} \quad & d(xy) + d(yz) - d(xz) + h(y) \geq 0 \quad (xz \in E_V, y \in V \setminus \{x, z\}), \\ & d(st) + h(s) + h(t) \geq \mu(s, t) \quad (st \in E_S), \\ & (h, d) \in \mathbf{R}_+^V \times \mathbf{R}_+^{E_V}. \end{aligned}$$

Indeed, take (h, d) feasible to (2.2), and project d to $\mathbf{R}^E$. Then, for any S-path $P = (x_0, x_1, x_2, \ldots, x_m)$, we have

$$\begin{aligned} h(VP) + d(EP) \quad = \quad & d(x_0x_m) + h(x_0) + h(x_m) \\ & + \sum_{i=0}^{m-2} [d(x_ix_{i+1}) + d(x_{i+1}x_m) - d(x_ix_m) + h(x_i)] \geq \mu(x_0, x_m). \end{aligned}$$

Therefore (h, d) is feasible to (2.1). Conversely, take (h, d) feasible to (2.1). Consider $\overline{d} : E_V \to \mathbf{R}_+$ defined by

$$\overline{d}(xy) = \min\{h(VP \setminus \{x, y\}) + d(EP) \mid P \text{ is a path connecting } x \text{ and } y\} \quad (xy \in E_V).$$

Then $\overline{d}(xy) \leq d(xy)$ for $xy \in E$, and $(h, \overline{d})$ is feasible to (2.2).

Fractional b-matching. We note a basic fact of fractional b-matchings. Let $G = (V, E)$ be a graph. Let $\underline{b}, \overline{b} : V \to \mathbf{R}_+$ be a lower- and a upper-bound functions on V. A *fractional b-matching* is a nonnegative function $\zeta : E \to \mathbf{R}_+$ on edges satisfying

$$\underline{b}(y) \leq \zeta(\delta y) \leq \overline{b}(y) \quad (y \in V).$$

Let $P(\underline{b}, \overline{b})$ be the polyhedron formed by all fractional b-matchings.

Lemma 2.1. *Suppose that both $\underline{b}$ and $\overline{b}$ are integral. Then $P(\underline{b}, \overline{b})$ is half-integral. Moreover, for each extreme point ζ^* and each node $y \in V$, $\zeta^*(\delta y)$ is integral.*

The latter property, noted by [24], plays a key role in constructing a half-integral multiflow f^* from the flow-support ζ^*. We give a proof for completeness.

Proof. As is well known, any fractional b-matching can be obtained by a circulation in a directed network constructed from $(V, E, \underline{b}, \overline{b})$ as follows. For each node $x \in V$, consider two nodes x^+, x^- with directed edge x^+x^- from x^+ to x^- of lower capacity $\underline{b}(x)$ and upper capacity $\overline{b}(x)$. For each edge $xy \in V$, consider two directed edges y^-x^+ and x^-y^+ of lower capacity 0 and upper capacity ∞. Consider a circulation ϕ in the resulting directed network. Let $\zeta : E \to \mathbf{R}$ be defined by $\zeta(xy) := (\phi(x^-y^+) + \phi(y^-x^+))/2$ for $xy \in E$. Then ζ is a fractional b-matching. Conversely, every fractional b-matching can be obtained in this way. So every extreme point ζ^* of $P(\underline{b}, \overline{b})$ is the image of an extreme point ϕ^* of the circulation polyhedron defined by integral capacity, Thus ϕ^* is integral, and ζ^* is half-integral. By construction, we have $\zeta^*(\delta x) = \phi^*(x^+x^-) \in \mathbf{Z}$. □

On the complexity of tree distances. Here we briefly discuss the complexity of a realization $(\mathit{\Gamma}, \{R_s\}_{s\in S}; \gamma)$ of a tree distance μ. One can check in a strongly polynomial time whether a given weight $\mu : \binom{S}{2} \to \mathbf{Q}_+$ is a tree distance [5, 6]; also see Theorem 5.2 in Section 5. If μ is a tree distance, then one can decompose μ into a nonnegative sum of cut distances for a laminar cut family in a strongly polynomial time [5, Section 4]; in this reference, a laminar family of cuts is called a *compatible family of partial cuts*. It is known that the number of cuts is bounded by $O(|S|^2)$ [5, Remark 4.15]. From this decomposition, one can obtain a tree $\bar{\mathit{\Gamma}} = (V\bar{\mathit{\Gamma}}, E\bar{\mathit{\Gamma}})$, a positive edge-length l, and a family $\{\bar{R}_s\}_{s\in S}$ of subtrees with $\mu(s,t) = d_{\bar{\mathit{\Gamma}},l}(\bar{R}_s, \bar{R}_t)$ $(s,t \in S)$, where $d_{\bar{\mathit{\Gamma}},l}$ denotes the shortest path distance with respect to edge-length l. Here there is a one-to-one correspondence between cuts and edges in $\mathit{\Gamma}$; this is a variation of a tree representation of a laminar family. So the number of edges is bounded by $O(|S|^2)$. Note that these procedure are not fast, compared with existing algorithms for (well-studied) tree metrics. By subdivision, one obtain a realization $(\mathit{\Gamma}, \{R_s\}_{s\in S}; \gamma)$ of μ. The size of $\mathit{\Gamma}$ is not polynomially bounded in general. However one can easily manipulate these subdivided points by a polynomial expression $(\bar{\mathit{\Gamma}}, \{\bar{R}_s\}_{s\in S}, l)$.

Remark 2.2. If μ is an integer-valued tree distance, then μ can be realized by a tree of edge-length $\gamma = 1/2$. One can easily verify it; in a minimal expression $(\bar{\mathit{\Gamma}}, \{\bar{R}_s\}_{s\in S}, l)$, edge-length $l(e)$ can be represented as $l(e) = (\mu(s,t) + \mu(u,r) - \mu(s,u) - \mu(t,r))/2$ for (not necessarily distinct) $s, t, u, r \in S$.

3 Proof of combinatorial duality relation

Here we prove the former part of Theorem 1.1. We consider a natural *continuous relaxation* of the subtree location problem (1.2). Let μ be a tree distance realized by $(\mathit{\Gamma}, \{R_s\}_{s\in S}; \gamma)$. By scaling, we assume $\gamma = 1$. A *geometric realization* $\overline{\mathit{\Gamma}}$ of tree $\mathit{\Gamma}$ is a subset of some Euclidean space $\mathbf{R}^n$ such that for some injective map $\phi : V\mathit{\Gamma} \to \mathbf{R}^n$

(1) $\overline{\Gamma} = \bigcup_{uv \in E\Gamma} [\phi(u), \phi(v)]$,

(2) $\|\phi(u) - \phi(v)\|_2 = 1$ for $uv \in E\Gamma$, and

(3) for two edges $uv, u'v' \in E\Gamma$, $[\phi(u), \phi(v)] \cap [\phi(u'), \phi(v')] \neq \emptyset$ implies $\{u, v\} \cap \{u', v'\} \neq \emptyset$.

By a *subtree* in $\overline{\Gamma}$ we mean a closed connected subset in $\overline{\Gamma}$. The set of all subtrees in $\overline{\Gamma}$ is denoted by $\mathcal{F}\overline{\Gamma}$. For a subtree R_s in Γ, the corresponding subtree in $\overline{\Gamma}$ is denoted by $\overline{R}_s$, i.e., $\overline{R}_s := \bigcup_{uv \in E\Gamma, u,v \in R_s} [\phi(u), \phi(v)]$. We identify $u \in V\Gamma$ with $\phi(u) \in \overline{\Gamma}$; in particular $V\Gamma \subseteq \overline{\Gamma}$. For two points $p, q \in \overline{\Gamma}$, the distance $d_\Gamma(p, q)$ is the Euclidean length of a unique path connecting p and q in $\overline{\Gamma}$. Also, for two subtrees F, F' in $\overline{\Gamma}$, the Euclidean length of a unique shortest path connecting F and F' is denoted by $d_\Gamma(F, F')$; if $F \cap F' \neq \emptyset$, then $d_\Gamma(F, F') = 0$. The diameter $\mathrm{diam}F$ of a subtree F is the maximum distance of points in F.

Given a network (V, E, S, b, c), consider the following continuous relaxation of (1.2):

$$
\begin{aligned}
(3.1) \qquad \text{Min.} \quad & \sum_{y \in V} b(y)\mathrm{diam}F(y) + \sum_{xy \in E} c(xy) d_\Gamma(F(x), F(y)) \\
\text{s.t.} \quad & F : VG \to \mathcal{F}\overline{\Gamma}, \\
& F(s) \cap \overline{R}_s \neq \emptyset \quad (s \in S).
\end{aligned}
$$

The goal of this section is to prove the following, that implies the former part of the main theorem.

Proposition 3.1. *The optimal values of three problems* (1.2), (2.2), *and* (3.1) *are same.*

Before the proof, we need further notions. A *leaf* of subtree F is a point u in F such that there are no $p, q \in F \setminus u$ with $d_\Gamma(p, q) = d_\Gamma(p, u) + d_\Gamma(u, q)$. Clearly the diameter $\mathrm{diam}F$ is attained by a pair of leaves. We call it a *maximum distant pair*. For a subtree F, let p_F be the midpoint of a unique path in F connecting a maximum distant pair u, v. We remark that the point p_F is determined independently on the choice of a maximum distant pair. Indeed, let p'_F be the midpoint of another maximum distant pair u', v'. If $p'_F \neq p_F$, then for some $(w, w') \in \{u, v\} \times \{u', v'\}$ we have $d_\Gamma(w, w') = d_\Gamma(w, p_F) + d_\Gamma(p_F, p'_F) + d_\Gamma(p'_F, w') > \mathrm{diam}F$; a contradiction. We call p_F the *center* of F. We will often use the following obvious property:

(3.2) Suppose $\mathrm{diam}F > 0$. A pair u, v of leaves in F is maximum distant if and only if $d_\Gamma(u, p_F) = d_\Gamma(v, p_F) = \mathrm{diam}F/2$, and u and v belong to distinct components of $\overline{\Gamma} \setminus p_F$.

For a positive integer k, a $1/k$-*point* p is a point in $\overline{\Gamma}$ represented as $l/ku + (k - l)/kv$ for some edge $uv \in E\Gamma$ and a nonnegative integer $l \leq k$. Let $V^k\Gamma$ denote the set of all $1/k$-points; in particular $V\Gamma = V^1\Gamma$. A subtree map $F : V \to \mathcal{F}\overline{\Gamma}$ is said to be $1/k$-*integral* if for each $x \in V$ each leaf of $F(x)$ belongs to $V^k\Gamma$.

3.1 Subtree lemma

The first step is to establish the relationship between LP-dual (2.2) and continuous subtree location (3.1). We recast the inequalities in (2.2):

$$
\begin{aligned}
(3.3) \qquad & d(xy) + d(yz) - d(xz) + h(y) \geq 0 \quad (xz \in E_V, y \in V \setminus \{x, z\}), \\
& d(st) + h(s) + h(t) \geq \mu(s, t) \quad (st \in E_S).
\end{aligned}
$$

For a subtree map $F : V \to \mathcal{F}\overline{\Gamma}$, define $(h^F, d^F) \in \mathbf{R}_+^V \times \mathbf{R}_+^{E_V}$ by

$$\begin{aligned} h^F(y) &= \mathrm{diam}F(y) \quad (y \in V), \\ d^F(xy) &= d_\Gamma(F(x), F(y)) \quad (xy \in E_V). \end{aligned} \tag{3.4}$$

The next lemma immediately implies the equivalence between (2.2) and (3.1).

Lemma 3.2. *Let $F : V \to \mathcal{F}\overline{\Gamma}$ be a subtree map satisfying $R_s \cap F(s) \neq \emptyset$ for each $s \in S$. Then (h^F, d^F) satisfies* (3.3)*. Conversely, for $(h, d) \in \mathbf{R}_+^V \times \mathbf{R}_+^{E_V}$ satisfying* (3.3)*, there exists a subtree map $F : V \to \mathcal{F}\overline{\Gamma}$ such that*

(1) $\mathrm{diam}F(y) \leq h(y)$ *for* $y \in V$,

(2) $d_\Gamma(F(x), F(y)) \leq d(xy)$ *for* $xy \in E_V$,

(3) $R_s \cap F(s) \neq \emptyset$ *for* $s \in S$, *and*

(4) *if* (h, d) *is* $1/k$*-integral, then* F *is* $1/k$*-integral.*

Moreover such a subtree map can be found in strongly polynomial time.

Proof. First we verify that (h^F, d^F) fulfills (3.3). Let $xz \in E_V$ and $y \in V \setminus \{x, z\}$. Take $(p_x, p_y) \in F(x) \times F(y)$ with $d_\Gamma(p_x, p_y) = d_\Gamma(F(x), F(y))$ and $(p_z, p'_y) \in F(z) \times F(y)$ with $d_\Gamma(p_z, p'_y) = d_\Gamma(F(z), F(y))$. Obviously $d_\Gamma(p_x, p_y) + d_\Gamma(p_y, p'_y) + d_\Gamma(p'_y, p_z) \geq d_\Gamma(p_x, p_z) \geq d_\Gamma(F(x), F(z)) = d^F(xz)$. Since $d_\Gamma(p_y, p'_y) \leq \mathrm{diam}F(y) = h^F(y)$, we have $d^F(xy) + h^F(y) + d^F(yz) \geq d^F(xz)$. Let $st \in E_S$. Take $p_s \in F(s) \cap R_s$, $p_t \in F(t) \cap R_t$, and $(q_s, q_t) \in F(s) \times F(t)$ with $d_\Gamma(q_s, q_t) = d^F(st)$. Then $\mathrm{diam}F(s) + d^F(st) + \mathrm{diam}F(t) \geq d_\Gamma(p_s, q_s) + d_\Gamma(q_s, q_t) + d_\Gamma(q_t, p_t) \geq d_\Gamma(p_s, p_t) \geq d_\Gamma(R_s, R_t) = \mu(s, t)$.

Next we show the latter part. For a subtree R and a nonnegative real r, let $B(R, r) = \{p \in \overline{\Gamma} \mid d_\Gamma(R, p) \leq r\}$, i.e., $B(R, r)$ is the closed ball around R with radius r. Let $V = \{x_1, x_2, \ldots, x_n\}$. We recursively define $F(x_j)$ by an subtree H satisfying following properties:

(i) $H \cap B(F(x_k), d(x_k x_j)) \neq \emptyset$ for $k = 1, 2, \ldots, j-1$.

(ii) $H \cap B(R_s, r^s_{x_j}) \neq \emptyset$ for $s \in S$, where $r^s_{x_j} := \begin{cases} h(s) + d(sx_j) & \text{if } x_j \neq s, \\ 0 & \text{if } x_j = s. \end{cases}$

(iii) $\mathrm{diam}H \leq h(x_j)$.

We show that such a subtree does exist for $j = 1, 2, \ldots, n$. Then the resulting $F : V \to \mathcal{F}\overline{\Gamma}$ fulfills (1-3). Indeed, $F(x_j) \cap B(F(x_k), d(x_k x_j)) \neq \emptyset$ implies $d_\Gamma(F(x_k), F(x_j)) \leq d(x_k x_j)$. Also, by definition of $r^s_{x_j}$ in (ii), $R_s \cap F(s) \neq \emptyset$ holds.

Let $H_k := B(F(x_k), d(x_k x_j))$ for $k = 1, 2, \ldots, j-1$, and let $H_s := B(R_s, r^s_{x_j})$ for $s \in S$. Consider the intersection graph $\mathcal{I} = (V\mathcal{I}, E\mathcal{I})$ of subtrees H_u ($u \in \{1, 2, \ldots, j-1\} \cup S$), i.e., $V\mathcal{I} := \{1, 2, \ldots, j-1\} \cup S$ and $E\mathcal{I} := \{uv \mid H_u \cap H_v \neq \emptyset\}$. Consider the set $\mathcal{K} \subseteq 2^{V\mathcal{I}}$ of all maximal cliques of $\mathcal{I}$. By the Helly property of subtrees, $H_K := \bigcap_{u \in K} H_u$ is nonempty for all $K \in \mathcal{K}$. By maximality, H_K ($K \in \mathcal{K}$) are disjoint subtrees. Let H be a (uniquely determined) minimal subtree meeting all H_K ($K \in \mathcal{K}$).

We show that this subtree H satisfies (i-iii). The properties (i) and (ii) are obvious. Indeed, for each $u \in \{1, 2, \ldots, j-1\} \cup S$, there is a maximal clique K containing u, and $H_K \subseteq H_u$. Therefore H meets H_u. Next we show (iii) $\mathrm{diam}H \leq h(x_k)$. If $|\mathcal{K}| = 1$, then H is a singleton, and hence this is obvious. Suppose not. The diameter of H is attained by some leaf pair (p, q). By minimality of H, there are $K, K' \in \mathcal{K}$ such that

$H \cap H_K = \{p\}$, $H \cap H_{K'} = \{q\}$, and $\mathrm{diam}H = d_\Gamma(H_K, H_{K'}) = d_\Gamma(p,q)$. Since each H_K is the intersection of subtrees H_u ($u \in K$), there is $(u,v) \in K \times K'$ with $(p,q) \in H_u \times H_v$ with $d_\Gamma(H_u, H_v) = d_\Gamma(p,q) = \mathrm{diam}H$. We verify $d_\Gamma(H_u, H_v) \leq h(x_k)$ for three cases: (a) $(u,v) = (i,k)$ for $1 \leq i < k \leq j-1$, (b) $(u,v) = (i,s)$ for $1 \leq i \leq j-1$ and $s \in S$, and (c) $(u,v) = (s,t)$ for $s,t \in S$. Here, recall that H_u is the closed ball around $F(x_k)$ (resp. R_s) with radius $d(x_k x_j)$ (resp. $r^s_{x_j}$). For case (a), we have $d_\Gamma(H_u, H_v) = d_\Gamma(F(x_i), F(x_k)) - d(x_i x_j) - d(x_k x_j) \leq d(x_i x_k) - d(x_i x_j) - d(x_k x_j) \leq h(x_j)$, where we use induction with property (i) on the second inequality. Suppose case (b). If $s \neq x_j$, then $d_\Gamma(H_u, H_v) = d_\Gamma(F(x_i), B(R_s, h(s))) - d(x_i x_j) - d(s x_j) \leq d(x_i s) - d(x_i x_j) - d(s x_j) \leq h(x_j)$, where we use induction with property (ii) on the second inequality. If $s = x_j$, then $d_\Gamma(H_u, H_v) = d_\Gamma(F(x_i), R_s) - d(x_i s) \leq d(x_i s) + h(s) - d(x_i s) \leq h(s) = h(x_j)$. For case (c), if $s \neq x_j$ and $t \neq x_j$, then $d_\Gamma(H_u, H_v) = d_\Gamma(R_s, R_t) - h(s) - h(t) - d(s x_j) - d(t x_j) = \mu(s,t) - h(s) - h(t) - d(s x_j) - d(t x_j) \leq d(st) - d(s x_j) - d(t x_j) \leq h(x_j)$. If $s = x_j$ and $t \neq x_j$, then $d_\Gamma(H_u, H_v) = d_\Gamma(R_s, R_t) - h(t) - d(ts) = \mu(s,t) - h(t) - d(st) \leq h(s) = h(x_j)$.

If (h,d) is $1/k$-integral, then each ball H_v and each intersection H_K are $1/k$-integral, and consequently H is $1/k$-integral. The construction of H can be done in strongly polynomial time. Indeed, we use a polynomial expression $(\bar{\Gamma}, \{\bar{R}_s\}_{s \in S}, l)$ in Section 2. We can subdivide $\bar{\Gamma}$ so that each ball H_v can be regarded as a subtree in a graph theoretical sence, and the size of $\bar{\Gamma}$ is polynomial in $|V|$ and $|S|$. Each H_K can be obtained by considering all subtrees H_v including a fixed edge in $\bar{\Gamma}$. In particular $|\mathcal{K}|$ is polynomially bounded. From this we obtain H in strongly polynomial time. □

3.2 Rounding

The second step is to *round* a map $F : V \to \mathcal{F}\overline{\Gamma}$ into $V \to \mathcal{F}\Gamma$, not increasing the objective value. Now suppose that $F : V \to \mathcal{F}\overline{\Gamma}$ is obtained from *rational* solution $(h,d) \in \mathbf{Q}^V_+ \times \mathbf{Q}^{E_V}_+$ satisfying (3.3) via Lemma 3.2. Then there is a positive integer k such that F is $1/k$-integral.

Our goal is to show the existence of maps $F_i : V \to \mathcal{F}\overline{\Gamma}$ ($i = 1, 2, \ldots, k$) with properties:

(i) F_i is integral for $i = 1, 2, \ldots, k$,

(ii) $d_\Gamma(F(x), F(y)) = \sum_{i=1}^k d_\Gamma(F_i(x), F_i(y))/k$ for each $x, y \in V$, and

(iii) $\mathrm{diam}F(x) = \sum_{i=1}^k \mathrm{diam}F_k(x)/k$ for each $x \in V$.

This means that (h^F, d^F) is a convex combination of (h^{F_i}, d^{F_i}). Therefore, if true, then we can always take an *integral* optimal solution F in (3.1), and get an optimal solution of (1.2) by restricting each $F(x)$ to $F(x) \cap V^1\Gamma$.

We fix an orientation of Γ with property that each node is a source or a sink, i.e., regard Γ as a bipartite graph of bipartition (A,B), and orient each edge from A to B. Let Γ^k be an undirected graph on the set of $1/k$-points with edge set $\{uv \mid d_\Gamma(u,v) = 1/k\}$. Namely Γ^k is the k-subdivision of Γ. We can regard $F : V \to \mathcal{F}\overline{\Gamma}$ as $V \to \mathcal{F}\Gamma^k$. For an edge $uv \in E\Gamma$, if oriented as $\overrightarrow{uv}$, then the subdivided edge $u'v'$ corresponding to segment $[(k-i+1)/ku + (i-1)/kv, (k-i)/ku + i/kv]$ is called the *i-th edge* of uv. For $1 \leq i \leq k$, we define $F_i : V \to \mathcal{F}\Gamma$ as follows. Contract all edges except i-th edges in Γ^k. Then the resulting graph coincides with Γ. So we obtain a map $\varphi_i : V\Gamma^k \to V\Gamma$ by defining $\varphi_i(x)$ to be a contracted node. Obviously $\varphi_i \circ F(x)$ is a subtree in Γ, and the image of a leaf in $F(x)$ is a leaf in $\varphi_i \circ F(x)$. Let $F_i := \varphi_i \circ F$. We show that $F_1, F_2, \ldots, F_k$ fulfill properties (i-iii). (i) is obvious by construction. For (ii) and (iii), we first remark the following property:

(3.5) For a shortest path P connecting u and v in Γ^k, $\varphi_i(P)$ is also a shortest path connecting $\varphi(u)$ and $\varphi(v)$ in Γ.

We verify (ii). If $F(x) \cap F(y) \neq \emptyset$, then $F_i(x) \cap F_i(y) \neq \emptyset$, and thus (ii) is obvious. Therefore $F(x) \cap F(y) = \emptyset$. Take a unique shortest path P connecting $F(x)$ and $F(y)$. Then $\varphi_i(P)$ is also shortest $F_i(x)$ and $F_i(y)$, and its length $d_\Gamma(F_i(x), F_i(y))$ is the number of i-th edges in P. From this we have (ii).

Next we verify (iii). For $x \in V$, take a maximum distant pair (u, u') of leaves in $F(x)$. It suffices to show

$$d_\Gamma(\varphi_i(u), \varphi_i(u')) = \mathrm{diam} F_i(x) \quad (i = 1, 2, \ldots, k). \tag{3.6}$$

Namely $(\varphi_i(u), \varphi_i(u'))$ is again a maximum distant pair in $F_i(x)$. If true, then we can apply the argument above to a shortest path connecting u and u'. In the sequel d_Γ is simply denoted by d. Take the center $p = p_{F(x)}$ of $F(x)$ and recall (3.2). Take an edge $qq' \in E\Gamma$ such that segment $[q, q']$ contains p. We may assume $d(q, p) \leq d(q', p) < 1$. If $\mathrm{diam} F(x)/2 \leq d(q, p)$, then $F(x)$ is a path in $[q, q']$, and hence (3.6) is obvious. Suppose $\mathrm{diam} F(x)/2 > d(q, p)$. For a leaf v in $F(x)$, let v_* denote its nearest (integral) point in $F(x) \cap V^1\Gamma$. Then $d(\varphi_i(v), q) = d(v_*, q) + 1$ if $[v, v_*]$ contains an i-th edge and $d(\varphi_i(v), q) = d(v_*, q)$ otherwise. Let K' be the connected component of $\overline{\Gamma} \setminus p$ containing q', and let $K = F(x) \setminus K'$. We may assume $(u, u') \in K \times K'$. Let $r = d(u_*, q)$. Take any leaf $v (\neq p)$ in K. Obviously $d(v_*, q) \leq r$. Recall the orientation and the numbering of edges; nodes at distance r from q are either all sources or all sinks. Thereofore, if $d(v_*, q) = r$ and $[v, v_*]$ contains an i-th edge, then so does $[u, u_*]$ by $d(v, v_*) \leq d(u, u_*)$. Thus we have $(*1)$ $d(\varphi_i(v), q) \leq d(\varphi_i(u), q)$. Similarly we have $(*2)$ $d(\varphi_i(v'), q) \leq d(\varphi_i(u'), q)$ for any leaf v' of $F(x)$ in K'. Let $r' = d(u'_*, q)$. Then $r \leq r' \leq r + 1$. Let (w, w') be an arbitrary pair of leaves in $F(x)$. It suffices to show $d(\varphi_i(w), \varphi_i(w')) \leq d(\varphi_i(u), \varphi_i(u'))$. If $(w, w') \in K \times K'$, then by $(*1\text{-}2)$ we have $d(\varphi_i(w), \varphi_i(w')) = d(\varphi_i(w), q) + d(\varphi_i(w'), q) \leq d(\varphi_i(u), q) + d(\varphi_i(u'), q) = d(\varphi_i(u), \varphi_i(u'))$. If $(w, w') \in K' \times K'$ (necessarily $r' \geq 1$), then $d(\varphi_i(w), \varphi_i(w')) \leq d(\varphi_i(w), q) + d(\varphi_i(w'), q) - 2 \leq 2d(\varphi_i(u'), q) - 2 \leq d(\varphi_i(u), q) + d(\varphi_i(u'), q) \leq d(\varphi_i(u), \varphi_i(u'))$, where we use $d(\varphi_i(u'), q) \leq r' + 1 \leq r + 2 \leq d(\varphi_i(u), q) + 2$. Suppose $(w, w') \in K \times K$. Then we have $d(\varphi_i(w), \varphi_i(w')) \leq d(\varphi_i(w), q) + d(\varphi_i(w'), q) \leq 2d(\varphi_i(u), q)$. Here suppose $r = r' - 1$. Then $d(\varphi_i(u), q) \leq r + 1 = r' \leq d(\varphi_i(u'), q)$. Suppose $r = r'$; u_* and u'_* are either both sinks or both sources. By $d(u, u_*) \leq d(u', u'_*)$, if $[u, u_*]$ contains an i-th edge, then so does $[u', u'_*]$. Thus we have $d(\varphi_i(u), q) \leq d(\varphi_i(u'), q)$. Consequently we have $d(\varphi_i(w), \varphi_i(w')) \leq d(\varphi_i(u), \varphi_i(u'))$. The proof is complete.

Corollary 3.3. *Suppose that μ is a tree distance on S realized by a tree of unit edge-length $\gamma = 1$. For any set V containing S, the following polyhedron is integral:*

$$\{(h, d) \in \mathbf{R}_+^V \times \mathbf{R}_+^{E_V} \mid (h, d) \text{ satisfies } (3.3)\} + \mathbf{R}_+^V \times \mathbf{R}_+^{E_V}.$$

Corollary 3.4. *There exists a strongly polynomial time algorithm to find an optimal subtree map in* (1.2).

Proof. Solving LP (2.2) by Tardos' method [29], we obtain an optimal minimal extreme solution (h, d). We may assume that $\gamma = 1$. Then (h, d) is integral by the previous corollary. By Lemma 3.2, we obtain an integral optimal subtree map $F : V \to \mathcal{F}\overline{\Gamma}$. These constructions can be done in strongly polynomial time. □

4 Half-integral minimum cost multiflows

In this section, we prove a minimum cost version of the half-integrality assertion in Theorem 1.1. Here a network (V, E, S, b, c) is supposed to be node-only-capacitated ($c \to +\infty$); the edge-capacity can be converted into node-capacity by subdividing each edge e and by defining the node-capacity of the subdivided node by $c(e)$. Further we are given a nonnegative *node-cost* $a : V \to \mathbf{Q}_+$. For a multiflow $f = (\mathcal{P}, \lambda)$, the *cost* $\mathrm{cost}(a, f)$ is defined by $\sum_{x \in V} a(x) \sum_{P \in \mathcal{P}: x \in VP} \lambda(P)$. We consider the following minimum cost multiflow problem:

(4.1) Maximize $\mathrm{val}(\mu, f) - \mathrm{cost}(a, f)$ over all multiflows f in (V, E, S, b, c).

The goal of this section is to prove the following:

Theorem 4.1. *Let μ be a tree distance on S. For any network (V, E, S, b, c) and any node-cost $a : V \to \mathbf{Q}_+$, there exists a half-integral optimal solution in* (4.1).

By letting $\mu \leftarrow p\mu$ for a sufficiently large positive $p > 0$, any optimal solution in (4.1) is also optimal solution in (1.1) having the smallest cost.

For simplicity, we further assume that

(4.2) each terminal s is incident to only one inner node, and has a sufficiently large node-capacity $b(s) \to +\infty$ and zero cost $a(s) = 0$.

We can always make the input network fulfill this assumption. Indeed, for each terminal s, add new inner node s^*, join s^* and s, replace each edge xs incident to s by xs^*, set $(b(s^*), a(s^*)) := (b(s), a(s))$, and reset $(b(s), a(s)) := (+\infty, 0)$. Then the LP-dual of (4.1) is given by

$$
\begin{aligned}
(4.3)\ \text{Min.} \quad & \sum_{y \in V \setminus S} b(y)\eta(y) \\
\text{s.t.} \quad & d(xy) + d(yz) - d(xz) + \eta(y) + a(y) \geq 0 \quad (xz \in E_V, y \in V \setminus \{x, z\}), \\
& d(st) + \eta(s) + a(s) + \eta(t) + a(t) \geq \mu(s, t) \quad (st \in E_S), \\
& \eta(s) = 0 \ (s \in S),\ d(xy) = 0 \ (xy \in E),\ (\eta, d) \in \mathbf{R}_+^V \times \mathbf{R}_+^{E_V}.
\end{aligned}
$$

Changing variable η by $h := \eta + a$, one obtain

$$
\begin{aligned}
(4.4)\ \text{Min.} \quad & \sum_{y \in V \setminus S} b(y) \max\{0, h(y) - a(y)\} \\
\text{s.t.} \quad & d(xy) + d(yz) - d(xz) + h(y) \geq 0 \quad (xz \in E_V, y \in V \setminus \{x, z\}), \\
& d(st) + h(s) + h(t) \geq \mu(s, t) \quad (st \in E_S), \\
& h(s) = 0 \ (s \in S),\ d(xy) = 0 \ (xy \in E),\ (h, d) \in \mathbf{R}_+^V \times \mathbf{R}_+^{E_V}.
\end{aligned}
$$

Suppose that μ is realized by $(\Gamma, \{R_s\}_{s \in S}; 1)$ (by scaling). Let $\overline{\Gamma}$ be a geometric realization of Γ. Note that $\min\{0, h(y) - a(y)\}$ is a monotone nondecreasing function on $h(y)$. Therefore, by subtree lemma (Lemma 3.2), this LP reduces to a convex-cost continuous subtree location:

$$
\begin{aligned}
(4.5)\ \text{Min.} \quad & \sum_{y \in V \setminus S} b(y) \max\{0, \mathrm{diam} F(y) - a(y)\} \\
\text{s.t.} \quad & F : V \to \mathcal{F}\overline{\Gamma}, \\
& F(x) \cap F(y) \neq \emptyset \quad (xy \in E), \\
& F(s) \text{ is a single point in } R_s \quad (s \in S).
\end{aligned}
$$

Our analysis is based on this problem. We need some notation. For a multiflow $f = (\mathcal{P}, \lambda)$, the *flow-support* $\zeta^f : E \to \mathbf{R}_+$ is defined by $\zeta^f(e) := \sum\{\lambda(P) \mid P \in \mathcal{P}, e \in EP\}$ for $e \in E$. Recall the notation δy (the set of edges incident to node y). The following obvious relations are useful for us:

$$\zeta^f(\delta y) = 2\sum\{\lambda(P) \mid P \in \mathcal{P}, y \in VP\} \quad (y \in V \setminus S),$$
$$\sum_{y \in V \setminus S} \mathrm{diam}F(y)\zeta^f(\delta y)/2 = \sum_{P \in \mathcal{P}} \lambda(P)(\mathrm{diam}F)(VP),$$

where $(\mathrm{diam}F)(VP) = \sum_{x \in VP} \mathrm{diam}F(x)$ and we use $\mathrm{diam}F(s) = 0$ $(s \in S)$ in the second equation.

We derive an optimality criterion of primal-dual type. For a multiflow $f = (\mathcal{P}, \lambda)$ and a subtree map F feasible to (4.5), the duality gap is given by

$$\begin{aligned} &\sum_{y \in V \setminus S} b(y) \max\{0, \mathrm{diam}F(y) - a(y)\} - (\mathrm{val}(\mu, f) - \mathrm{cost}(a, f)) \\ = &\sum_{x \in V \setminus S} b(y) \max\{0, \mathrm{diam}F(y) - a(y)\} - \zeta^f(\delta y)(\mathrm{diam}F(y) - a(y))/2 \\ &\quad + \sum_{P \in \mathcal{P}} \lambda(P)\{(\mathrm{diam}F)(VP) - d_\Gamma(R_{s_P}, R_{t_P})\}. \end{aligned}$$

Here we note $(\mathrm{diam}F)(VP) \geq d_\Gamma(R_{s_P}, R_{t_P})$. Indeed, let $P = (x_0, x_1, \ldots, x_m)$ be an S-path. We can take points $p_i \in F^*(x_i) \cap F^*(x_{i+1})$ for $i = 0, 1, \ldots, m-1$ since $F^*(x_i) \cap F^*(x_{i+1})$ is nonempty. Then we have

$$(\mathrm{diam}F)(VP) = \sum_{i=1}^{m-1} \mathrm{diam}F(x_i) \geq \sum_{i=1}^{m-1} d_\Gamma(p_{i-1}, p_i) \geq d_\Gamma(R_s, R_t). \tag{4.6}$$

An S-path P is said be *F-shortest* if $(\mathrm{diam}F)(VP) = d_\Gamma(R_{s_P}, R_{t_P})$. Thus we have:

Lemma 4.2. *A multiflow f and a feasible subtree map F are both optimal if and only if*

(1) $\zeta^f(\delta y) = \begin{cases} 2b(y) & \text{if } \mathrm{diam}F(y) > a(y) \\ 0 & \text{if } \mathrm{diam}F(y) < a(y) \end{cases}$ *for $y \in V \setminus S$, and*

(2) *each path in f with positive flow-value is F-shortest.*

Let F^* be an optimal subtree map in (4.5). In the sequel we construct an edge-weight $\zeta : E \to \mathbf{R}_+$ that can be represented as $\zeta = \zeta^f$ for some multiflow f satisfying (1-2). By a small perturbation to a, we may prove Theorem 4.1 under assumption:

$a(y)$ is positive for all inner nodes y. (4.7)

By Lemma 4.2 (1) and this positive cost assumption, there is no flow passing an inner node y with $\mathrm{diam}F^*(y) < a(y)$. Therefore we can delete all such inner nodes. In particular, $\mathrm{diam}F^*(y)$ is positive for all inner nodes y. The shape of $F^*(y)$ gives us the important information of how flow passes through y. The condition (1) in Lemma 4.2 is a b-matching condition for the flow-support ζ^f. We study the condition (2). For this purpose, we need some notions. Let y be an inner node. Recall definitions of the center of a subtree (Section 3). A leaf v of $F^*(y)$ is said to be *admissible* if $d_\Gamma(p_{F^*(y)}, v) = \mathrm{diam}F^*(y)/2$. By (3.2) and (4.6), we get a local criterion for an S-path to be F^*-shortest:

Lemma 4.3. *An S-path $(s = x_0, x_1, x_2, \ldots, x_{m-1}, x_m = t)$ is F^*-shortest if and only if*

(1) *$F^*(x_i) \cap F^*(x_{i+1})$ is an admissible leaf of both $F^*(x_i)$ and $F^*(x_{i+1})$ for $i = 0, 1, 2 \ldots, m-1$,*

(2) *$R_s \cap F^*(x_1) = F^*(s)$, $R_t \cap F^*(x_{m-1}) = F^*(t)$, and*

(3) *$F^*(x_{i-1})$ and $F^*(x_{i+1})$ belong to distinct connected components of $\overline{\Gamma} \setminus p_{F^*(x_i)}$ for $i = 1, 2, \ldots, m-1$.*

In particular, we can delete all edges xy such that $F^*(x) \cap F^*(y)$ is not an admissible leaf of both of $F^*(x)$ and $F^*(y)$; such an edge never has nonzero flow-support. Also we can delete a terminal s and its unique incident edge sx with $R_s \cap F^*(x) \neq F^*(s)$. Motivated by (3), for an inner node y, let $\mathcal{X}^{F^*}(y)$ be the partition of δy defined as: xy and zy belong to the same part in $\mathcal{X}^{F^*}(y)$ if and only if $F^*(x)$ and $F^*(z)$ belong to the same connected component of $\overline{\Gamma} \setminus p_{F^*(y)}$ (well-defined by $\mathrm{diam} F^*(y) > 0$). Now consider the polyhedron $\mathcal{Q}$ formed by $\zeta : E \to \mathbf{R}_+$ satisfying the following inequalities:

$$\zeta(\delta y) = 2b(y) \qquad (y \in V \setminus S : \mathrm{diam} F^*(y) > a(y)), \tag{4.8}$$

$$0 \leq \zeta(\delta y) \leq 2b(y) \qquad (y \in V \setminus S : \mathrm{diam} F^*(y) = a(y)), \tag{4.9}$$

$$\zeta(U) \leq \zeta(\delta y \setminus U) \qquad (y \in V \setminus S,\ U \in \mathcal{X}^{F^*}(y)). \tag{4.10}$$

For every optimal multiflow f, its flow-support ζ^f fulfills these inequalities; in particular $\mathcal{Q}$ is nonempty. Indeed the first and the second correspond to Lemma 4.2 (1) and the capacity constraint, respectively. The third means that flow coming from U must escape into $\delta y \setminus U$ by Lemma 4.3 (3). It will turn out that the converse also holds; see Lemma 4.5 below. Before it, we establish the half-integrality of polyhedron $\mathcal{Q}$:

Lemma 4.4. *Every extreme point ζ^* of $\mathcal{Q}$ is half-integral, and moreover $\zeta^*(\delta x)$ is integral for $x \in V \setminus S$.*

Proof. Let ζ^* be an extreme point of $\mathcal{Q}$. We adapt the idea of Pap [24] expressing ζ^* as the projection of an extreme point of a fractional b-matching polyhedron for another graph G' with lower- and upper-bound functions $\underline{b}, \overline{b}$ on its node set. Delete all edges e with $\zeta^*(e) = 0$. For each (inner) node y without active constraint in (4.10), set $(\underline{b}(y), \overline{b}(y)) = (0, 2b(y))$. Take an inner node $y \in V \setminus S$ having active constraint $\zeta^*(U) = \zeta^*(\delta y \setminus U)$ for some $U \in \mathcal{X}^{F^*}(y)$. Then other constraint in (4.10) is never active if $|\mathcal{X}^{F^*}(y)| > 2$; if $|\mathcal{X}^{F^*}(y)| = 2$, then (4.10) is one constraint $\zeta^*(U) = \zeta^*(U^c)$ for $U^c := \delta y \setminus U \in \mathcal{X}^{F^*}(y)$ and is always active. Then split y into two nodes y_U and y_{U^c}. Reconnect each edge in U to y_U, and reconnect each edge in U^c to y_{U^c}. Let $(\underline{b}(y_U), \overline{b}(y_U)) = (\underline{b}(y_{U^c}), \overline{b}(y_{U^c})) = (0, b(y))$. Apply it to all inner nodes having active constraint (4.10). Let G' be the resulting graph with $\underline{b}, \overline{b}$. Then ζ^* can be extended to a fractional b-matching ζ^* in the new graph by setting $\zeta^*(y_U y_{U^c}) = b(y) - \zeta^*(\delta y)/2$ for each inner node y having active constraint in (4.10). Therefore the original ζ^* is an extreme point of the projection of the polyhedron $P(\underline{b}, \overline{b})$ of fractional b-matching defined by integral node-capacity. By construction, the extended ζ is also extreme in $P(\underline{b}, \overline{b})$, and is half-integral by Lemma 2.1. Consequently the original ζ^* is also half-integral, and $\zeta^*(\delta y)$ is integral. □

The next lemma completes the proof of Theorem 4.1.

Lemma 4.5. *For every extreme point ζ^* of $\mathcal{Q}$, there exists a half-integral multiflow f^* such that $\zeta^{f^*} = \zeta^*$ and each path in f^* is F^*-shortest. Moreover such a multiflow can be found in strongly polynomial time.*

Proof. We can construct such a multiflow in a greedy way. Delete all edges e with $\zeta^*(e) = 0$. Take a terminal s and its unique incident edge sx with $\zeta^*(sx) > 0$. We extend sx to an S-path as follows. By constraint (4.10), there is a node x_1 incident to x such that $\zeta^*(xx_1) > 0$, and sx and xx_1 belong to distinct parts of $\mathcal{X}^{F^*}(x)$. Here, if some constraint $\zeta^*(U) \leq \zeta^*(\delta x \backslash U)$ in (4.10) is active with $sx \in \delta x \backslash U$, then take xx_1 from U. Extend path (s, x) to (s, x, x_1). Consider node x_1. Suppose that x_1 is not a terminal. Again we can take a node x_2 incident to x_1 such that $\zeta^*(x_1x_2) > 0$, and xx_1 and x_1x_2 belong to distinct parts of $\mathcal{X}^{F^*}(x_1)$, as above. Extend path (s, x, x_1) to (s, x, x_1, x_2), and consider node x_2. Repeat it. By construction, this path induces a sequence of subtrees with properties (1-3) in Lemma 4.3, and thus meets every node at most once. So after $O(|V|)$ steps, we arrive some terminal $t = x_m$ and obtain an F^*-shortest S-path $P_1 = (s, x, x_1, x_2, \ldots, x_m = t)$. Let $\lambda(P_1) := \max\{\lambda \geq 0 \mid (\zeta^* - \lambda\chi_{EP_1}) \text{ is nonnegative and keeps (4.10)}\}$. Then $\lambda(P_1)$ is positive half-integral since $\zeta^*(U) - \zeta^*(\delta y \setminus U)$ is integral by $\zeta^*(\delta y) \in \mathbf{Z}$ (Lemma 4.4). Let $\zeta^* \leftarrow \zeta^* - \lambda(P_1)\chi_{EP_1}$. Again $\zeta^*(\delta y)$ is integral for each inner node y. Repeat this process until there is no edge in G. After that, we obtain a required half-integral multiflow $f^* = (P_1, P_2, \ldots, P_k; \lambda(P_1), \lambda(P_2), \ldots, \lambda(P_k))$. These construction can be done in strongly polynomial time. Indeed, after getting path P_i, the number of edges decreases or some constraint (4.10) for $U \in \mathcal{X}^{F^*}(y)$ becomes active. Once active, it keeps active in the sequent process (as long as U is nonempty). □

Corollary 4.6. *There exists a strongly polynomial time algorithm to find a half-integral optimal solution in* (4.1).

Proof. Solving (4.4) by Tardos' method [29], we obtain an optimal dual solution (h^*, d^*). By subtree lemma (Lemma 3.2), we obtain an optimal subtree map F^*. Next take an extreme point ζ^* of $\mathcal{Q}$ again by Tardos' method. Finally, according to Lemma 4.5, we obtain a half-integral optimal multiflow f^* with $\zeta^{f^*} = \zeta^*$. These construction can be done in strongly polynomial time. □

Remark 4.7. Suppose that μ and a are integer-valued. By the standard TDI-argument, there exists a half-integral optimal dual solution (h, d) in (4.4). By the integrality of μ, we can realize μ by a tree Γ with edge-length $\gamma = 1/2$ (Remark 2.2). By Lemma 3.2, we can take an optimal subtree map F^* such that each leaf of each $F^*(y)$ belongs to $V^1\Gamma$. Thus the maximum value of (4.1) is equal to the minimum value of the following *nonlinear* discrete subtree location problem:

$$
\begin{aligned}
(4.11) \qquad \text{Min.} \quad & \sum_{y \in V \setminus S} b(y) \max\{0, \mathrm{diam} F(y)/2 - a(y)\} \\
\text{s.t.} \quad & F : V \to \mathcal{F}\Gamma, \\
& F(x) \cap F(y) \neq \emptyset \quad (xy \in E), \\
& F(s) \text{ is a single point in } R_s \quad (s \in S).
\end{aligned}
$$

See Section 6 for related arguments.

In the case where Γ is a path. Then $\mathcal{X}^{F^*}(y)$ is a bipartition and (4.10) is always active. From this, one can easily see that the graph G' in the proof of Lemma 4.4 is bipartite. Recall that the fractional b-matching polytope in a bipartite graph coincides with the (integral) b-matching polytope. Consequently any extreme solution ζ^* is integral, and moreover $\zeta^*(\delta y)$ is even for each inner node y. Thus we obtain an integral optimal multiflow. This explains max-flow min-cut theorem in Example 1 in the introduction.

Corollary 4.8. *Suppose that μ is a tree distance realized by a path. Then there exists an integral optimal multiflow in* (4.1).

However this is not a *fully* multiflow theorem. In fact there is a reduction to a minimum cost circulation problem. We omit it here; see [10] for an essence.

5 Unbounded fractionality

The goal of this section is to prove that the set of tree distances is the only class admitting half-integrality theorem:

Theorem 5.1. *Suppose that* $\mu : \binom{S}{2} \to \mathbf{Q}_+$ *is not a tree distance. Then there is no positive integer* k *such that* (1.1) *has a* $1/k$*-integral optimal multiflow for every network* (V, E, S, b, c).

Let k be an arbitrary positive integer. We construct a network in which its unique optimal multiflow is $1/k$-integral. Our construction is inspired by [13]; also see [15, section 3]. Let G_k be the grid graph of size k. Namely its node set consists of $x_{i,j}$ and edge set consists of $x_{i,j}x_{i+1,j}$ and $x_{i,j}x_{i,j+1}$ for $1 \le i, j \le k$. Add nodes $s, t, s', t', \bar{s}, \bar{t}, \bar{s}', \bar{t}'$ and edges $s\bar{s}, t\bar{t}, s'\bar{s}', t'\bar{t}'$, $\bar{s}x_{1,1}, \bar{s}x_{2,1}, \ldots, \bar{s}x_{k,1}$, $\bar{t}x_{1,k}, \bar{t}x_{2,k}, \ldots, \bar{t}x_{k,k}$, $\bar{s}'x_{1,1}, \bar{s}'x_{1,2}, \ldots, \bar{s}'x_{1,k}$, $\bar{t}'x_{k,1}, \bar{t}'x_{k,2}, \ldots, \bar{t}'x_{k,k}$. Let (V, E) be the resulting graph. Node-capacity b is defined as $b(x_{i,j}) = 1$, $b(\bar{s}) = b(\bar{t}) = 1$, $b(\bar{s}') = b(\bar{t}') = k - 1$, and $b(s) = b(t) = b(s') = b(t') = +\infty$. Edge-capacity c is sufficiently large. A terminal set is defined by $\{s, t, s', t'\}$. Let $(V, E, \{s, t, s', t'\}, b, c)$ be the resulting network. Consider the following (s, t)-paths P_i and (s', t')-paths Q_j:

$$\begin{aligned} P_i &= (s, \bar{s}, x_{i,1}, x_{i,2}, \ldots, x_{i,k}, \bar{t}, t) \quad (1 \le i \le k), \\ Q_j &= (s', \bar{s}', x_{1,j}, x_{2,j}, \ldots, x_{k,j}, \bar{t}', t') \quad (1 \le j \le k). \end{aligned}$$

Let $\mathcal{P} = \{P_1, P_2, \ldots, P_k, Q_1, Q_2, \ldots, Q_k\}$. The flow-value function λ is defined as $\lambda(P_i) = 1/k$ and $\lambda(Q_j) = (k-1)/k$ for each i, j. Then one can easily see:

(5.1) $f = (\mathcal{P}, \lambda)$ is a unique multiflow with properties that

(i) it consists of (s, t)-paths and (s', t')-paths, and

(ii) nodes $\bar{s}, \bar{t}, \bar{s}', \bar{t}'$ are saturated.

In particular, there is no positive integer k that every 2-commodity flow maximization has a $1/k$-integral optimal multiflow.

Now consider an arbitrary weight μ that is not tree distance. Here we use the following 4-point characterization of tree distances.

Theorem 5.2 ([6])**.** $\mu : \binom{S}{2} \to \mathbf{Q}_+$ *is a tree distance if and only if there is no* 4*-element subset* $\{s, t, s', t'\}$ *in* S *such that*

(5.2)
$$\begin{aligned} &\mu(s,t) + \mu(s',t') > \\ &\max \left\{ \begin{array}{c} \mu(s,t') + \mu(t,s'),\ \mu(s,s') + \mu(t,t'),\ \mu(s,t),\ \mu(s',t'), \\ \dfrac{\mu(t,s') + \mu(s',t') + \mu(t',t)}{2},\ \dfrac{\mu(s,s') + \mu(s',t') + \mu(t',s)}{2}, \\ \dfrac{\mu(s,t) + \mu(t,t') + \mu(t',s)}{2},\ \dfrac{\mu(s,t) + \mu(t,s') + \mu(s',s)}{2} \end{array} \right\}. \end{aligned}$$

So we may assume that S has a 4-element set $\{s, t, s', t'\}$ satisfying (5.2). Adding isolated terminals $S \setminus \{s, t, s', t'\}$ to $(V, E, \{s, t, s', t'\}, b, c)$, consider μ-weighted maximum

multiflow problem (1.1) on (V, E, S, c, b). We show that the above-constructed multiflow $f = (\mathcal{P}, \lambda)$ is a unique optimum in this problem.

Contract all nodes in grid G_k into one node g, delete loops, and identify multiple edges. Set $b(g) = +\infty$. Then we obtain a new network of 9 nodes $g, s, t, s', t', \bar{s}, \bar{t}, \bar{s}', \bar{t}'$ and of 6 edges $s\bar{s}, t\bar{t}, s'\bar{s}', t'\bar{t}', g\bar{s}, g\bar{t}, g\bar{s}', g\bar{t}'$. There are only six S-paths in the new network. So any multiflow f can be represented as a 6-tuple $(f_{st}, f_{s't'}, f_{ss'}, f_{st'}, f_{ts'}, f_{tt'})$, where f_{uv} is the flow-value of a unique (u, v)-path for $u, v \in \{s, t, s', t'\}$. According to this contraction, the above-constructed multiflow f is transformed into a multiflow $f = (f_{st}, f_{s't'}, f_{ss'}, f_{st'}, f_{ts'}, f_{tt'}) = (1, k-1, 0, 0, 0, 0)$ in the new network. By (5.1), it suffices to show that f is a unique optimum in the new network. Take an arbitrary optimal multiflow f' in the new network. Then $\min(f'_{st'}, f'_{ts'}) > 0$ is impossible. Otherwise, for small $\epsilon > 0$, decrease $f'_{st'}$ and $f'_{ts'}$ by ϵ and increase f'_{st} and $f'_{s't'}$ by ϵ. By $\mu(s,t) + \mu(s',t') > \mu(s,t') + \mu(t,s')$, the flow-value strictly increases. This contradicts to the optimality. Similarly $\min(f'_{ss'}, f'_{tt'}) = 0$. Also $\min(f'_{ss'}, f'_{st'}) > 0$ is impossible. Otherwise, for small $\epsilon > 0$, decrease $f'_{st'}$ and $f'_{ss'}$ by ϵ and increase $f'_{s't'}$ by ϵ, and increase f'_{st} by 2ϵ; $\bar{t}$ is not saturated by $f'_{ts'} = f'_{tt'} = 0$. By $\mu(s,t) + \mu(s',t') > (\mu(s,t') + \mu(t',s') + \mu(s,s'))/2$, the flow-value increases. A contradiction. Consequently, at most one of $f'_{ss'}, f'_{st'}, f'_{ts'}, f'_{tt'}$ is positive. Suppose, say, $f'_{ss'} > 0$. Then both $\bar{t}$ and $\bar{t}'$ are not saturated. For small $\epsilon > 0$, decrease $f'_{ss'}$ by ϵ, and increase f'_{st} and $f'_{s't'}$ by ϵ. By $\mu(s,t) + \mu(s',t') > \mu(s,s') + \mu(t,t') \geq \mu(s,s')$, the flow-value increases. A contradiction. Moreover both $\mu(s,t)$ and $\mu(s',t')$ are positive by $\mu(s,t)+\mu(s',t') > \max(\mu(s,t), \mu(s',t'))$. So $\bar{s}, \bar{t}, \bar{s}', \bar{t}'$ are saturated by (s,t)-flows and (s',t')-flows. Consequently $f' = f$.

Remark 5.3. One can also prove that if μ is not a tree distance, then there is no positive integer k such that the minimum cost μ-weighted maximum multiflow problem has a $1/k$-integral optimal solution for every edge-only-capacitated network and every edge-cost. Indeed, take 4-element set $\{s, t, s', t'\} \subseteq S$ with property (5.2). Consider the network depicted in [13, Figure 3] (or [15, Figure 4]) with replacing $(1, 1', 2, 2')$ by (s, t, s', t'). By the same idea as above, one can show that the solution given in [13, p. 79] is a unique $1/k$-integral optimal solution.

6 Concluding remarks

In concluding, we discuss some related topics including future research directions.

Convex-cost multiflows. Our method is applicable to the following class of *convex-cost* multiflow problems. Let (V, E) be an undirected graph with terminal set S. Instead of node- and edge-capacities, we are given monotone nondecreasing convex functions $g_x : \mathbf{R}_+ \to \mathbf{R} \cup \{+\infty\}$ $(x \in V)$ and $h_e : \mathbf{R}_+ \to \mathbf{R} \cup \{+\infty\}$ $(e \in E)$. We assume $g_x(0) < +\infty$ $(x \in V)$ and $h_e(0) < +\infty$ $(e \in E)$. A multiflow $f = (\mathcal{P}, \lambda)$ is a pair of a set $\mathcal{P}$ of S-paths and a flow-value function $\lambda : \mathcal{P} \to \mathbf{R}_+$ satisfying

$$g_x(\zeta^f(x)) < +\infty \quad (x \in V), \quad h_e(\zeta^f(e)) < +\infty \quad (e \in E),$$

where $\zeta^f : V \cup E \to \mathbf{R}_+$ is the flow-support defined by $\zeta^f(u) = \sum\{\lambda(P) \mid P \in \mathcal{P}, P \text{ contains } u\}$ for $u \in V \cup E$. Let μ be a terminal weight. Consider the following problem:

$$\text{(6.1)} \qquad \text{Maximize } \mathrm{val}(\mu, f) - \sum_{x \in V} g_x(\zeta^f(x)) - \sum_{e \in E} h_e(\zeta^f(e)) \text{ over all multiflows } f.$$

Suppose that μ is a tree distance realized by $(\Gamma, \{R_s\}_{s \in S}; 1)$. Let $\overline{\Gamma}$ be a geometric realization of Γ. Recall the Fenchel duality theory in convex cost network flows [11, 25].

For a function $\varphi : \mathbf{R}_+ \to \mathbf{R} \cup \{+\infty\}$, the *conjugate function* $\varphi^* : \mathbf{R}_+ \to \mathbf{R} \cup \{+\infty\}$ is defined by

$$\varphi^*(p) := \sup_{\zeta \in \mathbf{R}_+} \{p\zeta - \varphi(\zeta)\} \quad (p \in \mathbf{R}_+).$$

If φ is nondecreasing convex, then so does φ^*. By the Fenchel duality together with subtree lemma (Lemma 3.2), we obtain the following dual problem:

$$\begin{aligned} (6.2) \qquad \text{Min.} \quad & \sum_{x \in V} g_x^*(\mathrm{diam}F(x)) + \sum_{xy \in E} h_{xy}^*(d_\Gamma(F(x), F(y))) \\ \text{s.t.} \quad & F : V \to \mathcal{F}\overline{\Gamma}, \\ & F(s) \cap \overline{R_s} \neq \emptyset \quad (s \in S). \end{aligned}$$

Under a mild condition on convex functions g_x, h_e, e.g., they are closed proper convex, the maximum value of (6.1) is equal to the minimum value of (6.2). Moreover, if all g_x, h_e are piecewise linear and their non-differentiable points are all integral, then there exists a half-integral integral multiflow in (6.1). To prove it, replace Lemma 4.2 (1) by a general kilta condition $g_x(\zeta^f(x)) + g_x^*(\mathrm{diam}F(x)) = \zeta^f(x)\mathrm{diam}F(x)$, and replace (4.8) and (4.9) by b-matching conditions corresponding to a vertical segment and a horizontal segment in the characteristic curve $\{(\zeta, h) \mid g_x(\zeta) + g_x^*(h) = \zeta h\}$. The remaining argument is the same as in Section 4, and an algorithm can also be obtained by the same way. Moreover, if each gradient of g_x, h_e is integer-valued, then (6.2) is discretized into a nonlinear discrete subtree location problem, as in Remark 4.7. A detailed verification is left to readers.

Toward combinatorial polynomial algorithms. Our algorithm for problems (1.1), (1.2), (4.1) relies on a general LP solver (ellipsoid method or interior point method). So it is challenge to find a *purely combinatorial* polynomial time algorithm. Goldberg and Karzanov [3] developed a combinatorial weakly polynomial time algorithm for the edge-only-capacitated minimum cost maximum free multiflow problem. Also Babenko and Karzanov [1] presented a combinatorial weakly polynomial time algorithm for the node-capacitated maximum free multiflow problem. One possible approach is to extend their algorithms. Our subtree location model might be some help to design an algorithm of a simpler dual description, such as: *Find an augmenting path in an auxiliary graph. If fails, then move, expand, or shrink subtrees until there appear new edges in the auxiliary graph.* Nevertheless half-integral flow-augmentation is quite nontrivial.

What is discrete convexity theory for multiflows ? In closing, let us mention a possibility of discrete convex analysis for multiflows. *Discrete convex analysis* [21, 22] is a theory of convex functions defined on integer points, aiming at a unified framework for well-solvable discrete optimizations, including network flows, matroids, and submodular functions. One of the motivations comes from combinatorial properties of convex functions arising from convex cost network (edge-only-capacitated) flows.

Since the single commodity weight μ is realized by a path, the dual (6.2) is a location problem on a path, and therefore it is the minimization of a function defined on the product of paths. The product of paths can be identified with a subset of integer lattice points $\mathbf{Z}^V$. Hence (6.2) can also be regarded as a minimization of a function defined on $\mathbf{Z}^V$. The objective function $g(p) := \sum_{xy \in E} h_{xy}^*(|p(x) - p(y)|)$ fulfills a certain submodular-type condition and is a typical example of *L-convex functions*, a fundamental class of *discrete convex functions* in discrete convex analysis.

It would be interesting to extend the theory of discrete convexity to include multicommodity flows and well-solvable discrete location problems. In general tree distance weights, the dual is the minimization of the function $g(p) := \sum_{xy} h_{xy}^*(d_\Gamma(p(x), p(y)))$

defined on the *product of trees*; the similarity to the L-convex function above is notable. Moreover, in node-capacitated case, we need to treat a minimization of a function defined on a set of subsets in the product of trees (corresponding to the products of subtrees). Probably, discrete convex analysis for multiflows would be a theory of convex functions defined on the product of trees, more generally, graph structures, or polyhedral complexes, beyond integer lattice points in the Euclidean space. Such a theory, if exists, might bring a novel paradigm to discrete optimization. We will pursue this issue in the subsequent papers.

References

[1] M. A. Babenko and A. V. Karzanov, A scaling algorithm for the maximum node-capacitated multiflow problem, ESA (2008), 124-135.

[2] B. V. Cherkasski, A solution of a problem of multicommodity flows in a network, *Ekonomika i Matematicheskie Metody* **13** (1977), 143–151 (in Russian).

[3] A. V. Goldberg and A. V. Karzanov, Scaling methods for finding a maximum free multiflow of minimum cost, *Mathematics of Operations Research* **22** (1997), 90–109.

[4] S. L. Hakimi, E. F. Schmeichel, M. Labbe, On locating path or tree shaped facilities on networks, *Networks* **23** (1993), 543–555.

[5] H. Hirai, A geometric study of the split decomposition, *Discrete and Computational Geometry* **36** (2006), 331-361.

[6] H. Hirai, Characterization of the distance between subtrees of a tree by the associated tight span, *Annals of Combinatorics* **10** (2006), 111–128.

[7] H. Hirai, Tight spans of distances and the dual fractionality of undirected multiflow problems, *Journal of Combinatorial Theory, Series B* **99** (2009), 843–868.

[8] H. Hirai, Folder complexes and multiflow combinatorial dualities, RIMS-preprint 1675, (2009), `http://www.kurims.kyoto-u.ac.jp/preprint/file/RIMS1675.pdf`.

[9] H. Hirai, The maximum multiflow problems with bounded fractionality, RIMS-Preprint 1682, (2009), `http://www.kurims.kyoto-u.ac.jp/preprint/file/RIMS1682.pdf`.

[10] H. Hirai and S. Koichi, in preparation.

[11] M. Iri, *Network Flow, Transportation and Scheduling—Theory and Practice*, Academic Press, New York, 1969.

[12] A. V. Karzanov, A minimum cost maximum multiflow problem, in: *Combinatorial Methods for Flow Problems* (A. V. Karzanov, ed.), Institute for System Studies, Moscow, 1979, pp. 138–156 (In Russian).

[13] A. V. Karzanov, Maximum- and minimum-cost multicommodity flow problems having unbounded fractionality, in: *Selected Topics in Discrete Mathematics* (A. K. Kelmans, ed.), American Mathematical Society Translations Series 2, Volume 158, American Mathematical Society, Providence, 1994, pp. 71–80.

[14] A. V. Karzanov, Minimum cost multiflows in undirected networks, *Mathematical Programming* **66** (1994), 313–325.

[15] A. V. Karzanov, Multiflows and disjoint paths of minimum total cost, *Mathematical Programming* **78** (1997), 219–242.

[16] A. V. Karzanov, Minimum 0-extensions of graph metrics, *European Journal of Combinatorics* **19** (1998), 71–101.

[17] A. V. Karzanov, Metrics with finite sets of primitive extensions, *Annals of Combinatorics* **2** (1998), 211–241.

[18] L. Lovász, On some connectivity properties of Eulerian graphs, *Acta Mathematica Academiae Scientiarum Hungaricae* **28** (1976), 129–138.

[19] W. Mader, Über die Maximalzahl kreuzungsfreier H-Wege, *Archiv der Mathematik* **31** (1978/79), 387–402.

[20] E. Minieka, The optimal location of a path or tree in a tree network, *Networks* **15** (1985), 309–321.

[21] K. Murota, Discrete convex analysis, *Mathematical Programming* **83** (1998), 313–371.

[22] K. Murota, *Discrete Convex Analysis*, SIAM, Philadelphia, 2003.

[23] G. Pap, Some new results on node-capacitated packing of A-paths, *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, 599–604, ACM, New York, 2007.

[24] G. Pap, Strongly polynomial time solvability of integral and half-integral node-capacitated multiflow problems, EGRES Technical Report, TR-2008-12, (2008).

[25] R. T. Rockafellar, *Network Flows and Monotropic Optimization,* John Wiley & Sons, New York, 1984.

[26] C. Semple and M. Steel, *Phylogenetics*, Oxford University Press, Oxford, 2003.

[27] A. Tamir, and T. J. Lowe, The generalized p-forest problem on a tree network, *Networks* **22** (1992), 217–230.

[28] B. C. Tansel, R. L. Francis, and T. J. Lowe, Location on networks: a survey. II, *Management Science* **29** (1983), 498–511.

[29] É. Tardos, A strongly polynomial algorithm to solve combinatorial linear programs, *Operations Research* **34** (1986), 250–256.

[30] V. V. Vazirani, *Approximation Algorithms*, Springer, Berlin, 2001.