Discrete Applied Mathematics 112 (2001) 241–261

DISCRETE APPLIED MATHEMATICS

A comparison of Steiner tree relaxations

Tobias Polzin *, Siavash Vahdati Daneshmand

Theoretische Informatik, Universität Mannheim, 13-17 Raum 006, D-68131 Mannheim, Germany

Received 1 April 1998; revised 22 December 1998; accepted 1 August 2000

Abstract

There are many (mixed) integer programming formulations of the Steiner problem in networks. The corresponding linear programming relaxations are of great interest particularly, but not exclusively, for computing lower bounds; but not much has been known about the relative quality of these relaxations. We compare all classical and some new relaxations from a theoretical point of view with respect to their optimal values. Among other things, we prove that the optimal value of a flow-class relaxation (e.g. the multicommodity flow or the dicut relaxation) cannot be worse than the optimal value of a tree-class relaxation (e.g. degree-constrained spanning tree relaxation) and that the ratio of the corresponding optimal values can be arbitrarily large. Furthermore, we present a new flow-based relaxation, which is to the authors' knowledge the strongest linear relaxation of polynomial size for the Steiner problem in networks.

Keywords: Steiner problem; Relaxation; Lower bound

1. Introduction

The Steiner problem in networks is the problem of connecting together, at minimum cost, a set of required vertices in a weighted graph. This is a classical $\mathcal{NP}$-hard problem (see [11,10]) with many important applications in network design in general and VLSI design in particular. For more background information on this problem, its applications and its algorithmic aspects, we refer the reader to the second part of the book of Hwang et al. [10] on the Steiner problem. The primary goal of this paper is to compare the linear relaxations of all classical, frequently cited and some modified or new integer programming formulations of this problem with respect to their optimal values. We present several new results, establishing very clear relations between relaxations which have often been treated as unrelated or incomparable ones.

* Corresponding author.

E-mail addresses: polzin@informatik.uni-mannheim.de (T. Polzin), vahdati@informatik.uni-mannheim.de (S. Vahdati Daneshmand).

PII: S0166-218X(00)00318-8

We have also included some known results to provide the reader with a wider view at one sight.

The results in this paper are not explicitly presented as polyhedral ones; the relationship to results of this kind and polyhedral extensions of our results will be briefly discussed in Section 7.2.

Also, the empirical study of the relaxations and the algorithmic aspects of their application are not the subject of this paper. In another paper [19], we report on our empirical study of some of these relaxations and their algorithmic application, not only for computing lower bounds, but also as the basis of empirically successful heuristics for computing upper bounds and sophisticated reduction techniques, culminating in an exact algorithm which achieves impressive empirical results.

1.1. Definitions

The Steiner problem in networks can be stated as follows (see [10] for details):

Given an (undirected, connected) network $G=(V,E,c)$ (with vertices $V=\{v_1,\dots,v_n\}$, edges E and edge weights $c_{ij}=c((v_i,v_j))>0$) and a set R, $\emptyset \neq R \subseteq V$, of *required vertices* (or *terminals*), find a minimum weight tree in G that spans R.

For the ease of notation we assume $R=\{v_1,\dots,v_r\}$. If we want to stress that v_i is a terminal, we will write z_i instead of v_i.

We also look at two reformulations of this problem, because they are used in some relaxations. One uses the directed version: Given $G=(V,E,c)$ and R, find a minimum weight arborescence in $\vec{G}=(V,A,c)$ ($A{:=}\{[v_i,v_j],[v_j,v_i] \mid (v_i,v_j)\in E\}$, c defined accordingly) with a terminal (say z_1) as the root that spans $R_1{:=}R\setminus\{z_1\}$.

The problem can also be stated as finding a degree-constrained minimum spanning tree T_0 in a modified network $G_0=(V_0,E_0,c_0)$, produced by adding a new vertex v_0 and connecting it through zero cost edges to all vertices in $V\setminus R$ and to a fixed terminal (say z_1). The problem is now equivalent to finding a minimum spanning tree T_0 in G_0 with the additional restriction that in T_0 every vertex in $V\setminus R$ adjacent to v_0 must have degree one. For more details on this reformulation, see [2,3]. Again, a similar directed version for a network $\vec{G}_0$ can be defined, this time by adding zero cost arcs $[v_0,v_i]$ (for all $v_i\in V\setminus R$) and $[v_0,z_1]$ to $\vec{G}$.

A *cut* in $\vec{G}=(V,A,c)$ (or in $G=(V,E,c)$) is defined as a partition $C=\{\bar{W},W\}$ of V ($\emptyset\subset W\subset V$; $V=W\dot{\cup}\bar{W}$). We use $\delta^-(W)$ to denote the set of arcs $[v_i,v_j]\in A$ with $v_i\in\bar{W}$ and $v_j\in W$. For simplicity, we write $\delta^-(v_i)$ instead of $\delta^-(\{v_i\})$. The sets $\delta^+(W)$ and, for the undirected version, $\delta(W)$ are defined similarly. A cut $C=\{\bar{W},W\}$ is called a *Steiner cut*, if $z_1\in\bar{W}$ and $R_1\cap W\neq\emptyset$ (for the undirected version: $R\cap W\neq\emptyset$ and $R\cap\bar{W}\neq\emptyset$).

In the integer programming formulations, we use (binary) variables x_{ij} for each arc $[v_i,v_j]\in A$ (resp. X_{ij} for each edge $(v_i,v_j)\in E$), indicating whether an arc is in the solution ($x_{ij}=1$) or not ($x_{ij}=0$). Let P_1 be such a program. The corresponding linear relaxation is denoted by LP_1. The value of an optimal solution of the integer programming formulation (for given $\vec{G}$ and R), denoted by $v(P_1)$, is of course the value of an optimal solution of the corresponding Steiner arborescence problem in $\vec{G}$. Thus, in this context we are only interested in the optimal value $v(LP_1)$ of the corresponding linear relaxation, which can differ from $v(P_1)$.

In the following text, we will often define integer formulations or prove relationships between linear relaxations. The notations P_1 (or LP_1) always denote the integer (or linear) program corresponding to an arbitrary, but fixed instance (G,R) of the Steiner problem (with G replaced by $\vec{G}$, G_0 or $\vec{G}_0$ when appropriate).

We compare relaxations using the predicates *equivalent* and (*strictly*) *stronger*: We call a relaxation R_1 stronger than a relaxation R_2 if the optimal value of R_1 is no less than that of R_2 for all instances of the problem. If R_2 is also stronger than R_1, we call them equivalent, otherwise we say that R_1 is strictly stronger than R_2. If neither is stronger than the other, they are *incomparable*.

2. Cut and flow formulations

In this section, we state the basic flow- and cut-based formulations of the Steiner problem. There are some well-known observations concerning these formulations, which we cite without proof.

2.1. Cut formulations

The directed cut formulation was stated in [20].

$$\boxed{P_C} \quad \sum_{[v_i,v_j]\in A} c_{ij}x_{ij} \to \min,$$

$$\sum_{[v_i,v_j]\in\delta^-(W)} x_{ij} \geqslant 1 \quad (z_1 \notin W,\ W\cap R_1 \neq \emptyset), \tag{1.1}$$

$$x_{ij} \in \{0,1\} \quad ([v_i,v_j]\in A). \tag{1.2}$$

The constraints (1.1) are called Steiner cut constraints. They guarantee that in any arc set corresponding to a feasible solution, there is a path from z_1 to any other terminal.

A formulation for the undirected version was stated in [1]:

$$\boxed{P_{UC}} \quad \sum_{(v_i,v_j)\in E} c_{ij}X_{ij} \to \min,$$

$$\sum_{(v_i,v_j)\in\delta(W)} X_{ij} \geqslant 1 \quad (W\cap R \neq R,\ W\cap R \neq \emptyset), \tag{2.1}$$

$$X_{ij} \in \{0,1\} \quad ((v_i,v_j)\in E). \tag{2.2}$$

Lemma 1. *LP_C is strictly stronger than LP_{UC}; and* $\sup\{v(LP_C)/v(LP_{UC})\} = 2$ [4,6].

We just mention here that $v(P_{UC})/v(LP_{UC}) \leqslant 2$ [8]; and that when applied to undirected instances, the value $v(LP_C)$ is independent of the choice of the root [9]. For much more information on LP_C, LP_{UC} and their relationship, see [4]. Also, many related results are discussed in [17].

2.2. Flow formulations

Viewing the Steiner problem as a multicommodity flow problem leads to the following formulation (see [20]):

$$\boxed{P_F} \qquad \sum_{[v_i,v_j]\in A} c_{ij}x_{ij} \to \min,$$

$$\sum_{[v_j,v_i]\in A} y^t_{ji} - \sum_{[v_i,v_j]\in A} y^t_{ij} = \begin{cases} 1 & (z_t \in R_1;\ v_i = z_t), \\ 0 & (z_t \in R_1;\ v_i \in V \setminus \{z_1, z_t\}), \end{cases} \tag{3.1}$$

$$x_{ij} \geqslant y^t_{ij} \quad (z_t \in R_1;\ [v_i,v_j] \in A), \tag{3.2}$$

$$y^t_{ij} \geqslant 0 \quad (z_t \in R_1;\ [v_i,v_j] \in A), \tag{3.3}$$

$$x_{ij} \in \{0,1\} \quad ([v_i,v_j] \in A). \tag{3.4}$$

Each variable y^t_{ij} denotes the quantity of the commodity t flowing through $[v_i,v_j]$. Constraints (3.1) and (3.3) guarantee that for each terminal $z_t \in R_1$, there is a flow of one unit of commodity t from z_1 to z_t. Together with (3.2), they guarantee that in any arc set corresponding to a feasible solution, there is a path from z_1 to any other terminal.

Lemma 2. *LP_C is equivalent to LP_F* [20].

The correspondence is even stronger: Every feasible solution x for LP_C corresponds to a feasible solution (x,y) for LP_F.

The straightforward translation of P_F for the undirected version leads to LP_{UF} with $v(LP_{UF}) = v(LP_{UC})$ (see [9]). There are other undirected formulations (see [9]), leading to relaxations which are all equivalent to LP_F; so we use the notation LP_{FU} for all of them.

Of course, there is no need for different commodities in P_F. In an aggregated version, which we call $P_{F^{++}}$, one unit of a single commodity flows from z_1 to each terminal $z_t \in R_1$ (see [16]). This program has only $\Theta(|A|)$ variables and constraints, which is asymptotically minimal. But the corresponding linear relaxation $LP_{F^{++}}$ is not a strong one:

Lemma 3. *LP_F is strictly stronger than $LP_{F^{++}}$. The worst case ratio $v(LP_F)/v(LP_{F^{++}})$ is $r-1$* [16,6].

In [15], the two-terminal formulation was stated:

$$\boxed{P_{2T}} \qquad \sum_{[v_i,v_j]\in A} c_{ij}x_{ij} \to \min,$$

$$\sum_{[v_j,v_i]\in A} \check{y}^{kl}_{ji} - \sum_{[v_i,v_j]\in A} \check{y}^{kl}_{ij} \geqslant \begin{cases} -1 & (\{z_k,z_l\}\subseteq R_1;\ v_i = z_1), \\ 0 & (\{z_k,z_l\}\subseteq R_1;\ v_i \in V\setminus\{z_1\}), \end{cases} \tag{4.1}$$

$$\sum_{[v_j,v_i]\in A} (\check{y}^{kl}_{ji} + \grave{y}^{kl}_{ji}) - \sum_{[v_i,v_j]\in A} (\check{y}^{kl}_{ij} + \grave{y}^{kl}_{ij}) = \begin{cases} 1 & (\{z_k,z_l\}\subseteq R_1;\ v_i = z_k), \\ 0 & (\{z_k,z_l\}\subseteq R_1;\ v_i \in V\setminus\{z_1,z_k\}), \end{cases} \tag{4.2}$$

$$\sum_{[v_j,v_i]\in A} (\check{y}^{kl}_{ji} + \acute{y}^{kl}_{ji}) - \sum_{[v_i,v_j]\in A} (\check{y}^{kl}_{ij} + \acute{y}^{kl}_{ij}) = \begin{cases} 1 & (\{z_k,z_l\}\subseteq R_1;\ v_i = z_l), \\ 0 & (\{z_k,z_l\}\subseteq R_1;\ v_i \in V\setminus\{z_1,z_l\}), \end{cases} \tag{4.3}$$

$$\check{y}^{kl}_{ij} + \acute{y}^{kl}_{ij} + \grave{y}^{kl}_{ij} \leqslant x_{ij} \quad (\{z_k,z_l\}\subseteq R_1;\ [v_i,v_j]\in A), \tag{4.4}$$

$$\check{y}^{kl}_{ij}, \acute{y}^{kl}_{ij}, \grave{y}^{kl}_{ij} \geqslant 0 \quad (\{z_k,z_l\}\subseteq R_1;\ [v_i,v_j]\in A), \tag{4.5}$$

$$x_{ij} \in \{0,1\} \quad ([v_i,v_j]\in A). \tag{4.6}$$

The formulation P_F is based on the flow formulation of the shortest path problem (the special case of the Steiner problem with $|R_1| = 1$). The formulation P_{2T} is based on the special case with $|R_1| = 2$, namely, the two-terminal Steiner tree problem. In a Steiner tree, for any two terminals $z_k, z_l \in R_1$, there is a two-terminal tree consisting of a path from z_1 to a splitter node v_s and two paths from v_s to z_k and z_l (v_s can belong to $\{z_1, z_k, z_l\}$). In P_{2T}, $\check{y}$, $\grave{y}$ and $\acute{y}$ describe flows from z_1 to v_s, from v_s to z_k and from v_s to z_l. Note that the flow described by $\check{y}$ can have an excess at some vertices (because of the inequality in (4.1)), this excess is carried by the flows described by $\grave{y}$ and $\acute{y}$ to z_k and z_l (because of (4.2) and (4.3)).

Lemma 4. *LP_{2T} is strictly stronger than LP_F* [15].

3. Tree formulations

In this section, we state the basic tree-based formulations and prove that the corresponding linear relaxations are all equivalent. We also discuss some variants from the literature, which we prove to be weaker.

3.1. Degree-constrained tree formulations

In [3], the following program was suggested, which is a translation of the degree-constrained minimum spanning tree problem in G_0.

$$\boxed{P_{T_0}} \qquad \sum_{(v_i,v_j)\in E} c_{ij}X_{ij} \to \min,$$

$$\{(v_i,v_j)\,|\,X_{ij}=1\}\text{: builds a spanning tree for } G_0, \tag{5.1}$$

$$X_{0k}+X_{ki}\leqslant 1 \quad (v_k\in V\setminus R;\ (v_k,v_i)\in\delta(v_k)), \tag{5.2}$$

$$X_{ij}\in\{0,1\} \quad ((v_i,v_j)\in E_0). \tag{5.3}$$

The requirement (5.1) can be stated by linear constraints. In the following, we assume that (5.1) is replaced by the following constraints:

$$\sum_{(v_i,v_j)\in E_0} X_{ij}=n, \tag{5.4}$$

$$\sum_{(v_i,v_j)\in E_0;\ v_i,v_j\in W} X_{ij}\leqslant |W|-1 \quad (\emptyset\neq W\subset V_0). \tag{5.5}$$

The constraints (5.4) and (5.5), together with the nonnegativity of X, define a polyhedron whose extreme points are the incidence vectors of spanning trees in G_0 (see [7,17]). Thus, no other set of linear constraints replacing (5.1) can lead to a stronger linear relaxation.

A directed version can be stated as follows:

$$\boxed{P_{\vec{T}_0}} \qquad \sum_{[v_i,v_j]\in A} c_{ij}x_{ij} \to \min,$$

$$\sum_{[v_j,v_i]\in\delta^-(v_i)} x_{ji}=1 \quad (v_i\in V), \tag{6.1}$$

$$\sum_{[v_i,v_j]\in A_0;\ v_i,v_j\in W} x_{ij}\leqslant|W|-1 \quad (\emptyset\notin W\subseteq V_0), \tag{6.2}$$

$$x_{0i}+x_{ij}+x_{ji}\leqslant 1 \quad (v_i\in V\setminus R;\ [v_i,v_j]\in\delta^+(v_i)), \tag{6.3}$$

$$x_{ij}\in\{0,1\} \quad ([v_i,v_j]\in A_0). \tag{6.4}$$

Again, the constraints (6.1) and (6.2), together with the nonnegativity of x, define a polyhedron whose extreme points are the incidence vectors of spanning arborescences with root v_0 (see [17]). Note that $\delta^-(v_0)=\emptyset$ by the construction of $\vec{G}_0$.

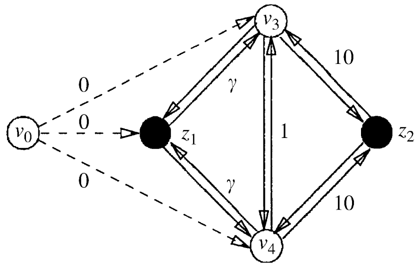

Fig. 1. Example with $v(L\dot{P}_{\vec{T}_0}) \ll v(LP_{\vec{T}_0}) = v(LP_{T_0}) \ll v(P_{T_0})$.

In the literature on the Steiner problem, one finds usually a directed variant $\dot{P}_{\vec{T}_0}$ that uses

$$x_{0i} + x_{ij} \leqslant 1 \quad (v_i \in V \setminus R;\ [v_i, v_j] \in \delta^+(v_i))$$

instead of the constraints (6.3) (see, for example, [10]). Obviously, $v(\dot{P}_{\vec{T}_0}) = v(P_{\vec{T}_0})$, and $v(L\dot{P}_{\vec{T}_0}) \leqslant v(LP_{\vec{T}_0})$. The following example shows that $LP_{\vec{T}_0}$ is strictly stronger than the version in the literature.

Example 1. Fig. 1 shows the network $\vec{G}$ with $R = \{z_1, z_2\}$, $\gamma \geqslant 100$ and the network $\vec{G}_0$. The minimum Steiner arborescence has the value $\gamma + 10$.

The following $\dot{x}$ is feasible (and optimal) for $L\dot{P}_{\vec{T}_0}$ and gives the value 11: $\dot{x}_{01} = 1$, $\dot{x}_{03} = \dot{x}_{04} = \dot{x}_{34} = \dot{x}_{43} = \dot{x}_{32} = \dot{x}_{42} = \frac{1}{2}$ and $\dot{x}_{ij} = 0$ (for all other arcs). But for $LP_{\vec{T}_0}$, $\dot{x}$ is infeasible. The optimal value here is: $v(LP_{\vec{T}_0}) = \gamma/3 + 14$ (this value is reached for example by $\hat{x}$ with $\hat{x}_{01} = 1$, $\hat{x}_{03} = \hat{x}_{04} = \hat{x}_{13} = \hat{x}_{23} = \hat{x}_{32} = \frac{1}{3}$, $\hat{x}_{42} = \hat{x}_{34} = \frac{2}{3}$ and $\hat{x}_{ij} = 0$ (for all other arcs)). So the ratio $v(L\dot{P}_{\vec{T}_0})/v(LP_{\vec{T}_0})$ can be arbitrarily close to 0.

3.2. Rooted tree formulation

The rooted tree formulation is stated, for example, in [13]:

$$\boxed{P_{\vec{T}}} \qquad \sum_{[v_i,v_j] \in A} c_{ij} x_{ij} \to \min,$$

$$\sum_{[v_j,v_i] \in \delta^-(v_i)} x_{ji} = 1 \quad (v_i \in R_1), \tag{7.1}$$

$$\sum_{[v_k,v_i] \in \delta^-(v_i);\ v_k \neq v_j} x_{ki} \geqslant x_{ij} \quad (v_i \in V \setminus R;\ [v_i, v_j] \in \delta^+(v_i)), \tag{7.2}$$

$$\sum_{[v_i,v_j] \in A;\ v_i, v_j \in W} x_{ij} \leqslant |W| - 1 \quad (\emptyset \neq W \subseteq V), \tag{7.3}$$

$$x_{ij} \in \{0, 1\} \quad ([v_i, v_j] \in A). \tag{7.4}$$

To get rid of the exponential number of constraints for avoiding cycles, many authors have considered replacing (7.3) by the subtour elimination constraints introduced in the TSP-context (known as the Miller–Tucker–Zemlin constraints [18]), allowing additional variables t_i for all $v_i \in V$:

$$t_i - t_j + nx_{ij} \leqslant n - 1 \quad ([v_i, v_j] \in A). \tag{7.5}$$

This leads to the program $\dot{P}_{\vec{T}}$ with $\Theta(|A|)$ variables and constraints, which is asymptotically minimal. The linear relaxation $L\dot{P}_{\vec{T}}$ was used by [12]. We will now prove the intuitive guess that $LP_{\vec{T}}$ is stronger than $L\dot{P}_{\vec{T}}$. Indeed, the ratio $v(L\dot{P}_{\vec{T}})/v(LP_{\vec{T}})$ can be arbitrarily close to 0 (see Fig. 2).

Lemma 5. $v(L\dot{P}_{\vec{T}}) \leqslant v(LP_{\vec{T}})$.

Proof. Let $\hat{x}$ denote an (optimal) solution for $LP_{\vec{T}}$. Obviously, $\hat{x}$ satisfies the constraints (7.1) and (7.2). We show now that it is possible to construct $\hat{t}$ such that $(\hat{x}, \hat{t})$ satisfies (7.5), as well.

We start with an arbitrary $\hat{t}$ (e.g. $\hat{t}_i = 0$ (for all $v_i \in V$)). We define for every arc $[v_i, v_j] \in A$: $s_{ij} := (n-1) - (\hat{t}_i - \hat{t}_j + n\hat{x}_{ij})$; and call an arc $[v_i, v_j]$ *good*, if $s_{ij} \geqslant 0$; *used*, if $s_{ij} \leqslant 0$; and *bad*, if $s_{ij} < 0$. Suppose $[v_i, v_j]$ is a bad arc (if no bad arcs exist, $(\hat{x}, \hat{t})$ satisfies (7.5)).

We show now how $\hat{t}_j$ (and perhaps some other $\hat{t}_p$) can be increased in a way that $[v_i, v_j]$ becomes good, but no good arc becomes bad. By repeating this procedure, we can make all arcs good and prove the lemma.

In each step, we denote by W_j the set of vertices $v_k \in V$ that can be reached from v_j through paths with only used arcs. We define Δ as $\min\{s_{kl} \mid [v_k, v_l] \in \delta^+(W_j)\}$, if this set is nonempty, and ∞ otherwise. Now, we increase for all vertices $v_p \in W_j$ the variables $\hat{t}_p$ by $\min\{-s_{ij}, \Delta\}$ (these values can change in every step). By doing this, no arc of $\delta^+(W_j)$ becomes bad. For arcs $[v_p, v_q]$ with $v_p, v_q \in W_j$ or $v_p, v_q \notin W_j$ the value of s_{pq} does not change; and for arcs $[v_q, v_p] \in \delta^-(W_j)$ s_{qp} does not decrease.

Because $\hat{t}_j$ is increased in every step, there is only one situation that could prevent that $[v_i, v_j]$ becomes good: In one step v_i is absorbed by W_j. But then, according to the definition of W_j, there exists a path $v_j \rightsquigarrow v_i$ with only used arcs. Thus, there exists a cycle $C := (v_i, v_j = v_{k_1}, \ldots, v_{k_l} = v_i)$, with $s_{k_l k_1} < 0$ and $s_{k_{t-1} k_t} \leqslant 0$ (for all $t \in \{2, \ldots, l\}$). Summation of the inequalities for arcs on the cycle C leads to: $n \sum_{[v_p, v_q] \in C} \hat{x}_{pq} > l(n-1)$. On the other hand, since $\hat{x}$ satisfies the constraints (7.3), $\sum_{[v_p, v_q] \in C} \hat{x}_{pq} \leqslant l - 1$. The consequence, $(l-1)/l > (n-1)/n$, is a contradiction. □

3.3. Equivalence of tree-class relaxations

We show now the equivalence of the tree-based relaxations $LP_{T_0}, LP_{\vec{T}_0}$, and $LP_{\vec{T}}$.

Lemma 6. $v(LP_{\vec{T}_0}) = v(LP_{T_0})$.

Proof. (I) $v(LP_{\vec{T}_0}) \geqslant v(LP_{T_0})$: Let x denote an (optimal) solution for $LP_{\vec{T}_0}$. Define X with $X_{ij} := x_{ij} + x_{ji}$ (for all $(v_i, v_j) \in E$), $X_{0i} := x_{0i}$ (for all $v_i \in V \setminus R$) and $X_{01} := x_{01}$. It is easy to check that X satisfies all constraints of LP_{T_0} and yields the same value as $v(LP_{\vec{T}_0})$.

(II) $v(LP_{T_0}) \geqslant v(LP_{\vec{T}_0})$: Now let X denote an (optimal) solution for LP_{T_0}. Define Δ with $\Delta_{ij} \in [0,1]$ arbitrarily (for all $(v_i, v_j) \in E$) and set x to $x_{ij} := \Delta_{ij} X_{ij}$, $x_{ji} := (1 - \Delta_{ij}) X_{ij}$ (for all $(v_i, v_j) \in E$), $x_{0i} := X_{0i}$ (for all $v_i \in V \setminus R$) and $x_{01} := X_{01}$. Again, it is easy to validate that x satisfies the constraints (6.2) and (6.3) and yields the same value as $v(LP_{T_0})$.

The only question is, whether there is a Δ such that x also satisfies the constraints (6.1). This question can be stated in the following way: Is it possible to distribute the "supply" X_{ij} of each edge (v_i, v_j) in such a way to its end-vertices that every vertex $v_i \in V$ gets one unit at the end? It is known that this problem can be viewed as a flow problem: Construct a flow network with source s, sink t, and vertices u_{ij} (for all $(v_i, v_j) \in E_0$) and u_i (for all $v_i \in V_0$). Every u_{ij} is connected with u_i and u_j through arcs $[u_{ij}, u_i]$ and $[u_{ij}, u_j]$ with capacity ∞. Furthermore, there are arcs $[s, u_{ij}]$ with capacity X_{ij} and arcs $[u_i, t]$ with capacity 1 (or 0, if $i = 0$). The question above is equivalent to the question, whether a flow from s to t with value n can be constructed. The max-flow min-cut theorem says that this is possible if and only if there is no cut $C = \{U, \bar{U}\}$ (with $s \in U$ and $t \notin U$) with capacity less than n (obviously $U = \{s\}$ and $U = V \setminus \{t\}$ correspond to cuts with capacity n). Suppose that U corresponds to a cut C with minimum capacity. Define $W := \{v_i \in V_0 \mid u_i \in U\}$, $E_W := \{(v_i, v_j) \in E_0 \mid v_i, v_j \in W\}$, and $E_U := \{(v_i, v_j) \in E_0 \mid u_{ij} \in U\}$. For every $[v_i, v_j] \in E_U$ ($u_{ij} \in U$), u_i and u_j must belong to U ($[v_i, v_j] \in E_W$), because otherwise the capacity of C would be ∞ which is not minimal. It follows that: $E_U \subseteq E_W$.

The capacity of C is

$$
\begin{aligned}
|W \setminus \{v_0\}| + \sum_{(v_i,v_j) \in E_0 \setminus E_U} X_{ij} &\geqslant |W \setminus \{v_0\}| + \sum_{(v_i,v_j) \in E_0 \setminus E_W} X_{ij} \quad (\text{since } E_U \subseteq E_W) \\
&\geqslant |W| - 1 + \sum_{(v_i,v_j) \in E_0} X_{ij} - \sum_{(v_i,v_j) \in E_W} X_{ij} \\
&= |W| - 1 + n - \sum_{(v_i,v_j) \in E_W} X_{ij} \quad (\text{because of (5.4)}) \\
&\geqslant n \quad (\text{because of (5.5)}). \qquad \square
\end{aligned}
$$

Lemma 7. $v(LP_{\vec{T}}) = v(LP_{\vec{T}_0})$.

Proof. (I) $v(LP_{\vec{T}_0}) \geqslant v(LP_{\vec{T}})$: Let $\hat{x}$ denote an (optimal) solution for $LP_{\vec{T}_0}$. Define $\tilde{x}$ with $\tilde{x}_{ij} := \hat{x}_{ij}$ (for all $[v_i, v_j] \in A$). Because $\hat{x}$ satisfies the constraints (6.1) and in $\vec{G}_0$ only arcs in A are incident with terminals in R_1, $\tilde{x}$ satisfies the constraints (7.1).

Furthermore, $\tilde{x}$ satisfies the constraints (7.2), because for every arc $[v_i, v_j] \in A$ with $v_i \in V \setminus R$ holds:

$$\begin{aligned}
\sum_{[v_k,v_i]\in\delta^-(v_i);v_k\neq v_j} \tilde{x}_{ki} &= \left(\sum_{[v_k,v_i]\in\delta^-(v_i)} \tilde{x}_{ki}\right) - \tilde{x}_{ji} \quad (\delta \text{ in } \vec{G})\\
&= \left(\sum_{[v_k,v_i]\in\delta^-(v_i)} \hat{x}_{ki}\right) - \hat{x}_{0i} - \hat{x}_{ji} \quad (\delta \text{ in } \vec{G}_0)\\
&= 1 - \hat{x}_{0i} - \hat{x}_{ji} \quad (\text{because of (6.1)})\\
&\geqslant \hat{x}_{ij} \quad (\text{because of (6.3)})\\
&= \tilde{x}_{ij}.
\end{aligned}$$

Finally, $\tilde{x}$ satisfies (7.3), because $\hat{x}$ satisfies (6.2).

(II) $v(LP_{\vec{T}}) \geqslant v(LP_{\vec{T}_0})$: Let $\tilde{x}$ denote an (optimal) solution for $LP_{\vec{T}}$. Define $\hat{x}$ with $\hat{x}_{ij} := \tilde{x}_{ij}$ (for all $[v_i, v_j] \in A$) and $\hat{x}_{0i} := 1 - \sum_{[v_j,v_i]\in\delta^-(v_i)} \tilde{x}_{ji}$ (for all $v_i \in V \setminus R_1$). Notice that for an optimal $\tilde{x}$, $\sum_{[v_j,v_i]\in\delta^-(v_i)} \tilde{x}_{ji} > 1$ could only be forced by (7.2) for some arc $[v_i, v_l]$ with $\sum_{[v_j,v_i]\in\delta^-(v_i),v_j\neq v_l} \tilde{x}_{ji} = \tilde{x}_{il}$, and it would follow that $1 < \sum_{[v_j,v_i]\in\delta^-(v_i)} \tilde{x}_{ji} = \tilde{x}_{li} + \tilde{x}_{il}$, but this is excluded by (7.3) (for $W = \{v_i, v_l\}$). So $\hat{x}$ satisfies (6.1) in a trivial way.

The constraints (6.2) are satisfied by $\hat{x}$ for every $W \subseteq V$, because $\tilde{x}$ satisfies (7.3). For $W \subseteq V_0$ with $v_0 \in W$ holds:

$$\begin{aligned}
\sum_{[v_i,v_j]\in A_0;\ v_i,v_j\in W} \hat{x}_{ij} &\leqslant \sum_{v_i\in W\setminus\{v_0\}} \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}_{ji} \quad (\text{in } \vec{G}_0)\\
&= \sum_{v_i\in W\setminus\{v_0\}} 1 \quad (\text{because of (6.1)})\\
&= |W| - 1.
\end{aligned}$$

Finally, for every $[v_i, v_j] \in A$ with $v_i \in V \setminus R$:

$$\begin{aligned}
\hat{x}_{0i} + \hat{x}_{ij} + \hat{x}_{ji} &= 1 - \left(\sum_{[v_k,v_i]\in\delta^-(v_i)} \tilde{x}_{ki}\right) + \tilde{x}_{ij} + \tilde{x}_{ji} \quad (\text{in } \vec{G})\\
&= 1 - \left(\sum_{[v_k,v_i]\in\delta^-(v_i);\ v_k\neq v_j} \tilde{x}_{ki}\right) + \tilde{x}_{ij}\\
&\leqslant 1 \quad (\text{because of (7.2)}).
\end{aligned}$$

Thus, $\hat{x}$ satisfies also the constraints (6.3). □

4. Relationship between the two classes

In this section, we settle the question of the relationship between flow and tree-based relaxations by proving that LP_C is strictly stronger than $LP_{\vec{T}}$. Our proofs show also that LP_C cannot be strengthened by adding constraints which are present in $LP_{\vec{T}_0}$ or $LP_{\vec{T}}$.

First, we show that every (optimal) solution $\hat{x}$ of LP_C has certain properties:

Lemma 8. *For every* (*optimal*) *solution* $\hat{x}$ *of* LP_C, $W \subseteq V \setminus \{z_1\}$ *and* $v_k \in W$ *holds*:

$$\sum_{[v_i,v_j]\in\delta^-(W)} \hat{x}_{ij} \geqslant \sum_{[v_i,v_k]\in\delta^-(v_k)} \hat{x}_{ik}.$$

Proof. Suppose that $\hat{x}$ violates the inequality for some W and v_k. Among all such inequalities, choose one for which $|W|$ is minimal. For this inequality to be violated, there must be an arc $[v_l, v_k] \in \delta^-(v_k) \setminus \delta^-(W)$ with $\hat{x}_{lk} > 0$. Because of the optimality of $\hat{x}$, $\hat{x}_{lk}$ cannot be decreased without violating a Steiner cut constraint, so there is a $U \subset V$ with $z_1 \notin U$, $U \cap R \neq \emptyset$, $[v_l, v_k] \in \delta^-(U)$, and $\sum_{[v_i,v_j]\in\delta^-(U)} \hat{x}_{ij} = 1$. Now one has the inequality †:

$$\begin{aligned}
&\sum_{[v_i,v_j]\in\delta^-(U)} \hat{x}_{ij} + \sum_{[v_i,v_j]\in\delta^-(W)} \hat{x}_{ij} \\
&\quad = \sum_{[v_i,v_j]\in\delta^-(U\cup W)} \hat{x}_{ij} + \sum_{[v_i,v_j]\in\delta^-(U\cap W)} \hat{x}_{ij} \\
&\qquad + \sum_{[v_i,v_j]\in A, v_i\in W\setminus U, v_j\in U\setminus W} \hat{x}_{ij} + \sum_{[v_i,v_j]\in A, v_i\in U\setminus W, v_j\in W\setminus U} \hat{x}_{ij} \\
&\quad \geqslant \sum_{[v_i,v_j]\in\delta^-(U\cup W)} \hat{x}_{ij} + \sum_{[v_i,v_j]\in\delta^-(U\cap W)} \hat{x}_{ij}.
\end{aligned}$$

Since $z_1 \notin U \cup W$ and $(U \cup W) \cap R \neq \emptyset$, $U \cup W$ corresponds to a Steiner cut, and $\sum_{[v_i,v_j]\in\delta^-(U\cup W)} \hat{x}_{ij} \geq 1 = \sum_{[v_i,v_j]\in\delta^-(U)} \hat{x}_{ij}$. Using †, one obtains: $\sum_{[v_i,v_j]\in\delta^-(W)} \hat{x}_{ij} \geqslant \sum_{[v_i,v_j]\in\delta^-(U\cap W)} \hat{x}_{ij}$. This implies that $\hat{x}$ also violates the lemma for $U \cap W$ and v_k. Since $v_l \in W \setminus U$, we have $|U \cap W| < |W|$, and this contradicts the minimality of W.[1] □

Lemma 9. *For every* (*optimal*) *solution* $\hat{x}$ *of* LP_C *and* $v_k \in V \setminus \{z_1\}$ *holds*:

$$\sum_{[v_i,v_k]\in\delta^-(v_k)} \hat{x}_{ik} \leqslant 1.$$

Proof. Suppose $\hat{x}$ violates the inequality for v_k. There is an arc $[v_l, v_k] \in \delta^-(v_k)$ with $\hat{x}_{lk} > 0$. Because of the optimality of $\hat{x}$, $\hat{x}_{lk}$ cannot be decreased without violating a Steiner cut constraint, so there is a $W \subset V$ with $z_1 \notin W$, $W \cap R \neq \emptyset$, $[v_l, v_k] \in$

[1] In a different context this argumentation was used in [9].

$\delta^-(W)$, and $\sum_{[v_i,v_j]\in\delta^-(W)} \hat{x}_{ij} = 1$. Together with Lemma 8 (for v_k and W), one gets a contradiction. $\square$

Lemma 10. *For every (optimal) solution $\hat{x}$ of LP_C, $v_l \in V \setminus \{z_1\}$, and $[v_l, v_k] \in A$ holds*:

$$\sum_{[v_i,v_l]\in\delta^-(v_l), v_i\neq v_k} \hat{x}_{il} \geqslant \hat{x}_{lk}.$$

Proof. This follows directly from Lemma 8 (for v_k and $W = \{v_l, v_k\}$) by subtracting $\sum_{[v_i,v_k]\in\delta^-(v_k), v_i\neq v_l} \hat{x}_{ik}$ from both sides. Note that the special case $v_k = z_1$ is trivial, because $\hat{x}_{l1} = 0$ in every optimal solution. $\square$

Theorem 11. $v(LP_{\vec{T}}) \leqslant v(LP_C)$.

Proof. Let $\hat{x}$ be an (optimal) solution for LP_C. We will show that $\hat{x}$ is feasible for $LP_{\vec{T}}$:

Because $\{v_i\}$ corresponds to a Steiner cut for $v_i \in R_1$ and using Lemma 9, $\hat{x}$ satisfies (7.1).

Because of Lemma 10, $\hat{x}$ satisfies (7.2).

Let $W \subseteq V$ be a nonempty set. If $z_1 \in W$:

$$\begin{aligned}
\sum_{[v_i,v_j]\in A;\ v_i,v_j\in W} \hat{x}_{ij} &\leqslant \sum_{v_i\in W} \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}_{ji} \\
&= \sum_{v_i\in W\setminus\{z_1\}} \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}_{ji} \quad \text{(optimality of } \hat{x}) \\
&\leqslant \sum_{v_i\in W\setminus\{z_1\}} 1 \quad \text{(Lemma 9)} \\
&= |W| - 1.
\end{aligned}$$

Now, we assume $z_1 \notin W$ and define $\Delta := \sum_{[v_k,v_l]\in\delta^-(W)} \hat{x}_{kl}$. There are two cases:

(I) $\Delta \geqslant 1$:

$$\begin{aligned}
\sum_{[v_i,v_j]\in A;\ v_i,v_j\in W} \hat{x}_{ij} &= \sum_{v_i\in W} \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}_{ji} - \sum_{[v_k,v_l]\in\delta^-(W)} \hat{x}_{kl} \\
&\leqslant \left(\sum_{v_i\in W} \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}_{ji} \right) - 1 \quad (\Delta \geqslant 1) \\
&\leqslant \sum_{v_i\in W} 1 - 1 \quad \text{(Lemma 9)} \\
&= |W| - 1.
\end{aligned}$$

(II) $\Delta < 1$:

$$\begin{aligned}
\sum_{[v_i,v_j]\in A;\ v_i,v_j\in W} \hat{x}_{ij} &= \sum_{v_i\in W}\ \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}_{ji} - \sum_{[v_k,v_l]\in\delta^-(W)} \hat{x}_{kl} \\
&\leqslant \sum_{v_i\in W}\ \sum_{[v_k,v_l]\in\delta^-(W)} \hat{x}_{kl} - \sum_{[v_k,v_l]\in\delta^-(W)} \hat{x}_{kl} \quad \text{(Lemma 8)} \\
&= (|W|-1) \sum_{[v_k,v_l]\in\delta^-(W)} \hat{x}_{kl} \\
&< |W|-1 \quad (\Delta < 1).
\end{aligned}$$

It follows that $\hat{x}$ also satisfies (7.3). $\square$

Corollary 11.1. *The proof shows that adding constraints of $LP_{\vec{T}}$ to LP_C cannot improve $v(LP_C)$.*

Corollary 11.2. *Because the proofs of the equivalence of the tree relaxations require the optimality only in one step of Lemma 7 to show that $\sum_{[v_j,v_i]\in\delta^-(v_i)} \tilde{x}_{ji} \leqslant 1$, which is forced by Lemma 9 for each (optimal) solution of LP_C, adding constraints of $LP_{\vec{T}_0}$ to LP_C cannot improve $v(LP_C)$, either.*

To show that LP_F and LP_C are strictly stronger than the tree-based relaxations $LP_{T_0}, LP_{\vec{T}_0}$, and $LP_{\vec{T}}$, it is sufficient to give the following example.

Example 2. For the network G (or in the directed view $\vec{G}$) in Fig. 2 set $\alpha \gg 1$ and $\gamma \gg \alpha$. Obviously, $v(P_F) = v(LP_F) = \gamma$. For $LP_{\vec{T}}$ is $\hat{x}$ with $\hat{x}_{23} = \hat{x}_{34} = \hat{x}_{42} = \frac{2}{3}$, $\hat{x}_{25} = \hat{x}_{56} = \hat{x}_{62} = \frac{1}{3}$, and $\hat{x}_{ij} = 0$ (otherwise) feasible, even optimal, and gives the value $v(LP_{\vec{T}}) = \alpha + 2$. Thus, there is no positive lower bound for the ratio $v(LP_{\vec{T}})/v(LP_F)$.

With respect to $L\dot{P}_{\vec{T}}$ and $LP_{\vec{T}}$, one observes that $(\dot{x}, \dot{t})$ with $\dot{t}_i = 0$ (for all $v_i \in V$), $\dot{x}_{23} = \dot{x}_{32} = \dot{x}_{34} = \dot{x}_{43} = \dot{x}_{24} = \dot{x}_{42} = \frac{1}{2}$, and $\dot{x}_{ij} = 0$ (otherwise) is an (optimal) solution for $L\dot{P}_{\vec{T}}$ with the value 3. So, there is no positive lower bound for the ratio $v(L\dot{P}_{\vec{T}})/v(LP_{\vec{T}})$.

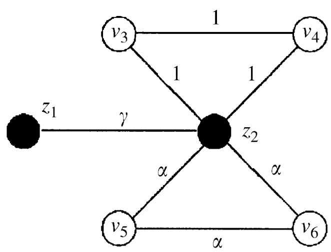

Fig. 2. Example for $v(L\dot{P}_{\vec{T}}) \ll v(LP_{\vec{T}}) \ll v(LP_F) = v(P_F)$.

5. Multiple trees and the relation to the flow model

In this section, we consider a relaxation based on multiple trees and prove its equivalence to an augmented flow relaxation. We also discuss some variants of the former relaxation.

5.1. Multiple trees formulation

In [13], a variant of $P_{\vec{T}}$ was stated, using the idea that an undirected Steiner tree can be viewed as $|R|$ different Steiner arborescences with different roots.

$$\boxed{P_{m\vec{T}}} \qquad \sum_{(v_i,v_j)\in E} c_{ij}X_{ij} \to \min,$$

$$\sum_{(v_i,v_j)\in\delta(v_i)} X_{ij} \geqslant 1 \quad (v_i \in R), \tag{8.1}$$

$$\sum_{(v_i,v_j)\in\delta(v_i)} X_{ij} \geqslant 2s_i \quad (v_i \in V \setminus R), \tag{8.2}$$

$$s_i \geqslant X_{ij} \quad (v_i \in V \setminus R;\ (v_i,v_j) \in \delta(v_i)), \tag{8.3}$$

$$x_{ij}^k + x_{ji}^k = X_{ij} \quad (v_k \in R;\ (v_i,v_j) \in E), \tag{8.4}$$

$$\sum_{[v_j,v_i]\in\delta^-(v_i)} x_{ji}^k = \begin{cases} 1 & (v_k \in R;\ v_i \in R \setminus \{v_k\}), \\ 0 & (v_k \in R;\ v_i = v_k), \end{cases} \tag{8.5}$$

$$\sum_{[v_j,v_i]\in\delta^-(v_i)} x_{ji}^k = s_i \quad (v_k \in R;\ v_i \in V \setminus R), \tag{8.6}$$

$$\{[v_i,v_j] \mid x_{ij}^k = 1\}\text{: contains no cycles } (v_k \in R), \tag{8.7}$$

$$X_{ij} \in \{0,1\} \quad ((v_i,v_j) \in E), \tag{8.8}$$

$$x_{ij}^k \in \{0,1\} \quad (v_k \in R;\ [v_i,v_j] \in A), \tag{8.9}$$

$$si \in \{0,1\} \quad (v_i \in V \setminus R). \tag{8.10}$$

In any feasible solution for $P_{m\vec{T}}$, each group of variables x^k describes an arborescence (with root z_k) spanning all terminals. The variables s describe the set of the other vertices used by these arborescences.

We will relate this formulation to the flow formulations. First, we have to present an improvement of LP_F.

5.2. Flow-balance constraints and an augmented flow formulation

There is a group of constraints (see, for example, [14]) that can be used to make LP_F stronger. We call them flow-balance constraints:

$$\sum_{[v_j,v_i]\in\delta^-(v_i)} x_{ji} \leqslant \sum_{[v_i,v_j]\in\delta^+(v_i)} x_{ij} \quad (v_i \in V \setminus R). \tag{9.1}$$

We denote the linear program that consists of LP_F and (9.1) by LP_{F+FB}. It is obvious that LP_{F+FB} is stronger than LP_F. The following example shows that it is even strictly stronger.

Example 3. The network $\vec{G}$ in Fig. 3 with z_1 as the root and $R_1 = \{z_2, z_3\}$ gives an example for $v(LP_F) < v(LP_{F+FB})$: $v(P_{F+FB}) = v(LP_{F+FB}) = 6$, $v(LP_F) = 5\frac{1}{2}$.

Now consider the following formulation:

$$\boxed{P_{F'+FB}} \qquad \sum_{(v_i,v_j)\in E} c_{ij} X_{ij} \to \min,$$

$$x_{ij} + x_{ji} = X_{ij} \quad ((v_i, v_j) \in E), \tag{10.1}$$

$$(x, y) : \text{is feasible for } P_{F+FB}. \tag{10.2}$$

Lemma 12. *If (X, x, y) is an (optimal) solution for $LP_{F'+FB}$ with root terminal z_a, then there exists an (optimal) solution $(X, \check{x}, \check{y})$ for $LP_{F'+FB}$ for any other root terminal $z_b \in R \setminus \{z_a\}$.*

Proof. One can verify that $(X, \check{x}, \check{y})$ with $\check{x}_{ij} := x_{ij} + y^b_{ji} - y^b_{ij}$, $\check{y}^t_{ij} := \max\{0, y^t_{ij} - y^b_{ij}\} + \max\{0, y^b_{ji} - y^t_{ji}\}$, $\check{y}^a_{ij} := y^b_{ji}$ (for all $[v_i, v_j] \in A$, $z_t \in R \setminus \{z_a, z_b\}$) satisfies (10.1), (3.2) and (3.3). Because of $\sum_{[v_j,v_i]\in\delta^-(v_i)} (\check{y}^t_{ji} - \check{y}^t_{ij}) = \sum_{[v_j,v_i]\in\delta^-(v_i)} (\max\{0, y^t_{ji} - y^b_{ji}\} +$

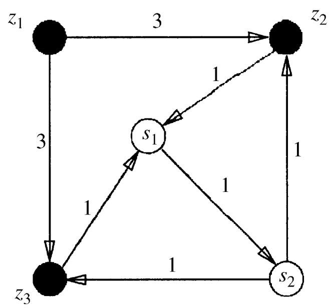

Fig. 3. Example with $v(LP_F) < v(LP_{F+FB}) = v(P_{F+FB})$.

$\max\{0, y^b_{ij} - y^t_{ij}\} + \min\{0, -y^t_{ij} + y^b_{ij}\} + \min\{0, -y^b_{ji} + y^t_{ji}\}) = \sum_{[v_j,v_i]\in\delta^-(v_i)} y^t_{ji} - y^b_{ji} + y^b_{ij} - y^t_{ij}$ (for all $v_i \in V$, $z_t \in R \setminus \{z_a, z_b\}$) the constraints (3.1) are satisfied, as well. From (3.1) for y^b follows that $\sum_{[v_j,v_i]\in\delta^-(v_i)} x_{ji} = \sum_{[v_j,v_i]\in\delta^-(v_i)} \check{x}_{ji}$ and $\sum_{[v_i,v_j]\in\delta^+(v_i)} x_{ij} = \sum_{[v_i,v_j]\in\delta^+(v_i)} \check{x}_{ij}$ for all $v_i \in V \setminus R$; therefore, $\check{x}$ satisfies the flow-balance constraints (9.1).

Because this translation could also be performed from any (optimal) solution with root terminal z_b to a feasible solution with root terminal z_a, the value $v(LP_{F'+FB})$ is independent of the choice of the root terminal and $(X, \check{x}, \check{y})$ is an (optimal) solution. □

It follows immediately that $LP_{F'+FB}$ is equivalent to LP_{F+FB}.

5.3. Relationship between the two models

We will now show that the linear relaxation $LP_{m\vec{T}}$ (where (8.7) is replaced by linear constraints of the form (7.3)) is equivalent to LP_{F+FB}.

Lemma 13. $v(LP_{m\vec{T}}) = v(LP_{F'+FB})$.

Proof. (I) $v(LP_{m\vec{T}}) \geqslant v(LP_{F'+FB})$: Let $(\hat{X}, \hat{x}, \hat{s})$ denote an (optimal) solution for $LP_{m\vec{T}}$. Define x with $x_{ij} := \hat{x}^1_{ij}$ (for all $[v_i, v_j] \in A$), and y with $y^t_{ij} := \max\{\hat{x}^1_{ij} - \hat{x}^t_{ij}, 0\}$ (for all $[v_i, v_j] \in A$, $z_t \in R_1$). Because of (8.4) and the definition of y, $y^t_{ij} = 0$ if $y^t_{ji} > 0$ (for all $(v_i, v_j) \in E$ and $z_t \in R_1$).

For all $z_t \in R_1$, $v_i \in V \setminus \{z_1, z_t\}$ holds:

$$\begin{aligned}
\sum_{[v_j,v_i]\in A} y^t_{ji} - \sum_{[v_i,v_j]\in A} y^t_{ij} &= \sum_{[v_j,v_i]\in A, \hat{x}^1_{ji} > \hat{x}^t_{ji}} (\hat{x}^1_{ji} - \hat{x}^t_{ji}) - \sum_{[v_i,v_j]\in A, \hat{x}^1_{ij} > \hat{x}^t_{ij}} (\hat{x}^1_{ij} - \hat{x}^t_{ij}) \\
&= \sum_{[v_j,v_i]\in A, \hat{x}^1_{ji} > \hat{x}^t_{ji}} (\hat{x}^1_{ji} - \hat{x}^t_{ji}) - \sum_{[v_j,v_i]\in A, \hat{x}^1_{ji} < \hat{x}^t_{ji}} (\hat{x}^t_{ji} - \hat{x}^1_{ji}) \\
&\quad \text{(because of (8.4))} \\
&= \sum_{[v_j,v_i]\in\delta^-(v_i)} (\hat{x}^1_{ji} - \hat{x}^t_{ji}) = 0 \quad \text{(because of (8.5) or (8.6))}.
\end{aligned}$$

With the same argumentation for $v_i = z_t$, it follows that y satisfies (3.1).

The other constraints (3.2) and (3.3) are satisfied in a trivial way. A substitution of (8.4) and (8.6) into (8.2) gives the flow-balance constraints (9.1). Thus, (X, x, y) is feasible for $LP_{F'+FB}$.

(II) $v(LP_{m\vec{T}}) \leqslant v(LP_{F'+FB})$: Let (X, x, y) denote an (optimal) solution for $LP_{F'+FB}$. From Lemma 12 we know that there is an (optimal) solution $(X, \hat{x}^r, \hat{y}^r)$ for each choice of the root vertex $z_r \in R$, with the property that $\hat{s}_i := \sum_{[v_j,v_i]\in\delta^-(v_i)} \hat{x}^t_{ji}$ (for any $v_i \in V \setminus R$) has the same value for any choice of $z_t \in R$. With the argumentation of Theorem 11 it follows that $(X, \hat{x}, \hat{s})$ is feasible for $LP_{m\vec{T}}$. □

Corollary 13.1. *The constraints* (8.1), (8.3) *and* (8.7) *are useless with respect to the value of the linear relaxation* $LP_{m\vec{T}}$.

Corollary 13.2. *The linear program* $LP_{m\vec{T}-}$ (*with the same objective function as* $LP_{m\vec{T}}$) *that contains only Eqs.* (8.4), (8.5) *and* (8.6) *is equivalent to* LP_F.

6. A new formulation

In this section, we introduce a new formulation and examine some of its properties. We call it common-flow formulation, because it embeds additional variables into the multicommodity flow formulation and these variables $\check{y}^{kl}$ correspond to the common flow from the root terminal to the terminals z_k and z_l. It can be stated in the following way:

$$\boxed{P_{F^2}} \qquad \sum_{[v_i,v_j]\in A} c_{ij}x_{ij} \to \min,$$

$$\sum_{[v_j,v_i]\in A} y^t_{ji} - \sum_{[v_i,v_j]\in A} y^t_{ij} = \begin{cases} 1 \ (z_t \in R_1;\ v_i = z_t), \\ 0 \ (z_t \in R_1;\ v_i \in V \setminus \{z_1, z_t\}), \end{cases} \tag{11.1}$$

$$\sum_{[v_j,v_i]\in A} \check{y}^{kl}_{ji} - \sum_{[v_i,v_j]\in A} \check{y}^{kl}_{ij} \geqslant \begin{cases} -1 \ (\{z_k, z_l\} \subseteq R_1;\ v_i = z_1), \\ 0 \quad (\{z_k, z_l\} \subseteq R_1;\ v_i \in V \setminus \{z_1\}), \end{cases} \tag{11.2}$$

$$\check{y}^{kl}_{ij} \leqslant y^k_{ij} \quad (\{z_k, z_l\} \subseteq R_1;\ [v_i, v_j] \in A), \tag{11.3}$$

$$\check{y}^{kl}_{ij} \leqslant y^l_{ij} \quad (\{z_k, z_l\} \subseteq R_1;\ [v_i, v_j] \in A), \tag{11.4}$$

$$y^k_{ij} + y^l_{ij} - \check{y}^{kl}_{ij} \leqslant x_{ij} \quad (\{z_k, z_l\} \subseteq R_1;\ [v_i, v_j] \in A), \tag{11.5}$$

$$\sum_{[v_j,v_i]\in A} x_{ji} - \sum_{[v_i,v_j]\in A} x_{ij} \leqslant 0 \quad (v_i \in V \setminus R), \tag{11.6}$$

$$\check{y}^{kl}_{ij},\ y^k_{ij} \geqslant 0 \quad (\{z_k, z_l\} \subseteq R_1;\ [v_i, v_j] \in A), \tag{11.7}$$

$$x_{ij} \in \{0, 1\} \quad ([v_i, v_j] \in A). \tag{11.8}$$

As in P_F, each set of variables y^t describes a flow from z_1 to z_t. The variables $\check{y}^{kl}$ describe the common flow from z_1 to z_k and z_l. The inequalities (11.2) guarantee that the common flow is nonincreasing; (11.5) state that the capacity of each arc must be at least the sum of each pair of flows minus the common flow through this arc. The idea behind this is to make it difficult for two flows to split up and rejoin again. The inequalities (11.6) are the flow-balance constraints (9.1).

Consider an (optimal) solution $(x, y, \check{y})$ for P_{F^2} and define $T{:=}\{[v_i, v_j] \mid x_{ij} = 1\}$. The constraints (11.1) guarantee that for each $z_t \in R_1$, there is a flow y^t of one unit (of the commodity t) from z_1 to z_t. For arcs $[v_i, v_j] \notin T$, the constraints (11.3)–(11.5) guarantee that $y^k_{ij} = y^l_{ij} = \check{y}^{kl}_{ij} = 0$ for all $\{z_k, z_l\} \subseteq R_1$. Therefore, there is no flow over arcs not in T and T contains ways from z_1 to each other terminal.

Now let x describe the arcs of an optimal (directed) tree $\hat{T}$. For each $z_t \in R_1$, there is a path from z_1 to z_t in $\hat{T}$. Set y^t to 1 along this path. For each $\{z_k, z_l\} \subseteq R_1$ set $\check{y}^{kl}_{ij}$ to 1, if $[v_i, v_j]$ is on the path from z_1 to z_k as well as on the path from z_1 to z_l. Obviously, $(x, y, \check{y})$ is feasible for P_{F^2}.

Thus, T itself is an optimal Steiner arborescence.

6.1. Comparing LP_{F^2} with other relaxations

We now compare LP_{F^2} with the two strongest relaxations presented before, namely LP_{2T} and LP_{F+FB}.

Lemma 14. $v(LP_{F^2}) \geqslant v(LP_{2T})$.

Proof. Let $(x, y, \check{y})$ be an (optimal) solution of LP_{F^2}. For all $\{z_k, z_l\} \subseteq R_1$, $k < l$, and $[v_i, v_j] \in A$ define $\grave{y}^{kl}_{ij} := y^k_{ij} - \check{y}^{kl}_{ij}$ and $\acute{y}^{kl}_{ij} := y^l_{ij} - \check{y}^{kl}_{ij}$. Obviously the constraints of LP_{2T} are satisfied. $\square$

Because LP_{F^2} contains the flow-balance constraints and is stronger than LP_{2T}, it is stronger than LP_{2T+FB} (constructed by adding (9.1) to LP_{2T}). It follows directly that LP_{F^2} is also stronger than LP_{F+FB}. The following example shows that it is even strictly stronger than the other stated relaxations.

Example 4. Fig. 4 shows an example for $v(LP_{2T}) < v(LP_{F^2}) = v(P_{F^2})$: Setting all x-variables to 0.5 leads to a feasible (and optimal) solution for LP_{2T} with the value 13.5. An optimal solution for LP_{F^2} is $x_{13} = x_{35} = x_{56} = x_{62} = x_{64} = 1$, which forms a Steiner tree with value 14. Notice that this is also an example with $v(LP_{2T}) < v(LP_{F+FB})$. On the other hand, if v_5 is moved to R, $v(LP_{F+FB}) = v(LP_F) = 12 < v(LP_{2T}) = v(LP_{2T+FB}) = 13.5 < v(LP_{F^2}) = v(P_{F^2})$ (The optimal value for LP_{F+FB} is reached by $\hat{x}$ with

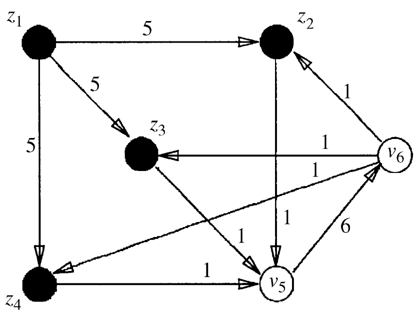

Fig. 4. Example with $v(LP_{2T}) < v(LP_{F+FB}) = v(LP_{F^2}) = v(P_{F^2})$.

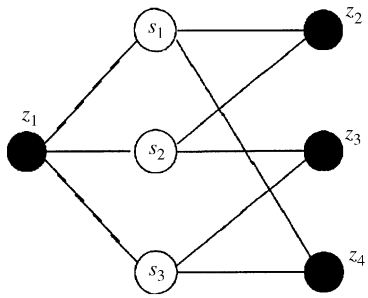

Fig. 5. The value $v(LP_{F^2})$ changes with different roots.

$\hat{x}_{12} = \hat{x}_{13} = \hat{x}_{14} = \hat{x}_{25} = \hat{x}_{35} = \hat{x}_{45} = \frac{1}{3}$, $\hat{x}_{56} = \hat{x}_{62} = \hat{x}_{63} = \hat{x}_{64} = \frac{2}{3}$). Thus, LP_{F+FB} and LP_{2T} are incomparable.

This example has been chosen because it is especially instructive. For $v(LP_{2T+FB}) < v(LP_{F^2})$, as for all other statements in this paper that one relaxation is *strictly* stronger than another, we know also (originally) undirected instances as examples.

Both LP_{F^2} and LP_{2T} make it difficult for flows to two different terminals to split up and rejoin again by increasing the x-variables on arcs with rejoined flow. One could say that rejoining has to be "payed". To get an intuitive impression why LP_{F^2} is strictly stronger than LP_{2T} (or even LP_{2T+FB}), notice that in LP_{F^2}, there is one flow to each terminal and rejoining of each pair of these flows has to be payed; while in LP_{2T}, it is just required that for each pair of terminals there are two flows and rejoining them has to be payed. The latter task is easier; for example it is possible (for given x-values) that for each pair of terminals there are two flows that do not rejoin, but there are not $|R_1|$ flows to all terminals in R_1 that do not rejoin pairwise; this is the case in Example 4 (setting all x-variables to 0.5).

6.2. Choice of the root

The following example shows that the value $v(LP_{F^2})$ is not independent of the choice of the root vertex.

Example 5. The value $v(LP_{F^2})$ changes for different roots: in Fig. 5, choosing z_1 as the root yields the value 4.5 (setting all x-variables in the direction away from z_1 to 0.5 leads to an optimal solution), while choosing z_2, z_3, or z_4 yields 5, which is the weight of a minimum Steiner tree.

7. Conclusion

7.1. A hierarchy of relaxations

Fig. 6 summarizes the relations stated in this paper. All relaxations in the same box are equivalent. A line between two boxes means that the relaxations in the upper box

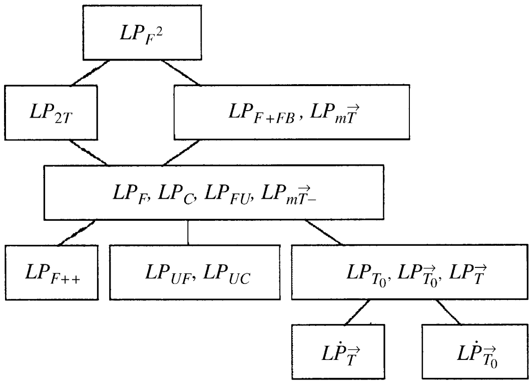

Fig. 6. A hierarchy of relaxations.

are strictly stronger than those in the lower box. Notice that the "strictly stronger" relation is transitive.

7.2. Remarks

It should be mentioned that some of the stated results on the relationship between the optimal values of linear relaxations extend directly to polyhedral results concerning the corresponding feasible sets. This is always the case if optimality is not used in the proofs (e.g. in Lemma 5 or Lemma 6); and hence the feasible set of one relaxation (projected into the x-space) is mapped into the corresponding set of some other. The situation is different in the other cases (e.g. the proofs of Lemma 7 or Theorem 11). Here the assumption of optimality of x can obviously be replaced by the assumption of minimality of x (a feasible x is minimal if there is no feasible $x' \neq x$ with $x' \leqslant x$). In such cases, the presented results extend directly to polyhedral results in the sense of inclusions between the dominants of the corresponding polyhedra (projected into the x-space). (The dominant of Q is $\{x' \mid x' \geqslant x \in Q\}$.)

Note also that polyhedral results concerning the facets of the Steiner tree polyhedron (like those in [4,5]) fall into a different category. Our line of approach in this paper has been studying linear relaxations of general, explicitly given (and frequently used) integer formulations; not methods for describing facet defining inequalities. Applying such descriptions is typically possible only if the graph has certain properties (e.g. that it contains a special substructure) and involves separation problems which are believed to be difficult.

Acknowledgements

We would like to thank the referees for their comments.

References

[1] Y.P. Aneja, An integer linear programming approach to the Steiner problem in graphs, Networks 10 (1980) 167–178.
[2] A. Balakrishnan, N.R. Patel, Problem reduction methods and a tree generation algorithm for the Steiner network problem, Networks 17 (1987) 65–85.
[3] J.E. Beasley, An SST-based algorithm for the Steiner problem in graphs, Networks 19 (1989) 1–16.
[4] S. Chopra, M.R. Rao, The Steiner tree problem I: formulations, compositions and extension of facets, Math. Program. (1994) 209–229.
[5] S. Chopra, M.R. Rao, The Steiner tree problem II: properties and classes of facets, Math. Program. (1994) 231–246.
[6] C.W. Duin, Steiner's Problem in Graphs, Ph.D. Thesis, Amsterdam University, 1993.
[7] J. Edmonds, Matroids and the greedy algorithm, Math. Program. 1 (1971) 127–136.
[8] M.X. Goemans, D.J. Bertsimas, Survivable networks, linear programming relaxations and the parsimonious property, Math. Program. 60 (1993) 145–166.
[9] M.X. Goemans, Y. Myung, A catalog of Steiner tree formulations, Networks 23 (1993) 19–28.
[10] F.K. Hwang, D.S. Richards, P. Winter, in: The Steiner tree problem, Annals of Discrete Mathematics, Vol. 53, North-Holland, Amsterdam, 1992.
[11] R.M. Karp, Reducibility among combinatorial problems, in: R.E. Miller, J.W. Thatcher (Eds.), Complexity of Computer Computations, Plenum Press, New York, 1972, pp. 85–103.
[12] B.N. Khoury, P.M. Pardalos, An exact branch and bound algorithm for the Steiner problem in graphs, in: D. Du, M. Li (Eds.), Proceedings of COCOON'95, Vol. 959, Lecture Notes in Computer Science, Springer, Berlin, 1995, pp. 582–590.
[13] B.N. Khoury, P.M. Pardalos, D.W. Hearn, Equivalent formulations for the Steiner problem in graphs, in: D.Z. Du, P.M. Pardalos (Eds.), Network Optimization Problems, World Scientific Publishing, Singapore, 1993, pp. 111–123.
[14] T. Koch, A. Martin, Solving Steiner tree problems in graphs to optimality, Networks 32 (1998) 207–232.
[15] W. Liu, A lower bound for the Steiner tree problem in directed graphs, Networks 20 (1990) 426–434.
[16] N. Maculan, The Steiner problem in graphs, Ann. Discrete Math. 31 (1987) 185–212.
[17] T.L. Magnanti, L.A. Wolsey, Optimal trees, in: M.O. Ball et al. (Eds.), Handbooks in Operations Research and Management Science, Vol. 7, Elsevier Science, Amsterdam, 1995 (Chapter 9).
[18] C.E. Miller, A.W. Tucker, R.A. Zemlin, Integer programming formulation of traveling salesman problems, J. ACM 7 (4) (1960) 326–329.
[19] T. Polzin, S. Vahdati Daneshmand, Improved Algorithms for the Steiner Problem in Networks, Technical Report 06/1998, Universität Mannheim, 1998, Discrete Appl. Math., to appear.
[20] R.T. Wong, A dual ascent approach for Steiner tree problems on a directed graph, Math. Program. 28 (1984) 271–287.