**Noname manuscript No.**
(will be inserted by the editor)

# An Efficient Cost-Sharing Mechanism for the Prize-Collecting Steiner Forest Problem

**Anupam Gupta · Jochen Könemann · Stefano Leonardi · R. Ravi · Guido Schäfer**

the date of receipt and acceptance should be inserted later

**Abstract** In an instance of the *prize-collecting Steiner forest* problem (PCSF) we are given an undirected graph $G = (V, E)$, non-negative edge-costs $c(e)$ for all $e \in E$, terminal pairs $R = \{(s_i, t_i)\}_{1 \leq i \leq k}$, and penalties $\pi_1, \ldots, \pi_k$. A feasible solution $(F, Q)$ consists of a forest $F$ and a subset $Q$ of terminal pairs such that for all $(s_i, t_i) \in R$ either $s_i, t_i$ are connected by $F$ or $(s_i, t_i) \in Q$. The objective is to compute a feasible solution of minimum cost $c(F) + \pi(Q)$.

A game-theoretic version of the above problem has $k$ players, one for each terminal-pair in $R$. Player $i$'s ultimate goal is to connect $s_i$ and $t_i$, and the player derives a privately held *utility* $u_i \geq 0$ from being connected. A service provider can connect the terminals $s_i$ and $t_i$ of player $i$ in two ways: (1) by

Anupam Gupta
School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: anupamg@andrew.cmu.edu.

Jochen Könemann
Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada. E-mail: jochen@math.uwaterloo.ca. Research supported by NSERC grant no. 288340-2004 and by an IBM Faculty Award.

Stefano Leonardi
Dipartimento di Informatica e Sistemistica, University of Rome "La Sapienza", Via Salaria 113, 00198 Rome, Italy. E-mail: leon@dis.uniroma1.it. Part of this work was done while the author was visiting the School of Computer Science at Carnegie Mellon University.

R. Ravi
Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA. E-mail: ravi@cmu.edu.

Guido Schäfer
Center for Mathematics and Computer Science (CWI), Algorithms, Combinatorics and Optimization, Science Park 123, 1098 XG Amsterdam, The Netherlands, and VU University Amsterdam, Department of Econometrics and Operations Research, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands. E-mail: g.schaefer@cwi.nl. Part of this work was done while the author was visiting the School of Computer Science at Carnegie Mellon University.

buying the edges of an $s_i, t_i$-path in $G$, or (2) by buying an alternate connection between $s_i$ and $t_i$ (maybe from some other provider) at a cost of $\pi_i$.

In this paper, we present a simple 3-budget-balanced and group-strategyproof mechanism for the above problem. We also show that our mechanism computes client sets whose social cost is at most $O(\log^2 k)$ times the minimum social cost of any player set. This matches a lower-bound that was recently given by Roughgarden and Sundararajan (STOC '06).

## 1 Introduction

In an instance of the *prize-collecting Steiner forest* problem (PCSF) we are given an undirected graph $G = (V, E)$ with edge costs $c : E \to \mathbb{R}^+$, a set of $k$ terminal pairs $R = \{(s_i, t_i)\}_{1 \leq i \leq k}$, and penalties $\pi : R \to \mathbb{R}^+$. A feasible solution $(F, Q)$ consists of a forest $F$ and a subset $Q$ of terminal pairs such that for all $(s_i, t_i) \in R$ either $s_i, t_i$ are connected by $F$ or $(s_i, t_i) \in Q$. The objective is to compute a feasible solution of minimum cost $c(F) + \pi(Q)$.

A game-theoretic version of the above problem has $k$ players, one for each terminal-pair in $R$. We use $U$ to denote the set of all players. Player $i$'s ultimate goal is to connect $s_i$ and $t_i$, and the player derives a privately held *utility* $u_i \geq 0$ from being connected. A service provider can connect the terminals $s_i$ and $t_i$ of player $i$ in two ways: (1) by buying the edges of an $s_i, t_i$-path in $G$, or (2) by buying an alternate connection between $s_i$ and $t_i$ (maybe from some other provider) at a cost of $\pi_i$.

Formally, we are interested in finding a *cost-sharing mechanism* that first solicits bids $\{b_i\}_{i \in U}$ from all players. The mechanism then determines a set $S \subseteq U$ of players to service and computes a prize-collecting Steiner forest for the terminal set of these players. Finally, the mechanism needs to determine a payment $x_i(S) \leq b_i$ for each of the players in $S$.

There are several desirable properties of a cost-sharing mechanism: a mechanism is called *strategyproof*, if bidding truthfully (i.e., announcing $b_i = u_i$) is a dominant strategy for all players. If this is true even if players are permitted to collude, then we call a mechanism *group-strategyproof*. A mechanism is *budget balanced* if the total cost $C(S)$ of servicing the players in $S$ is at most the sum of the costs charged to the players in $S$, and it is *competitive* if the sum of all costs charged to the players in $S$ does not exceed the cost of an optimal PCSF solution for $S$. A mechanism is called *efficient* if it selects a set $S$ of players that maximizes $u(S) - C(S)$.

Classical results in economics [12,26] state that budget balance and efficiency cannot be simultaneously achieved by any mechanism. Moreover, Feigenbaum et al. [10] recently showed that there is no strategyproof mechanism that always recovers a constant fraction of the maximum efficiency and a constant fraction of the incurred cost even for the simple fixed-tree multicast problem.

In light of these hardness results, most of the previous work on mechanism design concentrated on proper subsets of the above design goals. One notable class of such mechanisms are based on a framework due to Moulin and Shenker [24]. The authors showed that, given a budget balanced and *cross-monotonic cost sharing method* for the underlying problem, the well known Moulin mechanism [23] satisfies budget balance and group-strategyproofness. Moulin and Shenker's framework has recently been applied to game-theoretic variants of classical optimization problems such as fixed-tree multicast [2,9,10], submodular cost-sharing [24], Steiner trees [17,18], facility location, single-source rent-or-buy network design [25,22,14] and Steiner forests [21]. Lower bounds on the budget balance factor that is achievable by a cross-monotonic cost sharing mechanism are given in [16,20,21].

Recently, Roughgarden and Sundararajan [29] introduced an alternative measure of efficiency that circumvents the intractability results in [10,12,26] at least partially. Let $U$ be a universe of players and let $C$ be a cost function on $U$ that assigns to each subset $S \subseteq U$ a non-negative service cost $C(S)$. The authors define the *social cost* $\Pi(S)$ of a set $S \subseteq U$ as $\Pi(S) = u(U \setminus S) + C(S)$. A mechanism is said to be $\alpha$-*approximate* if the set it outputs has social cost at most $\alpha$ times the minimum over all sets $S \subseteq U$. The intuition for this definition loosely comes from the fact that $u(U) - \Pi(S) = u(S) - C(S)$, which is the traditional definition of efficiency; since $u(U)$ is a constant, a set $S$ has minimum social cost iff it has maximum efficiency.

Roughgarden and Sundararajan then developed a framework to quantify the extent to which a Moulin mechanism minimizes the social cost, and apply this framework to show that the Shapley mechanism is $O(\log k)$-approximate for submodular functions, and that the Steiner tree cost-shares of Jain and Vazirani [17] give a mechanism that is $O(\log^2 k)$-approximate. In a later result Chawla, Roughgarden and Sundararajan [7] applied the framework to show that the cost-shares of [21] are $O(\log^2 k)$-approximate for Steiner forests. In a following paper, Roughgarden and Sundararajan [28] proved polylogarithmic upper an lower bounds for uniform facility location, Steiner tree problems and rent-or-buy network design problems.

## 1.1 Prize-Collecting Steiner Problems

Computing minimum-cost prize-collecting Steiner trees or forests is APX-complete [3,5], and hence neither of the two problems admits a PTAS unless P = NP. The first constant-factor approximation for the prize-collecting Steiner tree problem was a LP-rounding based 3-approximation by Bienstock et al. [6], and this was improved to $2-1/k$ by Goemans and Williamson [11] using the primal-dual schema. One can easily modify the algorithm of Bienstock et al. to give a 3-approximation for the PCSF problem as well; in [15], Hajiaghayi and Jain refine Bienstock's LP rounding idea and obtain an LP-based 2.54 approximation for the problem. The authors also present a primal-dual combinatorial 3-approximation for the problem. This algorithm substantially deviates from the classical framework of Goemans and Williamson, requiring crucial use of Farkas' Lemma, wherein the dual variables are both increased and decreased along the execution of the algorithm.

## 1.2 Our Results and Techniques

The first contribution of this paper is the following:

**Theorem 1** *There is an efficiently computable cross-monotonic cost sharing method $\xi^{GKLRS}$ for the prize-collecting Steiner forest problem that is 3-budget balanced.*

Our algorithm *GKLRS* is a natural extension of the primal-dual algorithm of Goemans and Williamson [11] for prize-collecting Steiner trees and the cross-monotonic cost sharing method *KLS* for Steiner forests presented in [21]. Despite its simplicity, our algorithm achieves the same approximation guarantee as [15].

Our second result bounds the social cost of the mechanism associated with the cost-sharing method.

**Theorem 2** *The Moulin mechanism $M(\xi^{GKLRS})$ driven by the cross-monotonic cost sharing method $\xi^{GKLRS}$ is $\Theta(\log^2 k)$-approximate.*

This result is achieved in two steps. The first step is to show that if the Moulin mechanism $M(\xi^{KLS})$ is $\alpha$-approximate then the mechanism $M(\xi^{GKLRS})$ given by our cross-monotonic cost-sharing method $\xi^{GKLRS}$ is $3(1+\alpha)$-approximate for the prize-collecting Steiner forest game. As the second step, we show that the *KLS* mechanism is $O(\log^3 k)$-approximate for the Steiner Forest game. This is achieved by adding a novel methodological contribution to the framework proposed in [29]: we show that such a result can also be proved by embedding the graph distances into random HSTs [4,8] rather than using the construction proposed by Roughgarden and Sundararajan. Independently, Chawla, Roughgarden and Sundararajan [7] have shown (using a more involved analysis) that *KLS* is $O(\log^2 k)$-approximate. We are optimistic that the general idea of reductions between cost-sharing mechanisms that we use in our proof will extend to the prize-collecting versions of other optimization problems.

## 1.3 Organization of the Paper

In Section 2 we introduce some notations used in the paper. In Section 3 we present the linear programming formulation for PCSF. Section 4 presents the cross-monotonic cost-sharing scheme *GKLRS* for PCSF. In Section 5 we prove the bound on the social cost for the *GKLRS* mechanism, whereas in Section 6 we prove the bound on the social cost for the Steiner forest mechanism *KLS*.

## 2 Preliminaries

Let $U$ be a universe of players and let $C$ be a cost function on $U$ that assigns to each subset $S \subseteq U$ a non-negative cost $C(S)$. We assume that $C$ is non-decreasing, i.e., for all $S \subseteq T$, $C(S) \leq C(T)$, and $C(\emptyset) = 0$.

### 2.1 Moulin Mechanisms

A *cost sharing method* $\xi$ is an algorithm that, given any subset $S \subseteq U$ of players, computes a solution to service $S$ and for each $i \in S$ determines a non-negative cost share $\xi_i(S)$. We say that $\xi$ is *$\beta$-budget balance* if for every subset $S \subseteq U$,

$$\frac{1}{\beta} \cdot C(S) \leq \sum_{i \in S} \xi_i(S) \leq C(S).$$

A cost sharing method $\xi$ is *cross-monotonic* if for any two sets $S$ and $T$ such that $S \subseteq T$ and any player $i \in S$ we have $\xi_i(S) \geq \xi_i(T)$.

Moulin and Shenker [24] showed that, given a budget balanced and cross-monotonic cost sharing method $\xi$ for the underlying problem, the following cost sharing mechanism $M(\xi)$ satisfies budget-balance and group-strategyproofness: Initially, let $S = U$. If for each player $i \in S$ the cost share $\xi_i(S)$ is at most her bid $b_i$, we stop. Otherwise, remove from $S$ all players whose cost shares are larger than their bids, and repeat. Eventually, let $\xi_i(S)$ be the costs that are charged to players in the final set $S$.

### 2.2 Approximating Social Cost

Roughgarden and Sundararajan [29] recently introduced an alternative notion of efficiency for cost sharing mechanisms: Every player $i \in U$ has a private utility $u_i$. For a set $S \subseteq U$, define $u(S) = \sum_{i \in S} u_i$. Define the *social cost* $\Pi(S)$ of a set $S \subseteq U$ as

$$\Pi(S) = u(U \setminus S) + C(S).$$

**Definition 1** Suppose $S^M$ is the final set of players computed by the Moulin mechanism $M(\xi)$ on $U$. Then $M(\xi)$ is said to be *$\alpha$-approximate* if

$$\Pi(S^M) \leq \alpha \cdot \Pi(S) \quad \forall S \subseteq U.$$

Roughgarden and Sundararajan [29] proved that the Moulin mechanism $M(\xi)$ is $(\alpha + \beta)$-approximate and $\beta$-budget balanced if $\xi$ is *$\alpha$-summable* and $\beta$-budget balanced. The summability of a cost sharing method is defined as follows: Assume we are given an arbitrary permutation $\sigma$ on the players in $U$ and a subset $S \subseteq U$ of players. We assume that the players in $S$ are ordered according to $\sigma$, i.e., $S = \{i_1, \ldots, i_{|S|}\}$ where $i_j \prec_\sigma i_k$ if and only if $1 \leq j < k \leq |S|$. We define $S_j \subseteq S$ as the (ordered) set of the first $j$ players of $S$ according to the order $\sigma$.

**Definition 2** A cost sharing method $\xi$ is *$\alpha$-summable* if for every ordering $\sigma$ and every subset $S \subseteq U$

$$\sum_{j=1}^{|S|} \xi_{i_j}(S_j) \leq \alpha \cdot C(S). \tag{1}$$

where $S_j$ is the set of the first $j$ players, and $i_j$ is the $j^{th}$ player according to the ordering $\sigma$.

## 3 LP Formulation

Subsequently, we slightly abuse notation by using $R$ to refer to the set of terminal pairs and the set of terminals. For a terminal $u \in R$, let $\bar{u}$ be the mate of $u$, i.e., $(u, \bar{u}) \in R$. For a terminal pair $(u, \bar{u}) \in R$, define the *death time* as $\mathtt{d}(u, \bar{u}) = \frac{1}{2} d_G(u, \bar{u})$, where $d_G(u, \bar{u})$ is the cost of a shortest $u, \bar{u}$-path (with respect to $c$) in $G$.

Consider a cut $S \subseteq V$. We say $S$ *separates* a terminal pair $(u, \bar{u}) \in R$ iff $|\{u, \bar{u}\} \cap S| = 1$. We also write $(u, \bar{u}) \odot S$ iff $(u, \bar{u})$ is separated by $S$. A cut $S$ that separates at least one terminal pair is called a *Steiner cut*. Let $\mathcal{S}$ denote the set of all Steiner cuts. For a cut $S \subseteq V$, we use $\delta(S)$ to refer to the set of all edges $(u, v) \in E$ that *cross* $S$, i.e., $\delta(S) = \{(u, v) \in E : |\{u, v\} \cap S| = 1\}$.

A natural integer programming formulation for PCSF has a 0/1-variable $x_e$ for all edges $e \in E$ and a 0/1-variable $x_{u\bar{u}}$ for all terminal pairs $(u, \bar{u}) \in R$. Variable $x_e = 1$ iff $e \in F$ and $x_{u\bar{u}} = 1$ iff $(u, \bar{u}) \in Q$. The following is an integer programming formulation for PCSF:

$$\min \sum_{e \in E} c(e) \cdot x_e + \sum_{(u,\bar{u}) \in R} \pi(u, \bar{u}) \cdot x_{u\bar{u}} \tag{ILP}$$

$$\text{s.t.} \sum_{e \in \delta(S)} x_e + x_{u\bar{u}} \geq 1 \qquad \forall S \in \mathcal{S},\ \forall (u, \bar{u}) \odot S \tag{2}$$

$$x_e,\ x_{u\bar{u}} \in \{0, 1\} \quad \forall e \in E,\ \forall (u, \bar{u}) \in R.$$

We use $\mathit{OPT}_R$ to refer to the cost of an optimal solution to this LP. Constraint (2) ensures that each Steiner cut $S \in \mathcal{S}$ is either crossed by an edge of $F$, or all separated terminal pairs $(u, \bar{u}) \odot S$ are part of $Q$.

The dual of the linear programming relaxation (LP) of (ILP) is as follows. We have a non-negative dual variable $\xi_{S,u\bar{u}}$ for all Steiner cuts $S \in \mathcal{S}$ and all pairs $(u, \bar{u}) \in R$ such that $(u, \bar{u}) \odot S$:

$$\max \sum_{S \in \mathcal{S}} \sum_{(u,\bar{u}) \odot S} \xi_{S,u\bar{u}} \tag{D}$$

$$\text{s.t.} \sum_{S \in \mathcal{S}: e \in \delta(S)} \sum_{(u,\bar{u}) \odot S} \xi_{S,u\bar{u}} \leq c(e) \qquad \forall e \in E \tag{3}$$

$$\sum_{S \in \mathcal{S}: S \odot (u,\bar{u})} \xi_{S,u\bar{u}} \leq \pi(u,\bar{u}) \quad \forall (u,\bar{u}) \in R \tag{4}$$

$$\xi_{S,u\bar{u}} \geq 0 \qquad \forall S \in \mathcal{S},\ (u,\bar{u}) \odot S.$$

It is convenient to associate a dual solution $\{\xi_{S,u\bar{u}}\}_{S \in \mathcal{S},(u,\bar{u}) \odot S}$ with a set of dual values $\{y_S\}_{S \in \mathcal{S}}$ for all Steiner cuts $S \in \mathcal{S}$. To this aim, we define the dual $y_S$ of a Steiner cut $S \in \mathcal{S}$ simply as the total cost share of all its separated terminal pairs:

$$y_S = \sum_{(u,\bar{u}) \odot S} \xi_{S,u\bar{u}}.$$

We can think of $\xi_{S,u\bar{u}}$, $(u,\bar{u}) \odot S$, as a *cost share* that terminal pair $(u,\bar{u})$ receives from dual $y_S$ of $S$. Define the total cost share of $(u,\bar{u})$ as

$$\xi_{u\bar{u}} = \sum_{S \in \mathcal{S}: S \odot (u,\bar{u})} \xi_{S,u\bar{u}}.$$

Clearly, with these definitions

$$\sum_{S \in \mathcal{S}} y_S = \sum_{(u,\bar{u}) \in R} \xi_{u\bar{u}}.$$

Constraint (3) of LP (D) requires that for every edge $e \in E$, the total dual of all Steiner cuts $S \in \mathcal{S}$ that cross $e$ is at most the cost $c(e)$ of this edge. Constraint (4) states that the total cost share $\xi_{u\bar{u}}$ of terminal pair $(u,\bar{u})$ is at most its penalty $\pi(u,\bar{u})$.

## 4 A Cross-Monotonic Algorithm for the PCSF Problem

Our algorithm *GKLRS* for the prize-collecting Steiner forest problem is a primal-dual algorithm, that is, it maintains a primal solution $\{x_e, x_{u\bar{u}}\}_{e \in E,(u,\bar{u}) \in R}$ together with a set of dual values $\{y_S\}_{S \in \mathcal{U}}$ (the definition of the set $\mathcal{U}$ is given below). The primal solution is a 0/1-solution that is infeasible for (LP) initially. Throughout the execution of *GKLRS*, the degree of infeasibility of this solution is decreased successively until eventually, we obtain a feasible solution for (LP).

A subtle point of our algorithm is that it does not produce a set of dual values $\{y_S\}_{S \in \mathcal{U}}$ that corresponds to a feasible solution for (D). There are two reasons for this. First, we also raise dual values $y_S$ of cuts $S$ that do not correspond to Steiner cuts. We use $\mathcal{U}$ to refer to the set of all cuts that are raised throughout the execution of *GKLRS*. As a consequence, a terminal pair $(u, \bar{u})$ may receive cost share $\xi_{S,u\bar{u}}$ from a non-Steiner cut $S \in \mathcal{U} \setminus \mathcal{S}$. Second, a terminal pair $(u, \bar{u})$ may also receive cost share $\xi_{S,u\bar{u}}$ from a cut $S$ that does not separate $(u, \bar{u})$. However, *GKLRS* maintains the invariant that a terminal pair $(u, \bar{u})$ only receives cost share from cuts $S \in \mathcal{U}$ that either separate or entirely contain $(u, \bar{u})$, i.e., $(u, \bar{u}) \odot S$ or $\{u, \bar{u}\} \subseteq S$.

We can view the execution of *GKLRS* as a process over time. Initially, at time $\tau = 0$, $x_e^\tau = 0$ for all $e \in E$, $x_{u\bar{u}}^\tau = 0$ for all $(u, \bar{u}) \in R$ and $y_S^\tau = 0$ for all $S \in \mathcal{U}$. Let $F^\tau$ be the forest that corresponds to $\{x_e^\tau\}_{e \in E}$, i.e., $F^\tau = \{e \in E : x_e^\tau = 1\}$. Similarly, let $Q^\tau$ be the set of all terminal pairs $(u, \bar{u}) \in R$ such that $x_{u\bar{u}}^\tau = 1$.

We define $\bar{F}^\tau$ as the set of all edges that are tight at time $\tau$, i.e.,

$$\bar{F}^\tau = \{e \in E : \sum_{S \in \mathcal{U}} y_S^\tau = c(e)\}.$$

We use the term *moat* to refer to a connected component $M^\tau$ in $\bar{F}^\tau$. A moat $M^\tau$ defines a cut $S$ which is simply the set of vertices spanned by $M^\tau$. At time $\tau$, we increase the duals of all cuts defined by moats $M^\tau \in \bar{F}^\tau$ that are *active* at time $\tau$. The notion of activity will be defined shortly. These duals are increased simultaneously and by the same amount. Subsequently, we also say that we *grow* all active moats in $\bar{F}^\tau$ at time $\tau$. Moreover, it is convenient to regard the growing of moats as being identical to increasing the duals.

## 4.1 Activity Notion

We call a terminal pair $(u, \bar{u}) \in R$ *active* at time $\tau$ if

$$\xi_{u\bar{u}}^\tau < \pi(u, \bar{u}) \quad \text{and} \quad \tau < \mathtt{d}(u, \bar{u}). \tag{5}$$

If the above conditions do not hold, we say that $(u, \bar{u})$ is *inactive* at time $\tau$. Let $\tau_{u\bar{u}}$ be the first time when $(u, \bar{u})$ becomes inactive. Observe that by definition (5), a terminal pair $(u, \bar{u})$ remains inactive at all times $\tau > \tau_{u\bar{u}}$. A terminal $u \in R$ is active at time $\tau$ if its pair $(u, \bar{u})$ is active at this time. Let $\mathcal{A}^\tau$ be the set of all terminals that are active at time $\tau$.

We say that a moat $M^\tau \in \bar{F}^\tau$ is active at time $\tau$ if it contains at least one active terminal, i.e., $M^\tau \cap \mathcal{A}^\tau \neq \emptyset$. The growth of an active moat $M^\tau$ is shared evenly among all active terminals in $M^\tau$. Let $M^\tau(u)$ denote the moat in $\bar{F}^\tau$ that contains terminal $u \in R$. More formally, we define the cost share $\xi_u^{\tau'}$ of a terminal $u \in R$ at time $\tau' \leq \tau_{u\bar{u}}$ as follows:

$$\xi_u^{\tau'} = \int_0^{\tau'} \frac{1}{|M^\tau(u) \cap \mathcal{A}^\tau|} d\tau. \tag{6}$$

Let $\xi_u^{\tau'} = \xi_u^{\tau_{u\bar{u}}}$ for all $\tau' > \tau_{u\bar{u}}$. Moreover, we define $\xi_{u\bar{u}}^\tau = \xi_u^\tau + \xi_{\bar{u}}^\tau$.

Observe that the total contribution to the cost share of a terminal pair $(u, \bar{u})$ within $\epsilon$ time units is at most $2\epsilon$. Also, note that $(u, \bar{u})$ may receive cost share from a moat $M^\tau$ that contains $u$ and $\bar{u}$.

The following fact follows immediately from definitions (5) and (6).

**Fact 3** *For all terminal pairs* $(u, \bar{u}) \in R$, $\xi_{u\bar{u}} \leq \min\{\pi(u, \bar{u}), 2\mathtt{d}(u, \bar{u})\}$.

Since at any point of time, the growth of all active moats is shared among active terminals, the following must hold true.

**Fact 4** *For every time* $\tau \geq 0$,

$$\sum_{S \in \mathcal{U}} y_S^\tau = \sum_{(u,\bar{u}) \in R} \xi_{u\bar{u}}^\tau.$$

We say that two active moats $M_1$ and $M_2$ *collide* at time $\tau$ if their vertices are contained in the same connected component of $\bar{F}^{\tau'}$ iff $\tau' \geq \tau$. In this case, we add a cheapest collection of edges to $F^\tau$ s.t. all active vertices of $M_1$ and $M_2$ are in the same connected component of $F^{\tau'}$ for all $\tau' \geq \tau$.

Suppose a terminal pair $(u, \bar{u}) \in R$ becomes inactive at time $\tau = \tau_{u\bar{u}}$ because it reaches its penalty, i.e., $\xi_{u\bar{u}}^\tau = \pi(u, \bar{u})$. We then add $(u, \bar{u})$ to $Q^\tau$. Since $(u, \bar{u})$ remains inactive after time $\tau_{u\bar{u}}$, the following fact holds true.

**Fact 5** *Let $Q$ be the final set of terminal pairs computed by* GKLRS*. Then*

$$\sum_{(u,\bar{u}) \in Q} \pi(u, \bar{u}) = \sum_{(u,\bar{u}) \in Q} \xi_{u\bar{u}}$$

Suppose a terminal pair $(u, \bar{u})$ becomes inactive at time $\mathtt{d}(u, \bar{u})$. The next fact shows that $(u, \bar{u})$ must then be connected in $F$.

**Fact 6** *Let terminal pair $(u, \bar{u})$ become inactive just after time $\mathtt{d}(u, \bar{u})$. Then $u$ and $\bar{u}$ are connected in $F$.*

*Proof* Let $P_{u\bar{u}}$ be a shortest $u, \bar{u}$-path in $G$. Path $P_{u\bar{u}}$ becomes tight at time $\tau \leq \mathtt{d}(u, \bar{u})$ and both $u$ and $\bar{u}$ are active at this time. Thus either $u$ and $\bar{u}$ are already connected in $F^\tau$ or $P_{u\bar{u}}$ is added to $F^\tau$. □

Observe that the last fact also establishes correctness of *GKLRS*: The final solution $(F, Q)$ computed by *GKLRS* is a feasible solution for the given prize-collecting Steiner forest instance.

Subsequently, we use $\xi^{GKLRS}(S)$ to refer to final cost shares computed by *GKLRS* when run on terminal set $S \subseteq R$. We also identify the player set $U$ with the terminal-pair set $R$.

## 4.2 Cross-Monotonicity

We compare the execution of *GKLRS* on terminal set $R$ with the one on terminal set $R_{-st} = R \setminus \{(s,t)\}$ for any $(s,t) \in R$. We use $\mathcal{G}_{-st}$ ($\mathcal{G} =$ *GKLRS*, $F$, $\bar{F}$, $M$, etc.) to refer to $\mathcal{G}$ in the run of *GKLRS* on $R_{-st}$. For notational convenience, let $\xi_{-st}(u,\bar{u})$ refer to the cost share of $(u,\bar{u})$ in the run of *GKLRS* on $R_{-st}$ and let $\xi(u,\bar{u})$ refer to the respective cost share in *GKLRS* on $R$.

**Lemma 1** *Consider the execution of GKLRS on $R$ and $R_{-st}$, respectively. The following holds for every time $\tau \geq 0$:*

1. *$\bar{F}^{\tau}_{-st}$ is a refinement of $\bar{F}^{\tau}$, i.e., $\bar{F}^{\tau}_{-st} \subseteq \bar{F}^{\tau}$.*
2. *For all $(u,\bar{u}) \in R_{-st}$, $\xi^{\tau}_{-st}(u,\bar{u}) \geq \xi^{\tau}(u,\bar{u})$.*

*Proof* We prove the lemma by induction over time $\tau$. Clearly, the lemma holds at time $\tau = 0$. Suppose the lemma holds at time $\tau$.

The only moats that may potentially violate the claim $\bar{F}^{\tau+\epsilon}_{-st} \subseteq \bar{F}^{\tau+\epsilon}$ at time $\tau+\epsilon$ for some $\epsilon > 0$, are those that are active at time $\tau$ in *GKLRS*$_{-st}$. Let $M_{-st} \in \bar{F}^{\tau}_{-st}$ be a moat that is active at time $\tau$. By the induction hypothesis, there exists a moat $M \in \bar{F}^{\tau}$ such that $M_{-st} \subseteq M$. We argue that $M$ must be active at time $\tau$ in *GKLRS*.

Since $M_{-st}$ is active at time $\tau$, there must exist a terminal $u \in M_{-st}$ such that $\pi(u,\bar{u}) - \xi^{\tau}_{-st}(u,\bar{u}) > 0$ and $\tau < \mathtt{d}(u,\bar{u})$. By our induction hypothesis,

$$\pi(u,\bar{u}) - \xi^{\tau}(u,\bar{u}) \geq \pi(u,\bar{u}) - \xi^{\tau}_{-st}(u,\bar{u}) > 0.$$

Therefore, $M$ must be active at time $\tau$ too. This proves the first part of the lemma.

It remains to be shown that $\xi^{\tau+\epsilon}_{-st}(u,\bar{u}) \geq \xi^{\tau+\epsilon}(u,\bar{u})$ for all $(u,\bar{u}) \in R_{-st}$. Observe that all terminal pairs that are inactive at time $\tau$ do not receive any further cost share. Consider a terminal pair $(u,\bar{u}) \in R_{-st}$ that is active at time $\tau$ in *GKLRS*$_{-st}$ and let $M^{\tau}_{-st}(u)$ be the moat of $u$ at time $\tau$. From the discussion above, we know that every terminal pair $(v,\bar{v}) \in R_{-st}$ that is active at time $\tau$ in *GKLRS*$_{-st}$ must be active at time $\tau$ in *GKLRS*, i.e., $\mathcal{A}^{\tau}_{-st} \subseteq \mathcal{A}^{\tau}$. By our induction hypothesis, moat $M^{\tau}_{-st}(u)$ is contained in the moat $M^{\tau}(u) \in \bar{F}^{\tau}$ of $u$ in *GKLRS*. Therefore, $|M^{\tau}_{-st}(u) \cap \mathcal{A}^{\tau}_{-st}| \leq |M^{\tau}(u) \cap \mathcal{A}^{\tau}|$. Thus, the additional cost share that $(u,\bar{u})$ receives in the time interval $(\tau, \tau+\epsilon]$ in *GKLRS*$_{-st}$ is at least as large as the one it receives in *GKLRS*. □

## 4.3 Competitiveness

We next show that the total cost share of all terminal pairs is at most the cost of an optimal solution to the prize-collecting Steiner forest instance. The following proof is similar to the one presented in [21].

**Lemma 2** *Let $(F^*, Q^*)$ be an optimal solution to the prize-collecting Steiner forest instance with terminal pair set $R$. The cost shares $\xi$ computed by* GKLRS *for $R$ satisfy*

$$\sum_{(u,\bar{u})\in R} \xi_{u\bar{u}} \leq c(F^*) + \pi(Q^*).$$

*Proof* Consider a separated terminal pair $(u, \bar{u}) \in Q^*$. By Fact 3, we have

$$\sum_{(u,\bar{u})\in Q^*} \xi_{u\bar{u}} \leq \pi(Q^*).$$

It remains to be shown that the total cost share of all terminal pairs $(u, \bar{u}) \in R \setminus Q^*$ is bounded by $c(F^*)$.

Consider a connected component $T \in F^*$ and let $R(T)$ be the set of terminal pairs that are connected by $T$. We prove that

$$\sum_{(u,\bar{u})\in R(T)} \xi_{u\bar{u}} \leq c(T). \tag{7}$$

The lemma follows by summing over all connected components $T \in F^*$.

We define $\mathcal{M}^\tau(T) \subseteq \bar{F}^\tau$ as the set of moats at time $\tau$ that contain at least one active terminal of $R(T)$, i.e.,

$$\mathcal{M}^\tau(T) = \{M^\tau(u) \,:\, u \in R(T) \cap \mathcal{A}^\tau\}.$$

Among all terminal pairs in $R(T)$, let $(w, \bar{w})$ be a pair that is active longest. By our definition of activity in (5), all terminal pairs in $R(T)$ are inactive after time $\mathtt{d}(w, \bar{w})$. We show that the total growth of $\mathcal{M}^\tau(T)$ for all $\tau \in [0, \mathtt{d}(w, \bar{w})]$ is at most $c(T)$. This implies (7).

At any time $\tau$, the moats in $\mathcal{M}^\tau(T)$ are disjoint. Moreover, $T$ connects all terminals in $R(T)$. Thus, if there exists a moat $M^\tau \in \mathcal{M}^\tau(T)$ that intersects an edge of $T$ then each moat in $\mathcal{M}^\tau(T)$ must intersect an edge of $T$; we say that the moats in $\mathcal{M}^\tau(T)$ *load* $T$. Moreover, each moat $M^\tau$ loads a different part of $T$. Thus, the total growth of moats in $\mathcal{M}^\tau(T)$ for all $\tau$ at which $\mathcal{M}^\tau(T)$ loads $T$ is at most $c(T)$.

Let $\tau_0 \leq \mathtt{d}(w, \bar{w})$ be the first time such that $\mathcal{M}^{\tau_0}(T)$ does not load $T$. If $\mathcal{M}^{\tau_0}(T) = \emptyset$, we are done. Otherwise, we must have that $\mathcal{M}^{\tau_0}(T) = \{M^{\tau_0}\}$ and $T \subseteq M^{\tau_0}$. The additional growth of $M^\tau$ for all times $\tau \in [\tau_0, \mathtt{d}(w, \bar{w})]$ is at most $\mathtt{d}(w, \bar{w}) - \tau_0$. Since $w$ and $\bar{w}$ are connected by $T$, this additional growth is at most $\mathtt{d}(w, \bar{w}) \leq c(T)/2$. This gives an upper bound of $\frac{3}{2}c(T)$ on the total cost shares of pairs in $R(T)$.

The following refined argument proves (7). Let $P_{w\bar{w}}$ be the unique $w, \bar{w}$-path in $T$. Define $\mathcal{M}^\tau_{w\bar{w}} \subseteq \mathcal{M}^\tau(T)$ as the set of active moats different from $M^\tau(w)$ and $M^\tau(\bar{w})$ that load $P_{w\bar{w}}$ at time $\tau < \tau_0$, i.e.,

$$\mathcal{M}^\tau_{w\bar{w}} = \{M^\tau \in \mathcal{M}^\tau(T) \setminus \{M^\tau(w), M^\tau(\bar{w})\} \,:\, \delta(M^\tau) \cap P_{w\bar{w}} \neq \emptyset\}.$$

Define the *degree* $\mathtt{deg}(M^\tau)$ of a moat $M^\tau \in \mathcal{M}^\tau_{w\bar{w}}$ as

$$\mathtt{deg}(M^\tau) = |\delta(M^\tau) \cap P_{w\bar{w}}|.$$

**Proposition 1** *Consider a time $\tau < \tau_0$ and a moat $M^\tau \in \mathcal{M}^\tau_{w\bar{w}}$. Then* $\mathtt{deg}(M^\tau) \geq 2$.

*Proof* Both $M^\tau(w)$ and $M^\tau(\bar{w})$ are active at time $\tau < \tau_0$ and thus $\{M^\tau(w), M^\tau(\bar{w})\} \subseteq \mathcal{M}^\tau(T)$ (possibly $M^\tau(w) = M^\tau(\bar{w})$). By definition of $\mathcal{M}^\tau_{w\bar{w}}$, $M^\tau \in \mathcal{M}^\tau(T)$ and $M^\tau \notin \{M^\tau(w), M^\tau(\bar{w})\}$. Furthermore, $M^\tau$ is disjoint from all other moats in $\mathcal{M}^\tau(T)$. Suppose $|M^\tau \cap P_{w\bar{w}}| = 1$. But then, moat $M^\tau$ must contain $w$ or $\bar{w}$. This contradicts the disjointness of $M^\tau$ and $\{M^\tau(w), M^\tau(\bar{w})\}$. □

By our choice of $(w, \bar{w}) \in R(T)$ as the terminal pair with largest activity time and by our assumption that $\mathcal{M}^{\tau_0}(T) \neq \emptyset$ it follows that both, $M^\tau(w)$ and $M^\tau(\bar{w})$ are active for all $0 \leq \tau \leq \tau_0$. We define $l_{w\bar{w}}$ as the total dual growth of the moats containing $w$ and $\bar{w}$ up to time $\tau_0$. Formally, let

$$\delta^\tau_{w\bar{w}} = \begin{cases} 2 & : \quad M^\tau(w) \neq M^\tau(\bar{w}) \\ 1 & : \quad \text{otherwise} \end{cases}$$

and

$$l_{w\bar{w}} = \int_0^{\tau_0} \delta^\tau_{w\bar{w}} d\tau.$$

It follows that the cost of path $P_{w\bar{w}}$ is at least

$$l_{w\bar{w}} + \int_0^{\tau_0} \sum_{M^\tau \in \mathcal{M}^\tau_{w\bar{w}}} \mathtt{deg}(M^\tau) d\tau.$$

We let $\mathtt{slack}_{w\bar{w}}$ be the difference between $c(P_{w\bar{w}})$ and the above term and obtain

$$c(P_{w\bar{w}}) = l_{w\bar{w}} + \mathtt{slack}_{w\bar{w}} + \int_0^{\tau_0} \sum_{M^\tau \in \mathcal{M}^\tau_{w\bar{w}}} \mathtt{deg}(M^\tau) d\tau. \tag{8}$$

We define the total growth $y^{\tau_0}(T)$ produced by terminal pairs in $R(T)$ until time $\tau_0$ as follows:

$$y^{\tau_0}(T) = \int_0^{\tau_0} |\mathcal{M}^\tau(T)| d\tau.$$

At all times $\tau \leq \tau_0$, each moat in $\mathcal{M}^\tau(T)$ loads at least one distinct edge of $T$; those in $\mathcal{M}^\tau_{w\bar{w}}$ load at least two edges of $T$. Thus, we have

$$c(T) \geq y^{\tau_0}(T) + \mathtt{slack}_{w\bar{w}} + \int_0^{\tau_0} \sum_{M^\tau \in \mathcal{M}^\tau_{w\bar{w}}} (\mathtt{deg}(M^\tau) - 1) d\tau. \tag{9}$$

The additional growth between time $\tau_0$ and $\mathtt{d}(w, \bar{w})$ is at most $\mathtt{d}(w, \bar{w}) - \tau_0$. Using (8), we obtain

$$\begin{aligned} \mathtt{d}(w, \bar{w}) - \tau_0 &\leq \frac{l_{w\bar{w}}}{2} - \tau_0 + \frac{\mathtt{slack}_{w\bar{w}}}{2} \\ &\quad + \int_0^{\tau_0} \sum_{M^\tau \in \mathcal{M}^\tau_{w\bar{w}}} \frac{\mathtt{deg}(M^\tau)}{2} d\tau \\ &\leq \frac{\mathtt{slack}_{w\bar{w}}}{2} + \int_0^{\tau_0} \sum_{M^\tau \in \mathcal{M}^\tau_{w\bar{w}}} (\mathtt{deg}(M^\tau) - 1) d\tau, \end{aligned}$$

where we exploit that $\mathtt{deg}(M^\tau) \geq 2$ for all $M^\tau \in \mathcal{M}^\tau_{w\bar{w}}$ and the fact that $l_{w\bar{w}} \leq 2\tau_0$. The last inequality together with (9) proves that the total growth is at most $c(T)$. $\square$

### 4.4 Cost Recovery

**Lemma 3** *Let $(F, Q)$ be the solution and $\xi$ be the cost shares computed by GKLRS on terminal pair set $R$, respectively. Then*

$$c(F) + \pi(Q) \leq 3 \sum_{(u,\bar{u})\in R} \xi_{u\bar{u}}.$$

*Proof* Following the proof of Agrawal, Klein and Ravi [1], the cost of the constructed forest $F$ satisfies

$$c(F) \leq 2 \sum_{(u,\bar{u})\in R} \xi_{u\bar{u}}.$$

Moreover, by Fact 5

$$\pi(Q) = \sum_{(u,\bar{u})\in Q} \xi_{u\bar{u}}$$

and hence $c(F) + \pi(Q) \leq 3 \sum_{(u,\bar{u})\in R} \xi_{u\bar{u}}$. $\square$

## 5 Efficiency of *GKLRS*

In [7], Chawla et al. showed that the cost shares computed by *KLS* are also $O(\log^2 k)$-approximate. (A simple proof that they are $O(\log^3 k)$-approximate is given in Section 6.) In this paper, we extend this result to the prize-collecting Steiner forest (PCSF) game. We show that the approximability of *GKLRS* can be reduced to the one of *KLS*.

**Theorem 7** *If the mechanism $M(\xi^{KLS})$ is $\alpha$-approximate then the mechanism $M(\xi^{GKLRS})$ is $3(1+\alpha)$-approximate.*

We will prove this theorem in the rest of this section. The following fact will be useful, and is easily proved.

**Fact 8** *Given a cross-monotonic cost-sharing method $\xi$, the final set of players output by the Moulin mechanism $M(\xi)$ is independent of the order of eviction.*

The following lemma will allow us to partition the universe of players into two groups and to argue about each of them separately.

**Lemma 4** *Consider a universe $U$ of players, along with a non-decreasing cost function $C$ and a $\beta$-budget balanced and cross-monotonic cost-sharing method $\xi$. Given a partition of $U$ into two parts $U_1$ and $U_2$, if the Moulin mechanism on sub-universe $U_i$ is $\alpha_i$-approximate for all $i \in \{1, 2\}$ with respect to the induced cost-sharing method $\xi|_{U_i}$ and the cost function $C|_{U_i}$, then the Moulin mechanism is $(\alpha_1 + \alpha_2)\beta$-approximate for the entire set $U$ with respect to $\xi$ and $C$.*

*Proof* Let $A$ be the final set of players returned by the Moulin mechanism when run on $U$. Define $A_i = A \cap U_i$. Since $\xi$ is $\beta$-budget balanced we have $C(A)/\beta \le \sum_{i \in A} \xi_i(A) \le C(A)$ and hence,

$$\Pi(A) = C(A) + u(U \setminus A) \le \sum_{i \in A} \beta\, \xi_i(A) + \sum_{i \in U \setminus A} u_i \tag{10}$$

Consider a run of the Moulin mechanism on $U_j$ and let $B_j$ be the final set of players for $j \in \{1, 2\}$.

*Claim* $B_j \subseteq A_j$ for all $j \in \{1, 2\}$.

*Proof* Let $e_1, \ldots, e_p$ be the elements of $U_j$ in the order in which they are dropped by the Moulin mechanism when run on $U$. Assume for the sake of contradiction that $B_j \not\subseteq A_j$. In this case there must exist $1 \le i \le p$ s.t. $e_i \in B_j$. Choose $i$ smallest with this property and let $S$ be the set of players that are still part of the game in the Moulin run on $U$ just before $e_i$ is dropped. We have

$$u_i < \xi_i(S) \le \xi_i(\{e_i, \ldots, e_p\} \cup A_j) \le \xi_i(B_j)$$

where the last two inequalities use the cross-monotonicity of $\xi$. This contradicts the fact that $e_i$ is part of the final set of the Moulin run on $U_j$. □

Notice that the cost-share $\xi_i(A)$ of players $i \in A_j \setminus B_j$ is at most the utility $u_i$ of player $i$ by the termination condition of the Moulin mechanism. For a set $S_j \subseteq U_j$, define $\Pi_j(S_j) = C(S_j) + u(U_j \setminus S_j)$. As $B_j$ is an $\alpha_j$-approximate set of players, we then have

$$\Pi_j(B_j) \le \alpha_j \Pi_j(S_j)$$

for any set $S_j \subseteq U_j$.

We can now upper bound $\Pi(A)$:

$$\begin{aligned} \Pi(A) = C(A) + u(U \setminus A) &\le \sum_{i \in A} \beta\, \xi_i(A) + \sum_{i \in U \setminus A} u_i \\ &= \left( \sum_{i \in A_1} \beta\, \xi_i(A) + \sum_{i \in U_1 \setminus A_1} u_i \right) + \left( \sum_{i \in A_2} \beta\, \xi_i(A) + \sum_{i \in U_2 \setminus A_2} u_i \right) \end{aligned} \tag{11}$$

We upper-bound the first of the two parentheses on the right-hand side of the above inequality. An upper bound for the second parentheses is obtained analogously.

$$
\begin{aligned}
\sum_{i \in A_1} \beta\, \xi_i(A) + \sum_{i \in U_1 \setminus A_1} u_i &= \sum_{i \in B_1} \beta\, \xi_i(A) + \sum_{i \in A_1 \setminus B_1} \beta\, \xi_i(A) + \sum_{i \in U_1 \setminus A_1} u_i && (12)\\
&\le \sum_{i \in B_1} \beta\, \xi_i(A) + \sum_{i \in A_1 \setminus B_1} \beta\, u_i + \sum_{i \in U_1 \setminus A_1} u_i && (13)\\
&\le \sum_{i \in B_1} \beta\, \xi_i(B_1) + \sum_{i \in U_1 \setminus B_1} \beta\, u_i && (14)\\
&\le \beta\, C(B_1) + \sum_{i \in U_1 \setminus B_1} \beta\, u_i && (15)\\
&= \beta\, \Pi_1(B_1). && (16)
\end{aligned}
$$

Inequality (13) uses the fact that player $i \in A_1 \setminus B_1$ is part of the final set of players returned by the Moulin mechanism when run on $U_1$, and hence must have utility at least its cost-share. We then use cross-monotonicity of $\xi$ and the fact that $\beta \ge 1$ to get (14). Inequality (15) uses the competitiveness of $\xi$, and the final inequality follows from the definition of $\Pi_1$. Using the resulting inequality together with (11) yields

$$\Pi(A) \le \beta\,(\Pi_1(B_1) + \Pi_2(B_2)) \le \beta(\alpha_1 \Pi_1(S_1) + \alpha_2 \Pi_2(S_2)) \qquad (17)$$

for any $S_1 \subseteq U_1, S_2 \subseteq U_2$ where we use the fact that the Moulin mechanism when run on $U_j$ is $\alpha_j$-approximate for $j \in \{1, 2\}$.

Finally, for any set $S \subseteq U$ and for $i = 1, 2$, define $S_i = S \cap U_i$. Note that since $C$ is non-decreasing, $\Pi_j(S_j) = C(S_j) + u(U_j \setminus S_j) \le C(S) + u(U \setminus S) = \Pi(S)$. Putting these together with (17), we get that $\Pi(A) \le (\alpha_1 + \alpha_2)\beta\Pi(S)$ for any $S \subseteq U$, and hence the Moulin mechanism is $(\alpha_1 + \alpha_2)\beta$-approximate. □

Armed with the above lemma, let us consider the universe of players $U$ for the *GKLRS* instance, and divide them into two parts thus:

- The "high-utility" set $U_1$ are those players $i \in U$ with utility $u_i \ge \pi_i$.
- The "low-utility" set $U_2$ are the remaining players $i \in U$ with $u_i < \pi_i$.

We now show that $\xi^{GKLRS}$ on the sub-universes $U_1$ and $U_2$ is 1-approximate and $\alpha$-approximate, respectively. This together with Lemma 4 and the fact that *GKLRS* is 3-budget balanced (Lemma 3) proves that *GKLRS* is $3(1 + \alpha)$-approximate.

We first prove the following High-Utility-Lemma:

**Lemma 5** *The mechanism $M(\xi^{GKLRS})$ is 1-approximate when restricted to the players in the high-utility set $U_1$.*

*Proof* By Fact 3, $\xi_i^{GKLRS}(S) \le \pi_i$ for every set $S \subseteq U$ and every $i \in S$. Since $u_i \ge \pi_i \ge \xi_i^{GKLRS}(S)$ for any $S \subseteq U_1$ and $i \in S$, the players in $U_1$ will never be rejected by the mechanism $M(\xi^{GKLRS})$ when run on $U_1$. Moreover, the set achieving the optimal social cost is also $U_1$, and hence the Moulin mechanism gives the social optimum on the high-utility set. □

We show that for low-utility players $S \subseteq U_2$ the two runs of *GKLRS*$(S)$ and *KLS*$(S)$ are identical up to a certain point of time.

**Lemma 6** *Let $S \subseteq U_2$. Define $\tau_0$ as the first point of time $\tau$ at which $\xi_i^{\tau,GKLRS}(S) = \pi_i$ for some player $i \in S$; let $\tau_0 = \infty$ if no such time exists. Then for all $\tau \in [0, \tau_0)$ and every player $j \in S$ it holds that $j$ is active at time $\tau$ in* GKLRS$(S)$ *iff $j$ is active at time $\tau$ in* KLS$(S)$*; in particular, this implies*

$$\xi_j^{\tau,GKLRS}(S) = \xi_j^{\tau,KLS}(S) \quad \forall \tau \in [0, \tau_0),\ \forall j \in S.$$

*Proof* A necessary condition for $j$ being active at time $\tau$ in *GKLRS*$(S)$ is that $\tau \leq \mathtt{d}(s_j, t_j)$. Thus, $j$ is active at time $\tau$ in *KLS*$(S)$ if $j$ is active at this time in *GKLRS*$(S)$. Next, suppose $j$ is active at time $\tau$ in *KLS*$(S)$ and thus $\tau \leq \mathtt{d}(s_j, t_j)$. Since $\tau < \tau_0$, we have $\xi_i^{\tau,GKLRS}(S) < \pi_i$ for all $i \in S$; in particular this also holds for player $j$. Thus, $j$ is active at time $\tau$ in *GKLRS*$(S)$. □

Suppose we compare the runs of the Moulin mechanism corresponding to the two different cost-sharing mechanisms $\xi^{GKLRS}$ and $\xi^{KLS}$ with the same set of low-utility players $S \subseteq U_2$. An immediate consequence of Lemma 6 is that as long as some player is eliminated in either of the runs of the Moulin mechanisms, there must be a player that the mechanisms could eliminate in *both the runs*.

**Corollary 1** *Fix some $S \subseteq U_2$. Suppose there is a player $j \in S$ with $\xi_j^{GKLRS}(S) > u_j$ or $\xi_j^{KLS}(S) > u_j$. Then there is a player $i$ such that $\xi_i^{GKLRS}(S) > u_i$ and $\xi_i^{KLS}(S) > u_i$.*

*Proof* Let $\tau_0$ be as defined in Lemma 6. The claim clearly holds if $\tau_0 = \infty$ as all cost shares in *GKLRS*$(S)$ and *KLS*$(S)$ are the same. Otherwise, there exists some player $i \in S$ and some $\tau_0 = \tau$ such that $\xi_i^{\tau,GKLRS}(S) = \pi_i$. Lemma 6 then implies that $\xi_i^{\tau,GKLRS}(S) = \xi_i^{\tau,KLS}(S) = \pi_i > u_i$. □

The next lemma essentially shows that the prizes $\pi_i$ play no role for the low-utility players $U_2$.

**Lemma 7** *When starting with a set of low-utility players $U_2$, the final output $S^{M,GKLRS} \subseteq U_2$ of the Moulin mechanism $M(\xi^{GKLRS})$ is identical to the output $S^{M,KLS} \subseteq U_2$ of the Moulin mechanism $M(\xi^{KLS})$.*

*Proof* Corollary 1 states that we can always identify a player $i \in S$ that we may evict in both runs of $M(\xi^{GKLRS})$ and $M(\xi^{KLS})$ as long as some player is eliminated in either of the runs of the Moulin mechanism. We can then eliminate player $i$ in both the runs and use induction to show that both runs end with the same players if we make the right choices. However, Fact 8 implies that *any* choices would lead to the same outputs, as we claim. □

We can now prove the following Low-Utility Lemma:

**Lemma 8** *Restricting our attention to the low-utility set $U_2$, the mechanism $M(\xi^{GKLRS})$ is $\alpha$-approximate if the mechanism $M(\xi^{KLS})$ is $\alpha$-approximate.*

*Proof* On the low-utility players, the solution with the optimal social cost for PCSF would never service a player $i$ by paying her penalty $\pi_i$, since it would be better to reject the player and pay $u_i < \pi_i$. This implies that the optimal social cost $\Pi^*_{PCSF}$ for PCSF and and the optimal social cost $\Pi^*_{SF}$ for SF are the same on $U_2$. Also note that for every player set $S$ the cost $\textsc{opt}_{PCSF}(S)$ of an optimal PCSF solution for $S$ is at most the cost $\textsc{opt}_{SF}(S)$ of an optimal SF solution. Let $\Pi_{PCSF}$ and $\Pi_{SF}$ denote the social cost with respect to PCSF and SF, respectively. Given these facts together with the fact that $M(\xi^{GKLRS})$ and $M(\xi^{KLS})$ output the same set $S^M$ on the low-utility instances, we conclude that

$$\begin{aligned}\Pi_{PCSF}(S^M) &= u(U_2 \setminus S^M) + \textsc{opt}_{PCSF}(S^M)\\ &\le u(U_2 \setminus S^M) + \textsc{opt}_{SF}(S^M)\\ &= \Pi_{SF}(S^M) \le \alpha \cdot \Pi^*_{SF} = \alpha \cdot \Pi^*_{PCSF}.\end{aligned}$$

□

## 6 Efficiency of *KLS*

We prove the following theorem:

**Theorem 9** *The cost shares $\xi^{KLS}$ computed by* KLS *are $O(\log^3 k)$-summable.*

We drop the superscript KLS in the discussion below. Suppose we are given an arbitrary subset $S \subseteq U$ and an ordering $\sigma$. Recall that we assume that the terminal pairs in $S$ are ordered according to $\sigma$. Without loss of generality, we assume that the terminal pairs are labeled such that

$$S = \{(s_1, t_1), \ldots, (s_l, t_l)\}, \quad \text{where } l = |S|.$$

As before, let $S_i \subseteq S$ be the set of the first $i$ terminal pairs of $S$. We use $\xi_i(S_i)$ to refer to the cost share of terminal pair $(s_i, t_i)$, $1 \le i \le l$, computed by KLS when run on terminal pair set $S_i$. We need to prove that

$$\sum_{i=1}^{l} \xi_i(S_i) = O(\log^3 k \cdot \textsc{opt}(S)), \tag{18}$$

where $\textsc{opt}(S)$ is the minimum Steiner forest cost for terminal set $S$.

We assume that the distance between every two vertices in $G$ is at least 1, i.e., $d_G(u, v) \ge 1$ for all $u, v \in U$. This assumption is without loss of generality as we may scale the edge costs appropriately. Recall that in KLS each terminal pair $(s_i, t_i) \in U$ has a *death time* $\mathtt{d}(s_i, t_i)$ which is defined as half the distance between $s_i$ and $t_i$ in $G$. We partition terminal pairs in $S$ into classes, depending on their death times: A terminal pair $(s_i, t_i) \in S$ is of *class* $r \ge 0$ if $\mathtt{d}(s_i, t_i) \in$

$(2^{r-1}, 2^r]$. Let $r(i)$ be the class to which terminal pair $(s_i, t_i)$ belongs. We use $S^r$ to refer to the (ordered) set of terminal pairs in $S$ that belong to class $r$. Moreover, we define $S_i^r \subseteq S_i$ to be the set of class $r$ terminal pairs in $S_i$, i.e., $S_i^r = S_i \cap S^r$ for every $1 \leq i \leq l$. Let $\Delta_S$ be the maximum death time among all terminal pairs in $S$. Clearly, there are at most $\log(\Delta_S) + 1$ classes.

Since $\xi$ is cross-monotonic, we have for every $(s_i, t_i)$, $1 \leq i \leq l$,

$$\xi_i(S_i) \leq \xi_i(S_i^{r(i)}).$$

Thus,

$$\sum_{i=1}^{l} \xi_i(S_i) \leq \sum_{i=1}^{l} \xi_i(S_i^{r(i)}) = \sum_{r=0}^{\log(\Delta_S)+1} \sum_{(s_i,t_i)\in S^r} \xi_i(S_i^r). \tag{19}$$

We first consider all terminal pairs of classes $0, 1, \ldots, \log(\Delta_S/k) + 1$. Note that every such terminal pair has death time at most $2\Delta_S/k$. The cost share of a terminal pair is at most twice its death time and thus

$$\sum_{r=0}^{\log(\Delta_S/k)+1} \sum_{(s_i,t_i)\in S^r} \xi_i(S_i^r) \leq k \cdot \frac{4\Delta_S}{k} \leq 4\Delta_S \leq 2\textsc{opt}(S). \tag{20}$$

That is, all terminal pairs of class at most $\log(\Delta_S/k) + 1$ contribute at most $2\textsc{opt}(S)$ to the left-hand side of (18). We can therefore concentrate on terminal pairs in classes $\log(\Delta_S/k) + 2, \ldots, \log(\Delta_S) + 1$. Note that these are at most $\log k$ different classes. For each class $r > \log(\Delta_S/k) + 1$, we prove

$$\sum_{(s_i,t_i)\in S^r} \xi_i(S_i^r) = O(\log^2(|S^r|) \cdot \textsc{opt}(S)) = O(\log^2 k \cdot \textsc{opt}(S)).$$

This together with (19) and (20) proves (18). The next lemma states that for each class, we can assume that all death times are rounded up to the nearest power of 2.

**Lemma 9 (Rounding Lemma)** *Fix some $r$ and suppose we set all death times of terminal pairs in $S^r$ to $2^r$. Let $\tilde{\xi}$ be the cost shares computed by* KLS *with these modified death times. Then*

$$\sum_{(s_i,t_i)\in S^r} \xi_i(S_i^r) \leq 3 \sum_{(s_i,t_i)\in S^r} \tilde{\xi}_i(S_i^r)$$

The proof of Lemma 9 is deferred to the end of this section.

*Summability of KLS with identical death times.* We next show that the cost shares of *KLS* are $O(\log^2 k)$-summable if all death times are equal. Eventually, we apply the result presented in this section together with the Rounding Lemma to each class $r > \log(\Delta_S/k)+1$ separately. For notational convenience, we use $S$ instead of $S^r$ here.

Suppose that the death time of all terminal pairs in $S$ is $D$, i.e., $\mathtt{d}(s_i, t_i) = D$ for all $1 \leq i \leq l$; as before, we define $l = |S|$. Let $F^*$ be a minimum cost Steiner forest for terminal pair set $S$. For a tree $T \in F^*$, let $S(T)$ be the set of terminal pairs in $S$ that are spanned by $T$. Consider a terminal pair $(s_i, t_i)$, $1 \leq i \leq l$, of $S$ and let $T \in F^*$ be the tree that contains $s_i, t_i$, i.e., $(s_i, t_i) \in S(T)$. Define $S_i(T)$ as the set of terminal pairs in $S$ that precede $(s_i, t_i)$ (with respect to $\sigma$) and are also part of $T$; more precisely $S_i(T) = S_i \cap S(T)$. Run *KLS* on $S_i(T)$ and let $\xi_i(S_i(T))$ be the respective cost share of $(s_i, t_i)$. As $S_i(T) \subseteq S_i$ and the cost shares computed by *KLS* are cross-monotonic, we have

$$\xi_i(S_i(T)) \geq \xi_i(S_i). \tag{21}$$

We prove that for each tree $T \in F^*$, we have

$$\sum_{(s_i,t_i)\in S(T)} \xi_i(S_i(T)) = O(\log^2(|S(T)|) \cdot c(T)). \tag{22}$$

Summing over all trees $T \in F^*$ together with (21) then shows that

$$\sum_{i=1}^{l} \xi_i(S_i) = O(\log^2 k \cdot \mathit{OPT}(S)).$$

Given tree $T$, we construct a rooted tree $T' = (V', E')$, also called *Shapley tree* in the following, and a non-negative length function $\ell : E' \to \mathbb{R}^+$ on the edges of $T'$. We use $T'(e)$ to refer to the subtree of $T'$ below edge $e \in E'$. Moreover, for a vertex $u \in V'$ let $P_{ur}$ be the unique $u,r$-path from $u$ to the root $r$ of $T'$. We construct $T'$ such that the following conditions hold:

1. The leaves of $T'$ are the terminals in $S(T)$.
2. For every two terminals that are contained in the subtree $T'(e)$ for some $e \in E'$, their distance in $G$ is at most $\ell(e)$, i.e., $d_G(u, v) \leq \ell(e)$ for all $u, v \in S(T) \cap T'(e)$.
3. For every path $P_{ur} = (e_1, \ldots, e_m)$ from terminal $u \in S(T)$ to the root $r$ of $T'$, we have
   (a) $\ell(e_1) = 1$,
   (b) $\ell(e_j) = 2\ell(e_{j-1})$ for all $1 < j \leq m$, and
   (c) $\ell(e_m) \geq D$.
4. The total length of $T'$ is at most $O(\log |S(T)|)$ times the total cost of $T$, i.e., $\ell(T') = O(\log(|S(T)|) \cdot c(T))$.

We use tree $T'$ to define *Shapley cost shares* for all terminals in $S(T)$: Let $T'[S_i(T)]$ be the induced subtree of $T'$ on terminals pair set $S_i(T)$. For a terminal pair $(s_i, t_i) \in S(T)$, we define $\xi'_i(S_i(T))$ to be the sum of the respective Shapley cost shares of terminals $s_i$ and $t_i$ in $T'[S_i(T)]$.

**Lemma 10** *Let $T'$ be the Shapley tree of $T$ and let $\xi'$ be the respective Shapley cost shares. Then*

$$\sum_{(s_i,t_i)\in S(T)} \xi'_i(S_i(T)) \le H_{|S(T)|} \cdot \ell(T').$$

*Proof* As $T'[S_1(T)] \subseteq T'[S_2(T)] \subseteq \cdots \subseteq T'[S_l(T)]$, the cost share contribution of an edge $e \in E'$ to the left-hand side of the inequality is at most $H_{|S(T)|} \cdot \ell(e)$. Summing over all edges $e \in E'$ of tree $T'$ proves the lemma. □

We next show that the cost share $\xi_i(S_i(T))$ of terminal pair $(s_i, t_i)$ is upper bounded by its corresponding Shapley cost share $\xi'_i(S_i(T))$ in $T'[S_i(T)]$. This together with Lemma 10 and Property 4 establishes (22).

**Lemma 11** *The cost share $\xi_i(S_i(T))$ of terminal pair $(s_i, t_i) \in S(T)$ is at most its Shapley cost share $\xi'_i(S_i(T))$.*

*Proof* All terminals in $S(T)$ are active until time $D$. The cost share $\xi_u(S_i(T))$ of a terminal $u \in \{s_i, t_i\}$ in KLS is then defined as

$$\xi_u(S_i(T)) = \int_{\tau=0}^{D} \frac{d\tau}{a_i^\tau(u)}$$

where $a_i^\tau(u)$ is the number of active terminals in $u$'s moat at time $\tau$ in the run of KLS$(S_i(T))$. We bound the cost share that $u = s_i$ receives in KLS$(S_i(T))$ by its Shapley cost share. An analogous argument holds for $u = t_i$.

Consider the induced subtree $T'_i = T'[S_i(T)]$ on $S_i(T)$. Let $P_{ur} = (e_1, \ldots, e_m)$ be the unique $u,r$-path in $T'_i$. Consider an edge $e_j$, $1 < j \le m$ and let $T'_i(e_j)$ be the subtree of $T'_i$ below edge $e_j$. We use $m_i(e_j)$ to refer to the number of terminals in $T'_i(e_j)$; define $m_i(e_1) = 1$. The Shapley cost share that $u$ received for edge $e_j$ is $\ell(e_j)/m_i(e_j)$. Thus,

$$\xi'_u(S_i(T)) = \sum_{j=1}^{m} \frac{\ell(e_j)}{m_i(e_j)}.$$

Let $x$ be any terminal in $T'_i(e_j)$. By Property 2, we have $d_G(u,x) \le \ell(e_j)$. Since both $x$ and $u$ are active until time $D$, their respective moats in KLS$(S_i(T))$ must have met by time at most $d_G(u,x)/2 \le \ell(e_j)/2 = \ell(e_{j-1})$. Thus, $a_i^\tau(u) \ge m_i(e_j)$ for all $\tau \ge \ell(e_{j-1})$ for all $1 < j \le m$.

Note that the cost share that $u$ receives up to time 1 is at most 1. As $\ell(e_1) = 1$ and $\ell(e_m) \ge D$, we can write

$$\begin{aligned}
\xi_u(S_i(T)) = \int_{\tau=0}^{D} \frac{d\tau}{a_i^\tau(u)} &\le 1 + \sum_{j=2}^{m} \int_{\tau=\ell(e_{j-1})}^{\ell(e_j)} \frac{d\tau}{a_i^\tau(u)} \\
&\le 1 + \sum_{j=2}^{m} \int_{\tau=\ell(e_{j-1})}^{\ell(e_j)} \frac{d\tau}{m_i(e_j)} \\
&= 1 + \sum_{j=2}^{m} \frac{\ell(e_{j-1})}{m_i(e_j)} \le \xi'_u(S_i(T)).
\end{aligned}$$

□

*Tree construction:* There are several ways to obtain a tree $T'$ that satisfies Properties 1–4. For example, the HSTs construction given by Fakcharoenphol et al. [8] satisfies all Properties 1–3 and Property 4 on expectation.

Alternatively, using ideas similar to the one presented in [29], we may insert terminals one-by-one and obtain a tree $T'$ whose vertices are terminals in $S(T)$ and that satisfies Properties 2, 3(b), 3(c) and 4. In order to achieve Property 1 and 3(a), we simply replace each non-leaf terminal $u$ with parent edge $e$ in $T'$ by a path $(e_1, \ldots, e_m)$ with $\ell(e_m) = \ell(e)/2$ and $\ell(e_1) = 1$. Clearly, this construction will add an additional cost of at most $\ell(T')$.

*Rounding Lemma.* Consider the set $S = S^r$ of class $r$ terminals and let $l = |S|$. As before we assume that $S$ is ordered according to $\sigma$ and $S_i$ refers to the set of the first $i$ terminal pairs of $S$. For a terminal $u \in \{s_i, t_i\}$, we also use $S_u$ to refer to the corresponding set of terminal pairs $S_i$. Define $\mu = 2^{r-1}$, i.e., $\mathtt{d}(s_i, t_i) \in (\mu, 2\mu]$ for all $1 \leq i \leq l$.

Recall that in *KLS* a terminal $u \in \{s_i, t_i\}$ is called *active* at time $\tau$ if $\tau \leq \mathtt{d}(s_i, t_i)$; it is said to be *inactive* otherwise. A terminal receives cost share only if it is active. For a terminal $u$ that is active at time $\tau$ in *KLS*$(S)$, define $a_u^\tau(S)$ as the number of active terminals in $u$'s moat. The cost share that an active terminal $u$ receives at time $\tau$ is defined as $\xi_u^\tau(S) = 1/a_u^\tau(S)$. The cost share $\xi_{s_i t_i}^\tau(S)$ of terminal pair $(s_i, t_i)$ at time $\tau$ is defined as $\xi_{s_i}^\tau(S) + \xi_{t_i}^\tau(S)$.

Fix a point of time $\tau \in (\mu, 2\mu]$. Without loss of generality, let $\xi_{s_i}^\tau(S_i) \geq \xi_{t_i}^\tau(S_i)$ for every terminal pair $(s_i, t_i)$, $1 \leq i \leq l$. We say $s_i$ is the *dominating* terminal of $(s_i, t_i)$. Note that $\xi_{s_i t_i}^\tau(S_i) \leq 2\xi_{s_i}^\tau(S_i)$. Let $D^\tau$ be the set of all dominating terminals that are active at time $\tau$. The following technical lemma is the key to proving Lemma 9. It shows that for every terminal $s_i \in D^\tau$ the cost share $\xi_{s_i}^\tau(S_i)$ that $s_i$ receives at time $\tau$ can be charged to the cost share that some terminal $f^\tau(s_i)$ in $S_i$ received at time $\tau - \mu$. Moreover, the mapping $f^\tau$ is injective. This will enable us to charge the total cost share collected by terminals in $D^\tau$ at time $\tau$ in *KLS*$(S_i)$ to the total cost share of terminals in $S_i$ at time $\tau - \mu$.

**Lemma 12** *Let $D^\tau$ be the set of all dominating terminals that are active at time $\tau \in (\mu, 2\mu]$. There exists a mapping $f^\tau : D^\tau \to S$ such that the following conditions hold:*

1. *For each $s_i \in D^\tau$ we have $\xi_{s_i}^\tau(S_i) \leq \xi_{f^\tau(s_i)}^{\tau-\mu}(S_{f^\tau(s_i)})$.*
2. *For all $s_i, s_j \in D^\tau$, $i \neq j$, we have $f^\tau(s_i) \neq f^\tau(s_j)$.*

*Proof* We use $M_u^\tau(S)$ to refer to the moat of $u$ at time $\tau$ in the run of *KLS* on terminal set $S \subseteq R$. Subsequently, we exploit the following two properties of *KLS* (see [19,21] for the proofs of these facts).

**Fact 10** *Let $S \subseteq R$ and consider a terminal $u \in S$. For every $\tau' \leq \tau$ we have $M_u^{\tau'}(S) \subseteq M_u^\tau(S)$.*

**Fact 11** *Let $S' \subseteq S \subseteq R$ and consider a terminal $u \in S'$. For every $\tau$ we have $M_u^\tau(S') \subseteq M_u^\tau(S)$.*

We assume that the set of dominating terminals $D^\tau$ is ordered according to $\sigma$. We define $f^\tau$ inductively. Suppose $f^\tau$ satisfies Conditions 1 and 2 of the lemma for the first $n-1$ terminals in $D^\tau$. (Let $f^\tau$ be the empty mapping for $n=0$.) We define $f^\tau(s_i)$ of the $n$-th terminal $s_i$ of $D^\tau$ while maintaining Conditions 1 and 2.

Assume $\xi^\tau_{s_i}(S_i) = 1/x$. Let $C_{s_i} = M^\tau_{s_i}(S_i)$ be the set of all terminals that are contained in $s_i$'s moat at time $\tau$. Note that $|C_{s_i}| \geq x$. Order the set $C_{s_i}$ according to $\sigma$ and delete all terminals except the first $x$ ones. We call the resulting terminal set $C_{s_i}$ the *candidate set* of $s_i$. Note that $C_{s_i} \subseteq M^\tau_{s_i}(S_i)$. We will eventually define $f(s_i) = \hat{u}$ for some $\hat{u} \in C_{s_i}$.

Consider the $m$-th terminal $u$ of $C_{s_i}$, $1 \leq m \leq x$. Note that all terminals in $M^{\tau-\mu}_u(S_u)$ are active at time $\tau - \mu$ because all terminal death times are larger than $\mu$. By Facts 10 and 11 we have $M^{\tau-\mu}_u(S_u) \subseteq M^\tau_u(S_u) \subseteq M^\tau_u(S_i) = M^\tau_{s_i}(S_i)$. Therefore, the moat $M^{\tau-\mu}_u(S_u)$ contains at most $m$ terminals. Since $m \leq x$, we have $\xi^{\tau-\mu}_u(S_u) \geq 1/x$ for all $u \in C_{s_i}$.

Next we show that there always exists a choice of a terminal $\hat{u} \in C_{s_i}$ such that $f^\tau(s_j) \neq \hat{u}$ for all $s_j \in D^\tau$, $j < i$. The proof is by contradiction. Suppose that for each terminal $u \in C_{s_i}$ there exists a terminal $s_j \in D^\tau$, $j < i$, with $f^\tau(s_j) = u$. Note that by our induction hypothesis, $f^\tau(s_j) \neq f^\tau(s_k)$ for all $j \neq k$ and $j, k < i$. Consider some $u \in C_{s_i}$ and let $s_j \in D^\tau$, $j < i$, with $f^\tau(s_j) = u$. By our construction of the candidate set, we have $u = f^\tau(s_j) \in C_{s_j} \subseteq M^\tau_{s_j}(S_j)$. Moreover, $M^\tau_{s_j}(S_j) \subseteq M^\tau_{s_j}(S_i)$ by Fact 11. This implies that both $M^\tau_{s_j}(S_i)$ and $M^\tau_{s_i}(S_i)$ contain $u$ and therefore must be identical. As a consequence, $s_j$ is an active terminal of $M^\tau_{s_i}(S_i)$ (recall that $s_j$ is active because $s_j \in D^\tau$). Because this holds for every $u \in C_{s_i}$, this leads to a contradiction to the assumption that $\xi^\tau_{s_i}(S_i) = 1/x$ since we have identified $|C_{s_i}| \geq x$ active terminals in $M^\tau_{s_i}(S_i)$ that are different from $s_i$. □

We can then proof Lemma 9:

*Proof (Proof of Lemma 9)* First observe that the executions of KLS with and without rounded death times are identical until time $\mu$. Thus

$$\sum_{i=1}^{l} \int_{\tau=0}^{\mu} \xi^\tau_{s_i t_i}(S_i) d\tau = \sum_{i=1}^{l} \int_{\tau=0}^{\mu} \tilde{\xi}^\tau_{s_i t_i}(S_i) d\tau \leq \sum_{i=1}^{l} \tilde{\xi}_{s_i t_i}(S_i). \tag{23}$$

For time $\tau \in (\mu, 2\mu]$ let $f^\tau$ be a mapping as constructed in Lemma 12. Since $s_i$ is the dominating terminal of $(s_i, t_i)$, we have

$$\sum_{i=1}^{l} \xi^\tau_{s_i t_i}(S_i) d\tau \leq 2 \sum_{i=1}^{l} \xi^\tau_{s_i}(S_i) d\tau \leq 2 \sum_{i=1}^{l} \xi^{\tau-\mu}_{f^\tau(s_i)}(S_{f^\tau(s_i)}) d\tau \leq 2 \sum_{i=1}^{l} \xi^{\tau-\mu}_{s_i t_i}(S_i) d\tau,$$

where we used Condition 1 and 2 of Lemma 12 for the second and last inequality, respectively. Integrating over all time instants in $(\mu, 2\mu]$, we obtain

$$\sum_{i=1}^{l} \int_{\tau=\mu}^{2\mu} \xi^\tau_{s_i t_i}(S_i) d\tau \leq 2 \sum_{i=1}^{l} \int_{\tau=0}^{\mu} \xi^\tau_{s_i t_i}(S_i) d\tau \overset{(23)}{\leq} 2 \sum_{i=1}^{l} \tilde{\xi}_{s_i t_i}(S_i).$$

□

## 7 Conclusions

In this paper we presented a first cross-monotonic cost sharing mechanism for the price-collecting Steiner forest problem that is 3-budget balanced and $O(\log^2 k)$-approximate with respect to social cost. This result is obtained by developing a much simpler primal-dual approach than the one suggested in [15] while achieving the same approximation ratio and cross-monotonicity. We also present a general method that turns a mechanism that approximates social cost for an optimization problem into a mechanism that approximates social cost for the corresponding prize-collecting version. We hope that both techniques will help in achieving similar results for a wider class of network design problems.

## References

1. Agrawal, A., Klein, P., Ravi, R.: When trees collide: an approximation algorithm for the generalized Steiner problem on networks. SIAM J. Comput. **24**(3), 440–456 (1995). (Preliminary version in *23rd STOC*, 1991)
2. Archer, A., Feigenbaum, J., Krishnamurthy, A., Sami, R., Shenker, S.: Approximation and collusion in multicast cost sharing. Games and Economic Behavior **47**(1), 36–71 (2004)
3. Arora, S., Lund, C., Motwani, R., Sudan, M., Szegedy, M.: Proof verification and the hardness of approximation problems. J. ACM **45**(3), 501–555 (1998)
4. Bartal, Y.: Probabilistic approximations of metric spaces and its algorithmic applications. In: Proceedings of the 37th Symposium on the Foundations of Computer Science (FOCS), pp. 184–193 (1996)
5. Bern, M., Plassman, P.: The Steiner problem with edge lengths 1 and 2. Information Processing Letters **32**, 171–176 (1989)
6. Bienstock, D., Goemans, M.X., Simchi-Levi, D., Williamson, D.P.: A note on the prize collecting traveling salesman problem. Math. Program. **59**, 413–420 (1993)
7. Chawla, S., Roughgarden, T., Sundararajan, M.: Optimal cost-sharing mechanisms for steiner forest problems. In: Pooc. of the Second International Workshop on Internet and Network Economics, WINE 2006, pp. 112–123 (2006)
8. Fakcharoenphol, J., Rao, S., Talwar, K.: A tight bound on approximating arbitrary metrics by tree metrics. J. Comput. System Sci. **69**(3), 485–497 (2004)
9. Feigenbaum, J., Krishnamurthy, A., Sami, R., Shenker, S.: Hardness results for multicast cost-sharing. Theoretical Computer Science **304**, 215–236 (2003)
10. Feigenbaum, J., Papadimitriou, C.H., Shenker, S.: Sharing the cost of multicast transmissions. J. Comput. System Sci. **63**(1), 21–41 (2001). Special issue on internet algorithms
11. Goemans, M.X., Williamson, D.P.: A general approximation technique for constrained forest problems. SIAM J. Comput. **24**(2), 296–317 (1995). (Preliminary version in *5th SODA*, 1994)
12. Green, J., Kohlberg, E., Laffont, J.J.: Partial equilibrium approach to the free rider problem. Journal of Public Economics **6**, 375–394 (1976)
13. Gupta, A., Srinivasan, A., Tardos, É.: Cost-sharing mechanisms for network design. In: Proceedings of the Seventh International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (2004)
14. Gupta, A., Srinivasan, A., Tardos, É.: Cost-sharing mechanisms for network design. Algorithmica **50**(1), 98–119 (2008)
15. Hajiaghayi, M.T., Jain, K.: The prize-collecting generalized steiner tree problem via a new approach of primal-dual schema. In: Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2006), pp. 631–640 (2006)

16. Immorlica, N., Mahdian, M., Mirrokni, V.S.: Limitations of cross-monotonic cost sharing schemes. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 602–611. ACM Press (2005)
17. Jain, K., Vazirani, V.: Applications of approximation algorithms to cooperative games. In: Proceedings of the 33rd Annual ACM Symposium on the Theory of Computing (STOC), pp. 364–372 (2001)
18. Kent, K.J., Skorin-Kapov, D.: Population monotonic cost allocations on MSTs. In: Proceedings of the 6th International Conference on Operational Research (Rovinj, 1996), pp. 43–48. Croatian Oper. Res. Soc., Zagreb (1996)
19. Könemann, J., Leonardi, S., Schäfer, G.: A group-strategyproof mechanism for Steiner forests. In: Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 612–619. ACM Press (2005)
20. Könemann, J., Leonardi, S., Schäfer, G., van Zwam, S.: From primal-dual to cost shares and back: a stronger LP relaxation for the Steiner forest problem. In: Automata, languages and programming, *Lecture Notes in Comput. Sci.*, vol. 3580, pp. 930–942. Springer, Berlin (2005)
21. Könemann, J., Leonardi, S., Schäfer, G., van Zwam, S.H.M.: A group-strategyproof cost sharing mechanism for the Steiner forest game. SIAM J. Comput. **37**(5), 1319–1341 (2008)
22. Leonardi, S., Schäfer, G.: Cross-monotonic cost sharing methods for connected facility location games. Theor. Comput. Sci. **326**(1-3), 431–442 (2004)
23. Moulin, H.: Incremental cost sharing: Characterization by coalition strategy-proofness. Social Choice and Welfare **16**, 279–320 (1999)
24. Moulin, H., Shenker, S.: Strategyproof sharing of submodular costs: budget balance versus efficiency. Econom. Theory **18**(3), 511–533 (2001)
25. Pál, M., Tardos, E.: Group strategyproof mechanisms via primal-dual algorithms. In: Proceedings of the 44th Symposium on the Foundations of Computer Science (FOCS), pp. 584–593 (2003)
26. Roberts, K.: The characterization of implementable choice rules. In: J.J. Laffont (ed.) Aggregation and Revelation of Preferences. North-Holland (1979)
27. Roughgarden, T., Sundararajan, M.: New trade-offs in cost-sharing mechanisms. In: Proceedings of the 38th Annual ACM Symposium on Theory of Computing, STOC06, pp. 79–88 (2006)
28. Roughgarden, T., Sundararajan, M.: Optimal efficiency guarantees for network design mechanisms. In: Proceedings of the 12th International Conference on Integer Programming and Combinatorial Optimization, IPCO'07, pp. 469–483 (2007)
29. Roughgarden, T., Sundararajan, M.: Quantifying inefficiency in cost-sharing mechanisms. J. ACM **56**(4) (2009)