An improved approximation guarantee for Prize-Collecting TSP

Jannis Blauth*   Martin Nägele*

Abstract

We present a new approximation algorithm for the (metric) prize-collecting traveling salesperson problem (PCTSP). In PCTSP, opposed to the classical traveling salesperson problem (TSP), one may choose to not include a vertex of the input graph in the returned tour at the cost of a given vertex-dependent penalty, and the objective is to balance the length of the tour and the incurred penalties for omitted vertices by minimizing the sum of the two. We present an algorithm that achieves an approximation guarantee of 1.774 with respect to the natural linear programming relaxation of the problem. This significantly reduces the gap between the approximability of classical TSP and PCTSP, beating the previously best known approximation factor of 1.915. As a key ingredient of our improvement, we present a refined decomposition technique for solutions of the LP relaxation, and show how to leverage components of that decomposition as building blocks for our tours.

1   Introduction

The classical (metric) traveling salesperson problem (TSP) is one of the most well-known problems in Combinatorial Optimization. Given a complete undirected graph $G = (V, E)$ with metric edge lengths $c_e \in \mathbb{R}_{\geq 0}$ for all $e \in E$, the task is to find a shortest Hamiltonian cycle in G. TSP arises canonically in many applied settings, and it is thus not surprising that a vast number of variations of the classical problem are studied both from a theoretical and an applied perspective. One very natural such variation that we focus on in this paper is *prize-collecting TSP*.

> **Prize-Collecting TSP (PCTSP):** Given a complete undirected graph $G = (V, E)$ with metric edge lengths $c_e \geq 0$ for all $e \in E$, a root $r \in V$, and penalties $\pi_v \geq 0$ for all $v \in V \setminus \{r\}$, the task is to find a cycle $C = (V_C, E_C)$ in G that contains the root r and minimizes
>
> $$\sum_{e \in E_C} c_e + \sum_{v \in V \setminus V_C} \pi_v \ .$$

In other words, for each vertex other than the root, we may decide between covering the vertex by the output cycle or dropping it at cost π_v, and the goal is to find a balance between the tour cost and incurred penalties so as to minimize the sum of the two. We remark that this may lead to degenerate cycles C that may contain only two vertices (or even the root only). Many real-world applications of TSP variants (e.g., in logistics) share the property that it is allowed to not visit some of the vertices at certain extra costs, and PCTSP is one very natural way of modeling such degrees of freedom.

Clearly, PCTSP is at least as hard as TSP: Large enough penalties enforce visiting all vertices. As a consequence, PCTSP is APX-hard, and one research focus lies on approximation algorithms. For TSP, despite extensive research, the $3/2$-approximation algorithm of Christofides [Chr76] and Serdyukov [Ser87] (also see [Chr22; vS20]) persisted to be the best known for more than four decades, until a very recent breakthrough of Karlin, Klein, and Oveis Gharan [KKO21; KKO22], who improved the factor to $3/2 - \varepsilon$ for some $\varepsilon > 10^{-36}$. For PCTSP, the first constant factor approximation was given by Bienstock, Goemans, Simchi-Levi, and Williamson [BGSW93] through a threshold rounding approach based on the following linear programming relaxation, where we use $\pi_r := 0$ for convenience:[1]

$$
\begin{aligned}
\min \quad & \sum_{e \in E} c_e x_e + \sum_{v \in V} \pi_v (1 - y_v) \\
& \begin{aligned}
x(\delta(v)) &= 2 y_v && \forall v \in V \setminus \{r\} \\
x(\delta(r)) &\leq 2 \\
x(\delta(S)) &\geq 2 y_v && \forall S \subseteq V \setminus \{r\}, v \in S \\
y_r &= 1 \\
x_e &\geq 0 && \forall e \in E \\
y_v &\geq 0 && \forall v \in V \ .
\end{aligned}
\end{aligned}
\qquad \text{(PCTSP LP relaxation)}
$$

Here, it is implied that $y_v \leq 1$ for all $v \in V$, hence the variables y_v can be interpreted to indicate the extent to which v is covered by a fractional solution. Note that we are using a version of the linear relaxation that includes the constraint $x(\delta(r)) \leq 2$. This does not change the optimal value, but, if $|V| \geq 2$, it guarantees that this PCTSP LP relaxation generalizes the classical Held-Karp relaxation for TSP, which is recovered by setting $y_v = 1$ for all $v \in V$ (corresponding to the requirement that all vertices need to be fully connected to the root).

The threshold rounding algorithm of [BGSW93] considers an optimal solution of the PCTSP LP relaxation, and exploits the Christofides-Serdyukov algorithm to construct a tour on precisely those $v \in V$ with

---

[1] For $S \subseteq V$, we let $\delta(S) := \{e \in E \colon |e \cap S| = |e \setminus S| = 1\}$. Moreover, for $v \in V$, we use $\delta(v) := \delta(\{v\})$.

$y_v \geq 3/5$. This tour is shown to have cost within a factor $5/2$ of the value of the optimal LP solution, and thus also of the optimal tour. Relying on an equivalent LP, Goemans and Williamson [GW95] later gave a 2-approximation through a primal-dual algorithm. By exploiting that the latter algorithm in fact loses different factors on the two parts of the objective, Archer, Bateni, Hajiaghayi, and Karloff [ABHK11] could break the barrier of 2 and give a $97/49$-approximation algorithm. Finally, combining a randomized analysis of the threshold rounding approach with the improved analysis of the primal-dual method, Goemans [Goe09] obtained the currently best known guarantee of $1/\left(1 - \frac{2}{3}e^{-1/3}\right) \approx 1.915$. It is worth noting that the latter guarantee is still relative to the lower bound given by the PCTSP LP relaxation, and it allows using LP-relative TSP algorithms in a black-box way: Given an algorithm for TSP that outputs a β-approximate solution with respect to the standard Held-Karp relaxation, the approach of Goemans [Goe09] returns a solution of cost at most $1/(1 - \frac{1}{\beta}e^{1-2/\beta})$ times the cost of an optimum solution to the PCTSP LP relaxation.

1.1 Our results and contributions

Our main result is to improve the approximation guarantee for PCTSP, getting significantly closer to the best known approximation guarantee for TSP.

Theorem 1. *There is an LP-relative 1.774-approximation algorithm for PCTSP.*

Our algorithm is similar in spirit to the classical threshold rounding approach that was leveraged in [BGSW93] and [Goe09], but overcomes one of its critical weaknesses: In threshold rounding, the two decisions of (i) which vertices to visit, and (ii) how to build the tour are taken independently. Concretely, based on an optimal solution (x^*, y^*) of the PCTSP LP relaxation and a threshold $\gamma \in (0, 1]$, the set $V_\gamma := \{v \in V : y_v^* \geq \gamma\}$ of vertices that will be visited by the returned tour is fixed, and any LP-relative TSP algorithm can be used to blindly build a tour on V_γ. In our algorithm, to the contrary, choosing which vertices to visit and the actual construction of the tour are carefully intertwined.

Similarly to the threshold rounding algorithm, we also maintain a core set V_γ of vertices that are guaranteed to be on the output tour. Instead of building a tour from edges in $\binom{V_\gamma}{2}$, though, the building blocks of our tour are given by a set $\mathcal{W}$ of walks on V having precisely their endpoints in V_γ and that are pre-computed based on the solution (x^*, y^*) of the PCTSP LP relaxation. Through a procedure based on randomized pipage rounding [AS04; Sri01; CCPV11], we sample a subset of these walks from $\mathcal{W}$ such that the resulting multigraph H spans all vertices in V_γ, and its cost can be bounded with respect to the LP solution. Crucially, through our construction, the graph H will have odd degrees in V_γ only, hence the cost of constructing an actual tour based on H (typically called *parity correction*) is not more than in the classical threshold rounding approach. In comparison, the primal-dual approach by Goemans and Williamson [GW95] constructs a tree without taking care of the vertex parities and ends up with doubling the computed tree for parity correction. Finally, our randomized tour construction allows us to analyze probabilities for vertices in $V \setminus V_\gamma$ to appear on the tour, and thus get bounds on expected penalties incurred for such vertices. We also note that derandomizing our construction is possible using a deterministic variant of pipage rounding.

Due to the refined tour construction, our approach as sketched above does not allow to use any LP-relative approximation algorithm for TSP as a black box. We remark, though, that even if an LP-relative $4/3$-approximation for TSP (matching the current best lower bound on the integrality gap of the Held-Karp relaxation) was available, our approximation guarantee remains superior to the current best black-box algorithm by Goemans [Goe09]. Furthermore, as mentioned in Theorem 1, our analysis is still with respect to the optimal value of the PCTSP LP relaxation, thereby immediately giving the subsequent corollary.

Corollary 2. *The integrality gap of the PCTSP LP relaxation is less than 1.774.*

On the technical side, the main ingredient of our algorithm is the construction and composition of the set $\mathcal{W}$ of walks mentioned earlier. We obtain appropriate walks through a decomposition of the x-part of

a solution (x, y) of the PCTSP LP relaxation. Our decomposition of $x \in \mathbb{R}^E$ can equivalently be seen as a packing result on the multigraph obtained through scaling up x to an integral vector, and interpreting x_e as an edge multiplicity for every $e \in E$. Undoubtedly, the most prominent such packing theorem is Edmond's result on packing disjoint spanning arborescences [Edm73]: A digraph $D = (V, A)$ admits a packing of k spanning arborescences rooted at r if and only if it is rooted k-edge connected, i.e., all r-v cuts have size at least k for all $v \in V$. Due to the typically uniform connectivity requirement, such results may well be used to decompose solutions of the Held-Karp relaxation (see [GW17] for such an application). In order to decompose a solution of the PCTSP LP relaxation, where the variables y_v indicate the non-uniform connectivities to the root r, we generalize ideas behind one of few results with non-uniform connectivity requirements due to Bang-Jensen, Frank, and Jackson [BFJ95]. Concretely, we revise their splitting-off based construction to step away from decompositions into components that span the full subset of vertices with maximum connectivity (i.e., in our setting, those vertices with $y_v = 1$), and instead only anchor each of the components at exactly two vertices in this subset, which we call the *anchors* of the component. This anchoring has the advantage that we can split each component of the decomposition into a path connecting its anchors (its *backbone*) and trees incident to this path (its *limbs*). By deleting or doubling the limbs we get even degree on all vertices except the anchors, which will prove useful to bound the cost of parity correction as mentioned earlier.

Besides being key to our improved approximation factor for PCTSP, we remark that a common special case of our result and the one of Bang-Jensen, Frank, and Jackson [BFJ95] gives rise to simple algorithms with approximation guarantees matching those of the primal-dual approach of Goemans and Williamson [GW95], both for PCTSP and the related *prize-collecting Steiner tree* problem, which is the following.

Prize-Collecting Steiner Tree (PCST): Given a graph $G = (V, E)$ with edge lengths $c_e \geq 0$ for all $e \in E$, a root $r \in V$, and penalties $\pi_v \geq 0$ for all $v \in V \setminus \{r\}$, the task is to find a tree $T = (V_T, E_T)$ in G that contains the root r and minimizes

$$\sum_{e \in E_T} c_e + \sum_{v \in V \setminus V_T} \pi_v \ .$$

The classical LP formulation for PCST is almost identical to the PCTSP LP relaxation, except that the degree constraints are dropped, and the coefficients 2 on the right hand side of the cut constraints are replaced by a 1. For PCTSP and PCST, the algorithms of Goemans and Williamson based on their primal-dual framework return cycles and trees, respectively, with length at most twice the x-cost of an optimum LP solution, and total penalty no more than once the y-cost of an optimum LP solution. Matching the resulting guarantee on the sum of both, we present simple 2-approximations for either of the problems. We note that the currently best known approximation factor for PCST is 1.968 [ABHK11], hence the proposed algorithm is inferior. One particular reason why our new techniques for PCTSP do not lead to an improved approximation guarantee for PCST is that the LP formulation for PCST has a rather weak integrality gap of 2. Still, we believe that our techniques enrich the study and may prove useful again also for PCST, potentially in the context of provably stronger LP relaxations for PCST. One natural option is to adapt the bidirected cut relaxation for the Steiner tree problem to PCST. However, its a big open question whether the integrality gap of the bidirected cut relaxation is strictly less than 2 (see, e.g., [RV99; CDV08]).

1.2 Further related work

The prize collection traveling salesperson problem discussed in this paper can be considered the main variant among several problems of similar nature that emerged from work of Balas on a model for scheduling the daily operation of a steel rolling mill [Bal89]. In his original problem formulation, on top of the PCTSP

definition used here, there was a prize money $w_v \geq 0$ for every $v \in V \setminus \{r\}$ that the traveling salesperson collects upon visiting v, and a constraint that a prescribed amount w_0 of total prize money should be collected. To distinguish from the PCTSP formulation used here, we refer to Balas' formulation as the *Quota PCTSP*, also in view of the special case of *Quota TSP* that is obtained by setting $\pi_v = 0$ for all $v \in V \setminus \{r\}$. Quota TSP admits a 5-approximation [ALM00], and by concatenating solutions of the Quota TSP and the PCTSP subproblem, a constant factor approximation for Quota PCTSP is readily obtained [ABLM07]. The special case of Quota TSP where $w_v = 1$ for each $v \in V \setminus \{r\}$ and $w_0 = k$ for an integer k is typically referred to as *k-TSP* in the literature, and admits a 2-approximation [Gar05]. Another budget-driven version of PCTSP is *budgeted prize-collecting TSP*. In this problem, an upper bound on the allowed traveling distance is given, and the goal is to maximize the sum of all collected prizes. Paul, Freund, Ferber, Shmoys, and Williamson [PFFSW20] provide a 2-approximation algorithm for this problem, again assuming metric edge lengths as in all these variants. Searching for a path instead of a cycle, and again imposing an upper bound on the traveling distance, results in the *Orienteering* problem, which is as well widely studied in the literature in terms of approximation algorithms (e.g., [BCKLMM03; CKP12; DFPS22]).

PCTSP was also studied in special metric spaces. For graph metrics in planar graphs, Bateni, Chekuri, Ene, Hajiaghayi, Korula, and Marx [BCEHKM11] show the existence of a PTAS. For Euclidean distances, a PTAS is known as well [CJJ20]. For the asymmetric version of PCTSP, Nguyen [Ngu13] provides a $\lceil \log(n) \rceil$-approximation algorithm, where n denotes the number of vertices.

Prize-collecting variations are also popular with other network design problems, in particular the Steiner tree problem mentioned above. A further generalization of this problem is the *prize-collecting Steiner forest problem* (also called prize-collecting generalized Steiner tree problem), in which we are given penalties for vertex pairs that need to be paid in case the vertices are not connected in the output graph. By using a straightforward adaption of the threshold rounding approach from [Goe09], one can get a $1/(1 - e^{-1/2}) \approx 2.541$-approximation [HJ06]. This also upper bounds the integrality ratio of the natural LP relaxation for this problem. Könemann, Olver, Pashkovich, Ravi, Swamy, and Vygen [KOPRSV17] showed a lower bound of $9/4$ on the integrality gap, which exceeds the integrality gap of 2 for the corresponding LP relaxation of the *Steiner forest problem*.

1.3 Organization of the paper

In Section 2, we give a detailed overview of a randomized version of our new algorithm, including in particular our decomposition lemma, the key technical tool that allows for our improved approximation guarantee. Section 3 showcases the simple 2-approximation algorithms for PCTSP and PCST. In Section 4, we give a proof of the decomposition lemma. Section 5, besides providing details on a randomized step of our algorithm, shows how to transform our approach to a deterministic one at no qualitative loss. Finally, we conclude the paper with a deferred technical proof in Section 6.

2 Our approach

On a high level, we obtain our improved approximation guarantee for PCTSP by following a common recipe for constructing tours: (i) Construct a connected subgraph H of the input graph (typically a spanning tree), (ii) add a shortest $\mathrm{odd}(H)$-join J to H,[2] and (iii) return the cycle C obtained by shortcutting an Eulerian tour in $H \cup J$. In the final step, *shortcutting* is to skip vertices that have been visited already when traversing the tour, i.e., resulting in an actual cycle. To get bounds on the length of the thereby constructed cycle, it is enough to bound the length of H and J.

---

[2]We denote by $\mathrm{odd}(H)$ the set of all vertices of odd degree in H. Moreover, we remind the reader that for $Q \subseteq V$ with $|Q|$ even, a *Q-join* is a set of edges with odd degrees precisely at vertices in Q.

4

A classical incarnation of the above recipe is the Christofides-Serdyukov algorithm for TSP on a complete graph $G = (V, E)$ with metric edge lengths $c \colon E \to \mathbb{R}_{\geq 0}$. Wolsey [Wol80] proposed the following LP-based analysis: For any point x feasible for the *Held-Karp relaxation*, whose feasible region is given by

$$P_{\mathrm{HK}}(G) := \left\{ x \in \mathbb{R}_{\geq 0}^E \colon \begin{array}{ll} x(\delta(v)) = 2 & \forall v \in V \\ x(\delta(S)) \geq 2 & \forall S \subsetneq V,\ S \neq \emptyset \end{array} \right\},$$

one can prove that a shortest spanning tree on V has cost at most $c^\top x$, and a shortest $\mathrm{odd}(T)$-join has cost at most $\frac{1}{2} c^\top x$, therefore proving a $3/2$-approximation guarantee with respect to a shortest solution $x \in P_{\mathrm{HK}}$.

2.1 Review of the classical threshold rounding algorithm

Wolsey's analysis can be leveraged for an analysis of the threshold rounding algorithm by Bienstock, Goemans, Simchi-Levi, and Williamson [BGSW93] as follows. Starting from a solution (x, y) of the PCTSP LP relaxation, let $V_1 := \{v \in V \colon y_v = 1\}$. Through applying *splitting off* techniques to x (see Section 4 for details), we can obtain a point $z \in P_{\mathrm{HK}}(G_1)$ of total length at most $c^\top x$, where $G_1 = (V_1, E_1)$ is the complete graph on the vertex subset V_1. Hence, Wolsey's analysis shows that we can get a cycle on V_1 of length at most $\frac{3}{2} c^\top z \leq \frac{3}{2} c^\top x$. Because the loss in the penalty term incurred by this cycle may be arbitrarily large compared to the LP penalty $\sum_{v \in V} (1 - y_v) \pi_v$, the idea of threshold rounding is to not directly apply the above reasoning to an optimal solution of the PCTSP LP relaxation, but a scaled version thereof that can be obtained through the following lemma.

Lemma 3. *Let (x, y) be a feasible solution of the PCTSP LP relaxation. For every $\lambda \geq 0$, we can efficiently obtain a feasible solution (x_λ, y_λ) of the PCTSP LP relaxation such that $c^\top x_\lambda \leq \lambda \cdot c^\top x$ and*

$$y_{\lambda, v} = \min\{1, \lambda y_v\} \quad \text{for all } v \in V .$$

Simply speaking, Lemma 3 states that by investing a factor of λ more on the x-part of the solution, we can also boost the connectivities y by the same factor. We remark that Lemma 3 was already proved in [BGSW93] for optimal solutions of the (PCST LP relaxation) using the parsimonious property introduced by Goemans and Bertsimas [GB93]; we repeat it here for general feasible solutions (x, y), and provide a short proof in Section 4 for completeness. To complete the argument for threshold rounding started above, let (x^*, y^*) be an optimal solution of the PCTSP LP relaxation, and for a threshold $\gamma \in (0, 1]$, let (x, y) be the solution obtained from (x^*, y^*) through Lemma 3 with $\lambda = 1/\gamma$. We have seen that we can obtain a cycle $C = (V_C, E_C)$ of cost

$$c(E_C) \leq \frac{3}{2} c^\top x \leq \frac{3}{2\gamma} \cdot c^\top x^* \tag{1}$$

on $V_1 = \{v \in V \colon y_v = 1\} = \{v \in V \colon y_v^* \geq \gamma\}$. Additionally, we can now bound the penalty by $\pi(V \setminus V_C) = \sum_{v \notin V_1} \pi_v \leq \frac{1}{1-\gamma} \cdot \sum_{v \in V} (1 - y_v^*) \pi_v$. Choosing $\gamma = 3/5$ balances the coefficients $3/2\gamma$ and $1/(1-\gamma)$, and gives the upper bound of $5/2$ on the approximation factor, as showed in [BGSW93].

2.2 A refined tour construction

As already mentioned in the introduction, our refined tour construction allows us to include some vertices in the output cycle that are not contained in our core set V_1 while preserving the upper bound on the length of the cycle. To do so, we exploit a weak spot in the analysis of threshold rounding, namely the slack between the cost $c^\top z$ of the point $z \in P_{\mathrm{HK}}(G_1)$ defined in Section 2.1 and the cost $c^\top x$. Transforming x to $z \in P_{\mathrm{HK}}(G_1)$ is done by splitting off operations. In a single such operation, for some $\delta > 0$, we decrease the value of x on two well-chosen edges $\{v, u\}$ and $\{v, w\}$ by δ, and increase the value on $\{u, w\}$ by δ.

Thereby, the total cost decreases by $\delta c_{\{v,u\}} + \delta c_{\{v,w\}} - \delta c_{\{u,w\}}$, which is non-negative because c satisfies the triangle inequality. Viewed differently, we may exploit a budget of $c_{\{v,u\}} + c_{\{v,w\}}$ instead of $c_{\{u,w\}}$ for a δ-fraction of the edge $\{u,w\}$ in z, i.e., if $\{u,w\}$ is chosen to be part of a tour, we can afford to replace it by the path u-v-w with a certain probability, thereby gaining extra coverage of the vertex v while maintaining the total cost bound $c^\top x$ in expectation. This idea motivates backtracking the splitting off operations done to reach z, so that we can assign a certain budget to every component of z. Doing so in a careful way results in our main technical lemma, which is Lemma 5 below (a formal proof is deferred to Section 4). Before actually stating the lemma, let us discuss an assumption on feasible solutions of the PCTSP LP relaxation that we enforce repeatedly throughout the paper and which will prove useful later on.

Assumption 4. *For the feasible solution (x,y) of the PCTSP LP relaxation and $V_1 := \{v \in V : y_v = 1\}$, there is an edge $e_0 \in \binom{V_1}{2}$ with $x_{e_0} \geq 1$.*

We remark that Assumption 4 can be made without loss of generality. Indeed, we can leverage a trick used previously, for example by Karlin, Klein, and Oveis Gharan [KKO21], namely dividing the root node r and all incident edges into two (with penalty zero and equal edge costs, respectively), and adding an edge e_0 of length $c(e_0) = 0$ between the root and its copy. Note that an LP solution for the original instance can be transformed to an LP solution of the same value for the thereby obtained auxiliary instance by setting $y_v = 1$ for the copy of the root, $x_{e_0} = 2 - \frac{x(\delta(r))}{2}$, and distributing the x-weight on every edge incident to the root equally to its two copies in the auxiliary graph. (Note that $x_{e_0} > 1$ if $x(\delta(r)) < 2$.) Conversely, by shortcutting, we can transform a PCTSP solution in the auxiliary instance to a PCTSP solution of at most the same value in the original instance.

Lemma 5 (Decomposition lemma). *Let (x,y) be a feasible solution of the PCTSP LP relaxation satisfying Assumption 4 with edge e_0. We can in polynomial time construct a set $\mathcal{T}$ of trees and weights $\mu \in [0,1]^{\mathcal{T}}$ with the following properties:*

(i) The solution x is a conic combination of the trees in $\mathcal{T}$ with weights μ and the edge e_0, i.e.,

$$ x = \sum_{T \in \mathcal{T}} \mu_T \chi^{E[T]} + \chi^{e_0} \ . $$
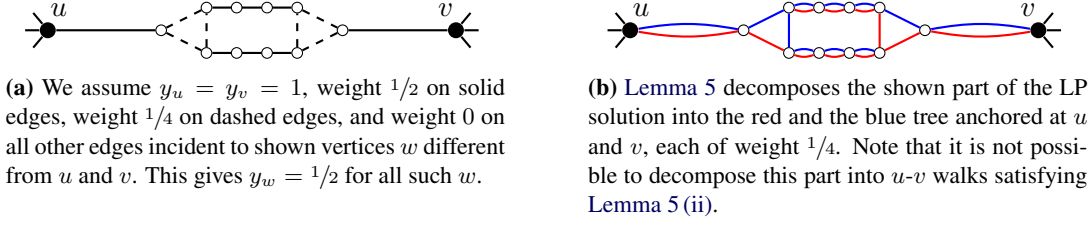
(ii) For every $v \in V \setminus V_1$,

$$ \sum_{T \in \mathcal{T} : v \in V[T]} \mu_T = y_v \ . $$

(iii) For every $T \in \mathcal{T}$, we have $|V[T] \cap V_1| = 2$, and we call the vertices in $V[T] \cap V_1$ the anchors of T.

(iv) For $T \in \mathcal{T}$, let $e_T := V[T] \cap V_1$ denote the edge joining the anchors of T, and let $G_1 := (V_1, E_1)$ be the multigraph with edge set $E_1 := \{e_0\} \cup \{e_T : T \in \mathcal{T}\}$. Then

$$ z := \sum_{T \in \mathcal{T}} \mu_T \chi^{e_T} + \chi^{e_0} \in P_{HK}(G_1) \ . $$

Note that Item (ii) in Lemma 5 guarantees that the connectivity y_v of vertices $v \in V \setminus V_1$ is reflected by the (weighted) number of trees incident to v in our decomposition. The latter will be crucial to obtain bounds on the penalties paid by our solutions for vertices in $V \setminus V_1$, and guaranteeing this property requires caution when constructing the trees in $\mathcal{T}$. We also remark that in our proof, the point $z \in P_{HK}(G_1)$ in Lemma 5 (iv) will be a point that can be obtained from x through splitting off operations. Figure 1 shows an example of our decomposition.

(a) We assume $y_u = y_v = 1$, weight $1/2$ on solid edges, weight $1/4$ on dashed edges, and weight 0 on all other edges incident to shown vertices w different from u and v. This gives $y_w = 1/2$ for all such w.

(b) Lemma 5 decomposes the shown part of the LP solution into the red and the blue tree anchored at u and v, each of weight $1/4$. Note that it is not possible to decompose this part into u-v walks satisfying Lemma 5 (ii).
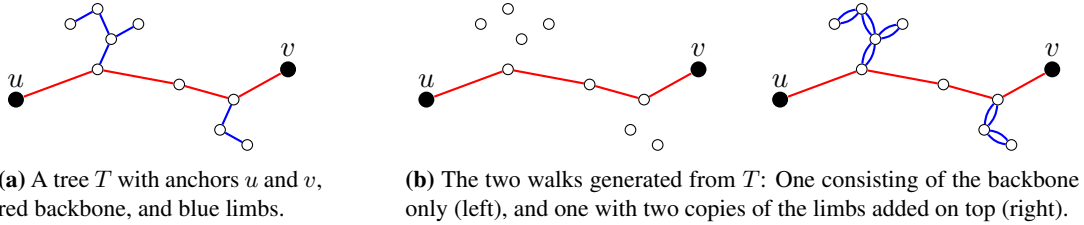
Figure 1: An excerpt of a solution of the PCTSP LP relaxation and its decomposition given by Lemma 5.

Our decomposition gives a more fine-grained analysis of the bound on the edge cost in threshold rounding than we obtained in (1): We have

$$c(E_C) \leq \frac{3}{2} c^\top z = \frac{3}{2} \sum_{T \in \mathcal{T}} \mu_T c_{e_T} \leq \frac{3}{2} \sum_{T \in \mathcal{T}} \mu_T c(E[T]) \leq \frac{3}{2} c^\top x \ . \tag{2}$$

Here, in the middle inequality, we are bounding c_{e_T} by $c(E[T])$. If T was a path connecting the endpoints of e_T, we could exploit this by, whenever e_T appears in a tour that we construct, replacing e_T with T, and thereby avoid penalty at all vertices on the path without affecting the upper bound in (2). In the general case, in order to maintain a tour, we can only plug in the path in T connecting the endpoints of e_T. This path and its complement are important notions that we use repeatedly, hence we define the following (also see Figure 2 for an illustration).

Definition 6. *Let T be a tree, and let V_1 be a vertex set such that $|V[T] \cap V_1| = 2$. The* anchors *of T in V_1 are the two vertices in $V[T] \cap V_1$. The* backbone *of T with respect to V_1, denoted by $\mathrm{backbone}_{V_1}(T)$, is the set of edges of T on the unique path connecting the two vertices in $V[T] \cap V_1$. The* limbs *of T with respect to V_1, denoted by $\mathrm{limbs}_{V_1}(T)$, is the set of edges of T that are not in the backbone of T with respect to V_1.*

(a) A tree T with anchors u and v, red backbone, and blue limbs.

(b) The two walks generated from T: One consisting of the backbone only (left), and one with two copies of the limbs added on top (right).

Figure 2: Anchors, backbone, and limb edges of a tree, and the two walks that can be constructed thereof.

Using this new notation, we can strengthen the bound in (2) to

$$c(E_C) \leq \frac{3}{2} c^\top z \leq \frac{3}{2} \sum_{T \in \mathcal{T}} \mu_T c(\mathrm{backbone}_{V_1}(T)) \ . \tag{3}$$

At the above cost, we can build a tour on the vertices in V_1 and replace edges of that tour with the corresponding backbone paths. This extends the set of vertices that are covered beyond V_1 by those lying on the added backbones. Moreover, compared to the upper bound $\frac{3}{2} c^\top x$, we still have a budget of $\frac{3}{2} \sum_{T \in \mathcal{T}} \mu_T c(\mathrm{limbs}_{V_1}(T))$ remaining. We can use this budget to add two copies of the limb edges of some trees to the cycle C in order to cover even more vertices. Note that by adding two copies, we make sure that the resulting edge set can again be shortcut to a cycle. Both of these observations contribute to saving penalties.

Illustrating how to translate the above intuition to actually improve coverage of vertices not in V_1 is easiest through exploiting randomization. Thus, we present a probabilistic approach here, and only later show that derandomization is possible. To build a basis for our tour, we exploit Assumption 4 and the structure obtained thereby: Lemma 5 (iv) implies that $z - \chi^{e_0} = \sum_{T \in \mathcal{T}} \mu_T \chi^T$ is in the *spanning tree polytope* $P_{\mathrm{ST}}(G_1)$ over G_1.[3] Being a matroid base polytope, we can exploit known marginal-preserving negatively correlated randomized rounding schemes in $P_{\mathrm{ST}}(G_1)$, concretely *randomized pipage rounding* (see Section 5 for details). Based on this, our approach is the following:

 (i) Sample a spanning tree of G_1 with marginals $z - \chi^{e_0}$, such that edges e_T appear with probability μ_T.
 (ii) Replace every sampled edge e_T by the path with edge set $\mathrm{backbone}_{V_1}(T)$, and with probability $3/4$ (independently for all sampled e_T), add two copies of $\mathrm{limbs}_{V_1}(T)$.
 (iii) Correct parities in the resulting graph using an appropriate join, and shortcut.

We remark that—instead of randomized pipage rounding—one could sample the spanning tree in Step (i) from a maximum entropy distribution and use the framework introduced by Karlin, Klein, and Oveis Gharan in their seminal $(1.5 - \varepsilon)$-approximation algorithm for TSP [KKO21; KKO22] to slightly reduce the cost of parity correction in Step (iii). However, the order of magnitude of this improvement (the authors show a bound of $\varepsilon > 10^{-36}$) is negligible in comparison with the loss that we incur through our analysis (see Remark 11). Additionally, derandomization of pipage rounding is rather straightforward in our setting using known tools (see Section 5).

Replacing each sampled tree edge e_T by the corresponding backbone together with parity correction leads to precisely the bound on the total edge cost in (3), and adding limb edges probabilistically as in Step (ii) uses up the remaining budget $\frac{3}{2} \sum_{T \in \mathcal{T}} \mu_T c(\mathrm{limbs}_{V_1}(T))$ in expectation. Adding two copies of $\mathrm{limbs}_{V_1}(T)$ implies that the edge set we replace e_T by can always be cast as a walk connecting the two endpoints of e_T, hence vertices in $V \setminus V_1$ will always have even degree. Consequently, in Step (iii), parities only need to be corrected at vertices in V_1, and hence at a cost that we can still bound by $\frac{1}{2} c^\top z$. Moreover, we can make use of the negative correlation property of the sampling procedure to prove bounds on the probability that a vertex $v \in V \setminus V_1$ is not covered by any of the used walks, and thereby obtain an upper bound on the expected penalty that we incur. (Note that in Step (iii), we may use a join that could possibly connect further vertices, but we do not exploit this in our analysis.) Summarizing Steps (i) and (ii) leads to the following lemma, whose formal proof we defer to Section 5.

Lemma 7. *Let (x, y) be a feasible solution of the PCTSP LP relaxation satisfying Assumption 4. Let $\mathcal{T}$ and $\mu \in [0, 1]^{\mathcal{T}}$ be a set of trees and weights with the properties guaranteed by Lemma 5. We can efficiently sample a multigraph H on a subset of V with the following properties:*
 (i) *H is connected, spans $V_1 := \{v \in V : y_v = 1\}$, and all vertices in $V \setminus V_1$ have even degree.*
 (ii) *We have*

$$\mathbb{E}\left[c(E[H]) + \pi(V \setminus V[H])\right]$$
$$\leq \sum_{T \in \mathcal{T}} \mu_T \cdot \left(c(\mathrm{backbone}_{V_1}(T)) + \frac{3}{2} \cdot c(\mathrm{limbs}_{V_1}(T)) \right) + \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3 y_v}{4} \right) .$$

With this result at hand, we can state a randomized version of our new algorithm, Algorithm 1. Note that the input only requires a feasible solution (x^*, y^*) of the PCTSP LP relaxation, even though typically, we will call Algorithm 1 on an optimal solution. Moreover, the choice of a suitable threshold $\gamma \in (0, 1]$ is left open for the moment; we show how to optimally exploit this remaining flexibility later on.

---

[3]Recall that the spanning tree polytope $P_{\mathrm{ST}}(G)$ over a graph $G = (V, E)$ is the convex hull of all incidence vectors of spanning trees in G, and can be described by $P_{\mathrm{ST}}(G) = \{x \in \mathbb{R}^E_{\geq 0} : x(E) = |V| - 1, \, x(E[S]) \leq |S| - 1 \,\forall S \subseteq V\}$.

Algorithm 1: Our new randomized algorithm for PCTSP.

Input: Feasible solution (x^*, y^*) of the PCTSP LP relaxation, threshold $\gamma \in (0, 1]$.

1. Guarantee that Assumption 4 is satisfied by modifying the instance accordingly if necessary.
2. Apply Lemma 3 to (x^*, y^*) with $\lambda = 1/\gamma$ to obtain another feasible solution (x, y).
3. Compute a set of trees $\mathcal{T}$ with weights $\mu \in [0, 1]^{\mathcal{T}}$ by applying Lemma 5 to (x, y).
4. Sample a multigraph H on V through Lemma 7 applied to (x, y) and $(\mathcal{T}, \mu)$.
5. Compute a shortest $\mathrm{odd}(H)$-join J.
6. Let C be a cycle obtained by shortcutting an Eulerian tour in $H \cup J$.

return C.

Using Lemma 7, we get the following guarantees. We remark that compared to classical threshold rounding, the factor $3/2\gamma$ on the term $c^\top x^*$ is unchanged, while at the same time, we get clearly improved factors on the penalty side.

Theorem 8. *Let (x^*, y^*) be a feasible solution of the PCTSP LP relaxation. On input (x^*, y^*) and $\gamma \in (0, 1]$, Algorithm 1 returns, in polynomial time, a cycle $C = (V_C, E_C)$ such that*

$$\mathbb{E}[c(E_C) + \pi(V \setminus V_C)] \leq \frac{3}{2\gamma} \cdot c^\top x^* + \sum_{v \in V \,:\, y_v^* < \gamma} \pi_v \cdot \exp\left(-\frac{3y_v^*}{4\gamma}\right) \ .$$

Proof. By Lemma 7, the multigraph H constructed in Algorithm 1 has odd degrees only at vertices in V_1, hence we can also correct parities by an $\mathrm{odd}(H)$-join in the multigraph $G_1 = (V_1, E_1)$ with edge set $E_1 = \{e_0\} \cup \{e_T : T \in \mathcal{T}\}$. By Item (iv) of Lemma 5, the point $z = \sum_{T \in \mathcal{T}} \mu_T \chi^{e_T} + \chi^{e_0}$ is in $P_{\mathrm{HK}}(G_1)$. Thus, following Wolsey's analysis, a shortest $\mathrm{odd}(H)$-join J in G_1 has cost at most

$$c(E[J]) \leq \frac{1}{2} c^\top z \leq \frac{1}{2} \sum_{T \in \mathcal{T}} \mu_T c(\mathrm{backbone}_{V_1}(T)) \ .$$

Note that J also is a shortest $\mathrm{odd}(H)$-join in the original complete graph G. Combining this with the bound of Item (ii) in Lemma 7 immediately gives the claimed guarantee

$$\mathbb{E}[c(E_C) + \pi(V \setminus V_C)] \leq \mathbb{E}[c(E[H]) + c(E[J]) + \pi(V \setminus V_C)]$$

$$\leq \frac{3}{2} c^\top x + \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3y_v}{4}\right) \leq \frac{3}{2\gamma} \cdot c^\top x^* + \sum_{v \in V \,:\, y_v^* < \gamma} \pi_v \exp\left(-\frac{3y_v^*}{4\gamma}\right) \ .$$

Here, in the last inequality, we use that by Lemma 3, we have $c^\top x \leq \frac{1}{\gamma} c^\top x^*$, and also $y_v = \frac{1}{\gamma} y_v^*$ for $v \in V \setminus V_1$ because $V \setminus V_1 = \{v \in V : y_v^* < \gamma\}$. To finish the proof, we note that indeed, all steps in Algorithm 1 can be implemented in polynomial time. $\square$

2.3 Choosing the right threshold

Directly balancing the penalty terms in Theorem 8 by a proper choice of the threshold γ gives the subsequent corollary. We remark that the resulting approximation guarantee does not yet beat the previously best known one of 1.915 by Goemans [Goe09], but it is significantly better than the standard analysis of threshold rounding, which gives a $\frac{5}{2}$-approximation.

Corollary 9. *Given a feasible solution (x^*, y^*) of the PCTSP LP relaxation and $\gamma = 1/(1 + \frac{2}{3} \exp(-3/4)) \approx 0.761$, Algorithm 1 returns in polynomial time a PCTSP solution of expected value at most $\alpha \cdot (c^\top x^* + \pi^\top(1 - y^*))$ with $\alpha = \frac{3}{2} + \exp(-3/4) < 1.973$.*

Proof. For general $\gamma \in (0, 1]$, we can bound the penalty term of the bound given in Theorem 8 as follows:

$$\sum_{v \in V: y_v^* < \gamma} \pi_v \exp\left(-\frac{3y_v^*}{4\gamma}\right) \le \max\left\{1, \frac{\exp(-3/4)}{1-\gamma}\right\} \cdot \sum_{v \in V: y_v^* < \gamma} \pi_v(1 - y_v^*) \ .$$

Here we used that for $y \in [0, \gamma]$, we have $\exp\left(-\frac{3y}{4\gamma}\right) \le \max\left\{1, \frac{\exp(-3/4)}{1-\gamma}\right\} \cdot (1 - y)$. The latter can be derived by noting that the inequality holds for $y \in \{0, \gamma\}$, and that $y \mapsto \exp\left(-\frac{3y}{4\gamma}\right)$ is convex in y. Thus, Theorem 8 implies

$$\mathbb{E}[c(E_C)] + \mathbb{E}[\pi(V \setminus V_C)] \le \max\left\{\frac{3}{2\gamma}, 1, \frac{\exp(-3/4)}{1-\gamma}\right\} \cdot \left(\sum_{e \in E} c_e x_e^* + \sum_{v \in V} \pi_v(1 - y_v^*)\right) \ ,$$

where the maximum expression is minimized for $\gamma = 1/(1 + \frac{2}{3}\exp(-3/4))$, giving the claimed approximation factor with respect to the objective value of the input solution (x^*, y^*) of the PCTSP LP relaxation. $\qquad\square$

When choosing the threshold γ with respect to a specific distribution (similarly to Goemans [Goe09]), we get a significant improvement of the best known approximation ratio. We remark that in contrast to the classical threshold rounding algorithm, it is no longer optimal to choose γ with respect to a uniform distribution over some interval in $[0, 1]$.

Theorem 10. *Let $b = 0.6945$. Sampling γ from the interval $[b, 1]$ using a distribution with density $f(\gamma) \propto \exp(-b/\gamma)$, and starting Algorithm 1 from a feasible solution (x^*, y^*) of the PCTSP LP relaxation, we get in polynomial time a PCTSP solution of expected value at most $\alpha \cdot (c^\top x^* + \pi^\top(1 - y^*))$, where $\alpha < 1.774$.*

The proof of Theorem 10 is postponed to Section 6. We remark that in contrast to the classical threshold rounding approach, we do not profit from combining our algorithm with the primal-dual algorithm by Goemans and Williamson [GW95]. This makes our algorithm self-contained and easy to state.

Note that the PCTSP LP relaxation can be solved in polynomial time through the ellipsoid method, as efficient separation over the constraints $x(\delta(S)) \ge y_v$ can be reduced to minimum r-v cut calculations for all $v \in V$ in the graph with capacities given by x. Thus, applying Theorem 10 to an optimal solution (x^*, y^*) of the PCTSP LP relaxation directly implies Theorem 1 if one allows randomization. In order to obtain a deterministic procedure, we show in Section 5 how to derandomize the proposed algorithm. Note that based on Theorem 10, there are two steps where we exploit randomization. First, we choose the threshold γ from some distribution. We will see that given an *optimal* LP-solution (x^*, y^*) of the PCTSP LP relaxation, it suffices to try all thresholds in the set $\{y_v^*: v \in V\}$ to get the claimed bound (cf. Section 5.3). Second, the construction of the multigraph H in Step 4 of Algorithm 1 is based on randomized pipage rounding in the spanning tree polytope of an auxiliary graph. We show in Section 5.2 that we can also use a deterministic version of pipage rounding to achieve the same guarantees. This will prove Theorem 1.

We make the reader aware that the first step above was immediate in the setting of Goemans [Goe09], because the values $\gamma = y_v^*$ for $v \in V$ already result in all possible supports of the tours that can be obtained through a threshold, and tours were subsequently built in a black-box way. In our refined approach, this is no longer the case; other values of γ may result in different tours, and it requires some thought to recover the conclusion that the best solution is obtained for some $\gamma = y_v^*$.

Remark 11. *Computational experiments (through discretizing the distribution to sample γ from) indicate that even by choosing an optimal distribution for γ, one cannot push the above analysis to prove a bound on the approximation factor of value less than or equal to 1.773. Consequently, the distribution over γ that we propose in Theorem 10 is very close to optimal for an analysis with respect to the bounds given by Theorem 8.*

10

3 Simple 2-approximations for PCTSP and PCST through decompositions

In this section, we show that for PCTSP and PCST, we can obtain an approximation guarantee similar to the one of the primal-dual approach by Goemans and Williamson [GW95] through very simple algorithms. More precisely, the proposed algorithms return cycles and trees, respectively, of cost at most twice the x-cost of an optimum LP solution plus once the y-cost of an optimum LP solution. The main ingredient giving rise to these algorithms is the following simplified variation of our decomposition lemma, which can also be obtained by utilizing an existential result on packing branchings in a directed multigraph by Bang-Jensen, Frank, and Jackson [BFJ95, Theorem 2.6].

Lemma 12. *Let (x, y) be a feasible solution of the PCTSP LP relaxation. We can in polynomial time compute a set of trees $\mathcal{T}$ containing the root r and weights $\mu \in [0, 1]^{\mathcal{T}}$ such that $\sum_{T \in \mathcal{T}} \mu_T = 1$,*

$$\sum_{T \in \mathcal{T}} \mu_T \chi^{E[T]} \leq x \ , \qquad and \qquad \forall v \in V \colon \sum_{T \in \mathcal{T} \colon v \in V[T]} \mu_T = y_v \ .$$

We prove Lemma 12 through a common generalization with Lemma 5 in Section 4. Based on Lemma 5, the simple algorithm we propose for PCTSP is the following.

Algorithm 2: A simple 2-approximation for PCTSP

Input: PCTSP instance (G, r, c, π) on $G = (V, E)$.

1. Compute an optimal solution (x^*, y^*) of the PCTSP LP relaxation.
2. Compute a set of trees $\mathcal{T}$ through Lemma 12 applied to (x^*, y^*).
3. For each tree $T \in \mathcal{T}$, obtain a cycle C_T through duplicating T and shortcutting.
4. Let $S = \arg\min_{T \in \mathcal{T}} c(E[C_T]) + \pi(V \setminus V[C_T])$.

return C_S.

Theorem 13. *Algorithm 2 returns, in polynomial time, a cycle C satisfying*

$$c(E[C]) + \pi(V \setminus V[C]) \leq 2 \cdot c^\top x^* + \pi^\top (1 - y^*) \ ,$$

where (x^*, y^*) is the optimal solution of the PCTSP LP relaxation computed in Step 1 of the algorithm.

Proof. We do a probabilistic analysis of the algorithm. The trees in $\mathcal{T}$ from Lemma 12 come with weights $\mu \in [0, 1]^{\mathcal{T}}$ such that $\sum_{T \in \mathcal{T}} \mu_T = 1$ and $\sum_{T \in \mathcal{T}} \mu_T \chi^{E[T]} \leq x^*$. Sampling a tree T from $\mathcal{T}$ with marginals μ gives

$$\mathbb{E}[c(E[C_T])] = \sum_{T \in \mathcal{T}} \mu_T \cdot 2c(E[T]) = 2 \cdot c^\top \left(\sum_{T \in \mathcal{T}} \mu_T \chi^{E[T]} \right) \leq 2 \cdot c^\top x^* \ .$$

Similarly, for the penalty term we get

$$\mathbb{E}[\pi(V \setminus V[C_T])] = \sum_{v \in V} \mathbb{P}[v \notin T] \cdot \pi_v$$

$$= \sum_{v \in V} \left(1 - \sum_{T \in \mathcal{T} \colon v \in V[T]} \mu_T \right) \cdot \pi_v = \sum_{v \in V} (1 - y_v^*) \cdot \pi_v = \pi^\top (1 - y^*) \ .$$

Hence, the expected objective value of C_T can be bounded by $2 \cdot c^\top x^* + \pi^\top (1 - y^*)$. Thus, also at least one of the cycles C_T for $T \in \mathcal{T}$ satisfies the desired guarantee. $\square$

For PCST, we consider the following very related LP relaxation. Recall that in PCST, we do not assume the triangle inequality for distances, and thus also the LP relaxation comes without degree constraints.

$$
\min \quad \sum_{e \in E} c_e x_e + \sum_{v \in V} \pi_v (1 - y_v)
$$

$$
\begin{aligned}
x(\delta(S)) &\geq y_v & \forall S \subseteq V \setminus \{r\}, v \in S \\
y_r &= 1 \\
x_e &\geq 0 & \forall e \in E \\
y_v &\in [0,1] & \forall v \in V \ .
\end{aligned}
\qquad \text{(PCST LP relaxation)}
$$

The main step towards mimicking Algorithm 2 is to transform a solution of the PCST LP relaxation to one of the PCTSP LP relaxation to be able to apply Lemma 12 in a black-box way, and later transform the obtained components back (see Step 4 of Algorithm 3). The latter is required because edge costs in PCST are not necessarily metric.

Algorithm 3: A simple 2-approximation for PCST

Input: PCST instance (G, r, c, π) on $G = (V, E)$.

1. Compute an optimal solution (x^*, y^*) of the PCST LP relaxation.
2. Consider the PCTSP instance $(\bar{G}, r, \bar{c}, \pi)$ where $\bar{c}$ denotes the metric closure of the edge lengths c on the complete graph $\bar{G}$ on V. Obtain a solution (x, y^*) of the PCTSP LP relaxation from $(2x^*, y^*)$ by splitting off until all degree constraints are satisfied.
3. Compute a set of trees $\mathcal{T}$ through Lemma 12 applied to (x, y^*).
4. For each $T \in \mathcal{T}$, let T_0 be obtained from T through replacing edges by shortest paths connecting their endpoints in G.
5. Let $S = \arg\min_{T \in \mathcal{T}} c(E[T_0]) + \pi(V \setminus V[T_0])$.

return *a minimum spanning tree on* $V[S_0]$.

Theorem 14. *Algorithm 3 returns, in polynomial time, a tree T satisfying*

$$
c(E[T]) + \pi(V \setminus V[T]) \leq 2 \cdot c^\top x^* + \pi^\top (1 - y^*) \ ,
$$

where (x^*, y^*) is the optimal solution of the PCST LP relaxation computed in Step 1 of the algorithm.

Proof. As in Algorithm 3, let $\bar{c}$ denote the metric closure of the (not necessarily metric) edge lengths c. Note that $\bar{c}^\top x \leq 2 c^\top x^*$, as splitting off does not increase the total length with respect to metric lengths. Additionally, for the connected multigraphs T_0 computed in Step 4, we have $\bar{c}(E[T]) = c(E[T_0])$ by definition. With this at hand, we can again do a randomized analysis, where we sample a tree T from $\mathcal{T}$ with marginals equal to the weights μ obtained through Lemma 12. This gives an expected objective value

$$
\mathbb{E}[c(E[T_0]) + \pi(V \setminus V[T_0])] = \mathbb{E}[\bar{c}(E[T])] + \sum_{v \in V} \Pr[v \notin T] \cdot \pi_v
$$

$$
\leq \sum_{T \in \mathcal{T}} \mu_T \bar{c}(E[T]) + \sum_{v \in V} \left(1 - \sum_{T \in \mathcal{T} : \, v \in V[T]} \mu_T \right) \cdot \pi_v
$$

$$
\leq \bar{c}^\top x + \sum_{v \in V} (1 - y_v^*) \pi_v
$$

$$
\leq 2 c^\top x^* + \pi^\top (1 - y^*) \ .
$$

Clearly, the same bound holds for a minimum spanning tree on $V[T_0]$. Thus again, we get the desired bound in expectation, hence for at least one $T \in \mathcal{T}$, we obtain the claimed guarantee. $\qquad \square$

4 From splitting off to tree decompositions

In this section, we prove Lemmas 3, 5 and 12. All three of these lemmas share a common technique in the background, namely *splitting off*, hence we start by recalling the concept. Splitting off was initially introduced as a fundamental tool in Graph Theory [Lov76; Mad78; Fra92] for modifying a graph while maintaining certain connectivity properties. We exploit a weighted version of splitting off. Here, given a complete graph $G = (V, E)$ and edge weights $w \colon E \to \mathbb{R}_{\geq 0}$, a splitting operation at a vertex $v \in V$ is the following: For two edges $e = \{v, u\}$ and $f = \{v, w\}$ incident to v and $\delta \in (0, \min\{w(e), w(f)\}]$, reduce the edge weights of e and f by δ, and increase the edge weight of $\{u, w\}$ by δ. In case $e = f = \{v, u\}$, the weight of e is reduced by δ, and we call the splitting operation *degenerate*. A splitting (e, f, δ) at v is called *feasible* if minimum s-t cut sizes in G with respect to the weights w are preserved under the splitting for all $s, t \in V \setminus \{v\}$. A *complete splitting* at $v \in V$ denotes a sequence of feasible splitting operations at v such that all edges incident to v have weight zero after performing the splittings. Frank [Fra92] showed that such a complete splitting always exists. It is straightforward to see that a complete splitting consisting of less than $|V|^2$ splitting operations can be found through a polynomial number of minimum s-t cut computations, implying the following theorem.

Theorem 15. *Let $G = (V, E)$ be a complete graph with edge weights $w \colon E \to \mathbb{R}_{\geq 0}$. Let $v \in V$ and $\beta \in [0, w(\delta(v))]$. There is a deterministic algorithm that computes in polynomial time a sequence of less than $\mathrm{poly}(|V|)$ many feasible splitting operations at v such that the edges incident to v have total weight β with respect to the resulting weight function. Note that if $\beta = 0$, this results in a complete splitting at v.*

We remark that it is known that already a linear number of splitting operations are sufficient to obtain a complete splitting. For further details, we refer to [NI97; Nag06] and references therein. Moreover, results of the above kind are typically proved for $\beta = 0$ only. The extension to $\beta > 0$ is immediate, since a sequence of splitting operations at a vertex v can simply be pruned as soon as the edges incident to v have remaining weight β. Having Theorem 15 at hand, we can readily prove Lemma 3.

Proof of Lemma 3. Consider the point $(\lambda x, y_\lambda)$. This point satisfies all cut constraints in the PCTSP LP relaxation, but might violate some of the degree constraints. If not, we are done; hence assume the contrary and fix a vertex $s \in V$ such that the degree constraint at s is violated. We then have $\lambda x(\delta(s)) = 2\lambda y_s > 2y_{\lambda,s} = 2$. Thus, we can apply Theorem 15 with $\beta = 2$ at the vertex s, resulting in a new weight function $\bar{x}$ after performing the splitting. All splitting operations are feasible, hence minimum r-v cut sizes are preserved for all $v \in V \setminus \{r, s\}$, and, in particular, equal to $\bar{x}(\delta(v)) = \lambda x(\delta(v))$ due to the initial degree constraints satisfied by x. Consequently, $(\bar{x}, y_\lambda)$ satisfies all cut constraints in the PCTSP LP relaxation. By the same argument, there cannot be degenerate splitting operations. Hence, degrees of vertices other than s are invariant under the splitting, i.e., degree constraints satisfied by λx are also satisfied by $\bar{x}$. Additionally, $\bar{x}$ satisfies the degree constraint at s by construction. Iterating this for all vertices $s \in V$ with violated degree constraints leads to a pair (x_λ, y_λ) feasible for the PCTSP LP relaxation. By the triangle inequality, applying a splitting off operation does not increase the total cost of the solution, hence $c^\top x_\lambda \leq \lambda c^\top x$, as desired. $\square$

Next, we show Lemmas 5 and 12. More precisely, we prove that the following common generalization of the two lemmas is true.

Lemma 16. *Let (x, y) be feasible for the PCTSP LP relaxation, and let $U \subseteq V_1 \coloneqq \{v \in V \colon y_v = 1\}$. Assume that there is an edge $e_0 \in \binom{U}{2}$ with $x_{e_0} \geq 1$. We can in polynomial time construct a set $\mathcal{T}$ of trees and weights $\mu \in [0, 1]^{\mathcal{T}}$ with the following properties:*

(i) The solution x is a conic combination of the trees in $\mathcal{T}$ with weights μ and the edge e_0, i.e.,

$$x = \sum_{T \in \mathcal{T}} \mu_T \chi^{E[T]} + \chi^{e_0} \ .$$

(ii) For every $v \in V \setminus U$,

$$\sum_{T \in \mathcal{T} \,:\, v \in V[T]} \mu_T = y_v \ .$$

(iii) For every $T \in \mathcal{T}$, we have $|V[T] \cap U| = 2$, and we call the vertices in $V[T] \cap U$ the anchors *of* T.

(iv) For $T \in \mathcal{T}$, let $e_T := V[T] \cap U$ denote the edge joining the anchors of T, and let $H := (U, F)$ be the multigraph with edge set $F := \{e_0\} \cup \{e_T : T \in \mathcal{T}\}$. Then

$$z := \sum_{T \in \mathcal{T}} \mu_T \chi^{e_T} + \chi^{e_0} \in P_{HK}(H) \ .$$

We start by showing that indeed, Lemma 16 is a common generalization of Lemmas 5 and 12.

Proof of Lemmas 5 and 12. Lemma 5 immediately follows from Lemma 16 by choosing $U = V_1$. In order to utilize Lemma 16 to prove Lemma 12, we note that we can guarantee that Assumption 4 is satisfied: If not, modify the given instance as described in Section 2 and transform the trees computed in the thereby obtained auxiliary graph into trees in the original graph by contracting e_0 and deleting some edge if necessary.

Now, choose U to be the vertex set containing the endpoints of e_0 only. In this case, the graph H in Item (iv) consists of parallel edges only, with one of them being e_0, and every tree in $\mathcal{T}$ corresponding to one of the other edges. Moreover, the total edge weight of any feasible solution in the Held-Karp polytope over a two-vertex graph is 2. In z, at least one unit is taken by the edge e_0; hence we must have $\sum_{T \in \mathcal{T}} \mu_T \leq 1$. On the other hand, by Item (ii)

$$\sum_{T \in \mathcal{T}} \mu_T \geq \sum_{T \in \mathcal{T} \,:\, r \in V[T]} \mu_T = y_r = 1 \ ,$$

implying $\sum_{T \in \mathcal{T}} \mu_T = 1$, and $r \in V[T]$ for each $T \in \mathcal{T}$. □

It remains to prove Lemma 16. On a high level, the proof is organized as follows. We apply Theorem 15 iteratively to find complete splittings at vertices $v \in V \setminus U$, in order of increasing connectivity y_v, until only vertices in U remain. In the resulting graph, it is trivial to find a decomposition meeting the requirements: We can simply consider each individual edge as a tree, with the corresponding weight being the edge weight. Carefully undoing the splitting off operations in a way similar to Bang-Jensen, Frank, and Jackson [BFJ95], and exploiting that we did them in increasing order of connectivity, we will show how to alter the initial set of trees so to obtain one with the desired properties for the original instance.

Proof of Lemma 16. We prove the statement by induction on $|V \setminus U|$. If $|V \setminus U| = 0$, there is nothing to be done: Creating a tree T from every single edge, with μ_T being equal to the edge weight (or equal to the edge weight minus 1 in case of e_0), clearly satisfies all the properties.

If $|V \setminus U| > 0$, consider a vertex $s \in V \setminus U$ of minimum connectivity y_s. By Theorem 15, we can efficiently compute a sequence of feasible splitting off operations resulting in a complete splitting at s. Note that the weights of minimum r-v-cuts are preserved under the splittings for all $v \in V \setminus \{r, s\}$. By the degree constraints of the PCTSP LP relaxation, this implies that the degrees of vertices different from s are preserved as well. The latter has two implications: First, none of the splitting operations at s can be degenerate, as degenerate splittings would reduce the degree of some vertex. Second, the graph on $V' = V \setminus \{s\}$ satisfies the assumptions of Lemma 16 with respect to the resulting edge weights. Consequently,

by the inductive assumption, we can compute in polynomial time edge-disjoint trees $\mathcal{T}$ with the desired properties; in particular, for every $v \in V' \setminus U$,

$$\sum_{T \in \mathcal{T} : v \in V[T]} \mu_T = y_v \ .$$

We now undo the splitting operations at s and modify the trees in $\mathcal{T}$ accordingly. We note that while intermediate steps may lead to $\mathcal{T}$ being a multiset of trees, we can merge identical trees at the very end by simply adding their weights. By Item (iv), this does not generate weights larger than 1. Before we start undoing the splitting operations, we initialize auxiliary variables $\text{spare}_v = 0$ for each $v \in V'$.

Let $(e = \{s, u\}, f = \{s, v\}, \delta)$ with $e \neq f$ be one of the splitting off operations at s that we want to revert. Let $\mathcal{T}' = \{T_1, \ldots, T_k\} \subseteq \mathcal{T}$ be a minimal subset of trees with $\{u, v\} \in E[T]$ for each $T \in \mathcal{T}'$ and $\sum_{T \in \mathcal{T}'} \mu_T \geq \delta$. Let $\varepsilon = \sum_{T \in \mathcal{T}'} \mu_T - \delta$. Note that $\varepsilon < \mu_{T_k}$ by the minimality of $\mathcal{T}'$. If $\varepsilon > 0$, add a copy T'_k of T_k to $\mathcal{T}$, set $\mu_{T'_k} = \varepsilon$, and reduce μ_{T_k} by ε. After this modification, $\sum_{T \in \mathcal{T}'} \mu_T = \delta$. For each $T \in \mathcal{T}'$ do the following:

(i) If $s \notin V[T]$, remove $\{u, v\}$ from $E[T]$ and add $\{s, u\}$ and $\{s, v\}$ to $E[T]$.

(ii) If $s \in V[T]$, remove $\{u, v\}$ from $E[T]$ and add either $\{s, u\}$ or $\{s, v\}$ to $E[T]$ such that T remains acyclic. If $\{s, u\}$ is added to $E[T]$ increase spare_v by μ_T. Otherwise, increase spare_u by μ_T.

Note that in case (i), the total weight of trees containing s increases by μ_T. Otherwise, either spare_u or spare_v increases by μ_T. Consequently, through the above operations, $\sum_{T \in \mathcal{T} : s \in V[T]} \mu_T + \sum_{v \in V'} \text{spare}_v$ increases by δ. Since the splitting off operation (e, f, δ) decreased the degree of s by 2δ, we get, after reverting all splitting off operations at s,

$$\sum_{T \in \mathcal{T} : s \in V[T]} \mu_T + \sum_{v \in V'} \text{spare}_v = \frac{x(\delta(s))}{2} = y_s \ .$$

Now, for each $w \in V'$ do the following: If $\text{spare}_w > 0$, find a minimal subset of trees $\mathcal{T}'' = \{T_1, \ldots, T_l\} \subseteq \mathcal{T}$ with $w \in V[T]$ but $s \notin V[T]$ for each $T \in \mathcal{T}''$, and $\sum_{T \in \mathcal{T}''} \mu_T \geq \text{spare}_w$. Note that such a subset of trees always exists since

$$\sum_{T \in \mathcal{T} : w \in V[T]} \mu_T \geq y_w \geq y_s = \sum_{T \in \mathcal{T} : s \in V[T]} \mu_T + \sum_{v \in V'} \text{spare}_v \ ,$$

where we used in the second inequality that s has minimum connectivity. Note that the first inequality is satisfied by the inductive assumption: If $w \notin U$, the bound is tight by Item (ii) of Lemma 16. If $w \in U$, then $\sum_{T \in \mathcal{T} : w \in V[T]} \mu_T \geq 1 = y_w$ by Item (iv) of Lemma 16 and the degree constraints in the Held-Karp polytope (note that e_0 may be incident to w). Let $\varepsilon = \sum_{T \in \mathcal{T}''} \mu_T - \text{spare}_w$. Note that $\varepsilon < \mu_{T_l}$ by the minimality of $\mathcal{T}''$. If $\varepsilon > 0$, add a copy T'_l of T_l to $\mathcal{T}$ and set $\mu_{T'_l} = \varepsilon$ and reduce μ_{T_l} by ε. For each $T \in \mathcal{T}''$, add $\{s, w\}$ to $E[T]$. Note that this increases the total weight of the trees containing s by spare_w. Hence, after using up all spares, we get

$$\sum_{T \in \mathcal{T} : s \in V[T]} \mu_T = y_s \ . \tag{4}$$

Note that throughout the above operations, the graphs $T \in \mathcal{T}$ are trees, and after reverting all splitting off operations, x is a conic combination of the trees in $\mathcal{T}$ with weights μ and the edge e_0 by construction. Furthermore, the above operations do not change the intersection of the trees in $\mathcal{T}$ with U, and splitting trees does not affect Item (iv). Hence, Item (iii) and Item (iv) still hold after the above operations. Moreover, for every vertex in V', the total weight of trees covering this vertex is unchanged. Hence, Item (ii) is still satisfied for each $v \in V' \setminus U$, and also for s by (4).

Finally, we note that our construction can be executed in polynomial time. Indeed, by Theorem 15, in each step of our inductive procedure, we have to revert less than $\mathrm{poly}(|V|)$ many splitting operations, which increases the total number of trees in $\mathcal{T}$ by an additive $\mathrm{poly}(|V|)$. This implies that the size of $\mathcal{T}$ remains polynomially bounded throughout. □

5 Sampling walks and derandomization

In this section, we prove Lemma 7, which is at the core of our new randomized algorithm, giving the basis of the returned tour. Furthermore, we show how to replace the randomized steps of our algorithm with deterministic procedures while obtaining the same guarantees. This includes derandomizing Lemma 7, but also derandomizing the random choice of a suitable threshold γ. Together with the analysis presented earlier, these ingredients prove our main result, Theorem 1.

Before going into the more technical parts, let us recall the pipage rounding procedure, which we use in a black-box way for both the proof of Lemma 7 and its deterministic analogue. Pipage rounding is a technique that goes back to Ageev and Sviridenko [AS04] and Srinivasan [Sri01]. Generalized to a matroid setting and put into a randomized framework by Calinescu, Chekuri, Pál, and Vondrák [CCPV11], randomized pipage rounding is one of several extensively studied efficient sampling procedures in matroid and matroid base polytopes (also see [CVZ10; HO14] and references therein). Generally, the idea is to start from a fractional point x, and modify this solution iteratively in a well-chosen way until an integral solution is found. In pipage rounding, each single modification is a maximal step along a direction parallel to some $e_i - e_j$, where e_k is the kth unit vector. Because the directions $e_i - e_j$ are precisely the edge directions of matroid base polytopes, such moves are enough to eventually end up in a vertex. In a randomized setting, probabilities for going either way are chosen such that in every step, marginals of every component are preserved. Concretely, if for a matroid $\mathcal{M}$, $B_{\mathcal{M}}$ denotes the associated matroid base polytope, i.e., the convex hull of all characteristic vectors of bases in $\mathcal{M}$, one can obtain the following.

Theorem 17 ([CVZ10, direct implication of Theorem 1.1]). *Let $x = (x_1, \ldots, x_n) \in B_{\mathcal{M}}$ be a fractional solution in the matroid base polytope, and let $X = (X_1, \ldots, X_n) \in \{0, 1\}^n$ be an integral solution obtained from x using randomized pipage rounding. Then, the following holds.*

(i) *For every $i \in [n]$, we have $\mathbb{E}[X_i] = x_i$.*

(ii) *For every $T \subseteq [n]$, we have*

$$\mathbb{E}\left[\prod_{i \in T} X_i\right] \leq \prod_{i \in T} x_i \qquad and \qquad \mathbb{E}\left[\prod_{i \in T}(1 - X_i)\right] \leq \prod_{i \in T}(1 - x_i) \ .$$

Thus, besides preserving marginals of a point in the spanning tree polytope exactly, randomized pipage rounding also guarantees that the events of edges being sampled as well as those of edges not being sampled are negatively correlated, a property that we crucially exploit to bound the total penalty that our PCTSP solutions incur. We remark that there are other known procedures achieving the same guarantees, in particular randomized swap rounding [CVZ10], or, at a $1 \pm \varepsilon$ error in the marginals, sampling from λ-uniform spanning tree distributions (see, for example, [AGMGS17; KKO21]).

For a deterministic version of pipage rounding, we use a formulation of Harvey and Olver [HO14] that is based on earlier work mentioned above. Here, the decisions on the direction of every single step are guided by a function $g \colon \mathcal{B}_{\mathcal{M}} \to \mathbb{R}$ that is required to be *concave under swaps*, i.e., for all $x \in \mathcal{B}_{\mathcal{M}}$ and all unit vectors e_i, e_j, we require that the function $t \mapsto g(x + t(e_i - e_j))$ is concave. This concavity under swaps allows to perform steps along decreasing values of g, and one can obtain the following.

Theorem 18 (Deterministic Pipage Rounding [HO14]). *There is a deterministic, polynomial-time algorithm that, given $x \in B_{\mathcal{M}}$ and a value oracle for a function g that is concave under swaps, outputs an extreme point $\hat{x}$ of $B_{\mathcal{M}}$ with $g(\hat{x}) \leq g(x)$.*

5.1 Proof of Lemma 7

As already outlined earlier, we prove Lemma 7 by exploiting the decomposition of a solution (x, y) of the PCTSP LP relaxation into a family $\mathcal{T}$ of trees through Lemma 5. Concretely, we will split each tree into two walks with corresponding weights, one of them consisting of backbone edges only, while the other contains two copies of the limb edges on top (see Figure 2 on Page 7). To sample walks, we leverage Item (iv) of Lemma 5, concretely the implication that $z - \chi^{e_0}$ is in the spanning tree polytope over the multigraph $G_1 = (V_1, E_1)$, where $E_1 := \{e_0\} \cup \{e_T : T \in \mathcal{T}\}$. In this graph, every edge corresponds to a tree; by duplicating every edge and splitting the edge weight accordingly, we can associate to every edge one of the walks constructed earlier. Also, the resulting edge weights still form a point in the spanning tree polytope of the blown-up graph, and we sample a spanning tree in this graph with the edge weights as marginals through pipage rounding. Sampled edges can then be replaced by the corresponding walks. The union of the edge sets of all these walks induces the multigraph H claimed by Lemma 7. Using the properties guaranteed by Theorem 17, we can prove Lemma 7.

Proof of Lemma 7. Let $\mathcal{T}$ and weights $\mu \in [0,1]^{\mathcal{T}}$ be obtained from Lemma 5. We construct a set $\mathcal{W}_0$ of walks as follows. For each $T \in \mathcal{T}$ with anchors $s, t \in V_1$, add one s-t walk $W_1(T)$ to $\mathcal{W}_0$ that consists of $\mathrm{backbone}_{V_1}(T)$ and two copies of $\mathrm{limbs}_{V_1}(T)$, and set $\nu_{W_1(T)} = \frac{3}{4}\mu_T$. As we add two copies of the limb edges, the resulting edge set has even degrees except at s and t, and can thus indeed be cast as an s-t walk. On top of that, add another s-t walk $W_2(T)$ to $\mathcal{W}_0$ that consists of $\mathrm{backbone}_{V_1}(T)$ only, and set $\nu_{W_2(T)} = \frac{1}{4}\mu_T$.

For each walk $W \in \mathcal{W}_0$, let e_W be the direct edge joining its start- and endpoint. Let F be the multiset of all these edges. Now, consider the point $z_0 \in [0,1]^F$ defined by $z_0(e_W) = \nu_W$ for each $W \in \mathcal{W}_0$. As z_0 and the point $z - \chi^{e_0}$ (where z is defined as in Item (iv) of Lemma 5) differ only by splitting edges into parallel edges of the same total weight, it immediately follows that $z_0 + \chi^{e_0} \in P_{\mathrm{HK}}(G_0)$, where $G_0 = (V_1, F)$. From the polyhedral descriptions of $P_{\mathrm{HK}}(G_0)$ and the spanning tree polytope $P_{\mathrm{ST}}(G_0)$ over G_0, one can see that this in turn implies that $z_0 \in P_{\mathrm{ST}}(G_0)$. Applying Theorem 17 with z_0, we thus obtain a spanning tree of (V_1, F_1) with $F_1 \subseteq F$. Let

$$\mathcal{W} := \{W \in \mathcal{W}_0 : e_W \in F_1\} \ .$$

Because $\mathcal{W} \subseteq \mathcal{W}_0$, each walk in $\mathcal{W}$ starts and ends in V_1. Since (V_1, F_1) is spanning, the multigraph H induced by the edge sets of the walks in $\mathcal{W}$ is connected and always spans V_1. Let $v \in V \setminus V_1$. Then

$$
\begin{aligned}
\mathbb{P}[v \notin H] &\leq \prod_{W \in \mathcal{W}_0 : v \in V[W]} \mathbb{P}[e_W \notin F] && \text{(by Theorem 17 (ii))} \\
&\leq \prod_{T \in \mathcal{T} : v \in V[T]} \mathbb{P}[e_{W_1(T)} \notin F] \\
&= \prod_{T \in \mathcal{T} : v \in V[T]} \left(1 - \frac{3}{4}\mu_T\right) && \text{(by Theorem 17 (i))} \\
&\leq \exp\left(-\frac{3}{4} \sum_{T \in \mathcal{T} : v \in V[T]} \mu_T\right) && \text{(using } 1 - x \leq \exp(-x)) \\
&= \exp\left(-\frac{3y_v}{4}\right) \ . && \text{(by Lemma 5)}
\end{aligned}
$$

17

Consequently,

$$\mathbb{E}[\pi(V \setminus V[H])] = \sum_{v \in V \setminus V_1} \pi_v \Pr[v \notin H] \leq \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3y_v}{4}\right) \; . \tag{5}$$

It remains to bound the expected total cost of the sampled walks. Using Theorem 17 (i) again, we get

$$
\begin{aligned}
\mathbb{E}[c(E[H])] &= \sum_{W \in \mathcal{W}_0} \nu_W \cdot c(E[W]) \\
&= \sum_{T \in \mathcal{T}} \left(\frac{3}{4}\mu_T \cdot c(E[W_1(T)]) + \frac{1}{4}\mu_T \cdot c(E[W_2(T)]) \right) \\
&= \sum_{T \in \mathcal{T}} \mu_T \left(c(\text{backbone}_{V_1}(T)) + \frac{3}{2} \cdot c(\text{limbs}_{V_1}(T)) \right) \; .
\end{aligned}
\tag{6}
$$

Together, (5) and (6) imply the bound claimed by Lemma 7. Finally, note that Theorem 17 guarantees that we can sample $\mathcal{W}$, and thus obtain H, efficiently. $\qquad\square$

5.2 A deterministic selection of walks

In this section, we show how to deterministically construct the graph H given by Lemma 7 while maintaining the overall guarantees on the performance of our algorithm. To be precise, we show the following deterministic analogue of Lemma 7.

Lemma 19. *Let (x, y) be a feasible solution of the PCTSP LP relaxation satisfying Assumption 4. Let $\mathcal{T}$ and $\mu \in [0, 1]^{\mathcal{T}}$ be a set of trees and weights with the properties guaranteed by Lemma 5. We can deterministically and in polynomial time obtain a multigraph H on a subset of V with the following properties:*

(i) H is connected, spans $V_1 := \{v \in V : y_v = 1\}$, and all vertices in $V \setminus V_1$ have even degree.

(ii) We have

$$
\begin{aligned}
&c(E[H]) + \pi(V \setminus V[H]) \\
&\qquad \leq \sum_{T \in \mathcal{T}} \mu_T \cdot \left(c(\text{backbone}_{V_1}(T)) + \frac{3}{2} \cdot c(\text{limbs}_{V_1}(T)) \right) + \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3y_v}{4}\right) \; .
\end{aligned}
$$

To prove Lemma 19, we recall the randomized construction from the proof of Lemma 7 (see Section 5.1): There, the graph H is based on a spanning tree sampled in an auxiliary graph through randomized pipage rounding. We now show that the deterministic analogue given through Theorem 18 can achieve the same guarantees.

Proof of Lemma 19. Starting from a solution (x, y) of the PCTSP LP relaxation and repeating the construction from the proof of Lemma 7, we define the multigraph $G_0 = (V_1, F)$, where for each $T \in \mathcal{T}$, there are two parallel edges in F joining the anchors of T: One corresponding to the walk $W_1(T)$ consisting of $\text{backbone}_{V_1}(T)$ and two copies of $\text{limbs}_{V_1}(T)$ with weight $\nu_{W_1(T)} = \frac{3}{4}\mu_T$, and one corresponding to the walk $W_2(T)$ consisting of $\text{backbone}_{V_1}(T)$ with weight $\nu_{W_2(T)} = \frac{1}{4}\mu_T$. Again, we denote by $\mathcal{W}_0$ the set of all walks constructed this way, and let e_W be the edge in F corresponding to the walk W. Moreover, we know that $z_0 \in [0, 1]^F$ defined by $z_0(e_W) = \nu_W$ satisfies $z_0 \in P_{\text{ST}}(G_0)$.

We want to apply deterministic pipage rounding, i.e., Theorem 18, in the spanning tree polytope $P_{\text{ST}}(G_0)$ starting from z_0 with the function $g \colon P_{\text{ST}}(G_0) \to \mathbb{R}$ defined by

$$g(z) := \sum_{W \in \mathcal{W}_0} c(E[W]) z_{e_W} + \sum_{v \in V \setminus V_1} \pi_v \prod_{W \in \mathcal{W}_0 \colon v \in W} (1 - z_{e_W}) \; .$$

18

To this end, we prove the following properties.

Claim 20.
  (i) *g is concave under swaps.*
  (ii) *If z is the incidence vector of a spanning tree in G_0, and H is the multigraph obtained from that spanning tree by replacing every edge e_W by the walk W, then $g(z) = c(E[H]) + \pi(V \setminus V[H])$.*
 (iii) *We have*

$$g(z_0) \le \sum_{T \in \mathcal{T}} \mu_T \cdot \left(c(\mathrm{backbone}_{V_1}(T)) + \frac{3}{2} \cdot c(\mathrm{limbs}_{V_1}(T)) \right) + \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3y_v}{4} \right) \ .$$

Clearly, Lemma 19 immediately follows from the claim: Item (i) allows for applying Theorem 18, while Items (ii) and (iii) show that the graph H satisfies the bound in Item (ii) of Lemma 19. Moreover, H is clearly connected, spans V_1, and has even degrees in $V \setminus V_1$ because it was obtained from a spanning tree on V_1 through replacing edges by walks connecting the same endpoints. Consequently, we are left with proving the claim.

Proof of Item (i). The first sum in the definition of $g(z)$ is linear in z, hence it suffices to show that $\bar{g} \colon P_{\mathrm{ST}}(G_0) \to \mathbb{R}$ defined by

$$\bar{g}(z) := \sum_{v \in V \setminus V_1} \pi_v \prod_{W \in \mathcal{W}_0 \colon v \in W} (1 - z_{e_W})$$

is concave under swaps. We show that $\bar{g}$ is in fact equal to (a restriction of) the multilinear extension of a supermodular set function $f \colon 2^{\mathcal{W}_0} \to \mathbb{R}$, which readily implies that $\bar{g}$ is concave under swaps [CCPV11]. For $S \subseteq \mathcal{W}_0$, let $V[S]$ denote the union of all vertices covered by the walks in S, and define the set function $f \colon 2^{\mathcal{W}_0} \to \mathbb{R}$ by

$$f(S) := \pi(V \setminus V[S]) = \sum_{v \in V} \pi_v \cdot \left(1 - \mathbb{1}_{v \in V[S]} \right) \ .$$

Supermodularity follows immediately from the sum representation above, because the indicator function $\mathbb{1}_{v \in V[S]}$ is clearly submodular. By definition, the multilinear extension $F \colon [0,1]^{\mathcal{W}_0} \to \mathbb{R}$ of f is given by

$$F(z) = \sum_{S \subseteq \mathcal{W}_0} f(S) \prod_{W \in S} z_W \prod_{W \in \mathcal{W}_0 \setminus S} (1 - z_W) \ .$$

It remains to show that $F(z) = \bar{g}(z)$. Indeed,

$$
\begin{aligned}
F(z) &= \sum_{S \subseteq \mathcal{W}_0} \pi(V \setminus V[S]) \prod_{W \in S} z_W \prod_{W \in \mathcal{W}_0 \setminus S} (1 - z_W) \\
&= \sum_{v \in V \setminus V_1} \pi_v \sum_{S \subseteq \{W \in \mathcal{W}_0 \colon v \notin W\}} \prod_{W \in S} z_W \prod_{W \in \mathcal{W}_0 \setminus S} (1 - z_W) \\
&= \sum_{v \in V \setminus V_1} \pi_v \prod_{W \in \mathcal{W}_0 \colon v \in W} (1 - z_W) \underbrace{\sum_{S \subseteq \{W \in \mathcal{W}_0 \colon v \notin W\}} \prod_{W \in S} z_W \prod_{W \in \{W \in \mathcal{W}_0 \colon v \notin W\} \setminus S} (1 - z_W)}_{=1} = \bar{g}(z) \ .
\end{aligned}
$$

Proof of Item (ii). If z is integral and corresponds to a spanning tree T of G_0, we may write

$$g(z) = \sum_{e_W \in T} c(E[W]) + \sum_{v \in V \setminus V_1} \pi_v \prod_{e_W \in E[G_0] \colon v \in W} \mathbb{1}_{e_W \notin E[T]} \ .$$

The first term equals $c(E[H])$ by definition of H. In the second term, the sum has nonzero terms only for $v \in V \setminus V_1$ that are not covered by any walk W for all $e_W \in E[T]$, i.e., precisely for the vertices not covered by H. As H always covers V_1, the second term thus equals $\pi(V \setminus V[H])$, as desired.

Proof of Item (iii). By definition of the walks in $\mathcal{W}_0$, we have

$$\sum_{W \in \mathcal{W}_0} c(E[W]) z_0(e_W) = \sum_{T \in \mathcal{T}} \left(c(E[W_1(T)]) \cdot \frac{3}{4} \mu_T + c(E[W_2(T)]) \cdot \frac{1}{4} \mu_T \right)$$

$$= \sum_{T \in \mathcal{T}} \mu_T \cdot \left(c(\text{backbone}_{V_1}(T)) + \frac{3}{2} \cdot c(\text{limbs}_{V_1}(T)) \right) \ ,$$

and

$$\sum_{v \in V \setminus V_1} \pi_v \prod_{W \in \mathcal{W}_0 \colon v \in W} (1 - z_0(e_W)) \leq \sum_{v \in V \setminus V_1} \pi_v \prod_{T \in \mathcal{T} \colon v \in V[T]} \left(1 - \frac{3}{4} \mu_T \right)$$

$$\leq \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3}{4} \sum_{T \in \mathcal{T} \colon r \in V[T]} \mu_T \right)$$

$$= \sum_{v \in V \setminus V_1} \pi_v \exp\left(-\frac{3}{4} y_v \right) \ ,$$

which together gives the claimed relation, finishes the proof of Claim 20, and thus also of Lemma 19. $\quad\square$

5.3 Deterministic selection of a threshold

In the analysis of Algorithm 1 in Theorem 10, we also exploited a random choice of the threshold parameter γ. We now show that if we start from an optimal solution (x^*, y^*) of the PCTSP LP relaxation, it suffices to try all values in the set $\{y_v^* \colon v \in V\}$ as thresholds.

Lemma 21. *Let (x^*, y^*) be an optimal solution of the PCTSP LP relaxation, and let $\alpha > 0$. If there exists $\gamma \in (0, 1]$ such that*

$$\frac{3}{2\gamma} \cdot c^\top x^* + \sum_{v \in V \colon y_v^* < \gamma} \exp\left(-\frac{3}{4\gamma} y_v^* \right) \cdot \pi_v \leq \alpha \cdot \left(c^\top x^* + \pi^\top (1 - y^*) \right) \ , \tag{7}$$

then there exists a vertex $v \in V$ such that the above also holds for $\gamma = y_v^*$.

Proof of Lemma 21. Fix $\gamma \in (0, 1]$ such that (7) holds. We claim that with $y := \min\{y_v^* \colon y_v^* \geq \gamma\}$, we have

$$\frac{3}{2y} \cdot c^\top x^* + \sum_{v \in V \colon y_v^* < y} \exp\left(-\frac{3}{4y} y_v^* \right) \cdot \pi_v \leq \frac{3}{2\gamma} \cdot c^\top x^* + \sum_{v \in V \colon y_v^* < \gamma} \exp\left(-\frac{3}{4\gamma} y_v^* \right) \cdot \pi_v \ , \tag{8}$$

which obviously implies the lemma. For the sake of deriving a contradiction, assume that (8) does not hold, and note that because the sums in (8) are over the same subset of vertices by definition of y, this assumption can be rewritten as

$$\frac{3}{2} \left(\frac{1}{\gamma} \cdot c^\top x^* - \frac{1}{y} \cdot c^\top x^* \right) < \sum_{v \in V \colon y_v^* < y} \left(\exp\left(-\frac{3 y_v^*}{4y} \right) - \exp\left(-\frac{3 y_v^*}{4\gamma} \right) \right) \cdot \pi_v \ .$$

As $x \mapsto \exp(-x) + x$ is increasing for $x \geq 0$, we have $\exp(-a) - \exp(-b) \leq b - a$ whenever $b \geq a \geq 0$, hence

$$\frac{3}{2}\left(\frac{1}{\gamma}\cdot c^\top x^* - \frac{1}{y}\cdot c^\top x^*\right) < \sum_{v\in V:\, y_v^* < y} \frac{3}{4}\left(\frac{y_v^*}{\gamma} - \frac{y_v^*}{y}\right)\pi_v = \frac{3}{4}\left(\pi^\top(1 - y_{1/y}^*) - \pi^\top(1 - y_{1/\gamma}^*)\right) \ ,$$

where we use the notation y_λ^* of Lemma 3 for $\lambda = 1/y$ and $\lambda = 1/\gamma$. Through scaling down the non-negative left-hand side by $\frac{1}{2}$ and rearranging terms, this further implies

$$\frac{1}{\gamma}\cdot c^\top x^* + \pi^\top(1 - y_{1/\gamma}^*) < \frac{1}{y}\cdot c^\top x^* + \pi^\top(1 - y_{1/y}^*) \ . \tag{9}$$

Now consider the function $f\colon \mathbb{R}_{\geq 1} \to \mathbb{R}$ defined by $\lambda \mapsto \lambda\cdot c^\top x^* + \pi^\top(1 - y_\lambda^*)$. By definition, f is convex and continuous. Moreover, (9) shows that f is not increasing. Thus, f must be decreasing at the left endpoint of its domain, i.e., at $\lambda = 1$, implying that there exists $\lambda_0 > 1$ such that $f(\lambda_0) < f(1)$. But $f(\lambda_0)$ is an upper bound on the objective value of the solution $(x_{\lambda_0}, y_{\lambda_0})$ of the PCTSP LP relaxation that can be constructed through Lemma 3, while $f(1)$ is the objective value of (x^*, y^*), hence $f(\lambda_0) < f(1)$ contradicts optimality of (x^*, y^*). This finishes the proof of Lemma 21. $\qquad\square$

6 Proof of Theorem 10

Proof of Theorem 10. The concrete choice of b is left to the end of this proof; with foresight, we only require $b \in [1 - e^{-3/4}, 1)$ for now. Let $I_b := \int_b^1 \exp(-b/\gamma)\,d\gamma$. Then $f(\gamma) = \exp(-b/\gamma)/I_b$. By Theorem 8, the expected edge cost of the returned cycle $C = (V_C, E_C)$ is at most

$$\mathbb{E}[c(E_C)] \leq \frac{3}{2}\int_b^1 \frac{f(\gamma)}{\gamma}\,d\gamma\cdot c^\top x^* = \frac{3}{2I_b}\int_b^1 \frac{\exp(-b/\gamma)}{\gamma}\,d\gamma\cdot c^\top x^* \ . \tag{10}$$

Moreover, the expected penalty cost can be bounded by

$$\mathbb{E}[\pi(V\setminus V_C)] \leq \sum_{v\in V:\, y_v^* < 1} \pi_v(1 - y_v^*)\cdot\frac{1}{1 - y_v^*}\int_{\max\{y_v^*, b\}}^1 \exp\left(-\frac{3y_v^*}{4\gamma}\right)f(\gamma)\,d\gamma$$

$$= \sum_{v\in V:\, y_v^* < 1} \pi_v(1 - y_v^*)\cdot\underbrace{\frac{1}{I_b(1 - y_v^*)}\cdot\int_{\max\{y_v^*, b\}}^1 \exp\left(-\frac{3y_v^* + 4b}{4\gamma}\right)d\gamma}_{=:f_b(y_v^*)} \ .$$

We investigate for which $y \in [0, 1)$ the function $y \mapsto f_b(y)$ attains its maximum value. If $y \in [0, b]$, convexity of the map $y \mapsto g_\gamma(y) := \frac{1}{1-y}\exp\left(-\frac{3y}{4\gamma}\right)$ implies that $g_\gamma(y) \leq \max\{g_\gamma(0), g_\gamma(b)\}$. Since $b \geq 1 - e^{-3/4}$ and because $\gamma \geq b$, the latter maximum is equal to $g_\gamma(b)$, and we can conclude that $f_b(y)$ attains its maximum for some $y \in [b, 1)$, where

$$f_b(y) = \frac{1}{I_b(1 - y)}\cdot\int_y^1 \exp\left(-\frac{3y + 4b}{4\gamma}\right)d\gamma \ .$$

Computational experiments suggest that $f_b(y)$ attains its maximum at $y = b$, but it seems hard to prove this analytically. For this reason, we aim for obtaining good bounds on the maximum value in what follows. Observe that the integrand $h_{b,y}(\gamma) := \exp\left(-\frac{3y+4b}{4\gamma}\right)$ is concave over $\gamma \in [y, 1]$. Hence,

$$\int_y^1 h_{b,y}(\gamma)\,d\gamma = \int_y^{\frac{y+1}{2}} h_{b,y}(\gamma)\,d\gamma + \int_{\frac{y+1}{2}}^1 h_{b,y}(\gamma)\,d\gamma \leq \frac{1 - y}{2}\cdot\left(h_{b,y}\left(\tfrac{3y+1}{4}\right) + h_{b,y}\left(\tfrac{1+3y}{4}\right)\right) \ .$$

Altogether, we therefore get that

$$\max_{y \in [0,1)} f_b(y) = \max_{y \in [b,1)} f_b(y)$$

$$= \max_{y \in [b,1)} \frac{1}{2I_b} \left(h_{b,y}\left(\tfrac{3y+1}{4}\right) + h_{b,y}\left(\tfrac{1+3y}{4}\right) \right)$$

$$= \frac{1}{2I_b} \cdot \max_{y \in [b,1)} \left(\underbrace{\exp\left(-\frac{3y+4b}{3y+1}\right) + \exp\left(-\frac{3y+4b}{y+3}\right)}_{=: \theta_b(y)} \right) \ .$$

Together with (10), this implies that we get in polynomial time a PCTSP solution of expected value at most $\alpha \cdot \left(c^\top x^* + \pi^\top (1 - y^*)\right)$, where

$$\alpha := \frac{1}{2I_b} \cdot \max \left\{ 3 \cdot \int_b^1 \frac{\exp(-b/\gamma)}{\gamma} \, d\gamma, \ \max_{y \in [b,1)} \theta_b(y) \right\} \ .$$

We numerically evaluate the latter expression and obtain $\alpha < 1.774$ for $b = 0.6945$, as claimed in Theorem 10. More concretely, to bound $\max_{y \in [b,1)} \theta_b(y)$, we note that the derivative of $\theta_b(y)$ on the interval $[b, 1]$ can be bounded by a constant, hence the maximum can be approximated up to a prescribed error by evaluating $\theta_b(y)$ on a sufficiently fine discretization of possible values of y. $\qquad\square$

Acknowledgments

References

[ABHK11]   A. Archer, M. Bateni, M. Hajiaghayi, and H. Karloff. "Improved Approximation Algorithms for Prize-Collecting Steiner Tree and TSP". In: *SIAM Journal on Computing* 40.2 (2011), pp. 309–332. DOI: 10.1137/090771429 (cit. on pp. 2, 3).

[ABLM07]   G. Ausiello, V. Bonifaci, S. Leonardi, and A. Marchetti-Spaccamela. "Prize-Collecting Traveling Salesman and Related Problems". In: *Handbook of Approximation Algorithms and Metaheuristics*. Chapman and Hall, 2007. Chap. 40, pp. 1–13. DOI: 10.1201/9781420010749 (cit. on p. 4).

[AGMGS17] A. Asadpour, M. X. Goemans, A. Madry, S. O. Gharan, and A. Saberi. "An $O(\log n / \log \log n)$-Approximation Algorithm for the Asymmetric Travelling Salesman Problem". In: *Operations Research* 65.4 (2017), pp. 1043–1061. DOI: 10.1287/opre.2017.1603 (cit. on p. 16).

[ALM00]   G. Ausiello, S. Leonardi, and A. Marchetti-Spaccamela. "On Salesmen, Repairmen, Spiders, and Other Traveling Agents". In: *Proceedings of the 4th Italian Conference on Algorithms and Complexity*. Vol. 1767. 2000, pp. 1–16. DOI: 10.1007/3-540-46521-9_1 (cit. on p. 4).

[AS04]   A. A. Ageev and M. I. Sviridenko. "Pipage Rounding: A New Method of Constructing Algorithms with Proven Performance Guarantee". In: *Journal of Combinatorial Optimization* 8.3 (2004), pp. 307–328. DOI: 10.1023/B:JOCO.0000038913.96607.c2 (cit. on pp. 2, 16).

[Bal89]   E. Balas. "The prize collecting traveling salesman problem". In: *Networks* 19.6 (1989), pp. 621–636. DOI: [10.1002/net.3230190602](https://doi.org/10.1002/net.3230190602) (cit. on p. 3).

[BCEHKM11]   M. Bateni, C. Chekuri, A. Ene, M. Hajiaghayi, N. Korula, and D. Marx. "Prize-collecting Steiner Problems on Planar Graphs". In: *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '11)*. 2011, pp. 1028–1049. DOI: [10.1137/1.9781611973082.79](https://doi.org/10.1137/1.9781611973082.79) (cit. on p. 4).

[BCKLMM03]   A. Blum, S. Chawla, D. Karger, T. Lane, A. Meyerson, and M. Minkoff. "Approximation algorithms for orienteering and discounted-reward TSP". In: *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (FOCS '03)*. 2003, pp. 46–55. DOI: [10.1109/SFCS.2003.1238180](https://doi.org/10.1109/SFCS.2003.1238180) (cit. on p. 4).

[BFJ95]   J. Bang-Jensen, A. Frank, and B. Jackson. "Preserving and Increasing Local Edge-Connectivity in Mixed Graphs". In: *SIAM Journal on Discrete Mathematics* 8.2 (1995), pp. 155–178. DOI: [10.1137/S0036142993226983](https://doi.org/10.1137/S0036142993226983) (cit. on pp. 3, 11, 14).

[BGSW93]   D. Bienstock, M. X. Goemans, D. Simchi-Levi, and D. Williamson. "A note on the prize collecting traveling salesman problem". In: *Mathematical Programming* 59 (1993), pp. 413–420. DOI: [10.1007/BF01581256](https://doi.org/10.1007/BF01581256) (cit. on pp. 1, 2, 5).

[CCPV11]   G. Calinescu, C. Chekuri, M. Pál, and J. Vondrák. "Maximizing a Monotone Submodular Function Subject to a Matroid Constraint". In: *SIAM Journal on Computing* 40.6 (2011), pp. 1740–1766. DOI: [10.1137/080733991](https://doi.org/10.1137/080733991) (cit. on pp. 2, 16, 19).

[CDV08]   D. Chakrabarty, N. R. Devanur, and V. V. Vazirani. "New Geometry-Inspired Relaxations and Algorithms for the Metric Steiner Tree Problem". In: *Integer Programming and Combinatorial Optimization (IPCO '08)*. 2008, pp. 344–358. DOI: [10.1007/978-3-540-68891-4_24](https://doi.org/10.1007/978-3-540-68891-4_24) (cit. on p. 3).

[Chr22]   N. Christofides. "Worst-Case Analysis of a New Heuristic for the Travelling Salesman Problem". In: *Operations Research Forum* 3 (2022). DOI: [10.1007/s43069-021-00101-z](https://doi.org/10.1007/s43069-021-00101-z) (cit. on p. 1).

[Chr76]   N. Christofides. *Worst-case analysis of a new heuristic for the Travelling Salesman Problem*. Technical Report 388. Graduate School of Industrial Administration, Carnegie Mellon University, 1976 (cit. on p. 1).

[CJJ20]   T.-H. H. Chan, H. Jiang, and S. H.-C. Jiang. "A Unified PTAS for Prize Collecting TSP and Steiner Tree Problem in Doubling Metrics". In: *ACM Transactions on Algorithms* 16.2 (2020), pp. 1–23. DOI: [10.1145/3378571](https://doi.org/10.1145/3378571) (cit. on p. 4).

[CKP12]   C. Chekuri, N. Korula, and M. Pál. "Improved Algorithms for Orienteering and Related Problems". In: *ACM Transactions on Algorithms* 8.3 (2012), pp. 1–27. DOI: [10.1145/2229163.2229167](https://doi.org/10.1145/2229163.2229167) (cit. on p. 4).

[CVZ10]   C. Chekuri, J. Vondrak, and R. Zenklusen. "Dependent Randomized Rounding via Exchange Properties of Combinatorial Structures". In: *Proceedings of the 51st IEEE Annual Symposium on Foundations of Computer Science (FOCS '10)*. 2010, pp. 575–584. DOI: [10.1109/FOCS.2010.60](https://doi.org/10.1109/FOCS.2010.60) (cit. on p. 16).

[DFPS22]   S. Dezfuli, Z. Friggstad, I. Post, and C. Swamy. "Combinatorial Algorithms for Rooted Prize-Collecting Walks and Applications to Orienteering and Minimum-Latency Problems". In: *Integer Programming and Combinatorial Optimization (IPCO '22)*. 2022, pp. 195–209. DOI: [10.1007/978-3-031-06901-7_15](https://doi.org/10.1007/978-3-031-06901-7_15) (cit. on p. 4).

[Edm73]   J. Edmonds. "Edge-disjoint branchings". In: *Combinatorial algorithms*. Academic Press, 1973, pp. 91–96 (cit. on p. 3).

[Fra92]   A. Frank. "On a theorem of Mader". In: *Discrete Mathematics* 101.1 (1992), pp. 49–57. DOI: 10.1016/0012-365X(92)90589-8 (cit. on p. 13).

[Gar05]   N. Garg. "Saving an Epsilon: A 2-Approximation for the k-MST Problem in Graphs". In: *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC '05)*. 2005, pp. 396–402. DOI: 10.1145/1060590.1060650 (cit. on p. 4).

[GB93]   M. X. Goemans and D. J. Bertsimas. "Survivable networks, linear programming relaxations and the parsimonious property". In: *Mathematical Programming* 60 (1993), pp. 145–166. DOI: 10.1007/BF01580607 (cit. on p. 5).

[Goe09]   M. X. Goemans. *Combining Approximation Algorithms for the Prize-Collecting TSP*. 2009. arXiv: 0910.0553 [cs.DS] (cit. on pp. 2, 4, 9, 10).

[GW17]   K. Genova and D. P. Williamson. "An Experimental Evaluation of the Best-of-Many Christofides' Algorithm for the Traveling Salesman Problem". In: *Algorithmica* 78 (2017), pp. 1109–1130. DOI: 10.1007/s00453-017-0293-5 (cit. on p. 3).

[GW95]   M. X. Goemans and D. P. Williamson. "A General Approximation Technique for Constrained Forest Problems". In: *SIAM Journal on Computing* 24.2 (1995), pp. 296–317. DOI: 10.1137/S0097539793242618 (cit. on pp. 2, 3, 10, 11).

[HJ06]   M. T. Hajiaghayi and K. Jain. "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 Algorithm (SODA '06)*. 2006, pp. 631–640. DOI: 10.1145/1109557.1109626 (cit. on p. 4).

[HO14]   N. J. A. Harvey and N. Olver. "Pipage Rounding, Pessimistic Estimators and Matrix Concentration". In: *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '14)*. 2014, pp. 926–945. DOI: 10.1137/1.9781611973402.69 (cit. on pp. 16, 17).

[KKO21]   A. R. Karlin, N. Klein, and S. Oveis Gharan. "A (Slightly) Improved Approximation Algorithm for Metric TSP". In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing (STOC '21)*. 2021, pp. 32–45. DOI: 10.1145/3406325.3451009 (cit. on pp. 1, 6, 8, 16).

[KKO22]   A. R. Karlin, N. Klein, and S. Oveis Gharan. *A (Slightly) Improved Deterministic Approximation Algorithm for Metric TSP*. 2022. arXiv: 2212.06296 [cs.DS] (cit. on pp. 1, 8).

[KOPRSV17]   J. Könemann, N. Olver, K. Pashkovich, R. Ravi, C. Swamy, and J. Vygen. "On the Integrality Gap of the Prize-Collecting Steiner Forest LP". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM '17)*. Vol. 81. 2017, 17:1–17:13. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2017.17 (cit. on p. 4).

[Lov76]   L. Lovász. "On some connectivity properties of Eulerian graphs". In: *Acta Mathematica Academiae Scientiarum Hungarica* 28.1 (1976), pp. 129–138. DOI: 10.1007/BF01902503 (cit. on p. 13).

[Mad78]   W. Mader. "A Reduction Method for Edge-Connectivity in Graphs". In: *Annals of Discrete Mathematics* 3 (1978), pp. 145–164. DOI: 10.1016/S0167-5060(08)70504-1 (cit. on p. 13).

[Nag06]   H. Nagamochi. "A Fast Edge-Splitting Algorithm in Edge-Weighted Graphs". In: *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences* E89-A.5 (2006), pp. 1263–1268. DOI: 10.1093/ietfec/e89-a.5.1263 (cit. on p. 13).

[Ngu13]   V. H. Nguyen. "A Primal-Dual Approximation Algorithm for the Asymmetric Prize-Collecting TSP". In: *Journal of Combinatorial Optimization* 25 (2013), pp. 265–278. DOI: 10.1007/s10878-012-9501-z (cit. on p. 4).

[NI97]   H. Nagamochi and T. Ibaraki. "Deterministic $\tilde{O}(nm)$ Time Edge-Splitting in Undirected Graphs". In: *Journal of Combinatorial Optimization* 1.6 (1997), pp. 5–46. DOI: 10.1023/A:1009739202898 (cit. on p. 13).

[PFFSW20]   A. Paul, D. Freund, A. Ferber, D. B. Shmoys, and D. P. Williamson. "Budgeted Prize-Collecting Traveling Salesman and Minimum Spanning Tree Problems". In: *Mathematics of Operations Research* 45.2 (2020), pp. 576–590. DOI: 10.1287/moor.2019.1002 (cit. on p. 4).

[RV99]   S. Rajagopalan and V. V. Vazirani. "On the Bidirected Cut Relaxation for the Metric Steiner Tree Problem". In: *Proceedings of the 10th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '99)*. 1999, pp. 742–751. DOI: 10.5555/314500.314909 (cit. on p. 3).

[Ser87]   A. I. Serdyukov. "O nekotorykh ekstremal'nykh obkhodakh v grafakh". In: *Upravlyaemye sistemy* 17 (1987), pp. 76–79 (cit. on p. 1).

[Sri01]   A. Srinivasan. "Distributions on level-sets with applications to approximation algorithms". In: *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science (FOCS '01)*. 2001, pp. 588–597. DOI: 10.1109/SFCS.2001.959935 (cit. on pp. 2, 16).

[vS20]   R. van Bevern and V. A. Slugina. "A historical note on the 3/2-approximation algorithm for the metric traveling salesman problem". In: *Historia Mathematica* 53 (2020), pp. 118–127. DOI: 10.1016/j.hm.2020.04.003 (cit. on p. 1).

[Wol80]   L. A. Wolsey. "Heuristic analysis, linear programming and branch and bound". In: *Combinatorial Optimization II*. Springer Berlin Heidelberg, 1980, pp. 121–134. DOI: 10.1007/BFb0120913 (cit. on p. 5).