An Improved Approximation Algorithm for the Minimum k-Edge Connected Multi-subgraph Problem

Anna R. Karlin
University of Washington
USA
karlin@cs.washington.edu

Nathan Klein
University of Washington
USA
nwklein@cs.washington.edu

Shayan Oveis Gharan
University of Washington
USA
shayan@cs.washington.edu

Xinzhi Zhang
University of Washington
USA
xinzhi20@cs.washington.edu

ABSTRACT

We give a randomized $1 + \frac{5.06}{\sqrt{k}}$-approximation algorithm for the minimum k-edge connected spanning multi-subgraph problem, k-ECSM.

CCS CONCEPTS

• **Theory of computation → Routing and network design problems**.

KEYWORDS

Approximation Algorithms, Edge Connectivity, Network Design, Randomized Rounding, Max Entropy, Strongly Rayleigh

1 INTRODUCTION

In an instance of the minimum k-edge connected spanning subgraph problem, or k-ECSS, we are given an (undirected) graph $G = (V, E)$ with $n := |V|$ vertices and a cost function $c : E \to \mathbb{R}_{\geq 0}$, and we want to choose a minimum cost set of edges $F \subseteq E$ such that the subgraph (V, F) is k-edge connected. In its most general form, k-ECSS generalizes several extensively-studied problems in network design such as tree augmentation or cactus augmentation, for which there has been recent exciting progress (e.g. [5, 7, 11, 28]). The k-edge-connected *multi*-subgraph problem, k-ECSM, is a close variant of k-ECSS in which we want to choose a k-edge-connected *multi*-subgraph of G of minimum cost, i.e., we can choose an edge $e \in E$ multiple times. Note that without loss of generality we can

assume the cost function c in k-ECSM is a metric, i.e., for any three vertices $x, y, z \in V$, we have $c(x, z) \leq c(x, y) + c(y, z)$.

Around four decades ago, Fredrickson and Jájá [12, 13] designed a 2-approximation algorithm for k-ECSS and a 3/2-approximation algorithm for k-ECSM. The latter essentially follows by a reduction to the well-known Christofides-Serdyukov approximation algorithm for the traveling salesperson problem (TSP). Over the last four decades, despite a number of papers on the problem [9, 14–16, 19, 22, 23, 26], the aforementioned approximation factors were only improved in the cases where the underlying graph is unweighted or $k \gg \log n$. Most notably, Gabow, Goemans, Tardos and Williamson [16] showed that if the graph G is unweighted then k-ECSS and k-ECSM admit $1 + 2/k$ approximation algorithms, i.e., as $k \to \infty$ the approximation factor approaches 1. The case of k-ECSM where $k = 2$ has received significant attention and (significantly) better than 3/2-approximation algorithms were designed for special cases [3, 4, 6, 27]. In the general $k = 2$ case, only a $3/2 - \epsilon$ approximation is known where $\epsilon = 10^{-36}$ [20]; we remark this also extends to all even k.

Motivated by [16], Pritchard posed the following conjecture:

Conjecture 1.1 ([26]). *The k-ECSM problem admits a $1 + O(1)/k$ approximation algorithm.*

In other words, if true, the above conjecture implies that the 3/2-classical factor can be substantially improved for large k, and moreover that it is possible to design an approximation algorithm whose factor gets arbitrarily close to 1 as $k \to \infty$. In this paper, we prove a weaker version of the above conjecture.

Theorem 1.2 (Main). *There is a polynomial time randomized algorithm for (weighted) k-ECSM with approximation factor (at most) $1 + \frac{5.06}{\sqrt{k}}$.*

We remark that our main theorem only improves the classical 3/2-approximation algorithm for k-ECSM when $k > 103$. However, the constants are not optimized and we expect our algorithm to beat 3/2 for much smaller values of k.

For a set $S \subseteq V$, let $\delta(S) = \{\{u, v\} : |\{u, v\} \cap S| = 1\}$ denote the set of edges with one endpoint in S. The following is the natural

linear programming relaxation for k-ECSM.

$$
\begin{aligned}
\min \quad & \sum_{e \in E} x_e c(e) \\
\text{s.t.} \quad & x(\delta(v)) = k && \forall v \in V \\
& x(\delta(S)) \geq k && \forall S \subseteq V, S \neq \emptyset \\
& x_e \geq 0 && \forall e \in E.
\end{aligned} \tag{1}
$$

Note that while in an optimum solution of k-ECSM the degree of each vertex is not necessarily equal to k, since the cost function satisfies the triangle inequality we may assume that in any optimum fractional solution each vertex has (fractional) degree k. This follows from the parsimonious property [17].

We prove Theorem 1.2 by rounding an optimum solution to the above linear program. So, as a corollary we also upper-bound the integrality gap of the above linear program.

Corollary 1.3. *The integrality gap of LP* (1) *is at most* $1 + \frac{5.06}{\sqrt{k}}$.

2 PROOF OVERVIEW

Before explaining our algorithm, we recall a randomized rounding approach of Karger [19]. Karger showed that if given a solution x to (1) we choose every edge e independently with probability x_e, then the sample is $k - O(\sqrt{k \log n})$-edge connected with probability close to 1. He then fixes the connectivity of the sample by adding $O(\sqrt{k \log n})$ copies of the minimum spanning tree of G. This gives a randomized $1 + O(\sqrt{\log n / k})$ approximation algorithm for the problem. While this is a very effective procedure for large k, it is not useful when k is a constant or grows slower than $\log n$. We view our result as a refinement of this method using random spanning trees which allows k to be independent of n.

First, we observe that when x is a solution to (1), the vector $2x/k$ is in the spanning tree polytope (after modifying x slightly, see Fact 3.3 for more details). Following a recent line of works on the traveling salesperson problem [21, 24] we write $2x/k$ as a so-called *max-entropy distribution* μ over spanning trees.

Warm-up algorithm and key idea. Our first algorithm, explained in Section 4, independently samples $k/2$ spanning trees $T_1, \ldots, T_{k/2}$ from μ. Call the (multi-set) union of these trees T^*. Since max entropy distributions are negatively correlated, it is easy to show using Chernoff bounds that any particular cut S has at least $k - O(\sqrt{k \ln k})$ edges with probability at least $1 - O(1/\sqrt{k})$[1]. So, in the second step of the algorithm, we add $O(\sqrt{k \ln k})$ additional spanning trees to fix the connectivity of every cut "with high probability." In other words, after this procedure (which has expected cost $1 + O(\sqrt{\ln k / k})$ times the cost of the LP), *every cut S has at least k edges with probability* $1 - O(1/\sqrt{k})$. One can think of this as a version of Karger's algorithm which does not fix *every cut* with high probability but instead fixes *each individual cut* with high probability.

A priori, this does not seem like a useful property, because there are exponentially many cuts to bound over. However, we show that

(perhaps somewhat surprisingly) there is a way to fix the connectivity of every cut simultaneously by only paying an additional factor of $O(1/k)$ times the cost of the LP in expectation.

To do so, we begin with the following simple observation: Fix a cut S. Then, if we ensure that every tree T_i has at least 2 edges in $\delta(S)$, the union of the trees T_i will have at least k edges across the cut and we are done. So, if a cut S turns out to have fewer than $k - O(\sqrt{k \ln k})$ edges in T^*, one can think of "blaming" the trees which had only one edge in $\delta(S)$; in particular, we will fix the cut by doubling the sole edge in $\delta(S)$ for each of those trees. This guarantees that every tree has at least two edges across the cut and therefore it has at least k edges total as desired. This is essentially the key idea of this paper.

Formally, after sampling $T_1, \ldots, T_{k/2}$, iterate over every edge e of every tree T_i and consider the unique cut S in which e appears as the only edge of T_i. This is the *only* cut for which e may be "blamed" and hence doubled. Now, we check if T^* has fewer than $k - O(\sqrt{k \ln k})$ edges across $\delta(S)$. Over the randomness of the remaining trees, this occurs with probability $O(1/\sqrt{k})$ (by Chernoff bounds, as argued above). This shows that every edge of T^* is doubled with probability $O(1/\sqrt{k})$, and therefore since T^* has expected cost at most $c(x) \leq OPT$, the approximation ratio of the algorithm is $1 + O(\sqrt{\ln k / k}) + O(1/\sqrt{k})$, i.e. $1 + O(\sqrt{\ln k / k})$, as desired.

Main algorithm. The above algorithm is suboptimal in a fairly obvious way. Suppose that a cut S is missing (for example) $\sqrt{k \ln^2 k}$ edges. Then, its connectivity is not fixed by the additional $O(\sqrt{k \ln k})$ spanning trees added by the algorithm. So, the warm-up algorithm simply adds an additional copy of every edge which appeared alone in this cut in its tree. However, it may be that the cut $\delta(S)$ has only one edge in as many as $\Omega(k)$ trees! Therefore, we will add $\Omega(k)$ extra edges to fix the cut instead of just the required $\sqrt{k \ln^2 k}$ edges: a huge overcorrection. Algorithm 2 simply avoids this overcorrection by only adding the number of edges actually missing from the cut, sampling them independently from the set of edges which appeared alone on this cut. It turns out this will let us add only $O(\sqrt{k})$ additional trees instead of $O(\sqrt{k \ln k})$, avoiding the extra $\sqrt{\ln k}$ factor[2].

However, this adds some difficulty to the analysis. For example, now that we only add $O(\sqrt{k})$ extra trees, it is not true that a cut only has to be fixed with probability $O(1/\sqrt{k})$. In fact, this probability may even be $O(1)$. To sharpen the analysis, for any fixed set S we study p_S, the probability that a random (max entropy) tree has exactly one edge in $\delta(S)$. In particular, we show that the expected number of edges that $\delta(S)$ is missing (below k) is at most $O(\sqrt{k} e^{-1/p_S})$.

This bound on the expectation is then enough to complete the argument as follows. Let $n(S)$ be the number of trees T_i for which $|T_i \cap \delta(S)| = 1$. Let e be the unique edge in $\delta(S)$ for some T_i. Then, the probability e needs to double is the expected number of edges missing from $\delta(S)$ divided by $n(S)$. Using the analysis above and

---

[1]Of course, one can make this probability much closer to 1 (say $1 - O(1/k^2)$) by only paying a constant factor in the $O(\sqrt{k \ln k})$ term, but it is sufficient to make it $1 - O(1/\sqrt{k})$.

[2]For a slightly tighter analysis we also include these additional trees in T^*, but this is mostly a superficial difference.

that $\mathbb{E}[n(S)] = \Omega(kp_S)$, this can be shown to be $O(\frac{\sqrt{k}e^{-1/p_S}}{kp_S}) = O(1/\sqrt{k})$ in expectation[3].

Remark 2.1. We note that we expect this algorithm to work for any distribution of spanning trees which is negatively correlated. So, one could for example apply swap rounding [8] to generate random spanning trees (instead of using the max entropy distribution).

However, while the analysis giving $1 + O(\sqrt{\frac{\ln k}{k}})$ approximation in Section 4 can easily be modified to give similar bounds for any negatively correlated distribution (since Chernoff bounds can be applied), the proof of Theorem 1.2 in Section 5 currently relies on the fact that the distribution of the number of edges in any cut can be written as a sum of independent Bernoullis. So, an extension of Theorem 1.2 to an arbitrary negatively correlated distribution would require a different analysis technique or a generalization of Lemma 5.2.

We also briefly remark that some form of concentration is necessary. In particular, consider a distribution over spanning trees in which a vertex v has degree 1 with probability $1 - 1/(n-2)$ and degree $n-1$ with probability $1/(n-2)$. In such a case, we expect to need to add $k/2$ edges from $\delta(v)$ to ensure v has degree at least k. If these edges have all the cost of the LP (or there are many vertices with this property), we can get an approximation ratio as bad as $3/2$ even for large k.

3 PRELIMINARIES

Definition 3.1 (G^0, u_0, v_0). We expand the graph $G = (V, E)$ to a graph G^0 by picking an arbitrary vertex $u \in V$, splitting it into two nodes u_0 and v_0, and then, for every edge $e = (u, w)$ incident to u, assigning fraction $\frac{x(e)}{2}$ to each of the two edges (u_0, w) and (v_0, w) in G^0. We set $x((u^0, v^0)) = 0$. Call this expanded graph G^0, its edge set E^0, and the resulting fractional solution x^0, where $x^0(e)$ and $x(e)$ are identical on all other edges. (Note that each of u_0 and v_0 now have fractional degree $k/2$ in x^0.) In Fact 3.3 below, we show that $\frac{2}{k} \cdot x^0$ is in the spanning tree polytope for the graph G^0. For ease of exposition, the algorithm is described as running on G^0 (and spanning trees[4] of G^0), which has the same edge set as G (when u_0 and v_0 are identified).

3.1 Basic Notation

For a subset of vertices $S \subseteq V$, we write $E(S) \subseteq E$ to denote the set of edges in the induced graph of G whose vertex set is S.

For two sets of edges $F, F' \subseteq E$, we write $F \uplus F'$ to denote the multi-set union of F and F' allowing multiple edges. Note that we always have $|F \uplus F'| = |F| + |F'|$.

For set of vertices $S \subseteq V$, let $\delta(S) = \{\{u, v\} : |\{u, v\} \cap S| = 1\}$ denote the set of edges with one endpoint in S and one endpoint in $\bar{S}$.

For any two sets of edges $F, T \subseteq E$, we write

$$F_T := |F \cap T|.$$

We will primarily use this notation to denote the number of edges F has in a spanning tree (or union of spanning trees) T.

Also, for any edge weight function $x : E \to \mathbb{R}$, we write $x(F) := \sum_{e \in F} x(e)$.

Definition 3.2 ($S_T(e)$, the "One-Cut" of e in T). For any spanning tree T on the vertex set V_0, and any edge $e \in T$, let $S_T(e) \subseteq V_0 \setminus \{u_0\}$ be the unique connected component of $T \setminus \{e\}$ which does not contain u_0. We will call this the **one-cut** of e in T.

Particular edges $e \in T$ of interest are those where both $u_0, v_0 \notin S_T(e)$.

Recall that the natural linear programming relaxation for k-ECSM is (1). The solution to this LP can be computed in polynomial time using the ellipsoid method.

For a real-valued random variable X, we write $X^+ = \max(0, X)$ to denote the positive part of X.

3.2 Random Spanning Trees

Edmonds [10] gave the following description for the convex hull of the spanning trees of any graph $G = (V, E)$, known as the *spanning tree polytope*.

$$
\begin{aligned}
z(E) &= |V| - 1 \\
z(E(S)) &\leq |S| - 1 && \forall S \subseteq V \\
z_e &\geq 0 && \forall e \in E.
\end{aligned}
\tag{2}
$$

Edmonds also [10] proved that the extreme point solutions of this polytope are the characteristic vectors of the spanning trees of G.

Fact 3.3 ([21]). Let x be the optimal solution of LP (1) and x^0 its extension to G^0 as described in Definition 3.1. Then $\frac{2}{k} \cdot x^0$ is in the spanning tree polytope (2) of G^0.

Proof. For any set $S \subseteq V(G^0)$, we have

$$x^0(E(S)) = \sum_{v \in S} x^0(\delta(v)) - x^0(E(S))$$

If $u_0, v_0 \notin S$, then $x^0(\delta(v)) = k \forall v \in S$, and $x^0(\delta(S)) = x(\delta(S))$. So $x^0(E(S)) = \frac{k|S| - x(\delta(S))}{2} \leq \frac{k}{2}(|S| - 1)$.

If $u_0 \in S, v_0 \notin S$, then $\sum_{v \in V} x^0(\delta(v)) = k(|S| - 1) + k/2 = k|S| - k/2$ and $x^0(\delta(S)) \geq k/2$, so $x^0(E(S)) \leq \frac{k|S| - k/2 - x^0(\delta(S))}{2} \leq \frac{k}{2}(|S| - 1)$. Similar analysis also holds for the case where $u_0 \notin S, v_i \in S$.

Finally, if $u_0, v_0 \in S$, then $\sum_{v \in V} x^0(\delta(v)) = k(|S| - 2) + 2 \cdot k/2 = k|S| - k$ and $x^0(\delta(S)) \geq k$. Thus, $x^0(E(S)) = \frac{k|S| - k - x^0(\delta(S))}{2} \leq \frac{k}{2}(|S| - 2)$. The claim follows as $x^0(E) = \frac{k|V(G^0)|}{2} = \frac{k}{2}(|V(G^0)| - 1)$. □

Given nonnegative edge weights $\lambda : E \to \mathbb{R}_{\geq 0}$, we say a distribution μ_λ over spanning trees of G is λ-uniform, if for any spanning tree T,

$$\mathbb{P}_{T \sim \mu_\lambda}[T] \propto \prod_{e \in T} \lambda(e).$$

It has been shown that the maximum entropy distribution over spanning trees is a λ-uniform distribution.

Theorem 3.4 ([1]). There is a polynomial-time algorithm that, given a connected graph $G = (V, E)$, and a point $z \in \mathbb{R}^{|E|}$ in the spanning

---

[3]This is not immediate since this is the ratio of the expectations, but we actually need to analyze the expectation of the ratio.

[4]A spanning tree in G^0 is a 1-tree in G, that is, a tree plus an edge.

tree polytope (2) *of* $G = (V, E)$, *returns* $\lambda : E \to \mathbb{R}_{\geq 0}$ *such that the corresponding* λ*-uniform spanning tree distribution* μ_λ *satisfies*

$$\sum_{T \in \mathcal{T} : e \in T} \mu_\lambda(T) \leq (1 + 2^{-n}) z_e, \ \forall e \in E,$$

i.e., the marginals are approximately preserved. In the above $\mathcal{T}$ *is the set of all spanning trees of* G.

3.3 Bernoulli-Sum Random Variables

In this section, we introduce several properties of the Bernoulli-sum random variable.

Definition 3.5 (Bernoulli-Sum Random Variable). *We say* $BS(q)$ *is a* Bernoulli-Sum *random variable if it has the law of a sum of independent Bernoullis, say* $B_1 + B_2 + \cdots + B_t$ *for some* $t \geq 1$, *with* $\mathbb{E}[B_1 + \cdots + B_t] = q$.

[2, 25] showed that the size of any set of edges on any λ-uniform random spanning tree, i.e. for a set $F \subseteq E$, F_T is distributed as a Bernoulli-sum random variable.

Lemma 3.6 (Random variables F_T are Bernoulli Sums [2, 25]). *Given* $G = (V, E)$ *and* $\lambda : E \to \mathbb{R}_{\geq 0}$, *let* μ_λ *be the* λ*-uniform spanning tree distribution of* G. *Let* T *be a sample from* μ_λ. *Then for any fixed* $F \subseteq E$, *the random variable* F_T *is distributed as* $BS(\mathbb{E}[F_T])$.

We start with a fact that comes directly from linearity of expectation and the definition of variance:

Fact 3.7. *If* $X = BS(q_1)$ *and* $Y = BS(q_2)$ *are two independent Bernoulli-sum random variables, then* $\mathbb{E}[X + Y] = q_1 + q_2$ *and* $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.

Theorem 3.8 (Optimize Expected Value of BS Random Variable, [18] Corollary 2.1). *Let* $g : \{0, 1, \cdots, m\} \to \mathbb{R}$ *and* $0 \leq p \leq m$ *for some integer* $m > 0$. *Let* $X_1, \cdots, X_m$ *be* m *independent Bernoulli random variables with success probabilities* $p_1, \cdots, p_m$ *that minimizes (or maximizes)*

$$\mathbb{E}[g(X_1 + \cdots + X_m)]$$

such that $X_1 + \cdots + X_m = BS(p)$. *Then* $p_1, \cdots, p_m \in \{0, x, 1\}$ *for some* $0 < x < 1$.

Corollary 3.9. *For any* $BS(q)$ *with* $q \geq 1$, $\mathbb{P}[BS(q) = 0] \leq 1/e$.

PROOF. Suppose $BS(q) = X_1 + \cdots + X_m$ for some $m \in \mathbb{Z}_+$ where $X_1, \cdots, X_m$ are independent Bernoullis with success probability $p_1, \cdots, p_m$. Let $g(x) = \mathbb{I}[x = 0]$. Then from Theorem 3.8, if we want to maximize $\mathbb{P}[BS(q) = 0] = \mathbb{E}[g(X_1 + \cdots + X_m)]$, then $p_1, \cdots, p_m \in \{0, x, 1\}$ for some $0 < x < 1$. Suppose m_1 of p_i's are 1, m_2 of p_i's are 0, and the rest of the $(m - m_1 - m_2)$ p_i's are $\frac{q - m_1}{m - m_1 - m_2}$. Then we have

$$\begin{aligned}
\Pr[BS(q) = 0] &= \prod_{i=1}^{m} (1 - p_i) \\
&\leq 0^{m_1} \cdot 1^{m_2} \cdot (1 - \frac{q - m_1}{m - m_1 - m_2})^{m - m_1 - m_2} \\
&\leq (1 - \frac{q}{m})^m \leq e^{-q}.
\end{aligned}$$

where maximization is reached when $m_1 = m_2 = 0$ and $m \to +\infty$. Notice that $q \geq 1$, we have $\Pr[BS(q) = 0] \leq 1/e$ as desired. $\quad\square$

Fact 3.10. *Given any* $0 \leq \epsilon < 1$, *let* $p_1 \geq p_2 \geq \ldots p_m$ *be the success probabilities of* $m \geq 2$ *independent Bernoullis such that* $\sum_{i=1}^{m} p_i = 1 + \epsilon$. *Suppose* $p_1 \leq \frac{1}{2}(1 + \epsilon)$. *Then* $\prod_{i=1}^{m}(1 - p_i) \geq \frac{1}{4}(1 - \epsilon)^2$.

PROOF. The first step is to see that $\prod_{i=1}^{m}(1 - p_i)$ is minimized when p_1 is as large as possible, i.e., $p_1 = \frac{1}{2}(1 + \epsilon)$. To see that, say $p_m > 0$ (for some $m > 1$) and observe that for any $0 < \delta \leq p_m$,

$$(1 - (p_1 + \delta))(1 - p_2)\ldots(1 - (p_m - \delta)) \leq \prod_{i=1}^{m}(1 - p_i).$$

Note that this operation does not change the order of p_i's. So, without loss of generality, assume $p_1 = \frac{1}{2}(1 + \epsilon)$. Now, by Weierstrass inequality we have

$$\begin{aligned}
\prod_{i=1}^{m}(1 - p_i) &\geq (1 - p_1)\left(1 - \sum_{i=2}^{m} p_i\right) \\
&= (1 - \frac{1}{2}(1 + \epsilon))(1 - \frac{1}{2}(1 - \epsilon)) \geq \frac{1}{4}(1 - \epsilon)^2
\end{aligned}$$

where the second to last identity uses that $\sum_i p_i = 1 + \epsilon$. $\quad\square$

Theorem 3.11 (Bernstein Inequality for BS Random Variables). *Let* $X = BS(q)$ *be a BS random variable with* $\mathbb{E}[X] = q$ *and* $\mathrm{Var}[X] = \sigma^2$. *Then* $\forall \lambda > 0$ *we have*

$$\mathbb{P}[X \leq q - \lambda] \leq \exp\left(-\frac{\lambda^2}{2(\sigma^2 + \lambda/3)}\right).$$

Theorem 3.12 (Multiplicative Chernoff-Hoeffding Bound for BS Random Variables). *Let* $X = BS(q)$ *be a Bernoulli-Sum random variable. Then, for any* $0 < \epsilon < 1$ *and* $q' \leq q$,

$$\mathbb{P}[X < (1 - \epsilon)q'] \leq e^{-\frac{\epsilon^2 q'}{2}},$$

and for any $\epsilon > 0$, $q' \geq q$,

$$\mathbb{P}[X > (1 + \epsilon)q'] \leq e^{-\frac{\epsilon^2 q'}{2 + \epsilon}}.$$

4 WARM-UP: A SIMPLE ALGORITHM WITH A $1 + O\left(\sqrt{\frac{\ln k}{k}}\right)$-APPROXIMATION RATIO

We first explain a simple algorithm (Algorithm 1) that has a slightly weaker $1 + O\left(\sqrt{\frac{\ln k}{k}}\right)$-approximation ratio. We defer our main algorithm (Algorithm 2) and the proof of our main result (Theorem 1.2) to Section 5.

In the first step of Algorithm 1, we solve (1) on the (slightly) extended graph G^0. Let x^0 to be the optimal solution. By Fact 3.3, $(2/k)x^0$ is in the spanning tree polytope. Then in line 2, we find the λ-uniform spanning tree distribution μ_λ where each edge has marginal probability $(2/k)x_e^0$ (ignoring the 2^{-n} relative errors). This step is guaranteed to be done in polynomial time by Theorem 3.4.

In line 3, we independently sample $k/2$ spanning trees[6] T_1 up to $T_{k/2}$ from μ_λ, and let $T^* = T_1 \uplus \cdots \uplus T_{k/2}$ to be the (multi-set) union of the samples. It follows that T^* satisfies many desirable properties of the λ-uniform spanning tree distribution:

   i) T^* has the same expectation as the LP solution x^0, since the marginal probability of each edge is exactly $x^0(e)$;

---

[6] If k is odd, we sample $\lceil k/2 \rceil$ trees. The bound remains unchanged relative to the analysis we give below as the potential cost of one extra tree is $O(OPT/k)$.

Algorithm 1 An Approximation Algorithm for k-ECSM

1: Let x^0 be an optimum solution of (1) extended to the graph G^0 as described above.
2: Find weights $\lambda : E^0 \to \mathbb{R}_{\geq 0}$ such that for any $e \in E^0$, $\mathbb{P}_{\mu_\lambda}[e] \leq \frac{2}{k}x_e^0(1 + 2^{-n})$. ▷ By Theorem 3.4
3: Sample $k/2$ spanning trees $T_1, \cdots, T_{k/2} \sim \mu_\lambda$ (in G^0) independently and let $T^* \leftarrow T_1 \uplus \cdots \uplus T_{k/2}$.
4: Let B be the disjoint union of an additional $\alpha\sqrt{k/2 - 1}$[5] spanning trees sampled from μ_λ. ▷ $\alpha = \Theta(\sqrt{\ln k})$ is a parameter we choose later.
5: **for** $i \in [\frac{k}{2}]$ and $e \in T_i$ **do**
6:     **if** $\delta(S_{T_i}(e))_{T^*} < k - \alpha\sqrt{k/2 - 1}$ and $(u_0, v_0) \notin \delta(S_{T_i}(e))$ **then**
7:         $F \leftarrow F \uplus \{e\}$.
8:     **end if**
9: **end for**
10: **Return** $T^* \uplus B \uplus F$.

---

ii) For any cut $\delta(S)$ in G, since $\delta(S)_{T^*}$ is distributed as a Bernoulli-sum random variable, Chernoff-type inequalities apply and $\delta(S)_{T^*}$ is highly concentrated around its mean;

iii) Since T^* is the union of $k/2$ trees, for all cuts we have $\delta(S)_{T^*} \geq k/2$. Moreover, if a cut $\delta(S)$ is not a tree cut of any of the $k/2$ trees, then each of the $k/2$ trees must have at least 2 edges crossing it. Therefore, the number of "bad" cuts of T^*, i.e. those with $\delta(S)_{T^*} < k$, is at most $(n-1)k/2$ (with probability 1).

To fix the potentially $O(nk)$ bad cuts, we divide them into two types: (i) Cuts S such that $\delta(S)_{T^*} \geq k - \alpha\sqrt{k/2 - 1}$ and (ii) Cuts S where $\delta(S)_{T^*} < k - \alpha\sqrt{k/2 - 1}$, for some $\alpha = \Theta(\sqrt{\ln k})$. We fix all cuts of type (i) by adding $B = \alpha\sqrt{k/2 - 1}$ additional spanning trees as in line 4 of the algorithm (note one could alternatively add $\alpha\sqrt{k/2 - 1}$ copies of the minimum spanning tree as in Karger's algorithm). To fix cuts S of type (ii), we employ the following procedure: for any tree T_i where $\delta(S)_{T_i} = 1$ and S is of type (ii), we add one extra copy of the unique edge of T_i in $\delta(S)$. This procedure is in line 5 to line 9 of the algorithm. Let F be the set of edges added in this step; then the output of our algorithm is $T^* \uplus B \uplus F$ as in line 10.

Now we analyze Algorithm 1.

Theorem 4.1 (Approximation Ratio for Algorithm 1). *Algorithm 1 outputs a (weighted) k-ECSM with approximation factor (at most) $1 + \sqrt{\frac{8\ln k}{k}}$.*

We begin by showing that the output of Algorithm 1 is k-edge connected (in G) with probability 1.

Lemma 4.2 (k-Connectivity of the Output). *For any $\alpha \geq 0$, the output of Algorithm 1, $F \uplus B \uplus T^*$ is a k-edge connected subgraph of G.*

PROOF. Fix spanning trees $T_1, \cdots, T_{k/2}$ in G^0 and a cut S where $(u_0, v_0) \notin \delta(S)$. We show that $\delta(S)_{T^* \uplus F \uplus B} \geq k$. If $\delta(S)_{T^*} \geq k - \alpha\sqrt{k/2 - 1}$, then since B has $\alpha\sqrt{k/2 - 1}$ copies of the minimum spanning tree, $\delta(S)_{T^* \uplus B} \geq k$ and we are done. Otherwise $\delta(S)_{T^*} <$

$k - \alpha\sqrt{k/2 - 1}$. Then, we know that for any tree T_i, if $\delta(S)_{T_i} = 1$, since $(u_0, v_0) \notin \delta(S)_{T_i}$, F has one extra copy of the unique edge of T_i in $\delta(S)$. Therefore, including those cases where an extra copy of the edge e is added, each T_i has at least two edges in $\delta(S)$, so $\delta(S)_{T^* \uplus F} \geq 2 \cdot \frac{k}{2} \geq k$ as desired since there are $\frac{k}{2}$ spanning trees T_i. □

To bound the expected cost of our rounded solution, we use the concentration property of λ-uniform trees on edges of T^* to show the probability that any fixed cut $\delta(s)$ is in type (ii), i.e. $\delta(S) < k - \alpha\sqrt{k/2 - 1}$, is exponentially small in α, i.e. $\leq e^{-\alpha^2/2}$, even if we condition on $\delta(S)_{T_i} = 1$ for a single tree T_i.

In our algorithm we sample $k/2$ trees $T_1, \ldots, T_{k/2}$. The following definition will be useful in this section as well as in Section 5. Note it is important to separate the case in which $(u_0, v_0) \in \delta(S)$ for a cut S because in this event, $x^0(\delta(S))$ may be as small as $k/2$, in which case our analysis is not valid. However, since the (u_0, v_0) edge has cost 0, we need not worry about such cuts since they can be trivially satisfied by adding many copies of this edge.

Definition 4.3 ($\mathcal{E}_e^i$). *For a tree T_i sampled in Algorithm 1 and an edge e, we define $\mathcal{E}_e^i$ to be the event that $e \in T_i \wedge (u_0, v_0) \notin \delta(S_{T_i}(e))$.*

Lemma 4.4. *For any $0 \leq \alpha \leq \sqrt{k}$, $1 \leq i \leq k/2$, and any $e \in E$,*

$$\mathbb{P}\left[\delta(S_{T_i}(e))_{T^*} \leq k - \alpha\sqrt{k/2 - 1} \mid \mathcal{E}_e^i\right] \leq e^{-\alpha^2/2}.$$

where the randomness is over the spanning trees $T_1, \cdots, T_{i-1}, T_{i+1}, \cdots, T_{k/2}$ independently sampled from μ_λ.

PROOF. Condition on tree T_i and the event $\mathcal{E}_e^i$. By Lemma 3.6, for any $1 \leq j \leq k/2$ such that $j \neq i$, $\delta(S_{T_i}(e))_{T_j}$ is a $BS(\mathbb{E}\left[\delta(S_{T_i}(e))_{T_j}\right])$ random variable, with $\mathbb{E}\left[\delta(S_{T_i}(e))_{T_j}\right] = \frac{2}{k}x(\delta(S_{T_i}(e))) \geq 2$. Also, by definition, $\delta(S_{T_i}(e))_{T_i} = 1$ (with probability 1). Since $T_1, \cdots, T_{k/2}$ are independently chosen, by Fact 3.7 the random variable $\delta(S_{T_i}(e))_{T^*}$ is distributed as $BS(q)$ for $q \geq k - 1$. Since each T_j has at least one edge in $\delta(S_{T_i}(e))$, $\delta(S_{T_i}(e))_{T^*} \geq k/2$ with probability 1. So, by Theorem 3.12, with $q' = k - 1 - k/2$, when $0 \leq \alpha \leq \sqrt{k/2 - 1}$,

$$\mathbb{P}\left[\delta(S_{T_i}(e))_{T^*} < k - \alpha\sqrt{k/2 - 1} \mid \mathcal{E}_e^i\right]$$
$$= \mathbb{P}\left[\delta(S_{T_i}(e))_{T^*} - k/2 < k/2 - \alpha\sqrt{k/2 - 1} \mid \mathcal{E}_e^i\right]$$
$$\leq e^{-\frac{(\alpha/\sqrt{k})^2(k/2-1)}{2}} = e^{-\alpha^2/2}.$$

Averaging over all realizations of T_i satisfying the required conditions proves the lemma. □

PROOF OF THEOREM 4.1. Let x be an optimum solution of LP (1). Since the output of the algorithm is always k-edge connected we just need to show $\mathbb{E}[c(F \cup T^* \cup B)] \leq \left(1 + \sqrt{\frac{8\ln k}{k}}\right)c(x)$. By linearity

of expectation,

$$\mathbb{E}\left[c(T^*)\right] = \sum_{i \in [\frac{k}{2}]} \mathbb{E}\left[c(T_i)\right]$$

$$= \frac{k}{2} \sum_{e \in E} c(e) \mathbb{P}_{\mu_\lambda}\left[e\right]$$

$$= \frac{k}{2} \sum_{e \in E} c(e) \cdot \frac{2}{k} \cdot x_e = c(x),$$

where for simplicity we ignored the $1 + 2^{-n}$ loss in the marginals. On the other hand, since by Fact 3.3, $\frac{2x}{k}$ is in the spanning tree polytope of G^0, $c(B) \le \frac{2c(x)}{k} \cdot \alpha \sqrt{k/2 - 1} \le \frac{\alpha c(x)}{\sqrt{k/2}}$. It remains to bound the expected cost of F. By Lemma 4.4, $\mathbb{E}\left[c(F)\right]$ is:

$$= \sum_{e \in E} c(e) \sum_{i=1}^{k/2} \mathbb{P}\left[\mathcal{E}_e^i\right] \mathbb{P}\left[\delta(S_{T_i}(e))_{T^*} < k - \alpha\sqrt{k/2 - 1} \mid \mathcal{E}_e^i\right]$$

$$\le \sum_{e \in E} c(e) x_e e^{-\alpha^2/2} \le e^{-\alpha^2/2} c(x).$$

Putting these together we get,

$$\mathbb{E}\left[c(T^* \cup B \cup F)\right] \le (1 + \alpha/\sqrt{k/2} + e^{-\alpha^2/2})c(x).$$

Setting $\alpha = \sqrt{\ln\left(\frac{k}{2}\right)}$ finishes the proof. □

5 IMPROVED ALGORITHM AND PROOF OF MAIN THEOREM

We now introduce our main algorithm that has an approximation ratio of $1 + O(\frac{1}{\sqrt{k}})$. Let x^0 be an optimal solution of LP (1) extended to G^0 as above. Our algorithm is given in Algorithm 2. Note for convenience we drop the ceiling in the expression $\frac{k}{2} + \alpha\sqrt{k}$ in all that follows.

Theorem 1.2 (Main). *There is a polynomial time randomized algorithm for (weighted) k-ECSM with approximation factor (at most) $1 + \frac{5.06}{\sqrt{k}}$.*

We remark that we may assume $k \ge 100$ without loss of generality because for smaller values of k our guarantee is worse than Christofides' algorithm.

Lemma 5.1 (k-Edge Connectivity of the Output). *The output of Algorithm 2, $F \uplus T^*$ is a k-edge connected subgraph of G.*

PROOF. First, note that for every set $S \subset V$ in G, the corresponding cut in G_0 has u_0, v_0 on the same side. Therefore, we may restrict our attention to sets $S \subset V$ such that $u_0, v_0 \notin S$. However for such an S, line 9 of the above algorithm ensures $\delta(S)_{T^* \uplus F} \ge k$, which completes the claim. □

Lemma 5.2 (Variance Upper Bound of Cuts in a Random Spanning Tree). *Let μ_λ be the max-entropy distribution in Algorithm 2. For any $0 \le p \le 1$, any $\epsilon \ge 0$ and any $S \subseteq V$ such that $\mathbb{P}_{T \sim \mu_\lambda}\left[\delta(S)_T = 1\right] = p$ and $\mathbb{E}_{T \sim \mu_\lambda}\left[\delta(S)_T\right] = 2 + \epsilon$, we have $\mathrm{Var}_{T \sim \mu_\lambda}\left[\delta(S)_T\right] \le 4p + 3\epsilon$.*

PROOF. By Theorem 3.4, $\delta(S)_T$ is distributed as a BS random variable with $\mathbb{E}_{T \sim \mu_\lambda}\left[\delta(S)_T\right] = 2 + \epsilon$ and $\mathbb{P}_{T \sim \mu_\lambda}\left[\delta(S)_T \ge 1\right] = 1$. Hence we can write $\delta(S)_T = 1 + X_1 + \cdots + X_m$ for some integer

---

Algorithm 2 Algorithm for k-ECSM with Approximation Ratio $1 + O(\frac{1}{\sqrt{k}})$)

1: Let x^0 be an optimum solution of (1) extended to the graph G^0 as described above.
2: Find weights $\lambda : E^0 \to \mathbb{R}_{\ge 0}$ such that for any $e \in E^0$, $\mathbb{P}_{\mu_\lambda}\left[e\right] \le \frac{2}{k} \cdot x_e^0 \cdot (1 + 2^{-n})$.
3: Initialize $F \leftarrow \emptyset$.
4: Sample $k/2 + \alpha\sqrt{k}$ spanning trees $T_1, \cdots, T_{k/2+\alpha\sqrt{k}} \sim \mu_\lambda$ (in G^0) independently and let $T^* \leftarrow T_1 \uplus \cdots \uplus T_{k/2+\alpha\sqrt{k}}$.
5: Let $\mathcal{S} \leftarrow \{S_{T_i}(e) : i \in [\frac{k}{2} + \alpha\sqrt{k}], e \in T_i, (u_0, v_0) \notin \delta(S)\}$. ▷ $\mathcal{S}$ is the set of one-cuts (see Definition 3.2) of $T_i \in T^*$.
6: **for** $S \in \mathcal{S}$ **do**
7:     $P(S) := \uplus_{i=1}^{k/2+\alpha\sqrt{k}}\{e \in T_i : S_{T_i}(e) = S\}$ ▷ $P(S)$ is the multi-set of $e \in T^*$ with one-cut S.
8:     **if** $\delta(S)_{T^*} < k$ **then**
9:         **for** $j = 1$ to $k - \delta(S)_{T^*}$ **do**
10:             Sample an edge from $P(S)$ uniformly at random and add into F.
11:         **end for**
12:     **end if**
13: **end for**
14: **Return** $T^* \uplus F$.

$m \ge 2$ [7], where $X_1, \cdots, X_m$ are independent Bernoulli random variables with success probabilities $b_1 \ge b_2 \ge \cdots \ge b_m$. Then from the assumption, $\sum_{i=1}^m b_i = 1 + \epsilon$. By Fact 3.7, we have

$$\mathrm{Var}_{T \sim \mu_\lambda}\left[\delta(S)_T\right] = \mathrm{Var}\left[\sum_{i=1}^m X_i\right] = \sum_{i=1}^m b_i(1 - b_i).$$

If $4p \ge 1 - 2\epsilon$, then we have

$$\mathrm{Var}_{T \sim \mu_\lambda}\left[\delta(S)_T\right] = \sum_{i=1}^m b_i(1 - b_i) \le \sum_{i=1}^m b_i = 1 + \epsilon \le 4p + 3\epsilon.$$

Otherwise, $4p < 1 - 2\epsilon$. Notice that

$$p = \Pr[\forall i, X_i = 0] = (1 - b_1)\prod_{i=2}^m (1 - b_i)$$

$$\ge (1 - b_1)(1 - \sum_{i=2}^m b_i) = (1 - b_1) \cdot (b_1 - \epsilon).$$

where the fourth step comes from Weierstrass Inequality, and the last step comes from $\sum_{i=1}^m b_i = 1 + \epsilon$. This gives $b_1 \le \frac{1}{2}(1 + \epsilon - \sqrt{(1 - \epsilon)^2 - 4p})$ or $b_1 \ge \frac{1}{2}(1 + \epsilon + \sqrt{(1 - \epsilon)^2 - 4p})$. Since $4p < 1 - 2\epsilon \le (1 - \epsilon)^2$, the solutions for b_1 are well-defined. By Fact 3.10 we have $b_1 \ge \frac{1}{2}(1 + \epsilon)$, so $b_1 \ge \frac{1}{2}(1 + \epsilon + \sqrt{(1 - \epsilon)^2 - 4p}) \ge 1 - 2p - \frac{\epsilon}{2}$ (using the square root inequality $\sqrt{1 - x} \ge 1 - x$ for $0 \le x \le 1$).

---

[7]We remark that the case for $m = 1$ is trivial.

Therefore, $\mathrm{Var}_{T\sim\mu_\lambda}\left[\delta(S)_T\right]$ is upper-bounded by:

$$\mathrm{Var}_{T\sim\mu_\lambda}\left[\delta(S)_T\right] = \sum_{i=1}^{m} b_i(1-b_i) \leq b_1(1-b_1) + \sum_{i=2}^{m} b_i$$
$$= b_1(1-b_1) + (1+\epsilon - b_1)$$
$$= 1 + \epsilon - b_1^2 \leq 1 + \epsilon - (1 - 2p - \frac{\epsilon}{2})^2 \leq 4p + 3\epsilon.$$

$\square$

As mentioned in Section 2, the following lemma is the key to analyzing Algorithm 2. Roughly speaking, it says that the probability a cut is "bad," i.e. has fewer than $k - \alpha\sqrt{k}$ edges in T^*, is exponentially small in the probability that $\delta(S)_T = 1$ for $T \sim \mu_\lambda$.

Lemma 5.3 (Expected Augmentation of a Cut). *For any $k \geq 100$ and integer $\alpha \geq 1$ let μ_λ be the max-entropy distribution and T^* be the union of $\frac{k}{2} + \alpha\sqrt{k}$ random spanning trees sampled from μ_λ in Algorithm 2. Then for any $S \subseteq V$,*

$$\mathbb{E}_{T^*}\left[(k-\delta(S)_{T^*})^+\right] \leq 1.8\sqrt{k}\exp\left(\frac{-0.6\alpha}{\max\{k^{-1/2}, \mathbb{P}\left[\delta(S)_T = 1\right]\}}\right)$$

PROOF. We can write the expectation as

$$\mathbb{E}_{T^*}\left[(k-\delta(S)_{T^*})^+\right] \leq \sum_{i=1}^{k} \mathbb{P}_{T^*}\left[\delta(S)_{T^*} \leq k - i\right]$$
$$\leq \sum_{i=0}^{\sqrt{k}}\sum_{j=1}^{\sqrt{k}} \mathbb{P}_{T^*}\left[\delta(S)_{T^*} \leq k - (i \cdot \sqrt{k} + j)\right]$$
$$\leq \sum_{i=2\alpha}^{\sqrt{k}} \sqrt{k} \cdot \mathbb{P}_{T^*}\left[\delta(S)_{T^*} \leq k - (i - 2\alpha)\sqrt{k}\right], \tag{3}$$

where we reindex for convenience in the following argument. Define $\beta \geq 0$ such that $x(\delta(S)) = k + \beta\sqrt{k}$, or equivalently that $\mathbb{E}\left[\delta(S)_T\right] = 2(1 + \beta/\sqrt{k})$. By Lemma 5.2, we have

$$\mathrm{Var}_{T^*}\left[\delta(S)_{T^*}\right] \leq (4p + 6\beta/\sqrt{k})(k/2 + \alpha\sqrt{k})$$

where $p = \mathbb{P}\left[\delta(S)_T = 1\right]$. Also, notice,

$$\mathbb{E}\left[\delta(S)_{T^*}\right] \geq (k/2 + \alpha\sqrt{k})\mathbb{E}\left[\delta(S)_T\right] = k + (2\alpha + \beta)\sqrt{k} + 2\alpha\beta.$$

Therefore, by Bernstein's inequality, for $i \geq 2\alpha$, we have that $\mathbb{P}_{T^*}\left[\delta(S)_{T^*} \leq k - (i - 2\alpha)\sqrt{k}\right]$ is equal to

$$\mathbb{P}_{T^*}\left[\delta(S)_{T^*} \leq \mathbb{E}\left[\delta(S)_{T^*}\right] - (i + \beta)\sqrt{k} - 2\alpha\beta\right]$$

Applying Theorem 3.11, this is at most

$$\exp\left(-\frac{(i+\beta)^2 k + 4\alpha\beta(i+\beta)\sqrt{k}}{2(2kp + 3\beta\sqrt{k} + 6\alpha\beta + 4\alpha p\sqrt{k} + (i+\beta)\sqrt{k}/3 + 2\alpha\beta/3)}\right)$$

Now, note that using that $\beta \geq 0$ and the mediant inequality (namely, that for $A, B, C, D \geq 0$ we have $\frac{A+B}{C+D} \geq \min\{A/C, B/D\}$), we can upper bound the term inside the exp by

$$-\min\left\{\frac{i^2 k}{4kp + 8\alpha p\sqrt{k} + 2i\sqrt{k}/3}, \frac{2ik + 4\alpha i\sqrt{k}}{40\alpha/3 + 20\sqrt{k}/3}\right\}$$

Therefore, we can upper bound this probability by

$$\exp\left(-\frac{1}{\max\{k^{-1/2}, p\}}\min\left\{\frac{i^2}{4 + 8\alpha/\sqrt{k} + 2i/3}, \frac{2i + 4\alpha i/\sqrt{k}}{20/3 + 40\alpha/3\sqrt{k}}\right\}\right)$$
$$\underset{i\geq 2\alpha, \alpha\geq 1, k\geq 100}{\leq} \exp\left(\frac{-0.3i}{\max\{k^{-1/2}, p\}}\right)$$

Therefore, $\mathbb{E}_{T^*}\left[(k-\delta(S)_{T^*})^+\right]$ is at most

$$\leq \sqrt{k}\sum_{i=2\alpha}^{\sqrt{k}}\exp\left(\frac{-0.3i}{\max\{k^{-1/2}, p\}}\right)$$
$$\underset{p, k^{-1/2}\leq 1/e}{\leq} \sqrt{k}\exp\left(\frac{-0.6\alpha}{\max\{k^{-1/2}, p\}}\right)\sum_{i=0}^{\infty} e^{-0.3ei}$$
$$\leq 1.8\sqrt{k}\exp\left(\frac{-0.6\alpha}{\max\{k^{-1/2}, p\}}\right)$$

$\square$

Given the above lemma, the expected cost of F follows from a relatively straightforward calculation:

Lemma 5.4 (Expected Payment of an Edge for Augmentation). *For any $k \geq 100$ and integer $\alpha \geq 1$, let T^* be the union of $\frac{k}{2} + \alpha\sqrt{k}$ random spanning trees $T_1, \cdots, T_{k/2+\alpha\sqrt{k}}$ in Algorithm 2. For any solution x to LP (1),*

$$\mathbb{E}_{T^*}\left[c(F)\right] \leq \left(1 + \frac{2\alpha}{\sqrt{k}}\right)\left(\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}\right)c(x)$$

where F is as defined in Algorithm 2.

PROOF. Fix any $i \in [\frac{k}{2} + \alpha\sqrt{k}]$, condition on T_i, fix an edge $e \in T_i$ such that $u_0, v_0 \notin S_{T_i}(e)$. (If $v_0 \in S_{T_i}(e)$, then this is not a cut in the original graph G, so there is nothing to prove). Let $S = S_{T_i}(e)$ and let $p = \mathbb{P}_{T\sim\mu_\lambda}\left[\delta(S)_T = 1\right]$. Recall $P(S) := \uplus_{j=1}^{k/2+\alpha\sqrt{k}}\{f \in T_j : S_{T_j}(f) = S\}$ denotes the multi-set of edges $f \in T_j$ for all $1 \leq j \leq k/2 + \alpha\sqrt{k}$, such that $S_{T_j}(f) = S$.

Let $X_{T_i, e}$ be the number of times that edge e from tree T_i is sampled in Algorithm 2, line 10. We will prove that, letting $\mathcal{E}_e^i$ denote the event $e \in T_i, (u_0, v_0) \notin \delta(S_{T_i}(e))$,

$$\mathbb{E}\left[X_{T_i, e} \mid \mathcal{E}_e\right] \leq \frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2} \tag{4}$$

Then, to prove the lemma,

$$\mathbb{E}\left[c(F)\right] = \sum_{e\in E} c(e)\sum_{i=1}^{k/2+\alpha\sqrt{k}}\mathbb{P}\left[\mathcal{E}_e^i\right] \cdot \mathbb{E}\left[X_{T_i, e} \mid \mathcal{E}_e^i\right]$$
$$\underset{(4)}{\leq} \sum_{e\in E} c(e)\sum_{i=1}^{k/2+\alpha\sqrt{k}}\mathbb{P}\left[\mathcal{E}_e^i\right]\left(\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}\right)$$
$$\leq \sum_{e\in E} c(e)\left(\frac{k}{2} + \alpha\sqrt{k}\right) \cdot \frac{2}{k}x_e \cdot \left(\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}\right)$$
$$= \left(1 + \frac{2\alpha}{\sqrt{k}}\right)\left(\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}\right)c(x)$$

In the rest of the proof we show (4). First, observe that

$$\mathbb{E}\left[X_{T_i,e} \mid T_i, e \in T_i, \delta(S)_{T^*}, |P(S)|\right] = \frac{(k - \delta(S)_{T^*})^+}{|P(S)|}.$$

So, it is enough to upper bound the expected value of the RHS conditioned on $T_i, e \in T_i$. Let $T_{-i}^* = T^* \smallsetminus T_i$ and $P(S)_{-i} := P(S) \smallsetminus T_i$ and note that $|P(S)| = |P(S)_{-i}| + 1$ and $\delta(S)_{T^*} = \delta(S)_{T_{-i}^*} + 1$. Define $Y_{e,T_i} = \frac{(k - \delta(S)_{T_{-i}^*} + 1)^+}{|P(S)_{-i}| + 1}$. Note that $\mathbb{E}\left[X_{T_i,e} | T_i, e \in T_i\right] = \mathbb{E}\left[Y_{e,T_i}\right]$. We drop the subscript of Y for readability.

First, assume $p \geq 2/k$. Then, we write,

$$\mathbb{E}[Y] = \mathbb{P}\left[|P(S)_{-i}| \geq \frac{pk}{4}\right] \cdot \mathbb{E}\left[Y \mid |P(S)_{-i}| \geq \frac{pk}{4}\right] \tag{5}$$
$$+ \mathbb{P}\left[|P(S)_{-i}| < \frac{pk}{4}\right] \cdot \mathbb{E}\left[Y \mid |P(S)_{-i}| < \frac{pk}{4}\right]$$

We upper bound each term in the RHS separately.

$$\mathbb{P}\left[|P(S)_{-i}| \geq \frac{pk}{4}\right] \mathbb{E}\left[Y \mid |P(S)_{-i}| \geq \frac{pk}{4}\right]$$
$$\leq \frac{4}{pk} \mathbb{P}\left[|P(S)_{-i}| \geq \frac{pk}{4}\right] \mathbb{E}\left[(k - \delta(S)_{T_{-i}^*})^+ \mid |P(S)_{-i}| \geq \frac{pk}{4}\right]$$
$$\leq \frac{4}{pk} \mathbb{E}\left[(k - \delta(S)_{T_{-i}^*})^+\right] \leq \frac{7.2}{p\sqrt{k}} \exp\left(\frac{-0.6\alpha}{\max\{p, k^{-1/2}\}}\right)$$

Now, if $p > k^{-1/2}$, the RHS is maximized when $p = 1/e$ since $\frac{1}{p}e^{-c/p}$ is an increasing function of p for $0 \leq p \leq 1/e$; note $p \leq 1/e$ (by Corollary 3.9) and $\alpha \geq 1$. We obtain a bound of $\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e}$.

Otherwise, the maximum is achieved by $k^{-1/2}$. Using $p \geq 2/k$, the above expression is at most $3.6\sqrt{k}e^{-0.6\alpha k^{1/2}}$. So, the first term is at most $\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e}$ for $k \geq 100$.

Next, we bound the second term of (5). First notice $Y \leq 1$ with probability 1; this is because if there are exactly ℓ trees which have $(S, \bar{S})$ as a one-cut then, $\delta(S)_{T^*} \geq k + 2\alpha\sqrt{k} - \ell$ whereas $|P(S)| = \ell$. Furthermore, $Y \neq 0$ only when $|P(S)_{-i}| \geq 2\alpha\sqrt{k} \geq 2\sqrt{k}$ (for $\alpha \geq 1$). Therefore,

$$\mathbb{P}\left[|P(S)_{-i}| < \frac{pk}{4}\right] \cdot \mathbb{E}\left[Y \mid |P(S)_{-i}| < \frac{pk}{4}\right]$$
$$\leq \mathbb{P}\left[|P(S)_{-i}| < \frac{pk}{4}\right] \mathbb{P}\left[|P(S)_{-i}| \geq 2\sqrt{k} \mid |P(S)_{-i}| < \frac{pk}{4}\right]$$
$$= \mathbb{P}\left[2\sqrt{k} \leq |P(S)_{-i}| \leq \frac{pk}{4}\right] \leq e^{-pk/16} \leq e^{-\sqrt{k}/2}$$

To see the last two inequalities notice we must have $p \geq 8k^{-1/2}$ or this event cannot occur. Therefore, since $\mathbb{E}[|P(S)_{-i}|] = p(\frac{k}{2} + \alpha\sqrt{k} - 1)$ the inequality follows by an application of the Chernoff bound (Theorem 3.12).

Putting these two terms together, if $p > 2/k$,

$$\mathbb{E}\left[X_{T_i,e} \mid T_i, e \in T_i\right] = \mathbb{E}[Y] \leq \frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}$$

Otherwise, suppose $p \leq 2/k$. Then since $\mathbb{E}[|P(S)|_{-i}] \leq 1 + 2\alpha k^{-1/2}$, by Theorem 3.12 (using $\alpha \geq 1$),

$$\mathbb{P}\left[|P(S)_{-i}| > (1 + 2\alpha\sqrt{k})(1 + 2\alpha k^{-1/2})\right] \leq e^{-\frac{4\alpha^2 k(1 + 2\alpha k^{-1/2})}{2 + 2\alpha\sqrt{k}}}$$
$$\underset{k \geq 10, \alpha \geq 1}{\leq} e^{-\sqrt{k}}.$$

Since $Y \leq 1$ as observed above, we obtain

$$\mathbb{E}[Y] \leq \mathbb{P}\left[\delta(S)_{T_{-i}^*} \leq k\right] \leq \mathbb{P}\left[|P(S)|_{-i} \geq 2\alpha\sqrt{k}\right] \leq e^{-\sqrt{k}}$$

which gives (4). Therefore we can bound $\mathbb{E}[Y]$ by $\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}$ for all values of p. $\qquad\square$

Proof of Theorem 1.2. Let x be the optimum solution of (1). From Lemma 5.1, the output of Algorithm 2 is always k-edge connected. Thus it suffices to show that $\mathbb{E}[c(T^* \uplus F)] \leq (1 + \frac{5.06}{\sqrt{k}})c(x)$. By linearity of expectation,

$$\mathbb{E}\left[c(T^*)\right] = \sum_{i \in [\frac{k}{2} + \alpha\sqrt{k}]} \mathbb{E}[c(T_i)]$$
$$= (\frac{k}{2} + \alpha\sqrt{k}) \sum_{e \in E} c(e) \mathbb{P}_{\mu_\lambda}[e]$$
$$= (\frac{k}{2} + \alpha\sqrt{k}) \sum_{e \in E} c(e) \cdot \frac{2}{k} \cdot x_e$$
$$= (1 + \frac{2\alpha}{\sqrt{k}})c(x),$$

where for simplicity we ignored the $1 + 2^{-n}$ loss in the marginals.

Therefore, by Lemma 5.4, $\mathbb{E}[c(T^* \uplus F)]$ is at most

$$\leq c(x) \cdot \left(1 + \frac{2\alpha}{\sqrt{k}} + \left(1 + \frac{2\alpha}{\sqrt{k}}\right)\left(\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}\right)\right)$$
$$\leq c(x) \cdot \left(1 + \frac{5.06}{\sqrt{k}}\right).$$

as desired, where in the last inequality we use $k \geq 100$ and set $\alpha = 2$. $\qquad\square$

6 CONCLUSION

We remark that the approximation factor $1 + O(1/\sqrt{k})$ is tight for any algorithm that starts by sampling $O(k)$ spanning trees independently from the max-entropy distribution and then fixes the union by adding edges. For a tight example, consider a complete graph with a unit metric on the edges and let x be uniform across all edges. In such a case, the max-entropy distribution μ_λ will be the uniform distribution over all spanning trees of a complete graph. A simple analysis shows that every vertex will have degree $k - \sqrt{k}$ in T^* with constant probability. Therefore, to fix T^* we need to add at least $\Omega(n\sqrt{k})$ edges.

It still remains open if the integrality gap of the LP is indeed $1 + O(1/k)$ or if there is an approximation algorithm with approximation factor $1 + O(1/k)$. It would also be interesting to find the optimal constant for Algorithm 2.

REFERENCES

[1] Arash Asadpour, Michel X. Goemans, Aleksander Mdry, Shayan Oveis Gharan, and Amin Saberi. 2017. An O (log n /log log n)-Approximation Algorithm for the Asymmetric Traveling Salesman Problem. *Operations Research* 65, 4 (2017).

[2] Julius Borcea, Petter Brändén, and Thomas M. Liggett. 2009. Negative dependence and the geometry of polynomials. *Journal of the American Mathematical Society* 22, 2 (2009), 521–567.

[3] Sylvia Boyd, Joseph Cheriyan, Robert Cummings, Logan Grout, Sharat Ibrahimpur, Zoltán Szigeti, and Lu Wang. 2020. A 4/3-Approximation Algorithm for the Minimum 2-Edge Connected Multisubgraph Problem in the Half-Integral Case. In *APPROX/RANDOM*, Jarosław Byrka and Raghu Meka (Eds.), Vol. 176. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 61:1–61:12.

[4] Sylvia Boyd, Yao Fu, and Yu Sun. 2016. A 5/4-approximation for subcubic 2EC using circulations and obliged edges. *Discrete Applied Mathematics* 209 (2016), 48–58.

[5] Jaroslaw Byrka, Fabrizio Grandoni, and Afrouz Jabal Ameli. 2020. Breaching the 2-approximation barrier for connectivity augmentation: a reduction to Steiner tree. In *STOC*, Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy (Eds.). ACM, 815–825.

[6] Robert Carr and R. Ravi. 1998. A New Bound for the 2-Edge Connected Subgraph Problem. In *IPCO*, Robert E. Bixby, E. Andrew Boyd, and Roger Z. Ríos-Mercado (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 112–125.

[7] Federica Cecchetto, Vera Traub, and Rico Zenklusen. 2021. Bridging the gap between tree and connectivity augmentation: unified and stronger approaches. In *STOC*, Samir Khuller and Virginia Vassilevska Williams (Eds.). ACM, 370–383.

[8] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. 2010. Dependent Randomized Rounding via Exchange Properties of Combinatorial Structures. In *FOCS*. 575–584.

[9] J. Cheriyan and R. Thurimella. 2000. Approximating minimum- size k-connected spanning subgraphs via matching. *SIAM J Comput* 30 (2000), 528–560.

[10] Jack Edmonds. 1970. Submodular functions, matroids and certain polyhedra. In *Combinatorial Structures and Their Applications*. Gordon and Breach, New York, NY, USA, 69–87.

[11] Samuel Fiorini, Martin Groß, Jochen Könemann, and Laura Sanità. 2018. Approximating Weighted Tree Augmentation via Chvátal-Gomory Cuts. In *SODA*. 817–831.

[12] G. N. Fredrickson and Joseph F. JáJá. 1981. Approximation Algorithms for Several Graph Augmentation Problems. *SIAM J. Comput.* 10, 2 (1981), 270–283.

[13] G. N. Fredrickson and Joseph F. JáJá. 1982. On the relationship between the biconnectivity augmentation and traveling salesman problem. *Theoretical Computer Science* 19 (1982), 189 – 201.

[14] H. Gabow. 2005. An improved analysis for approximating the smallest k-edge connected spanning subgraph of a multi-graph. *SIAM J Disc Math* 19 (2005), 1–18.

[15] H. Gabow and S. Gallagher. 2008. Iterated rounding algorithms for the smallest k-edge connected spanning subgraph. *SIAM J. Comput.* 41 (2008), 61–103.

[16] Harold N. Gabow, Michel X. Goemans, Éva Tardos, and David P. Williamson. 2009. Approximating the smallest k-edge connected spanning subgraph by LP-rounding. *Networks* 53, 4 (2009), 345–357.

[17] Michel X. Goemans and Dimitris Bertsimas. 1993. Survivable networks, linear programming relaxations and the parsimonious property. *Math. Program.* 60 (1993), 145–166.

[18] Wassily Hoeffding. 1956. On the distribution of the number of successes in independent trials. *Annals of Mathematical statistics* 27, 3 (1956), 713–721.

[19] D. Karger. 1999. Random sampling in cut, flow, and network design problems. *Math OR* 24 (1999), 383–413.

[20] Anna Karlin, Nathan Klein, and Shayan Oveis Gharan. 2021. A (Slightly) Improved Bound on the Integrality Gap of the Subtour LP for TSP. arXiv:2105.10043 [cs.DS]

[21] Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. 2021. A (Slightly) Improved Approximation Algorithm for Metric TSP. In *STOC*. ACM.

[22] S. Khuller and B. Raghavachari. 1996. Improved approximation algorithms for uniform connectivity problems. *J Algorithms* 21 (1996), 434–450.

[23] Bundit Laekhanukit, Shayan Oveis Gharan, and Mohit Singh. 2012. A Rounding by Sampling Approach to the Minimum Size k-Arc Connected Subgraph Problem. In *ICALP (Lecture Notes in Computer Science, Vol. 7391)*, Artur Czumaj, Kurt Mehlhorn, Andrew M. Pitts, and Roger Wattenhofer (Eds.). Springer, 606–616.

[24] Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. 2011. A Randomized Rounding Approach to the Traveling Salesman Problem. In *FOCS*. IEEE Computer Society, 550–559.

[25] Jim Pitman. 1997. Probabilistic bounds on the coefficients of polynomials with only real zeros. *Journal of Combinatorial Theory, Series A* 77, 2 (1997), 279–303.

[26] David Pritchard. 2011. k-Edge-Connectivity: Approximation and LP Relaxation. In *Approximation and Online Algorithms*, Klaus Jansen and Roberto Solis-Oba (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 225–236.

[27] András Sebö and Jens Vygen. 2014. Shorter tours by nicer ears: 7/5-Approximation for the graph-TSP, 3/2 for the path version, and 4/3 for two-edge-connected subgraphs. *Combinatorica* 34, 5 (2014), 597–629.

[28] Vera Traub and Rico Zenklusen. 2021. A Better-Than-2 Approximation for Weighted Tree Augmentation. (2021). to appear.