Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A (Slightly) Improved Approximation Algorithm for Metric TSP

Anna R. Karlin, Nathan Klein, Shayan Oveis Gharan

informs
https://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Vol. 72, No. 6, November–December 2024, pp. 2543–2594
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Methods

A (Slightly) Improved Approximation Algorithm for Metric TSP

Anna R. Karlin,[a] Nathan Klein,[a] Shayan Oveis Gharan[a,*]

[a] Department of Computer Science, University of Washington, Seattle, Washington 98103

*Corresponding author

Contact: karlin@cs.washington.edu (ARK); nwklein@cs.washington.edu, https://orcid.org/0009-0003-4052-5864 (NK); shayan@cs.washington.edu, https://orcid.org/0000-0002-5504-0849 (SOG)

Received: September 9, 2020
Revised: December 27, 2021
Accepted: May 31, 2022
Published Online in Articles in Advance: May 18, 2023

Area of Review: Optimization

https://doi.org/10.1287/opre.2022.2338

Abstract. For some $\epsilon > 10^{-36}$, we give a randomized $3/2 - \epsilon$ approximation algorithm for metric TSP.

Funding: This work was supported by the Air Force Office of Scientific Research [Grant FA9550-20-1-0212], the Alfred P. Sloan Foundation [Grant FG-2019-12196], and the Division of Computing and Communication Foundations [Grants CCF-1552097, CCF-1813135, and CCF-1907845].

Keywords: approximation algorithms • maximum entropy distributions • strongly Rayleigh distributions • random spanning trees • traveling salesperson problem

1. Introduction

One of the most fundamental problems in combinatorial optimization is the traveling salesperson problem (TSP), formalized as early as 1832 (Applegate et al. 2007, chapter 1). In an instance of TSP, we are given a set of n cities V along with their pairwise symmetric distances, $c : V \times V \to \mathbb{R}_{\geq 0}$. The goal is to find a Hamiltonian cycle of minimum cost. It is well known that for a general distance function, it is NP-hard to approximate TSP within any polynomial factor. Therefore, it is natural to study metric TSP, in which the distances satisfy the triangle inequality, that is,

$$c(u,w) \leq c(u,v) + c(v,w) \qquad \forall u,v,w \in V.$$

In this case, the problem is equivalent to finding a closed Eulerian connected walk of minimum cost.[1]

It is NP-hard to approximate metric TSP within a factor of $\frac{123}{122}$ (Karpinski et al. 2015). An algorithm of Christofides-Serdyukov (Christofides 1976, Serdyukov 1978) from four decades ago gives a $\frac{3}{2}$-approximation for TSP (see van Bevern and Slugina (2020) for a historical note about TSP). This remains the best known approximation algorithm for the general case of the problem despite significant work (Wolsey 1980, Shmoys and Williamson 1990, Boyd and Pulleyblank 1991, Goemans 1995, Carr and Vempala 2000, Gamarnik et al. 2005, Boyd and Elliott-Magwood 2010, Boyd and Carr 2011, Schalekamp et al. 2012, Haddadan et al. 2017, Haddadan and Newman 2019, Karlin et al. 2020).

In contrast, there have been major improvements to this algorithm for a number of special cases of TSP. For example, polynomial-time approximation schemes (PTAS) have been found for Euclidean (Arora 1998, Mitchell 1999), planar (Grigni et al. 1995, Arora et al. 1998, Klein 2008), and low-genus metric (Demaine et al. 2010) instances. In addition, the case of graph metrics has received significant attention. Oveis Gharan et al. (2011) found a $\frac{3}{2} - \epsilon_0$ approximation for this case. Mömke and Svensson (2016) then obtained a combinatorial algorithm for graphic TSP with an approximation ratio of 1.461. This ratio was later improved by Mucha (2012) to $\frac{13}{9} \approx 1.444$ and then by Sebö and Vygen (2014) to 1.4.

In this paper, we prove the following theorem.

Theorem 1.1. *For some absolute constant $\epsilon > 10^{-36}$, there is a randomized algorithm that outputs a tour with expected cost at most $\frac{3}{2} - \epsilon$ times the cost of the optimum solution.*

Although the algorithm makes use of the Held-Karp relaxation, we do not prove that the integrality gap of this polytope is bounded away from 3/2. We also remark that, although our approximation factor is only slightly better than Christofides-Serdyukov, we are not aware of any example where the approximation ratio of the algorithm we analyze exceeds 4/3 in expectation.

Following a new exciting result of Traub et al. (2020), we also get the following theorem.

Theorem 1.2. *For some absolute constant $\epsilon > 0$, there is a randomized algorithm that outputs a TSP path with expected cost at most $\frac{3}{2} - \epsilon$ times the cost of the optimum solution.*

1.1. Algorithm

First, we recall the classical Christofides-Serdyukov algorithm: Given an instance of TSP, choose a minimum spanning tree and then add the minimum cost matching on the odd degree vertices of the tree. The algorithm we study is very similar, except we choose a random spanning tree based on the standard linear programming relaxation of TSP.

Let x^0 be an optimum solution of the following TSP linear program relaxation (Dantzig et al. 1959, Held and Karp 1970):

$$\begin{aligned}
\min \quad & \sum_{u,v} x_{(u,v)} c(u,v) & \\
\text{s.t.,} \quad & \sum_{u} x_{(u,v)} = 2 & \forall v \in V, \\
& \sum_{u \in S, v \notin S} x_{(u,v)} \geq 2, & \forall S \subsetneq V, \\
& x_{(u,v)} \geq 0 & \forall u, v \in V.
\end{aligned} \qquad \text{(Held-Karp relaxation)}$$

Given x^0, we pick an arbitrary node, u, split it into two nodes u_0, v_0, and set $x_{(u_0,v_0)} = 1, c(u_0, v_0) = 0$, and we assign half of every edge incident to u to u_0 and the other half to v_0. This allows us to assume without loss of generality that x^0 has an edge $e_0 = (u_0, v_0)$ such that $x_{e_0} = 1, c(e_0) = 0$.

Let $E_0 = E \cup \{e_0\}$ be the support of x^0 and let x be x^0 restricted to E and $G = (V, E)$. x^0 restricted to E is in the spanning tree polytope (2.1).

For a vector $\lambda : E \to \mathbb{R}_{\geq 0}$, a λ-uniform distribution μ_λ over spanning trees of $G = (V, E)$ is a distribution where for every spanning tree

$$T \subseteq E, \mathbb{P}_\mu[T] = \frac{\prod_{e \in T} \lambda_e}{\sum_{T'} \prod_{e \in T'} \lambda_e}.$$

Now, find a vector λ such that for every edge $e \in E$, $\mathbb{P}_{\mu_\lambda}[e \in T] = x_e(1 \pm \epsilon)$, for some $\epsilon < 2^{-n}$. Such a vector λ can be found using the multiplicative weight update algorithm (Asadpour et al. 2017) or by applying interior point methods (Sebö and Vygen 2014) or the ellipsoid method (Asadpour et al. 2017). (The multiplicative weight update method can only guarantee $\epsilon < 1/\text{poly}(n)$ in polynomial time.) We will sometimes call such a distribution the *maximum entropy* distribution because a λ-uniform distribution has maximal entropy over all distributions with marginals x.[2]

Theorem 1.3 (Asadpour et al. 2017, Theorem 5.2). *Let z be a point in the spanning tree polytope (see (2.1)) of a graph $G = (V, E)$. For any $\epsilon > 0$, a vector $\lambda : E \to \mathbb{R}_{\geq 0}$ can be found such that the corresponding λ-uniform spanning tree distribution, μ_λ, satisfies*

$$\sum_{T \in \mathcal{T} : T \ni e} \mathbb{P}_{\mu_\lambda}[T] \leq (1 + \varepsilon) z_e, \quad \forall e \in E,$$

that is, the marginals are approximately preserved. In the previous notation, $\mathcal{T}$ is the set of all spanning trees of (V, E). The running time is polynomial in $n = |V|$, $-\log \min_{e \in E} z_e$ and $\log(1/\epsilon)$.

Finally, we sample a tree $T \sim \mu_\lambda$ and then add the minimum cost matching on the odd degree vertices of T. The previous algorithm is a slight modification of the algorithm proposed in Oveis Gharan et al. (2011). We refer the interested reader to exciting work of Genova and Williamson (2017) on the empirical performance of the max-entropy rounding algorithm. We also remark that, although the algorithm implemented in Genova and Williamson (2017) is slightly different from the previous algorithm, we expect the performance to be similar.

Algorithm 1 (Improved Approximation Algorithm for TSP)

Find an optimum solution x^0 of Held-Karp relaxation, and let $e_0 = (u_0, v_0)$ be an edge with $x^0_{e_0} = 1, c(e_0) = 0$.
Let $E_0 = E \cup \{e_0\}$ be the support of x^0 and x be x^0 restricted to E and $G = (V, E)$.
Find a vector $\lambda : E \to \mathbb{R}_{\geq 0}$ such that for any $e \in E$, $\mathbb{P}_{\mu_\lambda}[e] = x_e(1 \pm 2^{-n})$.
Sample a tree $T \sim \mu_\lambda$.
Let M be the minimum cost matching on odd degree vertices of T.
Output $T \cup M$.

1.2. New Techniques

Here we discuss new machinery and technical tools that we developed for this result that could be of independent interest.

1.2.1. Polygon Structure for Near Minimum Cuts Crossed on One Side.

Let $G = (V, E, x)$ be an undirected graph equipped with a weight function $x : E \to \mathbb{R}_{\geq 0}$ such that for any cut $(S, \overline{S})$ such that $u_0, v_0 \notin S$, $x(\delta(S)) \geq 2$.

For some (small) $\eta \geq 0$, consider the family of η-near min cuts of G. Let $\mathcal{C}$ be a connected component of crossing η-near min cuts. Given $\mathcal{C}$ we can partition vertices of G into sets $a_0, \ldots, a_{m-1}$ (called atoms); this is the coarsest partition such that for each a_i, and each $(S, \overline{S}) \in \mathcal{C}$, we have $a_i \subseteq S$ or $a_i \subseteq \overline{S}$. Here a_0 is the atom that contains u_0, v_0.

There have been several works studying the structure of edges between these atoms and the structure of cuts in a connected component of cuts $\mathcal{C}$ with respect to (w.r.t.) the a_is. The *cactus structure* (Dinits et al. 1976) shows that if $\eta = 0$, then we can arrange the a_is of a connected component around a cycle, say $a_1, \ldots, a_m$ (after renaming), such that $x(E(a_i, a_{i+1})) = 1$ for all i.

Benczúr (1995) and Benczúr and Goemans (2008) studied the case when $\eta \le 6/5$ and introduced the notion of *polygon representation*, in which case atoms can be placed on the sides of an equilateral polygon and some atoms placed inside the polygon, such that every cut in $\mathcal{C}$ can be represented by a diagonal of this polygon. Later, Oveis Gharan et al. (2011) studied the structure of edges of G in this polygon when $\eta < 1/100$.

In this paper, we show it suffices to study the structure of edges in a special family of polygon representations: Suppose we have a polygon representation for a connected component $\mathcal{C}$ of η-near min cuts of G such that

- No atom is mapped inside.
- If we identify each cut $(S, \overline{S}) \in \mathcal{C}$ with the interval along the polygon that does not contain a_0, then any interval is only crossed on one side (only on the left or only on the right).

Then, we have (i) for any atom a_i, $x(\delta(a_i)) \le 2 + O(\eta)$, and (ii) for any pair of atoms, a_i, a_{i+1}, $x(E(a_i, a_{i+1}) \ge 1 - \Omega(\eta)$ (see Theorem 4.3 for details).

We expect to see further applications of our theorem in studying variants of TSP.

1.2.2. Generalized Gurvits' Lemma.

Given a real stable polynomial $p \in \mathbb{R}_{\ge 0}[z_1, \ldots, z_n]$ (with nonnegative coefficients), Gurvits (2006, 2008) proved the following inequality:

$$\frac{n!}{n^n}\inf_{z>0}\frac{p(z_1,\ldots,z_n)}{z_1\ldots z_n} \le \partial_{z_1}\ldots\partial_{z_n}p|_{z=0} \le \inf_{z>0}\frac{p(z_1,\ldots,z_n)}{z_1\ldots z_n}. \tag{1.1}$$

As an immediate consequence, one can prove the following theorem about strongly Rayleigh (SR) distributions.

Theorem 1.4. *Let $\mu : 2^{[n]} \to \mathbb{R}_{\ge 0}$ be SR and $A_1, \ldots, A_m$ be random variables corresponding to the number of elements sampled in m disjoint subsets of $[n]$ such that $\mathbb{E}[A_i] = n_i$ for all i. If $n_i = 1$ for all $1 \le i \le n$, then $\mathbb{P}[\forall i, A_i = 1] \ge \frac{m!}{m^m}$.*

One can ask what happens if the vector $(n_1, \ldots, n_m)$ in the previous theorem is not equal but close to the all ones vector, **1**.

A related theorem was proved in Oveis Gharan et al. (2011).

Theorem 1.5. *Let $\mu : 2^{[n]} \to \mathbb{R}_{\ge 0}$ be SR and A, B be random variables corresponding to the number of elements sampled in two disjoint sets. If $\mathbb{P}[A + B = 2] \ge \epsilon$, $\mathbb{P}[A \le 1], \mathbb{P}[B \le 1] \ge \alpha$ and $\mathbb{P}[A \ge 1], \mathbb{P}[B \ge 1] \ge \beta$ then $\mathbb{P}[A = B = 1] \ge \epsilon\alpha\beta/3$.*

We prove a generalization of both of the previous statements; roughly speaking, we show that as long as $\sum_{i=1}^m |n_i - 1| < 1 - \epsilon$, then $\mathbb{P}[\forall i, A_i = 1] \ge f(\epsilon, m)$, where $f(\epsilon, m)$ has no dependence on n, the number of underlying elements in the support of μ.

Theorem 1.6 (Informal Version of Proposition 5.1). *Let $\mu : 2^{[n]} \to \mathbb{R}_{\ge 0}$ be SR and let $A_1, \ldots, A_m$ be random variables corresponding to the number of elements sampled in m disjoint subsets of $[n]$. Suppose that there are integers $n_1, \ldots, n_m$ such that for any set $S \subseteq [m]$, $\mathbb{P}[\sum_{i\in S} A_i = \sum_{i\in S} n_i] \ge \epsilon$. Then,*

$$\mathbb{P}[\forall i, A_i = n_i] \ge f(\epsilon, m).$$

The previous statement is even stronger than Theorem 1.4 as we only require $\mathbb{P}[\sum_{i\in S} A_i = \sum_{i\in S} n_i]$ to be bounded away from zero for any set $S \subseteq [m]$, and we do not need a bound on the expectation. Our proof of the previous theorem has double exponential dependence on ϵ. We leave it an open problem to find the optimum dependency on ϵ. Furthermore, our proof of the previous theorem is probabilistic in nature; we expect that an algebraic proof based on the theory of real stable polynomials will provide a significantly improved lower bound. Unlike the previous theorem, such a proof may possibly extend to the more general class of completely log-concave distributions (Anari et al. 2018). In an independent work, Gurvits and Leake (2021) proved a variant of the previous theorem with a much better dependence on ϵ and m for a homogeneous strong Rayleigh distribution.

1.2.3. Conditioning While Preserving Marginals.

Consider a SR distribution $\mu : 2^{[n]} \to \mathbb{R}_{\ge 0}$ and let $x : [n] \to \mathbb{R}_{\ge 0}$, where for all i, $x_i = \mathbb{P}_{T\sim\mu}[i \in T]$, be the marginals.

Let $A, B \subseteq [n]$ be two disjoint sets such that $\mathbb{E}[A_T]$, $\mathbb{E}[B_T] \approx 1$. It follows from Theorem 1.6 that $\mathbb{P}[A_T = B_T = 1] \ge \Omega(1)$. Here, however, we are interested in a stronger event; let $\nu = \mu | A_T = B_T = 1$ and let $y_i = \mathbb{P}_{T\sim\mu}[i \in T]$. It turns out that the y vector can be very different from the x vector, in particular, for some is we can have $|y_i - x_i|$ bounded away from zero. We show that there is an event of nonnegligible probability that is a subset of $A_T = B_T = 1$ under which the marginals of elements in A, B are almost preserved.

Theorem 1.7 (Informal Version of Proposition 5.2). *Let $\mu : 2^{[n]} \to \mathbb{R}_{\ge 0}$ be a SR distribution and let $A, B \subseteq [n]$ be two disjoint subsets such that $\mathbb{E}[A_T], \mathbb{E}[B_T] \approx 1$. For any $\alpha \ll 1$ there is an event $\mathcal{E}_{A,B}$ such that $\mathbb{P}[\mathcal{E}_{A,B}] \ge \Omega(\alpha^2)$ and*

- $\mathbb{P}[A_T = B_T = 1 | \mathcal{E}_{A,B}] = 1$,
- $\sum_{i\in A} |\mathbb{P}[i] - \mathbb{P}[i|\mathcal{E}_{A,B}]| \le \alpha$,
- $\sum_{i\in B} |\mathbb{P}[i] - \mathbb{P}[i|\mathcal{E}_{A,B}]| \le \alpha$.

We remark that the quadratic lower bound on α is necessary in the previous theorem for a sufficiently small $\alpha > 0$. The previous theorem can be seen as a generalization of Theorem 1.4 in the special case of two sets.

We leave it an open problem to extend the previous theorem to arbitrary k disjoint sets. We suspect that in such a case the ideal event $\mathcal{E}_{A_1,\ldots,A_k}$ occurs with probability $\Omega(\alpha)^k$ and preserves all marginals of elements in each of the sets $A_1,\ldots,A_k$ up to a total variation distance of α.

2. Preliminaries

2.1. Notation

We write $[n] := \{1,\ldots,n\}$ to denote the set of integers from one to n. For a set of edges $A \subseteq E$ and (a tree) $T \subseteq E$, we write

$$A_T = |A \cap T|.$$

For a set $S \subseteq V$, we write

$$E(S) = \{(u,v) \in E : u,v \in S\}$$

to denote the set of edges in S, and we write

$$\delta(S) = \{(u,v) \in E : |\{u,v\} \cap S| = 1\}$$

to denote the set of edges that leave S.

For two *disjoint* sets of vertices $A, B \subseteq V$, we write

$$E(A,B) = \{(u,v) \in E : u \in A, v \in B\}.$$

For a set $A \subseteq E$ and a function $x : E \to \mathbb{R}$, we write

$$x(A) := \sum_{e \in A} x_e.$$

For two sets $A, B \subseteq V$, we say A *crosses* B if all of the following sets are nonempty:

$$A \cap B, A \smallsetminus B, B \smallsetminus A, \overline{A \cup B}.$$

We write $G = (V,E,x)$ to denote an (undirected) graph G together with special vertices u_0, v_0 and a weight function $x : E \to \mathbb{R}_{\geq 0}$ such that

$$x(\delta(S)) \geq 2, \qquad \forall S \subsetneq V : u_0, v_0 \notin S.$$

For such a graph, we say a cut $S \subseteq V$ is an *η-near min cut w.r.t., x* (or simply η-near min cut when x is understood) if $x(\delta(S)) \leq 2 + \eta$. Unless otherwise specified, in any statement about a cut $(S, \overline{S})$ in G, we assume $u_0, v_0 \notin S$.

2.2. Polyhedral Background

For any graph $G = (V,E)$, Edmonds (1970) gave the following description for the convex hull of spanning trees of a graph $G = (V,E)$, known as the *spanning tree polytope*:

$$\begin{aligned} & z(E) = |V| - 1 & & \\ & z(E(S)) \leq |S| - 1 & & \forall S \subseteq V \\ & z_e \geq 0 & & \forall e \in E. \end{aligned} \tag{2.1}$$

Edmonds (1970) proved that the extreme point solutions of this polytope are the characteristic vectors of the spanning trees of G.

Fact 2.1. *Let x^0 be a feasible solution of the Held-Karp relaxation such that $x^0_{e_0} = 1$ with support $E_0 = E \cup \{e_0\}$. Let x be x^0 restricted to E; then x is in the spanning tree polytope of $G = (V,E)$.*

For any set $S \subseteq V$ such that $u_0, v_0 \notin S$, $x(E(S)) = \frac{2|S| - x^0(\delta(S))}{2} \leq |S| - 1$. If $u_0 \in S, v_0 \notin S$, then $x(E(S)) = \frac{2|S|-1-(x^0(\delta(S))-1)}{2} \leq |S| - 1$. Finally, if $u_0, v_0 \in S$, then $x(E(S)) = \frac{2|S|-2-x^0(\delta(S))}{2} \leq |S| - 2$. The claim follows because $x(E) = x^0(E_0) - 1 = n - 1$.

Because $c(e_0) = 0$, the following fact is immediate.

Fact 2.2. *Let $G = (V,E,x)$ where x is in the spanning tree polytope. Let μ be any distribution of spanning trees with marginals x, then $\mathbb{E}_{T \sim \mu}[c(T \cup e_0)] = c(x)$.*

To bound the cost of the min-cost matching on the set O of odd degree vertices of the tree T, we use the following characterization of the O-join polytope[3] due to Edmonds and Johnson (1973).

Proposition 2.1. *For any graph $G = (V,E)$, cost function $c : E \to \mathbb{R}_+$, and a set $O \subseteq V$ with an even number of vertices, the minimum weight of an O-join equals the optimum value of the following integral linear program.*

$$\begin{aligned} \min \quad & c(y) & & \\ s.t. \quad & y(\delta(S)) \geq 1 & & \forall S \subseteq V, |S \cap O| \text{ odd} \\ & y_e \geq 0 & & \forall e \in E \end{aligned} \tag{2.2}$$

Definition 2.1 (Satisfied Cuts). For a set $S \subseteq V$ such that $u_0, v_0 \notin S$ and a spanning tree $T \subseteq E$ we say a vector $y : E \to \mathbb{R}_{\geq 0}$ satisfies S if $\delta(S)_T$ is even or $y(\delta(S)) \geq 1$.

To analyze our algorithm, we will see that the main challenge is to construct a (random) vector y that satisfies all cuts and $\mathbb{E}[c(y)] \leq (1/2 - \epsilon)OPT$, where OPT is the cost of an optimal solution.

2.3. Structure of Near Minimum Cuts

Lemma 2.1 (Oveis Gharan et al. 2011). *For $G = (V,E,x)$, let $A, B \subsetneq V$ be two crossing ϵ_A, ϵ_B near min cuts respectively. Then, $A \cap B, A \cup B, A \smallsetminus B, B \smallsetminus A$ are $\epsilon_A + \epsilon_B$ near min cuts.*

We prove the lemma only for $A \cap B$; the rest of the cases can be proved similarly. By submodularity,

$$\begin{aligned} x(\delta(A \cap B)) + x(\delta(A \cup B)) &\leq x(\delta(A)) + x(\delta(B)) \\ &\leq 4 + \epsilon_A + \epsilon_B. \end{aligned}$$

Because $x(\delta(A \cup B)) \geq 2$, we have $x(\delta(A \cap B)) \leq 2 + \epsilon_A + \epsilon_B$, as desired.

The following lemma is proved in Benczúr (1997).

Lemma 2.2 (Benczúr 1997, Lemma 5.3.5). *For $G=(V,E,x)$, let $A,B\subsetneq V$ be two crossing ϵ-near minimum cuts. Then,*

$$x(E(A\cap B,A-B)),x(E(A\cap B,B-A)),$$
$$x(E(\overline{A\cup B},A-B)),x(E(\overline{A\cup B},B-A))\geq(1-\epsilon/2).$$

Lemma 2.3. *For $G=(V,E,x)$, let $A,B\subsetneq V$ be two ϵ near min cuts such that $A\subsetneq B$. Then*

$$x(\delta(A)\cap\delta(B))=x(E(A,\overline{B}))\leq 1+\epsilon,\ \ and$$
$$x(\delta(A)\smallsetminus\delta(B))\geq 1-\epsilon/2.$$

Notice

$$2+\epsilon\geq x(\delta(A))=x(E(A,B\smallsetminus A))+x(E(A,\overline{B}))$$
$$2+\epsilon\geq x(\delta(B))=x(E(B\smallsetminus A,\overline{B}))+x(E(A,\overline{B}))\ .$$

Summing these up, we get

$$2x(E(A,\overline{B}))+x(E(A,B\smallsetminus A))+x(E(B\smallsetminus A,\overline{B}))$$
$$=2x(E(A,\overline{B}))+x(\delta(B\smallsetminus A))\leq 4+2\epsilon.$$

Because $B\smallsetminus A$ is nonempty, $x(\delta(B\smallsetminus A))\geq 2$, which implies the first inequality. To see the second one, let $C=B\smallsetminus A$ and note

$$4\leq x(\delta(A))+x(\delta(C))=2x(E(A,C))+x(\delta(B))$$
$$\leq 2x(E(A,C))+2+\epsilon,$$

which implies $x(E(A,C))\geq 1-\epsilon/2$.

2.4. Strongly Rayleigh Distributions and λ-Uniform Spanning Tree Distributions

Let $\mathcal{B}_E$ be the set of all probability measures on the Boolean algebra 2^E. Let $\mu\in\mathcal{B}_E$. The generating polynomial $g_\mu:\mathbb{R}[\{z_e\}_{e\in E}]$ of μ is defined as follows:

$$g_\mu(z):=\sum_S\mu(S)\prod_{e\in S}z_e.$$

We say μ is a strongly Rayleigh distribution if $g_\mu\neq 0$ over all $\{y_e\}_{e\in E}\in\mathbb{C}^E$, where $\mathrm{Im}(z_e)>0$ for all $e\in E$. We say μ is d-homogenous if for any $\lambda\in\mathbb{R}$, $g_\mu(\lambda\mathbf{z})=\lambda^d g_\mu(\mathbf{z})$. Strongly Rayleigh (SR) distributions were defined in Borcea et al. (2009) where it was shown any λ-uniform spanning tree distribution is strongly Rayleigh. In this section, we recall several properties of SR distributions proved in Borcea et al. (2009) and Oveis Gharan et al. (2011) that will be useful to us.

2.4.1. Closure Operations of SR Distributions.

SR distributions are closed under the following operations.

- **Projection.** For any $\mu\in\mathcal{B}_E$, and any $F\subseteq E$, the projection of μ onto F is the measure μ_F where for any $A\subseteq F$,

$$\mu_F(A)=\sum_{S:S\cap F=A}\mu(S).$$

- **Conditioning.** For any $e\in E$, $\{\mu|e\text{ out}\}$ and $\{\mu|e\text{ in}\}$.
- **Truncation.** For any integer $k\geq 0$ and $\mu\in\mathcal{B}_E$, truncation of μ to k, is the measure μ_k where for any $A\subseteq E$,

$$\mu_k(A)=\begin{cases}\dfrac{\mu(A)}{\sum_{S:|S|=k}\mu(S)} & \text{if } |A|=k\\ 0 & \text{otherwise.}\end{cases}$$

- **Product.** For any two disjoint sets E, F, and $\mu_E\in\mathcal{B}_E,\mu_F\in\mathcal{B}_F$ the product measure $\mu_{E\times F}$ is the measure where for any $A\subseteq E,B\subseteq F$, $\mu_{E\times F}(A\cup B)=\mu_E(A)\mu_F(B)$.

Throughout this paper, we will repeatedly apply the previous operations. We remark that SR distributions are *not* necessarily closed under truncation of a subset, that is, if we require exactly k elements from $F\subsetneq E$.

Because λ-uniform spanning tree distributions are special classes of SR distributions, if we perform any of the previous operations on a λ-uniform spanning tree distribution μ we get another SR distribution. Later, we see that by performing the following particular operations, we still have a λ-uniform spanning tree distribution (perhaps with a different λ).

2.4.2. Closure Operations of λ-uniform Spanning Tree Distributions.

For $G=(V,E)$, a spanning tree distribution $\mu\in\mathcal{B}_E$, and $T\sim\mu$, we have the following:

- **Conditioning**. For any $e\in E$, $\{\mu|e\notin T\},\{\mu|e\in T\}$.
- **Tree Conditioning**. For $S\subseteq V$, $\{\mu||E(S)\cap T|=|S|-1\}$, that is, T restricted to S is a tree. We will often just write S is a tree to denote such an event.

Arbitrary spanning tree distributions are not necessarily closed under truncation and projection. We remark that SR measures are also closed under an analogue of tree conditioning, that is, for a set $F\subseteq E$, let $k=\max_{S\in\text{supp }\mu}|S\cap F|$. Then, $\{\mu||S\cap F|=k\}$ is SR. However, if μ is a spanning tree distribution, we get an extra *independence* property. The following independence is crucial to several of our proofs.

Fact 2.3. *For a graph $G=(V,E)$, and a vector $\lambda(G):E\to\mathbb{R}_{\geq 0}$, let $\mu_{\lambda(G)}$ be the corresponding λ-uniform spanning tree distribution. Then for any $S\subsetneq V$,*

$$\{\mu_{\lambda(G)}|S\text{ is a tree}\}=\mu_{\lambda(G[S])}\times\mu_{\lambda(G/S)}.$$

Intuitively, this holds because in the max entropy distribution (recall a λ-uniform distribution maximizes entropy subject to matching the marginals of x), conditioned on S being a tree, any tree chosen inside S can be composed with any tree chosen on G/S to obtain a spanning tree on G. Therefore, to maximize the entropy these trees should be chosen independently. More formally, for any

$T_1 \in G[S]$ and $T_2 \in G/S$,

$$\begin{aligned}
&\mathbb{P}[T = T_1 \cup T_2 | S \text{ is a tree}] \\
&\quad = \frac{\lambda^{T_1}\lambda^{T_2}}{\sum_{T_1' \in G[S], T_2' \in G/S} \lambda^{T_1'}\lambda^{T_2'}} \\
&\quad = \frac{\lambda^{T_1}}{\sum_{T_1' \in G[S]} \lambda^{T_1'}} \cdot \frac{\lambda^{T_2}}{\sum_{T_2' \in G/S} \lambda^{T_2'}} \\
&\quad = \mathbb{P}_{T_1' \sim G[S]}[T_1' = T_1]\mathbb{P}_{T_2' \sim G/S}[T_2' = T_2],
\end{aligned}$$

giving independence.

2.4.3. Negative Dependence Properties. An *upward event*, $\mathcal{A}$, on 2^E is a collection of subsets of E that is closed under upward containment, that is, if $A \in \mathcal{A}$ and $A \subseteq B \subseteq E$, then $B \in \mathcal{A}$. Similarly, a *downward event* is closed under downward containment. An *increasing function* $f : 2^E \to \mathbb{R}$, is a function where for any $A \subseteq B \subseteq E$, we have $f(A) \le f(B)$. We also say $f : 2^E \to \mathbb{R}$ is a *decreasing function* if $-f$ is an increasing function. Therefore, an indicator of an upward event is an increasing function. For example, if E is the set of edges of a graph G, then the existence of a Hamiltonian cycle is an increasing function, and the three-colorability of G is a decreasing function.

Definition 2.2 (Negative Association). A measure $\mu \in \mathcal{B}_E$ is *negatively associated* if for any increasing functions $f, g : 2^E \to \mathbb{R}$, that depend on *disjoint* sets of edges,

$$\mathbb{E}_\mu[f] \cdot \mathbb{E}_\mu[g] \ge \mathbb{E}_\mu[f \cdot g].$$

It is shown in Borcea et al. (2009) and Feder and Mihail (1992) that strongly Rayleigh measures are negatively associated.

2.4.4. Stochastic Dominance. For two measures $\mu, \nu : 2^E \to \mathbb{R}_{\ge 0}$, we say $\mu \preceq \nu$ if there exists a *coupling* $\rho : 2^E \times 2^E \to \mathbb{R}_{\ge 0}$ such that

$$\sum_B \rho(A, B) = \mu(A), \ \forall A \in 2^E,$$

$$\sum_A \rho(A, B) = \nu(B), \ \forall B \in 2^E,$$

and for all A, B such that $\rho(A, B) > 0$, we have $A \subseteq B$ (coordinate-wise).

Theorem 2.1 (Borcea et al. 2009). *If μ is strongly Rayleigh and μ_k, μ_{k+1} are well defined, then $\mu_k \preceq \mu_{k+1}$.*

In the previous particular case, the coupling ρ satisfies the following: For any $A, B \subseteq E$ where $\rho(A, B) > 0$, $B \supseteq A$ and $|B \smallsetminus A| = 1$, that is, B has exactly one more element.

Let μ be a strongly Rayleigh measure on edges of G. Recall that for a set $A \subseteq E$, we write $A_T = |A \cap T|$ to denote the random variable indicating the number of edges in A chosen in a random sample T of μ. The following facts immediately follow from the negative association and stochastic dominance properties. We will use these facts repeatedly in this paper.

Fact 2.4 (Borcea et al. 2009, Theorems 4.8 and 4.19). *Let μ be any SR distribution on E, then for any $F \subset E$, and any integer k*

1. *(Negative Association) If $e \notin F$, then $\mathbb{P}_\mu[e | F_T \ge k] \le \mathbb{P}_\mu[e]$ and $\mathbb{P}_\mu[e | F_T \le k] \ge \mathbb{P}_\mu[e]$.*
2. *(Stochastic Dominance) If $e \in F$, then $\mathbb{P}_\mu[e | F_T \ge k] \ge \mathbb{P}_\mu[e]$ and $\mathbb{P}_\mu[e | F_T \le k] \le \mathbb{P}_\mu[e]$.*

The following fact is a direct consequence of the previous notation (see corollary 6.10 in Oveis Gharan et al. (2011)).

Fact 2.5. *Let μ be a homogenous SR distribution on E. Then,*

- *(Negative Association with Homogeneity) For any $A \subseteq E$, and any $B \subseteq \overline{A}$*

$$\mathbb{E}_\mu[B_T | A_T = 0] \le \mathbb{E}_\mu[B_T] + \mathbb{E}_\mu[A_T] \tag{2.3}$$

- *Suppose that μ is a spanning tree distribution. For $S \subseteq V$, let $q := |S| - 1 - \mathbb{E}_\mu[E(S)_T]$. For any $A \subseteq E(S), B \subseteq \overline{E(S)}$,*

$$\mathbb{E}_\mu[B_T] - q \le \mathbb{E}_\mu[B_T | S \text{ is a tree}] \le \mathbb{E}_\mu[B_T] \quad \text{(Negative association and homogeneity)}$$

$$\mathbb{E}_\mu[A_T] \le \mathbb{E}_\mu[A_T | S \text{ is a tree}] \le \mathbb{E}_\mu[A_T] + q \quad \text{(Stochastic dominance and tree)}$$

2.4.5. Rank Sequence. The *rank sequence* of μ is the sequence

$$\mathbb{P}[|S| = 0], \mathbb{P}[|S| = 1], \ldots, \mathbb{P}[|S| = m],$$

where $S \sim \mu$. Let $g_\mu(\mathbf{z})$ be the generating polynomial of μ. The *diagonal specialization* of μ is the univariate polynomial

$$\overline{g}_\mu(z) := g_\mu(z, z, \ldots, z).$$

Observe that $\overline{g}(.)$ is the generating polynomial of the rank sequence of μ. It follows that if μ is SR, then $\overline{g_\mu}$ is real rooted.

It is not hard to see that the rank sequence of μ corresponds to sum of independent Bernoullis if and only if (iff) $\overline{g_\mu}$ is real rooted. It follows that the rank sequence of an SR distributions has the law of a sum of independent Bernoullis. As a consequence, it follows (Hardy et al. 1952, Darroch 1964, Borcea et al. 2009) that the rank sequence of any strongly Rayleigh measure is log concave (see later for the definition), unimodal, and its mode differs from the mean by less than one.

Definition 2.3 (Log-Concavity (Borcea et al. 2009, Definition 2.8)). A real sequence $\{a_k\}_{k=0}^m$ is log-concave if $a_k^2 \ge a_{k-1}$

$\cdot a_{k+1}$ for all $1 \le k \le m-1$, and it is said to have no internal zeros if the indices of its nonzero terms form an interval (of nonnegative integers).

2.5. Sum of Bernoullis

In this section, we collect a number of properties of sums of Bernoulli random variables.

Definition 2.4 (Bernoulli Sum Random Variable). We say $BS(q)$ is a *Bernoulli-Sum* random variable if it has the law of a sum of independent Bernoulli random variables, say $B_1 + B_2 + \cdots + B_n$ for some $n \ge 1$, with $\mathbb{E}[B_1 + \cdots + B_n] = q$.

We start with the following theorem of Hoeffding.

Theorem 2.2 (Hoeffding 1956, Corollary 2.1). *Let $g : \{0, 1, \ldots, n\} \to \mathbb{R}$ and $0 \le q \le n$ for some integer $n \ge 0$. Let $B_1, \ldots, B_n$ be n independent Bernoulli random variables with success probabilities $p_1, \ldots, p_n$, where $\sum_{i=1}^n p_n = q$ that minimizes (or maximizes)*

$$\mathbb{E}[g(B_1 + \cdots + B_n)]$$

over all such distributions. Then, $p_1, \ldots, p_n \in \{0, x, 1\}$ for some $0 < x < 1$. In particular, if only m of p_is are nonzero and ℓ of p_is are 1, then the remaining $m-\ell$ are $\frac{q-\ell}{m-\ell}$.

Fact 2.6. *Let $B_1, \ldots, B_n$ be independent Bernoulli random variables each with expectation $0 \le p \le 1$. Then*

$$\mathbb{P}\left[\sum_i B_i \text{ even}\right] = \frac{1}{2}(1 + (1-2p)^n).$$

Note that

$$(p + (1-p))^n = \sum_{k=0}^{n} p^k (1-p)^{n-k} \binom{n}{k} \quad \text{and}$$

$$((1-p) - p)^n = \sum_{k=0}^{n} (-p)^k (1-p)^{n-k} \binom{n}{k}.$$

Summing them up we get

$$1 + (1-2p)^n = \sum_{0 \le k \le n,\, k \text{ even}} 2p^k (1-p)^{n-k} \binom{n}{k}.$$

Corollary 2.1. *Given a $BS(q)$ random variable with $0 < q \le 1.2$, then*

$$\mathbb{P}[BS(q) \text{ even}] \le \frac{1}{2}(1 + e^{-2q}).$$

First, if $q \le 1$, then by Hoeffding's theorem we can write $BS(q)$ as sum of n Bernoullis with success probability $p = q/n$. If $n = 1$, then the statement obviously holds. Otherwise, by the previous fact, we have (for some n),

$$\mathbb{P}[BS(q) \text{ even}] \le \frac{1}{2}(1 + (1-2p)^n)) \le \frac{1}{2}(1 + e^{-2q}),$$

where we used that $|1 - 2p| \le e^{-2p}$ for $p \le 1/2$.

Therefore, now assume $q > 1$. Write $BS(q)$ as the sum of n Bernoullis, each with success probabilities 1 or p. First assume we have no ones. Then, either we only have two nonzero Bernoullis with success probability $q/2$ in which case $\mathbb{P}[BS(q) \text{ even}] \le 0.6^2 + 0.4^2$ and we are done. Otherwise, $n \ge 3$ so $p \le 1/2$ and similar to the previous case we get $\mathbb{P}[BS(q) \text{ even}] \le \frac{1}{2}(1 + e^{-2q})$.

Finally, if $q > 1$ and one of the Bernoullis is always one, that is, $BS(q) = BS(q-1) + 1$, then we get

$$\mathbb{P}[BS(q) \text{ even}] = \mathbb{P}[BS(q-1) \text{ odd}] = \frac{1}{2}(1 - (1-2p)^{n-1}) \le 1/2,$$

where we used that $p \le 0.5$ (because $q \le 1.2$).

Lemma 2.4. *Let $p_0, \ldots, p_n$ be a log-concave sequence. If for some i, $\gamma p_i \ge p_{i+1}$ for some $\gamma < 1$, then,*

$$\sum_{j=k}^{n} p_j \le \frac{p_k}{1-\gamma}, \quad \forall k \ge i$$

$$\sum_{j=i+1}^{n} p_j \cdot j \le \frac{p_{i+1}}{1-\gamma}\left(i + 1 + \frac{\gamma}{1-\gamma}\right).$$

Because we have a log-concave sequence we can write

$$\frac{1}{\gamma} \le \frac{p_i}{p_{i+1}} \le \frac{p_{i+1}}{p_{i+2}} \le \cdots \tag{2.4}$$

Because all of the previous ratios are at least $1/\gamma$, for all $l \ge 1$, we can write

$$p_{i+l} \le \gamma^{l-1} p_{i+1} \le \gamma^l p_i.$$

Therefore, the first statement is immediate and the second one follows:

$$\sum_{j=i+1}^{n} p_j j \le \sum_{l=0}^{\infty} \gamma^l p_{i+1}(i + l + 1) = p_{i+1}\left(\frac{i+1}{1-\gamma} + \frac{\gamma}{(1-\gamma)^2}\right).$$

Corollary 2.2. *Let X be a $BS(q)$ random variable such that $\mathbb{P}[X = k] \ge 1 - \epsilon$ for some integer $k \ge 1, \epsilon < 1/10$. Then, $k(1-\epsilon) \le q \le k(1+\epsilon) + 3\epsilon$.*

The left inequality simply follows because $X \ge 0$. Because $\mathbb{P}[X = k+1] \le \epsilon$, we can apply Lemma 2.4 with $\gamma = \epsilon/(1-\epsilon)$ to get

$$\mathbb{E}[X | X \ge k+1]\mathbb{P}[X \ge k+1] \le \frac{\epsilon(1-\epsilon)}{1-2\epsilon}\left(k + 1 + \frac{\epsilon}{1-2\epsilon}\right).$$

Therefore,

$$q = \mathbb{E}[X] \le k(1-\epsilon) + \frac{\epsilon(1-\epsilon)}{1-2\epsilon}\left(k+1+\frac{\epsilon}{1-2\epsilon}\right)$$
$$\le k(1+\epsilon)+3\epsilon$$

as desired.

Fact 2.7. *For integers $k<t$ and $k-1 \le p \le k$,*

$$\prod_{i=1}^{k-1}(1-i/t)(1-p/t)^{t-k} \ge e^{-p}.$$

We show that the left-hand side (LHS) is a decreasing function of t. Because ln is monotone, it is enough to show

$$0 \ge \partial_t \ln(\text{LHS}) = \partial_t\left(\sum_{i=1}^{k-1}\ln(1-i/t)+(t-k)\ln(1-p/t)\right)$$
$$= \frac{1}{t^2}\sum_{i=1}^{k-1}\frac{1}{\frac{1}{i}-\frac{1}{t}} + \ln(1-p/t) + \frac{(t-k)p}{t(t-p)}.$$

Using $\sum_{i=1}^{k-2}\frac{1}{t^2/i-t} \le \int_0^{k-1}\frac{dx}{t^2/x-t} = -(k-1)/t - \ln(1-(k-1)/t)$, it is enough to show

$$0 \ge -\frac{k-1}{t} - \ln\left(1-\frac{k-1}{t}\right) + \ln(1-p/t) + \frac{(t-k)p}{t(t-p)} + \frac{1}{t^2\left(\frac{1}{k-1}-\frac{1}{t}\right)}$$
$$= \ln\frac{t-p}{t-k+1} + \frac{p-k}{t-p} + \frac{1}{t} + \frac{k-1}{t(t-k+1)}.$$

Rearranging, it is equivalent to show

$$\ln\left(1+\frac{p-k+1}{t-p}\right) \ge \frac{p-k}{t-p} + \frac{1}{t-k+1}.$$

Because $p>k-1$, using Taylor series of ln, to prove the previous expression, it is enough to show

$$\frac{p-k+1}{t-p} - \frac{(p-k+1)^2}{2(t-p)^2} \ge \frac{p-k}{t-p} + \frac{1}{t-k+1}.$$

This is equivalent to show

$$\frac{p-k+1}{(t-p)(t-k+1)} \ge \frac{(p-k+1)^2}{2(t-p)^2} \Leftrightarrow \frac{1}{t-k+1} \ge \frac{p-k+1}{2(t-p)}.$$

Finally, the latter holds because $(t-k+1)(p-k+1) \le (t-k+1) \le 2(t-p)$, where we use $t \ge k+1$ and $p \le k$.

Let $\text{Poi}(p,k) = e^{-p}p^k/k!$ be the probability that a Poisson random variable with rate p is exactly k; similarly, define $\text{Poi}(p, \le k), \text{Poi}(p, \ge k)$ as the probability that a Poisson with rate p is at most k or at least k.

Lemma 2.5. *Let X be a Bernoulli sum $BS(p)$ for some n. For any integer $k \ge 0$ such that $k-1<p<k+1$, the following holds true:*

$$\mathbb{P}[X=k] \ge \min_{0\le \ell \le p,k} \text{Poi}(p-\ell, k-\ell)\left(1-\frac{p-\ell}{k-\ell+1}\right)^{(p-k)_+},$$

where the minimum is over all nonnegative integers $\ell \le p,k$, and for $z\in\mathbb{R}$, $z_+ = \max\{z,0\}$.

Let $X = B_1 + \cdots + B_n$, where B_i is a Bernoulli. Applying Hoeffding's theorem, if ℓ of them have success probability 1, it suffices to prove a lower bound of $\text{Poi}(p-\ell, k-\ell)\left(1-\frac{p-\ell}{k-\ell+1}\right)^{(p-k)_+}$. Because without loss of generality none have success probability 1, it follows that each has success probability p/n. If $k \ge p$,

$$\mathbb{P}[X=k] = \binom{n}{k}\left(\frac{p}{n}\right)^k(1-p/n)^{n-k}$$
$$= \prod_{i=1}^{k-1}(1-i/n)\frac{p^k}{k!}(1-p/n)^{n-k} \ge \frac{p^k}{k!}e^{-p}$$
$$= \text{Poi}(p,k),$$

where in the inequality we used Fact 2.7 (also, if $n=k$, the inequality follows from Stirling's formula and that $p \ge k-1$). If $k<p<k+1$, then as previously shown:

$$\mathbb{P}[X=k] = \prod_{i=1}^{k-1}(1-i/n)\frac{p^k}{k!}(1-p/n)^{n-p}(1-p/n)^{p-k}$$
$$\underset{p\ge k}{\ge} \prod_{i=1}^{k-1}(1-i/n)\frac{p^k}{k!}(1-p/n)^{n-k}(1-p/n)^{p-k}$$
$$\ge \text{Poi}(p,k)(1-p/n)^{p-k},$$

where we used Fact 2.7 in the last inequality.

If we further know $X \ge a$ with probability 1, we can restrict ℓ in the statement to be in the interval $[a, \min(p,k)]$.

Lemma 2.6. *Let X be a Bernoulli sum $BS(p)$, where for some integer $k = \lceil p \rceil$, Then,*

$$\mathbb{P}[X \ge k] \ge \min_{0\le\ell\le p}\text{Poi}(p-\ell, \ge k-\ell),$$

where the minimum is over all nonnegative integers $\ell \le p$.

Suppose that X is a $BS(p)$ with n Bernoullis with probabilities $p_1,\ldots,p_n$. If $p-1<k-1<p$, by Hoeffding (1956, theorem 4, (25)),

$$\mathbb{P}[X \le k-1] \le \max_{0\le\ell<p}\sum_{i=0}^{k-1-\ell}\binom{n-\ell}{i}q^i(1-q)^{n-\ell-i}, \tag{2.5}$$

where $q = \frac{p-\ell}{n-\ell}$.

If Y is a $BS(p)$ with $m > n$ Bernoullis with probabilities $q_1, \ldots, q_m$, the same upper bound applies of course, with m replacing n. Also,

$$\max_{p_1 \ldots p_n} \mathbb{P}[X \le k-1] \le \max_{q_1, \ldots, q_m} \mathbb{P}[Y \le k-1],$$

because it is always possible to set $q_i = p_i$ for $i \le n$ and $q_j = 0$ for $j > n$.

Therefore, the upper bound in (2.5) obtained by taking the limit as n goes to infinity applies, from which it follows that

$$\mathbb{P}[X \le k-1] \le \max_{0 \le \ell < p} \sum_{i=0}^{k-1-\ell} \mathrm{Poi}(p-\ell, i),$$

and therefore

$$\mathbb{P}[X \ge k] \ge \min_{0 \le \ell < p} \mathrm{Poi}(p-\ell, \ge k-\ell).$$

2.6. Random Spanning Trees

Lemma 2.7. *Let $G = (V, E, x)$, and let μ be any distribution over spanning trees with marginals x. For any ϵ-near min cut $S \subseteq V$ (such that none of the endpoints of $e_0 = (u_0, v_0)$ are in S), we have*

$$\mathbb{P}_{T \sim \mu}[T \cap E(S) \text{ is tree}] \ge 1 - \epsilon/2.$$

Moreover, if μ is a max-entropy distribution with marginals x, then for any set of edges $A \subseteq E(S)$ and $B \subseteq E \smallsetminus E(S)$,

$$\mathbb{E}[A_T] \le \mathbb{E}[A_T | S \text{ is tree}] \le \mathbb{E}[A_T] + \epsilon/2, \mathbb{E}[B_T] - \epsilon/2 \le \mathbb{E}[B_T | S \text{ is tree}] \le \mathbb{E}[B_T].$$

First, observe that

$$\mathbb{E}[E(S)_T] = x(E(S)) \ge \frac{2|S| - x(\delta(S))}{2} \ge |S| - 1 - \epsilon/2,$$

where we used that because $u_0, v_0 \notin S$, and that for any $v \in S$, $\mathbb{E}[\delta(v)_T)] = x(\delta(v)) = 2$.

Let $p_S = \mathbb{P}[S \text{ is tree}]$. Then, we must have

$$|S| - 1 - (1 - p_S) = p_S(|S| - 1) + (1 - p_S)(|S| - 2) \ge \mathbb{E}[E(S)_T] \ge |S| - 1 - \epsilon/2.$$

Therefore, $p_S \ge 1 - \epsilon/2$.

The second part of the claim follows from Fact 2.5.

Corollary 2.3. *Let $A, B \subseteq V$ be disjoint sets such that A, $B, A \cup B$ are $\epsilon_A, \epsilon_B, \epsilon_{A\cup B}$-near minimum cuts w.r.t., x, respectively, where none of them contain endpoints of e_0. Then for any distribution μ of spanning trees on E with marginals x,*

$$\mathbb{P}_{T \sim \mu}[E(A, B)_T = 1] \ge 1 - (\epsilon_A + \epsilon_B + \epsilon_{A\cup B})/2.$$

By the union bound, with probability at least $1 - (\epsilon_A + \epsilon_B + \epsilon_{A\cup B})/2$, A, B, and $A \cup B$ are trees. However, this implies that we must have exactly one edge between A, B.

The following simple fact also holds by the union bound.

Fact 2.8. *Let $G = (V, E, x)$ and let μ be a distribution over spanning trees with marginals x. For any set $A \subseteq E$, we have*

$$\mathbb{P}_{T \sim \mu}[T \cap A = \emptyset] \ge 1 - x(A).$$

Lemma 2.8. *Let $G = (V, E, x)$, and let μ be a λ-uniform random spanning tree distribution with marginals x. For any edge $e = (u, v)$ and any vertex $w \ne u, v$ we have*

$$\mathbb{E}[W_T | e \notin T] \le \mathbb{E}[W_T] + \mathbb{P}[w \in P_{u,v} | e \notin T] \cdot \mathbb{P}[e \in T],$$

where $W_T = |T \cap \delta(w)|$ and for a spanning tree T and vertices $u, v \in V$, $P_{u,v}(T)$ is the set of vertices on the path from u to v in T.

Define $E' = E \smallsetminus \{e\}$. Let $\mu' = \mu|_{E'}$ be μ projected on all edges except e. Define $\mu_{in} = \mu'_{n-2}$ (corresponding to e in the tree) and $\mu_{out} = \mu'_{n-1}$ (corresponding to e out of the tree). Observe that any tree T has positive measure in exactly one of these distributions.

By Theorem 2.1, $\mu_{in} \preceq \mu_{out}$ so there exists a coupling $\rho : 2^{E'} \times 2^{E'}$ between them such that for any T_{in}, T_{out} such that $\rho(T_{in}, T_{out}) > 0$, the tree T_{out} has exactly one more edge than T_{in}. Also, observe that T_{out} is always a spanning tree whereas $T_{in} \cup \{e\}$ is a spanning tree. The added edge (i.e., the edge in $T_{out} \smallsetminus T_{in}$) is always along the unique path from u to v in T_{out}.

For intuition for the rest of the proof, observe that if w is not on the path from u to v in T_{out}, then the same set of edges is incident to w in both T_{in} and T_{out}. Therefore, if w is almost never on the path from u to v, the distribution of W_T is almost independent of e. On the other hand, whenever w is on the path from u to v, then in the worst case, we may replace e with one of the edges incident to w, so conditioned on e out, W_T increases by at most the probability that e is in the tree.

Say x_e is the marginal of e. Then,

$$\begin{aligned}\mathbb{E}[W_T] &= \mathbb{E}[W_T | e \notin T](1 - x_e) + \mathbb{E}[W_T | e \in T] x_e \\ &= \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out}) W_o (1 - x_e) \\ &\quad + \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out}) W_i x_e \\ &= \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out})((1 - x_e) W_o + x_e W_i),\end{aligned} \tag{2.6}$$

where we write W_i/W_o instead of $W_{T_{in}}/W_{T_{out}}$

$$\begin{aligned}
\mathbb{E}[W_T|e \notin T] &= \sum_{T_{in},T_{out}} \rho(T_{in},T_{out})W_o \\
&= \sum_{T_{in},T_{out}:w\in P_{u,v}(T_{out})} \rho(T_{in},T_{out})W_o \\
&\quad + \sum_{T_{in},T_{out}:w\notin P_{u,v}(T_{out})} \rho(T_{in},T_{out})W_o \\
&\leq \sum_{T_{in},T_{out}:w\in P_{u,v}(T_{out})} \rho(T_{in},T_{out})(x_e(W_i+1) \\
&\quad + (1-x_e)W_o) + \sum_{T_{in},T_{out}:w\notin P_{u,v}(T_{out})} \rho(T_{in},T_{out}) \\
&\qquad (x_e W_i + (1-x_e)W_o) \\
&= \mathbb{E}[W_T] + \sum_{T_{in},T_{out}:w\in P_{u,v}(T_{out})} \rho(T_{in},T_{out})x_e \\
&= \mathbb{E}[W_T] + \sum_{T_{out}:w\in P_{u,v}(T_{out})} \mu_{out}(T_{out})x_e \\
&= \mathbb{E}[W_T] + \mathbb{P}[w\in P_{u,v}|e \text{ out}]\cdot\mathbb{P}[e \text{ in}],
\end{aligned}$$

where in the inequality, we used the following: When $w \notin P_{u,v}(T_{out})$, we have $W_i = W_o$ and when $w \in P_{u,v}(T_{out})$, we have $W_o \leq W_i + 1$. Finally, in the third to last equality, we used (2.6).

Lemma 2.9. *Let $G=(V,E,x)$, and let μ be a λ-uniform spanning tree distribution with marginals x. For any pair of edges $e=(u,v), f=(v,w)$ such that $|\mathbb{P}[e]-1/2|, |\mathbb{P}[f]-1/2| < \epsilon$ (see Figure 1), if $\epsilon < 1/1{,}000$, then*

$$\mathbb{E}[W_T|e\notin T] + \mathbb{E}[U_T|f\notin T] \leq \mathbb{E}[W_T + U_T] + 0.81,$$

where $U=\delta(u)_{-e}$ and $W=\delta(w)_{-f}$.

All probabilistic statements are with respect to ν, so we drop the subscript. First, by Lemma 2.8 and negative association, we can write

$$\begin{aligned}
\mathbb{E}[W_T|e\notin T] &\leq \mathbb{E}[W_T] + \mathbb{P}[w\in P_{u,v}|e\notin T]\mathbb{P}[e\in T] \\
&\leq \mathbb{E}[W_T] + \mathbb{P}[w\in P_{u,v}\wedge e\notin T] + 2\epsilon.
\end{aligned}$$

The lemma only implies $\mathbb{E}[\delta(w)_T|e\notin T] \leq \mathbb{E}[\delta(w)_T] + \mathbb{P}[w\in P_{u,v}|e\notin T]\mathbb{P}[e\in T]$. To derive the first inequality, we also exploit negative association that asserts that the marginal of every edge only goes up under $e\notin T$, so any subset of $\delta(w)$ (in particular W) also goes up by at most $\mathbb{P}[e\notin T\wedge w\in P_{u,v}]$. Also, the second inequality uses $\mathbb{P}[e\in T]\leq\mathbb{P}[e\notin T]+2\epsilon$. Using a similar inequality for U_T, to prove the lemma, it is enough to show that

$$\mathbb{P}[w\in P_{u,v}\wedge e\notin T] + \mathbb{P}[u\in P_{v,w}\wedge f\notin T] \leq 0.806,$$

or that when this inequality fails, a different argument yields the lemma.

The main observation is that in any tree it cannot be that both u is on the $v-w$ path, and w is on the $u-v$ path. Therefore,

$$\mathbb{P}[u\in P_{v,w}|e,f\notin T] + \mathbb{P}[w\in P_{u,v}|e,f\notin T] \leq 1.$$

Therefore, we have

$$\begin{aligned}
&\mathbb{P}[e\notin T\wedge w\in P_{u,v}] + \mathbb{P}[f\notin T\wedge u\in P_{v,w}] \\
&\leq \mathbb{P}[e,f\notin T\wedge w\in P_{u,v}] + \mathbb{P}[e\notin T, f\in T] \\
&\quad + \mathbb{P}_\nu[e,f\notin T\wedge u\in P_{v,w}] + \mathbb{P}[f\notin T, e\in T] \\
&\leq \mathbb{P}[e,f\notin T] + \mathbb{P}[e\notin T,f\in T] + \mathbb{P}[f\notin T, e\in T] \\
&= 1-\mathbb{P}[e,f\in T].
\end{aligned}$$

It remains to upper bound the right-hand side. Let $\alpha = \mathbb{P}[f\in T|e\notin T]$. Observe that

$$\begin{aligned}
\mathbb{P}[e,f\in T] &= \mathbb{P}[f\in T] - \mathbb{P}[f\in T, e\notin T] \\
&\geq 1/2-\epsilon-(1/2+\epsilon)\alpha.
\end{aligned}$$

If $\alpha\leq 0.6$, then $\mathbb{P}[e,f\in T]\geq 0.198$ (using $\epsilon<0.001$) and the claim follows. Otherwise, $\mathbb{P}[f|e\notin T]\geq 0.6$. Similarly, $\mathbb{P}[e|f\notin T]\geq 0.6$. However, by negative association,

$$\begin{aligned}
\mathbb{E}[W_T|e\notin T] &\leq \mathbb{E}[W_T] + \mathbb{P}[e] - (\mathbb{P}[f|e\notin T]-\mathbb{P}[f]) \\
&\leq \mathbb{E}[W_T] + 2\epsilon + 0.4 \leq \mathbb{E}[W_T] + 0.405,
\end{aligned}$$

and similarly, $\mathbb{E}[U_T|f\notin T]\leq\mathbb{E}[U_T]+0.405$, so the claim follows.

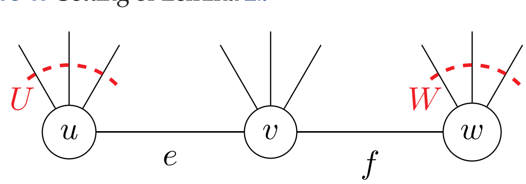
3. Overview of Proof

As alluded to earlier, the crux of the proof of Theorem 1.1 is to show that the expected cost of the minimum cost matching on the odd degree vertices of the sampled tree is at most $OPT(1/2-\epsilon)$. We do this by showing the existence of a cheap feasible O-join solution to (2.2).

First, recall that if we only wanted to get an O-join solution of value at most $OPT/2$, to satisfy all cuts, it is enough to set $y_e := x_e/2$ for each edge Wolsey (1980). To do better, we want to take advantage of the fact that we only need to satisfy a constraint in the O-join for S when $\delta(S)_T$ is odd. Here, we are aided by the fact that the sampled tree is likely to have many even cuts because it is drawn from a strong Rayleigh distribution.

If an edge e is exclusively on even cuts, then y_e can be reduced below $x_e/2$. This, more or less, was the approach in Oveis Gharan et al. (2011) for graphic TSP, where it was shown that a constant fraction of LP edges will be exclusively on *even* near min cuts with constant

Figure 1. Setting of Lemma 2.9



probability. The difficulty in implementing this approach in the metric case comes from the fact that a high cost edge can be on many cuts and it may be exceedingly unlikely that *all* of these cuts will be even simultaneously. Overall, our approach to addressing this is to start with $y_e := x_e/2$ and then modify it with a random[4] *slack vector* $s : E \to \mathbb{R}$: When certain special (few) cuts that e is on are even, we let $s_e = -x_e\beta$ (for a carefully chosen constant $\beta > 0$); for other cuts that contain e, whenever they are odd, we will increase the slack of other edges on that cut to satisfy them. The bulk of our effort is to show that we can do this while guaranteeing that $\mathbb{E}[s_e] < -\epsilon\beta x_e$ for some $\epsilon > 0$.

By carefully choosing β smaller than η, we do not need to worry about the reduction breaking a constraint for any cut S such that $x(\delta(S)) > 2(1+\eta)$. In particular, if we choose $\beta \le \eta/4.1$, any such cut is always satisfied, even if every edge in $\delta(S)$ is decreased and no edge is increased.

Let OPT be the optimum TSP tour, that is, a Hamiltonian cycle, with set of edges E^*; throughout the paper, we write e^* to denote an edge in E^*. To bound the expected cost of the O-join for a random spanning tree $T \sim \mu_\lambda$, we also construct a random slack vector $s^* : E^* \to \mathbb{R}_{\ge 0}$ such that $(x + OPT)/4 + s + s^*$ is a feasible for Equation (2.2) with probability 1. In Section 3.1, we explain how to use s^* to satisfy all but a linear number of near min cuts.

Theorem 3.1 (Main Technical Theorem). *Let x^0 be a solution of the Held-Karp relaxation with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E. Let $z := (x + OPT)/2$, $\eta \le 10^{-12}$, $\beta > 0$, and let μ be the max-entropy distribution with marginals x. Also, let E^* denote the support of OPT. There are two functions $s : E_0 \to \mathbb{R}$ and $s^* : E^* \to \mathbb{R}_{\ge 0}$ (as functions of $T \sim \mu$), such that*

(i) *For each edge $e \in E$, $s_e \ge -x_e\beta$.*

(ii) *For each η-near-min cut S of z, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \ge 0$.*

(iii) *For every OPT edge e^*, $\mathbb{E}[s^*_{e^*}] \le 218\eta\beta$ and for every LP edge $e \ne e_0$, $\mathbb{E}[s_e] \le -\frac{1}{3}x_e\epsilon_P\beta$ for $\epsilon_P = 3.12 \cdot 10^{-16}$ (defined in (7.4)).*

In the next section, we explain the main ideas needed to prove this technical theorem. However, first, we show how our main theorem follows readily from Theorem 3.1.

Proof of Theorem 1.1. Let x^0 be an extreme point solution of the Held-Karp relaxation, with support E_0 and let x be x^0 restricted to E. By Fact 2.1, x is in the spanning tree polytope. For $\mu = \mu_{\lambda^*}$ the max entropy distribution with marginals x and $\beta > 0$ a parameter we choose, let s, s^* be as defined in Theorem 3.1. We will define $y : E_0 \to \mathbb{R}_{\ge 0}$ and $y^* : E^* \to \mathbb{R}_{\ge 0}$. Let

$$y_e = \begin{cases} x_e/4 + s_e & \text{if } e \in E \\ \infty & \text{if } e = e_0; \end{cases}$$

we also let $y^*_{e^*} = 1/4 + s^*_{e^*}$ for any edge $e^* \in E^*$. We will show that $y + y^*$ is a feasible solution[5] to (2.2). First, observe that for any S where $e_0 \in \delta(S)$, we have $y(\delta(S)) + y^*(\delta(S)) \ge 1$. Otherwise, we assume $u_0, v_0 \notin S$. If S is an η-near min cut w.r.t., z and $\delta(S)_T$ are odd, then by property (ii) of Theorem 3.1, we have

$$y(\delta(S)) + y^*(\delta(S)) = \frac{z(\delta(S))}{2} + s(\delta(S)) + s^*(\delta(S)) \ge 1.$$

On the other hand, if S is not an η-near min cut (w.r.t., z):

$$\begin{aligned} y(\delta(S)) + y^*(\delta(S)) &\ge \frac{z(\delta(S))}{2} - \beta x(\delta(S)) \\ &\ge \frac{z(\delta(S))}{2} - \beta 2(z(\delta(S)) - 1) \\ &\ge z(\delta(S))(1/2 - 2\beta) + 2\beta \\ &\ge (2+\eta)(1/2 - 2\beta) + 2\beta, \end{aligned}$$

where in the first inequality we used property (i) of Theorem 3.1 that says that $s_e \ge x_e\beta$ with probability 1 for all LP edges and that $s^*_{e^*} \ge 0$ with probability 1. In the second inequality, we used that $z = (x + OPT)/2$, so, because $OPT \ge 2$ across any cut, $x(\delta(S)) \le 2(z(\delta(S)) - 1)$. Finally, if we choose

$$\beta = \eta/4.1, \tag{3.1}$$

then the RHS is at least one, so $y + y^*$ is a feasible O-join solution.

Finally, using $c(e_0) = 0$ and part (iii) of Theorem 3.1,

$$\begin{aligned} \mathbb{E}[c(y) + c(y^*)] &= OPT/4 + c(x)/4 + \mathbb{E}[c(s) + c(s^*)] \\ &\le OPT/4 + c(x)/4 + 218\eta\beta OPT - \frac{1}{3}\epsilon_P\beta c(x) \\ &\le \left(1/2 - \frac{1}{6}\epsilon_P\beta\right) OPT, \end{aligned}$$

choosing η such that

$$218\eta = \frac{1}{6}\epsilon_P \tag{3.2}$$

and using $c(x) \le OPT$.

Now, we are ready to bound the approximation factor of our algorithm. First, because x^0 is an extreme point solution of the Held-Karp relaxation, $\min_{e \in E_0} x^0_e \ge \frac{1}{n!}$. Therefore, by Theorem 1.3, in polynomial time, we can find $\lambda : E \to \mathbb{R}_{\ge 0}$ such that for any $e \in E$, $\mathbb{P}_{\mu_\lambda}[e] \le x_e(1+\delta)$ for some δ that we fix later. It follows that

$$\sum_{e \in E} |\mathbb{P}_\mu[e] - \mathbb{P}_{\mu_\lambda}[e]| \le n\delta.$$

By stability of maximum entropy distributions (Straszak and Vishnoi 2019, theorem 4, and references therein), we have that $\|\mu - \mu_\lambda\|_1 \le O(n^4\delta) =: q$. Therefore, for some

$\delta \ll n^{-4}$, we get $\|\mu - \mu_\lambda\|_1 = q \le \frac{\epsilon_P \beta}{100}$. That means that

$$\mathbb{E}_{T\sim\mu_\lambda}[\text{min cost matching}] \le \mathbb{E}_{T\sim\mu}[c(y) + c(y^*)] + q(OPT/2) \le \left(\frac{1}{2} - \frac{1}{6}\epsilon_P\beta + \frac{\epsilon_P\beta}{100}\right)OPT,$$

where we used that for any spanning tree the cost of the minimum cost matching on odd degree vertices is at most $OPT/2$. Finally, because $\mathbb{E}_{T\sim\mu_\lambda}[c(T)] \le OPT(1+\delta), \epsilon_P = 3.12 \cdot 10^{-16}$, and $\beta = \eta/4.1 = \epsilon_p/5{,}362.8$ (from (3.2)) we get a $3/2 - 3\cdot 10^{-36}$ approximation algorithm for TSP.

3.1. Ideas Underlying Proof of Theorem 3.1

The first step of the proof is to show that it suffices to construct a slack vector s for a "cactus-like" structure of near min cuts that we call a *hierarchy*. Informally, a hierarchy $\mathcal{H}$ is a laminar family of min cuts,[6] consisting of two types of cuts: *triangle cuts* and *degree cuts*. A triangle S is the union of two min cuts X and Y in $\mathcal{H}$ such that $x(E(X,Y)) = 1$. Figure 2 provides an example of a hierarchy with three triangles.

We will refer to the set of edges $E(X,\overline{S})$ (respectively, $E(Y,\overline{S})$) as A (respectively, B) for a triangle cut S. In addition, we say a triangle cut S is *happy* if A_T and B_T are both odd. All nontriangle cuts are called degree cuts. A degree cut S is *happy* if $\delta(S)_T$ is even.

Theorem 3.2 (Main Payment Theorem (Informal)). *Let $G = (V,E,x)$ for LP solution x and let μ be the max-entropy distribution with marginals x and $\beta > 0$. Given a hierarchy $\mathcal{H}$, there is a slack vector $s : E \to \mathbb{R}$ such that*

(i) *For each edge $e \in E$, $s_e \ge -x_e\beta$.*
(ii) *For each cut $S \in \mathcal{H}$ if S is not happy, then $s(\delta(S)) \ge 0$.*
(iii) *For every LP edge $e \ne e_0$, $\mathbb{E}[s_e] \le -\beta\epsilon_P x_e$ for $\epsilon_P > 0$.*

In the following section, we discuss how to prove this theorem. Here we explain at a high level how to define the hierarchy and reduce Theorem 3.1 to this theorem. The details are in Section 4.

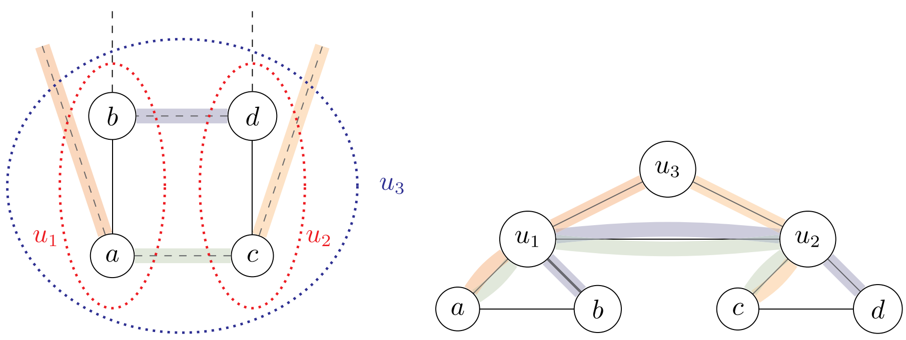
First, observe that, given Theorem 3.2, cuts in $\mathcal{H}$ will automatically satisfy (ii) of Theorem 3.1. The approach we take to satisfying all other cuts is to introduce additional slack, the vector s^*, on OPT edges.

Consider the set of all near-min cuts of z, where $z := (x + OPT)/2$. Starting with z rather than x allows us to restrict attention to a significantly *more structured* collection of near-min cuts. The key observation here is that in OPT, all min cuts have value 2, and any non-min cut has value of *at least* 4. Therefore, averaging x with OPT guarantees that every η-near min cut of z must consist of a *contiguous sequence of vertices (an interval) along the OPT cycle*. Moreover, each of these cuts is a 2η-near min cut of x. Arranging the vertices in the OPT cycle around a circle, we identify every such cut with the interval of vertices that does not contain (u_0, v_0). Also, we say that a cut is crossed on both sides if it is crossed on the left and on the right.

To ensure that any cut S that is *crossed on both sides* is satisfied, we first observe that S is odd with probability $O(\eta)$. To see this, let S_L and S_R be the cuts crossing S on the left and right with minimum intersection with S and consider the two (bad) events $\{E(S\cap S_L, S_L \smallsetminus S))_T \ne 1\}$ and $\{E(S \cap S_R, S_R \smallsetminus S))_T \ne 1\}$. Recall that if A, B and $A\cup B$ are all near-min cuts, then $\mathbb{P}[E(A,B)_T \ne 1] = O(\eta)$ (see Corollary 2.3). Applying this fact to the two aforementioned bad events implies that each of them has probability $O(\eta)$. Therefore, we will let the two OPT edges in $\delta(S)$ be responsible for these two events; that is, we will increase the slack s^* on these two OPT edges by $O(\eta)$ when the respective bad events happens. This gives $\mathbb{E}[s^*(e^*)] = O(\eta^2)$ for each OPT edge e^*. As we will see, this simple step will reduce the number of near-min cuts of z that we need to worry about satisfying to $O(n)$.

Next, we consider the set of near-min cuts of z that are crossed on at most one side. Partition these into maximal

Figure 2. Example of Part of a Hierarchy with Three Triangles



Notes. The graph on the left shows part of a feasible LP solution where dashed (and sometimes colored) edges have fraction 1/2 and solid edges have fraction 1. The dotted ellipses on the left show the min cuts u_1, u_2, u_3 in the graph. (Each vertex is also a min cut). On the right is a representation of the corresponding hierarchy. Triangle u_1 corresponds to the cut $\{a,b\}$, u_2 corresponds to $\{c,d\}$ and u_3 corresponds to $\{a,b,c,d\}$. For example, the edge (a,c), represented in green, is in $\delta(u_1)$, $\delta(u_3)$, and inside u_3. For triangle u_1, we have $A = \delta(a) \smallsetminus (a,b)$ and $B = \delta(b) \smallsetminus (b,d)$.

connected components of crossing cuts. Each such component corresponds to an interval along the OPT cycle and, by definition, these intervals form a laminar family.

A single connected component $\mathcal{C}$ of at least two crossing cuts is called a *polygon*. We prove the following structural theorem about the polygons induced by z.

Theorem 3.3 (Polygons Look Like Cycles (Informal Version of Theorem 4.3)). *Given a connected component $\mathcal{C}$ of near-min cuts of z that are crossed on one side, consider the coarsest partition of vertices of the OPT cycle into a sequence $a_1, \ldots, a_{m-1}$ of sets called atoms (together with a_0 which is the set of vertices not contained in any cut of $\mathcal{C}$). Then*

- *Every cut in $\mathcal{C}$ is the union of some number of consecutive atoms in $a_1, \ldots, a_{m-1}$.*
- *For each i such that $0 \le i < m-1$, $x(E(a_i, a_{i+1})) \approx 1$ and similarly $x(E(a_{m-1}, a_0)) \approx 1$.*
- *For each $i > 0$, $x(\delta(a_i)) \approx 2$.*

The main observation used to prove Theorem 3.3 is that the cuts in $\mathcal{C}$ crossed on one side can be partitioned into two laminar families $\mathcal{L}$ and $\mathcal{R}$, where $\mathcal{L}$ (respectively, $\mathcal{R}$) is the set of cuts crossed on the left (respectively, right). This immediately implies that $|\mathcal{C}|$ is linear in m. Because cuts in $\mathcal{L}$ cannot cross each other (and similarly for $\mathcal{R}$), the proof boils down to understanding the interaction between $\mathcal{L}$ and $\mathcal{R}$.

The approximations in Theorem 3.3 are correct up to $O(\eta)$. Using additional slack in *OPT*, at the cost of an additional $O(\eta^2)$ for edge, we can treat these approximate equations as if they are exact. Observe that if $x(E(a_i, a_{i+1})) = 1$, and $x(\delta(a_i)) = x(\delta(a_{i+1})) = 2$ for $1 \le i \le m-2$, then with probability 1, $E(a_i, a_{i+1})_T = 1$. Therefore, any cut in $\mathcal{C}$ that does not include a_1 or a_{m-1} is even with probability 1. The cuts in $\mathcal{C}$ that contain a_1 are even precisely[7] when $E(a_0, a_1)_T$ is odd and similarly the cuts in $\mathcal{C}$ that contain a_{m-1} are even when $E(a_0, a_{m-1})_T$ is odd. These observations are what allow us to imagine that each polygon is a triangle, that is, assume $m = 3$. (Often it is convenient to look at the event in which $E(a_0, a_1)_T = 1$ and $E(a_0, a_{m-1})_T = 1$ becuase these are simple criteria that imply that all cuts in $\mathcal{C}$ are even.)

The hierarchy $\mathcal{H}$ is the set of all η-near min cuts of z that are not crossed at all (these will be the degree cuts), together with a triangle for every polygon. In particular, for a connected component $\mathcal{C}$ of size more than one, the corresponding triangle cut is $a_1 \cup \cdots \cup a_{m-1}$, with $A = E(a_0, a_1)$ and $B = E(a_0, a_{m-1})$. Observe that from the previous discussion, when a triangle cut is happy, then all the cuts in the corresponding polygon $\mathcal{C}$ are even.

Summarizing, we show that if we can construct a good slack vector s for a hierarchy of degree cuts and triangles, then there is a nonnegative slack vector s^* that satisfies all near-minimum cuts of z not represented in the hierarchy while maintaining slack for each OPT edge e^* such that $\mathbb{E}[s^*(e^*)] = O(\eta^2)$.

3.1.1. Remarks. The reduction that we sketched previously only uses the fact that μ is an arbitrary distribution of spanning trees with marginals x and not necessarily a maximum-entropy distribution.

We also observe that to prove Theorem 1.1, we crucially used that $28\eta \ll \epsilon$. This forces us to take η very small, which is why we get only a "very slightly" improved approximation algorithm for TSP. Furthermore, because we use OPT edges in our construction, we do not get a new upper bound on the integrality gap. We leave it as an open problem to find a reduction to the "cactus" case that does not involve using a slack vector for OPT (or a completely different approach).

3.2. Proof Ideas for Theorem 3.2

We now address the problem of constructing a good slack vector s for a hierarchy of degree cuts and triangle cuts. For each LP edge f, consider the lowest cut in the hierarchy that contains both endpoints of f. We call this cut $\mathsf{p}(f)$. If $\mathsf{p}(f)$ is a degree cut, then we call f a *top edge*, and otherwise, it is a *bottom edge*.[8] We will see that bottom edges are easier to deal with, so we start by discussing the slack vector s for top edges.

Let S be a degree cut and let $\mathbf{e} = (u, v)$ (where u and v are children of S in $\mathcal{H}$) be the set of all top edges $f = (u', v')$ such that $u' \in u$ and $v' \in v$. We call $\mathbf{e}$ a *top edge bundle* and say that u and v are the *top cuts* of each $f \in \mathbf{e}$. We will also sometimes say that $\mathbf{e} \in S$.

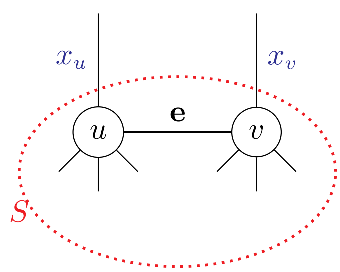
Ideally, our plan is to reduce the slack of every edge $f \in \mathbf{e}$ when it is *happy*, that is, both of its top cuts are even in T. Specifically, we will set $s_f := -\eta x_f$ when $\delta(u)_T$ and $\delta(v)_T$ are even. When this happens, we say that f is *reduced*, and refer to the event $\{\delta(u)_T, \delta(v)_T \text{ even}\}$ as the *reduction event* for f. Because this latter event does not depend on the actual endpoints of f, we view this as a simultaneous reduction of $s_{\mathbf{e}}$.

Now consider the situation from the perspective of the degree cut u (where $\mathsf{p}(u) = S$) and consider any incident edge bundle in S, for example, $\mathbf{e} = (u, v)$. Either its top cuts are both even and $s_{\mathbf{e}} := -\eta x_{\mathbf{e}}$, or they are not even, because, for example, $\delta(u)_T$ is odd. In this latter situation, edges in $\delta^{\uparrow}(u) := \delta(u) \cap \delta(S)$ might have been reduced (because *their* top two cuts are even), which a priori could leave $\delta(u)$ unsatisfied. In such a case, we *increase* $s_{\mathbf{e}}$ for edge bundles in $\delta^{\rightarrow}(u) := \delta(u) \smallsetminus \delta(S)$ to compensate for this reduction. Our main goal is then to prove is that for any edge bundle its expected reduction is greater than its expected increase. The next example shows this analysis in an ideal setting.

Example 3.1 (Simple Case). Fix a top edge bundle $\mathbf{e} = (u, v)$ with $\mathsf{p}(\mathbf{e}) = S$. Let $x_u := x(\delta^{\uparrow}(u))$ and let $x_v := x(\delta^{\uparrow}(v))$. Suppose we have constructed a (fractional) *matching* between edges whose top two cuts are children of S in $\mathcal{H}$ and the edges in $\delta(S)$, and this matching satisfies the following three conditions: (a) $\mathbf{e} = (u, v) \in S$ is matched (only) to edges going higher from its top two cuts (i.e., to edges in $\delta^{\uparrow}(u)$ and $\delta^{\uparrow}(v)$), (b) $\mathbf{e}$ is matched to an $m_{\mathbf{e},u}$ fraction of every edge in $\delta^{\uparrow}(u)$ and to an $m_{\mathbf{e},v}$ fraction of each edge in $\delta^{\uparrow}(v)$, where

$$m_{\mathbf{e},u} + m_{\mathbf{e},v} = x_{\mathbf{e}},$$

and (c) the fractional value of edges in $\delta^{\rightarrow}(u) := \delta(u) \smallsetminus \delta^{\uparrow}(u)$ matched to edges in $\delta^{\uparrow}(u)$ is equal to x_u. That is, for each $u \in S$, $\sum_{\mathbf{f} \in \delta^{\rightarrow}(u)} m_{\mathbf{f},u} = x_u$.



The plan is for $\mathbf{e} \in S$ to be tasked with part of the responsibility for *fixing the cuts* $\delta(u)$ and $\delta(v)$ when they are odd and edges going higher are reduced. Specifically, $s_{\mathbf{e}}$ is increased to compensate for an $m_{\mathbf{e},u}$ fraction of the reductions in edges in $\delta^{\uparrow}(u)$ when $\delta(u)_T$ is odd (and similarly for reductions in v). Thus,

$$\begin{aligned}\mathbb{E}[s_{\mathbf{e}}] = &-\mathbb{P}[\mathbf{e} \text{ reduced}]\eta x_{\mathbf{e}} \\ &+ m_{\mathbf{e},u} \sum_{g \in \delta^{\uparrow}(u)} \mathbb{P}[\delta(u)_T \text{ odd} \mid g \text{ reduced}] \\ &\mathbb{P}[g \text{ reduced}]\eta \frac{x_g}{x(\delta^{\uparrow}(u))} \\ &+ m_{\mathbf{e},v} \sum_{g \in \delta^{\uparrow}(v)} \mathbb{P}[\delta(v)_T \text{ odd} \mid g \text{ reduced}] \\ &\mathbb{P}[g \text{ reduced}]\eta \frac{x_g}{x(\delta^{\uparrow}(v))}. \end{aligned} \tag{3.3}$$

We will lower bound $\mathbb{P}[\delta(u)_T \text{even} \mid g \text{ reduced}]$. We can write this as

$$\mathbb{P}[\delta^{\rightarrow}(u)_T \text{ and } \delta^{\uparrow}(u)_T \text{ have same parity} \mid g \text{ reduced}].$$

Unfortunately, we do not currently have a good handle on the parity of $\delta^{\uparrow}(u)_T$ conditioned on g reduced. However, we can use the following simple but crucial property: Because $x(\delta(S)) = 2$, by Lemma 2.7, T consists of two independent trees, one on S and one on $V \smallsetminus S$, each with the corresponding marginals of x. Therefore, we can write

$$\begin{aligned}&\mathbb{P}[\delta(u)_T \text{ even} \mid g \text{ reduced}] \\ &\geq \min(\mathbb{P}[(\delta^{\rightarrow}(u))_T \text{ even}], \mathbb{P}[(\delta^{\rightarrow}(u))_T \text{ odd}]).\end{aligned}$$

This gives us a reasonable bound when $\epsilon \leq x_u, x_v \leq 1 - \epsilon$ because $x(\delta(u)) = x(\delta(v)) = 2$, by the SR property, $(\delta^{\rightarrow}(u))_T$ (and similarly $(\delta^{\rightarrow}(v))_T$) is the sum of Bernoulis with expectation in $[1+\epsilon, 2-\epsilon]$. From this, it follows that

$$\min(\mathbb{P}[(\delta^{\rightarrow}(u))_T \text{ even}], \mathbb{P}[(\delta^{\rightarrow}(u))_T \text{ odd}]) = \Omega(\epsilon).$$

We can therefore conclude that $\mathbb{P}[\delta(u)_T \text{ odd} \mid g \text{ reduced}] \leq 1 - O(\epsilon)$.

The rest of the analysis of this special case follows from (a) the fact that our construction will guarantee that for *all* edges g, the probability that g is reduced is *exactly* p, that is, it is the same for all edges, and (b) the fact that $m_{\mathbf{e},u}x_u + m_{\mathbf{e},v}x_v = x_{\mathbf{e}}$. Plugging these facts back into (3.3) gives

$$\begin{aligned}\mathbb{E}[s_{\mathbf{e}}] &\leq -p\eta x_{\mathbf{e}} + m_{\mathbf{e},u}(1-\epsilon)p\eta + m_{\mathbf{e},v}(1-\epsilon)p\eta \\ &\leq -p\eta x_{\mathbf{e}} + (1-\epsilon)p\eta x_{\mathbf{e}} = -\epsilon p \eta x_{\mathbf{e}}.\end{aligned} \tag{3.4}$$

If we could prove (3.4) for *every* edge f in the support of x, that would complete the proof that the expected cost of the min O-join for a random spanning tree $T \sim \mu$ is at most $(1/2 - \epsilon)OPT$.

3.2.1. Remark. Throughout this paper, we repeatedly use a mild generalization of the previous "independent trees fact": that if S is a cut with $x(\delta(S)) \leq 2 + \epsilon$, then S_T is very likely to be a tree. Conditioned on this fact, marginals inside S and outside S are nearly preserved and the trees inside S and outside S are sampled independently (see Lemma 2.7).

3.2.2. Ideal Reduction. In the example, we were able to show that $\mathbb{P}[\delta(u)_T \text{ odd} \mid g \text{ reduced}]$ was bounded away from one for every edge $g \in \delta^{\uparrow}(u)$, and this is how we proved that the expected reduction for each edge was greater than the expected increase on each edge, yielding negative expected slack.

This motivates the following definition: A reduction for an edge g is *k-ideal* if, conditioned on g reduced, every cut S that is in the top k levels of cuts containing g is odd with probability that is bounded away from one.

3.2.3. Moving Away from an Idealized Setting. In Example 3.1, we oversimplified in four ways:

(a) We assumed that it would be possible to show that each top edge is *good*. That is, that its top two cuts are even *simultaneously* with constant probability.

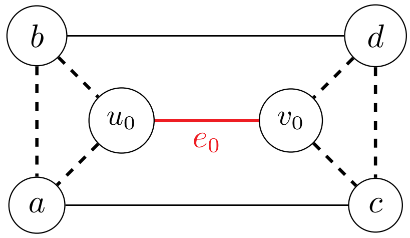
(b) We considered only top edge bundles (i.e., edges whose top cuts were inside a degree cut).

(c) We assumed that $x_u, x_v \in [\epsilon, 1 - \epsilon]$.

(d) We assumed the existence of a nice matching between edges whose top two cuts were children of S and the edges in $\delta(S)$.

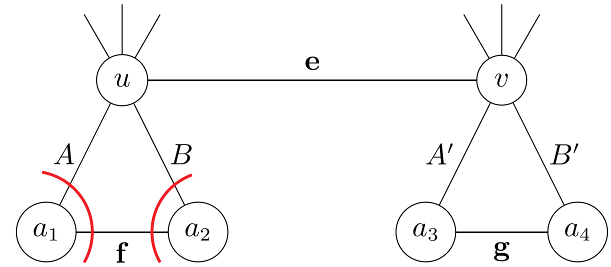
Our proof needs to address all four anomalies that result from deviating from these assumptions.

Figure 3. Example with Bad Edges



Notes. A feasible solution of the Held-Karp relaxation is shown; dashed edges have fraction 1/2 and solid edges have fraction 1. Writing $E = E_0 \setminus \{e_0\}$ as a maximum entropy distribution μ we get the following: Edges $(a,b),(c,d)$ must be completely negatively correlated (and independent of all other edges). So, $(b,u_0),(a,u_0)$ are also completely negatively correlated. This implies (a,b) is a bad edge.

Figure 4. Triangle Example



Notes. In this representation of the cut hierarchy (as in Figure 2), for the triangle u corresponding to the cut $\delta(a_1 \cup a_2)$, when A_T and B_T are odd, all three cuts $(\delta(a_1)_T, \delta(a_2)_T$ and $\delta(a_1 \cup a_2)_T = \delta(u)_T$ are even (because $\mathbf{f}_T$ is always one). (Recall also that the edges in the bundle $\mathbf{e}$ must have one endpoint in $\{a_1 \cup a_2\}$ and one endpoint in $\{a_3 \cup a_4\}$, as was the case, for example, for the edge (a,c) in Figure 2.)

3.2.3.1. Bad Edges. Consider first (a). Unfortunately, it is not the case that all top edges are good. Indeed, some are *bad*. However, it turns out that bad edges are rare in the following senses: First, for an edge to be bad, it must be a half edge, where we say that an edge $\mathbf{e}$ is a half edge if $x_\mathbf{e} \in 1/2 \pm \epsilon_{1/2}$ for a suitably chosen constant $\epsilon_{1/2}$. Second, of any two half edge bundles sharing a common endpoint in the hierarchy, at least one is good. For example, in Figure 3, (a,u_0) and (b,u_0) are good half-edge bundles. We advise the reader to ignore half edges in the first reading of the paper. Correspondingly, we note that our proofs would be much simpler if half-edge bundles never showed up in the hierarchy. It may not be a coincidence that half edges are hard to deal with, as it is conjectured that TSP instances with half-integral LP solutions are the hardest to round (Schalekamp et al. 2012, 2013).

Our solution is to *never* reduce bad edges. However, this in turn poses two problems. First, it means that we need to address the possibility that the bad edges constitute most of the cost of the LP solution. Second, our objective is to get negative expected slack on each good edge and nonpositive expected slack on bad edges. Therefore, if we never reduce bad edges, we cannot increase them either, which means that the responsibility for fixing an odd cut with reduced edges going higher will have to be split amongst fewer edges (the incident good ones).

We deal with the first problem by showing that in every cut u in the hierarchy at least 3/4 of the fractional mass in $\delta(u)$ is good and these edges suffice to compensate for reductions on the edges going higher. Moreover, because there are sufficiently many good edges incident to each cut, we can show that either using the slack vector $\{s_e\}$ gives us a low-cost O-join, or we can average it out with another O-join solution concentrated on bad edges to obtain a reduced cost matching of odd degree vertices.

We deal with the second problem by proving Lemma 6.1, which guarantees a matching between *good* edge bundles $\mathbf{e} = (u,v)$ and fractions $m_{\mathbf{e},u}, m_{\mathbf{e},v}$ of edges in $\delta^\uparrow(u)$, $\delta^\uparrow(v)$ such that, roughly, $m_{\mathbf{e},u} + m_{\mathbf{e},v} = (1 + O(\epsilon_{1/2}))x_\mathbf{e}$.

3.2.3.2. Dealing with Triangles. Turning to (b), consider a triangle cut S, for example $\delta(a_1 \cup a_2)$ in Figure 4. Recall that in a triangle, we can assume that there is an edge of fractional value 1 connecting a_1 and a_2 in the tree, and this is why we defined the cut to be happy when A_T and B_T are odd: this guarantees that all three cuts defined by the triangle $(\delta(a_1), \delta(a_2), \delta(a_1 \cup a_2)$ are even.

Now suppose that $\mathbf{e} = (u,v)$ is a top edge bundle, where u and v are both triangles, as shown in Figure 4. Then we would like to reduce $s_\mathbf{e}$ when both cuts u and v are happy. However, this would require more than simply both cuts being even. This would require *all* A_T, B_T, A'_T, B'_T to be odd. If, for whatever reason, $\mathbf{e}$ is reduced only when $\delta(u)_T$ and $\delta(v)_T$ are both even, then it could be, for example, that this only happens when A_T and B_T are both even. In this case, both $\delta(a_1)_T$ and $\delta(a_2)_T$ will be odd with probability 1 (recalling that $\mathbf{f}_T = 1$), which would then necessitate an increase in $s_\mathbf{f}$ whenever $\mathbf{e}$ is reduced. In other words, the reduction will not even be 1-ideal.

It turns out to be easier for us to get a 1-ideal reduction rule for $\mathbf{e}$ as follows: Say that $\mathbf{e}$ is *2-1-1 happy with respect to* u if $\delta(u)_T$ is even and both A'_T, B'_T are odd. We reduce $\mathbf{e}$ with probability $p/2$ when it is 2-1-1 happy with respect to u and with probability $p/2$ when it is 2-1-1 happy with respect to v. This means that when $\mathbf{e}$ is reduced, half of the time no increase in $s_\mathbf{f}$ is needed because u is happy. Similarly for v.

The 2-1-1 criterion for reduction introduces a new kind of bad edge: a half edge that is good, but not 2-1-1 good. We are able to show that non-half-edge bundles are 2-1-1 good (Lemmas 5.7 and 5.8) and that if there are two half edges that are both in A or are both in B, then at least one of them is 2-1-1 good (Lemma 5.9). Finally, we show that if there are two half edges, where one is in A and the other is in B, and neither is 2-1-1 good, then we can apply a different reduction criterion that we call *2-2-2 good*. When the latter applies, we are guaranteed to decrease both of the half edge bundles simultaneously. All together, the various considerations discussed in this paragraph force us to come up with a relatively more complicated set of rules under which we reduce $s_{\mathbf{e}}$ for a top edge bundle $\mathbf{e}$ whose children are triangle cuts. Section 5 focuses on developing the relevant probabilistic statements.

3.2.3.3. Bottom Edge Reduction. Next, consider a bottom edge bundle $\mathbf{f} = (a_1, a_2)$ where $\mathsf{p}(a_1) = \mathsf{p}(a_2)$ is a triangle. Our plan is to reduce $s_{\mathbf{f}}$ (i.e., set it to $-\eta x_{\mathbf{f}}$) when the triangle is happy, that is, $A_T = B_T = 1$. The good news here is that every triangle is happy with constant probability. However, when a triangle is *not* happy, $s_{\mathbf{f}}$ may need to increase to make sure that the O-join constraint for $\delta(a_1)$ and $\delta(a_2)$ are satisfied, if edges in A and B going higher are reduced. Because $x_{\mathbf{f}} = x(A) = x(B) = 1$, this means that $\mathbf{f}$ may need to compensate at *twice* the rate at which it is getting reduced. This would result in $\mathbb{E}[s_{\mathbf{f}}] > 0$, which is the opposite of what we seek.

We use two key ideas to address this problem. First, we reduce top edges and bottom edges by different amounts: Specifically, when the relevant reduction event occurs, we reduce a bottom edge $\mathbf{f}$ by $\beta x_{\mathbf{f}}$ and top edges $\mathbf{e}$ by $\tau x_{\mathbf{e}}$, where $\beta > \tau$ (and τ is a multiple of η).

Thus, the expected reduction in $s_{\mathbf{f}}$ is $p\beta x_{\mathbf{f}} = p\beta$, whereas the expected increase (due to compensation of, say, top edges going higher) is $p\tau(x(A) + x(B))q = p\tau 2q$, where

$$q = \mathbb{P}[\text{triangle not happy} \mid \text{reductions in } A \text{ and } B].$$

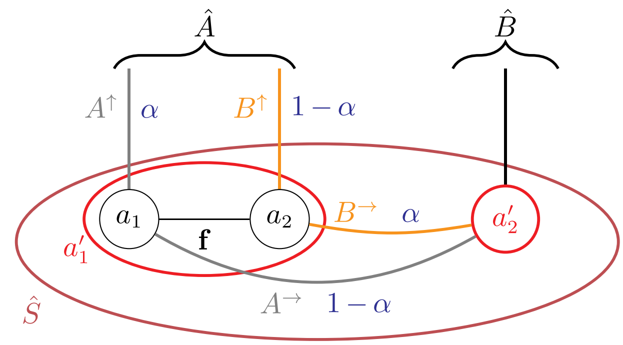
Thus, as long as $2\tau q < \beta - \epsilon$, we get the expected reduction in $s_{\mathbf{f}}$ that we seek.

The discussion thus far suggests that we need to take τ smaller than $\beta/2q$, which is $\beta/2$ if q is 1, for example. On the other hand, if $\tau = \beta/2$, then when a top edge needs to fix a cut due to reductions on bottom edges, we have the opposite problem – their expected increase will be greater than their expected reduction, and we are back to square one.

Coming to our aid is the second key idea, already discussed in Section 1.2.3. We reduce bottom edges only when $A_T = B_T = 1$ *and* the marginals of edges in A, B are approximately preserved (conditioned on $A_T = B_T = 1$). This allows us to get much stronger upper bounds on the probability that a lower cut a bottom edge is on is odd, given that the bottom edge is reduced, and enables us to show that bottom edge reduction is ∞-ideal.

It turns out that the combined effects of (a) choosing $\tau = 0.571\beta$ and (b) getting better bounds on the probability that a lower cut is even given that a bottom edge is reduced, suffice to deal with the interaction between the reductions and the increases in slack for top and bottom edges.

Figure 5. Setting of Example 3.2



Notes. The set $A = \delta(a_1) \cap \delta(a_1')$ decomposes into two sets of edges, $A^{\uparrow}$, those that are also in $\delta(S)$, and the rest, which we call $A^{\rightarrow}$. Similarly for B.

Example 3.2 (Bottom-Bottom Case). To see how preserving marginals helps us handle the interaction between bottom edges at consecutive levels, consider a triangle cut $a_1' = \{a_1, a_2\}$ whose parent cut $\hat{S} = \{a_1', a_2'\}$ is also a triangle cut (as shown in Figure 5). Let's analyze $\mathbb{E}[s_{\mathbf{f}}]$ where $\mathbf{f} = (a_1, a_2)$. Observe first that $A^{\rightarrow} \cup B^{\rightarrow}$ is a bottom edge bundle in the triangle $\hat{S}$ and all edges in this bundle are reduced simultaneously when $\hat{A}_T = \hat{B}_T = 1$ and marginals of all edges in $\hat{A} \cup \hat{B}$ are approximately preserved. (For the purposes of this overview, we'll assume they are preserved exactly). Furthermore, because the tree inside $\hat{S}$ is picked independently of the tree on $G/\hat{S}$ (using Lemma 2.7 and assuming $\epsilon = 0$ for this overview), exactly one edge in $A^{\rightarrow} \cup B^{\rightarrow}$ is selected independently of the reduction event $\hat{A}_T = \hat{B}_T = 1$. Let $x(A^{\uparrow}) = \alpha$. Then because $A = A^{\uparrow} \cup A^{\rightarrow}$ and $x(A) = 1$, we have $x(A^{\rightarrow}) = 1 - \alpha$. Moreover, because $\hat{A} = A^{\uparrow} \cup B^{\uparrow}$ and $x(\hat{A}) = 1$, we also have $x(B^{\uparrow}) = 1 - \alpha$ and $x(B^{\rightarrow}) = \alpha$.

Therefore, using the fact that when $A^{\rightarrow} \cup B^{\rightarrow}$ is reduced, exactly one edge in $A^{\uparrow} \cup B^{\uparrow}$ is selected (and also exactly one edge in $A^{\rightarrow} \cup B^{\rightarrow}$ is selected independently because it is a bottom edge bundle, as mentioned previously), and marginals are preserved given the reduction, we conclude that

$$\begin{aligned}&\mathbb{P}[a_1' \text{ happy} \mid A^{\rightarrow} \cup B^{\rightarrow} \text{ reduced}]\\&= \mathbb{P}[A_T = B_T = 1 \mid A^{\rightarrow} \cup B^{\rightarrow} \text{ reduced}] = \alpha^2 + (1-\alpha)^2.\end{aligned}$$

Now, we calculate $\mathbb{E}[s_{\mathbf{f}}]$. First, $\mathbf{f}$ may have to increase to compensate either for reduced edges in $A^{\uparrow} \cup B\uparrow$ or in $A^{\rightarrow} \cup B^{\rightarrow}$. For the sake of this discussion, suppose that $A^{\uparrow} \cup B^{\uparrow}$ is a set of top edges. Then, in the worst case we need to increase $\mathbf{f}$ by $p\tau$ in expectation to fix the cuts a_1, a_2 due to the reduction in $A^{\uparrow} \cup B^{\uparrow}$. Now, we calculate the expected increase due to the reduction in $A^{\rightarrow} \cup B^{\rightarrow}$. The crucial observation is that edges in $A^{\rightarrow} \cup B^{\rightarrow}$ are reduced simultaneously, so both cuts $\delta(a_1)$ and $\delta(a_2)$ can be fixed simultaneously by an increase in $s_{\mathbf{f}}$. Therefore, when they are both odd, it suffices for $\mathbf{f}$ to increase by

$$\max\{x(A^{\rightarrow}), x(B^{\rightarrow})\}\beta = \max\{\alpha, 1-\alpha\}\beta,$$

to fix cuts a_1, a_2. Putting this together, we get

$$\begin{aligned}\mathbb{E}[s_{\mathbf{f}}] &= -p\beta + \mathbb{E}[\text{increase due to } A^{\rightarrow} \cup B^{\rightarrow}] \\ &\quad + \mathbb{E}[\text{increase due to } A^{\uparrow} \cup B^{\uparrow}] \\ &\leq -p\beta + p\beta \max_{\alpha \in [1/2,1]} \alpha[1-\alpha^2-(1-\alpha)^2] + p\tau\end{aligned}$$

which, since $\max_{\alpha\in[1/2,1]}\alpha[1-\alpha^2-(1-\alpha)^2] = 8/27$ and $\tau = 0.571\beta$ is

$$= p\beta\left(-1 + \frac{8}{27} + 0.571\right) = -0.13p\beta.$$

3.2.4. Dealing with x_u Close to One.[9]

Now, suppose that $\mathbf{e} = (u, v)$ is a top edge bundle with $x_u := x(\delta^{\uparrow}(u))$ is close to one. Then, the analysis in Example 3.1, bounding $r := \mathbb{P}[\delta(u)_T \text{ odd} | g \text{ reduced}]$ away from one for an edge $g \in \delta^{\uparrow}(u)$ does not hold. To address this, we consider two cases: The first case, is that the edges in $\delta^{\uparrow}(u)$ break up into many groups that end at different levels in the hierarchy. In this case, we can analyze r separately for the edges that end at any given level, taking advantage of the independence between the trees chosen at different levels of the hierarchy.

The second case is when nearly all of the edges in $\delta^{\uparrow}(u)$ end at the same level, for example, they are all in $\delta^{\rightarrow}(u')$, where $\mathsf{p}(u')$ is a degree cut. In this case, we introduce a more complex (2-1-1) reduction rule for these edges. The observation is that from the perspective of these edges u' is a "pseudo-triangle." That is, it looks like a triangle cut, with atoms u and $u' \smallsetminus u$, where $\delta(u) \cap \delta(u')$ corresponds to the "A"-side of the triangle.

Now, we define this more complex 2-1-1 reduction rule: Consider a top edge $\mathbf{f} = (u', v') \in \delta^{\rightarrow}(u')$. Thus far, we only considered the following reduction rule for $\mathbf{f}$: If both u', v' are degree cuts, $\mathbf{f}$ reduces when they are both even in the tree; otherwise if say u' is a triangle cut, $\mathbf{f}$ reduces when it is 2-1-1 good w.r.t., u' (and similarly for v'). However, clearly these rules ignore the pseudo triangle. The simplest adjustment is, if u' is a pseudo triangle with partition $(u, u' \smallsetminus u)$, to require $\mathbf{f}$ to reduce when $A_T = B_T = 1$ and v' is happy. However, as stated, it is not clear that the sets A and B are well defined. For example, u' could be an actual triangle or there could be multiple ways to see u' as a pseudo triangle only one of which is $(u, u' \smallsetminus u)$. Our solution is to find the *smallest* disjoint pair of cuts $a, b \subset u'$ in the hierarchy such that $x(\delta(a) \cap \delta(u')), x(\delta(b) \cap \delta(u')) \geq 1 - \epsilon_{1/1}$, where $\epsilon_{1/1}$ is a fixed universal constant, and then let $A = \delta(a) \cap \delta(u'), B = \delta(b) \cap \delta(u')$ and $C = \delta(u') \smallsetminus A \smallsetminus B$ (see Figure 6 for an example). Then, we say $\mathbf{f}$ is 2-1-1 happy w.r.t., u' if $A_T = B_T = 1$ and $C_T = 0$.

A few observations are in order:

- Because u is a candidate for, say a, it must be that a is a descendent of u in the hierarchy (or equal to u). In addition, b cannot simultaneously be in u, because $a \cap b = \emptyset$ and $x(\delta(u) \cap \delta(u')) \leq 1$ by Lemma 2.3. Therefore, when $\mathbf{f}$ is 2-1-1 happy w.r.t. u', we get $(\delta(u) \cap \delta(u'))_T = 1$.
- If $u' = (X, Y)$ is a actual triangle cut, then we must have $a \subseteq X, b \subseteq Y$. Therefore, when $\mathbf{f}$ is 2-1-1 happy w.r.t. u', we know that u' is a happy triangle, that is, $(\delta(X) \cap \delta(u'))_T = 1$ and $(\delta(Y) \cap \delta(u'))_T = 1$.

Now, suppose for simplicity that all top edges in $\delta(u')$ are 2-1-1 good w.r.t. u'. Then, when an edge $g \in \delta(u) \cap \delta(u')$ is reduced, $(\delta(u) \cap \delta(u'))_T = 1$, so

$$\begin{aligned}&\mathbb{P}[\delta(u)_T \text{ odd} | g \text{ reduced}] \\ &\leq \mathbb{P}[E(u, u' \smallsetminus u)_T \text{ even} | g \text{ reduced}] \leq 0.57,\end{aligned}$$

becuase edges in $E(u, u' \smallsetminus u)$ are in the tree independent of the reduction and $\mathbb{E}[E(u, u' \smallsetminus u)_T] \approx 1$.

3.2.3.4. Dealing with x_u Close to Zero and the Matching.

We already discussed how the matching is modified to handle the existence of bad edges. We now observe that we can handle the case $x_u \approx 0$ by further modifying the matching. The key observation is that in this case, $x(\delta^{\rightarrow}(u)) \gg x(\delta^{\uparrow}(u))$. Roughly speaking, this enables us to find a matching in which each edge in $\delta^{\rightarrow}(u)$ has to increase about half as much as would normally be expected to fix the cut of u. This eliminates the need to prove a nontrivial bound on $\mathbb{P}[\delta(u)_T \text{ odd} | g \text{ reduced}]$. The details of the matching are in Section 6.

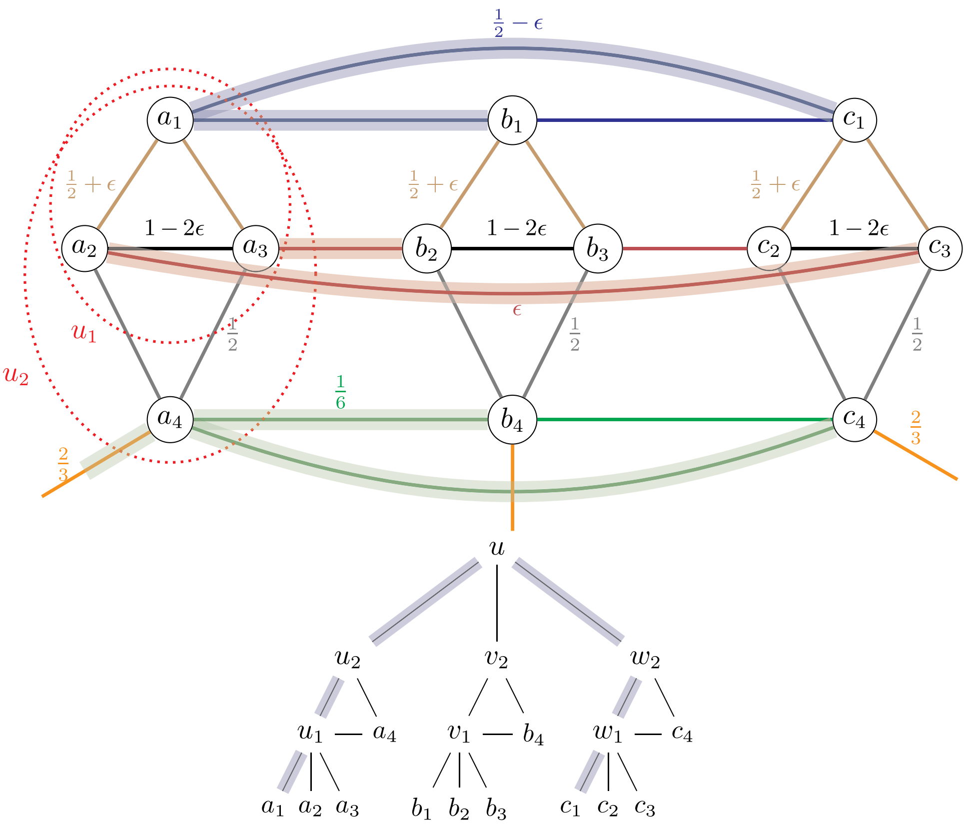
4. Polygons and the Hierarchy of Near Minimum Cuts

Let OPT be a minimum TSP solution, that is, minimum cost Hamiltonian cycle and without loss of generality assume it visits u_0 and v_0 consecutively (recall that $c(u_0, v_0) = 0$). We write E^* to denote the edges of OPT, and we write e^* to denote an edge of OPT. Analogously, we use $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ to denote the slack vector that we will construct for OPT edges.

Throughout this section, we study η-near minimum cuts of $G = (V, E, z)$ These cuts are 2η-near minimum cuts w.r.t., x. For every such near minimum cut, $(S, \overline{S})$, we identify the cut with the side, say S, such that $u_0, v_0 \notin S$. Equivalently, we can identify these cuts with an interval along the optimum cycle, OPT, that does not contain u_0, v_0.

We will use "left" synonymously with "clockwise" and "right" synonymously with "counterclockwise."

Figure 6. Part of the Hierarchy of the Graph is Shown on Top



Notes. Edges of the same color have the same fraction and $\epsilon \gg \eta$ is a small constant. u_1 corresponds to the degree cut $\{a_1, a_2, a_3\}$, u_2 corresponds to the triangle cut $\{u_1, a_4\}$ and u corresponds to the degree cut containing all of the vertices shown. Observe that edges in $\delta^{\uparrow}(a_1)$ are top edges in the degree cut u. If $\epsilon < \frac{1}{2}\epsilon_{1/1}$ then the (A, B, C)-degree partitioning of edges in $\delta(u_2)$ is as follows: $A = \delta(a_1) \cap \delta(u_2)$ are the blue highlighted edges each of fractional value $1/2 - \epsilon$, $B = \delta(a_4) \cap \delta(u_2)$ are the green highlighted edges of total fractional value 1, and C are the red highlighted edges each of fractional value ϵ. The cuts that contain edge (a_1, c_1) are highlighted in the hierarchy at the bottom.

We say a vertex is to the left of another vertex if it is to the left of that vertex and to the right of edge $e_0 = (u_0, v_0)$. Otherwise, we say it is to the right (including the root itself in this case).

Definition 4.1 (Crossed on the Left/Right, Crossed on Both Sides). For two crossing near minimum cuts S, S', we say S *crosses* S' *on the left* if the leftmost endpoint of S on the optimal cycle is to the left of the leftmost endpoint of S. Otherwise, we say S *crosses* S' *on the right*.

A near minimum cut is *crossed on both sides* if it is crossed on both the left and the right. We also say a a near minimum cut is *crossed on one side* if it is either crossed on the left or on the right but not both.

4.1. Cuts Crossed on Both Sides

The following theorem is the main result of this section.

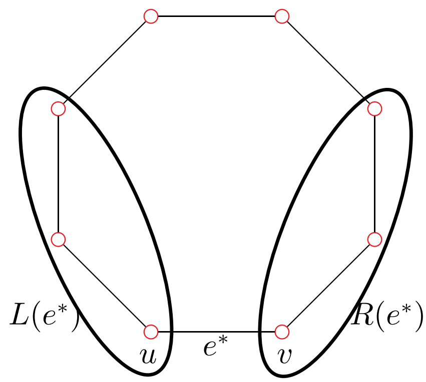
Theorem 4.1. *Given OPT TSP tour with set of edges E^*, and a feasible LP solution x^0 of the Held-Karp relaxation with support $E_0 = E \cup \{e_0\}$ and let x be x^0 restricted to E. For any distribution μ of spanning trees with marginals x and $\beta > 0$, if $\eta < 1/100$, then there is a random vector $s^*: E^* \to \mathbb{R}_{\geq 0}$ (the randomness in s^* depends exclusively on $T \sim \mu$) such that*

- *For any vector $s: E \to \mathbb{R}$ where $s_e \geq -x_e\beta$ for all e and for any η-near minimum cut S w.r.t., $z = (x + OPT)/2$ crossed on both sides where $\delta(S)_T$ is odd, we have $s(\delta(S)) + s^*(\delta(S)) \geq 0$;*
- *For any $e^* \in E^*$, $\mathbb{E}[s^*_{e^*}] \leq 37\eta\beta$.*

For an OPT edge $e^* = (u, v)$, let $L(e^*)$ be the largest η-near minimum cut (w.r.t. z) containing u and not v that is crossed on both sides. Let $R(e^*)$ be the largest near minimum cut containing v and not u that is crossed on both sides ($L(e^*), R(e^*)$ do not necessarily exist). For example, see Figure 7.

Definition 4.2. For a near minimum cut S that is crossed on both sides let S_L be the near minimum cut crossing S on the left which minimizes the intersection with S, and similarly for S_R; if there are multiple sets crossing S on the left with the same minimum intersection, choose the smallest one to be S_L (and similar do for S_R).

Figure 7. L and R for an OPT Edge e^*



We partition $\delta(S)$ into three sets $\delta(S)_L, \delta(S)_R$ and $\delta(S)_O$ as in Figure 8 such that

$$\delta(S)_L = E(S \cap S_L, S_L \smallsetminus S)$$

$$\delta(S)_R = E(S \cap S_R, S_R \smallsetminus S)$$

$$\delta(S)_O = \delta(S) \smallsetminus (\delta(S)_L \cup \delta(S)_R).$$

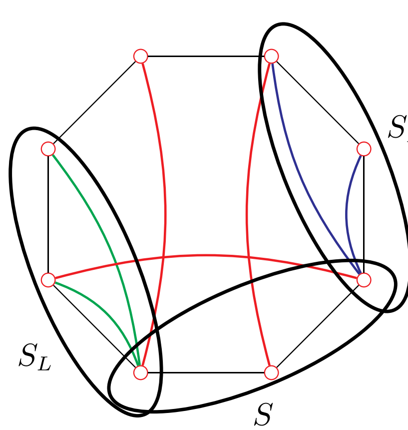
For an OPT edge e^* define an (increase) event (of second type) $\mathcal{I}_2(e^*)$ as the event that at least one of the following *does not* hold. (If $L(e^*)$ does not exist, assume the first and third events always hold; similarly if $R(e^*)$ does not exist, assume the second and fourth events always hold.)

$$|T \cap \delta(L(e^*))_R| = 1, |T \cap \delta(R(e^*))_L| = 1, \quad T \cap \delta(L(e^*))_O = \emptyset, \text{ and } T \cap \delta(R(e^*))_O = \emptyset. \tag{4.1}$$

In the proof of Theorem 4.1, we will increase an OPT edge e^* whenever $\mathcal{I}_2(e^*)$ occurs.

Lemma 4.1. *For any OPT edge e^*, $\mathbb{P}[\mathcal{I}_2(e^*)] \le 18\eta$.*

Figure 8. A Cut Crossed on Both Sides



Notes. S is crossed on the left by S_L and on the right by S_R. In green are edges in $\delta(S)_L$, in blue edges in $\delta(S)_R$, and in red are edges in $\delta(S)_O$.

Fix e^*. To simplify notation we abbreviate $L(e^*), R(e^*)$ to L, R. Because L is crossed on both sides, L_L, L_R are well defined. Because by Lemma 2.1 $L_L \cap L, L_L \smallsetminus L$ are 4η-near min cuts and L is 2η-near min cut with respect to x, by Corollary 2.3, $\mathbb{P}[|T \cap \delta(L)_L)| = 1] \ge 1 - 5\eta$. Similarly, $\mathbb{P}[|T \cap \delta(R)_L| = 1] \ge 1 - 5\eta$. On the other hand, because L, L_L, L_R are 2η-near min cuts, by Lemma 2.2, $x(E(L \cap L_R, L_R)), x(E(L \cap L_L, L_L)) \ge 1 - \eta$. Therefore,

$$\begin{aligned} x(\delta(L)_O) \le 2 + 2\eta - x(E(L \cap L_R, L_R)) \\ - x(E(L \cap L_L, L_L)) \le 4\eta. \end{aligned}$$

It follows that $\mathbb{P}[T \cap \delta(L)_O = \emptyset] \ge 1 - 4\eta$. Similarly, $\mathbb{P}[T \cap \delta(R)_O = \emptyset] \ge 1 - 4\eta$. Finally, by the union bound, all events occur simultaneously with probability at least $1 - 18\eta$. Therefore, $\mathbb{P}[\mathcal{I}_2(e^*)] \le 18\eta$ as desired.

Lemma 4.2. *Let S be a cut that is crossed on both sides and let e_L^*, e_R^* be the OPT edges on its interval where e_L^* is the edge further clockwise. Then, if $\delta(S)_T \ne 2$, at least one of $\mathcal{I}_2(e_L^*), \mathcal{I}_2(e_R^*)$ occurs.*

We prove by contradiction. Suppose none of $\mathcal{I}_2(e_L^*)$, $\mathcal{I}_2(e_R^*)$ occur; we will show that this implies $\delta(S)_T = 2$ (See Figure 9).

Let $R = R(e_L^*)$; S is a candidate for $R(e_L^*)$, so $S \subseteq R$. Therefore, $S_L = R_L$, and we have

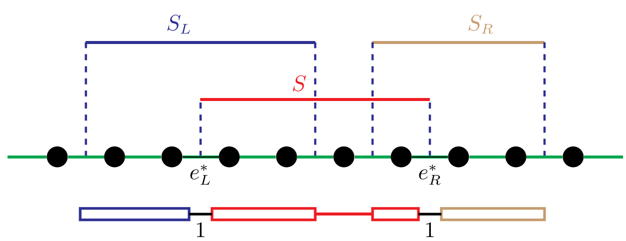
$$\delta(R)_L = E(R \cap R_L, R_L \smallsetminus R) = E(R \cap S_L, S_L \smallsetminus R) = \delta(S)_L.$$

We used $S \cap S_L = R \cap S_L$ and that $S_L \smallsetminus S = S_L \smallsetminus R$. Similarly let $L = L(e_R^*)$, and, we have $\delta(L)_R = \delta(S)_R$.

Now, because $\mathcal{I}_2(e_L^*)$ has not occurred, $1 = |T \cap \delta(R)_L| = |T \cap \delta(S)_L|$, and because $\mathcal{I}_2(e_R^*)$ has not occurred, $1 = |T \cap \delta(L)_R| = |T \cap \delta(S)_R|$, where $L = L(e_R^*)$. Therefore, to get $\delta(S)_T = 2$, it remains to show that $T \cap \delta(S)_O = \emptyset$. Consider any edge $e = (u, v) \in \delta(S)_O$ where $u \in S$. We need to show $e \notin T$. Assume that v is to the left of S (the other case can be proven similarly). Then $e \in \delta(R)$. Therefore, because e goes to the left of R, either $e \in E(R \cap R_L, R_L \smallsetminus R)$ or $e \in \delta(R)_O$. However, because $e \notin \delta(S)_L = \delta(R)_L$, we must have $e \in \delta(R)_O$. Therefore, because $\mathcal{I}_2(e_L^*)$ has not occurred, $e \notin T$ as desired. □

Proof of Theorem 4.1. For any OPT edge e^* whenever $\mathcal{I}_2(e^*)$ occurs, define $s_{e^*}^* = 2.02\beta$. Then, by Lemma 4.1,

Figure 9. Setting of Lemma 4.2



Notes. Here we zoom in on a portion of the optimal cycle and assume the root is not shown. If $\mathcal{I}_2(e_L^*)$ does not occur then $E(S \cap S_L, S_L \smallsetminus S)_T = 1$.

$\mathbb{E}[s_{e^*}] \le 18 \cdot 2.02\beta$ and for any 2η-near min cut S (w.r.t., x) that is crossed on both sides if $\delta(S)_T$ is odd, then at least one of $\mathcal{I}_2(e^*_L), \mathcal{I}_w(e^*_R)$ occurs, so

$$s(\delta(S)) + s^*(\delta(S)) \ge -x(\delta(S))\beta + s^*_{e^*_L} + s^*_{e^*_R} \ge -(2+2\eta)\beta + 2.02\beta \ge 0$$

for $\eta < 1/100$ as desired.

4.2. Proof of the Main Technical Theorem (Theorem 3.1)

The following theorem is the main result of this section.

Theorem 4.2. *Let x^0 be a feasible solution of the Held-Karp relaxation with support $E_0 = E \cup \{e_0\}$ and x be x^0 restricted to E. Let μ be the max entropy distribution with marginals x. For $\eta \le 10^{-12}$, $\beta > 0$, there is a set $E_g \subset E \smallsetminus \delta(\{u_0, v_0\})$ of good edges and two functions $s : E_0 \to \mathbb{R}$ and $s^* : E^* \to \mathbb{R}_{\ge 0}$ (as functions of $T \sim \mu$) such that*

(i) *For each edge $e \in E_g$, $s_e \ge -x_e\beta$ and for any $e \in E \smallsetminus E_g$, $s_e = 0$.*

(ii) *For each η-near-min cut S w.r.t. z, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \ge 0$.*

(iii) *We have $\mathbb{E}[s_e] \le -\epsilon_P\beta x_e$ for all edges $e \in E_g$ and $\mathbb{E}[s^*_{e^*}] \le 218\eta\beta$ for all OPT edges $e^* \in E^*$. for ϵ_P defined in (7.4).*

(iv) *For every η-near minimum cut S of z crossed on (at most) one side such that $S \ne V \smallsetminus \{u_0, v_0\}$, $x(\delta(S) \cap E_g) \ge 3/4$.*

Before proving this theorem, we use it to prove the main technical theorem from the previous section.

Theorem 3.1 (Main Technical Theorem). *Let x^0 be a solution of the Held-Karp relaxation with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E. Let $z := (x + OPT)/2$, $\eta \le 10^{-12}$, $\beta > 0$, and let μ be the max-entropy distribution with marginals x. Also, let E^* denote the support of OPT. There are two functions $s : E_0 \to \mathbb{R}$ and $s^* : E^* \to \mathbb{R}_{\ge 0}$ (as functions of $T \sim \mu$), such that*

(i) *For each edge $e \in E$, $s_e \ge -x_e\beta$.*

(ii) *For each η-near-min cut S of z, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \ge 0$.*

(iii) *For every OPT edge e^*, $\mathbb{E}[s^*_{e^*}] \le 218\eta\beta$ and for every LP edge $e \ne e_0$, $\mathbb{E}[s_e] \le -\frac{1}{3}x_e\epsilon_P\beta$ for $\epsilon_P = 3.12 \cdot 10^{-16}$ (defined in (7.4)).*

Proof of Theorem 3.1. Let E_g be the good edges defined in Theorem 4.2 and let $E_b := E \smallsetminus E_g$ be the set of bad edges; in particular, edges in $\delta(\{u_0, v_0\})$ are bad edges. We define a new vector $\tilde{s} : E \cup \{e_0\} \to \mathbb{R}$ as follows:

$$\tilde{s}(e) \leftarrow \begin{cases} \infty & \text{if } e = e_0 \\ -x_e(4\beta/5)(1-2\eta) & \text{if } e \in E_b, \\ x_e(4\beta/3) & \text{otherwise.} \end{cases} \tag{4.2}$$

Let $\tilde{s}^*$ be the vector s^* from Theorem 4.1. We claim that for any η-near minimum cut S such that $\delta(S)_T$ is odd, we have

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \ge 0.$$

To check this note by (iv) of Theorem 4.2 for every set $S \ne V \smallsetminus \{u_0, v_0\}$ crossed on at most one side, we have $x(E_g \cap \delta(S)) \ge \frac{3}{4}$, so

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \ge \tilde{s}(\delta(S)) = \frac{4\beta}{3}x(E_g \cap \delta(S)) - \frac{4\beta}{5}(1-2\eta)x(E_b \cap \delta(S)) \ge 0. \tag{4.3}$$

For $S = V \smallsetminus \{u_0, v_0\}$, we have $\delta(S)_T = \delta(u_0)_T + \delta(v_0)_T = 2$ with probability 1, so condition (ii) is satisfied for these cuts as well. Finally, consider cuts S which are crossed on both sides. By Theorem 4.1,

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \ge 0 \tag{4.4}$$

because $\tilde{s}_e \ge -\frac{4}{5}\beta x_e \ge -\beta x_e$ for all e.

Now, we are ready to define s, s^*. Let $\hat{s}, \hat{s}^*$ be the s, s^* of Theorem 4.2, respectively. Define $s = \gamma\tilde{s} + (1-\gamma)\hat{s}$ and similarly define $s^* = \gamma\tilde{s}^* + (1-\gamma)\hat{s}^*$ for some γ that we choose later. We prove all three conclusions for s, s^*. (i) follows by (i) of Theorem 4.2 and Equation (4.2). (ii) follows by (ii) of Theorem 4.2 and Equation (4.3). It remains to verify (iii). For any OPT edge e^*, $\mathbb{E}[s^*_{e^*}] \le 218\eta\beta$ by (iii) of Theorem 4.2 and the construction of $\tilde{s}^*$. On the other hand, by (iii) of Theorem 4.2 and Equation (4.2),

$$\mathbb{E}[s_e] \begin{cases} \le x_e\left(\gamma\frac{4}{3}\beta - (1-\gamma)\epsilon_P\beta\right) & \forall e \in E_g, \\ = -x_e\gamma \cdot \left(\frac{4}{5}\beta\right)(1-2\eta) & \forall e \in E_b. \end{cases}$$

Setting $\gamma = \frac{15}{32}\epsilon_P$ we get $\mathbb{E}[s_e] \le -\frac{1}{3}\epsilon_P\beta x_e$ for $e \in E_g$ and $\mathbb{E}[s_e] \le -\frac{1}{3}x_e\beta\epsilon_P$ for $e \in E_b$ as desired. □

4.3. Structure of Polygons of Cuts Crossed on One Side

Definition 4.3 (Connected Component of Crossing Cuts). Given a family of cuts crossed on at most one side, construct a graph where two cuts are connected by an edge if they cross. Partition this graph into *maximal connected components*. We call a path in this graph, a *path of crossing cuts*.

In the rest of this section, we will focus on a single connected component $\mathcal{C}$ of cuts crossed on (at most) one side.

Definition 4.4 (Polygon). For a connected component $\mathcal{C}$ of crossing near min cuts that are crossed on one side, let $a_0, \ldots, a_{m-1}$ be the coarsest partition of the vertices V, such that for all $0 \le i \le m-1$ and for any $A \in \mathcal{C}$ either $a_i \subseteq A$ or $a_i \cap A = \emptyset$. These are called atoms. We assume a_0 is the atom that contains the special edge e_0, and we call it the *root*. For any $A \in \mathcal{C}, a_0 \cap A = \emptyset$.

Because every cut $A \in \mathcal{C}$ corresponds to an interval of vertices in V in the optimum Hamiltonian cycle, we can arrange $a_0, \ldots, a_{m-1}$ around a cycle (in the counter clockwise order). We label the arcs in this cycle from one to m, where $i+1$ is the arc connecting a_i and a_{i+1} (and m is the name of the arc connecting a_{m-1} and a_0). Then every cut $A \in \mathcal{C}$ can be identified by the two arcs surrounding its atoms. Specifically, A is identified with arcs i, j (where $i < j$) if A contains atoms $a_i, \ldots, a_{j-1}$, and we write $\ell(A) = i, r(A) = j$. Note that A does not contain the root a_0.

By construction for every arc $1 \le i \le m$, there exists a cut A such that $\ell(A) = i$ or $r(A) = i$. Furthermore, $A, B \in \mathcal{C}$ (with $\ell(A) \le \ell(B)$) cross iff $\ell(A) < \ell(B) < r(A) < r(B)$.

See Figure 10 for a visual example.

Every atom of a polygon is an interval of the optimal cycle. In this section, we prove the following structural theorem about polygons of near minimum cuts crossed on one side.

Theorem 4.3 (Polygon Structure). *For $\epsilon_\eta \ge 14\eta$ and any polygon with atoms $a_0 \ldots a_{m-1}$ (where a_0 is the root) the following holds:*

- *For all adjacent atoms a_i, a_{i+1} (also including a_0, a_{m-1}), we have $x(E(a_i, a_{i+1})) \ge 1 - \epsilon_\eta$.*
- *All atoms a_i (including the root) have $x(\delta(a_i)) \le 2 + \epsilon_\eta$.*
- *We have $x(E(a_0, \{a_2, \ldots, a_{m-2}\})) \le \epsilon_\eta$.*

The interpretation of this theorem is that the structure of a polygon converges to the structure of an actual integral cycle as $\eta \to 0$. The proof of the theorem follows from the lemmas in the rest of this section.

Definition 4.5 (Left and Right Hierarchies). For a polygon u corresponding to a connected component $\mathcal{C}$ of cuts crossed on one side, let $\mathcal{L}$ (the *left hierarchy*) be the set of all cuts $A \in \mathcal{C}$ that are not crossed on the left. We call any cut in $\mathcal{L}$ *open* on the left. Similarly, we let $\mathcal{R}$ be the set of cuts that are open on the right. Therefore, $\mathcal{L}, \mathcal{R}$ is a partitioning of all cuts in $\mathcal{C}$.

For two distinct cuts $A, B \in \mathcal{L}$ we say A is an *ancestor* of B in the left polygon hierarchy if $A \supseteq B$. We say A is a *strict* ancestor of B if, in addition, $\ell(A) \neq \ell(B)$. We define the right hierarchy similarly: A is a strict ancestor of B if $A \supseteq B$ and $r(A) \neq r(B)$.

We say B is a *strict parent* of A if among all strict ancestors of A in the (left or right) hierarchy, B is the one closest to A.

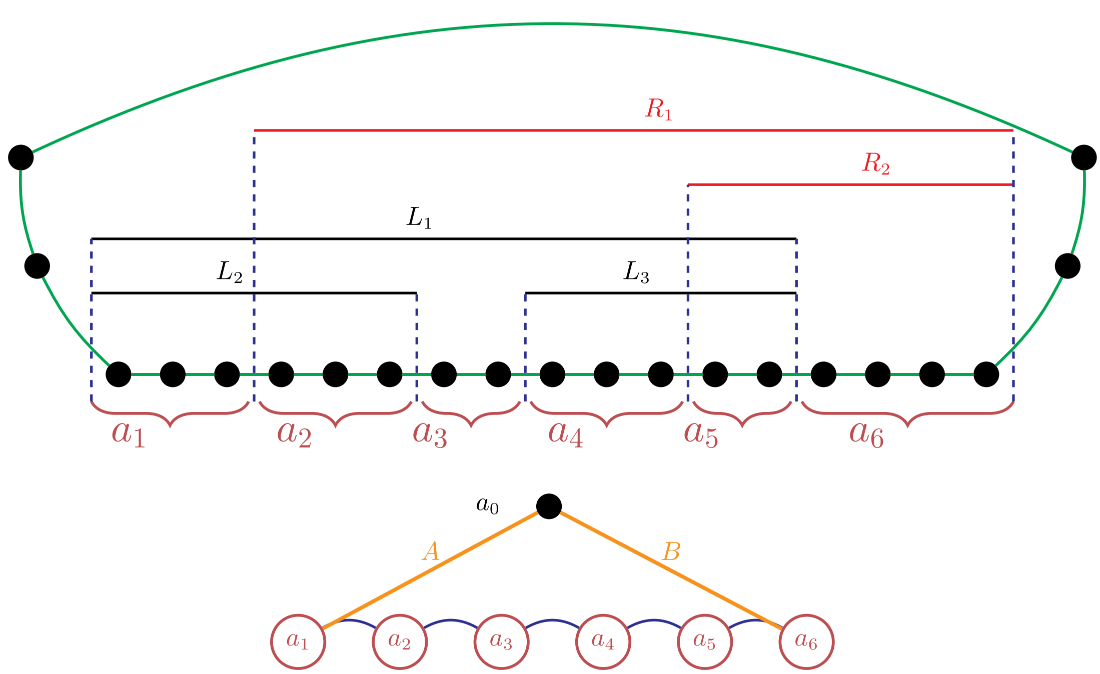
Figure 10 provides examples of sets and their parent/ancestor relationships.

Fact 4.1. *If A, B are in the same hierarchy and they are not ancestors of each other, then $A \cap B = \emptyset$.*

If $A \cap B \neq \emptyset$ then they cross. Therefore, they cannot be open on the same side.

This lemma immediately implies that the cuts in each of the left (and right) hierarchies form a laminar family.

Figure 10. Example of a Polygon with Contracted Atoms



Notes. In black are the cuts in the left polygon hierarchy; in red the cuts in the right polygon hierarchy. OPT edges around the cycle are shown in green. Here R_1 is an ancestor of R_2; however, it is not a strict ancestor of R_2 because they have the same right endpoint. L_1 is a strict ancestor and the strict parent of L_3. By Theorem 4.3, every edge in the bottom picture represents a set of LP edges of total fraction at least $1 - \epsilon_\eta$.

Lemma 4.3. *For $A, B \in \mathcal{R}$ where B is a strict parent of A, there exists a cut $C \in \mathcal{L}$ that crosses both A, B. Similarly, if $A, B \in \mathcal{L}$ and B is a strict parent of A, there exists a cut $C \in \mathcal{R}$ that crosses A, B.*

Because we have a connected component of near min cuts, there exists a path of crossing cuts from A to B. Let $P = (A = C_0, C_1, \ldots, C_k = B)$ be the shortest such path. We need to show that $k = 2$.

First, because C_1 crosses C_0 and C_0 is open on right, we have

$$\ell(C_1) < \ell(C_0) < r(C_1) < r(C_0).$$

Let I be the *closed* interval $[\ell(C_1), r(C_0)]$. Note that $C_k = B$ has an endpoint that does not belong to I. Let C_i be the *first* cut in the path with an endpoint not in I (definitely $i > 1$). This means $C_{i-1} \subseteq I$; so, because C_{i-1} crosses C_i, exactly one of the endpoints of C_i is strictly inside I. We consider two cases.

Case 1: $r(C_i) > r(C_0)$. In this case, C_i must be crossed on the left (by C_{i-1}) and $C_i \in \mathcal{R}$ and it does not cross C_0. Therefore, $C_0 \subsetneq C_i$ and

$$\ell(C_1) < \ell(C_i) \le \ell(C_0),$$

where the first inequality uses that the left endpoint of C_i is strictly inside I. Therefore, C_1 crosses both of C_0, C_i, and C_i is a strict ancestor of $A = C_0$. If $C_i = B$ we are done, otherwise, $A \subseteq B \subseteq C_i$, but because C_1 crosses both A and C_i, it also crosses B and we are done.

Case 2: $\ell(C_i) < \ell(C_1)$. In this case, C_i must be crossed on the right (by C_{i-1}) and $C_i \in \ell$ and it does not cross C_1. Therefore, we must have

$$r(C_1) \le r(C_i) < r(C_0),$$

where the second inequality uses that the right endpoint of C_i is strictly inside I. However, this implies that C_i also crosses C_0. Therefore, we can obtain a shorter path by excluding all cuts $C_1, \ldots, C_{i-1}$ and that is a contradiction. □

Lemma 4.4. *Let $A, B \in \mathcal{R}$ such that $A \cap B = \emptyset$, that is, they are not ancestors of each other. Then, they have a common ancestor, that is, there exists a set $C \in \mathcal{R}$ such that $A, B \subseteq C$.*

Without loss of generality (WLOG) assume $r(A) \le \ell(B)$. Let C be the highest ancestor of A in the hierarchy, that is, C has no ancestor. For the sake of contradiction, suppose $B \cap C = \emptyset$ (otherwise, C is an ancestor of B, and we are done). Therefore, $r(C) \le \ell(B)$. Consider the path of crossing cuts from C to B, say $C = C_0, \ldots, C_k = B$.

Let C_i be the first cut in this path such that $r(C_i) > r(C_0)$. Such a cut always exists as $r(B) > r(C)$. Because C_{i-1} crosses C_i and $r(C_{i-1}) \le r(C_0)$, C_{i-1} crosses C_i on the left and C_i is open on the right. We show that C_i is an ancestor of $C = C_0$ and we get a contradiction to C_0 having no ancestors (in $\mathcal{R}$). If $\ell(C_0) < \ell(C_i)$, then C_i crosses C_0 on the right and that is a contradiction. Therefore, we must have $C_0 \subseteq C_i$, that is, C_i is an ancestor of C_0.

It follows from the previous lemma that each of the left and right hierarchies have a unique cut with no ancestors.

Lemma 4.5. *If A is a cut in $\mathcal{R}$ such that $r(A) < m$, then A has a strict ancestor. Similarly, if $A \in \mathcal{L}$ satisfies $\ell(A) > 1$, then it has a strict ancestor.*

Fix a cut $A \in \mathcal{R}$. If there is a cut in $B \in \mathcal{R}$ such that $r(B) > r(A)$, then either B is a strict ancestor of A in which case we are done, or $A \cap B = \emptyset$, but then by Lemma 4.4 A, B have a common ancestor C, and C must be a strict ancestor of A and we are done.

Now, suppose for any $R \in \mathcal{R}$, $r(R) \le r(A)$. Therefore, there must be a cut $B \in \mathcal{L}$ such that $r(B) > r(A)$ (otherwise we should have less than m atoms in our polygon). The cut B must be crossed on the right by a cut $C \in \mathcal{R}$. However, then, we must have $r(C) > r(B) > r(A)$ which is a contradiction. □

Corollary 4.1. *If $A \in \mathcal{C}$ has no strict ancestor, then $r(A) = m$ if $A \in \mathcal{R}$ and $\ell(A) = 1$ otherwise.*

Lemma 4.6 (Polygons Are Near Minimum Cuts). *For any polygon with atoms $a_0, \ldots a_{m-1}$ (where a_0 is the root), we have $x(\delta(a_1 \cup \cdots \cup a_{m-1})) \le 2 + 4\eta$.*

Let $A \in \ell$ and $B \in \mathcal{R}$ be the unique cuts in the left/right hierarchy with no ancestors. Note that A and B are crossing (because there is a cut C that crosses A on the right, and B is an ancestor of C). Therefore, because A, B are both 2η near min cuts (with respect to x), by Lemma 2.1, $A \cup B$ is a 4η near min cut. □

Lemma 4.7 (Root Neighbors). *For any polygon with atoms $a_0, \ldots a_{m-1}$ (where a_0 is the root), we have $x(E(a_0, a_1))$, $x(E(a_0, a_{m-1})) \ge 1 - 2\eta$.*

Here we prove $x(E(a_0, a_1)) \ge 1 - 2\eta$. One can prove $x(E(a_0, a_{m-1})) \ge 1 - 2\eta$ similarly. Let $A \in \mathcal{L}$ and $B \in \mathcal{R}$ be the unique cuts in the left/right hierarchy with no ancestors. First, observe that if $\ell(B) = 2$, then becaus A, B are crossing, by Lemma 2.2, we have

$$x(E(A \smallsetminus B, \overline{A \cup B})) = x(E(a_1, a_0)) \ge 1 - \eta.$$

By definition of atoms, there exists a cut $C \in \mathcal{C}$ such that either $\ell(C) = 2$ or $r(C) = 2$; but if $r(C) = 2$ we must have $\ell(C) = 1$ in which case C cannot be crossed, so this does not happen. Therefore, we must have $\ell(C) = 2$. If $C \in \mathcal{R}$, then because C is a descendent of B, we must have $\ell(B) = 2$, and we are done by the previous paragraph.

Otherwise, suppose $C \in \mathcal{L}$. We claim that B crosses C. This is because, C is crossed on the right by some cut B' and B is an ancestor of B', so $B \cap C \ne \emptyset$ and $C \not\subseteq B$ because $\ell(B) > 2$. Therefore, by Lemma 2.1, $B \cup C$ is a 4η near min cut. Because A crosses $B \cup C$, by Lemma 2.2, we have

$$x(E(A \smallsetminus (B \cup C), \overline{A \cup B \cup C})) = x(E(a_1, a_0)) \ge 1 - 2\eta$$

as desired.

Lemma 4.8. *For any pair of atoms a_i, a_{i+1} where $1 \le i \le m-2$ we have $x(\delta(\{a_i, a_{i+1}\})) \le 2 + 12\eta$, so $x(E(a_i, a_{i+1})) \ge 1 - 6\eta$.*

We prove the following claim: There exists $j \le i$ such that $x(\delta(\{a_j, \dots, a_{i+1}\})) \le 2 + 6\eta$. Then, by a similar argument we can find $j' \ge i+1$ such that $x(\delta(\{a_i, \dots, a_{j'}\})) \le 2 + 6\eta$. By Lemma 2.1, it follows that $x(\delta(\{a_i, a_{i+1}\})) \le 2 + 12\eta$. Because $x(\delta(a_i)), x(\delta(a_{i+1})) \ge 2$, we have

$$x(\delta(\{a_i, a_{i+1}\})) + 2x(E(a_i, a_{i+1})) \ge 4.$$

However, because of the bound on $x(\delta(\{a_i, a_{i+1}\}))$, we must have $x(E(a_i, a_{i+1})) \ge 1 - 6\eta$ as desired.

It remains to prove the claim. First, observe that there is a cut A separating a_{i+1}, a_{i+2} (If $i+1 = m-1$, then $a_{i+2} = a_0$); therefore, either $\ell(A) = i+2$ or $r(A) = i+2$. If $r(A) = i+2$ then, A is the cut we are looking for and we are done. Therefore, assume $\ell(A) = i+2$.

Case 1: $A \in \mathcal{L}$. Let $L \in \mathcal{L}$ be the strict parent of A. If $\ell(L) \le i$ then we are done (because there is a cut $R \in \mathcal{R}$ crossing A, L on the right so $L \smallsetminus (A \cup R)$ is the cut that we want. If $\ell(L) = i+1$, then let L' be the strict parent of L). Then, there is a cut $R \in \mathcal{R}$ crossing A, L and a cut R' crossing L, L'. First, because both R, R' cross L (on the right) they have a nonempty intersection, so one of them say R' is an ancestor of the other (R) and therefore R' must intersect A. On the other hand, because R' crosses L and $\ell(L) = i+1$, $\ell(R') \ge i+2 = \ell(A)$. Because R' intersect A, either they cross, or $A \subseteq R'$, so we must have $x(\delta(A \cup R)) \le 2 + 4\eta$. Finally, because R' crosses L' (on the right), we have $x(\delta(L' \smallsetminus (A \cup R))) \le 2 + 6\eta$, and $L' \smallsetminus (A \cup R)$ is our desired set.

Case 2: $A \in \mathcal{R}$. We know that A is crossed on the left by, say, $L \in \ell$. If $\ell(L) \le i$, we are done, since then $L \smallsetminus A$ is the cut that we seek and we get $x(\delta(L \smallsetminus A)) \le 2 + 4\eta$.

Suppose then that $\ell(L) = i+1$. Let L' be the strict parent of L, which must have $\ell(L') \le i$. If L' crosses A, then $L' \smallsetminus A$ is the cut we seek and we get $x(\delta(L \smallsetminus A)) \le 2 + 4\eta$.

Finally, if L' does not cross A, that is, $r(A) \le r(L')$, then consider the cut $R \in \mathcal{R}$ that crosses L and L' on the right. Because $r(L) < r(A)$, and A is not crossed on the right, it must be that $\ell(R) = i+2$. In this case, $L' \smallsetminus R$ is the cut we want, and we get $x(\delta(L' \smallsetminus R)) \le 2 + 4\eta$. □

Lemma 4.9 (Atoms Are Near Minimum Cuts). *For any $1 \le i \le m-1$, we have $x(\delta(a_i)) \le 2 + 14\eta$.*

By Lemma 4.8, $x(\delta(\{a_i, a_{i+1}\})) \le 2 + 12\eta$ (in the special case $i = m-1$, we take the pair a_{i-1}, a_i). There must be a 2η-near minimum cut C (w.r.t., x) separating a_i from a_{i+1}. Then either $a_i = C \cap \{a_i, a_{i+1}\}$ or $a_i = \{a_i, a_{i+1}\} \smallsetminus C$. In either case, we get $x(\delta(a_i)) \le 2 + 14\eta$ by Lemma 2.1.

4.4. Happy Polygons

Definition 4.6 (*A*, *B*, *C*-Polygon Partition). Let u be a polygon with atoms $a_0, \dots, a_{m-1}$ with root a_0 where a_1, a_{m-1} are the atoms left and right of the root. The *A*, *B*, *C*-polygon partition of u is a partition of edges of $\delta(u)$ into sets $A = E(a_1, a_0)$ and $B = E(a_{m-1}, a_0)$, $C = \delta(u) \smallsetminus A \smallsetminus B$.

By Theorem 4.3, $x(A), x(B) \ge 1 - \epsilon_\eta$ and $x(C) \le \epsilon_\eta$ where we set

$$\epsilon_\eta = 14\eta \tag{4.5}$$

as needed for Theorem 4.3.

Definition 4.7 (Leftmost and Rightmost Cuts). Let u be a polygon with atoms $a_0, \dots, a_{m-1}$ and arcs labeled $1, \dots, m$ corresponding to a connected component $\mathcal{C}$ of η-near minimum cuts (w.r.t., z). We call any cut $C \in \mathcal{C}$ with $\ell(C) = 1$ a *leftmost* cut of u and any cut $C \in \mathcal{C}$ with $r(C) = m$ a *rightmost* cut of u. We also call a_1 the leftmost atom of u (respectively, a_{m-1} the rightmost atom).

Observe that by Corollary 4.1, any cut that is not a leftmost or a rightmost cut has a strict ancestor.

Definition 4.8 (Happy Polygon). Let u be a polygon with polygon partition *A*, *B*, *C*. For a spanning tree T, we say that u is *happy* if

$$A_T \text{ and } B_T \ \text{ odd}, C_T = 0.$$

We say that u is *left-happy* (respectively *right-happy*. if

$$A_T \,\text{odd}, C_T = 0,$$

(respectively, B_T odd, $C_T = 0$).

Definition 4.9 (Relevant Cuts). Given a polygon u corresponding to a connected component $\mathcal{C}$ of cuts crossed on one side with atoms $a_0, \dots, a_{m-1}$, define a family of relevant cuts

$$\mathcal{C}' = \mathcal{C} \cup \{a_i : 1 \le i \le m-1, z(\delta(a_i)) \le 2 + \eta\}.$$

Atoms of u are always $\epsilon_\eta/2$-near minimum cuts w.r.t., z but not necessarily η-near minimum cuts. The following theorem is the main result of this section.

Theorem 4.4 (Happy Polygons and Cuts Crossed on One Side). *Let $G = (V, E, x)$ for x be an LP solution and $z = (x + OPT)/2$. For a connected component $\mathcal{C}$ of near minimum cuts of z, let u be the polygon with atoms $a_0, a_1 \dots a_{m-1}$ with polygon partition A, B, C. For μ an arbitrary distribution of spanning trees with marginals x, $\beta > 0$, there is a random vector $s^* : E^* \to \mathbb{R}_{\ge 0}$ (as a function of $T \sim \mu$) such that for any vector $s : E \to \mathbb{R}$, where $s_e \ge -\beta x_e$ for all $e \in E$ the following holds:*

- *If u is happy then, for any cut $S \in \mathcal{C}'$ if $\delta(S)_T$ is odd then we have $s(\delta(S)) + s^*(\delta(S)) \ge 0$,*
- *For any $S \in \mathcal{C}'$ that is not a rightmost/leftmost cut or rightmost/leftmost atom, if $\delta(S)_T$ is odd, then we have $s(\delta(S)) + s^*(\delta(S)) \ge 0$.*
- *For all OPT edges $e_2^*, \dots, e_{m-1}^*$ with respect to the above polygon, $\mathbb{E}[s^*_{e_i^*}] \le 181\eta\beta$. $\mathbb{E}[s^*_{e^*}] = 0$ for all other OPT edges.*

Before proving the previous theorem, we study a special case.

Lemma 4.10 (Triangles as Degenerate Polygons). *Let $S = X \cup Y$ where X, Y, S are ϵ_η-near min cuts (w.r.t., x) and each of these sets is a contiguous interval around the OPT cycle. Then, viewing X as a_1 and Y as a_2 (and $a_0 = \overline{X \cup Y}$), the previous theorem holds viewing S as a degenerate polygon.*

In this case, $A = E(a_1, a_0), B = E(a_2, a_0), C = \emptyset$. For the OPT edge e^* between X, Y we define $\mathcal{I}_1(e^*)$ to be the event that at least one of $T \cap E(X)$, $T \cap E(Y)$, $T \cap E(S)$ is not a tree. Whenever this happens we define $s^*_{e^*} = 2.05 \cdot \beta$. If S is left-happy we need to show when $\delta(X)_T$ is odd, then $s(\delta(X)) + s^*(\delta(X)) \geq 0$. This is because when S is left-happy we have A_T is odd (and $C_T = 0$), so either $\mathcal{I}_1(e^*)$ does not happen and $\delta(X)_T$ is even, or it happens in which case $s(\delta(X)) + s^*(\delta(X)) \geq 0$ as $s(\delta(X)) \geq -(2 + 2\eta)\beta$ and $s^*_{e^*} = 2.05\beta$. Finally, observe that by Corollary 2.3, $\mathbb{P}[\mathcal{I}_1(e^*)] \leq 3\epsilon_\eta$, so $\mathbb{E}[s^*_{e^*}] = 3\epsilon_\eta \cdot 2.05\beta \leq 87\eta\beta$ using $\eta < 1/100$ and ϵ_η as defined in Equation (4.5).

Lemma 4.11. *For every cut $A \in \mathcal{C}$ that is not a leftmost or a rightmost cut, $\mathbb{P}[\delta(A)_T = 2] \geq 1 - 22\eta$.*

Assume $A \in \mathcal{R}$; the other case can be proven similarly. Let B be the strict parent of A. By Lemma 4.3 there is a cut $C \in \mathcal{L}$ that crosses A, B on their left. It follows by Lemma 2.1 that $C \smallsetminus A, C \cap A$ are 4η near minimum cuts (w.r.t., x). Therefore, by Corollary 2.3, $\mathbb{P}[E(A \cap C, C \smallsetminus A)_T = 1] \geq 1 - 5\eta$. On the other hand, $B \smallsetminus (A \cup C)$ is a 6η near minimum cut and $A \smallsetminus C, B \smallsetminus C$ are 4η near min cuts (w.r.t., x). Therefore, by Corollary 2.3, $\mathbb{P}[E(A \smallsetminus C, B \smallsetminus (A \cup C))_T = 1] \geq 1 - 7\eta$.

Finally, by Lemma 2.2, $x(E(A \cap C, C \smallsetminus A)), x(E(A \smallsetminus C, B \smallsetminus (A \cup C))) \geq 1 - 3\eta$. Because A is a 2η near min cut (w.r.t., x), all remaining edges have fractional value at most 8η, so with probability $1 - 8\eta$, T does not choose any of them. Taking a union bound over all of these events, $\mathbb{P}[\delta(A)_T = 2] \geq 1 - 22\eta$. □

Lemma 4.12. *For any atom $a_i \in \mathcal{C}'$ that is not the leftmost or the rightmost atom, we have*

$$\mathbb{P}[\delta(a_i)_T = 2] \geq 1 - 42\eta.$$

By Lemma 4.8, $x(\delta(\{a_i, a_{i+1}\})) \leq 2 + 12\eta$, and by Lemma 4.9, $x(\delta(a_{i+1})) \leq 2 + 14\eta$ (also recall by the assumption of lemma $x(\delta(a_i)) \leq 2 + 2\eta$, Therefore, by Corollary 2.3,

$$\mathbb{P}[E(a_i, a_{i+1})_T = 1], \mathbb{P}[E(a_{i-1}, a_i)_T = 1] \geq 1 - 14\eta,$$

where the second inequality holds similarly. Also, by Lemma 4.8, $x(E(a_{i-1}, a_i)), x(E(a_i, a_{i+1})) \geq 1 - 6\eta$. Because $x(\delta(a_i)) \leq 2 + 2\eta$, $x(E(a_i, \overline{a_{i-1} \cup a_i \cup a_{i+1}})) \leq 14\eta$. Therefore,

$$\mathbb{P}[T \cap E(a_i, \overline{a_{i-1} \cup a_i \cup a_{i+1}}) = \emptyset] \geq 1 - 14\eta.$$

Finally, by the union bound, all events occur with probability at least $1 - 42\eta$. □

Let $e^*_1, \ldots, e^*_m$ be the OPT edges mapped to the arcs $1, \ldots, m$ of the component $\mathcal{C}$, respectively.

Lemma 4.13. *There is a mapping[10] of cuts in $\mathcal{C}'$ to OPT edges $e^*_2, \ldots e^*_{m-1}$ such that each OPT edge has at most four cuts mapped to it, an OPT edge e^* is mapped to a cut S only if $e^* \in \delta(S)$, and every atom of the polygon in $\mathcal{C}'$ gets mapped to two (not necessarily distinct) OPT edges.*

Consider first the set of cuts in $\mathcal{C}'_\mathcal{R} := \mathcal{R} \cup \{a_i : 1 \leq i \leq m-1, z(\delta(a_i)) \leq 2 + \eta\}$ and similarly $\mathcal{C}'_\ell := \mathcal{L} \cup \{a_i : 1 \leq i \leq m-1, z(\delta(a_i)) \leq 2 + \eta\}$. Observe that this is also a laminar family. Atoms are in both $\mathcal{C}'_\mathcal{R}$ and $\mathcal{C}'_\mathcal{L}$. We define a map from cuts in $\mathcal{C}'_R$ to OPT edges such that every OPT edge $e^*_2, \ldots, e^*_{m-1}$ gets at most two cuts mapped to it. A similar argument works for cuts in $\mathcal{C}'_\ell$.

For any $2 \leq i \leq m-1$, we map

$$\arg\max_{A \in \mathcal{C}'_\mathcal{R} : \ell(A) = i} |A| \quad \text{and} \quad \arg\max_{A \in \mathcal{C}'_\mathcal{R} : r(A) = i} |A|$$

to e^*_i, where recall $\ell(A)$ is the OPT edge leaving A on the left side and $r(A)$ the OPT edge leaving on the right. By construction, each OPT edge gets at most two cuts mapped to it.

Furthermore, we claim every cut $A \in \mathcal{C}'_\mathcal{R}$ gets mapped to at least one OPT edge. For the sake of contradiction let $A \in \mathcal{C}'_\mathcal{R}$ be a cut that is not mapped to any OPT edge. First, a_1 is mapped to edge e^*_2 (in both hierarchies) and a_{m-1} is mapped to edge e^*_{m-1}. Otherwise, if $A \in \mathcal{R}$, $\ell(A) \neq 1$. Furthermore, if $A \in \mathcal{R}$ and $r(A) = m$, then A is definitely the largest cut with left endpoint $\ell(A)$. Therefore, assume $1 < \ell(A) < r(A) < m$. Let $B = \arg\max_{B \in \mathcal{C}'_\mathcal{R} : \ell(B) = \ell(A)} |B|$ and let $C = \arg\max_{B \in \mathcal{C}'_\mathcal{R} : r(C) = r(A)} |C|$. Because A is not mapped to any OPT edge but B, C are mapped by previous definition, we must have $B, C \neq A$. However, this implies $A \subsetneq B, C$. Also, this means B, C cross; but this is a contradiction with $\mathcal{R}$ being a laminar family. □

Definition 4.10 (Happy Cut). We say a leftmost cut $L \in \mathcal{L}$ is *happy* if

$$E(L, \overline{a_0 \cup L})_T = 1.$$

Similarly, the leftmost atom a_1 is *happy* if $E(a_1, \overline{a_0 \cup a_1})_T = 1$. Define rightmost cuts in u or the rightmost atom in u to be happy, similarly.

By definition, if leftmost cut L is happy and u is left happy then L is even, that is, $\delta(L)_T = 2$. Similarly, a_1 is even if it is happy and u is left-happy.

Lemma 4.14. *For every leftmost or rightmost cut A in u that is an η-near min cut w.r.t. z, $\mathbb{P}[A \text{ happy}] \geq 1 - 10\eta$, and for the leftmost atom a_1 (respectively, rightmost atom a_{m-1}), if it is an η-near min cut then $\mathbb{P}[a_1 \text{ happy}] \geq 1 - 24\eta$ (respectively, $\mathbb{P}[a_{m-1} \text{ happy}] \geq 1 - 24\eta$).*

Recall that if A is a η-near min cut w.r.t. z then it is a 2η-near min cut w.r.t. x. Also, recall for a cut $L \in \mathcal{L}$, L_R is the near minimum cut crossing L on the right that minimizes the intersection (Definition 4.2). We prove this for the leftmost cuts and the leftmost atom; the other case can be proven similarly. Consider a cut $L \in \mathcal{L}$. Because by

Lemma 2.1, $L_R \cap L, L_R \smallsetminus L$ are 4η near min cuts (w.r.t., x) and L_R is a 2η near min cut, by Corollary 2.3, $\mathbb{P}[E(L_R \cap L, L_R \smallsetminus L)_T = 1] \geq 1 - 5\eta$. On the other hand, by Lemma 2.2, $x(E(L_R \cap L, L_R \smallsetminus L)) \geq 1 - \eta$, and by Lemma 4.7, $x(E(L, a_0)) \geq 1 - 2\eta$. It follows that

$$x(\delta(L) \smallsetminus E(L_R \cap L, L_R \smallsetminus L) \smallsetminus E(L, a_0)) \leq 5\eta.$$

Therefore, by the union bound, $\mathbb{P}[L \text{ happy}] \geq 1 - 10\eta$, because if $(\delta(L) \smallsetminus E(L_R \cap L, L_R \smallsetminus L) \smallsetminus E(L, a_0))_T = 0$ and $E(L_R \cap L, L_R \smallsetminus L)_T = 1$, then $E(L, \overline{a_0 \cup L})_T = 1$ and therefore L is happy.

Now consider the atom a_1, and suppose it is an η near min cut. By Lemma 4.8, $x(\delta(\{a_1, a_2\})) \leq 2 + 12\eta$, and by Lemma 4.9, $x(\delta(a_2)) \leq 2 + 14\eta$. Therefore, by Corollary 2.3, $\mathbb{P}[E(a_1, a_2)_T = 1] \geq 1 - 14\eta$. On the other hand, by Lemma 4.8, $x(E(a_1, a_2)) \geq 1 - 6\eta$, and by Lemma 4.7, $x(E(a_1, a_0)) \geq 1 - 2\eta$. Therefore,

$$\begin{aligned} &x(E(a_1, a_3 \cup \cdots \cup a_{m-1})) \\ &\quad \leq 2 + 2\eta - (1 - 6\eta) - (1 - 2\eta)) \leq 10\eta. \end{aligned}$$

Observe, a_1 is happy when both of these events occur; so, by the union bound, $\mathbb{P}[a_1 \text{ happy}] \geq 1 - 24\eta$ as desired. □

Proof of Theorem 4.4. Consider an OPT edge e_i^* for $1 < i < m$. For the at most four cuts mapped to e_i^* in Lemma 4.13, we define the following three events:

(i) A leftmost cut assigned to e_i^* is not happy. (Equivalently, a leftmost cut $L \in \mathcal{L} \cap \mathcal{C}'$ with $r(L) = i$ is not happy.)

(ii) A rightmost cut assigned to e_i^* is not happy. (Equivalently, a rightmost cut $R \in \mathcal{R} \cap \mathcal{C}'$ with $l(R) = i$ is not happy.[11])

(iii) A cut that is not leftmost or rightmost assigned to e_i^* is odd.

Observe that the cuts in (i) and (ii) are assigned to e_i^* in Lemma 4.13. We say an atom a is singly-mapped to e_i^* if in the matching a is only mapped to e_i^* once; otherwise, we say it is doubly mapped to e_i^*.

We say an event $\mathcal{I}_1(e_i^*)$ occurs if either (i), (ii), or (iii) occurs. If $\mathcal{I}_1(e_i^*)$ occurs then we set

$$s^*_{e_i^*} = \begin{cases} 2.05\beta & \text{If (i), (ii), or (iii) occurred for at least} \\ & \text{one non-atom cut in } \mathcal{C}', \text{or for an atom} \\ & \text{which is doubly-mapped to } e_i^* \\ 2.05\beta/2 & \text{Otherwise.} \end{cases}$$

If $\mathcal{I}_1(e_i^*)$ does not occur we set $s^*_{e_i^*} = 0$. First, observe that for any nonatom cut $S \in \mathcal{C}'$ that is not a leftmost or a rightmost cut, if $\delta(S)_T$ is odd, then if e_i^* is the OPT edge that S is mapped to, it satisfies $s^*_{e_i^*} = 2.05\beta$, so

$$\begin{aligned} s(\delta(S)) + s^*(\delta(S)) &\geq -x(\delta(S))\beta + s^*(e_i^*) \\ &\geq -(2 + 2\eta)\beta + 2.05\beta \geq 0, \end{aligned}$$

for $\eta < 1/100$. The same inequality holds for nonleftmost/rightmost atom cuts $a \in \mathcal{C}'$, which are doubly mapped to e_i^*. For nonleftmost/rightmost atom cuts $a \in \mathcal{C}'$, which are singly mapped to e_i^*, a is mapped (possibly even twice) to another edge e_j^* (note $j = i - 1$ or $i + 1$), and in this case, $s^*(e_i^*) + s^*(e_j^*) \geq 2.05\beta$, and again the previous inequality holds.

Now, suppose for a leftmost cut $S \in \mathcal{L} \cap \mathcal{C}'$ with $r(S) = i$ has $\delta(S)_T$ odd. If u is not left-happy, there is nothing to prove. If u is left-happy, then we must have S is not happy (as otherwise $\delta(S)_T$ would be even), so $\mathcal{I}_1(e_i^*)$ occurs, so similar to the previous inequality $s(\delta(S)) + s^*(\delta(S)) \geq 0$. The same holds for rightmost cuts and the leftmost/rightmost atoms in $\mathcal{C}'$ (leftmost/rightmost atoms are always doubly-mapped: a_1 to e_2^* and a_{m-1} to e_{m-1}^*).

It remains to upper bound $\mathbb{E}[s^*(e_i^*)]$ for $1 < i < m$. By Lemma 4.13, at most four cuts are mapped to e_i^*. Then, either there is an atom that is doubly mapped to e_i^* or there is not.

First suppose exactly one atom is doubly mapped to e_i^*. Then there are at most three cuts mapped to e_i^*, including that atom. The probability of an event of type (i) or (ii) occurring for the leftmost or rightmost atom is at most $1 - 24\eta$ by Lemma 4.14. Atoms that are not leftmost or rightmost are even with probability at least $1 - 42\eta$ by Lemma 4.12. Therefore, in the worst case, the doubly mapped atom is not leftmost or rightmost. For the remaining two cuts, leftmost and rightmost cuts are happy with probability at least $1 - 10\eta$ by Lemma 4.14, and (nonatom) non leftmost/rightmost cuts are even with probability at least $1 - 22\eta$ by Lemma 4.11. Therefore, in the worst case, the remaining two (nonatom) cuts mapped to e_i^* are not leftmost/rightmost. Therefore, if an atom is doubly mapped to e_i^*,

$$\mathbb{E}[s^*(e_i^*)] \leq 42\eta \cdot 2.05\beta + 2 \cdot 22\eta \cdot 2.05\beta \leq 177\eta\beta.$$

If two atoms are doubly mapped to e_i^*,

$$\mathbb{E}[s^*(e_i^*)] \leq 2 \cdot 42\eta \cdot 2.05\beta \leq 173\eta\beta.$$

Otherwise, any atoms mapped to e_i^* are singly mapped. In this case, if only an atom cut is odd/unhappy, we set $s^*(e_i^*) = 2.05\beta/2$. The probability of an event of type (i) or (ii) occurring for the leftmost or rightmost atom is at most $1 - 24\eta$ by Lemma 4.14, so we can bound the contribution of this event to $\mathbb{E}[s^*(e_i^*)]$ by $24\eta \cdot 2.05\beta/2$. Atoms that are not leftmost or rightmost are even with probability at least $1 - 42\eta$ by Lemma 4.12, and so we can bound their contribution by $42\eta \cdot 2.05\beta/2$. Therefore, in the worst case four nonleftmost/rightmost *nonatom* cuts are mapped to e_i^*, in which case,

$$\mathbb{E}[s^*(e_i^*)] \leq 4 \cdot 22\eta \cdot 2.05\beta = 181\eta\beta$$

as desired.

4.5. Hierarchy of Cuts and Proof of Theorem 4.2

Definition 4.11 (Hierarchy). For an LP solution x^0 with support $E_0 = E \cup \{e_0\}$ and x be x^0 restricted to E, a hierarchy

$\mathcal{H}$ is a *laminar* family of ϵ_η-near min cuts of $G=(V,E,x)$ with root $V \smallsetminus \{u_0,v_0\}$, where every cut $S \in \mathcal{H}$ is either a polygon cut (including triangles) or a degree cut and $u_0, v_0 \notin S$. Furthermore, every cut S is a union of its children. For any (nonroot) cut $S \in \mathcal{H}$, define the parent of S, $\mathsf{p}(S)$, to be the smallest cut $S' \in \mathcal{H}$ such that $S \subsetneq S'$.

For a cut $S \in \mathcal{H}$, let $\mathcal{A}(S) := \{u \in \mathcal{H} : \mathsf{p}(u) = S\}$. If S is a polygon cut, then we can order cuts in $\mathcal{A}(S)$, $u_1, \dots, u_{m-1}$ such that

- Each of $A = E(\overline{S}, u_1)$ and $B = E(u_{m-1}, \overline{S})$ satisfy $x(A)$, $x(B) \ge 1-\epsilon_\eta$.
- For any $1 \le i < m-1$, $x(E(u_i,u_{i+1})) \ge 1-\epsilon_\eta$.
- For $C = \cup_{i=2}^{m-2} E(u_i, \overline{S})$ we have $x(C) \le \epsilon_\eta$.

We call the sets A, B, C the polygon partition of edges in $\delta(S)$. We say S is left-happy when A_T is odd and $C_T = 0$ and right happy when B_T is odd and $C_T = 0$ and happy when A_T, B_T are odd and $C_T = 0$.

We abuse notation, and for an (LP) edge $e=(u,v)$ that is not a neighbor of u_0, v_0, let $\mathsf{p}(e)$ denote the smallest[12] cut $S' \in \mathcal{H}$ such that $u, v \in S'$. We say edge e is a *bottom edge* if $\mathsf{p}(e)$ is a polygon cut and we say it is a *top edge* if $\mathsf{p}(e)$ is a degree cut.

When S is a polygon cut, $u_1, \dots, u_{m-1}$ will be the atoms $a_1, \dots, a_{m-1}$ that we defined in the previous section, but a reader should understand this definition independent of the polygon definition that we discussed before; in particular, the reader no longer needs to worry about the details of specific cuts $\mathcal{C}$ that make up a polygon. Also, because $V \smallsetminus \{u_0, v_0\}$ is the root of the hierarchy, for any edge $e \in E$ that is not incident to u_0 or v_0, $\mathsf{p}(e)$ is well defined; therefore, all those edges are either bottom or top, and edges that are incident to u_0 or v_0 are neither bottom edges nor top edges.

The following observation is immediate from the above definition.

Observation 4.1. For any polygon cut $S \in \mathcal{H}$, and any cut $S' \in \mathcal{H}$ that is a descendant of S let $D = \delta(S') \cap \delta(S)$. If $D \neq \emptyset$, then exactly one of the following is true: $D \subseteq A$ or $D \subseteq B$ or $D \subseteq C$.

Theorem 4.5 (Main Payment Theorem). *For an LP solution x^0 and x be x^0 restricted to E and a hierarchy $\mathcal{H}$ for some $\epsilon_\eta \le 10^{-10}$ and any $\beta > 0$, the maximum entropy distribution μ with marginals x satisfies the following:*

(i) *There is a set of good edges $E_g \subseteq E \smallsetminus \delta(\{u_0,v_0\})$ such that any bottom edge e is in E_g and for any (nonroot) $S \in \mathcal{H}$ such that $\mathsf{p}(S)$ is a degree cut, we have $x(E_g \cap \delta(S)) \ge 3/4$.*

(ii) *There is a random vector $s : E_g \to \mathbb{R}$ (as a function of $T \sim \mu$) such that for all e, $s_e \ge -x_e\beta$ (with probability 1).*

(iii) *If a polygon cut u with polygon partition A, B, C is not left happy, then for any set $F \subseteq E$ with $\mathsf{p}(e) = u$ for all $e \in F$ and $x(F) \ge 1-\epsilon_\eta/2$, we have*

$$s(A) + s(F) + s^-(C) \ge 0,$$

where $s^-(C) = \sum_{e\in C} \min\{s_e, 0\}$. A similar inequality holds if u is not right happy.

(iv) *For every cut $S \in \mathcal{H}$ such that $\mathsf{p}(S)$ is not a polygon cut, if $\delta(S)_T$ is odd, then $s(\delta(S)) \ge 0$.*

(v) *For a good edge $e \in E_g$, $\mathbb{E}[s_e] \le -\epsilon_P \beta x_e$ (see Equation (7.4) for definition of ϵ_P).*

The previous theorem is the main part of the paper in which we use that μ is a SR distribution. See Section 7 for the proof. We use this theorem to construct a random vector s such that essentially for all cuts $S \in \mathcal{H}$ in the hierarchy $z/2 + s$ is feasible; furthermore, for a large fraction of "good" edges, we have that $\mathbb{E}[s_e]$ is negative and bounded away from zero.

As we will see in the this section, using part (iii) of the theorem we will be able to show that every leftmost and rightmost cut of any polygon is satisfied.

In the rest of this section, we use the previous theorem to prove Theorem 4.2. We start by explaining how to construct $\mathcal{H}$. Given the vector $z = (x + OPT)/2$ run the following procedure on the OPT cycle with the family of η-near minimum cuts of z that are crossed on at most one side.

For every connected component $\mathcal{C}$ of η near minimum cuts (w.r.t., z) crossed on at most one side, if $|\mathcal{C}| = 1$ then add the unique cut in $\mathcal{C}$ to the hierarchy. Otherwise, $\mathcal{C}$ corresponds to a polygon u with atoms $a_0, \dots, a_{m-1}$ (for some $m > 3$). Add $a_1, \dots, a_{m-1}$[13] and $\cup_{i=1}^{m-1} a_i$ to $\mathcal{H}$. Because every vertex except u_0, v_0 has degree 2, they all appear in the hierarchy as singletons. Therefore, every set in the hierarchy is the union of its children. Because $z(\delta(\{u_0,v_0\})) = 2$, the root of the hierarchy is always $V \smallsetminus \{u_0, v_0\}$.

Now, we name every cut in the hierarchy. For a cut S if there is a connected component of at least two cuts with union equal to S, then call S a polygon cut with the A, B, C partitioning as defined in Definition 4.6. If S is a cut with exactly two children X, Y in the hierarchy, then also call S a polygon cut,[14] $A = E(X, \overline{X} \smallsetminus Y)$, $B = E(Y, \overline{Y} \smallsetminus X)$, and $C = \emptyset$. Otherwise, call S a degree cut.

Fact 4.2. *The previous procedure produces a valid hierarchy for $\epsilon_\eta \ge 14\eta$.*

First observe that whenever $|\mathcal{C}| = 1$ the unique cut in $\mathcal{C}$ is a 2η near min cut (w.r.t, x) that is not crossed. For a polygon cut S in the hierarchy, by Lemma 4.6, the set S is a ϵ_η near min cut w.r.t., x. If S is an atom of a polygon, then by Lemma 4.9, S is a ϵ_η near min cut.

Now, it remains to show that for a polygon cut S we have a valid ordering $u_1, \dots, u_k$ of cuts in $\mathcal{A}(S)$. If S is a nontriangle polygon cut, the $u_1, \dots, u_k$ are exactly atoms of the polygon of S and $x(A), x(B) \ge 1-\epsilon_\eta$ and $x(C) \le \epsilon_\eta$ and $x(E(u_i,u_{i+1})) \ge 1-\epsilon_\eta$ follow by Theorem 4.3. For a triangle cut $S = X \cup Y$ because S, X, Y are ϵ_η-near min cuts (by the previous paragraph), we get $x(A), x(B) \ge 1-\epsilon_\eta$ as desired, by Lemma 2.3. Finally, because $x(\delta(X)), x(\delta(Y)) \ge 2$ we have $x(E(X,Y)) \ge 1-\epsilon_\eta$. □

The following observation is immediate.

Observation 4.2. Each cut $S \in \mathcal{H}$ corresponds to a contiguous interval around OPT cycle. For a polygon u (or a triangle) with atoms $a_0, \ldots, a_{m-1}$ for $m \geq 3$, we say an OPT edge e^* is *interior* to u if $e^* \in E^*(a_i, a_{i+1})$ for some $1 \leq i \leq m-2$. Any OPT edge e^* is interior to at most one polygon.

Theorem 4.2. *Let x^0 be a feasible solution of the Held-Karp relaxation with support $E_0 = E \cup \{e_0\}$ and x be x^0 restricted to E. Let μ be the max entropy distribution with marginals x. For $\eta \leq 10^{-12}$, $\beta > 0$, there is a set $E_g \subset E \setminus \delta(\{u_0, v_0\})$ of good edges and two functions $s : E_0 \to \mathbb{R}$ and $s^* : E^* \to \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$) such that*

(i) *For each edge $e \in E_g$, $s_e \geq -x_e\beta$ and for any $e \in E \setminus E_g$, $s_e = 0$.*

(ii) *For each η-near-min cut S w.r.t. z, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.*

(iii) *We have $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ for all edges $e \in E_g$ and $\mathbb{E}[s^*_{e^*}] \leq 218\eta\beta$ for all OPT edges $e^* \in E^*$. for ϵ_P defined in (7.4).*

(iv) *For every η-near minimum cut S of z crossed on (at most) one side such that $S \neq V \setminus \{u_0, v_0\}$, $x(\delta(S) \cap E_g) \geq 3/4$.*

For ϵ_η as in Equation (4.5), let E_g, s be as defined in Theorem 4.5, and let $s_{e_0} = \infty$. Also, let s^* be the sum of the s^* vectors from Theorem 4.1 and Theorem 4.4. (i) follows (ii) of Theorem 4.5. $\mathbb{E}[s^*_{e^*}] \leq 218\eta\beta$ follows from Theorem 4.1 and Theorem 4.4 and the fact that every OPT edge is interior to at most one polygon. Also, $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ for edges $e \in E_g$ follows from (v) of Theorem 4.5.

Now, we verify (iv): For any (nonroot) cut $S \in \mathcal{H}$ such that $\mathsf{p}(S)$ is not a polygon cut $x(\delta(S) \cap E_g) \geq 3/4$ by (i) of Theorem 4.5. The only remaining η-near minimum cuts are sets S that are either atoms or near minimum cuts in the component $\mathcal{C}$ corresponding to a polygon u. Therefore, by Lemma 2.3, $x(\delta(S) \cap \delta(u)) \leq 1 + \epsilon_\eta$. By (i) of Theorem 4.5 all edges in $\delta(S) \setminus \delta(u)$ are in E_g. Therefore, $x(\delta(S) \cap E_g)) \geq 1 - \epsilon_\eta \geq 3/4$.

It remains to verify (ii): We consider four groups of cuts.

Type 1: Near minimum cuts S such that $e_0 \in \delta(S)$. Then, because $s_{e_0} = \infty$, $s(\delta(S)) + s^*(\delta(S)) \geq 0$.

Type 2: Near minimum cuts $S \in \mathcal{H}$ where $\mathsf{p}(S)$ is not a polygon cut. By (iv) of Theorem 4.5 and that $s^* \geq 0$ the inequality follows.

Type 3: Near minimum cuts S crossed on both sides. Then, the inequality follows by Theorem 4.1 and the fact that $s_e \geq -\beta x_e$ for all $e \in E$.

Type 4: Near minimum cuts S that are crossed on one side (and not in $\mathcal{H}$) or $S \in \mathcal{H}$ and $\mathsf{p}(S)$ is a (nontriangle) polygon cut. In this case S must be an atom or a η-near minimum cut (w.r.t., z) in some polygon $u \in \mathcal{H}$. If S is not a leftmost cut/atom or a rightmost cut/atom, then the inequality follows by Theorem 4.4. Otherwise, say S is a leftmost cut. If u is left-happy then by Theorem 4.4 the inequality is satisfied. Otherwise, for $F = \delta(S) \setminus \delta(u)$, by Lemma 2.3, we have $x(F) \geq 1 - \epsilon_\eta/2$. Therefore, by (iii) of Theorem 4.5, we have

$$s(\delta(S)) + s^*(\delta(S)) \geq s(A) + s(F) + s^-(C) \geq 0$$

as desired. Because S is a leftmost cut, we always have $A \subseteq \delta(S)$. However, C may have an unpredictable intersection with $\delta(S)$; in particular, in the worst case only edges of C with negative slack belong to $\delta(S)$. A similar argument holds when S is the leftmost atom or a rightmost cut/atom.

Type 5: Near min cut S is the leftmost atom or the rightmost atom of a triangle u. This is similar to the previous case except we use Lemma 4.10 to argue that the inequality is satisfied when u is left happy. □

4.6. Hierarchy Notation

In the rest of the paper, we will not work with z, OPT edges, or the notion of polygons. Therefore, practically, by Definition 4.11, from now on, a reader can just think of every polygon as a triangle. In the rest of the paper we adopt the following notation.

We abuse notation and call any $u \in \mathcal{A}(S)$ an atom of S.

Definition 4.12 (Edge Bundles, Top Edges, and Bottom Edges). For every degree cut S and every pair of atoms $u, v \in \mathcal{A}(S)$, we define a *top edge bundle* $\mathbf{f} = (u, v)$ such that

$$\mathbf{f} = \{e = (u', v') \in E : \mathsf{p}(e) = S, u' \in u, v' \in v\}.$$

In the previous definition, u', v' are actual vertices of G.

For every polygon cut S, we define the *bottom edge bundle* $\mathbf{f} = \{e : \mathsf{p}(e) = S\}$.

We will always use bold letters to distinguish top edge bundles from actual LP edges. Also, we abuse notation and write $x_{\mathbf{e}} := \sum_{f \in \mathbf{e}} x_f$ to denote the total fractional value of all edges in this bundle.

In the rest of the paper, unless otherwise specified, we work with edge bundles and sometimes we just call them edges.

For any $u \in \mathcal{H}$ with $\mathsf{p}(u) = S$, we write

$$\delta^{\uparrow}(u) := \delta(u) \cap \delta(S),$$
$$\delta^{\rightarrow}(u) := \delta(u) \setminus \delta(S).$$
$$E^{\rightarrow}(S) := \{e = (u_i, u_j) : u_i, u_j \in \mathcal{A}(S), u_i \neq u_j\}.$$

Also, for a set of edges $A \subseteq \delta(u)$, we write $A^{\rightarrow}, A^{\uparrow}$ to denote $A \cap \delta^{\rightarrow}(u), A \cap \delta^{\uparrow}(u)$, respectively (when u is clear in context). Note that $E^{\rightarrow}(S) \subseteq E(S)$ includes only edges between atoms of S and not all edges between vertices in S.

Finally, for a set of edges F and an edge bundle $\mathbf{e}$, we define $F_{-\mathbf{e}} = F \setminus \mathbf{e}$, and similarly $F_{+\mathbf{e}} = F \cup \mathbf{e}$.

5. Probabilistic Statements

5.1. Gurvits' Machinery and Generalizations

The following is the main result of this section.

Proposition 5.1. *Given a SR distribution $\mu : 2^{[n]} \to \mathbb{R}_+$, let $A_1, \dots, A_m$ be random variables corresponding to the number of elements sampled from m disjoint sets, and let integers $n_1, \dots, n_m \geq 0$ be such that for any $S \subseteq [m]$,*

$$\mathbb{P}\left[\sum_{i\in S} A_i \geq \sum_{i\in S} n_i\right] \geq \epsilon,$$
$$\mathbb{P}\left[\sum_{i\in S} A_i \leq \sum_{i\in S} n_i\right] \geq \epsilon,$$

it follows that,

$$\mathbb{P}[\forall i : A_i = n_i] \geq f(\epsilon)\mathbb{P}[A_1 + \cdots + A_m = n_1 + \cdots + n_m],$$

where $f(\epsilon) \geq \epsilon^{2^m} \prod_{k=2}^{m} \frac{1}{\max\{n_k, n_1+\cdots+n_{k-1}\}+1}$.

We remark that in applications of the previous statement, it is enough to know that for any set $S \subseteq [m]$, $\sum_{i\in S} n_i - 1 < \mathbb{E}[\sum_{i\in S} A_i] < \sum_{i\in S} n_i + 1$. Because, then, by Lemma 2.5, we can prove a lower bound on the probability that $\sum_{i\in S} A_i = \sum_{i\in S} n_i$.

We also remark the previous lower bound of $f(\epsilon)$ is not tight; in particular, we expect the dependency on m should only be exponential (not doubly exponential). We leave it as an open problem to find a tight lower bound on $f(\epsilon)$.

Let $\mathcal{E}$ be the event $A_1 + \cdots + A_m = n_1 + \cdots + n_m$.

$$\begin{aligned}\mathbb{P}[1 \leq i \leq m : A_i = n_i] &= \mathbb{P}[\mathcal{E}]\mathbb{P}[A_m = n_m | \mathcal{E}]\mathbb{P}[A_{m-1} = n_{m-1} | A_m = n_m, \mathcal{E}]\\ &\quad \dots \mathbb{P}[A_2 = n_2 | A_3 = n_3, \dots, S_{A_m} = n_m, \mathcal{E}]\end{aligned}$$

Therefore, to prove the statement, it is enough to prove that for any $2 \leq k \leq n$,

$$\mathbb{P}[A_k = n_k | A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E}] \geq \epsilon^{2^{m-k+1}} \frac{1}{\max\{n_k, n_1 + \cdots + n_{k-1}\} + 1}. \tag{5.1}$$

By the following Claim 5.1,

$$\mathbb{P}[A_k \geq n_k | A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E}] \geq \epsilon^{2^{m-k+1}},$$
$$\mathbb{P}[A_k \leq n_k | A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E}] \geq \epsilon^{2^{m-k+1}}.$$

Therefore, (5.1) simply follows by Lemma 5.1. Now we prove this claim.

Claim 5.1. *Let $[k] := \{1, \dots, k\}$. For any $2 \leq k \leq m$, and any set $S \subsetneq [k]$,*

$$\mathbb{P}\left[\sum_{i\in S} A_i \geq \sum_{i\in S} n_i | A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E}\right] \geq \epsilon^{2^{m-k+1}},$$
$$\mathbb{P}\left[\sum_{i\in S} A_i \leq \sum_{i\in S} n_i | A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E}\right] \geq \epsilon^{2^{m-k+1}}.$$

We prove by induction. First, notice for $k = m$, the statement holds just by lemma's assumption and Lemma 5.2. Now, suppose the statement holds for $k+1$. Now, fix a set $S \subsetneq [k]$. Let $\overline{S} = [k] \setminus S$. Define $A = \sum_{i\in S} A_i$ and $B = \sum_{i\in \overline{S}} A_i$, and similarly define n_A, n_B. By the induction hypothesis,

$$\epsilon^{2^{m-k}} \leq \mathbb{P}[A \leq n_A | A_{k+2} = n_{k+2}, \dots, A_m = n_m, \mathcal{E}].$$

The same statement holds for events $A \geq n_A, B \leq n_B, B \geq n_B, A + B \geq n_A + n_B, A + B \leq n_A + n_B$. Let $\mathcal{E}_{k+1}$ be the event $A_{k+2} = n_{k+2}, \dots, A_m = n_m, \mathcal{E}$. Conditioned on $\mathcal{E}_{k+1}$, $A + B = n_A + n_B$ if and only if $A_{k+1} = n_{k+1}$. By Lemma 5.1, $\mathbb{P}[A + B = n_A + n_B | \mathcal{E}_{k+1}] > 0$. Therefore, by Lemma 5.2,

$$\mathbb{P}[A \geq n_A | A + B = n_A + n_B, \mathcal{E}_{k+1}],$$
$$\mathbb{P}[A \leq n_A | A + B = n_A + n_B, \mathcal{E}_{k+1}] \geq (\epsilon^{2^{m-k}})^2 = \epsilon^{2^{m-k+1}}$$

as desired.

This finishes the proof of Proposition 5.1.

Lemma 5.1. *Let $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$ be a d-homogeneous SR distribution. If for an integer $0 \leq k \leq d$, $\mathbb{P}_{S\sim\mu}[|S| \geq k] \geq \epsilon$ and $\mathbb{P}_\mu[|S| \leq k] \geq \epsilon$. Then,*

$$\mathbb{P}[|S| = k] \geq \min\left\{\frac{\epsilon}{k+1}, \frac{\epsilon}{d-k+1}\right\},$$
$$\mathbb{P}[|S| = k] \geq \min\left\{p_m, \epsilon\left(1 - \left(\frac{\epsilon}{p_m}\right)^{1/\max\{k,d-k\}}\right)\right\}.$$

Here $p_m \leq \max_{0\leq i\leq d}\mathbb{P}[|S| = i]$ is a lower bound on the mode of $|S|$.

Because μ is SR, the sequence $s_0, s_1, \dots, s_d$ where $s_i = \mathbb{P}[|S| = i]$ is log-concave and unimodal. Therefore, either the mode is in the interval $[0, k]$ or in $[k, d]$. We assume the former and prove the lemma; the latter can be proven similarly. First, observe that because $s_k \geq s_{k+1} \geq \cdots \geq s_d$, we get $s_k \geq \epsilon/(d - k + 1)$. In the rest of the proof, we show that $s_k \geq \epsilon(1 - (\epsilon/p_m)^{1/k})$ or $s_k \geq p_m$.

Suppose s_i is the mode. It follows that there is $i \leq j \leq k-1$ such that $\frac{s_j}{s_{j+1}} \geq \left(\frac{s_i}{s_k}\right)^{1/(k-i)}$. Therefore, by Lemma 2.4,

$$\epsilon \leq s_k + \cdots + s_d \leq \frac{s_k}{1 - \left(\frac{s_k}{s_i}\right)^{1/(k-i)}}.$$

If $s_k \geq p_m$ or $s_k \geq \epsilon$, then we are done. Otherwise,

$$s_k \geq \epsilon\left(1 - (s_k/p_m)^{1/(k-i)}\right) \geq \epsilon\left(1 - (\epsilon/p_m)^{1/k}\right),$$

where we used $s_i \geq p_m$ and $s_k \leq \epsilon$. □

Lemma 5.2. *Given a strongly Rayleigh distribution $\mu : 2^{[n]} \to \mathbb{R}_{\geq 0}$, let A, B be two (nonnegative) random variables corresponding to the number of elements sampled from two disjoint sets such that $\mathbb{P}[A + B = n] > 0$ where $n = n_A + n_B$.*

Then,

$$\mathbb{P}[A \ge n_A | A+B=n] = \mathbb{P}[B \le n_B | A+B=n] \ge \mathbb{P}[A \ge n_A]\mathbb{P}[B \le n_B], \quad (5.2)$$

$$\mathbb{P}[A \le n_A | A+B=n] = \mathbb{P}[B \ge n_B | A+B=n] \ge \mathbb{P}[A \le n_A]\mathbb{P}[B \ge n_B]. \quad (5.3)$$

We prove the second statement. The first one can be proven similarly. First, notice

$$\mathbb{P}[A \le n_A, A+B \ge n] + \mathbb{P}[B \ge n_B, A+B < n]$$
$$= \mathbb{P}[B \ge n_B, A \le n_A, A+B \ge n] + \mathbb{P}[A \le n_A, B \ge n_B, A+B < n]$$
$$= \mathbb{P}[B \ge n_B, A \le n_A] \ge \mathbb{P}[B \ge n_B]\mathbb{P}[A \le n_A] =: \alpha,$$

where the last inequality follows by negative association. Say $q = \mathbb{P}[A+B \ge n]$. From the previous statement, either $\mathbb{P}[A \le n_A, A+B \ge n] \ge \alpha q$ or $\mathbb{P}[B \ge n_B, A+B < n] \ge \alpha(1-q)$. In the former case, we get $\mathbb{P}[A \le n_A | A+B \ge n] \ge \alpha$, and in the latter, we get $\mathbb{P}[B \ge n_B | A+B < n] \ge \alpha$. Now the lemma follows by the stochastic dominance property

$$\mathbb{P}[A \le n_A | A+B=n] \ge \mathbb{P}[A \le n_A | A+B \ge n]$$
$$\mathbb{P}[B \ge n_B | A+B=n] \ge \mathbb{P}[B \ge n_B | A+B < n] .$$

In the special case that $A+B<n$ never happens, the lemma holds trivially. □

Combining the previous two lemmas, we get the following.

Corollary 5.1. *Let $\mu : 2^{[n]} \to \mathbb{R}_{\ge 0}$ be a SR distribution. Let A, B be two random variables corresponding to the number of elements sampled from two disjoint sets of elements such that $A \ge k_A$ with probability 1 and $B \ge k_B$ with probability 1. If $\mathbb{P}[A \ge n_A], \mathbb{P}[B \ge n_B] \ge \epsilon_1$ and $\mathbb{P}[A \le n_A], \mathbb{P}[B \le n_B] \ge \epsilon_2$, then, letting $n'_A = n_A - k_A, n'_B = n_B - k_B$,*

$$\mathbb{P}[A = n_A | A+B = n_A+n_B] \ge \epsilon \min\left\{\frac{1}{n'_A+1}, \frac{1}{n'_B+1}\right\},$$

$$\mathbb{P}[A = n_A | A+B = n_A+n_B] \ge \min\left\{p_m, \epsilon\left(1-(\epsilon/p_m)^{1/\max\{n'_A,n'_B\}}\right)\right\},$$

where $\epsilon = \epsilon_1\epsilon_2$ and $p_m \le \max_{k_A \le k \le n_A+n_B-k_B} \mathbb{P}[A=k | A+B = n_A+n_B]$ is a lower bound on the mode of A.

In the special case that $n_A = 1, n_B = 1, k_A = 0, k_B = 0$, if $\mathbb{P}[A=1|A+B=2] \le \epsilon$, $p_m \ge 1-2\epsilon$. If $\epsilon \le 1/3$,

$$\mathbb{P}[A=1|A+B=2] \ge \max\left\{\epsilon/2, \epsilon\left(1-\frac{\epsilon}{1-2\epsilon}\right)\right\}.$$

To get the first statement, we construct a new SR distribution from μ as follows. First, we symmetrize g_μ by setting all $x_a \in A$ to x and all $x_b \in B$ to y; call the resulting polynomial q_μ. Then, notice $q'_\mu = q_\mu/(x^{k_A}x^{k_B})$ is real stable. Therefore, we can apply the above corollary to a distribution with generating polynomial q'_μ.[15]

To get the second statement, because the distribution of A is unimodal,

$$\min\{\mathbb{P}[A=0], \mathbb{P}[A=2]\} \le \epsilon.$$

5.2. Max Flow

This proposition and the max flow event are crucially used in the analysis of the bottom-bottom case in the payment theorem (Theorem 4.5). See Example 3.2 and the preceding discussion for more high-level intuition. The main consequences of this section are Corollary 5.3 and Corollary 5.4.

Proposition 5.2. *Let $\mu : 2^E \to \mathbb{R}_{\ge 0}$ be a homogeneous SR distribution. For any $303\epsilon < \zeta < 0.003$ and disjoint sets $A, B \subseteq E$ such that $\mathbb{E}[A_T], \mathbb{E}[B_T] \in [1-\epsilon, 1+\epsilon]$ (where $T \sim \mu$) there is an event $\mathcal{E}_{A,B}(T)$ such that $\mathbb{P}[\mathcal{E}_{A,B}(T)] \ge 0.002\zeta^2(1-\zeta/3-\epsilon)$ and it satisfies the following three properties.*

(i) $\mathbb{P}[A_T = B_T = 1 | \mathcal{E}_{A,B}(T)] = 1$,
(ii) $\sum_{e \in A} |\mathbb{P}[e] - \mathbb{P}[e|\mathcal{E}_{A,B}(T)]| \le \zeta$, *and*
(iii) $\sum_{e \in B} |\mathbb{P}[e] - \mathbb{P}[e|\mathcal{E}_{A,B}(T)]| \le \zeta$.

In other words, under event $\mathcal{E}_{A,B}$ which has a constant probability, $A_T = B_T = 1$ and the marginals of all edges in A, B are preserved up to total variation distance ζ. We also remark that above statement holds for a much larger value of ζ at the expense of a smaller lower bound on $\mathbb{P}[\mathcal{E}_{A,B}(T)]$.

Before, proving the above statement we prove the following lemma.

Lemma 5.3. *Let $\mu : 2^E \to \mathbb{R}_{\ge 0}$ be a homogeneous SR distribution. Let $A, B \subseteq E$ be two disjoint sets such that $\mathbb{E}[A_T]$, $\mathbb{E}[B_T] \in [1-\epsilon, 1+\epsilon]$ (where $T \sim \mu$), $A' \subset A$ and $B' \subseteq B$ and $\mathbb{E}[A'_T \cup B'_T] \ge 1+\alpha$ for some $\alpha > 100\epsilon$. If $\alpha < 0.001$, we have*

$$\mathbb{P}[A'_T = B'_T = A_T = B_T = 1] \ge 0.1\alpha^3.$$

First, condition on $(A \setminus A')_T = (B \setminus B')_T = 0$. This happens with probability at least $\alpha - 2\epsilon \ge 0.98\alpha$ because $\mathbb{E}[A_T] + \mathbb{E}[B_T] \le 2+2\epsilon$ and $\mathbb{E}[A'_T] + \mathbb{E}[B'_T] \ge 1+\alpha$. Call this measure ν. It follows by negative association that

$$\mathbb{E}_\nu[A'_T], \mathbb{E}_\nu[B'_T] \in [\alpha - \epsilon, 2+3\epsilon-\alpha]. \quad (5.4)$$

- Case 1: $\mathbb{E}_\nu[A'_T + B'_T] > 1.5$. Because $\mathbb{E}_\nu[A'_T + B'_T] \le 2+2\epsilon$, by Lemma 2.5, $\mathbb{P}_\nu[A'_T + B'_T = 2] \ge 0.25$. Furthermore, by

$$\mathbb{P}_\nu[A'_T \ge 1], \mathbb{P}_\nu[B'_T \ge 1] \ge 1 - e^{-(\alpha-\epsilon)} \ge 0.98\alpha \quad \text{(By Lemma 2.6, } \alpha < 0.001\text{)},$$

$$\mathbb{P}_\nu[A'_T \le 1], \mathbb{P}_\nu[B'_T \le 1] \ge \alpha/2 - 1.5\epsilon \quad \text{(Markov's Inequality)}.$$

Therefore, by Corollary 5.1 and using $\alpha \le 0.001$, $\mathbb{P}[A'_T = 1 | A'_T + B'_T = 2] \ge 0.45\alpha^2$. It follows that

$$\mathbb{P}[A_T = B_T = A'_T = B'_T = 1] \ge (0.98\alpha)\mathbb{P}_\nu[A'_T = B'_T = 1] \ge (0.98\alpha)0.25(0.45\alpha^2) \ge 0.1\alpha^3.$$

- Case 2: $\mathbb{E}[A'_T + B'_T] \le 1.5$. Because $\mathbb{E}_\nu[A'_T + B'_T] \ge 1 + \alpha$, by Lemma 2.5, $\mathbb{P}[A'_T + B'_T = 2] \ge \alpha e^{-\alpha} \ge 0.99\alpha$. However, now $\mathbb{E}[A'_T], \mathbb{E}[B'_T] \le 1.5$ and therefore by Markov's inequality,

$$\mathbb{P}_\nu[A'_T \le 1], \mathbb{P}_\nu[B'_T \le 1] \ge 0.25.$$

On the other hand, by Lemma 2.6 (similar to case 1) $\mathbb{P}_\nu[A'_T \ge 1], \mathbb{P}_\nu[B'_T \ge 1] \ge 1 - e^{-\alpha+\epsilon} \ge 0.98\alpha$. It follows by Corollary 5.1 that $\mathbb{P}[A'_T = 1 | A'_T + B'_T = 2] \ge 0.2\alpha$. Therefore,

$$\mathbb{P}[A_T = B_T = A'_T = B'_T = 1] \ge (0.98\alpha)$$
$$\mathbb{P}_\nu[A'_T = B'_T = 1] \ge (0.98\alpha)(0.2\alpha)(0.99\alpha) \ge 0.1\alpha^3$$

as desired.

It is worth noting that α^3 dependency is necessary in the previous example. For an explicit strongly Rayleigh distribution consider the following product distribution:

$$(\alpha x_1 + (1-\alpha)y_2)(\alpha y_1 + (1-\alpha)z_2)(\alpha z_1 + (1-\alpha)x_2),$$

and let $A = \{x_1, x_2\}$, $B' = B = \{y_1, y_2\}$, and $A' = \{x_1\}$. Observe that

$$\mathbb{P}[A_T = B_T = A'_T = B'_T = 1] = \mathbb{P}[x_1 = 1, y_1 = 1, z_1 = 1] = \alpha^3.$$

Proof of Proposition 5.2. To prove the lemma, we construct an instance of the max-flow, min cut problem. Consider the following graph with vertex set $\{s, A, B, t\}$. For any $e \in A, f \in B$ connect e to f with a directed edge of capacity $y_{e,f} = \mathbb{P}[e,f \in T | A_T = B_T = 1]$. For any $e \in E$, let $x_e := \mathbb{P}[e \in T]$. Connect s to $e \in A$ with an arc of capacity βx_e and similarly connect $f \in B$ to t with arc of capacity βx_f, where β is a parameter that we choose later. We claim that the min cut of this graph is at least $\beta(1 - \epsilon - \zeta/3)$. Assuming this, we can prove the lemma as follows: let $\mathbf{z}$ be the maximum flow, where $z_{e,f}$ is the flow on the edge from e to f. We define the event $\mathcal{E}_{A,B}(T) = \mathcal{E}(T)$ to be the union of events $z_{e,f}$. More precisely, conditioned on $A_T = B_T = 1$ the events $e,f \in T | A_T = B_T = 1$ are disjoint for different pairs $e \in A, f \in B$, so we know that we have a specific e, f in the tree T with probability $y_{e,f}$. Of course, $\sum_{e\in A, f\in B} y_{e,f} = 1$. Therefore, for $e \in A, f \in B$ we include a $z_{e,f}$ measure of trees, T, such that $A_T = B_T = 1, e,f \in T$. First, observe that

$$\mathbb{P}[\mathcal{E}] = \sum_{e\in A, f\in B} z_{e,f}\mathbb{P}[A_T = B_T = 1] \ge \beta(1 - \zeta/3 - \epsilon)\mathbb{P}[A_T = B_T = 1]. \tag{5.5}$$

Part (i) of the proposition follows from the definition of $\mathcal{E}$. Now, we check part (ii): Say $z = \sum_{e\in A, f\in B} z_{e,f}$, and the flow into e is z_e. Then,

$$\sum_{e\in A} |x_e - \mathbb{P}[e \in T | \mathcal{E}]| = \sum_{e\in A} \left| x_e - \sum_f \frac{z_{e,f}}{z} \right| = \sum_{e\in A} \left| x_e - \frac{z_e}{z} \right|.$$

Both x and z_e/z define a probability distribution on edges in A; therefore, the RHS is just the total variation distance between these two distributions. We can write

$$\begin{aligned}\sum_{e\in A} |x_e - \mathbb{P}[e \in T | \mathcal{E}]| &= 2 \sum_{e\in A: z_e/z > x_e} \left(\frac{z_e}{z} - x_e\right) \\ &\le 2 \sum_{e\in A: z_e/z > x_e} \left(\frac{\beta x_e}{\beta(1 - \zeta/3 - \epsilon)} - x_e\right) \\ &\le 2 \cdot \sum_e x_e \frac{\zeta/3 + \epsilon}{1 - \zeta/3 - \epsilon} \\ &\le 2\frac{(1+\epsilon)(\zeta/3 + \epsilon)}{1 - \zeta/3 - \epsilon} \le \zeta.\end{aligned}$$

The first inequality uses that the max-flow is at least $\beta(1 - \zeta/3)$ and that the incoming flow of e is at most βx_e, and the last inequality follows by $\zeta < 1/20$ and $\epsilon < \zeta/300$. (iii) can be checked similarly.

It remains to lower bound the max-flow or equivalently the min cut. Consider an s, t-cut $S, \overline{S}$, that is, assume $s \in S$ and $t \notin S$. Define $S_A = A \cap S$, $S_B = B \cap S$, and similarly $\overline{S}_A = A \cap \overline{S}$, $\overline{S}_B = B \cap \overline{S}$. We write

$$\begin{aligned}\mathrm{cap}(S, \overline{S}) &= \beta x(\overline{S}_A) + \beta x(S_B) + \sum_{e\in S_A, f\in \overline{S}_B} y_{e,f} \\ &= \beta x(\overline{S}_A \cup S_B) + \mathbb{P}[(S_A)_T \\ &= (\overline{S}_B)_T = 1 | A_T = B_T = 1].\end{aligned}$$

If $x(S_B) \ge x(S_A) - \zeta/3$, then

$$\begin{aligned}\mathrm{cap}(S, \overline{S}) &\ge \beta x(\overline{S}_A \cup S_B) \ge \beta(x(\overline{S}_A \cup S_A) - \zeta/3) \\ &\ge \beta(1 - \epsilon - \zeta/3),\end{aligned}$$

and we are done. Otherwise, say $x(S_B) + \gamma = x(S_A)$, for some $\gamma > \zeta/3$. Therefore,

$$x(\overline{S}_B) + x(S_A) = x(\overline{S}_B) + x(S_B) + \gamma \ge 1 - \epsilon + \gamma.$$

Therefore, by Lemma 5.3 with ($\alpha = \gamma - \epsilon > \zeta/3 - \epsilon > 100\epsilon$),

$$\begin{aligned}\mathbb{P}[(S_A)_T = (\overline{S}_B)_T = 1 | A_T = B_T = 1] &\ge \frac{\mathbb{P}[(S_A)_T = (\overline{S}_B)_T = A_T = B_T = 1]}{\mathbb{P}[A_T = B_T = 1]} \\ &\ge \frac{0.1(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1]}.\end{aligned}$$

It follows that

$$\begin{aligned}\mathrm{cap}(S,\overline{S}) &\geq \beta x(\overline{S}_A \cup S_B) + \frac{0.1(\gamma-\epsilon)^3}{\mathbb{P}[A_T = B_T = 1]}\\ &\geq \beta(x(\overline{S}_A \cup S_A) - \gamma) + \frac{0.1(\gamma-\epsilon)^3}{\mathbb{P}[A_T = B_T = 1]}\\ &\geq \beta(1-\epsilon-\gamma) + \frac{0.1(\gamma-\epsilon)^3}{\mathbb{P}[A_T = B_T = 1]}.\end{aligned}$$

To prove the lemma, we just need to choose β such that RHS is at least $\beta(1-\epsilon-\zeta/3)$. Or equivalently,

$$\frac{0.1(\gamma-\epsilon)^3}{\mathbb{P}[A_T = B_T = 1]} \geq \beta(\gamma - \zeta/3).$$

In other words, it is enough to choose $\beta \leq \frac{0.1(\gamma-\epsilon)^3}{\mathbb{P}[A_T=B_T=1](\gamma-\zeta/3)}$. Because $\gamma \geq \zeta/3$ and $\zeta/3 > 100\epsilon$, we certainly have $\gamma - \epsilon \geq \zeta/6$. Therefore, we can set $\beta = \frac{0.1\zeta^2/6^2}{\mathbb{P}[A_T=B_T=1]}$. Finally, this plus (5.5) gives

$$\begin{aligned}\mathbb{P}[\mathcal{E}] &\geq (1-\zeta/3-\epsilon)\beta\mathbb{P}[A_T = B_T = 1]\\ &= 0.1(\zeta^2/6^2)(1-\zeta/3-\epsilon) \geq 0.002\zeta^2(1-\zeta/3-\epsilon)\end{aligned}$$

as desired.

Definition 5.1 (Max-Flow Event). For a polygon cut $S \in \mathcal{H}$ with polygon partition A, B, C, let ν be the max-entropy distribution conditioned on S is a tree and $C_T = 0$. By Lemma 2.7, we can write $\nu : \nu_S \times \nu_{G/S}$, where ν_S is supported on trees in $E(S)$ and $\nu_{G/S}$ on trees in $E(G/S)$. For a sample $(T_S, T_{G/S}) \sim \nu_S \times \nu_{G/S}$, we say $\mathcal{E}_S$ occurs if $\mathcal{E}_{A,B}(T_{G/S})$ occurs, where $\mathcal{E}_{A,B}(.)$ is the event defined in Proposition 5.2 for sets A, B and $\zeta = \epsilon_M := \frac{1}{4{,}000}$ and $\epsilon = 2\epsilon_\eta$.

Corollary 5.2. *For a polygon cut $S \in \mathcal{H}$ with polygon partition A, B, C, we have*

(i) $\mathbb{P}[\mathcal{E}_S] \geq 0.001\epsilon_M^2$.

(ii) *For any set $F \subseteq \delta(S)$ conditioned on $\mathcal{E}_S$ marginals of edges in F are preserved up to $\epsilon_M + \epsilon_\eta$ in total variation distance.*

(iii) *For any $F \subseteq E(S) \cup \delta(S)$ where either $F \cap A = \emptyset$ or $F \cap B = \emptyset$, there is some $q \in x(F) \pm (\epsilon_M + 2\epsilon_\eta)$ such that the law of $F_T | \mathcal{E}_S$ is the same as a $BS(q)$.*

Condition S to be a tree and $C_T = 0$ and let ν be the resulting measure. It follows that

$$\begin{aligned}\mathbb{P}[\mathcal{E}_S] &= \mathbb{P}_\nu[\mathcal{E}_S]\mathbb{P}[C_T = 0, S \text{ tree}]\\ &\geq 0.002\epsilon_M^2(1-\epsilon_M/3-\epsilon)\mathbb{P}[C_T = 0, S \text{ tree}]\\ &\geq 0.001\epsilon_M^2.\end{aligned}$$

This proves (i).

Now, we prove (ii). By Proposition 5.2, the marginals of edges in $\delta(S)$ are preserved up to a total variation distance of ϵ_M, so

$$\mathbb{E}_\nu[(F\cap\delta(S))_T | \mathcal{E}_{A,B}(T_{G/S})] = \mathbb{E}_\nu[(F\cap\delta(S))_T] \pm \epsilon_M.$$

Because $x(C) \leq \epsilon_\eta$ and $x(\delta(S)) \leq 2 + \epsilon_\eta$, by negative association,

$$x(F\cap\delta(S)) - \epsilon_\eta/2 \leq \mathbb{E}_\nu[(F\cap\delta(S))_T] \leq x(F\cap\delta(S)) + \epsilon_\eta.$$

This proves (ii). Also observe that because conditioned on $\mathcal{E}_S$, we choose at most one edge of $F\cap\delta(S)$, $(F\cap\delta(S))_T$ is a $BS(q_{G/S})$ for some $q_{G/S} = x(F\cap\delta(S)) \pm (\epsilon_M + \epsilon_\eta)$.

On the other hand, observe that conditioned on $\mathcal{E}_S$, S is a tree, so

$$x(F\cap E(S)) \leq \mathbb{E}[(F\cap E(S))_T | \mathcal{E}_S] \leq x(F\cap E(S)) + \epsilon_\eta/2.$$

Because the distribution of $(F\cap E(S))_T$ under $\nu|\mathcal{E}_S$ is SR, there is a random variable $BS(q_S) = (F\cap E(S))_T$ where $x(F\cap E(S)) \leq q_S \leq x(F\cap E(S)) + \epsilon_\eta/2$.

Finally, $F_T|\mathcal{E}_S$ is exactly $BS(q_S) + BS(q_{G/S}) = BS(q)$ for $q = x(F) \pm (\epsilon_M + 2\epsilon_\eta)$.

Normally, conditioning on $\delta(S)_T$ for a polygon $S \in \mathcal{H}$ may dramatically change the distribution of any random variable $\delta(u)_T$ for any u that is an ancestor of S and for that $\delta(u)\cap\delta(S) \neq \emptyset$. For example, it may essentially determine the parity of $\delta(u)_T$. On the other hand, the following two corollaries show that after conditioning on $\mathcal{E}_S$ the probability $\delta(u)$ is even remains a (large) constant. Therefore, in some sense, conditioning on the max-flow event $\mathcal{E}_S$ decouples the random variables $\delta(S)_T$ and $\delta(u)_T$.

Corollary 5.3. *For $u \in \mathcal{H}$ and a polygon cut $S \in \mathcal{H}$ that is an ancestor of u,*

$$\mathbb{P}[\delta(u)_T \ odd | \mathcal{E}_S] \leq 0.5678.$$

First, notice by Observation 4.1, $\delta(u)\cap\delta(S)$ is either a subset of A, B, or C. Therefore, by (iii) of Corollary 5.2, we can write $\delta(u)_T|\mathcal{E}_S$ as a $BS(q)$ for $q \in 2 \pm [0.001]$ (where we use that $\epsilon_M + 3\epsilon_\eta < 0.001$). Furthermore, because $\delta(u)_T \neq 0$ with probability 1, we can write this as a $1 + BS(q-1)$. Therefore, by Corollary 2.1,

$$\begin{aligned}\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{E}_S] = \mathbb{P}[BS(q-1) \text{ even}] &\leq \frac{1}{2}(1+e^{-2(q-1)})\\ &\leq \frac{1}{2}(1+e^{-1.999}) \leq 0.5678\end{aligned}$$

as desired.

Corollary 5.4. *For a polygon cut $u \in \mathcal{H}$ and a polygon cut $S \in \mathcal{H}$ that is an ancestor of u,*

$$\mathbb{P}[u \textit{ not left happy} | \mathcal{E}_S] \leq 0.56797.$$

The same follows for right happy.

Let A, B, C be the polygon partition of u. Recall that for u to be left-happy, we need $C_T = 0$ and A_T odd.

Similar to the previous statement, we can write $A_T|\mathcal{E}_S$ as a $BS(q_A)$ for $q_A \in 1 \pm [0.00026]$ (where we used that $\epsilon_M = 1/4,000$ and $\epsilon_\eta \le \epsilon_M/300$). Therefore, by Corollary 2.1,

$$\mathbb{P}[A_T \text{ even}|\mathcal{E}_S] \le \frac{1}{2}(1+e^{-2q_A}) \le \frac{1}{2}(1+e^{-1.99948}) \le 0.56771.$$

Finally, $\mathbb{E}[C_T|\mathcal{E}_S] \le x(C_T) + \epsilon_M + 2\epsilon_\eta \le 0.00026$. Now using the union bound,

$$\mathbb{P}[u \text{ not left happy}|\mathcal{E}_S] \le 0.56771 + 0.00026 \le 0.56797$$

as desired.

5.3. Good Edges

Definition 5.2 (Half Edges). We say an edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ in a degree cut $S \in \mathcal{H}$, that is, $\mathsf{p}(\mathbf{e}) = S$, is a *half edge* if $|x_{\mathbf{e}} - 1/2| \le \epsilon_{1/2}$, where $\epsilon_{1/2}$ is defined in global constants.

Definition 5.3 (Good Edges). We say a top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ in a degree cut $S \in \mathcal{H}$ is (2-2) good, if one of the following holds:

1. $\mathbf{e}$ is not a half edge or
2. $\mathbf{e}$ is a half edge and $\mathbb{P}[\delta(u)_T = \delta(v)_T = 2|u, v \text{ trees}] \ge 3\epsilon_{1/2}$.

We say a top edge $\mathbf{e}$ is bad otherwise. We say every bottom edge bundle is good (but generally do not refer to bottom edges as good or bad). We say any edge e that is a neighbor of u_0 or v_0 is bad.

In the next section we will see that for any top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ that is not a half edge, $\mathbb{P}[(\delta(u))_T = (\delta(v))_T = 2|u, v \text{ trees}] = \Omega(1)$. The following theorem is the main result of this section.

Theorem 5.1. *For $\epsilon_{1/2} \le 0.0002$, $\epsilon_\eta \le \epsilon_{1/2}^2$, a top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ is bad only if the following three conditions hold simultaneously:*

- *The edge bundle $\mathbf{e}$ is a half edge,*
- *We have $x(\delta^\uparrow(u)), x(\delta^\uparrow(v)) \le 1/2 + 9\epsilon_{1/2}$,*
- *Every other half edge bundle incident to u or v is (2-2) good.*

The proof of this theorem follows from Lemma 5.5 and Lemma 5.6.

In this section, we use repeatedly that for any atom u in a degree cut S, $x(\delta(u)) \le 2 + \epsilon_\eta$. We also repeatedly use that for a half edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ in a degree cut, conditioned on u, v trees, $\mathbf{e}$ is in or out with probability at least $1/2 - \epsilon_{1/2} - 3\epsilon_\eta > 0.49$.

Lemma 5.4. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle in a degree cut $S \in \mathcal{H}$. Let $A = \delta(u)_{-\mathbf{e}}$ and $B = \delta(v)_{-\mathbf{e}}$. If $\epsilon_{1/2} \le 0.001$ and $\epsilon_\eta < \epsilon_{1/2}/100$, then*

$$\mathbb{P}[A_T + B_T \le 2|u, v \text{ trees}],$$
$$\mathbb{P}[A_T + B_T \ge 4|u, v \text{ trees}] \ge 0.4\epsilon_{1/2}$$

Throughout the proof all probabilistic statements are with respect to the measure μ conditioned on u, v trees.

Let $p_{\le 2} = \mathbb{P}[A_T + B_T \le 2]$ and similarly define $p_{\ge 4}$. Observe that whenever $\delta(u)_T = \delta(v)_T = 2$, we must have $A_T + B_T \ne 3$. Because $\mathbf{e}$ is 2-2 good, this event happens with probability at least $3\epsilon_{1/2}$, that is,

$$p_{\le 2} + p_{\ge 4} \ge 3\epsilon_{1/2}. \tag{5.6}$$

By Lemma 2.5, using the fact that $p_0 = 0$, we get $p_{=3} \ge 1/4$.

First, we show that $p_{\le 2} \ge 0.4\epsilon_{1/2}$. We have

$$\begin{aligned} 3 + 2\epsilon_{1/2} \ge \mathbb{E}[A_T + B_T] &\ge 4p_{\ge 4} + 2p_{=2} + 3(1 - p_{\ge 4} - p_{\le 2}) \\ &= 3 + p_{\ge 4} - p_{=2} - 3p_{=1}. \end{aligned}$$

Again, we are using $p_0 = 0$. By log-concavity $p_{=2}^2 \ge p_{=3}\, p_{=1}$, so because $p_{=3} \ge 1/4$, $p_{=1} \le 4p_{=2}^2 \le 4p_{\le 2}^2$. Therefore,

$$p_{\ge 4} - 2\epsilon_{1/2} \le p_{=2} + 3p_{=1} = p_{\le 2} + 2p_{=1} \le p_{\le 2}(1 + 8p_{\le 2}).$$

Finally, because $\epsilon_{1/2} < 0.001$, plugging this upper bound on $p_{\ge 4}$ into Equation (5.6) we get $p_{\le 2} \ge 0.4\epsilon_{1/2}$.

Now, we show $p_{\ge 4} \ge 0.4\epsilon_{1/2}/2$. Assume $p_{\ge 4} < \epsilon_{1/2}/2$ (otherwise we are done). Because $p_{=3} \ge 1/4$ by Lemma 2.4 with $\gamma \le (\epsilon_{1/2}/2)/(1/4) = 2\epsilon_{1/2}$:

$$\mathbb{E}[A_T + B_T|A_T + B_T \ge 4] \cdot p_{\ge 4} \le \frac{p_{\ge 4}}{1 - 2\epsilon_{1/2}}(4 + 3\epsilon_{1/2}).$$

Therefore,

$$\begin{aligned} 3 - 2\epsilon_{1/2} - 2\epsilon_\eta \le \mathbb{E}[A_T + B_T] &\le 2p_{\le 2} \\ &+ \frac{p_{\ge 4}}{1 - 2\epsilon_{1/2}}(4 + 3\epsilon_{1/2}) + 3(1 - p_{\le 2} - p_{\ge 4}). \end{aligned}$$

Therefore, $1.01p_{\ge 4} \ge p_{\le 2} - 2.02\epsilon_{1/2}$ where we used $\epsilon_{1/2} \le 0.001$ and $\epsilon_\eta < \epsilon_{1/2}/100$. Now, $p_{\ge 4} \ge 0.4\epsilon_{1/2}$ follows by Equation (5.6). □

Lemma 5.5. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a half edge bundle in a degree cut $S \in \mathcal{H}$, and suppose $x(\delta^\uparrow(u)) \ge 1/2 + k\epsilon_{1/2}$. If $k \ge 9$, $\epsilon_{1/2} \le 0.0002$, and $\epsilon_\eta \le \epsilon_{1/2}^2$, then, $\mathbf{e}$ is 2-2 good.*

First, condition u, v, S to be trees. Let $W = S \smallsetminus \{u\}$ as in Figure 11. Because S is a near min cut,

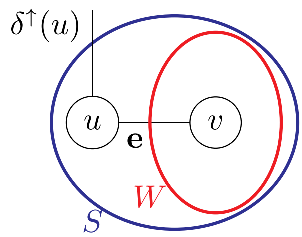
$$\begin{aligned} x(\delta(W)) &= x(\delta(S)) + x(\delta(u)) - 2x(\delta^\uparrow(u)) \\ &\le 2(2 + \epsilon_\eta) - 2(1/2 + k\epsilon_{1/2}) = 3 - 2k\epsilon_{1/2} + 2\epsilon_\eta. \end{aligned}$$

Therefore, by Lemma 2.7, $\mathbb{P}[W \text{ is tree}] \ge 1/2 + k\epsilon_{1/2} - \epsilon_\eta - \epsilon_\eta$. The extra $-\epsilon_\eta$ comes from the fact that conditioning u be a tree can decrease marginals of edges in $E(W)$ by at most ϵ_η.

Let ν be the resulting measure, namely the measure obtained by first conditioning u, v, S to be trees and then W to be a tree. Note that ν is a strongly Rayleigh distribution on the set of edges in $E(W) \cup E(u, W) \cup E(G/S)$; This is because ν is a product of three SR distributions each supported on one of the aforementioned sets.

Let $X = \delta^\uparrow(u)_T$ and $Y = \delta(v)_T - 1$. Observe that, under ν, $X = Y = 1$ iff $\delta(u)_T = \delta(v)_T = 2$. Furthermore, $Y \ge 0$ with probability 1 because v is connected to the rest of the

Figure 11. Setting of Lemma 5.5



graph. Therefore, we just need to lower bound $\mathbb{P}_\nu[X = Y = 1]$. First, notice

$$\mathbb{E}_\nu[X] \in [0.5 + k\epsilon_{1/2} - \epsilon_\eta, 1 + \epsilon_\eta]$$
$$\mathbb{E}_\nu[Y] \in [0.5 + k\epsilon_{1/2} - 4\epsilon_\eta, 1.5 - k\epsilon_{1/2} + 3\epsilon_\eta] \tag{5.7}$$

We will give a brief explanation of this: first, $\frac{1}{2} + k\epsilon_{1/2} \le \mathbb{E}[X] \le 1 + \epsilon_\eta$ before conditioning. By conditioning u, v, S to be trees, we can increase $\mathbb{E}[|E(S)|_T]$ by at most ϵ_η; therefore, this may decrease $\mathbb{E}[X]$ by at most ϵ_η. Under this measure, $E(S)$ is independent of X; therefore, conditioning on W to be a tree cannot change $\mathbb{E}[X]$. Second, $1 \le \mathbb{E}[Y] \le 1 + \epsilon_\eta$ before conditioning. Now, conditioning on u, v, S to be trees may decrease $\mathbb{E}[Y]$ by at most $2\epsilon_\eta$ and increase by at most ϵ_η. Conditioning on W to be a tree may increase or decrease $\mathbb{E}[Y]$ by a most $1/2 - k\epsilon_{1/2} + 2\epsilon_\eta$.

Using Proposition 5.1, we can immediately argue that $\mathbb{P}_\nu[X = Y = 1] \ge \Omega(\epsilon_{1/2})$. We do the following more refined analysis to make sure that this probability is at least $6\epsilon_{1/2}$ (for $\epsilon_{1/2} \le 0.0005$) and $k \ge 9$. Once we prove this, we obtain the lemma:

$$\begin{aligned}\mathbb{P}[\delta(u)_T = \delta(v)_T = 2 | u, v \text{ trees}] &\ge \mathbb{P}[S, W \text{ trees} | u, v \text{ trees}]\mathbb{P}_\nu[X = Y = 1] \\ &\ge 0.5 \cdot 6\epsilon_{1/2}.\end{aligned}$$

Case 1: $\mathbb{P}_\nu[X + Y = 2] \ge 48\epsilon_{1/2}$. By Lemma 2.6, $\mathbb{P}_\nu[X \ge 1], \mathbb{P}_\nu[Y \ge 1] \ge 1 - e^{-0.5}$. On the other hand, by Theorem 2.2, $\mathbb{P}_\nu[X \le 1], \mathbb{P}_\nu[Y \le 1] \ge 7/16$. This is because if we have one Bernoulli of value 1, $\mathbb{P}_\nu[X \le 1] \ge \left(1 - \frac{0.5}{n}\right)^n$ is minimized at $n = 1$, whereas if we have no Bernoullis of value 1, $\mathbb{P}_\nu[X \le 1] \ge \left(1 - \frac{1.5}{n}\right)^n + 1.5\left(1 - \frac{1.5}{n}\right)^{n-1}$ which is minimized at $n = 2$. Therefore, by Corollary 5.1, $\mathbb{P}_\nu[X = 1 | X + Y = 2] \ge 0.1269$. Therefore, we get

$$\mathbb{P}_\nu[X = 1, Y = 1] \ge 48\epsilon_{1/2} \cdot 0.1269 \ge 6\epsilon_{1/2}.$$

Case 2: $\mathbb{P}_\nu[X + Y = 2] < 48\epsilon_{1/2} < 0.01$. By Lemma 2.5, $\mathbb{P}_\nu[X + Y = 1] \ge 0.25$ (if $\mathbb{E}_\nu[X + Y] \ge 1.2$ then the assumption of this case obviously fails). So, since $\mathbb{P}_\nu[X + Y = 2] < 0.01$, by log concavity, $\mathbb{P}_\nu[X + Y = 3] \le 0.01/25$. Furthermore, by Lemma 2.4 (with $\gamma = 1/25, i = 1, k = 3$), $\mathbb{P}_\nu[X + Y > 2] < 0.0005$.

Now, assume that $\mathbb{P}_\nu[X \ge 1], \mathbb{P}_\nu[Y \ge 1] \ge 0.47$ (we will prove this shortly). Now, applying stochastic dominance, we have

$$\begin{aligned}\mathbb{P}_\nu[X \ge 1 | X + Y = 2] &\ge \mathbb{P}_\nu[X \ge 1 | X + Y \le 2] \\ &\ge \mathbb{P}_\nu[X \ge 1, X + Y \le 2] \\ &\ge \mathbb{P}_\nu[X \ge 1] - \mathbb{P}_\nu[X + Y > 2] \\ &\ge \mathbb{P}_\nu[X \ge 1] - 0.0005 \ge 0.469.\end{aligned}$$

Similarly, $\mathbb{P}_\nu[X \le 1 | X + Y = 2] = \mathbb{P}_\nu[Y \ge 1 | X + Y = 2] \ge \mathbb{P}_\nu[Y \ge 1] - 0.0005 \ge 0.469$. Finally because the distribution of X conditioned on $X + Y = 2$ is the same as the number of successes in two independent Bernoulli trials, with probabilities, say, p_1 and p_2, we can minimize $p_1(1 - p_2) + (1 - p_1)p_2$ subject to $1 - p_1 p_2 \ge 0.469$ and $1 - (1 - p_1)(1 - p_2) \ge 0.469$. Solving this yields $\mathbb{P}_\nu[X = 1 | X + Y = 2] \ge 0.395$.

Lastly, observe that because by Equation (5.7) $1.2 \ge \mathbb{E}_\nu[X + Y] \ge 1 + (2k - 1)\epsilon_{1/2}$, by Lemma 2.5, we can write

$$\mathbb{P}_\nu[X + Y = 2] \ge (2k - 1)\epsilon_{1/2} e^{-(2k-1)\epsilon_{1/2}} \ge (2k - 2)\epsilon_{1/2}.$$

Therefore,

$$\begin{aligned}\mathbb{P}_\nu[X = Y = 1] &= \mathbb{P}_\nu[X = 1 | X + Y = 2]\mathbb{P}_\nu[X + Y = 2] \\ &\ge 0.395(2k - 2)\epsilon_{1/2}.\end{aligned}$$

To get the RHS to be at least $6\epsilon_{1/2}$ it suffices that $k \ge 9$.

Now we prove that $\mathbb{P}_\nu[X \ge 1] \ge 0.47$; $\mathbb{P}_\nu[Y \ge 1] \ge 0.47$ follows similarly.

$$\mathbb{P}_\nu[X = 2] \le \mathbb{P}_\nu[X + Y \ge 2] \le 0.01 + 0.00042 \le 0.0105$$

Also notice that $\mathbb{P}_\nu[X = 1] \ge 0.3$ by Lemma 2.5. Now, using Lemma 2.4 we can write, for $\gamma = 1/25$ and $i = 1$,

$$\mathbb{E}_\nu[X | X \ge 2]\mathbb{P}_\nu[X \ge 2] \le 0.0224.$$

Therefore, because X is integer valued,

$$\begin{aligned}\mathbb{P}_\nu[X \ge 1] &\ge \mathbb{E}_\nu[X] - \mathbb{E}_\nu[X | X \ge 2]\mathbb{P}_\nu[X \ge 2] \\ &\ge \mathbb{E}_\nu[X] - 0.0224 \ge 0.47,\end{aligned}$$

as desired.

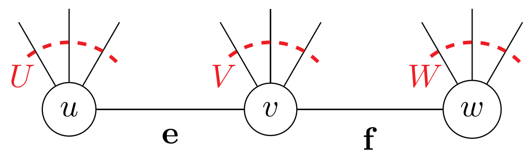
Lemma 5.6. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be two half edge bundles in a degree cut $S \in \mathcal{H}$. If $\epsilon_{1/2} < 0.0005$ and $\epsilon_\eta \le \epsilon_{1/2}^2$, then one of $\mathbf{e}$ or $\mathbf{f}$ is good.*

We use the following notation $V = \delta(v)_{-\mathbf{e}-\mathbf{f}}, U = \delta(u)_{-\mathbf{e}}, W = \delta(w)_{-\mathbf{f}}$ (see Figure 12 for an illustration). For a set A of edges and an edge bundle $\mathbf{e}$ we write $A_{+\mathbf{e}} = A \cup \{\mathbf{e}\}$. Furthermore, for a measure ν we write $\nu_{-\mathbf{e}}$ to denote ν conditioned on $\mathbf{e} \notin T$.

Condition u, v, w to be trees. This occurs with probability at least $1 - 3\epsilon_\eta$. Let ν be this measure. By Lemma 2.9, without loss of generality, we can assume

$$\mathbb{E}_\nu[W_T | \mathbf{e} \notin T] \le \mathbb{E}_\nu[W_T] + 0.405. \tag{5.8}$$

Figure 12. Setting of Lemma 5.6



Now, if $\mathbb{E}_\nu[V_T|\mathbf{e}\notin T]\geq \mathbb{E}_\nu[V_T]+0.03$, then we will show $\mathbf{e}$ is 2-2 good. First,

$$\mathbb{E}_{\nu_{-\mathbf{e}}}[(V_{+\mathbf{f}})_T]\in[1.53-\epsilon_{1/2}-3\epsilon_\eta, 2+\epsilon_\eta],$$
$$\mathbb{E}_{\nu_{-\mathbf{e}}}[U_T]\in[1.5-\epsilon_{1/2}-3\epsilon_\eta, 2+\epsilon_\eta],$$
$$\mathbb{E}_{\nu_{-\mathbf{e}}}[(V_{+\mathbf{f}})_T+U_T]\in[3.03-2\epsilon_{1/2}-3\epsilon_\eta, 3.5+2\epsilon_{1/2}+2\epsilon_\eta],$$

where we may decrease the marginals by $3\epsilon_\eta$ due to conditioning u, v, w to be trees.

Therefore, by Lemma 2.5, $\mathbb{P}_{\nu_{-\mathbf{e}}}[(V_{+\mathbf{f}})_T+U_T=4]\geq 0.029$, where we use the fact that $U_T\geq 1$ and $(V_{+\mathbf{f}})_T\geq 1$ with probability 1 under $\nu_{-\mathbf{e}}$ and apply this and the remaining calculations to $U_T-1, (V_{+\mathbf{f}})_T-1$. In addition, we have

$$\mathbb{P}_{\nu_{-\mathbf{e}}}[U_T\leq 2], \mathbb{P}_{\nu_{-\mathbf{e}}}[(V_{+\mathbf{f}})_T\leq 2]\geq 0.499 \quad \text{(by Markov's Inequality)},$$
$$\mathbb{P}_{\nu_{-\mathbf{e}}}[U_T\geq 2], \mathbb{P}_{\nu_{-\mathbf{e}}}[(V_{+\mathbf{f}})_T\geq 2]\geq 0.39 \quad \text{(By Lemma 2.6)}.$$

It follows by Corollary 5.1 applied to U_T-1 and $(V_{+\mathbf{f}})_T-1$ (with $\epsilon=0.194$ and $p_m=0.6$) that

$$\mathbb{P}_{\nu_{-\mathbf{e}}}[U_T=2|U_T+(V_{+\mathbf{f}})_T=4]\geq 0.13,$$

where we use that $U_T\geq 1, (V_{+\mathbf{f}})_T\geq 1$ with probability 1 under $\nu_{-\mathbf{e}}$ because otherwise the tree would be disconnected.

Therefore,

$$\begin{aligned}&\mathbb{P}[\delta(u)_T=\delta(v)_T=2|u,v\ \text{trees}]\\ &\geq \mathbb{P}[w\ \text{is a tree}, \mathbf{e}\notin T]\mathbb{P}_{\nu_{-\mathbf{e}}}[U_T=(V_{+\mathbf{f}})_T=2]\\ &\geq (0.49)(0.029)(0.13)\geq 0.0018.\end{aligned}$$

The lemma follows (i.e., e is 2-2 good) because $0.0018\geq 3\epsilon_{1/2}$ for $\epsilon_{1/2}\leq 0.0005$.

Otherwise, if $\mathbb{E}_\nu[V_T|\mathbf{e}\notin T]\leq \mathbb{E}_\nu[V_T]+0.03$, then we will show that $\mathbf{f}$ is 2-2 good. We have

$$\mathbb{E}_{\nu_{+\mathbf{f}}}[(V_{+\mathbf{e}})_T], \mathbb{E}_{\nu_{+\mathbf{f}}}[W_T]\in[1-2\epsilon_{1/2}-3\epsilon_\eta, 1.5+2\epsilon_{1/2}+\epsilon_\eta],$$

$$\mathbb{P}_{\nu_{+\mathbf{f}}}[(V_{+\mathbf{e}})_T\leq 1], \mathbb{P}_{\nu_{+\mathbf{f}}}[W_T\leq 1]\geq 0.249 \quad \text{(By Markov's Inequality)},$$
$$\mathbb{P}_{\nu_{+\mathbf{f}}}[(V_{+\mathbf{e}})_T\geq 1], \mathbb{P}_{\nu_{+\mathbf{f}}}[W_T\geq 1]\geq 0.63 \quad \text{(By Lemma 2.6)}.$$

Therefore, by Corollary 5.1 (with $\epsilon=0.15, p_m=0.7$), we get $\mathbb{P}_{\nu_{+\mathbf{f}}}[W_T=1|(V_{+\mathbf{e}})_T+W_T=2]\geq 0.11$. On the other hand,

$$\mathbb{P}_{\nu_{+\mathbf{f}}}[(V_{+\mathbf{e}})_T+W_T=2]\geq \mathbb{P}_{\nu_{+\mathbf{f}}}[\mathbf{e}\notin T]$$
$$\mathbb{P}_{\nu_{+\mathbf{f}-\mathbf{e}}}[(V_{+\mathbf{e}})_T+W_T=2]\geq (0.49)(0.0582)\geq 0.0285.$$

To derive the last inequality, we show $\mathbb{P}_{\nu_{+\mathbf{f}-\mathbf{e}}}[(V_{+\mathbf{e}})_T+W_T=2]\geq 0.0582$. This is because by negative association and Equation (5.8):

$$\begin{aligned}\mathbb{E}_{\nu_{+\mathbf{f}-\mathbf{e}}}[(V_{+\mathbf{e}})_T+W_T]&=\mathbb{E}_{\nu_{+\mathbf{f}-\mathbf{e}}}[V_T+W_T]\leq \mathbb{E}_{\nu_{-\mathbf{e}}}[V_T+W_T]\\ &\leq \mathbb{E}_\nu[W_T]+0.405+\mathbb{E}_\nu[V_T]+0.03\\ &\leq 2.94;\end{aligned}$$

therefore, because $(V_{+\mathbf{e}})_T+W_T$ is always at least on, by Theorem 2.2, in the worst case, $\mathbb{P}_{\nu_{-\mathbf{e}+\mathbf{f}}}[(V_{+\mathbf{e}})_T+W_T=2]$ is the probability that the sum of two Bernoullis with success probability $1.94/2$ is 1, which is 0.0582.

Therefore, similar to the previous case,

$$\begin{aligned}&\mathbb{P}[\delta(v)_T=\delta(w)_T=2|v,w\ \text{trees}]\\ &\geq \mathbb{P}[u\ \text{is a tree}, f\in T]\mathbb{P}_{\nu_{+\mathbf{f}}}[(V_{+\mathbf{e}})_T+W_T=2]\\ &\quad\cdot \mathbb{P}_{\nu_{+\mathbf{f}}}[W_T=1|(V_{+\mathbf{e}})_T+W_T=2]\\ &\geq (0.49)(0.0285)(0.11)\geq 3\epsilon_{1/2}\end{aligned}$$

for $\epsilon_{1/2}\leq 0.0005$ as desired.

5.4. 2-1-1 and 2-2-2 Good Edges

Consider a cut $u\in\mathcal{H}$, and recall that $x(\delta(u))\approx 2$. Normally, it is sufficient to have $\delta(u)_T=2$ when an edge $e\in\delta(u)$ is reduced. In the worst case, the edges of $\delta(u)$ essentially come from two of its descendants u', v', that is, $x(\delta(u')\cap\delta(u))\approx 1$ and $x(\delta(v')\cap\delta(u))\approx 1$. Let $A=\delta(u')\cap\delta(u), B=\delta(v')\cap\delta(u), C=\delta(u)\smallsetminus(A\cup B)$. In such a case, if we condition on reducing an edge in A, we may have A_T to be even with probability close to one, and it will be very expensive to fix the constraint coming from $\delta(u')$, as $(\delta(u')\smallsetminus(\delta(u)))_T$ is one, that is, odd, with probability close to one. Therefore, it is crucial to make sure that when we reduce an edge in A (B), we have A_T (B_T) is odd with some probability. Because when $\delta(u)$ is even and A_T is odd, B_T will be odd as well (discounting the leftovers C, which have negligible expectation), a natural criterion is to ask for $A_T=B_T=1$, hence motivating the upcoming definition of 2-1-1 happy. To get a more high level understanding of how we use these events, see the following two sections of the overview: dealing with x_u close to one and dealing with triangles.

Definition 5.4 (*A*, *B*, *C* Degree Partitioning). For $u\in\mathcal{H}$ and $\epsilon_{1/1}$ defined in global constants, we define a partitioning of edges in $\delta(u)$: Let $a,b\subsetneq u$ be *minimal* cuts in the hierarchy, that is, $a,b\in\mathcal{H}$, such that $a\neq b$ and $x(\delta(a)\cap\delta(u)), x(\delta(b)\cap\delta(u))\geq 1-\epsilon_{1/1}$. Because the hierarchy is laminar, a, b cannot cross. Let $A = \delta(a) \cap \delta(u), B = \delta(b) \cap \delta(u), C = \delta(u) \setminus A \setminus B$.

If there is no cut $a \subsetneq u$ (in the hierarchy) such that $x(\delta(a) \cap \delta(u)) \geq 1 - \epsilon_{1/1}$, we just let A, B be two arbitrary disjoint sets of edges in $\delta(u)$ for which $x(A), x(B) \geq 1 - \epsilon_{1/1}$. As previously shown, set $C = \delta(u) \setminus A \setminus B$. This exists WLOG because we may split any edge into an arbitrary number of parallel copies.

If there is just one minimal cut $a \subsetneq u$ (in the hierarchy) with $x(\delta(a) \cap \delta(u)) \geq 1 - \epsilon_{1/1}$, that is, b does not exist in the previous definition, then we define $A = \delta(a) \cap \delta(u)$. Let $a' \in \mathcal{H}$ be the unique child of u such that $a \subseteq a'$, that is, a is equal to a' or a descendant of a'. Then we define $C = \delta(a') \cap \delta(u) \setminus \delta(a)$ and $B = (\delta(u) \setminus A) \setminus C$. In this case, because $x(\delta^{\uparrow}(a')) \leq 1 + \epsilon_\eta$, we have $x(B) \geq 1 - \epsilon_\eta \geq 1 - \epsilon_{1/1}$.

See Figure 6 for an example. The following inequalities on A, B, C degree partitioning will be used in this section:

$$\begin{aligned} x(A), x(B) &\in [1 - \epsilon_{1,1}, 1 + \epsilon_\eta], \\ x(C) &\leq 2\epsilon_{1/1} + \epsilon_\eta. \end{aligned} \tag{5.9}$$

In this section, we will define a constant $p > 0$ that is the minimum probability that a good edge bundle is happy.

Definition 5.5 (2-1-1 Happy/Good). Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle. Let $A, B, C \subseteq \delta(u)$ be a degree partitioning of edges $\delta(u)$ as defined in Definition 5.4. We say that $\mathbf{e}$ is 2-1-1 happy with respect to u if the event

$$A_T = 1, B_T = 1, C_T = 0, \delta(v)_T = 2, \text{and} \quad u \text{ and } v \text{ are both trees}$$

occurs.

We say $\mathbf{e}$ is *2-1-1 good with respect to u* if

$$\mathbb{P}[\mathbf{e} \text{ is 2-1-1 happy wrt } u] \geq p.$$

Remark 5.1. We also use this A, B, C partitioning to help deal with the triangle cut case. In the special case that u is a polygon cut with A, B, C-polygon partitioning, let A', B', C' be the degree partitioning of $\delta(u)$. Then, by Definition 4.11, we have $A' \subseteq A$, $B' \subseteq B$, $C \subseteq C'$. Therefore, if an edge in $\delta(u)$ is reduced and is 2-1-1 happy with respect to u, the polygon u is also happy. See the overview for an example.

Many of the lemmas in this section are proved in Appendix A. In the following, we assume that $\epsilon_\eta \leq \epsilon_{1/2}^2$ and $12\epsilon_{1/1} \leq \epsilon_{1/2}$.

Lemma 5.7. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \leq 1/2 - \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$ then, $\mathbf{e}$ is 2-1-1 happy with probability at least $0.005\epsilon_{1/2}^2$.*

Lemma 5.8. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \geq 1/2 + \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$, then, $\mathbf{e}$ is 2-1-1 happy with respect to u with probability at least $0.006\epsilon_{1/2}^2$.*

Fix u in the hierarchy with degree partitioning A, B, C. The previous two lemmas show that any edge bundle $\mathbf{e} \in \delta(u)$ that is not a half edge bundle is 2-1-1 good, so the difficult case is when the majority of $x(\delta^{\rightarrow}(u))$ comes from half edge bundles. In Theorem 5.1, we showed that $\delta(u)$ can have at most one 2-2 bad edge. Oddly enough, one of the simplest cases of the reduction argument is when there *is* a bad edge in $\delta(u)$. This is because we never reduce bad edges, and therefore, we never need to increase edges which are matched to them.[16] Therefore, the main problem is good edges that are not 2-1-1 good. The following key statement, Lemma 5.11, shows that these problematic edges are rare in the sense that there is at most one good half edge bundle in A (respectively, B) that is not 2-1-1 good.

To prove this, we need the following two lemmas. In the first one, we show that if $\mathbf{e}, \mathbf{f}$ are two half edge bundles that almost entirely land in A (or B), at least one of them is 2-1-1 good. In the second, we show that if a good half edge bundle does not entirely land in A (or B), then it is 2-1-1 good. This is the main tool we use to upper bound the expected increase of good top edges in Section 7.

For a set of edges D, and an edge bundle $\mathbf{e}$, let $\mathbf{e}(D) := \mathbf{e} \cap D$. Note that $\mathbf{e}(D)$ is not really an edge bundle.

Lemma 5.9. *Let $\mathbf{e} = (\mathbf{v}, \mathbf{u})$ and $\mathbf{f} = (\mathbf{v}, \mathbf{w})$ be good half top edge bundles and let A, B, C be the degree partitioning of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \leq \epsilon_{1/2}$. Then, one of $\mathbf{e}, \mathbf{f}$ is 2-1-1 happy with probability at least $0.005\epsilon_{1/2}^2$.*

Lemma 5.10. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle and let A, B, C be the degree partitioning of $\delta(u)$ (Figure A.3). If $\epsilon_{1/2} \leq 0.001$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \geq \epsilon_{1/2}$, then*

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 } \textit{happy w.r.t } u] \geq 0.02\epsilon_{1/2}^2.$$

Lemma 5.11. *For a degree cut $S \in \mathcal{H}$, and $u \in \mathcal{A}(S)$, let A, B, C be the degree partition of u. Then, $A \cap \delta^{\rightarrow}(u) =: A^{\rightarrow}$ has fraction at most $1/2 + 4\epsilon_{1/2}$ of good edges that are not 2-1-1 good (w.r.t., u).*

Suppose by way of contradiction that there is a set $D \subseteq A^{\rightarrow}$ of good edges that are not 2-1-1 good w.r.t. u with $x(D) \geq \frac{1}{2} + 4\epsilon_{1/2}$. By Lemma 5.7 and Lemma 5.8, every edge in D is part of a half edge bundle.

There are at least two half edge bundles $\mathbf{e}, \mathbf{f}$ such that $x(D \cap \mathbf{e}), x(D \cap \mathbf{f}) \geq \epsilon_{1/2}$, as there are at most four half edge bundles in $\delta^{\rightarrow}(u)$ (and using that for any half edge bundle $\mathbf{e}$, $x_{\mathbf{e}} \leq \frac{1}{2} + \epsilon_{1/2}$). Because $D \subseteq A^{\rightarrow}$, we have

$$x(A \cap \mathbf{e}), x(A \cap \mathbf{f}) \geq \epsilon_{1/2}.$$

Because $x(A \cap \mathbf{e}) \geq \epsilon_{1/2}$, if $x(B \cap \mathbf{e}) \geq \epsilon_{1/2}$ then, by Lemma 5.10, $\mathbf{e}$ is 2-1-1 good. However, because every edge in D is not 2-1-1 good w.r.t u, we must have $x(B \cap \mathbf{e}) < \epsilon_{1/2}$. The same also holds for $\mathbf{f}$. Finally, because $x(B \cap \mathbf{e}) < \epsilon_{1/2}$ and $x(B \cap \mathbf{f}) < \epsilon_{1/2}$ by Lemma 5.9, at least one of $\mathbf{e}, \mathbf{f}$ is 2-1-1 good w.r.t u. This is a contradiction. □

5.4.1. 2-2-2 Good Edges. Although Lemma 5.11 is sufficient for bounding the increase of top edges, it is not sufficient for bottom edges. Fix a polygon u with partition A, B, C and suppose $\mathsf{p}(u) = S$ is a degree cut (recall that by Remark 5.1, the degree partitioning and polygon partitioning of u are essentially the same). Roughly speaking, a bottom edge $g \in E(u)$ is "matched" to all edges in $\delta(u)$, and needs to increase for edges $f \in A$ when f is reduced and A_T is even, and for edges $f \in B$ when f is reduced and B_T is even. Therefore, g is matched to essentially twice its fraction. If most of the edges in $\delta(u)$ are 2-1-1 good, this is sufficient to bound the expected increase of g because when such an edge is reduced and 2-1-1 happy with respect to u, g does not need to increase.

It turns out that the previous lemmas are sufficient to bound the expected increase of $g \in E(u)$ except when $A \cap \delta(S) \approx B \cap \delta(S) \approx 1/2$ and $\mathbf{e} \approx A \cap \delta^{\rightarrow}(u)$ and $\mathbf{f} \approx B \cap \delta^{\rightarrow}(u)$ are both good edge bundles that are not 2-1-1 good. In this extreme case, we use a new strategy. In Lemma 5.12, we prove that the two edge bundles $\mathbf{e}, \mathbf{f}$ are 2-2 happy *simultaneously* with a constant probability. We call such a pair 2-2-2 good. Later, in Section 7, we use this to ensure that $\mathbf{e}$ and $\mathbf{f}$ are always reduced simultaneously. The point is that because $\mathbf{e}, \mathbf{f}$ do not both come from A (or B), no cut inside u contains $\mathbf{e}$ and $\mathbf{f}$. Therefore, g only needs to increase by the maximum of the decrease of $\mathbf{e}, \mathbf{f}$ (not the sum), effectively saving a factor of 2.

Definition 5.6 (2-2-2 Happy/Good). Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be top half-edge bundles (with $\mathsf{p}(\mathbf{e}) = \mathsf{p}(\mathbf{f})$). We say $\mathbf{e}, \mathbf{f}$ are 2-2-2 happy (with respect to v) if $\delta(u)_T = \delta(v)_T = \delta(w)_T = 2$ and u, v, w are all trees.

We say $\mathbf{e}, \mathbf{f}$ are 2-2-2 good with respect to v if $\mathbb{P}[\mathbf{e}, \mathbf{f}\, \text{2-2-2 happy}] \ge p$.

Lemma 5.12. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be two good top half edge bundles and let A, B, C be degree partitioning of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \le \epsilon_{1/2}$. If $\mathbf{e}, \mathbf{f}$ are not 2-1-1 good with respect to v, and $\epsilon_{1/2} \le 0.0002$, then $\mathbf{e}, \mathbf{f}$ are 2-2-2 happy with probability at least* 0.01.

The following theorem summarizes the previous results in a compact form. This is the main result used in the analysis of the increase for bottom edges in Section 7.

Theorem 5.2. *Let $v, S \in \mathcal{H}$ where $\mathsf{p}(v) = S$, and let A, B, C be the degree partitioning of $\delta(v)$. For $p \ge 0.005\epsilon_{1/2}^2$, with $\epsilon_{1/2} \le 0.0002$, $\epsilon_{1/1} \le \epsilon_{1/2}/12$ and $\epsilon_\eta \le \epsilon_{1/2}^2$, at least one of the following is true*:

(i) *The set $\delta^{\rightarrow}(v)$ has at least $1/2 - \epsilon_{1/2}$ fraction of bad edges,*

(ii) *The set $\delta^{\rightarrow}(v)$ has at least $1/2 - \epsilon_{1/2} - \epsilon_\eta$ fraction of 2-1-1 good edges with respect to v.*

(iii) *There are two (top) half edge bundles $\mathbf{e}, \mathbf{f} \in \delta^{\rightarrow}(v)$ such that $x_{\mathbf{e}(B)} \le \epsilon_{1/2}$, $x_{\mathbf{f}(A)} \le \epsilon_{1/2}$, and $\mathbf{e}, \mathbf{f}$ are 2-2-2 good (with respect to v).*

Suppose case (i) does not happen. Because every bad edge has fraction at least $1/2 - \epsilon_{1/2}$ this means that $\delta(v)$ has no bad edges. First, notice by Lemma 5.7 and Lemma 5.8, any non-half-edge in $\delta^{\rightarrow}(v)$ is 2-1-1 good (with respect to v). (Recall we define $\delta^{\rightarrow}(v) = \delta(v) \smallsetminus \delta(\mathsf{p}(v))$, where $\mathsf{p}(v)$ is the immediate parent of v in the hierarchy.) If there is only one half edge in $\delta^{\rightarrow}(v)$, then we have at least fraction $1 - \epsilon_\eta - (1/2 + \epsilon_{1/2})$ fraction of 2-1-1 good edges, and we are done with case (ii). Otherwise, there are two good half edges $\mathbf{e}, \mathbf{f} \in \delta^{\rightarrow}(v)$.

First, by Lemma 5.10, if $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \ge \epsilon_{1/2}$, then $\mathbf{e}$ is 2-1-1 good (w.r.t., v) and we are done. Similarly, if $x_{\mathbf{f}(A)}, x_{\mathbf{f}(B)} \ge \epsilon_{1/2}$, then $\mathbf{f}$ is good. Therefore, assume none of these happens.

Furthermore, by Lemma 5.9, if $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \le \epsilon_{1/2}$ (or $x_{\mathbf{e}(A)}, x_{\mathbf{f}(A)} \le \epsilon_{1/2}$) then one of $\mathbf{e}, \mathbf{f}$ is 2-1-1 good.

Therefore, the only remaining case is when $\mathbf{e}, \mathbf{f}$ are not 2-1-1 good and $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \le \epsilon_{1/2}$. However, in this case, by Lemma 5.12, $\mathbf{e}, \mathbf{f}$ are 2-2-2 good; so (iii) holds.

6. Matching

The main result of this section is to construct a matching that we use to decide which edges will have positive slack to compensate for the negative slack of edges going higher. Refer to Example 3.1 for a high-level motivation to construct a matching.

Definition 6.1 (ϵ_F Fractional Edge). For $z \ge 0$, we say that z is ϵ_F-fractional if $\epsilon_F \le z \le 1 - \epsilon_F$.

The following lemma is the main result of this section

Lemma 6.1 (Matching Lemma). *For any $S \in \mathcal{H}$, $\epsilon_F \le 1/10$, $\epsilon_B \ge 21\epsilon_{1/2}, \alpha \ge 2\epsilon_\eta$, $\epsilon_{1/2} \le 0.0002$, there is a matching from good edges (see Definition 5.3) in $E^{\rightarrow}(S)$ to edges in $\delta(S)$ where every good edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ (where $u, v \in \mathcal{A}(S)$) is matched to a fraction $m_{\mathbf{e},u}$ of edges in $\delta^{\uparrow}(u)$ and a fraction $m_{\mathbf{e},v}$ of $\delta^{\uparrow}(v)$, and*

$$m_{\mathbf{e},u} F_u + m_{\mathbf{e},v} F_v \le x_{\mathbf{e}}(1 + \alpha), \tag{6.1}$$

$$\sum_{e \in \delta^{\rightarrow}(u)} m_{\mathbf{e},u} = x(\delta^{\uparrow}(u)) Z_u, \tag{6.2}$$

where for every atom $u \in \mathcal{A}(S)$, define

$$F_u = 1 - \epsilon_B \mathbb{I}\{x(\delta^{\uparrow}(u)) \text{ is } \epsilon_F \text{ fractional}\},$$
$$Z_u := \left(1 + \mathbb{I}\{|\mathcal{A}(S)| \ge 4, x(\delta^{\uparrow}(u)) \le \epsilon_F\}\right).$$

Roughly speaking, the intention of the previous lemma is to match good edges in $E^{\rightarrow}(S)$ to a similar fraction of edges that go higher (such that an edge bundle e adjacent to atoms u, v is only matched to edges in $\delta^{\uparrow}(u), \delta^{\uparrow}(v)$). Because we never "reduce" bad edges in the proof of payment theorem (Theorem 4.5), we do not use them in the matching. That inherently can cause a problem, as there could not be "enough" good edges in $E^{\rightarrow}(S)$ to saturate the edges going higher in the matching. The parameter F_u help us in this regard; in particular, it allows us to match some of the (good) edges in $E^{\rightarrow}(S)$ to more than their fraction in $\delta(S)$.

Next, we motivate the parameter Z_u. If $x(\delta^{\uparrow}(u)) \approx 0$, when those edges are reduced the conditional probability that $\delta(u)_T$ is even could be very close to zero. The parameter Z_u lets us match twice as many edges to $\delta^{\uparrow}(u)$; therefore, there will be only half a burden to fix the parity of $\delta(u)_T$. See the discussion in overview section for more details.

Throughout this section, we adopt the following notation: For a cut $S \in \mathcal{H}$ and a set $W \subseteq \mathcal{A}(S)$, we write

$$E(W, S \smallsetminus W) := \cup_{u \in W, v \in \mathcal{A}(S) \smallsetminus W} E(u,v),$$

$$\delta^{\uparrow}(W) := \cup_{u \in W} \delta^{\uparrow}(u) = \delta(W) \cap \delta(S),$$

$$\delta^{\rightarrow}(W) := \cup_{u \in W} \delta^{\rightarrow}(u).$$

In $\delta^{\rightarrow}(W) \not\subseteq \delta(W)$, because it includes edge bundles between atoms in W.

Before proving the main lemma we record the following facts.

Lemma 6.2. *For any $S \in \mathcal{H}$ and $W \subsetneq \mathcal{A}(S)$ (recall $\mathcal{A}(S)$ is the set of $u \in \mathcal{H}$ with $\mathsf{p}(u) = S$), we have*

$$x(\delta^{\rightarrow}(W)) \geq \frac{1}{2}\sum_{u \in W} x(\delta(a)) - \epsilon/2 \geq |W| - \epsilon/2.$$

We have

$$x(\delta^{\rightarrow}(W)) = \frac{1}{2}\left(\sum_{u \in W}(x(\delta(u)) + x(E(W, S \smallsetminus W)) - x(\delta^{\uparrow}(W))\right).$$

Because $x(\delta(S \smallsetminus W)) \geq 2$ and $x(\delta(S)) \leq 2 + \epsilon$, we have

(a) $x(E(W, S \smallsetminus W)) + x(\delta^{\uparrow}(S \smallsetminus W)) \geq 2$ *and*

(b) $x(\delta^{\uparrow}(W)) + x(\delta^{\uparrow}(S \smallsetminus W)) \leq 2 + \epsilon$.

Subtracting (b) from (a), we get

$$x(E(W, S \smallsetminus W)) - x(\delta^{\uparrow}(W)) \geq -\epsilon,$$

which after substituting into the previous equation, completes the proof of the first inequality in the lemma statement. The second inequality follows from the fact that $\delta(u) \geq 2$ for each atom u. □

Lemma 6.3. *For $S \in \mathcal{H}$, if $|\mathcal{A}(S)| = 3$, then there are no bad edges in $E^{\rightarrow}(S)$.*

Suppose $\mathcal{A}(S) = \{u, v, w\}$ and $\mathbf{e} = (u, v)$ is a bad edge bundle. Then $\left|x_{\mathbf{e}} - \frac{1}{2}\right| \leq \epsilon_{1/2}$. In addition, by Theorem 5.1, $x(\delta^{\uparrow}(u)), x(\delta^{\uparrow}(v)) \leq 1/2 + 9\epsilon_{1/2}$. Therefore,

$$x_{(\mathbf{u},\mathbf{w})} = x(\delta(u)) - x_{\mathbf{e}} - x(\delta^{\uparrow}(u)) \geq 1 - 10\epsilon_{1/2}.$$

Similarly, $x_{(\mathbf{v},\mathbf{w})} \geq 1 - 10\epsilon_{1/2}$. Finally, because $x(\delta(S)) \geq 2$, and $x(\delta^{\uparrow}(u)), x(\delta^{\uparrow}(v)) \leq 1/2 + 9\epsilon_{1/2}$, we must have $x(\delta^{\uparrow}(w)) \geq 1 - 18\epsilon_{1/2}$. However, this contradicts the assumption that $w \in \mathcal{H}$ must satisfy $x(\delta(w)) \leq 2 + \epsilon_\eta$.

Proof of Lemma 6.1. We will prove this by setting up a max-flow min cut problem. Construct a graph with vertex set $\{s, X, Y, t\}$, where s, t are the source and sink. We identify X with the set of good edge bundles in $E^{\rightarrow}(S)$ and Y with the set of atoms in $\mathcal{A}(S)$. For every edge bundle $\mathbf{e} \in X$, add an arc from s to $\mathbf{e}$ of capacity $c(s, \mathbf{e}) := (1 + \alpha)x_{\mathbf{e}}$. For every $u \in \mathcal{A}(S)$, there is an arc (u, t) with capacity

$$c(u, t) = x(\delta^{\uparrow}(u))F_u Z_u.$$

Finally, connect $\mathbf{e} = (u, v) \in X$ to nodes u and $v \in Y$ with a directed edge of infinite capacity, that is, $c(\mathbf{e}, u) = c(\mathbf{e}, v) = \infty$. We will show later that there is a flow saturating t, that is, there is a flow of value

$$c(t) := \sum_{u \in \mathcal{A}(S)} c(u, t) = \sum_{u \in \mathcal{A}(S)} x(\delta^{\uparrow}(u))F_u Z_u.$$

Suppose that in the corresponding max-flow, there is a flow of value $f_{\mathbf{e},u}$ on the edge $(\mathbf{e}, u)$. Define

$$m_{\mathbf{e},u} := \frac{f_{\mathbf{e},u}}{F_u}.$$

Then (6.1) follows from the fact that the flow leaving $\mathbf{e}$ is at most the capacity of the edge from s to $\mathbf{e}$, and (6.2) follows by conservation of flow on the node u (after cancelling out F_u from both sides).

We have left to show that for any s-t cut $A, \overline{A}$ where $s \in A, t \in \overline{A}$ that the capacity of this cut is at least $c(t)$.

Claim 6.1. *If $A = \{s\}$, then capacity of $(A, \overline{A})$ is at least $c(t)$.*

First, note that

$$\begin{aligned} c(t) &= \sum_{u \in \mathcal{A}(S)} x(\delta^{\uparrow}(u))F_u Z_u \leq \sum_{u \in \mathcal{A}(S)} x(\delta^{\uparrow}(u))Z_u \\ &\leq \mathbb{I}\{|\mathcal{A}(S)| \geq 4\} \cdot |\{u \in \mathcal{A}(S) : x(\delta^{\uparrow}(u)) \leq \epsilon_F\}| \cdot \epsilon_F \\ &\quad + x(\delta(S)) \\ &\leq 2 + \epsilon_\eta + \epsilon_F \mathbb{I}\{|\mathcal{A}(S)| \geq 4\}|\mathcal{A}(S)|, \end{aligned} \tag{6.3}$$

because $F_u \leq 1$ and $Z_u = 1 + \mathbb{I}\{|\mathcal{A}(S)| \geq 4, x(\delta^{\uparrow}(u)) \leq \epsilon_F\}$.

Second, note that

$$\begin{aligned} x(E^{\rightarrow}(S)) &= \frac{1}{2}\sum_{u \in \mathcal{A}(S)}(x(\delta(u)) - x(\delta^{\uparrow}(u))) \\ &\geq \frac{2|\mathcal{A}(S)| - (2 + \epsilon_\eta)}{2} = |\mathcal{A}(S)| - 1 - \epsilon_\eta/2. \end{aligned}$$

Therefore, if there are k bad edges in $E^{\rightarrow}(S)$, then

$$x_G \geq |\mathcal{A}(S)| - 1 - \epsilon_\eta/2 - k\left(\frac{1}{2} + \epsilon_{1/2}\right). \tag{6.4}$$

Case 1: $|\mathcal{A}(S)| = 3$. Then $Z_u = 1$ for all $u \in \mathcal{A}(S)$ and by Lemma 6.3 all edges are good. Therefore, by Equation (6.4), $x(E^{\rightarrow}(S)) \geq 2 - \epsilon_\eta/2$. Thus, for $\alpha \geq 2\epsilon_\eta$ we have

$$c(s) = (1+\alpha)x_G \geq (2-\epsilon_\eta/2)(1+\alpha) \geq 2+\epsilon_\eta \underset{Eq.~(6.3)}{\geq} c(t)$$

as desired.

Case 2: $|\mathcal{A}(S)| \geq 5$. By Theorem 5.1 there is at most one bad half edge adjacent to every vertex. Therefore, there are at most $|\mathcal{A}(S)|/2$ bad edges, so by Equation (6.4),

$$\begin{aligned}&(1+\alpha)x_G\\ &\geq (1+\alpha)\left(|\mathcal{A}(S)| - 1 - \epsilon_\eta/2 - \frac{1}{2}|\mathcal{A}(S)|\left(\frac{1}{2}+\epsilon_{1/2}\right)\right)\\ &\geq 2+\epsilon_\eta+\epsilon_F|\mathcal{A}(S)| \underset{Eq.~(6.3)}{\geq} c(t),\end{aligned}$$

where the second to last inequality holds, using $\alpha \geq 2\epsilon_\eta$, $|\mathcal{A}(S)| \geq 5$, $\epsilon_{1/2} \leq 0.01$, and $\epsilon_F \leq 0.1$.

Case 3: $|\mathcal{A}(S)| = 4$, and we have zero or one bad edges. Then by Equation (6.4), $x_G \geq 2.5 - \epsilon_\eta/2 - \epsilon_{1/2}$, so by Equation (6.3), $(1+\alpha)x_G \geq 2+\epsilon_\eta+4\epsilon_F \geq c(t)$ for $\epsilon_F \leq 0.1$, $\alpha \geq 2\epsilon_\eta, \epsilon_{1/2} \leq 0.01$.

Case 4: $|\mathcal{A}(S)| = 4$, and there are two bad edges. Then they form a perfect matching inside S and for each $u \in \mathcal{A}(S)$, $x(\delta^\uparrow(u)) \leq 1/2 + 9\epsilon_{1/2}$ (see Theorem 5.1).

Therefore, it must also be the case that $x(\delta^\uparrow(u)) \geq \epsilon_F$ for each $u \in \mathcal{A}(S)$. If not, there would have to be a node $u' \in \mathcal{A}(S)$ such that $x(\delta^\uparrow(u')) \geq (2-\epsilon_F)/3 > 1/2 + 9\epsilon_{1/2}$, which is a contradiction to u' having an incident bad edge. Thus, for each $u \in \mathcal{A}(S)$, $x(\delta^\uparrow(u))$ is ϵ_F-fractional, that is, $F_u = 1-\epsilon_B$ and $Z_u = 1$ implying that $c(t) \leq (2+\epsilon_\eta)(1-\epsilon_B)$. Therefore, by Equation (6.4),

$$c(s) = (1+\alpha)x_G \geq (1+\alpha)(2-2\epsilon_{1/2}-\epsilon_\eta/2),$$

and the rightmost quantity is at least $c(t)$ for $\epsilon_B \geq 2\epsilon_{1/2}$ and $\alpha \geq 2\epsilon_\eta$.

From now on, we assume that the min s-t cut $A \neq \{s\}$. In the following, we will prove that for any set of *atoms* $W \subsetneq S$, we have

$$c(s, \delta^\rightarrow(W)) = (1+\alpha)x_G(\delta^\rightarrow(W)) \geq c(\delta^\uparrow(W), t), \tag{6.5}$$

where for a set F of edges, we write $x_G(F)$ to denote the total fractional value of good edges in F.

Let $A_X = A \cap X, A_Y = A \cap Y$ and so on. Assuming the previously inequality, let us prove the lemma: First, for the set of edges A_X chosen from X, let Q be the set of endpoints of all edge bundles in A_X (in $\mathcal{A}(S)$).

Observe that we must choose all atoms in Q inside A_Y due to the infinite capacity arcs, that it, $Q \subseteq A_Y$. Let $W = S \smallsetminus Q$. Note that $W \neq S$. Then,

$$\begin{aligned}c(A, \overline{A}) &= c(A_Y, t) + c(s, \overline{A}_X)\\ &\geq c(\delta^\uparrow(Q), t) + c(s, \delta^\rightarrow(W))\\ &= c(\delta^\uparrow(S), t) - c(\delta^\uparrow(W)) + c(s, \delta^\rightarrow(W)) \geq c(\delta^\uparrow(S), t),\end{aligned}$$

where the last inequality follows by (6.5).

Finally, we prove (6.5). Suppose atoms in W are adjacent to k bad edges. Then

$$x_G(\delta^\rightarrow(W)) = x(\delta^\rightarrow(W)) - x_B(\delta^\rightarrow(W)),$$

which by Fact 6.2 and the fact that each bad edge has fraction at most $1/2+\epsilon_{1/2}$, is

$$\geq |W| - \epsilon_\eta/2 - k(1/2+\epsilon_{1/2}). \tag{6.6}$$

To upper bound $c(\delta^\uparrow(W), t)$, we observe that for any $u \in \mathcal{A}(S)$,

$$c(u,t) \leq \begin{cases} x(\delta^\uparrow(u))Z_u \leq 1/5 & \text{if } x(\delta^\uparrow(u)) < \epsilon_F\\ (1/2+9\epsilon_{1/2})(1-\epsilon_B) & \text{if } x(\delta^\uparrow(u)) > \epsilon_F \text{ and } u \text{ incident}\\ & \text{to bad edge}\\ 1+\epsilon_\eta & \text{otherwise, using Lemma 2.3.}\end{cases}$$

Therefore, we can write

$$c(\delta^\uparrow(W), t) \leq k(1/2+9\epsilon_{1/2})(1-\epsilon_B) + (|W|-k)(1+\epsilon_\eta).$$

Now, to prove (6.5), using (6.6), it is enough to choose α and ϵ_B such that

$$\begin{aligned}&(1+\alpha)\left(|W| - \epsilon_\eta/2 - k(1/2+\epsilon_{1/2})\right)\\ &\quad\geq k(1/2+9\epsilon_{1/2})(1-\epsilon_B) + (|W|-k)(1+\epsilon_\eta),\end{aligned}$$

or equivalently,

$$\begin{aligned}&|W|(\alpha-\epsilon_\eta)\\ &\geq k(\alpha/2 + 10\epsilon_{1/2} + \alpha\epsilon_{1/2} - \epsilon_B/2 - 9\epsilon_B\epsilon_{1/2} - \epsilon_\eta)\\ &\quad + \frac{\epsilon_\eta}{2}(1+\alpha).\end{aligned}$$

Because every atom is adjacent to at most one bad edge, $k \leq |W|$ and $|W| \geq 1$, the inequality follows using $\epsilon_B \geq 21\epsilon_{1/2}$ and $\alpha > 2\epsilon_\eta$ and $\epsilon_{1/2} \leq 0.0002$ and $\epsilon_\eta \leq \epsilon_{1/2}^2$.

7. Reduction and Payment

In this section, we prove Theorem 4.5.

In Section 5, we defined a number of happy events, such as 2-1-1 happy or 2-2-2 happy and showed that each of these events occurs with probability at least p. In this section, we will subsample these events to define a corresponding decrease event that occurs with probability *exactly*[17] p. See Table 1 for a list of all relevant constants used in this section.

7.1. Reduction Events

- **Bottom edges.** For each polygon cut $S \in \mathcal{H}$, let $\mathcal{R}_S$ be the indicator of a uniformly random subset of measure p of the max flow event $\mathcal{E}_S$. When $\mathcal{R}_S = 1$ then in particular we know that the polygon S is happy.

- **Top edges.** For a top edge bundle $\mathbf{e} = (u, v)$ define

$$\mathcal{H}_{\mathbf{e},u} = \begin{cases} 1 & \text{if } \mathbf{e} \text{ is 2-1-1 happy and good w.r.t. } u \\ 1 & \text{if } \mathbf{e} \text{ is 2-2 happy and good, but not 2-1-1} \\ & \quad \text{good with respect to } u \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathcal{H}_{\mathbf{e},v}$ be defined similarly. Because p is a lower bound on the probability a good edge is happy, we may now let $\mathcal{R}_{\mathbf{e},u}$ and $\mathcal{R}_{\mathbf{e},v}$ be indicators of subsets of measure p of $\mathcal{H}_{\mathbf{e},u}$ and $\mathcal{H}_{\mathbf{e},v}$, respectively (note $\mathcal{R}_{\mathbf{e},u}$ and $\mathcal{R}_{\mathbf{e},v}$ may overlap). In this way every top edge bundle $\mathbf{e} = (u, v)$ is associated with indicators $\mathcal{R}_{\mathbf{e},u}$ and $\mathcal{R}_{\mathbf{e},v}$. In the special case that u is in case 3 (and not case 1 or 2) of Theorem 5.2, fix two half edge bundles $\mathbf{e}, \mathbf{f}$ that are neighbors of u that satisfy the conditions of case 3. For these edges, by Theorem 5.2, $\mathcal{H}_{\mathbf{e},u} \cap \mathcal{H}_{\mathbf{f},u}$ has measure at least p. This is because $\mathcal{H}_{\mathbf{e},u} \cap \mathcal{H}_{\mathbf{f},u}$ happens if and only if $\mathbf{e}, \mathbf{f}$ are 2-2-2 happy with respect to u. Here, we choose $\mathcal{R}_{\mathbf{e},u}, \mathcal{R}_{\mathbf{f},u}$ to be the same subset of measure p of $\mathcal{H}_{\mathbf{e},u} \cap \mathcal{H}_{\mathbf{f},u}$.

Define $r : E \to \mathbb{R}_{\geq 0}$ as follows: For any (nonbundle) edge e,

$$r_e = \begin{cases} \beta x_e \mathcal{R}_S & \text{if } \mathsf{p}(e) = S \text{ for a polygon} \\ & \quad \text{cut } S \in \mathcal{H} \\ \dfrac{1}{2}\tau x_e(\mathcal{R}_{\mathbf{f},u} + \mathcal{R}_{\mathbf{f},v}) & \text{if } e \in \mathbf{f} \text{ for a top edge} \\ & \quad \text{bundle } \mathbf{f} = (u, v), \end{cases}$$

for β, the parameter of Theorem 4.5 and τ as defined in global constants.

7.2. Increase Events

Let $\mathbf{E}$ be the set of edge bundles, that is, top/bottom edge bundles. Now, we define the increase vector $I : \mathbf{E} \to \mathbb{R}_{\geq 0}$ as follows:

- **Bottom edges.** For each polygon $S \in \mathcal{H}$ (and corresponding bottom edge bundle) with polygon partition A, B, C, let $r(A) := \sum_{f \in A} r_f$, $r(B) := \sum_{f \in B} r_f$, and $r(C) := \sum_{f \in C} r_f$. Then set

$$\begin{aligned} I_S := (1 + \epsilon_\eta)(\max\{r(A) \cdot \mathbb{I}\{S \text{ not left happy}\}, r(B) \\ \cdot \mathbb{I}\{S \text{ not right happy}\}\} \\ + r(C)\mathbb{I}\{S \text{ not happy}\}). \end{aligned} \tag{7.1}$$

- **Top edges.** For every degree cut $S \in \mathcal{H}$, invoke Lemma 6.1 with

$$\alpha = 2\epsilon_\eta, \epsilon_B = 21\epsilon_{1/2}, \epsilon_F = 1/10 \qquad \text{(Matching parameters)}$$

and let $m_{\mathbf{e},u}$ be the resulting matching for every $u \in \mathcal{A}(S)$. For each top edge bundle $\mathbf{e} = (u, v)$, let

$$I_{\mathbf{e},u} := \sum_{g \in \delta^{\uparrow}(u)} r_g \cdot \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f} \in \delta^{\rightarrow}(u)} m_{\mathbf{f},u}} \mathbb{I}\{u \text{ is odd}\}, \tag{7.2}$$

and define $I_{\mathbf{e},v}$ analogously. Let $I_{\mathbf{e}} = I_{\mathbf{e},u} + I_{\mathbf{e},v}$.

The following theorem is the main technical result of this section.

Theorem 7.1. *For any good top edge bundle* $\mathbf{e}$, $\mathbb{E}[I_{\mathbf{e}}] \leq (1 - \frac{\epsilon_{1/1}}{6}) p \tau x_{\mathbf{e}}$, *and for any bottom edge bundle* S, $\mathbb{E}[I_S] \leq 0.99994 \beta p$.

Using this theorem, we can prove the desired theorem.

Theorem 4.5 (Main Payment Theorem). *For an LP solution* x^0 *and* x *be* x^0 *restricted to* E *and a hierarchy* $\mathcal{H}$ *for some* $\epsilon_\eta \leq 10^{-10}$ *and any* $\beta > 0$, *the maximum entropy distribution* μ *with marginals* x *satisfies the following:*

(i) *There is a set of* good *edges* $E_g \subseteq E \setminus \delta(\{u_0, v_0\})$ *such that any bottom edge* e *is in* E_g *and for any (nonroot)* $S \in \mathcal{H}$ *such that* $\mathsf{p}(S)$ *is a degree cut, we have* $x(E_g \cap \delta(S)) \geq 3/4$.

(ii) *There is a random vector* $s : E_g \to \mathbb{R}$ *(as a function of* $T \sim \mu$*) such that for all* e, $s_e \geq -x_e \beta$ *(with probability 1), and*

(iii) *If a polygon cut* u *with polygon partition* A, B, C *is not left happy, then for any set* $F \subseteq E$ *with* $\mathsf{p}(e) = u$ *for all* $e \in F$ *and* $x(F) \geq 1 - \epsilon_\eta/2$, *we have*

$$s(A) + s(F) + s^-(C) \geq 0,$$

where $s^-(C) = \sum_{e \in C} \min\{s_e, 0\}$. *A similar inequality holds if* u *is not right happy.*

(iv) *For every cut* $S \in \mathcal{H}$ *such that* $\mathsf{p}(S)$ *is not a polygon cut, if* $\delta(S)_T$ *is odd, then* $s(\delta(S)) \geq 0$.

(v) *For a good edge* $e \in E_g$, $\mathbb{E}[s_e] \leq -\epsilon_P \beta x_e$ *(see Equation (7.4) for definition of* ϵ_P*).*

Proof of Theorem 4.5. First, we set the constants:

$$\epsilon_{1/2} = 0.0002, \epsilon_{1/1} = \frac{\epsilon_{1/2}}{12}, p = 0.005\epsilon_{1/2}^2, \epsilon_M = 0.00025, \\ \tau = 0.571\beta. \qquad \text{(Global constants)}$$

Define E_g to be the set of bottom edges together with any edge e which is part of a good top edge bundle. Now, we verify (i), we show for any $S \in \mathcal{H}$ such that $\mathsf{p}(S)$ is a degree cut, $x(E_g \cap \delta(S)) \geq 3/4$. First, by Theorem 5.1, if $x(\delta^{\uparrow}(S)) \geq 1/2 + 9\epsilon_{1/2}$ then all edges in $\delta^{\rightarrow}(S)$ are good, so the claim follows because by Lemma 2.3, $x(\delta^{\rightarrow}(S)) \geq 1 - \epsilon_\eta \geq 3/4$. Otherwise, $x(\delta^{\uparrow}(S)) \leq 1/2 + 9\epsilon_{1/2}$. Then, by Theorem 5.1, there is at most one bad edge in $\delta^{\rightarrow}(S)$. Therefore, there is a fraction at least $x(\delta^{\rightarrow}(S)) - (1/2 + \epsilon_{1/2}) \geq 3/4$ of good edges in $\delta^{\rightarrow}(S)$.

For any edge $e \in E'$, define

$$s_e = -r_e + \begin{cases} I_{\mathbf{f}} \dfrac{x_e}{x_{\mathbf{f}}} & \text{if } e \in \mathbf{f} \text{ for a top edge bundle } \mathbf{f}, \\ I_S x_e & \text{if } \mathsf{p}(e) = S \text{ for a polygon cut } S \in \mathcal{H}. \end{cases} \tag{7.3}$$

Now, we verify (ii): First, we observe that $s_e = 0$ (with probability 1) if e is part of a bad edge bundle because we defined reduction events only for good edges and $m_{\mathbf{e},u}$ is nonzero only for good edge bundles. Because

Table 1. All Constants Used in the Paper

Name	Value	Set in	Explanation
$\epsilon_{1/2}$	0.0002	Global constants	Half edge threshold, Definition 5.2
$\epsilon_{1/1}$	$\frac{\epsilon_{1/2}}{12}$	Global constants	A, B, C partitioning threshold, Definition 5.4
p	$0.005\epsilon_{1/2}^2$	Global constants	Min prob. of happiness for a (2-*) good edge
ϵ_M	0.00025	Global constants	Marginal errors due to max flow, Definition 5.1
τ	0.571β	Global constants	Top edge decrease
ϵ_P	$\frac{\epsilon_{1/1}}{6}p\frac{\tau}{\beta}$	(7.4)	Expected decrease constant, Theorem 4.5
α	$2\epsilon_\eta$	Matching parameters	Parameter of Lemma 6.1
ϵ_B	$21\epsilon_{1/2}$	Matching parameters	Parameter of Lemma 6.1
ϵ_F	1/10	Matching parameters	Parameter of Lemma 6.1
ϵ_η	14η	(4.5)	Definition 4.11
η	$\frac{1}{1{,}308}\epsilon_P$	(3.2)	Near min cut constant
β	$\eta/4.1$	(3.1)	Slack shift constant; e.g., Theorems 4.1, 4.4, and 4.5

$r_e \le \beta x_e$ for bottom edges and $r_e \le \tau x_e$ for top edges, and $\tau \le \beta$, it follows that $s_e \ge -x_e\beta$ with probability 1.

Now, we verify (iii): Suppose a polygon cut u is not left-happy. Because u is not happy, we must have $\mathcal{R}_u = 0$ and $r_e = 0$ for any $e \in F$. Therefore,

$$\begin{aligned} s(A) + s(F) + s^-(C) &= s(A) + I_S x(F) + s^-(C) \\ &\ge -r(A) + (1+\epsilon_\eta)(r(A) + r(C)) \\ &\quad (1 - \epsilon_\eta/2) - r(C) \ge 0. \end{aligned}$$

Here, we used that $x(F) \ge 1 - \epsilon_\eta/2$.

Now, we verify (iv): Let $S \in \mathcal{H}$, where $\mathsf{p}(S)$ is a degree cut. If S is odd, then $r_e = 0$ for all edges $e \in \delta^{\rightarrow}(S)$; therefore, by Equation (7.2)

$$\begin{aligned} s(\delta(S)) &\ge -\sum_{g \in \delta^{\uparrow}(S)} r_g + \sum_{\mathbf{e} \in \delta^{\rightarrow}(S)} I_{\mathbf{e},S} \\ &= -\sum_{g \in \delta^{\uparrow}(S)} r_g + \sum_{\mathbf{e} \in \delta^{\rightarrow}(S)} \sum_{g \in \delta^{\uparrow}(S)} r_g \frac{m_{\mathbf{e},S}}{\sum_{\mathbf{f} \in \delta^{\rightarrow}(S)} m_{\mathbf{f},S}} = 0. \end{aligned}$$

Finally, we verify (v): Here, we use Theorem 7.1. For a good top edge e that is part of a top edge bundle $\mathbf{f}$, we have

$$\begin{aligned} \mathbb{E}[s_e] &= -\mathbb{E}[r_e] + \mathbb{E}[I_{\mathbf{f}}]\frac{x_e}{x_{\mathbf{f}}} \le -\tau p x_e + \left(1 - \frac{\epsilon_{1/1}}{6}\right)p\tau \\ x_e &= -\frac{\epsilon_{1/1}}{6}p\tau x_e. \end{aligned}$$

On the other hand, for a bottom edge e with $\mathsf{p}(e) = S$, then

$$\begin{aligned} \mathbb{E}[s_e] &= -\mathbb{E}[r_e] + \mathbb{E}[I_S]x_e \le -\beta p x_e + 0.99994 p\beta x_e \\ &\le -0.00006 p\beta x_e. \end{aligned}$$

Finally, we can let

$$\begin{aligned} \epsilon_P &:= \frac{\epsilon_{1/1}}{6}p\frac{\tau}{\beta} = \frac{\epsilon_{1/2}}{72}0.005\epsilon_{1/2}^2 0.571 \ge 0.000039\epsilon_{1/2}^3 \\ &\ge 3.12 \cdot 10^{-16} \end{aligned} \tag{7.4}$$

as desired.

In the rest of this section, we prove Theorem 7.1. Throughout the proof, we will repeatedly use the following facts proved in Section 5: If a top edge $e = (u, v)$ that is part of a bundle $\mathbf{f}$ is reduced (equivalently $\mathcal{H}_{\mathbf{f},u} = 1$ or $\mathcal{H}_{\mathbf{f},v} = 1$), then u and v are trees, which means that tree sampling inside u and v is independent of the reduction of e.

However, that conditioning on a near-min cut or atom to be a tree increases marginals inside and reduces marginals outside as specified by Lemma 2.7. Because for any $S \in \mathcal{H}$, $x(\delta(S)) \le 2 + \epsilon_\eta$, the overall change is $\pm\epsilon_\eta/2$.

The proof of Theorem 7.1 simply follows from Lemma 7.1 and Lemma 7.4 that we will prove in the following two sections.

7.3. Increase for Good Top Edges

The following lemma is the main result of this section.

Lemma 7.1 (Top Edge Increase). *Let $S \in \mathcal{H}$ be a degree cut and $\mathbf{e} = (u, v)$ a good edge bundle with $\mathsf{p}(\mathbf{e}) = S$. If $\epsilon_{1/2} \le 0.0002$, $\epsilon_{1/1} \le \epsilon_{1/2}/12$ and $\epsilon_\eta \le \frac{\epsilon_{1/1}}{100}$, $\epsilon_F = 1/10$ then*

$$\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] \le p\tau x_{\mathbf{e}}\left(1 - \frac{\epsilon_{1/1}}{6}\right).$$

We will use the following technical lemma to prove the previous lemma.

Lemma 7.2. *Let $S \in \mathcal{H}$ be a degree cut with an atom $u \in \mathcal{A}(S)$. If $x(\delta^{\uparrow}(u)) > \epsilon_F$, $\epsilon_{1/2} \le 0.0002$, $\epsilon_{1/1} \le \epsilon_{1/2}/12$, $\epsilon_\eta \le \frac{\epsilon_{1/1}}{100}$, then we have*

$$\begin{aligned} &\sum_{\substack{g \in \delta^{\uparrow}(u), \\ g \in \mathbf{f} = (u', v') \text{ good top}}} \frac{1}{2}\tau x_g \cdot (\mathbb{P}[\delta(u)_T \text{ odd} \mid \mathcal{R}_{\mathbf{f},u'}] \\ &\qquad + \mathbb{P}[\delta(u)_T \text{ odd} \mid \mathcal{R}_{\mathbf{f},v'}]) \\ &\qquad + \sum_{g \in \delta^{\uparrow}(u), \mathsf{p}(g) = S' \text{ polygon}} \beta x_g \cdot \mathbb{P}[\delta(u)_T \text{ odd} \mid \mathcal{R}_{S'}] \\ &\le \tau\left(1 - \frac{\epsilon_{1/1}}{5}\right)x(\delta^{\uparrow}(u))F_u, \end{aligned} \tag{7.5}$$

where recall we set $F_u := 1 - \epsilon_B \mathbb{I}\{x(\delta^{\uparrow}(u))$ is ϵ_F fractional$\}$ in Lemma 6.1, where $\epsilon_B := 21\epsilon_{1/2}$ and $\epsilon_F = 1/10$ as in matching parameters.

Proof of Lemma 7.1. By linearity of expectation and using Equation (7.2):

$$\mathbb{E}[I_{\mathbf{e},u}] = \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f}\in\delta^{\rightarrow}(u)} m_{\mathbf{f},u}} \mathbb{E}\left[\sum_{g\in\delta^{\uparrow}(u)} r_g \cdot \mathbb{I}\{u \text{ is odd}\}\right]$$

$$= \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f}\in\delta^{\rightarrow}(u)} m_{\mathbf{f},u}} \left(\sum_{\substack{g\in\delta^{\uparrow}(u):\\ g\in\mathbf{f}=(u',v')\text{ good top}}} \frac{1}{2}\tau x_g (\mathbb{P}[\mathcal{R}_{\mathbf{f},u'}, \delta(u)_T \text{ odd}] + \mathbb{P}[\mathcal{R}_{\mathbf{f},v'}, \delta(u)_T \text{ odd}]) + \sum_{g\in\delta^{\uparrow}(u):\mathsf{p}(g)=S' \text{ polygon}} \beta x_g \mathbb{P}[\mathcal{R}_{S'}, \delta(u)_T \text{ odd}] \right). \quad (7.6)$$

A similar equation holds for $\mathbb{E}[I_{e,v}]$.

The case where $x(\delta^{\uparrow}(u)) \le \epsilon_F$ or $x(\delta^{\uparrow}(v)) \le \epsilon_F$ is dealt with in Lemma 7.3. Therefore, consider the case where $x(\delta^{\uparrow}(u)), x(\delta^{\uparrow}(v)) > \epsilon_F$. Now recall that from (6.2),

$$\sum_{\mathbf{f}\in\delta^{\rightarrow}(u)} m_{\mathbf{f},u} = Z_u x(\delta^{\uparrow}(u)), \quad (7.7)$$

where $Z_u = 1 + \mathbb{I}\{|S| \ge 4, x(\delta^{\uparrow}(u)) \le \epsilon_F\}$. In this case, $Z_u = Z_v = 1$.

Using $\mathbb{P}[\mathcal{R}_{\mathbf{f},u'}, \delta(u)_T \text{ odd}] = p\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}]$, and plugging (7.5) into (7.6) for u and v, we get (and using Equation (7.7)):

$$\begin{aligned}
&\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] \\
&\le p\tau\left(1 - \frac{\epsilon_{1/1}}{5}\right)\left(x(\delta^{\uparrow}(u))F_u \frac{m_{\mathbf{e},u}}{x(\delta^{\uparrow}(u))} + x(\delta^{\uparrow}(v))F_v \frac{m_{\mathbf{e},v}}{x(\delta^{\uparrow}(v))}\right) \\
&= p\tau\left(1 - \frac{\epsilon_{1/1}}{5}\right)(F_u m_{\mathbf{e},u} + F_v m_{\mathbf{e},v}) \\
&\le p\tau\left(1 - \frac{\epsilon_{1/1}}{5}\right)(1 + 2\epsilon_\eta)x_{\mathbf{e}} < p\tau x_{\mathbf{e}}\left(1 - \frac{\epsilon_{1/1}}{6}\right). \quad (7.8)
\end{aligned}$$

Here, on the final line, we used (6.1) and $\epsilon_\eta < \frac{\epsilon_{1/1}}{100}$.

Proof of Lemma 7.2. Suppose that $S_i \in \mathcal{H}$ are the ancestors of S in the hierarchy (in order) such $S_1 = S$ and for each i, $S_{i+1} = \mathsf{p}(S_i)$. Let

$$\delta^{\ge i} := \delta(u) \cap \delta(S_i) \qquad \text{and} \qquad \delta^i := \delta(u) \cap \delta^{\rightarrow}(S_i).$$

Each group of edges δ^i is either entirely top edges or entirely bottom edges. First, if $g \in \delta^i$ and g is a bottom edge, that is, S_{i+1} is a polygon cut, then by Corollary 5.3,

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{S_{i+1}}] = \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{E}_{S_{i+1}}] \le 0.5678$$

(see Definition 5.1 and Section 7 for definition of $\mathcal{E}_{S_{i+1}}$, $\mathcal{R}_{i+1}$), where in the equality we used that $\mathcal{R}_{S_{i+1}}$ is a uniformly random event chosen in $\mathcal{E}_{S_{i+1}}$. Therefore, to prove Equation (7.5), it is enough to show

$$\begin{aligned}
&\sum_{\substack{g\in\delta^{\uparrow}_{\text{good}}(u):\\ g\in\mathbf{f}=(u',v')\text{ top},}} \frac{1}{2}\tau x_g (\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] + \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},v'}]) \\
&\le \tau\left(\left(1 - \frac{\epsilon_{1/1}}{5}\right)F_u\left(x(\delta^{\uparrow}_{\text{good}}(u)) + x(\delta^{\uparrow}_{\text{bad}}(u))\right) + 0.0014 x(\delta^{\uparrow}_{\beta}(u))\right), \quad (7.9)
\end{aligned}$$

where we write $\delta_\beta(u), \delta_{\text{good}}(u), \delta_{\text{bad}}(u)$ to denote the set of bottom edges, good top edges, and bad (top) edges in $\delta(u)$, respectively, and we used that

$$\tau\left(1 - \frac{\epsilon_{1/1}}{5}\right)(1 - \epsilon_B) - 0.5678\beta \ge 0.0014\tau$$

because $\tau = 0.571\beta$, $\epsilon_{1/1} \le \frac{\epsilon_{1/2}}{12}$, $\epsilon_{1/2} \le 0.0002$, and $\epsilon_B = 21\epsilon_{1/2}$ as defined in matching parameters.

Because $h(\mathbf{f}) := \frac{1}{2}(\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] + \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},v'}]) \le 1$ and $(1 - \frac{\epsilon_{1/1}}{5})F_u$ is nearly one, in each of the following cases,

$$x(\delta^{\uparrow}_{\beta}(u)) \ge \begin{cases} 0.003 & \text{when } F_u = 1 \\ \frac{4}{5}x(\delta^{\uparrow}(u)) & \text{when } F_u = 1 - \epsilon_B \end{cases} \quad \text{or} \quad x(\delta^{\uparrow}_{\text{bad}}(u)) \ge 0.006 \text{ when } F_u \ge 1 - \epsilon_B, \quad (7.10)$$

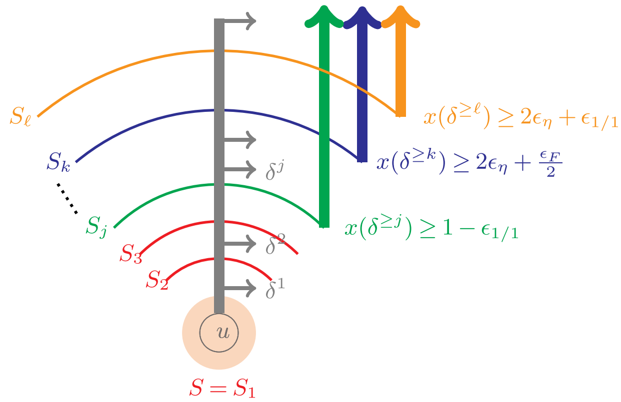
(7.9) holds. To see this, just plug in $\epsilon_{1/1} \le \frac{\epsilon_{1/2}}{12}$, $\epsilon_{1/2} \le 0.0002$, $\epsilon_B = 21\epsilon_{1/2}$, $\epsilon_\eta \le 10^{-10}$, $x(\delta^{\uparrow}(u)) \le 1 + \epsilon_\eta$, and any inequality from (7.10) into (7.9), using the upper bound $h(\mathbf{f}) = 1$.

Alternatively, for $\delta_{\text{top}}(u) = \delta_{\text{good}}(u) \cup \delta_{\text{bad}}(u)$ be the set of top edges in $\delta(u)$, if we can show the existence of a set $D \subseteq \delta^{\uparrow}_{\text{top}}(u)$ such that

$$x(D) \cdot \min_{\substack{g\in D:\\ g\in\mathbf{f}=(u',v')\text{ good}}} 1 - \frac{\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] + \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},v'}]}{2} \ge \left(\frac{\epsilon_{1/1}}{5} + 1 - F_u\right)x(\delta^{\uparrow}_{\text{top}}(u)), \quad (7.11)$$

then again, (7.9) holds.

In the rest of the proof, we will consider a number of cases and show that in each of them, either one of the inequalities in (7.10) or the inequality in (7.11) for some set D is true, which will imply the lemma.



First, let

$$j = \max\{i : x(\delta^{\geq i}) \geq 1 - \epsilon_{1/1}\}$$
$$k = \max\{i : x(\delta^{\geq i}) \geq 2\epsilon_\eta + \epsilon_F/2\},$$
$$\ell = \max\{i : x(\delta^{\geq i}) \geq 2\epsilon_\eta + \epsilon_{1/1}\}$$

noting that $j \leq k \leq \ell$. Levels ℓ and k exist because $x(\delta^\uparrow(u)) \geq \epsilon_F$, whereas level j may not exist (if $x(\delta^\uparrow(u)) < 1 - \epsilon_{1/1}$). We consider three cases.

Case 1: $x(\delta^\uparrow(u)) \geq 1 - \epsilon_{1/1}$**:** Then j exists and S_j has a valid A, B, C degree partitioning (Definition 5.4), where $A = \delta(v) \cap \delta(S_j)$ such that either $u = v$ or v is a descendant of u in $\mathcal{H}$. Note that, $x(\delta(u) \cap \delta(S_j)) \geq 1 - \epsilon_{1/1}$, and by Definition 5.4, $B \cap \delta(u) = \emptyset$. In addition, in this case, $x(\delta^\uparrow(u))$ is not ϵ_F fractional (see Lemma 6.1), so $F_u = 1$.

Case 1a: $x(\delta^j) \geq 3/4$. If δ^j are bottom edges, then (7.10) holds. Therefore, suppose that δ^j is a set of top edges. By Lemma 5.11, at most $1/2 + 4\epsilon_{1/2}$ fraction of edges in $A \cap \delta^j$ are good but not 2-1-1 good (w.r.t., u). Therefore, the rest of the edges in $A \cap \delta^j$ are either bad or 2-1-1 good. Because

$$x(A \cap \delta^j) \geq 3/4 - x(C) \geq 3/4 - 2\epsilon_{1/1} - \epsilon_\eta,$$

δ^j either has a mass of $\frac{1}{2}(1/4 - 2\epsilon_{1/1} - \epsilon_\eta - 4\epsilon_{1/2}) > 1/8 - 3\epsilon_{1/2}$ of bad edges or of 2-1-1 good edges.[18] The former case implies that (7.10) holds. In the latter case, by Claim 7.1, for any 2-1-1 good edge $g \in \delta^j$ with $g \in \mathbf{f} = (u', v')$, we have $\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] \leq 2\epsilon_\eta + \epsilon_{1/1}$; therefore, (7.11) holds for D defined as the set of 2-1-1 good edges in δ^j.

Case 1b: $x(\delta^j) < 3/4$. If $x(\delta^\uparrow_\beta(u)) \geq 0.003$, then (7.10) holds. Otherwise, we apply Claim 7.2 with $\epsilon = \epsilon_{1/1}$ to all good top edge bundles $\mathbf{f} \in D = \delta^{\geq j+1} \smallsetminus \delta^{\geq \ell+1}$, and we get that

$$\frac{1}{2}(\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] + \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},v'}]) \leq 1 - \epsilon_{1/1} + \epsilon_{1/1}^2.$$

Because $x(D) \geq 1 - \epsilon_{1/1} - 3/4 - 2\epsilon_\eta - \epsilon_{1/1} - 0.003 > 0.24$, (7.11) holds.

Case 2: $1 - \epsilon_F < x(\delta^\uparrow(u)) < 1 - \epsilon_{1/1}$. Again we have $F_u = 1$. Therefore, we can either show that $x(\delta^\uparrow_\beta(u)) \geq 0.003$ or take D to be the top edges in $\delta^\uparrow(u) \smallsetminus \delta^{\geq \ell+1}$ and use Claim 7.2 with $\epsilon = \epsilon_{1/1}$. This will enable us to show that (7.11) holds as in the previous case.

Case 3: $\epsilon_F < x(\delta^\uparrow(u)) < 1 - \epsilon_F$**:** In this case, $F_u = 1 - \epsilon_B$. If at least 4/5 of the edges in $\delta^\uparrow(u)$ are bottom edges, then we are done by (7.10).

Otherwise, let $u' = \mathsf{p}(u)$. For any top edge $e \in \delta^\uparrow(u)$, where $e \in \mathbf{f} = (u'', v'')$, we have

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u''}] \leq \mathbb{P}[u' \text{ tree} | \mathcal{R}_{\mathbf{f},u''}]$$
$$\mathbb{P}[\delta(u)_T \text{ odd} | u' \text{ tree}, \mathcal{R}_{\mathbf{f},u''}] + \mathbb{P}[u' \text{ not tree} | \mathcal{R}_{\mathbf{f},u''}].$$

Using that $u' \subseteq u''$ is a tree under $|\mathcal{R}_{\mathbf{f},u''}$ with probability at least $1 - \epsilon_\eta/2$, and applying Claim 7.2 (to u and u') with $\epsilon = \epsilon_F$ we have $\mathbb{P}[\delta(u)_T \text{ odd} | u' \text{ tree}, \mathcal{R}_{\mathbf{f},u''}] \leq 1 - \epsilon_F + \epsilon_F^2$, and we get

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u''}] \leq 1 - \epsilon_F + \epsilon_F^2 + \epsilon_\eta/2.$$

Now, let D be all top edges in $\delta^\uparrow(u)$. Then, we apply Equation (7.11) to this set of mass at least $x(\delta^\uparrow(u))/5$, and we are done, using that $(\epsilon_F - 2\epsilon_F^2)/5 \geq \left(\frac{\epsilon_{1/1}}{5} + \epsilon_B\right)$ that holds for $\epsilon_F \geq 1/10$, $\epsilon_B = 21\epsilon_{1/2}$, and $\epsilon_{1/2} \leq 0.0002$. □

Claim 7.1. *For $u \in \mathcal{H}$ and a top edge $e \in \mathbf{f} = (u', v')$ for some $u' \in \mathcal{H}$ that is an ancestor of u, if $x(\delta(u) \cap \delta(u')) \geq 1 - \epsilon_{1/1}$ and $\mathbf{f}$ is 2-1-1 good, then*

$$\mathbb{P}[\delta(u)_T \textit{ odd} | \mathcal{R}_{\mathbf{f},u'}] \leq 2\epsilon_\eta + \epsilon_{1/1}.$$

Let A, B, C be the degree partitioning of $\delta(u')$. By the assumption of the claim, without loss of generality, assume $A \subseteq \delta(u) \cap \delta(u')$. Furthermore, by definition, $B \cap \delta(u) = \emptyset$. This means that if $\mathcal{R}_{\mathbf{f},u'} = 1$, then u' is a tree and $A_T = 1 = (\delta(u) \cap \delta(u'))_T$ (also using $C_T = 0$ and $B \cap \delta(u) = \emptyset$). Therefore,

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] = \mathbb{P}[(\delta(u) \smallsetminus \delta(u'))_T \text{ even} | \mathcal{R}_{\mathbf{f},u'}].$$

To upper bound the RHS, first observe that

$$\begin{aligned}\mathbb{E}[(\delta(u) \smallsetminus \delta(u'))_T | \mathcal{R}_{\mathbf{f},u'}] &\leq \epsilon_\eta/2 + x(\delta(u) \smallsetminus \delta(u')) \\ &\leq \epsilon_\eta/2 + x(\delta(u)) - x(A) \\ &< 1 + 2\epsilon_\eta + \epsilon_{1/1}.\end{aligned}$$

Under the conditional measure $|\mathcal{R}_{\mathbf{f},u'}$, u' is a tree, so u must be connected inside u', that is, $(\delta(u) \smallsetminus \delta(u'))_T \geq 1$ with probability 1. Therefore,

$$\mathbb{P}[(\delta(u) \smallsetminus \delta(u'))_T \text{ even} | \mathcal{R}_{\mathbf{f},u'}] \leq \mathbb{P}[(\delta(u) \smallsetminus \delta(u'))_T - 1 \neq 0 | \mathcal{R}_{\mathbf{f},u'}] \leq 2\epsilon_\eta + \epsilon_{1/1}$$

as desired.

Claim 7.2. *For $u, u' \in \mathcal{H}$ such that u' is an ancestor of u. Let $\nu = \nu_{u'} \times \nu_{G/u'}$ be the measure resulting from conditioning*

u' to be a tree. if $x(\delta(u)\cap\delta(u'))\in[\epsilon,1-\epsilon]$, then

$$\mathbb{P}_\nu[\delta(u)\ odd\,|\,(\delta(u)\cap\delta(u'))_T]\le 1-\epsilon+\max\{2\epsilon_\eta,\epsilon^2\}. \tag{7.12}$$

In other words, for any integer $k\ge 0$, we have $\mathbb{P}_\nu[\delta(u)\ odd\,|\,(\delta(u)\cap\delta(u'))_T=k]\le 1-\epsilon+\max\{2\epsilon_\eta,\epsilon^2\}$.

Let $D=\delta(u)\smallsetminus\delta(u')$. By assumption, u' is a tree, so $D_T\ge 1$ with probability 1. Therefore, because we have no control over the parity of $(\delta(u)\cap\delta(u'))_T$,

$$\begin{aligned}&\mathbb{P}_\nu[\delta(u)_T\ \text{even}\,|\,(\delta(u)\cap\delta(u'))_T]\\&\ge\min\{\mathbb{P}[D_T-1\ \text{odd}\,|\,u'\ \text{tree}],\mathbb{P}[D_T-1=0\,|\,u'\ \text{tree}]\},\end{aligned}$$

where we removed the conditioning by taking the worst case over $(\delta(u)\cap\delta(u'))_T$ even, $(\delta(u)\cap\delta(u'))_T$ odd. First, observe by the assumption of the claim and that $x(\delta^\uparrow(u))\le 2+\epsilon_\eta$ we have

$$\mathbb{E}[D_T-1\,|\,u'\ \text{tree}]\in[\epsilon,1-\epsilon+2\epsilon_\eta].$$

Furthermore, because we have a SR distribution on $G[u'],D_T-1$ is a Bernoulli sum random variable. Therefore,

$$\mathbb{P}[D_T-1=0\,|\,u'\ \text{tree}]\ge\epsilon-2\epsilon_\eta,$$

and by Corollary 2.1,

$$\mathbb{P}[D_T-1\ \text{odd}\,|\,u'\ \text{tree}]\ge 1-1/2(1+e^{-2\epsilon})\ge\epsilon-\epsilon^2$$

as desired.

Lemma 7.3. *Let $S\in\mathcal{H}$ be a degree cut and $\mathbf{e}=(u,v)$ a good edge bundle with $\mathsf{p}(\mathbf{e})=S$. If $x(\delta^\uparrow(u))<\epsilon_F$, $\epsilon_{1/2}\le 0.0002$, $\epsilon_{1/1}\le\epsilon_{1/2}/10$, then*

$$\mathbb{E}[I_{\mathbf{e},u}]+\mathbb{E}[I_{\mathbf{e},v}]\le p\tau x_{\mathbf{e}}\left(1-\frac{\epsilon_{1/1}}{6}\right).$$

First notice, by Corollary 5.3, for any bottom edge $g\in\delta^\uparrow(u)$ with $\mathsf{p}(g)=S'$, we have

$$\mathbb{P}[\delta(u)_T\ \text{odd}\,|\,\mathcal{R}_{S'}]=\mathbb{P}[\delta(u)_T\ \text{odd}\,|\,\mathcal{E}_{S'}]\le 0.5678,$$

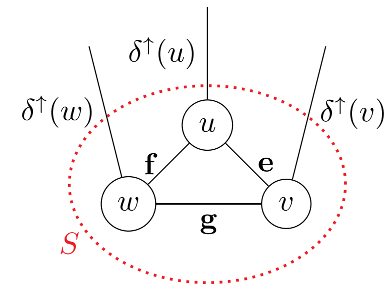
using $0.5678\beta\le\tau$ and $F_u=1$ (as $x(\delta^\uparrow(u))\le\epsilon_F$), we can write,

$$\mathbb{E}[I_{\mathbf{e},u}]\le\sum_{h\in\delta^\uparrow(u)}x_h p\tau F_u\cdot\frac{m_{\mathbf{e},u}}{Z_u x(\delta^\uparrow(u))}. \tag{7.13}$$

Second, if $x(\delta^\uparrow(v))\ge\epsilon_F$, applying (7.5) and (7.6) to $I_{\mathbf{e},v}$, and using $Z_v\ge 1$, we get

$$\begin{aligned}\mathbb{E}[I_{\mathbf{e},v}]&\le\frac{m_{\mathbf{e},v}}{\sum_{\mathbf{f}\in\delta^\rightarrow(v)}m_{\mathbf{f},v}}p\tau\left(1-\frac{\epsilon_{1/1}}{5}\right)x(\delta^\uparrow(v))\\F_v&=m_{\mathbf{e},v}p\tau\left(1-\frac{\epsilon_{1/1}}{5}\right)F_v.\end{aligned} \tag{7.14}$$

Case 1: $|\mathcal{A}(S)|=3$, **where** $\mathcal{A}(S)=\{u,v,w\}$. Let $\mathbf{f}=(u,w)$, $\mathbf{g}=(v,w)$ (and of course $\mathbf{e}=(u,v)$). We will use the following facts:



$$\begin{aligned}&x_{\mathbf{e}}+x_{\mathbf{f}}\ge 2-\epsilon_F &&(x(\delta(u))\ge 2\text{ and }x(\delta^\uparrow(u))\le\epsilon_F)\\&x(\delta^\uparrow(v))+x(\delta^\uparrow(w))\ge 2-\epsilon_F &&(x(\delta(S))\ge 2)\\&x_{\mathbf{f}},x(\delta^\uparrow(w))\le 1+\epsilon_\eta, &&(\text{Lemma 2.3}),\end{aligned}$$

so we have

$$x_{\mathbf{e}},x(\delta^\uparrow(v))\ge 1-\epsilon_F-\epsilon_\eta. \tag{7.15}$$

Now we bound $\mathbb{E}[I_{\mathbf{e},u}]+\mathbb{E}[I_{\mathbf{e},v}]$. By Equation (7.13) and Equation (7.14) (which we may apply to $\mathbb{E}[I_{\mathbf{e},v}]$ because $x(\delta^\uparrow(v))\ge\epsilon_F$),

$$\begin{aligned}&\mathbb{E}[I_{\mathbf{e},u}]+\mathbb{E}[I_{\mathbf{e},v}]\\&\le\sum_{h\in\delta^\uparrow(u)}x_h p\tau F_u\cdot\frac{m_{\mathbf{e},u}}{Z_u x(\delta^\uparrow(u))}+p\tau\left(1-\frac{\epsilon_{1/1}}{5}\right)F_v m_{\mathbf{e},v}\\&=p\tau F_u m_{\mathbf{e},u}+p\tau\left(1-\frac{\epsilon_{1/1}}{5}\right)F_v m_{\mathbf{e},v}\\&\qquad\qquad(Z_u=1\ \text{as}\,|\mathcal{A}(S)|=3)\\&=p\tau(F_u m_{\mathbf{e},u}+F_v m_{\mathbf{e},v})-\frac{\epsilon_{1/1}}{5}p\tau F_v m_{\mathbf{e},v}\\&\le p\tau(1+2\epsilon_\eta)x_{\mathbf{e}}-\frac{\epsilon_{1/1}}{5}p\tau F_v m_{\mathbf{e},v},\end{aligned} \tag{7.16}$$

where the final inequality follows from (6.1). To complete the proof, we lower bound $m_{\mathbf{e},v}$.

Using (6.2) for v and w, we can write

$$\begin{aligned}&x(\delta^\uparrow(v))+x(\delta^\uparrow(w))\\&=m_{\mathbf{e},v}+m_{\mathbf{g},v}+m_{\mathbf{f},w}+m_{\mathbf{g},w}\\&\le m_{\mathbf{e},v}+\frac{(1+2\epsilon_\eta)}{(1-\epsilon_B)}(x_{\mathbf{f}}+x_{\mathbf{g}}))\\&=m_{\mathbf{e},v}+\frac{(1+2\epsilon_\eta)}{(1-\epsilon_B)}\left(\sum_{a\in\mathcal{A}(S)}\frac{x(\delta(a))}{2}-\frac{x(\delta(S))}{2}-x_{\mathbf{e}}\right)\\&\le m_{\mathbf{e},v}+\frac{(1+2\epsilon_\eta)}{(1-\epsilon_B)}(2+3\epsilon_\eta-x_{\mathbf{e}}) \qquad (\text{using (6.1)}),\end{aligned}$$

and using the fact that $x(\delta^\uparrow(v))+x(\delta^\uparrow(w))\ge 2-\epsilon_F$, we get

$$m_{\mathbf{e},v}\ge x_{\mathbf{e}}-\epsilon_F-4\epsilon_B\ge(1-1.2\epsilon_F)x_{\mathbf{e}},$$

where the second inequality follows from (7.15) and $\epsilon_B = 21\epsilon_{1/2}$ and $\epsilon_\eta < \epsilon_{1/2}^2$ and $\epsilon_F \geq 1/10$. Plugging this back into (7.16) and using $F_v \geq 1 - \epsilon_B = 1 - 21\epsilon_{1/2}$, we get

$$\begin{aligned}&\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}]\\ &\leq p\tau x_{\mathbf{e}}\left(1 + 2\epsilon_\eta - \frac{\epsilon_{1/1}}{5}(1 - 1.2\epsilon_F)(1 - 21\epsilon_{1/2})\right)\\ &\leq p\tau x_{\mathbf{e}}\left(1 - \frac{\epsilon_{1/1}}{6}\right)\end{aligned}$$

as desired. In the last inequality, we used $\epsilon_F \leq 1/10$ and $\epsilon_{1/2} \leq 0.0002$.

Case 2: $|S| \geq 4$. In this case, $Z_u = 2$. Therefore, by Equation (7.13),

$$\mathbb{E}[I_{\mathbf{e},u}] \leq \sum_{e \in \delta^\uparrow(u)} x_e p\tau F_u \frac{m_{\mathbf{e},u}}{Z_u x(\delta^\uparrow(u))} = \frac{1}{2} p\tau F_u m_{\mathbf{e},u}.$$

If $x(\delta^\uparrow(v)) < \epsilon_F$, we get the same inequality for $I_{\mathbf{e},v}$. Then,

$$\begin{aligned}\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] &\leq \frac{1}{2}p\tau(F_u m_{\mathbf{e},u} + F_v m_{\mathbf{e},v})\\ &\underset{(6.1)}{\leq} \frac{1}{2}p\tau x_e(1 + 2\epsilon_\eta),\end{aligned}$$

which is clearly sufficient for the lemma statement.

Otherwise, $x(\delta^\uparrow(v)) \geq \epsilon_F$, in which case by (7.14), we get $\mathbb{E}[I_{\mathbf{e},v}] \leq m_{\mathbf{e},v} p\tau F_v (1 - \epsilon_{1/1}/5)$. We conclude the lemma similar to the previous case. □

7.4. Increase for Bottom Edges

The following lemma is the main result of this section.

Lemma 7.4 (Bottom Edge Increase). *If $\epsilon_{1/2} \leq 0.0002$, $\epsilon_\eta \leq \epsilon_{1/2}^2$, for any polygon cut $S \in \mathcal{H}$,*

$$\mathbb{E}[I_S] \leq 0.99994\beta p.$$

For a set of edges $D \subseteq \delta(S)$ define the random variable:

$$\begin{aligned}I_S(D) := (1 + \epsilon_\eta)(\max\{&r(A \cap D)\mathbb{I}\{S \text{ not left happy}\},\\ &r(B \cap D)\mathbb{I}\{S \text{ not right happy}\}\}\\ &+ r(C \cap D)\mathbb{I}\{S \text{ not happy}\}).\end{aligned} \quad (7.17)$$

By definition $I_S(\delta(S)) = I_S$ and for any two disjoint sets D_1, D_2, $I_S(D_1 \cup D_2) \leq I_S(D_1) + I_S(D_2)$. Also, define $I_S^\uparrow = I_S(\delta^\uparrow(S))$ and $I_S^\rightarrow = I_S(\delta^\rightarrow(S))$.

First, we upper bound $\mathbb{E}[I_S^\uparrow]$. Let $f \in \delta^\uparrow(S)$ and suppose that f with $\mathsf{p}(f) = S'$ is a bottom edge. Say we have $f \in A^\uparrow(S)$ ($f \in B^\uparrow(S)$ is similar). We write

$$\begin{aligned}\mathbb{E}[I_S(f)] &= (1 + \epsilon_\eta)\beta x_f \mathbb{P}[\mathcal{R}_{S'}]\mathbb{P}[S \text{ not left happy} | \mathcal{R}_{S'}]\\ &\leq 0.568 x_f p\beta \leq x_f p\tau,\end{aligned}$$

where in the inequality we used Corollary 5.4 and that

$$\mathbb{P}[S \text{ not left happy} | \mathcal{R}_S] = \mathbb{P}[S \text{ not left happy} | \mathcal{E}_S]$$

because $\mathcal{R}_S$ is a uniformly random subset of $\mathcal{E}_S$. If $f \in C^\uparrow(S)$, we use the trivial guarantee $\mathbb{E}[I_S(f)] \leq (1 + \epsilon_\eta) x_f p\beta$.

On the other hand, if f is a top edge, then we use the trivial bound

$$\mathbb{E}[I_S(f)] \leq (1 + \epsilon_\eta)\tau p x_f. \quad (7.18)$$

Therefore,

$$\begin{aligned}\mathbb{E}[I_S^\uparrow] &\leq (1 + \epsilon_\eta)\tau p x(\delta^\uparrow(S)) + (1 + \epsilon_\eta)\epsilon_\eta p\beta\\ &\leq (1 + \epsilon_\eta)(0.571)\beta p x(\delta^\uparrow(S)) + 2\epsilon_\eta p\beta\end{aligned} \quad (7.19)$$

because $x(C) \leq \epsilon_\eta$.

Now, we consider three cases.

Case 1: $\hat{S} = \mathsf{p}(S)$ **is a degree cut.** Combining (7.19) and Lemma 7.5, we get

$$\begin{aligned}\mathbb{E}[I_S] &\leq (1 + \epsilon_\eta)p(0.571)\beta(7/4 + 6\epsilon_{1/2} + \epsilon_\eta) + 2\epsilon_\eta p\beta\\ &\leq 0.99994\beta p\end{aligned}$$

using $\epsilon_{1/2} \leq 0.0002$ and $\epsilon_\eta \leq \epsilon_{1/2}^2$.

Case 2: $\hat{S} = \mathsf{p}(S)$ **is a polygon cut with ordering** $u_1, \ldots, u_k$ **of** $\mathcal{A}(\hat{S})$, $S = u_1$ or $S = u_k$ Then, by Lemma 7.6,

$$\begin{aligned}\mathbb{E}[I_S] &\leq (1 + \epsilon_\eta)\beta p(0.571 x(\delta^\uparrow(S)) + 0.31) + 2\epsilon_\eta p\beta\\ &\leq 0.89\beta p,\end{aligned}$$

where we used $x(\delta^\uparrow(S)) \leq 1 + \epsilon_\eta$.

Case 3: $\hat{S} = \mathsf{p}(S)$ **is a polygon cut with ordering** $u_1, \ldots, u_k$ **of** $\mathcal{A}(\hat{S})$, $S \neq u_1, u_k$ Then, by Lemma 7.8,

$$\mathbb{E}[I_S] \leq (1 + \epsilon_\eta)\beta p(0.571 x(\delta^\uparrow(S)) + 0.85) + 2\epsilon_\eta p\beta \leq 0.86\beta p,$$

where we use that $x(\delta^\uparrow(S)) \leq \epsilon_\eta$ because we have a hierarchy. This concludes the proof.

7.4.1. Case 1: $\hat{S}$ is a Degree Cut

Lemma 7.5. *Let $S \in \mathcal{H}$ be a polygon cut with parent $\hat{S}$ which is a degree cut. Then*

$$\mathbb{E}[I_S^\rightarrow] \leq (1 + \epsilon_\eta)p\tau(x(\delta^\rightarrow(S)) - (1/4 - 6\epsilon_{1/2})).$$

Let A, B, C be the polygon partition of S. We will show that for a constant fraction of the edges in $\delta^\rightarrow(S)$, we can improve over the trivial bound in (7.18). To this end, consider the cases given by Theorem 5.2.

Case 1: **There is a bad half edge e in** $\delta^\rightarrow(S)$. Because bad edges never decrease, no corresponding increase occurs, so by the trivial bound Equation (7.18)

$$\mathbb{E}[I_S^\rightarrow] \leq (1 + \epsilon_\eta)p\tau(x(\delta^\rightarrow(S)) - (1/2 - \epsilon_{1/2})).$$

This concludes the proof.

Case 2: **There is a set of 2-1-1 good edges (w.r.t., S)** $D \subseteq \delta^\rightarrow(S)$**, such that** $x_D \geq 1/2 - \epsilon_{1/2} - \epsilon_\eta$. For any (top) edge $e \in \mathbf{f} = (S, u)$ such that $e \in D$, if $\mathcal{R}_{\mathbf{f},S}$, then S is happy, that is $A_T = B_T = 1, C_T = 0$ by Remark 5.1. Therefore,

$$\begin{aligned}\mathbb{E}[I_S(D)] &\leq \sum_{e \in D: e \in \mathbf{f} = (S,u)} \frac{1 + \epsilon_\eta}{2}\tau x_e \mathbb{P}[S \text{ not happy} | \mathcal{R}_{\mathbf{f},u}]\mathbb{P}[\mathcal{R}_{\mathbf{f},u}]\\ &\leq \frac{1 + \epsilon_\eta}{2} p\tau x(D).\end{aligned}$$

Using the trivial inequality Equation (7.18) for edges in $\delta^{\rightarrow}(S) \setminus D$, we get

$$\begin{aligned}\mathbb{E}[I_S^{\rightarrow}] &\leq (1+\epsilon_\eta)p\tau\left(\frac{x(D)}{2} + x(\delta^{\rightarrow}(S)) - x(D)\right)\\ &\leq (1+\epsilon_\eta)p\tau(x(\delta^{\rightarrow}(S)) - (1/4 - \epsilon_{1/2}))\end{aligned}$$

as desired. In the last inequality, we used $x(D) \geq 1/2 - \epsilon_{1/2} - \epsilon_\eta$.

Case 3: **Cases 1 and 2 do not hold.** Therefore, by Theorem 5.2 there are least two 2-2-2 good top half edge bundles. In this case, S has chosen a fixed pair of 2-2-2 good edges $\mathbf{e} = (S, v)$, $\mathbf{f} = (S, w)$ in $\delta^{\rightarrow}(S)$ (as defined in the reduction events) such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$ and $\mathcal{R}_{\mathbf{e},S} = \mathcal{R}_{\mathbf{f},S}$ with probability 1. (Recall that $\mathbf{e}(A) = \mathbf{e} \cap A$.) Let $D = \mathbf{e}(A) \cup \mathbf{f}(B)$. In this case, $\mathbf{e}$ and $\mathbf{f}$ are reduced simultaneously by τ when they are 2-2-2 happy (w.r.t., S), thatis, when $\mathcal{R}_{\mathbf{e},S} = \mathcal{R}_{\mathbf{f},S} = 1$. In such a case, we have $\delta(S)_T = \delta(v)_T = \delta(w)_T = 2$. Therefore,

$$\begin{aligned}\mathbb{E}[I_S(D)] &\leq (1+\epsilon_\eta)\mathbb{E}[\max\{r(A\cap D), r(B\cap D)\}]\\ &\leq (1+\epsilon_\eta)\frac{\tau}{2}\max\{x_{\mathbf{e}(A)}, x_{\mathbf{f}(B)}\}(\mathbb{P}[\mathcal{R}_{\mathbf{e},S} \wedge \mathcal{R}_{\mathbf{f},S}]\\ &\quad + \mathbb{P}[\mathcal{R}_{\mathbf{e},v}] + \mathbb{P}[\mathcal{R}_{\mathbf{f},w}])\\ &\leq (1+\epsilon_\eta)\tau\frac{3p}{2}x(D)\left(\frac{1}{2} + 3\epsilon_{1/2}\right)\\ &= (1+\epsilon_\eta)\tau p x(D)\left(\frac{3}{4} + 4.5\epsilon_{1/2}\right),\end{aligned}$$

where we used that $1/2 - 2\epsilon_{1/2} - x(C) \leq x_{\mathbf{e}(A)}, x_{\mathbf{f}(B)} \leq 1/2 + \epsilon_{1/2}$ and that $x(C) \leq \epsilon_\eta$. Using the trivial inequality Equation (7.18) for edges in $\delta^{\rightarrow}(S) \setminus D$, we get

$$\begin{aligned}\mathbb{E}[I_S^{\rightarrow}] &\leq (1+\epsilon_\eta)p\tau(x(D)(3/4 + 4.5\epsilon_{1/2})\\ &\quad + x(\delta^{\rightarrow}(S)) - x(D))\\ &\leq (1+\epsilon_\eta)p\tau(x(\delta^{\rightarrow}(S)) - (1/4 - 6\epsilon_{1/2})),\end{aligned}$$

where we used $x(D) \geq 1 - 4\epsilon_{1/2} - \epsilon_\eta$. □

7.4.2. Case 2: S and its Parent $\hat{S}$ are Both Polygon Cuts.

In this section, we prove two lemmas: Lemma 7.6, which bounds $\mathbb{E}[I_S^{\rightarrow}]$ when S is the leftmost or rightmost atom of $\hat{S}$, and Lemma 7.8, which bounds this quantity when S is not leftmost or rightmost.

Lemma 7.6. *Let $S \in \mathcal{H}$ be a polygon cut with $\mathsf{p}(S) = \hat{S}$ also a polygon cut. Let $u_1, \ldots, u_k$ be the ordering of cuts in $\mathcal{A}(\hat{S})$ (as defined in Definition 4.11). If $\epsilon_M \leq 0.001$, $\epsilon_\eta \leq \epsilon_M^2$, $S = u_1$ or $S = u_k$, then*

$$\mathbb{E}[I_S^{\rightarrow}] \leq 0.31\beta p.$$

Let S be the leftmost atom of $\hat{S}$ and let A, B, C be the polygon partition of $\delta(S)$. First, note

$$\begin{aligned}\mathbb{E}[I_S^{\rightarrow}] \leq (1+\epsilon_\eta)(&\mathbb{E}[\max(r(A^{\rightarrow}), r(B^{\rightarrow}))\cdot\mathbb{I}\{S \text{ not happy}\}]\\ &+ \mathbb{E}[r(C^{\rightarrow})\mathbb{I}\{S \text{ not happy}\}]).\end{aligned} \tag{7.20}$$

Here, recall that $A^{\rightarrow} = A \cap \delta^{\rightarrow}(S)$. WLOG assume $x(A^{\rightarrow}) \geq x(B^{\rightarrow})$. Then,

$$\begin{aligned}&\mathbb{E}[\max\{r(A^{\rightarrow}), r(B^{\rightarrow})\}\mathbb{I}\{S \text{ not happy}\}]\\ &\quad = \beta p x(A^{\rightarrow})\cdot\mathbb{P}[S \text{ not happy} | \mathcal{R}_{\hat{S}}]\end{aligned}$$

By Lemma 7.7 we have

$$\begin{aligned}&x(A^{\rightarrow})\cdot\mathbb{P}[S \text{ not happy}|\mathcal{R}_{\hat{S}}]\\ &\leq x(A^{\rightarrow})(1 - ((1-x(A^{\rightarrow}))^2 + (x(A^{\rightarrow}))^2 - 2\epsilon_M - 17\epsilon_\eta))\\ &\leq (2x(A^{\rightarrow})^2 - 2x(A^{\rightarrow})^3 + 2\epsilon_M x(A^{\rightarrow}) + 17\epsilon_\eta x(A^{\rightarrow}))\\ &\leq (8/27 + 2\epsilon_M + 17\epsilon_\eta),\end{aligned}$$

where in the final inequality we used that the function $x \longmapsto x^2(1-x)$ is maximized at $x = 2/3$, and using $\epsilon_M \leq 0.001$, $\epsilon_\eta < \epsilon_M^2$.

Plugging this back into (7.20), and using $x(C) \leq \epsilon_\eta$, we get

$$\mathbb{E}[I_S^{\rightarrow}] \leq (1+\epsilon_\eta)\beta p\left(\frac{8}{27} + 2\epsilon_M + 18\epsilon_\eta\right) \leq 0.31\beta p,$$

where the last inequality follows because $\epsilon_M \leq 0.001$ and $\epsilon_\eta < \epsilon_M^2$. □

Lemma 7.7. *Let $S \in \mathcal{H}$ be a polygon cut with $\mathsf{p}(S) = \hat{S}$ also a polygon cut. Let $u_1, \ldots, u_k$ be the ordering of cuts in $\mathcal{A}(\hat{S})$. If $S = u_1$, (or $S = u_k$), then*

$$\begin{aligned}&\mathbb{P}[S \text{ happy}|\mathcal{R}_{\hat{S}}]\\ &\quad \geq (1 - x(A^{\rightarrow}))^2 + (x(A^{\rightarrow}))^2 - 2\epsilon_M - 17\epsilon_\eta.\end{aligned}$$

Let $A, B, C, \hat{A}, \hat{B}, \hat{C}$ be the polygon partition of $S, \hat{S}$, respectively. Observe that because $S = u_1$, we have $\hat{A} = E(u_1, \overline{\hat{S}}) = A^{\uparrow} \cup B^{\uparrow} \cup C^{\uparrow}$ and $\hat{B}, \hat{C} \cap (A \cup B \cup C) = \emptyset$. Conditioned on $\mathcal{R}_{\hat{S}}$, $\hat{S}$ is a tree, and marginals of all edges in $\hat{A}$ is changed by a total variation distance at most $\epsilon'_M := \epsilon_M + 2\epsilon_\eta$ from x (see Corollary 5.2), and they are independent of edges inside $\hat{S}$. The tree conditioning increases marginals inside by at most $\epsilon_\eta/2$. Because after the changes just described

$$\mathbb{E}[C_T] \leq x_C + \epsilon_\eta + \epsilon'_M \leq 4\epsilon_\eta + \epsilon_M,$$

it follows that $\mathbb{P}[C_T = 0|\mathcal{R}_{\hat{S}}] \geq 1 - 4\epsilon_\eta - \epsilon_M$. Therefore,

$$\begin{aligned}&\mathbb{P}[S \text{ happy}|\mathcal{R}_{\hat{S}}]\\ &\quad \geq (1 - 4\epsilon_\eta - \epsilon_M)\mathbb{P}[A_T = B_T = 1|C_T = 0, \mathcal{R}_{\hat{S}}].\end{aligned} \tag{7.21}$$

Let ν be the conditional measure $C_T = 0$, $\mathcal{R}_{\hat{S}}$. We see that

$$\begin{aligned}\mathbb{P}_\nu[A_T = B_T = 1] &= \mathbb{P}_\nu[A_T^{\uparrow} = 1, B_T^{\uparrow} = 0, A_T^{\rightarrow} = 0, B_T^{\rightarrow} = 1]\\ &\quad + \mathbb{P}_\nu[A_T^{\uparrow} = 0, B_T^{\uparrow} = 1, A_T^{\rightarrow} = 1, B_T^{\rightarrow} = 0],\end{aligned}$$

so, using independence of $(\delta^{\uparrow}(S))_T$ and $(\delta^{\rightarrow}(S))_T$:

$$= \mathbb{P}_\nu[A_T^{\uparrow}=1, B_T^{\uparrow}=0]\mathbb{P}_\nu[A_T^{\rightarrow}=0, B_T^{\rightarrow}=1]$$
$$+ \mathbb{P}_\nu[A_T^{\uparrow}=0, B_T^{\uparrow}=1]\mathbb{P}_\nu[A_T^{\rightarrow}=1, B_T^{\rightarrow}=0]$$
$$\geq (x(A^{\uparrow}) - \epsilon'_M)\mathbb{P}_\nu[A_T^{\rightarrow}=0, B_T^{\rightarrow}=1] + (x(B^{\uparrow}) - \epsilon'_M)$$
$$\mathbb{P}_\nu[A_T^{\rightarrow}=1, B_T^{\rightarrow}=0].$$

In the final inequality, we used the fact that conditioned on $\mathcal{R}_{\hat{S}}$, $\hat{A} = (A^{\uparrow} \cup B^{\uparrow} \cup C^{\uparrow})_T = 1$ and marginals in $A^{\uparrow}$ and $B^{\uparrow}$ are approximately preserved. Now, we lower bound $\mathbb{P}_\nu[A_T^{\rightarrow}=1, B_T^{\rightarrow}=0]$. Let ϵ_A, ϵ_B be such that

$$\mathbb{E}_\nu[A_T^{\rightarrow}] = \mathbb{P}_\nu[A_T^{\rightarrow}=1, B_T^{\rightarrow}=0] + \epsilon_A,$$
$$\mathbb{E}_\nu[B_T^{\rightarrow}] = \mathbb{P}_\nu[A_T^{\rightarrow}=0, B_T^{\rightarrow}=1] + \epsilon_B.$$

First $\mathbb{P}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow} \geq 1] = 1$, and so $\mathbb{P}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow} \geq 2] \leq \mathbb{E}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow}] - 1$. Therefore,

$$\epsilon_A + \epsilon_B = \mathbb{E}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow}] - \mathbb{P}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow} = 1]$$
$$= \mathbb{E}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow}] - (1 - \mathbb{P}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow} \geq 2])$$
$$\leq 2(\mathbb{E}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow}] - 1) \leq 5\epsilon_\eta.$$

To see the last inequality, first, by Definition 4.11, $x(\delta^{\uparrow}(S)) \geq 1 - \epsilon_\eta$. Because $x(\delta(S)) \leq 2 + \epsilon_\eta$, we get that $x(\delta^{\rightarrow}(S)) \leq 1 + 2\epsilon_\eta$. Therefore,

$$\mathbb{E}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow}] \leq \mathbb{E}[\delta^{\rightarrow}(S) | \mathcal{R}_{\hat{S}}] \leq x(\delta^{\rightarrow}(S)) + \epsilon_\eta/2$$
$$\leq 1 + 2.5\epsilon_\eta.$$

Therefore,

$$\mathbb{P}_\nu[A_T = B_T = 1]$$
$$\geq (x(A^{\uparrow}) - \epsilon'_M)(\mathbb{E}_\nu[B_T^{\rightarrow}] - \epsilon_B) + (x(B^{\uparrow}) - \epsilon'_M)(\mathbb{E}_\nu[A_T^{\rightarrow}] - \epsilon_A)$$
$$\geq (x(A^{\uparrow}) - \epsilon'_M)(x(B^{\rightarrow}) - 5\epsilon_\eta) + (x(B^{\uparrow}) - \epsilon'_M)(x(A^{\rightarrow}) - 5\epsilon_\eta),$$

where the second inequality uses that the tree conditioning and $C_T^{\rightarrow} = 0$ can only increase the marginals of edges in $A^{\rightarrow}$ and $B^{\rightarrow}$. Simplify the previous expression using $x(A^{\uparrow}) + x(A^{\rightarrow}) \geq 1 - \epsilon_\eta$, and similarly for B,

$$\mathbb{P}_\nu[A_T = B_T = 1]$$
$$\geq (1 - x(A^{\rightarrow}) - \epsilon_\eta - \epsilon'_M)(x(B^{\rightarrow}) - 5\epsilon_\eta) + (1 - x(B^{\rightarrow})$$
$$- \epsilon_\eta - \epsilon'_M)(x(A^{\rightarrow}) - 5\epsilon_\eta)$$

and since $x(A^{\rightarrow}) + x(B^{\rightarrow}) \geq 1 - 2\epsilon_\eta$ (because $x(A^{\uparrow}) + x(B^{\uparrow}) \leq 1 + \epsilon_\eta$ and $x_C \leq \epsilon_\eta$), this is

$$\geq (1 - x(A^{\rightarrow}) - \epsilon_\eta - \epsilon'_M)(1 - x(A^{\rightarrow}) - 7\epsilon_\eta)$$
$$+ (x(A^{\rightarrow}) - 3\epsilon_\eta - \epsilon'_M)(x(A^{\rightarrow}) - 5\epsilon_\eta)$$
$$\geq (1 - x(A^{\rightarrow}))^2 + (x(A^{\rightarrow}))^2 - \epsilon'_M - 8\epsilon_\eta.$$

Plugging this into Equation (7.21), we obtain

$$\mathbb{P}[A_T = B_T = 1, C_T = 0 | \mathcal{R}_{\hat{S}}]$$
$$\geq (1 - 2\epsilon_\eta - \epsilon'_M)\mathbb{P}[A_T = B_T = 1 | C_T = 0, \mathcal{R}_{\hat{S}}]$$
$$\geq (1 - 2\epsilon_\eta - \epsilon'_M)((1 - x(A^{\rightarrow}))^2 + (x(A^{\rightarrow}))^2 - \epsilon'_M - 8\epsilon_\eta)$$
$$\geq (1 - x(A^{\rightarrow}))^2 + (x(A^{\rightarrow}))^2 - 2\epsilon'_M - 10\epsilon_\eta,$$

which noting $\epsilon'_M = \epsilon_M + 2\epsilon_\eta$ completes the proof of the lemma. □

Lemma 7.8. *Let $S \in \mathcal{H}$ be a polygon cut with $\mathsf{p}(S) = \hat{S}$ also a polygon cut with $u_1, \ldots, u_k$ be the ordering of cuts in $\mathcal{A}(\hat{S})$. If $S \neq u_1, u_k$, then*

$$\mathbb{E}[I_S^{\rightarrow}] \leq 0.85\beta p.$$

Let $S = u_i$ for some $2 \leq i \leq k-1$. Let A, B, C be the polygon partitioning of $\delta(u_i)$ and $\hat{A}, \hat{B}, \hat{C}$ be the polygon partition of $\hat{S}$. Because u_i is in the hierarchy $A^{\uparrow} \cup B^{\uparrow} \cup C^{\uparrow} \subseteq \hat{C}$. Therefore, conditioned on $\mathcal{R}_{\hat{S}}$, $A_T^{\uparrow} = B_T^{\uparrow} = C_T^{\uparrow} = 0$.

Once again, let ν be the conditional measure $C_T = 0$, $\mathcal{R}_{\hat{S}}$. Similar to the previous case, we will lower bound

$$\mathbb{P}[S \text{ happy} | \mathcal{R}_{\hat{S}}]$$
$$\geq (1 - 2\epsilon_\eta)\mathbb{P}[A_T^{\rightarrow} = 1, B_T^{\rightarrow} = 1, | C_T = 0, \mathcal{R}_{\hat{S}}]$$
$$= (1 - 2\epsilon_\eta)\mathbb{P}_\nu[A_T^{\rightarrow} = 1 | A_T^{\rightarrow} + B_T^{\rightarrow} = 2]\mathbb{P}_\nu[A_T^{\rightarrow} + B_T^{\rightarrow} = 2], \tag{7.22}$$

where we used $\mathbb{E}[C_T^{\rightarrow} | \mathcal{R}_{\hat{S}}] \leq 2\epsilon_\eta$ in the first inequality. Therefore, it remains to lower bound each of the two terms in the RHS.

We start with the first one. Because $x(A) \in [1 - \epsilon_\eta, 1 + \epsilon_\eta]$ and $x(A^{\uparrow}) \leq \epsilon_\eta$, we have

$$\mathbb{E}_\nu[A_T^{\rightarrow}] \in [1 - 2\epsilon_\eta, 1 + 3\epsilon_\eta].$$

The same bounds hold for $\mathbb{E}_\nu[x(B^{\rightarrow})]$.

Therefore,

$$\mathbb{P}_\nu[A_T^{\rightarrow} \geq 1], \mathbb{P}_\nu[B_T^{\rightarrow} \geq 1] \geq 1 - e^{-1+2\epsilon_\eta} \quad \text{(By Lemma 2.6)}$$

$$\mathbb{P}_\nu[A_T^{\rightarrow} \leq 1], \mathbb{P}_\nu[B_T^{\rightarrow} \leq 1] \geq 0.495 \quad \text{(By Markov)}.$$

Therefore, by Corollary 5.1 (with $\epsilon = 0.495(1 - e^{-1+2\epsilon_\eta}) \geq 0.31$), we have

$$\mathbb{P}_\nu[A_T^{\rightarrow} = 1 | A_T^{\rightarrow} + B_T^{\rightarrow} = 2] \geq 0.155.$$

By Corollary 2.3, $\mathbb{P}_\nu[E(u_{i-1}, u_i)_T = 1] \geq 1 - 4\epsilon_\eta$. Similarly, $\mathbb{P}_\nu[E(u_i, u_{i+1})_T = 1] \geq 1 - 4\epsilon_\eta$. Also,

$$\mathbb{P}_\nu[\delta^{\rightarrow}(u_i)_T - E(u_{i-1}, u_i)_T - E(u_i, u_{i+1})_T = 0] \geq 1 - 4\epsilon_\eta.$$

Therefore, by a union bound, all of these events happen simultaneously, and we get $\mathbb{P}_\nu[\delta^{\rightarrow}(u_i)_T = 2] \geq 1 - 12\epsilon_\eta$. Therefore,

$$\mathbb{P}_\nu[(A^{\rightarrow})_T = (B^{\rightarrow})_T = 1] \geq 0.155(1 - 12\epsilon_\eta) \geq 0.153.$$

Plugging this back into (7.22), we get

$$\mathbb{P}[S \text{ happy} | \mathcal{R}_{\hat{S}}] \geq 0.153(1 - 2\epsilon_\eta) \geq 0.152.$$

Plugging this in (7.20), we get

$$\begin{aligned}\mathbb{E}[I_S^{\rightarrow}] &\leq (1+\epsilon_\eta)\beta p \mathbb{P}[S \text{ not happy} | \mathcal{R}_{\hat{S}}] \\ &\quad (\max\{x(A^{\rightarrow}), x(B^{\rightarrow})\} + x(C^{\rightarrow})) \\ &\leq (1+\epsilon_\eta)\beta p(1 - 0.152)(1 + \epsilon_\eta + \epsilon_\eta) \leq 0.85\beta p\end{aligned}$$

as desired.

Acknowledgments

The authors thank the anonymous reviewers for detailed comments on the earlier versions of this manuscript and the students of the graduate seminar on this paper at the University of Bonn in 2021 for many helpful comments.

Appendix A. Proofs from Section 5

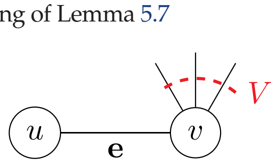
In all of the following lemmas, we assume that $\epsilon_\eta \leq \epsilon_{1/2}^2$ and $12\epsilon_{1/1} \leq \epsilon_{1/2}$.

Lemma 5.7. *Let* $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ *be a top edge bundle such that* $x_{\mathbf{e}} \leq 1/2 - \epsilon_{1/2}$. *If* $\epsilon_{1/2} \leq 0.001$, *then* $\mathbf{e}$ *is 2-1-1 happy with probability at least* $0.005\epsilon_{1/2}^2$.

Let A, B, C be the degree partitioning of $\delta(u)$. Let $V := \delta(v)_{-\mathbf{e}}$ (Figure A.1). Condition u, v be trees, $\mathbf{e}$ and C to zero, let ν be the resulting measure. This happens with probability at least 0.5 and increases marginals in $A_{-\mathbf{e}}, B_{-\mathbf{e}}, V$ by at most $x_{\mathbf{e}} + 2\epsilon_{1/1} + \epsilon_\eta \leq x_{\mathbf{e}} + 2.1\epsilon_{1/1}$ and by tree conditioning decreases marginals by at most $2\epsilon_\eta$. After conditioning, we have

$$\begin{aligned}&\mathbb{E}_\nu[A_T] \in x(A) - x_{\mathbf{e}(A)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \\ &\qquad \subset [0.5, 1.5], \text{ similarly } \mathbb{E}_\nu[B_T] \subset [0.5, 1.5] \\ &\mathbb{E}_\nu[V_T] \in x(\delta(v)) - x_{\mathbf{e}} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \subset [1.5, 2.01] \\ &\mathbb{E}_\nu[B_T + V_T] \in x(B) + x(\delta(v)) - x_{\mathbf{e}} - x_{\mathbf{e}(B)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \\ &\qquad \subset [2 + 1.8\epsilon_{1/2}, 3.01], \\ &\mathbb{E}_\nu[A_T + B_T] \in x(A) + x(B) - x_{e(A)} - x_{\mathbf{e}(B)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \\ &\qquad \subset [1.5, 2.01], \\ &\mathbb{E}_\nu[A_T + B_T + V_T] \in x(A) + x(B) + x(\delta(v)) - x_{\mathbf{e}} - x_{e(A)} - x_{\mathbf{e}(B)} \\ &\qquad + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \subset [3 + 1.75\epsilon_{1/2}, 4.01].\end{aligned}$$

Figure A.1. Setting of Lemma 5.7



Here, we used $\epsilon_{1/2} \leq 0.001$ and $12\epsilon_{1/1} < \epsilon_{1/2}$ and $x_{e(A)}, x_{\mathbf{e}(B)}, x_{e(A)} + x_{\mathbf{e}(B)} \leq x_{\mathbf{e}} \leq 1/2 - \epsilon_{1/2}$. It immediately follows from Proposition 5.1 that $\mathbb{P}_\nu[A_T = B_T = 1, V_T = 2]$ is at least a constant. In the rest of the proof, we do a more refined analysis. Using $A_T + B_T \geq 1, V_T \geq 1$,

$$\mathbb{P}_\nu[A_T + B_T + V_T = 4] \geq (1.75\epsilon_{1/2})e^{-1.75\epsilon_{1/2}} \geq 1.7\epsilon_{1/2} \qquad \text{(By Lemma 2.5)},$$

$$\mathbb{P}_\nu[A_T + B_T \geq 2], \mathbb{P}_\nu[V_T \geq 2] \geq 0.39 \qquad \text{(By Lemma 2.6)},$$

$$\mathbb{P}_\nu[A_T + B_T \leq 2], \mathbb{P}_\nu[V_T \leq 2] \geq 0.5 \qquad (\text{Markov, } A_T + B_T \geq 1, V_T \geq 1 \text{ under } \nu),$$

$$\mathbb{P}_\nu[A_T \leq 1] \geq 0.25, \mathbb{P}_\nu[B_T + V_T \leq 3] \geq 0.33 \qquad (\text{Markov's Inequality and } V_T \geq 1 \text{ under } \nu).$$

$$\mathbb{P}_\nu[A_T \geq 1] \geq 0.39, \mathbb{P}_\nu[B_T + V_T \geq 3] \geq 1.75\epsilon_{1/2} \quad \text{(By Lemma 2.6)},$$

It follows by Corollary 5.1 (with $\epsilon = 0.195, p_m \geq 1 - 2\epsilon \geq 0.6$) that

$$\mathbb{P}_\nu[V_T = 2 | A_T + B_T + V_T = 4] \geq 0.13.$$

Because $V_T \geq 1, A_T + B_T \geq 1$ with probability 1, we apply Corollary 5.1 to $V_T - 1, A_T + B_T - 1$.

Furthermore, by Lemma 5.2, $\mathbb{P}_\nu[A_T \geq 1 | A_T + B_T + V_T = 4] \geq 0.128$, $\mathbb{P}_\nu[A_T \leq 1 | A_T + B_T + V_T = 4] \geq 0.43\epsilon_{1/2}$. The same holds for B_T. Therefore, by Corollary 5.1 (with $\epsilon = 0.055\epsilon_{1/2}$), using that $\epsilon_{1/2} < 0.001$,

$$\mathbb{P}_\nu[A_T = 1 | A_T + B_T = 2, V_T = 2] \geq 0.05\epsilon_{1/2}.$$

Putting these together, we have

$$\begin{aligned}\mathbb{P}[e \text{ 2-1-1 happy}] &\geq 0.5\mathbb{P}_\nu[A_T = B_T = 1, V_T = 2] \\ &= 0.5\mathbb{P}_\nu[A_T + B_T + V_T = 4] \\ &\quad \mathbb{P}_\nu[V_T = 2 | A_T + B_T + V_T = 4] \\ &\quad \cdot \mathbb{P}_\nu[A_T = 1 | V_T = 2, A_T + B_T = 2] \\ &\geq 0.5(1.7\epsilon_{1/2})(0.13)(0.05\epsilon_{1/2}) \geq 0.005\epsilon_{1/2}^2\end{aligned}$$

as desired.

Lemma 5.8. *Let* $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ *be a top edge bundle such that* $x_{\mathbf{e}} \geq 1/2 + \epsilon_{1/2}$. *If* $\epsilon_{1/2} \leq 0.001$, *then,* $\mathbf{e}$ *is 2-1-1 happy with respect to* u *with probability at least* $0.006\epsilon_{1/2}^2$.

Let A, B, C be the degree partitioning of the edges in $\delta(u)$, $V = \delta_{-\mathbf{e}}(v)$. Condition u, v be trees, $C_T = 0$ and $u \cup v$ to be a tree (in order). This happens with probability at least $\frac{1}{2} + \epsilon_{1/2} - 3\epsilon_\eta - 2\epsilon_{1/1} \geq 0.5$. Let ν be the resulting measure restricted to edges in A, B, V. Note that ν on edges in A, B, V is SR. This is because ν is a product of two strongly Rayleigh distribution on the following two disjoint set of edges (i) the edges between u, v and (ii) the edges in $A_{-\mathbf{e}}, B_{-\mathbf{e}}, V$.

Furthermore, observe that under ν, every set of edges in $A_{-\mathbf{e}}$, $B_{-\mathbf{e}}, V$ increases by at most $2\epsilon_{1/1} + \epsilon_\eta < 0.2\epsilon_{1/2}$ (using $12\epsilon_{1/1}$

$\le \epsilon_{1/2}$), and decreases by at most $1 - x_{\mathbf{e}} + 2\epsilon_\eta$. Therefore,

$$\begin{aligned}
&\mathbb{E}_\nu[A_T] \in x(A) + [-(1-x_{\mathbf{e}}) - 2\epsilon_\eta, 1 - x_{\mathbf{e}} + 0.2\epsilon_{1/2}] \\
&\qquad \subset [0.5, 1.5], \text{ similarly, } \mathbb{E}_\nu[B_T] \in [0.5, 1.5] \\
&\mathbb{E}_\nu[V_T] \in x(\delta(v)) - x_{\mathbf{e}} + [-(1-x_{\mathbf{e}}) - 2\epsilon_\eta, 0.2\epsilon_{1/2}] \subset [0.995, 1.5]. \\
&\mathbb{E}_\nu[A_T + B_T] \in x(A) + x(B) + 1 - x_{e(A)} - x_{\mathbf{e}(B)} \\
&\qquad + [-(1-x_{\mathbf{e}}) - 2\epsilon_\eta, 0.2\epsilon_{1/2}] \subset [1.995, 2.5], \\
&\mathbb{E}_\nu[B_T + V_T] \in x(B) + x(\delta(v)) - x_{\mathbf{e}} + [-(1-x_{\mathbf{e}}) - 2\epsilon_\eta, 1 - x_{\mathbf{e}} \\
&\qquad + 0.2\epsilon_{1/2}] \subset [1.99, 3 - 1.75\epsilon_{1/2}]. \\
&\mathbb{E}_\nu[A_T + B_T + V_T] \in x(A) + x(B) + x(\delta(v)) + 1 - x_{\mathbf{e}} - x_{e(A)} - x_{\mathbf{e}(B)} \\
&\qquad + [-(1-x_{\mathbf{e}}) - 2\epsilon_\eta, 0.2\epsilon_{1/2}] \\
&\qquad \subset [2.99, 4 - 1.75\epsilon_{1/2}].
\end{aligned}$$

Here, in the upper bound on $\mathbb{E}_\nu[A_T]$, $\mathbb{E}_\nu[B_T]$, $\mathbb{E}_\nu[B_T + V_T]$, we used that the marginals of edges in the bundle $\mathbf{e}$ can only increase by $1 - x_{\mathbf{e}}$ (in total) when conditioning $u \cup v$ to be a tree. Therefore,

$$\begin{aligned}
&\mathbb{P}_\nu[A_T + B_T + V_T = 3] \ge \epsilon_{1/2} && \text{(By Theorem 2.2)}, \\
&\mathbb{P}_\nu[A_T + B_T \ge 2] \ge 0.63, \mathbb{P}_\nu[V_T \ge 1] \ge 0.63 && \text{(By Lemma 2.6, } A_T + B_T \ge 1) \\
&\mathbb{P}_\nu[A_T + B_T \le 2] \ge 0.25, \mathbb{P}_\nu[V_T \le 1] \ge 0.25 && \text{(By Markov Inequality, } A_T + B_T \ge 1), \\
&\mathbb{P}_\nu[A_T \ge 1] \ge 0.39, \mathbb{P}_\nu[B_T + V_T \ge 2] \ge 0.59 && \text{(By Lemma 2.6)} \\
&\mathbb{P}_\nu[A_T \le 1] \ge 0.25, \mathbb{P}_\nu[B_T + V_T \le 2] \ge 1.75\epsilon_{1/2}, && \text{(By Markov, In worst case } \mathbb{P}[B_T + V_T < 2] = 0).
\end{aligned}$$

It follows by Corollary 5.1 (with $\epsilon = 0.157, p_m = 0.68$) that

$$\mathbb{P}_\nu[A_T + B_T = 2 | A_T + B_T + V_T = 3] \ge 0.12.$$

Because $A_T + B_T \ge 1$ with probability 1, we apply Corollary 5.1 to $A_T + B_T - 1, V_T$.

Furthermore, by Lemma 5.2, $\mathbb{P}_\nu[A_T \ge 1 | A_T + B_T + V_T = 3] \ge 0.68\epsilon_{1/2}$ and $\mathbb{P}_\nu[A_T \le 1 | A_T + B_T + V_T = 3] \ge 0.147$. By symmetry, the same holds for B_T. Therefore, by Corollary 5.1,

$$\mathbb{P}_\nu[A_T = 1 | A_T + B_T = 2, V_T = 1] \ge 0.09\epsilon_{1/2}.$$

Here, we used $\epsilon_{1/2} < 0.001$.

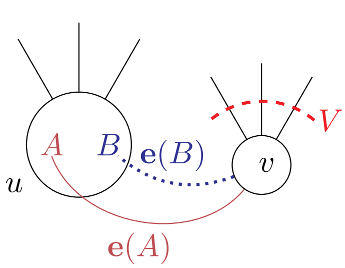
Finally,

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy}] \ge (0.09\epsilon_{1/2}) 0.12 (\epsilon_{1/2}) 0.5 \ge 0.005\epsilon_{1/2}^2,$$

as desired.

Lemma A.1. *For a good half top edge bundle* $\mathbf{e} = (\mathbf{u}, \mathbf{v})$, *let* A, B, C *be the degree partitioning of* $\delta(u)$, *and let* $V = \delta(v)_{-\mathbf{e}}$ *(Figure A.2).*

Figure A.2. Setting of Lemma A.1



If $\epsilon_{1/2} \le 0.001$, $x_{\mathbf{e}(B)} \le \epsilon_{1/2}$, *and* $\mathbb{P}[(A_{-\mathbf{e}})_T + V_T \le 1] \ge 5\epsilon_{1/2}$ *then* $\mathbf{e}$ *is 2-1-1 good,*

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 } \textit{happy w.r.t. } u] \ge 0.005\epsilon_{1/2}^2.$$

The proof is similar to Lemma 5.8. We condition u, v to be trees, $C_T = 0$, $u \cup v$ to be a tree. Let ν be the resulting SR measure on edges in A, B, V. The main difference is because $x_{\mathbf{e}} \not\ge 1/2 + \epsilon_{1/2}$, we use the lemma's assumptions to lower bound $\mathbb{P}_\nu[A_T + B_T + V_T = 3], \mathbb{P}_\nu[A_T + V_T \le 2], \mathbb{P}_\nu[B_T + V_T \le 2]$.

First, because $\mathbf{e}$ is 2-2 good, by Lemma 5.4 and negative association,

$$\begin{aligned}
\mathbb{P}_\nu[(\delta(u)_{-\mathbf{e}})_T + V_T \le 2] &\ge \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + V_T \le 2] - \mathbb{P}[C_T = 0] \\
&\ge 0.4\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta \ge 0.22\epsilon_{1/2},
\end{aligned}$$

where we used $\epsilon_{1/1} \le \epsilon_{1/2}/12$. Letting $p_i = \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + V_T = i]$, we therefore have $p_{\le 2} \ge 0.22\epsilon_{1/2}$. In addition, by Lemma 2.5, $p_3 \ge 1/4$. If $p_2 < 0.2\epsilon_{1/2}$, then from $p_2/p_3 \le 0.8\epsilon_{1/2}$, we could use log-concavity to derive a contradiction to $p_{\le 2} \ge 0.22\epsilon_{1/2}$ (analogously to what's done in the proof of Lemma 2.4). Therefore, we must have

$$\mathbb{P}_\nu[A_T + B_T + V_T = 3] = \mathbb{P}_\nu[(\delta(u)_{-\mathbf{e}})_T + V_T = 2] \ge 0.2\epsilon_{1/2}.$$

Next, because $\mathbb{P}[u, v, u \cup v \text{ trees}, C_T = 0] \ge 0.49$, by the lemma's assumption, $\mathbb{P}_\nu[\mathbf{e}(B)] \le 2.01\epsilon_{1/2}$. Therefore,

$$\mathbb{E}_\nu[B_T + V_T] \le x(V) + x(B) + 1.01\epsilon_{1/2} + 2\epsilon_{1/1} + \epsilon_\eta \le 2.51.$$

Therefore, by Markov, $\mathbb{P}_\nu[B_T + V_T \le 2] \ge 0.15$. Finally, by negative association,

$$\begin{aligned}
\mathbb{P}_\nu[A_T + V_T \le 2] &\ge \mathbb{P}_\nu[(A_{-\mathbf{e}})_T + V_T \le 1] \\
&\ge \mathbb{P}[(A_{-\mathbf{e}})_T + V_T \le 1] - \mathbb{P}[C_T = 0] \ge 4.8\epsilon_{1/2},
\end{aligned}$$

where we used the lemma's assumption.

Now, following the same line of arguments as in Lemma 5.8, we have $\mathbb{P}_\nu[A_T + B_T = 2 | A_T + B_T + V_T = 3] \ge 0.12$. Also, $\mathbb{P}_\nu[A_T \ge 1 | A - T + B_T + V_T = 3] \ge 3.02$, which implies $\mathbb{P}_\nu[A_T = 1 | A_T + B_T = 2, V_T = 1] \ge 0.42\epsilon$. This implies

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy}] \ge (0.42\epsilon_{1/2}) 0.12 (0.2\epsilon_{1/2}) 0.498 \ge 0.005\epsilon_{1/2}^2$$

as desired.

Lemma 5.9. *Let* $\mathbf{e} = (\mathbf{v}, \mathbf{u})$ *and* $\mathbf{f} = (\mathbf{v}, \mathbf{w})$ *be good half top edge bundles and let* A, B, C *be the degree partitioning of* $\delta(v)$ *such that* $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \le \epsilon_{1/2}$. *Then, one of* $\mathbf{e}, \mathbf{f}$ *is 2-1-1 happy with probability at least* $0.005\epsilon_{1/2}^2$.

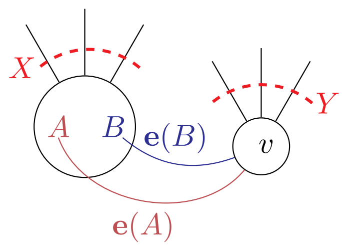
Let $U = \delta(u)_{-\mathbf{e}}$. By Lemma 2.9, we can assume, without loss of generality, that

$$\mathbb{E}[U_T | \mathbf{f} \notin T, u, v, w \text{ tree}] \le x(U_T) + 0.405 + 3\epsilon_\eta. \tag{A.1}$$

On the other hand,

$$\begin{aligned}
\mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T] &\ge \mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T | \mathbf{f} \notin T, u, v, w \text{ tree}] \\
&\quad \mathbb{P}[\mathbf{f} \notin T, u, v, w, \text{ tree}] \\
&\ge \mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T | \mathbf{f} \notin T, u, v, w \text{ tree}] 0.49\,.
\end{aligned}$$

Figure A.3. Setting of Lemma 5.10



Therefore,

$$\mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T \mid \mathbf{f} \notin T, u, v, w, \text{ tree}] \le \frac{1}{0.49} x(A_{-\mathbf{e}-\mathbf{f}}) \le \frac{1}{0.49}(4\epsilon_{1/2} + \epsilon_\eta) \le 8.2\epsilon_{1/2}. \tag{A.2}$$

Combining (A.1) and (A.2), we get $\mathbb{E}[U_T + (A_{-\mathbf{e}}) \mid \mathbf{f} \notin T, u, v, w \text{ tree}] \le 1.91$ where we used $\epsilon_{1/2} \le 0.001$. Therefore, using Lemma 2.5, we get

$$\mathbb{P}[U_T + (A_{-\mathbf{e}})_T \le 1] \ge 0.49\mathbb{P}[U_T + (A_{-\mathbf{e}})_T \le 1 \mid \mathbf{f} \notin T, u, v, w \text{ tree}] \ge 0.01.$$

Because $\epsilon_{1/2} \le 0.001$, by Lemma A.1, $\mathbf{e}$ is 2-1-1 good.

Lemma 5.10. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle and let A, B, C be the degree partitioning of $\delta(u)$ (Figure A.3). If $\epsilon_{1/2} \le 0.001$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \ge \epsilon_{1/2}$, then*

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 } \textit{happy w.r.t } u] \ge 0.02\epsilon_{1/2}^2.$$

Condition C_T to be zero, u, v and $u \cup v$ be trees. This happens with probability at least 0.49. Let ν be the resulting measure. Let $X = A_{-\mathbf{e}} \cup B_{-\mathbf{e}}, Y = \delta(v)_{-\mathbf{e}}$ Because $\mathbf{e}$ is 2-2 good by Lemma 5.4 and stochastic dominance,

$$\mathbb{P}_\nu[X_T + Y_T \le 2] \ge \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + Y_T \le 2] - \mathbb{P}[C_T = 0] \ge 0.4\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta \ge 0.22\epsilon_{1/2},$$

where we used $\epsilon_{1/1} < 12\epsilon_{1/2}$. It follows by log-concavity of $X_T + Y_T$ that $\mathbb{P}_\nu[X_T + Y_T = 2] \ge 0.2\epsilon_{1/2}$. Now,

$$\mathbb{E}_\nu[X_T], \mathbb{E}_\nu[Y_T] \in [1 - 3\epsilon_{1/1}, 1.5 + \epsilon_{1/2} + 2\epsilon_{1/1} + 3\epsilon_\eta] \subset [0.995, 1.51].$$

Therefore,

$$\mathbb{P}_\nu[X_T \ge 1], \mathbb{P}_\nu[Y_T \ge 1] \ge 0.63 \qquad \text{(Lemma 2.6)},$$

$$\mathbb{P}_\nu[X_T \le 1], \mathbb{P}_\nu[Y_T \le 1] \ge 0.245 \qquad \text{(Markov)}.$$

Therefore, by Corollary 5.1, $\mathbb{P}_\nu[X_T = 1 \mid X_T + Y_T = 2] \ge 0.119$:

$$\mathbb{P}_\nu[X_T = Y_T = 1] \ge (0.2\epsilon_{1/2})0.119 \ge 0.023\epsilon_{1/2},$$

and let $\mathcal{E}$ be the event $\{X_T = Y_T = 1 \mid \nu\}$. In ν, we always choose exactly one edge from the $\mathbf{e}$ bundle, and that is independent of edges in X, Y, in particular the above event. Therefore, we can correct the parity of A, B by choosing from e_A *or* e_B. It follows that

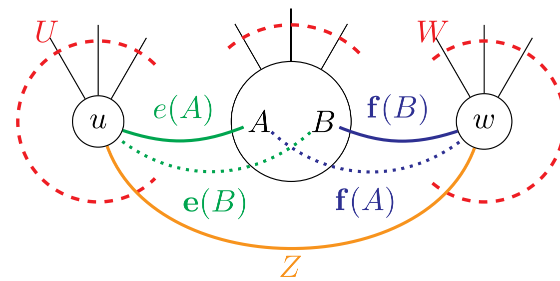
$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t } u] \ge \mathbb{P}_\nu[\mathcal{E}](1.99\epsilon_{1/2})0.49 \ge 0.02\epsilon_{1/2}^2,$$

where we used that $\mathbb{E}_\nu[\mathbf{e}(A)_T] \ge 1.99\epsilon_{1/2}$, and the same fact for $\mathbf{e}(B)_T$. To see why this latter fact is true, observe that conditioned on u, v trees, we always sample at most one edge between u, v. Therefore, because under ν, we choose exactly one edge between u, v, the probability of choosing from $e(A)$ (and similarly choosing from $\mathbf{e}(B)$) is at least

$$\frac{\mathbb{E}[\mathbf{e}(A)_T \mid u, v \text{ trees}, C_T = 0]}{\mathbb{P}[\mathbf{e} \mid u, v \text{ trees}, C_T = 0]} \ge \frac{x_{\mathbf{e}(A)} - 2\epsilon_\eta}{x_{\mathbf{e}} + 3\epsilon_{1/1}} \ge \frac{\epsilon_{1/2} - 2\epsilon_\eta}{1/2 + 1.3\epsilon_{1/2}} \ge 1.99\epsilon_{1/2}$$

as desired.

Figure A.4. Setting of Lemma 5.12



Notes. We assume that the dotted green/blue edges are at most $\epsilon_{1/2}$. The edges of C are not shown.

Lemma 5.12. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be two good top half edge bundles and let A, B, C be degree partitioning of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \le \epsilon_{1/2}$. If $\mathbf{e}, \mathbf{f}$ are not 2-1-1 good with respect to v, and $\epsilon_{1/2} \le 0.0002$, then $\mathbf{e}, \mathbf{f}$ are 2-2-2 happy with probability at least* 0.01.

See Figure A.4 for the setting of this lemma. First, observe that by Lemma A.1 if $\mathbb{P}[U_T + (A_{-\mathbf{e}})_T \le 1] \ge 0.25\epsilon$, where $\epsilon \ge 20\epsilon_{1/2}$ is a constant that we fix later, then $\mathbf{e}$ is 2-1-1 good, which is a contradiction. Therefore, assume $\mathbb{P}[U_T + (A_{-\mathbf{e}})_T \ge 2] \ge 1 - 0.25\epsilon$. Furthermore, let $q = \mathbb{P}[U_T + (A_{-\mathbf{e}})_T \ge 3]$. Because $x(U) + x(A_{-\mathbf{e}}) \le 2 + 3\epsilon_{1/2} + 2\epsilon_{1/1} + 3\epsilon_\eta \le 2 + 3.2\epsilon_{1/2}$ (where we used $x_{\mathbf{e}(A)} \ge x_{\mathbf{e}} - x_{\mathbf{e}(B)} - x_C \ge 1/2 - 2\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta$ and where we used $12\epsilon_{1/1} \le \epsilon_{1/2}$),

$$2(1 - q - 0.25\epsilon) + 3q \le 2 + 3.2\epsilon_{1/2}.$$

This implies that $q \le 0.5\epsilon + 3.2\epsilon_{1/2} \le 0.75\epsilon$ (for $\epsilon \ge 13\epsilon_{1/2}$). Therefore,

$$\mathbb{P}[U_T + (A_{-\mathbf{e}})_T = 2], \mathbb{P}[W_T + (B_{-\mathbf{f}})_T = 2] \ge 1 - \epsilon, \tag{A.3}$$

where the second inequality follows by a similar argument.

Claim A.1. *Let $Z = \delta(u) \cap \delta(w)$. If $\epsilon < 1/15$, then either $\mathbb{E}[Z \mid u, v, w \text{ tree}] \le 3\epsilon$ or $\mathbb{E}[Z \mid u, v, w \text{ tree}] \ge (1 - 3\epsilon)$.*

For the whole proof we work with μ conditioned on u, v, w are trees. Let $z = \mathbb{E}[Z]$. Let $D = U \cup W \cup A_{-\mathbf{e}} \cup B_{-\mathbf{f}} \smallsetminus Z$. Note that $D_T + 2Z_T = U_T \cup W_T \cup (A_{-\mathbf{e}})_T \cup (B_{-\mathbf{f}})_T$. By Equation (A.3) and a union bound $\mathbb{P}[D_T + 2Z_T = 4] \ge 1 - 2\epsilon - 3\epsilon_\eta$. Therefore,

$$2.1\epsilon \ge 2\epsilon + 3\epsilon_\eta \ge \mathbb{P}[D_T + 2Z_T \ne 4] \ge \mathbb{P}[D_T = 3] \ge \sqrt{\mathbb{P}[D_T = 2]\mathbb{P}[D_T = 4]},$$

where the last inequality follows by log-concavity. On the other hand,

$$z = \mathbb{P}[Z = 1] \le \mathbb{P}[D_T = 2, Z = 1] + \mathbb{P}[D_T + 2Z_T \ne 4] \le \mathbb{P}[D_T = 2] + 2.1\epsilon,$$

$$1 - z = \mathbb{P}[Z = 0] \le \mathbb{P}[D_T = 4, Z = 0] + \mathbb{P}[D_T + 2Z_T \ne 4] \le \mathbb{P}[D_T = 4] + 2.1\epsilon.$$

Putting everything together,

$$(2.1\epsilon)^2 \geq (z - 2.1\epsilon)(1 - z - 2.1\epsilon) = z(1-z) - 2.1\epsilon + 2.1\epsilon^2.$$

Therefore, using $\epsilon \leq 1/15$, we get that either $z \leq 3\epsilon$ or $z \geq 1 - 3\epsilon$. □

Therefore, for the rest of proof, we assume $\mathbb{E}[Z_T | u, v, w \text{ trees}] < 3\epsilon$. A similar proof shows $\mathbf{e}, \mathbf{f}$ are 2-2-2 good when $\mathbb{E}[Z_T | u, v, w \text{ trees}] > 1 - 3\epsilon$. We run the following conditionings in order: u, v, w trees, $Z_T = 0$, $C_T = 0$, $\mathbf{e}(B), \mathbf{f} \notin T, \mathbf{e}(A) \in T$. Note that $\mathbf{e}(A) \in T$ is equivalent to $u \cup v$ be a tree. Call this event $\mathcal{E}$ (i.e., the event that all things we conditioned on happen). First, notice

$$\mathbb{P}[\mathcal{E}] \geq (1 - 3\epsilon_\eta)(1 - 3\epsilon - 2\epsilon_{1/1} - \epsilon_\eta - \epsilon_{1/2} - (1/2 + \epsilon_{1/2}))(1/2 - 3\epsilon_{1/2}) \geq 0.22 \geq 1/5. \quad (A.4)$$

Moreover, because all these conditionings correspond to upward/downward events, $\mu | \mathcal{E}$ is strongly Rayleigh. The main statement we will show is that

$$\mathbb{P}[\mathbf{e}, \mathbf{f} \text{ 2-2-2 happy} | \mathcal{E}] \geq \mathbb{P}[U_T = (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0, W_T = 2 | \mathcal{E}] = \Omega(1).$$

The main insight of the proof is that Equation (A.3) holds (up to a larger constant of ϵ), even after conditioning $\mathcal{E}, B_{-\mathbf{f}} = 0$, $A_{-\mathbf{e}} = 1$; therefore, we can bound the preceding event by just a union bound. The main nontrivial statement is to argue thatthe expectations of $B_{-\mathbf{f}}$ and $A_{-\mathbf{e}}$ do not change so much under $\mathcal{E}$.

Combining (A.3) and (A.4),

$$\mathbb{P}[U_T + (A_{-\mathbf{e}})_T = 2 | \mathcal{E}], \mathbb{P}[W_T + (B_{-\mathbf{f}})_T = 2 | \mathcal{E}] \geq 1 - 5\epsilon. \quad (A.5)$$

We claim that

$$\mathbb{E}[B_T | \mathcal{E}] = \mathbb{E}[(B_{-\mathbf{f}})_T | \mathcal{E}] \leq x(B_{-\mathbf{f}}) + 3\epsilon_\eta + 3\epsilon_{1/1} + \epsilon_{1/2} + 35\epsilon \leq 0.66 \quad (A.6)$$

using $\epsilon_{1/2} < 0.0002$ and $\epsilon = 20\epsilon_{1/2}$. To see this, observe that after each conditioning in $\mathcal{E}$ either all marginals increase or all decrease. Furthermore, the events $C_T = 0, Z_T = 0, \mathbf{e}(B)_T = 0$ can increase marginals by at most $3\epsilon_\eta + 3\epsilon_{1/1} + \epsilon_{1/2}$; the only other event that can increase $B_{-\mathbf{f}}$ is $\mathbf{f} \notin T$. Now we know $\mathbb{P}[(B_{-\mathbf{f}})_T + W_T = 2 | \mathcal{E}] \geq 1 - 5\epsilon$ before and after conditioning $\mathbf{f} \notin T$. Therefore, by Corollary 2.19, $2 - 10\epsilon \leq \mathbb{E}[(B_{-\mathbf{f}})_T + W_T] \leq 2 + 25\epsilon$. However, if $\mathbb{E}[(B_{-\mathbf{f}})_T]$ increased by more than 35ϵ, then either before conditioning $\mathbf{f} \notin T$, $\mathbb{E}[(B_{-\mathbf{f}}) + W_T] < 2 - 10\epsilon$ or afterward it is more than $2 + 25\epsilon$, which is a contradiction, and completes the proof of (A.6). A similar argument shows that $\mathbb{E}[(A_{-\mathbf{e}})_T | \mathcal{E}] \leq 0.66$.

We also claim that

$$\mathbb{E}[(A_{-e})_T | \mathcal{E}] \geq x(A_{-e}) - 3\epsilon_\eta - 35\epsilon \geq 0.33.$$

As previously, everything conditioned on $\mathcal{E}$ increases $\mathbb{E}[(A_{-e})_T]$ except for possibly $\mathbf{e}(A) \in T$. As previously, we know that $\mathbb{P}[U_T + (A_{-e})_T = 2 | \mathcal{E}] \geq 1 - 5\epsilon$ before and after $\mathbf{e}(A) \notin T$. Therefore, again applying Corollary 2.19, we see that it cannot decrease by more than 35ϵ.

It follows that

$$0.33 \leq \mathbb{E}[(A_{-\mathbf{e}})_T | \mathcal{E}] \leq \mathbb{E}[(A_{-\mathbf{e}})_T | \mathcal{E}, (B_{-\mathbf{f}})_T = 0] \leq 0.66 + 0.66 \leq 1.32.$$

Therefore, by Lemma 2.5 and Theorem 2.2, $\mathbb{P}[(A_{-\mathbf{e}})_T = 1 | \mathcal{E}, (B_{-f})_T = 0] \geq 0.33e^{-.33} \geq 0.237$.

Therefore, by Lemma 2.5,

$$\mathbb{P}[\mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0] \geq (0.22)(0.39)(0.23) \geq 0.019.$$

Therefore, by (A.5)

$$\mathbb{P}[U_T = 1 | \mathcal{E}, (A_{-\mathbf{e}})_T$$
$$\mathbb{P}[W_T = 2 | \mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0] \geq 1 - 5\epsilon/0.019.$$

Finally, by union bound

$$\mathbb{P}[U_T = 1, W_T = 2 | \mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0] \geq 1 - \epsilon/0.009.$$

Using $\epsilon = 20\epsilon_{1/2}$ and $\epsilon_{1/2} \leq 0.0002$, this means both of the above events happens, so $\mathbf{e}, \mathbf{f}$ are 2-2-2-happy with probability $0.019(1 - \epsilon/0.009) > 0.01$ as desired.

Endnotes

[1] Given such an Eulerian cycle, we can use the triangle inequality to shortcut vertices visited more than once to get a Hamiltonian cycle.

[2] Although, because we do not always find a distribution that preserves marginals exactly (as one does not necessarily exist), for precision we generally refer to the distribution used by the algorithm as λ-uniform instead.

[3] The standard name for this is the T-join polytope. Because we reserve T to represent our tree, we call this the O-join polytope, where O represents the set of odd vertices in the tree.

[4] Where the randomness comes from the random sampling of the tree.

[5] Recall that we merely need to prove the *existence* of a cheap O-join solution. The actual optimal O-join solution can be found in polynomial time.

[6] This is really a family of near-min cuts, but for the purpose of this overview, assume $\eta = 0$.

[7] Roughly, this corresponds to the definition of the polygon being left-happy.

[8] For example, in Figure 2, $\mathsf{p}(a, c) = u_3$, and (a, c) is a bottom edge.

[9] Some portions of this discussion might be easier to understand after reading the rest of the paper.

[10] Each cut will be mapped to one or two OPT edges.

[11] In the special case that $i = 2$, L in (i) will be the leftmost atom if it is a near min cut, and similarly in (ii) when $i = m - 1$, R will be the rightmost atom if it is a near min cut.

[12] In the sense of the number of vertices that it contains.

[13] An atom may already correspond to a connected component; in such a case, we do not add it in this step.

[14] Think about such set as a *degenerate* polygon with atoms $a_1 := X, a_2 := Y, a_0 := \overline{X \cup Y}$. Therefore, for the rest of this section, we call them triangles and in later section we just think of them as polygon cuts.

[15] To be precise, we apply the previous corollary to the polarization of q'_μ, where x, y are polarized by a disjoint set of variables of size equal to their maximum degree.

[16] The main problem with bad edges is that we cannot match them to edges going higher in the matching Lemma 6.1. Therefore, to prove the matching lemma we need to justify that there are not too many bad edges in any cut. Therefore, we cannot simply "pretend" that one half edge bundle of $\delta(u)$ is bad.

[17] Suppose that under the distribution μ on spanning trees, some event $\mathcal{D}'$ has probability $q \geq p$ and we seek to define an event $\mathcal{D} \subseteq$

$\mathcal{D}'$ that has probability *exactly* p. To this end, one can copy every tree T in the support of μ, exactly $\lfloor \frac{kq}{p} \rfloor$ times for some integer $k > 0$, and whenever we sample T, we choose a copy uniformly at random. Therefore, to get a probability exactly p for an event, we say this event occurs if for a "feasible" tree T one of the first k copies are sampled. Now, as $k \to \infty$ the probability that $\mathcal{D}$ occurs converges to p. Now, for a number of decreasing events, $\mathcal{D}_1, \mathcal{D}_2, \ldots,$ that occur with probabilities $q_1, q_2, \ldots$ (respectively), we just need to let k be the least common multiple of $p/q_1, p/q_2, \ldots$ and follow the above procedure. Another method is to choose an independent Bernoulli with success probability p / q for any such event $\mathcal{D}$.

[18] We are using the fact that $\epsilon_{1/1} = \epsilon_{1/2}/12$ and that ϵ_η is tiny by comparison with these.

References

Anari N, Oveis Gharan S, Vinzant C (2018) Log-concave polynomials, entropy, and a deterministic approximation algorithm for counting bases of matroids. Thorup M, ed. *Sympos. Foundations Comput. Sci.* (IEEE Computer Society, Washington, DC), 35–46.

Applegate DL, Bixby RE, Chvatal V, Cook WJ (2007) *The Traveling Salesman Problem: A Computational Study (Princeton Series in Applied Mathematics)* (Princeton University Press, Princeton, NJ).

Arora S (1998) Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. *J. Assoc. Comput. Machinery* 45(5):753–782.

Arora S, Grigni M, Karger D, Klein P, Woloszyn A (1998) A polynomial-time approximation scheme for weighted planar graph tsp. *Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia), 33–41.

Asadpour A, Goemans MX, Madry A, Oveis Gharan S, Saberi A (2017) An $O(\log n/\log\log n)$-approximation algorithm for the asymmetric traveling salesman problem. *Oper. Res.* 65(4):1043–1061.

Benczúr AA (1995) A representation of cuts within 6/5 times the edge connectivity with applications. *Sympos. Foundations Comput. Sci.* (IEEE Computer Society, Piscataway, NJ), 92–102.

Benczúr AA (1997) Cut structures and randomized algorithms in edge-connectivity problems. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA.

Benczúr AA, Goemans MX (2008) Deformable polygon representation and near-mincuts. Groetschel M, Katona GOH, eds. *Building Bridges: Between Mathematics and Computer Science*. Bolyai Society Mathematical Studies, vol. 19 (Springer, Berlin), 103–135.

Borcea J, Branden P, Liggett TM (2009) Negative dependence and the geometry of polynomials. *J. Amer. Math. Soc.* 22:521–567.

Boyd S, Carr R (2011) Finding low cost tsp and 2-matching solutions using certain half-integer subtour vertices. *Discrete Optim.* 8(4):525–539.

Boyd S, Elliott-Magwood P (2010) Structure of the extreme points of the subtour elimination polytope of the stsp. *Combin. Optim. Discrete Algorithms* B23:33–47.

Boyd SC, Pulleyblank WR (1991) Optimizing over the subtour polytope of the travelling salesman problem. *Math. Programming* 49:163–187.

Carr RD, Vempala S (2000) Toward a 4/3 approximation for the asymmetric traveling salesman problem. *Sympos. Discrete Algorithms* (Society for Industrial and Applied Mathematics, Philadelphia), 116–125.

Christofides N (1976) Worst case analysis of a new heuristic for the traveling salesman problem. Report 388, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA.

Dantzig G, Fulkerson D, Johnson S (1959) On a linear programming combinatorial approach to the traveling salesman problem. *Oper. Res.* 7:58–66.

Darroch JN (1964) On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* 36:1317–1321.

Demaine ED, Hajiaghayi M, Mohar B (2010) Approximation algorithms via contraction decomposition. *Combinatorica* 30:533–552.

Dinits E, Karzanov A, Lomonosov M (1976) On the structure of a family of minimal weighted cuts in graphs. Fridman AA, ed. *Studies in Discrete Mathematics* (Nauka, Moskva), 290–306.

Edmonds J (1970) Submodular functions, matroids and certain polyhedra. *Combinatorial Structures and Their Applications* (Gordon and Breach, New York), 69–87.

Edmonds J, Johnson EL (1973) Matching, euler tours and the chinese postman. *Math. Programming* 5(1):88–124.

Feder T, Mihail M (1992) Balanced matroids. *Proc. 24th Annual ACM Sympos. on Theory of Comput.* (ACM, New York), 26–38.

Gamarnik D, Lewenstein M, Sviridenko M (2005) An improved upper bound for the tsp in cubic 3-edge-connected graphs. *Oper. Res. Lett.* 33(5):467–474.

Genova K, Williamson DP (2017) An experimental evaluation of the best-of-many christofides' algorithm for the traveling salesman problem. *Algorithmica* 78(4):1109–1130.

Goemans MX (1995) Worst-case comparison of valid inequalities for the tsp. *Math. Programming* 69:335–349.

Grigni M, Koutsoupias E, Papadimitriou C (1995) An approximation scheme for planar graph tsp. *Proc. 36th Annual Sympos. on Foundations of Comput. Sci.* (IEEE Computer Society, Washington, DC), 640.

Gurvits L (2006) Hyperbolic polynomials approach to van der waerden/schrijver-valiant like conjectures: Sharper bounds, simpler proofs and algorithmic applications. Kleinberg JM, ed. *Sympos. Theory Comput.* (ACM, New York), 417–426.

Gurvits L (2008) Van der waerden/schrijver-valiant like conjectures and stable (aka hyperbolic) homogeneous polynomials: One theorem for all. *Electronic J. Combin.* 15(1):R66.

Gurvits L, Leake J (2021) Capacity lower bounds via productization. *Sympos. Theory Comput.* (Association for Computing Machinery, New York), 847–858.

Haddadan A, Newman A (2019) Toward improving christofides algorithm for half-integer TSP. Bender MA, Svensson O, Herman G, eds. *ESA*, vol. 144 of LIPIcs (Schloss Dagstuhl-Leibniz-Zentrum für Informatik), 56:1–56:12.

Haddadan A, Newman A, Ravi R (2017) Cover and conquer: Augmenting decompositions for connectivity problems. Preprint, submitted July 17, https://arxiv.org/abs/1707.05387.

Hardy GH, Littlewood JE, Polya G (1952) *Inequalities* (Cambridge University Press, Cambridge, UK).

Held M, Karp R (1970) The traveling salesman problem and minimum spanning trees. *Oper. Res.* 18:1138–1162.

Hoeffding W (1956) On the distribution of the number of successes in independent trials. *Ann. Math. Statist.* 27:713–721.

Karlin AR, Klein N, Oveis Gharan S (2020) An improved approximation algorithm for TSP in the half integral case. Makarychev K, Makarychev Y, Tulsiani M, Kamath G, Chuzhoy J, eds. *Sympos. Theory Comput.* (ACM, New York), 28–39.

Karpinski M, Lampis M, Schmied R (2015) New inapproximability bounds for tsp. *J. Comput. System Sci.* 81(8):1665–1677.

Klein PN (2008) A linear-time approximation scheme for TSP in undirected planar graphs with edge-weights. *SIAM J. Comput*. 37(6).

Mitchell JS (1999) Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric tsp, k-mst, and related problems. *SIAM J. Comput.* 28(4):1298–1309.

Mömke T, Svensson O (2016) Removing and adding edges for the traveling salesman problem. *J. ACM* 63(1):2:1–2:28.

Mucha M (2012) $\frac{13}{9}$-approximation for graphic tsp. *Sympos. Theoret. Aspects Comput. Sci.* 30–41.

Oveis Gharan S, Saberi A, Singh M (2011) A randomized rounding approach to the traveling salesman problem. *Sympos. Foundations Comput. Sci.* (IEEE Computer Society, Washington, DC), 550–559.

Schalekamp F, Williamson DP, van Zuylen A (2012) A proof of the boyd-carr conjecture. *Sympos. Discrete Algorithms*, 1477–1486.

Schalekamp F, Williamson DP, van Zuylen A (2013) 2-matchings, the traveling salesman problem, and the subtour lp: A proof of the boyd-carr conjecture. *Math. Oper. Res.* 39(2):403–417.

Sebö A, Vygen J (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* 34:597–629.

Serdyukov AI (1978) O nekotorykh ekstremal'nykh obkhodakh v grafakh. *Upravlyaemye Sistemy* 17:76–79.

Shmoys DB, Williamson DP (1990) Analyzing the held-karp tsp bound: A monotonicity property with application. *Inform. Processing Lett.* 35(6):281–285.

Straszak D, Vishnoi NK (2019) Maximum entropy distributions: Bit complexity and stability. Beygelzimer A, Hsu D, eds. *COLT*, vol. 99 (PMLR, New York), 2861–2891.

Traub V, Vygen J, Zenklusen R (2020) Reducing path TSP to TSP. Makarychev K, Makarychev Y, Tulsiani M, Kamath G, Chuzhoy J, eds. *Sympos. Theory Comput.* (ACM, New York), 14–27.

van Bevern R, Slugina VA (2020) A historical note on the 3/2-approximation algorithm for the metric traveling salesman problem. *Historia Math.* 53:118–127.

Wolsey LA (1980) Heuristic analysis, linear programming and branch and bound. *Combinatorial Optimization II*, vol. 13 of Mathematical Programming Studies (Springer, Berlin), 121–134.

Anna R. Karlin is a professor in the Paul G. Allen School of Computer Science and Engineering and the Bill & Melinda Gates Chair of Computer Science and Engineering at the University of Washington. Her research is primarily in the design and analysis of algorithms and in algorithmic game theory. Karlin is coauthor of the book *Game Theory, Alive* published by the American Mathematical Society and a co-winner of the 2021 ACM Paris Kanellakis Theory and Practice Award.

Nathan Klein is a PhD student in computer science at the University of Washington studying approximation algorithms.

Shayan Oveis Gharan is an associate professor in the Paul Allen School of Computer Science and Engineering at the University of Washington. His research exploits several tools in mathematics such as theory of real stable and log-concave polynomials and spectral graph theory to design and analyze algorithms for discrete objects. He is a recipient of the Presburger Award in 2021.