A Randomized Rounding Approach to the Traveling Salesman Problem

Shayan Oveis Gharan [*]   Amin Saberi[†]   Mohit Singh [‡]

Abstract

For some positive constant ϵ_0, we give a $(\frac{3}{2} - \epsilon_0)$-approximation algorithm for the following problem: given a graph $G_0 = (V, E_0)$, find the shortest tour that visits every vertex at least once. This is a special case of the metric traveling salesman problem when the underlying metric is defined by shortest path distances in G_0. The result improves on the $\frac{3}{2}$-approximation algorithm due to Christofides [13] for this special case.

Similar to Christofides, our algorithm finds a spanning tree whose cost is upper bounded by the optimum, then it finds the minimum cost Eulerian augmentation (or T-join) of that tree. The main difference is in the selection of the spanning tree. Except in certain cases where the solution of LP is nearly integral, we select the spanning tree randomly by sampling from a maximum entropy distribution defined by the linear programming relaxation.

Despite the simplicity of the algorithm, the analysis builds on a variety of ideas such as properties of strongly Rayleigh measures from probability theory, graph theoretical results on the structure of near minimum cuts, and the integrality of the T-join polytope from polyhedral theory. Also, as a byproduct of our result, we show new properties of the near minimum cuts of any graph, which may be of independent interest.

1 Introduction

The Traveling Salesman Problem (TSP) is a central and perhaps the most well-known problem in combinatorial optimization. TSP has been a source of inspiration and intrigue. In the words of Schrijver [36, Chapter 58], "it belongs to the most seductive problems in combinatorial optimization, thanks to a blend of complexity, applicability, and appeal to imagination".

In an instance of the TSP, we are given a set of vertices with their pairwise distances and the goal is to find the shortest Hamiltonian cycle which visits every vertex. It is typically assumed that the distance function is a metric.

The best known approximation algorithm for TSP has an approximation factor of $\frac{3}{2}$ and is due to Christofides [13]. Polynomial-time approximation schemes (PTAS) have been found for Euclidean [2], planar [24, 3, 28], or low-genus metrics [16, 15] instances. However, the problem is known to be MAX SNP-hard [33] even when the distances one or two (a.k.a $(1, 2)$-TSP). It is also

---

proved that there is no polynomial-time algorithm with an approximation factor better than $\frac{220}{219}$ for this problem, unless $P = NP$ [32].

It has been conjectured that there is a polynomial-time algorithm for this problem with an approximation ratio of $\frac{4}{3}$. We make some progress towards proving this conjecture for *graph metrics*. These are metrics defined by shortest path distances in an arbitrary undirected graph. In other words, we develop an approximation algorithm for the following problem: Given a graph $G_0 = (V, E_0)$, find the shortest tour that visits every vertex at least once.

Theorem 1.1. *The approximation ratio of Algorithm 2 on graph metrics is at most* $\frac{3}{2} - \epsilon_0$, *where* $\epsilon_0 > 0$ *is a constant.*

A corollary of the analysis of the above theorem is that the integrality gap of the natural linear programming relaxation (due to Held and Karp [26]) is also strictly below $\frac{3}{2}$ on graph metrics. This ratio is also conjectured to be $\frac{4}{3}$ in general. Recently, the conjecture has been proved in the special case when the underlying graph G_0 is cubic [21, 1, 9].

1.1   Overview of the Algorithm and Techniques

We propose the same algorithm as in Asadpour et al. [4] for TSP. Let x be the optimum solution of the Held-Karp linear programming relaxation. We sample a tree F from a maximum entropy distribution in which for every T, $\mathbb{P}[T] \propto \prod_{e \in T} \lambda_e$. We find non-negative λ_e's in a such a way that for every edge $e \in E$ and tree F sampled from μ, $\mathbb{P}[e \in F]$ is proportional to x_e. The details are described in Section 3.

It is not hard to see that the expected cost of the above tree is bounded by the cost of x. We conjecture that the expected cost of the minimum cost Eulerian augmentation of T is strictly less than half of the cost of x for every metric.

Conjecture 1.2. *The approximation ratio of Algorithm 1 on any metric is at most* $\frac{3}{2} - \epsilon_0'$, *where* $\epsilon_0' > 0$ *is a constant.*

However, in this paper, we analyze this algorithm only for graphical metrics and after a slight modification. In our algorithm, we handle the case where x is nearly integral separately using a deterministic algorithm. In fact, when x is nearly integral, it is not hard to find a rounding scheme with an approximation ratio close to $\frac{4}{3}$. When x is not nearly integral, we follow Algorithm 1. See Algorithm 2 for the details.

The analysis of the algorithm has three major ingredients: (i) polyhedral structure of T-join polytope (ii) structure of near minimum cuts, and (iii) properties of random spanning trees. In Part (i), we use the integrality of the T-join polytope to relate the cost of the Eulerian augmentation to the distribution of near minimum cuts and the parity of the edges of F across them. This builds on the work of Wolsey [39] who showed that Christofides algorithm returns a solution whose cost is at most $\frac{3}{2}$ times the optimum value of Held and Karp LP. Part (ii) on the structure of near minimum cuts builds on the cactus structure [17] and polygon representation [5] of minimum and near-minimum cuts, respectively. Finally, the last part uses techniques from a recent and very interesting study of strongly Rayleigh measures [8] and their properties to prove results on the joint distribution of the parity of the number of edges across multiple cuts.

Structure of Near Minimum Cuts. Let $G(V, E)$ be the weighted (or fractional) graph defined by x. G is fractionally 2-regular and 2-edge connected. For some δ, consider all $(1 + \delta)$ near minimum

2

cuts or equivalently all cuts of size at most $2(1 + \delta)$. We show that for δ small enough, either a constant fraction of edges appear in a constant number of $(1+\delta)$-near minimum cuts, or x is nearly integral.

If x is integral, that is if G is a cycle, then every edge belongs to $\theta(n)$ minimum cuts where n is the number of vertices. Our characterization proves an approximate converse of this statement: for some large constant τ, if almost all the edges are in more than τ near minimum cuts, then the graph is close to a Hamiltonian cycle in the sense that almost all of its edges are nearly integral.

The above theorem is proved by a careful characterization of the structure of near minimum cuts for *any* graph and it could be of independent interest. For stating this characterization, we need to define a few things. Let C be a collection of cuts in graph G. Define a cross graph $\mathcal{G}$ on vertex set C where an edge between two vertices denotes that their corresponding cuts cross. Every connected component of $\mathcal{G}$ partitions the vertices of G into a set of "atoms". We show that if C is a collection of near minimum cuts, the graph resulting from contracting the atoms of any connected component is very close to a cycle. In particular, the weight of nearly all the edges in the resulting graph is very close to half of the size of a minimum cut of G.

Stated in the above form, our result is a generalization of Dinits et al. [17] from minimum cuts to near-minimum cuts. The main technical tool behind the proof is the structure called *polygon representation* of near-minimum cuts as defined by Benczur [5, 6, 7]. In some sense, our theorem adds a characterization of the placement of the edges between the atoms to the polygon representation, when δ is sufficiently small. We refer the reader to section 5 for the proof.

Random Spanning Trees and Strongly Rayleigh Measures. In the analysis of this algorithm for asymmetric TSP [4], Asadpour et al. use the negative correlation between the edges of random spanning trees to obtain concentration results on the distribution of edges across a cut. For this work, we have to use even stronger virtues of negative dependence [34]. In particular, we use the fact that the distribution of spanning trees belongs to a more general class of measures called Strongly Rayleigh. These measures maintain negative association and log concavity of the rank sequence similar to random spanning trees. In addition, they are closed under projection and truncation and conditioning in certain scenarios.

For an edge e, let C be the set of near minimum cuts of G that contain e. We prove that for a constant fraction of edges e, with constant probability, all of the cuts in C that contain e have an even intersection with F. Note that the expected number of edges of F across any near minimum cut in C is very close to 2 and it follows simply that a particular cut in C contains two edges of F with constant probability. Our proof shows the stronger property that with constant probability, the number of edges of F across all cuts in C is even. It is instructive to look at the case where C contains only two degree cuts corresponding to the endpoints of an edge $e = \{u, v\}$. It seems that even in this special case, there is no direct combinatorial argument to prove that with constant probability, both u and v have an even degree in F. We refer the reader to section 6 for more details.

2 Notation and the Linear Program Relaxation

We will use the following linear programming relaxation called $LP_{subtour}$, known as subtour elimination or Held-Karp linear program. Let $c(\{u, v\})$ denote the distance between u and v or the cost of choosing edge $\{u, v\}$ for each $u, v \in V(G_0)$.

$$(LP_{subtour}) \qquad \text{minimize} \qquad \sum_{u,v \in V} c(\{u,v\}) x_{\{u,v\}}$$

$$\text{subject to} \qquad \sum_{u \in S, v \in \bar{S}} x_{\{u,v\}} \geq 2 \qquad \forall\, S \subsetneq V$$

$$\sum_{u \in V} x_{\{u,v\}} = 2 \qquad \forall\, v \in V$$

$$x_{\{u,v\}} \geq 0 \qquad \forall\, u, v \in V.$$

With a slight abuse of notation, let x be an optimal solution of this LP. Define $G = (V, E, x)$ to be the fractional support graph corresponding to the optimal vector x, i.e., $E = \{e : x_e > 0\}$.

Throughout the paper, we will refer to x_e as the fraction of edge e in G and to G as a fractional or weighted graph. In this sense, the degree of a node in G is the sum of the fractions of edges incident to that node. Therefore, G is fractionally 2-regular and 2-edge connected.

The following notations will be adopted. For a set $E' \subseteq E$, and any function f defined on the edges of G, let

$$f(E') = \sum_{e \in E'} f(e).$$

For example, $c(E') = \sum_{e \in E'} c(e)$. Similarly, let $x(E') = \sum_{e \in E'} x_e$, and $c(x(E')) = \sum_{e \in E'} c(e) x_e$. In particular, we use $c(x) := c(x(E))$.

For a set $S \subseteq V$, let $E(S) = \{\{u, v\} : u, v \in S\}$ be the set of edges inside S. For two non-crossing sets $S, S' \subset V$, let $E(S, S') = \{\{u, v\} : u \in S, v \in S'\}$ be the set of edges between the vertices in S and S'. In particular, if $S \subset S'$, we use $E(S, S') := \{\{u, v\} : u \in S, v \in S' \setminus S\}$. Also let $\bar{S} = V \setminus S$, and $d(S) = E(S, \bar{S})$ for any $F \subseteq E$.

3 The Algorithm

Our algorithm is quite similar to Christofides algorithm: first it finds a spanning tree whose cost is upper bounded by the optimum, then it finds the minimum cost Eulerian augmentation of that tree.

The main difference is in the selection of the spanning tree. Here, our idea is similar to Asadpour et al. [4]. The algorithm selects a spanning tree randomly from G, the support graph of the solution of $LP_{subtour}$. The tree is sampled from a distribution μ defined over $\mathcal{T}$, the set of spanning trees of G. This distribution is called λ-uniform or maximum entropy because for every $T \in \mathcal{T}$,

$$\mathbb{P}[T] \propto \prod_{e \in T} \lambda_e.$$

The algorithm finds non-negative λ_e's in a such a way that for every edge $e \in E$ and tree F sampled from μ, $\mathbb{P}[e \in F]$ is (approximately) equal to $(1 - \frac{1}{n}) x_e$. We refer the reader to [4] for more details.

After selecting the spanning tree, the algorithm finds the minimum cost Eulerian augmentation or T-join on the odd-degree vertices of F and constructs a Hamiltonian cycle by short cutting. The details are described in Algorithm 1.

Algorithm 1 Algorithm for TSP for general metrics

Input: A set V of vertices and a cost function $c : V \times V \to \mathbb{R}^+$ satisfying the triangle inequality.
Output: A hamiltonian tour on V.
  1: Solve the $LP_{subtour}$ to get an optimum solution x. Let $G = (V, E, x)$ be the support graph of x.
  2: Define $z := (1 - 1/n)x$. Let μ denote the maximum entropy distribution over spanning trees of G such that for a spanning tree F sampled from μ, $\mathbb{P}[e \in F] = z_e$ for each edge $e \in E$.
  3: Sample a spanning tree F from μ.
  4: Let T denote the set of odd-degree nodes in F. Compute the cheapest T-join J.
  5: **return** the tour $J \cup F$.

In Conjecture 1.2, we conjecture that the expected cost of the tour returned by Algorithm 1 is strictly less than $\frac{3}{2}$ of the cost of OPT for general metrics. However, we can analyze this algorithm only for graphical metrics and after a slight modification. In a special case, where a large fraction of edges in x are nearly integral, we choose the tree deterministically. More specifically, we say an edge $e \in E$ is *nearly integral* if $x_e \geq 1 - \gamma$, where $\gamma > 0$ is a constant. Also x is a nearly integral solution of $LP_{subtour}$ if it has many nearly integral edges, i.e., $|\{e : x_e \geq 1 - \gamma\}| \geq (1 - \epsilon_2)n$ for certain constants $\gamma, \epsilon_2 > 0$. If x is a nearly integral solution of $LP_{subtour}$, we find the minimum cost spanning tree that contains as many nearly integral edges as possible. In other words, we find F' the minimum cost spanning subgraph of G_0 that contains all of the nearly integral edges and define F to be the minimum cost spanning tree of F'. Then we simply add minimum T-join on odd-degree vertices of F. The details of our final algorithm are described in Algorithm 2.

Algorithm 2 Improved approximation algorithm for graphic TSP

Input: A set V of vertices and a cost function $c : V \times V \to \mathbb{R}^+$ satisfying the triangle inequality.
Output: A hamiltonian tour on V.
  1: Let $\epsilon_2 = 2 \cdot 10^6 \sqrt{\delta}, \gamma = 4\sqrt{\delta}, \delta = 6.25 \cdot 10^{-16}$.
  2: Solve the $LP_{subtour}$ to get an optimum solution x. Let $G = (V, E, x)$ be the support graph of x.
  3: **if** x contains $(1 - \epsilon_2)n$ edges of fraction greater than $1 - \gamma$ **then**
  4:     Find a minimum cost spanning subgraph F' in G_0 that contains all the edges of fraction greater than $1 - \gamma$, and let F be the minimum cost spanning tree in F'.
  5:     Let T denote the set of odd-degree nodes in F. Compute the cheapest T-join J.
  6:     **return** the tour $J \cup F$.
  7: **else**
  8:     **return** output of Algorithm 1.
  9: **end if**

3.1 Analysis and the Structure Theorem

In the analysis, we handle the cases considered in Algorithm 1 and Algorithm 2 differently. If x is nearly integral, then a simple polyhedral argument bounds the cost of the tree F and the T-join J. The argument is presented in Section B, Case 2. Indeed the approximation factor is close to $\frac{4}{3}$ in this case.

The more interesting case is when x is not nearly integral, and F is sampled from the distribution

$$(LP_{T-join}) \qquad \text{minimize} \qquad \sum_{e \in E} c(e) y_e$$

$$\text{subject to} \qquad \sum_{e \in E(S,\overline{S})} y_e \;\geq\; 1 \qquad \forall\, S \subseteq V, |S \cap T| \text{ odd}$$

$$y_e \;\geq\; 0 \qquad \forall\, e \in E$$

Figure 1: Linear program for the T-join problem.

μ in Step 3 of Algorithm 1. In that case, first observe that the expected cost of F is at most $c(x)$ since the probability of choosing each edge e is at most x_e. The main part of the argument is to show that the expected cost of the T-join J is smaller than $(1 - \epsilon_0)\frac{c(x)}{2}$.

In order to bound the cost of the T-join, first observe that half of any solution of $LP_{subtour}$, the vector $\frac{x}{2}$, is a feasible fractional solution to the LP_{T-join} (see Figure 1) for any set $T \subseteq V$. This is because across any cut, the sum of the fractions of $\frac{x}{2}$ is at least 1. This observation, made originally by Wolsey [39], also implies that the solution of Christofides is at most $\frac{3}{2}c(x)$.

In order to get a factor better than $\frac{3}{2}$, it is sufficient to construct a feasible solution of smaller cost for the T-join polytope, when T is the set of odd degree vertices of the sampled spanning tree F. When T in LP_{T-join} is set to the odd-degree vertices of F, the constraints present are exactly for the cuts which intersect in odd number of edges with F.

A cut is a $(1 + \delta)$ near minimum cut of G if the total fraction of the edges in the cut is at most $(1 + \delta)$ times the minimum cut of G. In other words, cuts $(S, \overline{S})$ for which $x(E(S,\overline{S})) \leq 2(1 + \delta)$ are called near minimum cuts. Also, a cut $(S, \overline{S})$ is odd with respect to F iff $F \cap E(S,\overline{S})$ is odd, i.e., F contains an odd number of edges of the cut $(S, \overline{S})$. The following two definitions are crucial.

We say an edge e is **even** with respect to F if any near minimum cut that includes e is even with respect to F, i.e., for all $(S, \overline{S})$ such that $e \in E(S,\overline{S})$ and $x(E(S,\overline{S})) \leq 2(1 + \delta)$, $|F \cap E(S,\overline{S})|$ is even. Given a tree F, setting $y_e = \frac{x_e}{2(1+\delta)}$ for each edge e which is even with respect to F and $y_e = \frac{x_e}{2}$ for every other edge e, we obtain a feasible solution to the LP_{T-join} when T is the set of odd-degree vertices of F. Thus it is enough to find a tree F for which the set of even edges is large.

Let $\mathcal{E}(e)$ be the event that e is even with respect to F where F is sampled from the distribution μ. We say e is **good** if the probability of this event is bounded from zero by some constant. More precisely, if for a fixed constant $\rho > 0$,

$$\mathbb{P}\Big[\exists (S,\overline{S}) : e \in E(S,\overline{S}) \text{ and } x(E(S,\overline{S})) \leq 2(1 + \delta) \text{ and } |F \cap E(S,\overline{S})| \text{ is odd}\Big] \leq 1 - \rho.$$

Our strategy is to identify a large number of good edges in the graph. We will use these edges to show that the cost of T-join is strictly less than $\frac{c(x)}{2}$. The following Theorem shows that it is indeed possible to find such edges if the algorithm samples the tree F in Step 3.

Theorem 3.1 (Structure Theorem). *Let x be an optimal solution of $LP_{subtour}$, and let μ be the λ-uniform measure defined based on x. There exist sufficiently small constants ϵ_1, ρ bounded away from zero such that at least one of the following is satisfied by x:*

1. ***there is an abundance of good edges in x:*** *There exists a set $E^* \subset E$ such that $x(E^*) \geq \epsilon_1 n$, and*

$$\forall e \in E^* : \; \mathbb{P}[\mathcal{E}(e)] \geq \rho.$$

6

2. x **is nearly integral:** x contains at least $(1 - \epsilon_2)n$ edges of fraction greater than $1 - \gamma$.

We note that the Structure Theorem is valid for *all* feasible solutions to the Held-Karp relaxation and not just for vertex solutions of the linear program which have been studied intensely [12, 10, 23]. We also remark that we have not tried to optimize the constants but do explain the necessary dependence between them in Appendix H. Theorem 3.1 implies Theorem 1.1 by constructing feasible solutions to LP_{T-join} of small cost. We show that in Appendix B.

4 Proof of the Structure Theorem (Main Ideas)

The rest of the paper is dedicated to proving the Structure Theorem. For proving this theorem, we have to establish several results about the structure of near minimum cuts in graphs as well as properties of random spanning trees. In this section, we will explain the main ideas.

4.1 Atoms, Cut Classes and Cross Graphs

Definition 4.1 (Atom). *For a collection $\mathcal{C}$ of cuts of a graph $G = (V, E)$, the atoms of $\mathcal{C}$ are the members of a partition $\mathcal{P}$ of the vertex set V such that*

- *no cut of $\mathcal{C}$ divides any of the atoms of $\mathcal{C}$, and*

- *$\mathcal{P}$ is the coarsest partition with this property.*

We say an atom is **singleton** *if it is a set of a single vertex of V.*

Definition 4.2 (Cross Graph). *A pair of cuts $(A, \overline{A})$ and $(B, \overline{B})$ is said to* cross *if $A \cap B, A \setminus B, B \setminus A, V \setminus (A \cup B)$ are all non-empty. For a collection $\mathcal{C}$ of cuts of a graph $G = (V, E)$,* **cross graph $\mathcal{G}$** *is a graph on vertex set $\mathcal{C}$ and that has an edge between two cuts in $\mathcal{C}$ if they cross. Each connected component of $\mathcal{G}$ is called a* **cut class**.

Consider the cross graph corresponding to $(1 + \delta)$-near minimum cuts of G and let $\mathcal{C}_1, \mathcal{C}_2, \cdots \mathcal{C}_l$ be its cut classes. Denote the set of atoms of any of these families of cuts by $\phi(\mathcal{C}_i)$ for $1 \leq i \leq l$. We say a cut in $\mathcal{C}_i$ is **trivial**, if it separates an atom of $\phi(\mathcal{C}_i)$ from the rest of the atoms. Since no cut of $\mathcal{C}_i$ can cross its trivial cuts, $\mathcal{C}_i$ has a trivial cut iff it has exactly one near minimum cut, or equivalently two atoms. Moreover, no cut class can have 3 atoms.

Definition 4.3. *Let $\tau = \frac{1}{20\sqrt{\delta}} = 2 \cdot 10^6$. We say a cut class $\mathcal{C}_i$ is* **large** *if $|\phi(\mathcal{C}_i)| \geq \tau$, and* **small** *otherwise.*

Let $L(\tau)$ be the set of all atoms of the large cut classes, i.e.

$$L(\tau) = \bigcup_{\mathcal{C}_i : |\phi(\mathcal{C}_i)| \geq \tau} \phi(\mathcal{C}_i).$$

The size of $L(\tau)$ plays an important role. It can be shown that we always have $|L(\tau)| \leq n(1 + \frac{2}{\tau - 2})$. Now, if $|L(\tau)|$ is far from its maximum possible value, i.e. $|L(\tau)| < (1 - \epsilon)n$, then case 1 of Theorem 3.1 holds. Otherwise case 2 holds. In order to understand this intuitively, think about the cross graph defined by the minimum cuts of a cycle of length n. Observe that this graph contains $\binom{n}{2}$ (near) minimum cuts, and the cross graph $\mathcal{G}$ contains only one connected component or equivalently, one

7

large cut class with n atoms. Therefore, if G is a cycle of length n, $|L(\tau)| = n$. The following Lemma is a weak inverse of this observation: if $|L(\tau)|$ is large, then the LP solution is in a sense close to a Hamiltonian cycle.

Lemma 4.4. *For any $\epsilon \geq \frac{1}{\tau-2}$, if $|L(\tau)| \geq (1-\epsilon)n$ then G contains at least $(1 - 20\sqrt{\delta} - 17\epsilon)n$ edges of fraction greater than $1 - 4\sqrt{\delta}$.*

On the other hand, if $L(\tau)$ is small, we would like to show that G contains many good edges. Our best hope for finding good edges are among the edges that do not belong to too many near minimum cuts. Because of that, we will try to find edges that are not contained in any large cut classes. Let us make that precise.

Definition 4.5. *An edge e is incident to an atom a, if exactly one of its endpoints is contained in a. An edge e is said to be contained in a cut class C_i if e is incident to some atom of C_i.*

Let E_S be the set of edges that are not contained in any of the large cut classes. In the next lemma we show that if $|L(\tau)| < (1-\epsilon)n$, then $x(E_S)$ is large:

Lemma 4.6. *If $|L(\tau)| < (1-\epsilon)n$ then $x(E_S) \geq n(\epsilon - 3\delta)$.*

Being in E_S does not automatically guarantee that an edge is good (see Figure 2 for a counter example). We will identify three types of good edges in E_S in Section 4.3. We do this after we establish a few properties of near minimum cuts in the next subsection.

Before that, we are ready to assign the exact values of the constants. Lemma 4.4 places a lower bound on the value of ϵ in terms of τ. A similar bound on ϵ is also given in Theorem 4.26 when $L(\tau)$ is small. We set $\epsilon = \frac{5000}{\tau} = 2.5 \cdot 10^{-3}$ so as to satisfy all the conditions. This already implies appropriate values for ϵ_2 and γ in the algorithm. We set $\epsilon_2 = 2 \cdot 10^6 \sqrt{\delta} \geq 20\sqrt{\delta} + 17\epsilon$ and $\gamma = 4\sqrt{\delta}$. Finally, from Lemma B.3, $2\epsilon_2 + 4\gamma \leq 0.11$ is enough to give a better than $\frac{3}{2}$ bound on the performance of the algorithm. This implies $\delta = 6.25 \cdot 10^{-16}$ suffices to satisfy all the conditions.

4.2 Near Minimum Cuts and their Cactus-like Structure

In this section, we prove crucial lemmas about the structure of near minimum cuts of any graph. Applying these lemmas to the solution of the Held-Karp linear program directly yields Lemma 4.4 and Lemma 4.6.

Let H be an unweighted graph and let c denote the minimum cut of H. For a partitioning $\mathcal{P} = \{P_1, P_2, \ldots, P_k\}$ of vertices in H, let $H(\mathcal{P})$ be the graph obtained by identifying the vertex set of each part P_i, and removing the self-loops afterwards. For example, for a cut class C_i, each vertex of $H(\phi(C_I))$ is an atom of C_i.

The following lemma about the structure of minimum cuts follows from the cactus representation [17] (also see Fleiner and Frank [20] for a short proof).

Lemma 4.7. *[17] Let C_i denote a cut class of minimum cuts of H. Then $H(\phi(C_i))$ is a cycle where weight of every edge is exactly $\frac{c}{2}$ and every pair of edges of the cycle corresponds to a minimum cut of H.*

Our main result in this section is that the above lemma generalizes to the structure of near minimum cuts in an approximate sense. In the following definition we define cactaceous structures as a cycle-like structure. Then we show that any cut class of a collection of near minimum cuts is cactaceous.

Definition 4.8 ((α, α', β)-cactaceous). *A graph $H = (V, E)$ with minimum cut c is (α, α', β)-cactaceous if for some $\delta \geq 0$:*

- *There exists at least $m := (1 - \alpha \sqrt{\delta})|V(H)|$ pairs of vertices of H, $\{(v_1, u_1), (v_2, u_2), \ldots, (v_m, u_m)\}$ such that for each $1 \leq i \leq m$, $E(v_i, u_i) \geq \frac{c}{2}(1 - \alpha' \sqrt{\delta})$, and each vertex $v \in V(H)$ is contained in at most two such pairs.*

- *The number of edges of the graph H satisfies the following:*

$$\frac{c}{2}|V(H)| \leq |E(H)| \leq (1 + \beta\delta)\frac{c}{2}|V(H)|.$$

Theorem 4.9. *For any $\delta < 1/100$, let C_i denote a cut class of $(1 + \delta)$ near minimum cuts of H. Then $H(\phi(C_i))$ is $(20, 4, 3)$-cactaceous.*

If we let $\delta = 0$ in the description of Theorem 4.9, we obtain that for any cut class C_i of the collection of minimum cuts of H, the graph $H(\phi(C_i))$ is a cycle where the weight of each edge is $\frac{c}{2}$, thus we obtain Lemma 4.7.

Theorem 4.9 is proved in section 5. The main technical tool behind the proof is the structure called *polygon representation* of near-minimum cuts as defined by Benczur [5, 6, 7]. Benczur showed that for $\delta \leq 1/5$, the near minimum cuts of any graph H can be represented using polygon representation (see section 5 for more information). Our theorem uses this representation heavily. However, the emphasis of [7] (and results before that) were on representing the vertex sets of minimum cuts. Instead, here we focus on the edge sets and observe several interesting properties that could be of independent interest.

Lemma 4.7 and Theorem 4.9 show connections between a single cut class of minimum cuts and a single cut class of near minimum cuts. Observe that one particular edge might be included in various cuts occurring in distinct cut classes. We first discuss the representation of different cut classes of minimum cuts as represented by a cactus graph [17] and then the representation of different cut classes of an arbitrary collection of cuts in a tree hierarchy [5].

Definition 4.10 (Cactus Graph). *A **cactus graph** is a graph with no cut edges in which no two simple cycles share an edge. Every cactus graph K can be represented as a tree T where the vertices of T are the cycles of K and there is an edge between two cycle vertices if they share a vertex.*

The cactus representation of H consists of a partition $\mathcal{P}$ of vertices of H and the cactus graph K whose vertices may have either an element of $\mathcal{P}$ or the empty set, each element of $\mathcal{P}$ appearing exactly once.

Lemma 4.11. *[17] Let H be an unweighted graph with minimum cut c. There is a cactus $K = (U, F)$ and a mapping $f : V \to U$ so that the preimages $f^{-1}(U_1)$ and $f^{-1}(U_2)$ are the two shores of a minimum cut of H for every 2-element cut of K with shores U_1 and U_2. Moreover, every minimum cut of H arises this way.*

In his thesis, Benczur [6] generalized this concept to *any* collection of cuts. He proved that any collection of cuts possesses a tree hierarchy [6, Theorem 4.1.6].

For the purpose of the proofs, Benczur gives a slightly modified description of the cactus representation such that unlike the usual definitions, it is uniquely defined (e.g. in his definition unlike the usual description of a cactus cycles of length 3 are now allowed). Here, we also use this new representation.

Definition 4.12 (Tree Hierarchy). *A collection of cuts C with cut classes $C_1, \ldots, C_l$ possesses a* ***tree hierarchy*** *if*

- *a cactus graph $K(C)$ is given, with cycles $C_1, \ldots, C_l$ such that C_i is of length $|\phi(C_i)|$, and the elements of $\phi(C_i)$ are mapped to the vertices of C_i; and*

- *all pairs of atoms $a \in \phi(C_i)$ and $b \in \phi(C_j)$ with $i \neq j$ that are mapped to a coinciding vertex of the cactus satisfy $a \cup b = V$. We call them* ***connecting*** *atoms.*

In Figure 3 an example of a feasible solution of $LP_{subtour}$ with its cactus representation is shown.

Remark 4.13. *Unlike the old definitions for the cactus structure, the vertices of G are not mapped to just a single vertex of $K(C)$. Instead, as described above they are mapped to exactly one vertex of each cycle of $K(C)$.*

There is a simple way to transform the original cactus representation to the new one. For a cactus vertex a of a cycle C, we may get the vertices of the new definition from the original one by erasing the edges of C, and taking the union of the vertices in all cactus vertices reachable from a by the remaining cactus edges.

Theorem 4.14. *[6] Any collection of cuts possess a tree hierarchy.*

Observe that the above results can be simply applied to the fractionally weighted graphs, by multiplying the weight of each edge by some large number N such that Nx_e is an integer for all edges e, then inserting Nx_e parallel edges between the endpoints of e and applying the above theorems to the obtained multi-graph.

We now show that applying the above results to the structure of near minimum cuts of the Held-Karp linear programming solution x leads to the proofs of Lemma 4.6 and Lemma 4.4. Firstly, Theorem 4.9 implies the following Corollary about the structure of any cut class of near minimum cuts of the weighted graph $G = (V, E, x)$.

Corollary 4.15. *For any $\delta < 1/100$, let C_i be a cut class of the $(1 + \delta)$ near minimum cuts of the weighted graph $G = (V, E, x)$. Then $G(\phi(C_i))$ satisfies the following:*

- *There exists at least $m := (1 - 20\sqrt{\delta})|\phi(C_i)|$ pairs of vertices of $G(\phi(C_i))$, $\{(a_1, b_1), (a_2, b_2), \ldots, (a_m, b_m)\}$ such that for each $1 \leq i \leq m$, $x(a_i, b_i) \geq 1 - 4\sqrt{\delta}$, and each vertex $a \in V(G(\phi(C_i)))$ is contained in at most two such pairs.*

- $|\phi(C_i)| \leq x(E(G(\phi(C_i)))) \leq (1 + 3\delta)|\phi(C_i)|$.

The proof of Lemmas 4.4 and 4.6 simply follows from the above corollary and tree representation of near min-cuts as given by Theorem 4.14. We show that in Appendix C.

4.3 Good Edges and Random Spanning Trees

In this section, we show that at least a constant fraction of the edges in E_S (as defined before Lemma 4.6) are good. It is easy to see (and it is proved in Lemma C.1) that each small cut class can only have a constant number of near minimum cuts. Therefore all edges in E_S that are only in a constant number of small cut classes appear only in a constant number of near minimum cuts. We will identify three types of good edges from these. The first of these are "trivial edges".

Definition 4.16 (Trivial Edge). *We call an edge **trivial** if it occurs in only two near minimum cuts, which are the degree constraint of its endpoints.*

Trivial edges are the simplest possible candidate edges for being good. In fact, besides certain extreme cases, most trivial edges are good (see Figure 2 for an example of a trivial edge which is not good).

Proposition 4.17. *Any trivial edge e such that $x_e < \frac{1}{2} - \frac{1}{8000}$ or $x_e > \frac{1}{2} + \frac{1}{8000}$ is good. Furthermore, of any pair of adjacent trivial edges, one of them is good.*
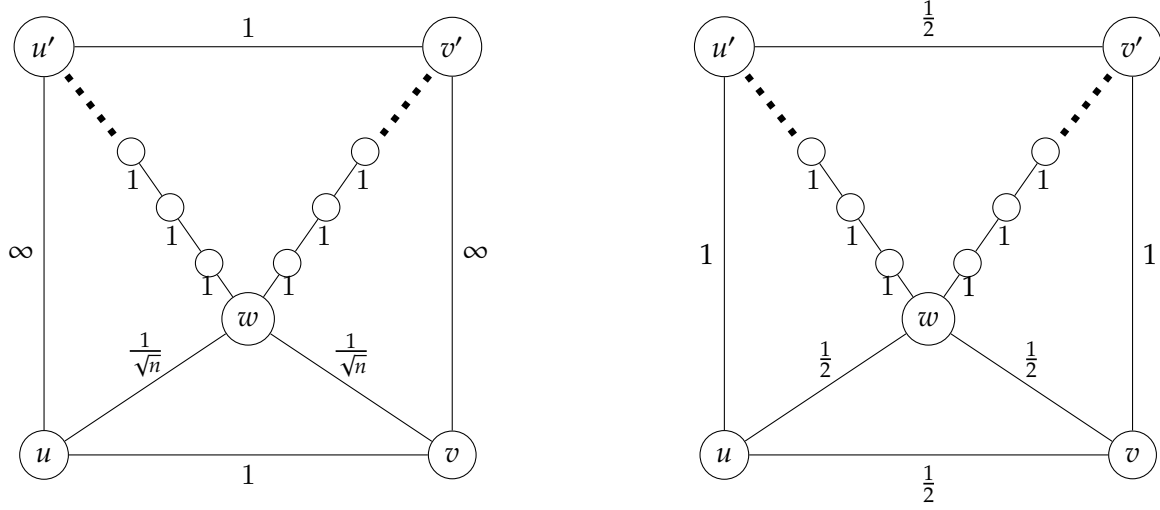

Figure 2: The left diagram represents the λ values of the edges, while the right diagram represents the approximate probability of each edge. The example shows that although $\mathbb{P}\left[(u,v) \in T\right] \simeq \frac{1}{2}$, and the expected degree of u and v is 2, $\mathbb{P}\left[deg_T(u) + deg_T(v) = 3\right] = 1 - o(1)$. Therefore (u,v) is not good. Note that this is not exactly a solution to $LP_{subtour}$ but still points to difficulties in proving Proposition 4.17.

The above Proposition is proved in Section 6.2 in two parts under Proposition 6.18 and Lemma 6.23. As showed in Figure 2, if e is a trivial edge but $x_e = 1/2 \pm o(1)$, then e might be contained in an odd near minimum cut with high probability and therefore it will not be good. Because of this, we have the pairing of trivial edges in the second part of the proposition.

Definition 4.18 (Inside Edge). *An edge $e = (u,v)$ is an inside edge of a non-trivial small cut class C_i if C_i is the only non-trivial small cut class that contains e, and atoms of C_i containing u and v are singletons.*

The second type of good edges are inside edges.

Proposition 4.19. *Let C_i be a small cut class which contains an atom with more than $n/2$ vertices of G and $200|\phi(C_i)|^2\delta \le 1$. Then any inside edge of C_i is good.*

Again, the above statement is directly implied by Lemma 6.27 proved in Section 6.3. In most cases, the sum of the fractions of good trivial and inside edges covered in the above lemmas add

11

up to $\Omega(n)$. However, there are graphs that do not have any trivial or inside edges, even though they do not have any large cut classes (see Figure 3). In those graphs, we identify another class of potentially good edges.
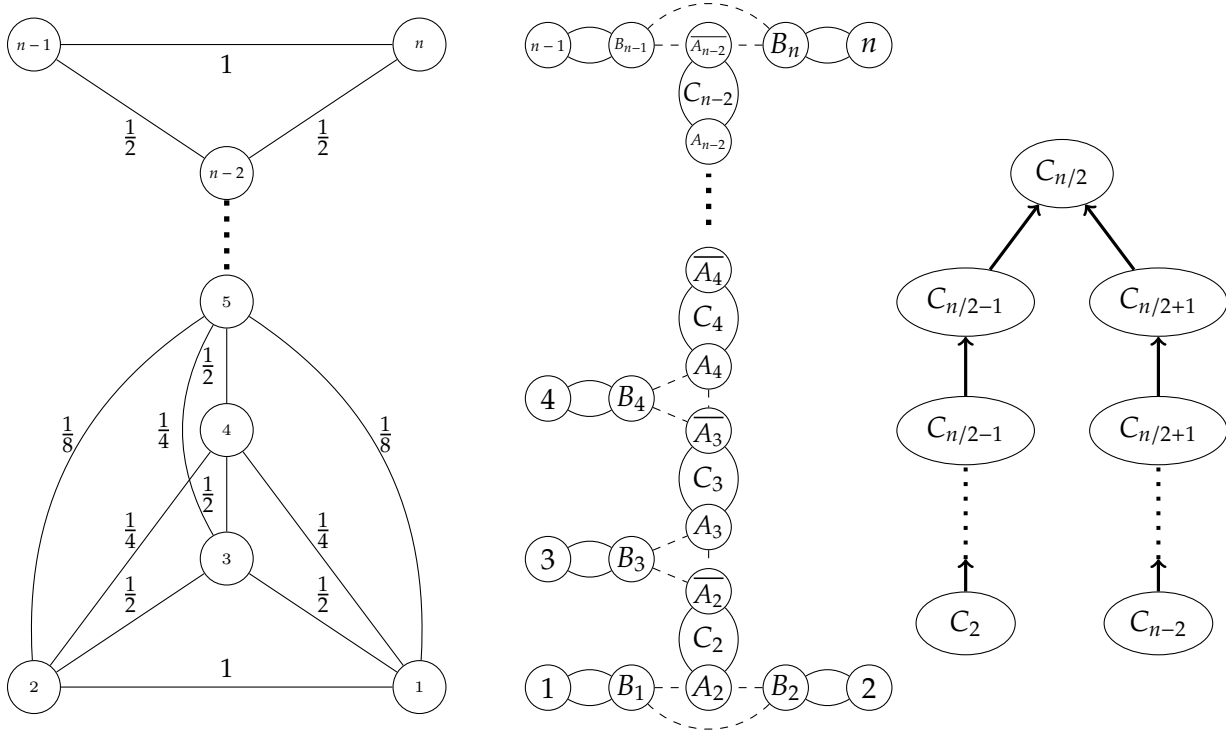

Figure 3: The left diagram represents a feasible solution of $LP_{subtour}$, the middle diagram represents the tree hierarchy of the cut classes, and the right diagram shows the tree $\Gamma(\mathcal{C})$. The graph is constructed by connecting each pair of vertices i, j by $x_{i,j} := 2^{-|i-j|}$, unless i or $j = 1, 2, n-1, n$; in which the fractions are selected precisely to get a feasible LP solution. In the middle diagram the cactus representation of the graph is shown. Here, $\mathcal{C}_2, \mathcal{C}_3, \ldots, \mathcal{C}_{n-2}$ are the non-trivial cut classes, and the rest are trivial. Each cut class $\mathcal{C}_i$ contains two atoms $\phi(\mathcal{C}_i) := \{A_i, \overline{A_i}\}$, where $A_i := \{1, 2, \ldots, i\}$, and each trivial-cut class contain the atoms $\{\{i\}, B_i := \overline{\{i\}}\}$. The dashed edges is used to represent the connecting pairs of atoms. $\Gamma(\mathcal{C})$, however, contains only the non-trivial cut classes. Observe that $\mathcal{C}_{n/2}$ is chosen to be the root, and the tree is made of two long threads. In particular, note that all of the edges except $\{1, 2\}, \{n-1, n\}$, are non-trivial, and the graph does not contain any inside edges.

Definition 4.20 (Trivial Cut Classes). *We say a cut class is trivial if it contains exactly two atoms, and one of them is a singleton.*

Let $\Gamma(\mathcal{C})$ be the tree defined by the cactus graph representing the tree hierarchy of all near minimum cuts of G, such that each node of $\Gamma(\mathcal{C})$ corresponds to a **non-trivial** cut class of $\mathcal{C}$ (since the trivial cut classes are the leaves of tree defined on $K(\mathcal{C})$ the tree remains connected after removing those cut classes). Lemma F.1 shows that for all cut classes, except possibly one, there is an atom which has more than $n/2$ vertices. Let $\mathcal{C}_r$ be that cut class. We root the tree $\Gamma(\mathcal{C})$ at $\mathcal{C}_r$

(if C_r does not exist, we just pick an arbitrary cut class) and direct it away from the root. Figure 3 represents an example of the directed tree $\Gamma(C)$.

Definition 4.21 (Father-Connector). *For each non-root cut class C_i in $\Gamma(C)$, we call the unique connecting atom that contains at least $n/2$ vertices the **father-connector** of C_i. Observe that father-connector of C_i coincides with an atom of the father of C_i in a vertex of $K(C)$.*

Definition 4.22 (Thread). *Let P be a directed path in $\Gamma(C)$. We call P a **thread** iff all of its vertices are small cut classes, and they have degree exactly 2 in $\Gamma(C)$. The length of the thread is the number of vertices of P. We call a thread **unbalanced** with parameter $\epsilon \geq 0$ if the father-connector atoms of all of its cut classes contain at least $n(\frac{1}{2} + \epsilon)$ vertices. Observe that unbalanced threads do not contain the root.*

Proposition 4.25 shows that in each unbalanced thread, a constant fraction of edges are good edges. We will also show if the sum of the fraction of good edges in E_S that are either trivial or inside is small, then $\Gamma(C)$ contains many long threads. However, these two statement do *not* necessarily imply that we always have a sufficient number of good edges; because we may count a good edge several times. In order to avoid this, we *assign* every good edge to a cut class.

Definition 4.23 (Edge-Assignment). *Let $e = \{u, v\} \in E_S$ be a good edge, not contained in C_r. We assign e to the farthest cut class C_i from the root that contains e. For example, if e is an inside edge of some small cut class, it will be assigned to that class. Otherwise, if e is not contained in any cut class, we assign e to the farthest cut class C_i from the root that has u and v in one of its atoms other than the father-connector. For example, if a small cut class C_i contains a non-connecting non-singleton atom a, then all of the good edges between the vertices of a will be assigned to C_i. In both cases, if C_i is not a small cut class, or C_i is the root of $\Gamma(C)$, we do not assign e.*

Remark 4.24. *It is worth noting that the set of cut classes that contain an edge $e \in E$ make a path $C_1, C_2, \ldots, C_k$, in the underlying (undirected) graph of $\Gamma(C)$. This path can be decomposed into two directed paths, say $C_1, C_2, \ldots, C_i$ and $C_k, C_{k-1} \ldots, C_i$; therefore, the farthest cut class from the root that contains e might not be well-defined (i.e. it can be either C_1 or C_k if $i - 1 = k - i$). In this special case we assign e to one of the two possible farthest cut classes arbitrarily. However, all three types of edges that we consider in the proof of Theorem 4.26 are contained in a single **directed** path of cut classes in $\Gamma(C)$.*

We say a good edge is assigned to a thread if it is assigned to one of the cut classes of the thread. For any $\beta > 0$, let π_β be the set of unbalanced threads with parameter β. In the next proposition, we show that any thread $P \in \pi_{\frac{\epsilon}{5120}}$ of length at least 10 is assigned at least a fraction $\frac{\epsilon}{20480}$ of good edges. See Subsection 6.4 for the proof.

Proposition 4.25. *Let $\pi_{\frac{\epsilon}{5120}}$ be the set of unbalanced threads with parameter $\frac{\epsilon}{5120}$. Then, the above method assigns at least $\frac{\epsilon}{20480}$ fraction of (good) edges that are even with probability at least $\epsilon^3 10^{-16}$ to any thread $P \in \pi_{\frac{\epsilon}{5120}}$ of length at least 10.*

In the next theorem, we show that if the fractional value of trivial or inside edges in $\Gamma(C)$ is less than $\frac{\epsilon n}{3840}$, then it contains at least $\frac{\epsilon n}{2560}$ unbalanced threads with parameter $\frac{\epsilon}{5120}$ of length 10.

Let us define some notations. Let $E_{S_T} \subset E_S$ be the set of trivial good edges, $E_{S_N} \subset E_S$ be the set of edges (not necessarily good) that are contained in at least one small non-trivial cut class, and let $E_{S_I} \subset E_{S_N}$ be the set of inside edges of small cut classes. Also let π^* be the maximum disjoint set of unbalanced threads with parameter $\epsilon/5120$, each of length at least 10.

Theorem 4.26. *If* $|L(\tau)| < (1-\epsilon)n$, *and* $\epsilon \geq \frac{4920}{\tau}$ *then one of the following is always true:*

1. $x(E_{S_I}) + x(E_{S_T}) \geq \frac{\epsilon n}{3840}$; *or*

2. *The fractional value of edges assigned to threads in* π^* *is at least* $\frac{\epsilon^2 n}{6 \cdot 10^7}$.

The proof appears in the appendix Section D

5 The Cactus-like Structure of Near Minimum Cuts

In this section, we will prove Theorem 4.9. For the sake of brevity, we do not work with the graph H, instead we prove the theorem for a fixed cut class C_i of H. Since C_i is fixed in this section, with a little abuse of notation let H be the graph obtained by identifying the vertex set of each atom of $\phi(C_i)$, and removing the self-loops (i.e. we use H to represent $H(\phi(C_i))$). Theorem 4.9 can be seen as an immediate corollary of the following lemma:

Lemma 5.1. *Let* $c \geq 1$, *be an integer,* $\delta < 1/100$, *and* $H = (V, E)$ *be an unweighted* c-*edge connected graph such that for any non-trivial cut* $(S, \overline{S})$ *there exists a* $(1 + \delta)$ *near minimum cut that crosses* $(S, \overline{S})$. *If the cross graph of the* $(1 + \delta)$ *near minimum cuts of* H *is connected, then* H *is* $(20, 4, 3)$-*cactaceous.*

Observe that the two properties of graph H in the Lemma follows by the definition of atoms and cut classes. In the rest of this section we prove the above Lemma. Therefore, we always assume H is a c-edge connected unweighted graph that satisfies the assumptions of Lemma 5.1.

Let us start by some basic lemmas regarding the properties of near minimum cuts. The cut of a set $X \subset V$ is the set of edges $d(X)$ that separates X from its complement. Let X and Y be two crossing sets of V. It was shown in [17] that if their corresponding cuts are minimum cuts, then the cuts defined by any of the four sets $X \cap Y, X \cup Y, X \setminus Y, Y \setminus X$ is also a minimum cut. The following lemma proves an extension of this property for near minimum cuts:

Lemma 5.2. *Let* $(X, \overline{X})$ *and* $(Y, \overline{Y})$ *be two crossing cuts of* H *and let* $(X, \overline{X})$ *be a* $(1 + \delta)$ *near minimum cut. Then the size of the cuts of any of the four sets* $X \cap Y, X \cup Y, X \setminus Y, Y \setminus X$ *is at most* δc *more than* $|d(Y)|$.

Proof. We prove the lemma only for $X \cap Y$; the rest of the cases can be proved similarly. Since the cut function $|d(.)|$ is a submodular function we have

$$|d(X \cap Y)| + |d(X \cup Y)| \leq |d(X)| + |d(Y)| \leq (1 + \delta)c + |d(Y)|.$$

Since $|d(X \cup Y)| \geq c$, we have $|d(X \cap Y)| \leq |d(Y)| + \delta c$. This completes the proof of the lemma. □

For two disjoint sets $X, Y \subset V$ recall that $E(X, Y) := \{(u, v) : u \in X, v \in Y\}$. Benczur in [6, Lemma 5.3.5] proved the following useful lemma:

Lemma 5.3 (Benczur [6]). *Let* $(X, \overline{X})$ *and* $(Y, \overline{Y})$ *be two crossing* $(1 + \delta)$ *near minimum cuts of* H. *Then* $|E(X \cap Y, X \cap \overline{Y})| \geq (1 - \delta)\frac{c}{2}$.

Note that it can be derived from the above lemma that $|E(X \cap Y, Y \cap \overline{X})| \geq (1-\delta)\frac{c}{2}$, $|E(X \cap \overline{Y}, \overline{X} \cap \overline{Y})| \geq (1-\delta)\frac{c}{2}$ and $|E(Y \cap \overline{X}, \overline{X} \cap \overline{Y})| \geq (1-\delta)\frac{c}{2}$.

We start by proving that H satisfies the second property of cactaceous structures for $\beta = 3$. Since H is c-edge connected, the degree each vertex is at least c. Therefore, it remains to show that the average degree of vertices of H is not far from c. Interestingly, there are examples where some vertices in H can have very large degrees. The difficulty of the proof follows from this fact by noting that we are not able to use proof techniques that bound the maximum degree of the vertices.

Lemma 5.4. *The average degree of the vertices of H is at most $(1 + 3\delta)c$.*

Proof. Let $(X, \overline{X})$ be a near minimum cut of H. We maintain a partition $\mathcal{S}$ of vertices of H. We initialize $\mathcal{S} = \{X, \overline{X}\}$ and iteratively break one of the sets in $\mathcal{S}$ into two sets. In particular, we pick a subset $S_j \in \mathcal{S}$ of size at least 2, and we find a cut $(S'_j, S_j \setminus S'_j)$ separating the vertices of S_j minimizing $|E(S'_j, S_j \setminus S'_j)|$. Note these minimum cuts are restricted to the subsets, $S_j \subset V$, and can be different from the near minimum cuts of H. We then substitute S_j with S'_j and $S_j \setminus S'_j$. We keep doing this operation until all of the elements of $\mathcal{S}$ become singletons. Note that each edge between the vertices of H has contributed to exactly one of the cuts considered during this procedure. Therefore, if we show that the size of each cut $|E(S'_j, S_j \setminus S'_j)| \leq (1 + 3\delta)\frac{c}{2}$, then the average degree of the vertices of H will be at most:

$$\frac{|E|}{|V|} \leq \frac{|E(X, \overline{X})|}{|V|} + \frac{|V| - 2}{|V|}(1 + 3\delta)\frac{c}{2} \leq (1 + 3\delta)c.$$

We say a near minimum cut $(Y, \overline{Y})$ is *outside* of a set S, if both of its cut sides are not a subset of S (i.e., $Y, \overline{Y} \not\subseteq S$). Observe that if a cut $(Y, \overline{Y})$ is outside of S then either one of the cut sides $Y, \overline{Y}$ contains S or the cut $(Y, \overline{Y})$ crosses S, i.e., both the cut sides intersect S non-trivially. In Claim 5.5 we show that that if a set $S \subset V$ is highly connected (i.e., $\min_{T \subset S} |E(T, S \setminus T)| \geq 1 + 3\delta)\frac{c}{2}$), then any near minimum cut $(Y, \overline{Y})$ of H that is outside of S crosses it. But, all S_j considered throughout the procedure are either a subset of X, or $\overline{X}$ and $(X, \overline{X})$ does not cross any of them. Since $(X, \overline{X})$ is outside of any set S_j, the size of the minimum cut of each S_j must be at most $(1 + 3\delta)\frac{c}{2}$, and we are done.

Claim 5.5. *Let $S \subset V$ such that $(S, \overline{S})$ is a non-trivial cut, and for any set $T \subset S$, $|E(T, S \setminus T)| > (1 + 3\delta)\frac{c}{2}$, then any near minimum cut of H that is outside of S, crosses S.*

Proof. We prove by contradiction. Suppose there exists a near minimum cut $(X_0, \overline{X_0})$ outside of S, and not crossing S. Since $(S, \overline{S})$ is a non-trivial cut of H, there exists a near minimum cut $(X_1, \overline{X_1})$ crossing $(S, \overline{S})$. Moreover, since the set of near minimum cuts of H are connected, there is a path of crossing cuts which connects $(X_0, \overline{X_0})$, and $(X_1, \overline{X_1})$.

Let $(Y_0, \overline{Y_0}), \ldots, (Y_k, \overline{Y_k})$ be a sequence of cuts such that $Y_0 = X_0$, for each $0 \leq i < k$, $(Y_i, \overline{Y_i})$ crosses $(Y_{i+1}, \overline{Y_{i+1}})$, and $(Y_k, \overline{Y_k})$ is the only cut that crosses S. We reach to a contradiction by showing that $(Y_k, \overline{Y_k})$ can not be a near minimum cut. Since $(Y_0, \overline{Y_0})$ is outside of S, and all cuts except $(Y_k, \overline{Y_k})$ do not cross S, all of the cuts of the path must be outside of S. Therefore, since $(Y_{k-1}, \overline{Y_{k-1}})$ does not cross S, without loss of generality we may assume that $Y_{k-1} \cap S = \emptyset$.

Since $(Y_k, \overline{Y_k})$ crosses S, by Claim's assumption we have $|E(Y_k \cap S, \overline{Y_k} \cap S)| > (1+3\delta)\frac{c}{2}$. On the other hand, since $(Y_k, \overline{Y_k})$ crosses $(Y_{k-1}, \overline{Y_{k-1}})$ by Lemma 5.3 we have $|E(Y_k \cap Y_{k-1}, \overline{Y_k} \cap Y_{k-1})| \geq (1-\delta)\frac{c}{2}$.

Therefore:

$$
\begin{aligned}
|E(Y_k)| \;&\geq\; |E(Y_k \cap S, \overline{Y_k})| + |E(Y_k \cap Y_{k-1}, \overline{Y_k})| \\
&\geq\; |E(Y_k \cap S, \overline{Y_k} \cap S)| + |E(Y_k \cap Y_{k-1}, \overline{Y_k} \cap Y_{k-1})| \\
&>\; (1 + 3\delta)\frac{c}{2} + (1 - \delta)\frac{c}{2} = (1 + \delta)c,
\end{aligned}
$$

where the first inequality holds by the fact that $Y_{k-1} \cap S = \emptyset$. Therefore, $(Y_k, \overline{Y_k})$ can not be a near minimum cut of H which is a contradiction. $\square$

$\square$

It remains to prove H satisfies the first property of the cactaceous structures for $\alpha = 20, \alpha' = 4$. This is the most technical part of the proof of Lemma 5.1.

Lemma 5.6. *There exists at least $m = (1 - 20\sqrt{\delta})|V|$ pairs of vertices of H, $\{(v_1, u_1), (v_2, u_2), \ldots, (v_m, u_m)\}$, such that for each $1 \leq i \leq m$, $|E(v_i, u_i)| \geq (1 - 4\sqrt{\delta})\frac{c}{2}$, and each vertex $v \in V$ is contained in at most two such pairs.*

To prove this lemma we need to use the polygon representation defined in [5] and then generalized in [7] to represent the set of near minimum cuts of each of the cut classes of H.

Definition 5.7. *The* polygon representation *of a cut class C_i possesses the following properties:*

- *A* **representing polygon** *is a regular polygon with a collection of distinguished* **representing diagonals***, with all polygon-edges and diagonals drawn by straight lines in the plane. These diagonals divide the polygon into* cells.

- *Each atom of $\phi(C_i)$ is mapped to a (different) cell of this polygon; some cells may contain no atoms.*

- *No cell has more than one incident polygon edge; each cell incident to the polygon boundary contains an atom which we call an* outside *atom. The rest of the atoms are called* inside *atoms.*

- *Each representing diagonal defines a cut, with sides being the union of the atoms contained by cells on each side of the diagonal; The collection of cuts C_i is equal to the collection of cuts defined by representing diagonals.*

We learn from Benczur [6, Theorem 5.2.2] that if C_i represents a connected set of $(1 + \delta)$ near minimum cuts and $\delta < 1/5$, then it possesses a polygon representation. Moreover, he gives an explicit representation that determines which atoms will be inside or outside.

Recall that by the assumptions of Lemma 5.1, H contains exactly one cut class, and the atoms of that cut class are exactly the vertices of H. Therefore, in the rest of the proof we use the fact that $(1 + \delta)$ near minimum cuts of H can be represented by a polygon representation where each atom is a single vertex of H.

We start by proving Lemma 5.6, in the special case where the polygon representation does not have any inside atoms. In subsection 5.1 we extend the proof to the general case where inside atoms may exist. So for now, we may assume that we have a (regular) convex $|V|$-gon, such that each vertex is mapped to a distinguished edge of the polygon, and diagonals represent the near minimum cuts.

Lemma 5.8. *Suppose that the polygon representation of H does not have any inside atoms. Then, for any integer $\sigma \geq 12$, there exists at least $m = (1 - \frac{8}{\sigma - 10})|V|$ pairs of vertices of H, $\{(v_1, u_1), (v_2, u_2), \ldots, (v_m, u_m)\}$ such that for each $1 \leq i \leq m$, $|E(v_i, u_i)| \geq (1 - 2\sigma\delta)\frac{c}{2}$, and each vertex $v \in V$ is contained in at most two such pairs.*

Proof. Let $n := |V|$, and let us label the vertices of the polygon representation cyclically as $p_1, p_2, \ldots, p_n$. Moreover, assume that the n vertices $v_1, v_2, \ldots, v_n$ of H are placed on the edges of the polygon such that v_i is placed on the edge (p_i, p_{i+1}). We show that we can select the m highly connected pairs, from the pairs of vertices mapped to adjacent edges of the polygon. In particular, we show that there is a set A of vertices of H of size $|A| = (1 - \frac{8}{\sigma - 10})n$ such that for any $v_i \in A$, we have $|E(v_i, v_{i+1})| \geq (1 - 2\sigma\delta)\frac{c}{2}$. Note that the lemma easily follows from the existence of A, we can construct the m pairs based on each vertex v_i of A and v_{i+1}. Therefore, each $v_j \in V$ may be included in at most two pairs (v_{j-1}, v_j), and (v_j, v_{j+1}).

In order to show there is a large number of edges connecting v_i and v_{i+1} it is sufficient to show that $|d(\{v_i, v_{i+1}\})|$ is small. In particular, since H is c-edge connected, if $|d(\{v_i, v_{i+1}\})| \leq (1 + \delta')c$, then $|E(v_i, v_{i+1})| \geq (1 - \delta')\frac{c}{2}$ for any $\delta' > 0$. Therefore it is sufficient to show that there is a large set A, such that for any $v_i \in A$, the size of the cut $d(\{v_i, v_{i+1}\})$ is at most $(1 + 2\sigma\delta)c$.

Before defining the set A, we need to define some notations. Since each near minimum cut of H is corresponding to a representing diagonal (and a consecutive sequence of vertices), we will use intervals to represent (near minimum) cuts. For example, for any two vertices p_i and p_j the interval $[p_i, p_j] := \{p_i, p_{i+1}, \ldots, p_j\}$, or $[p_i, p_j) := \{p_i, p_{i+2}, \ldots, p_{j-1}\}$. Also let $C(p_i, p_j) := d(\{v_i, v_{i+1}, \ldots, v_{j-1}\})$ be the cut corresponding to the representing diagonal that connects p_i to p_j if it existed. We say two intervals cross, if their corresponding cuts cross. For example, the intervals $[p_i, p_{i+2}]$ and $[p_{i+2}, p_{i+4}]$ do not cross, while $[p_i, p_{i+2}]$ and $[p_{i+1}, p_{i+3}]$ cross each other. We say an interval $[p_{i'}, p_{j'}]$ is a *subinterval* of $[p_i, p_j]$ if the set of polygon vertices contained in $[p_{i'}, p_{j'}]$ is a subset of $[p_i, p_j]$. For example, $[p_2, p_4]$ is a subinterval of $[p_1, p_5]$, but $[p_4, p_2]$ is not a subinterval of $[p_1, p_5]$.

For each vertex p_i, let q_i be the nearest vertex to p_i (in terms of the (clockwise) cyclic distance), such that $C(p_i, q_i)$ is a near minimum cut. Note that each vertex p_i is adjacent to at least one representing diagonal; otherwise, there is no near minimum cut crossing $\{v_{i-1}, v_i\}$, which is a contradiction. Thus q_i is well defined. Since we only consider the non-trivial near minimum cuts of H we have $q_i \neq p_{i+1}$ and $q_i \neq p_{i-1}$. Moreover, note that if $q_i = p_{i+2}$, then we can easily add v_i to A, since $|d(\{v_i, v_{i+1}\})| \leq (1 + \delta)\frac{c}{2}$. Hence, we may only focus on the vertices of the polygon where $q_i \notin [p_{i-1}, p_{i+2}]$.

For any vertex p_i of the polygon, we define a *chain* as a sequence of vertices $q_{i_0}, q_{i_1}, q_{i_2}, \ldots, q_{i_l}$ satisfying the following properties:

1. $q_{i_0} = q_i$,

2. $q_{i_l} = p_{i+2}$,

3. for all $j \geq 1$, we have $q_{i_j} \in [p_{i+2}, q_{i_{j-1}})$; and

4. for all $j \geq 1$, there exists a vertex r_j such that the cut $C(q_{i_j}, r_j)$ is a near-minimum cut, and it crosses the cut $C(p_i, q_{i_{j-1}})$.

The length of the chain is the number of its vertices. For any vertex p_i, let P_i be its shortest length chain.
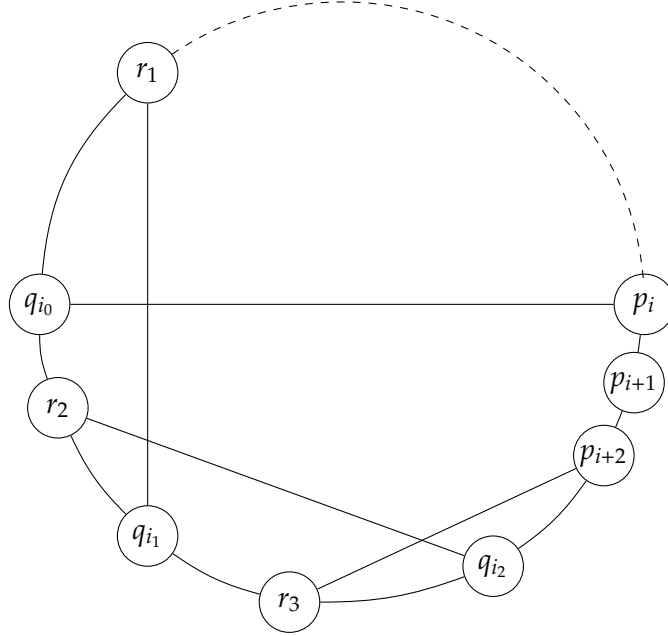
Figure 4: An example of a chain of length 4 with vertices $q_{i_0}, q_{i_1}, q_{i_2}, p_{i+2}$ for a vertex p_i.

Chains will be very useful constructions in the rest of the proof. Figure 4, and the construction given in the proof of the first part of the next claim should help with a better understanding of the definition. The last statement of the Claim will be useful later in the proof of Lemma 5.10.

Claim 5.9. *The following properties are satisfied for any vertex p_i of the polygon representation:*

1. *There exists at least one chain for p_i.*

2. *If there exists a chain of length σ for p_i, then $|E(v_i, v_{i+1})| \geq (1 - 2\sigma\delta)\frac{c}{2}$.*

3. *Let $P'_i := \{q'_{i_0}, q'_{i_1}, \ldots, q'_{i_l}\}$ be a chain for vertex p_i, and let $C(p_{i'}, p_{j'})$ be a near minimum cut such that $p_{i'} \in [p_{i+2}, q_i)$, and more than two vertices of P'_i are contained in the interval $[p_{i'}, p_{j'})$, If $[p_{i'}, p_{j'}]$ crosses $[p_i, q_i]$, or is a subinterval of $[p_i, q_i]$, then there is a shorter chain for p_i.*

Proof. We start by proving the first property. We construct the chain starting at $q_{i_0} := q_i$, and inductively adding vertices that are closer to p_i, such that for each new vertex q_{i_j} there is a near minimum cut $C(q_{i_j}, r_j)$ that crosses $C(p_i, q_{i_{j-1}})$. Since $q_{i_{j-1}} \neq p_{i+1}$, the cut $C(p_i, q_{i_{j-1}})$ is always a non-trivial cut, thus by the assumptions on H, there exists a near minimum $C(p_{i'}, p_{j'})$ crossing it. Trivially one of $p_{i'}$ or $p_{j'}$ (say $p_{j'}$) is closer to p_i than $q_{i_{j-1}}$. Therefore, we can let $q_{i_j} := p_{j'}$, unless $j' = i + 1$ (i.e., $p_{j'} \notin [p_{i+2}, q_{i_{j-1}})$). To avoid this problem we let $C(p_{i'}, p_{j'})$ to be a near minimum crossing $C(p_{i+1}, q_{i_{j-1}})$, and we exploit the assumption that q_i is the closest vertex to p_i that makes a near minimum cut.

Consider the cut $C(p_{i+1}, q_{i_{j-1}})$. Since $q_{i_{j-1}} \in [p_{i+3}, q_i]$ (i.e., the chain has not completed yet), $C(p_{i+1}, q_{i_{j-1}})$ is a non-trivial cut. Thus there is a near minimum cut $C(p_{i'}, p_{j'})$ that crosses $C(p_{i+1}, p_{i_{j-1}})$. Without loss of generality, suppose that $p_{j'} \in (p_{i+1}, p_{i_{j-1}})$. Therefore, the cut $C(p_{i'}, p_{j'})$ would also cross $C(p_i, q_{i_{j-1}})$, unless $j' = i$. But this means that q_i was not the closest vertex to p_i that makes a

near minimum cut which is a contradiction. Therefore, the near minimum cut $C(p_{i'}, p_{j'})$ also crosses $C(p_i, q_{i_{j-1}})$, and we can simply set $q_{i_j} := p_{j'}$.

For the second property we use a simple induction to show that $|C(p_i, q_{i_j})| \leq c + (2j + 1)\delta c$. Note that if this is true, then we have $|C(p_i, p_{i+2})| = |C(p_i, q_{i_{\sigma-1}})| \leq c + (2\sigma - 1)\delta c$ which implies that $|E(v_i, v_{i+1})| \geq (1 - \sigma\delta)\frac{c}{2}$, and we are done. The induction hypothesis trivially holds for the base case since $|C(p_i, q_{i_0})| = |C(p_i, q_i)| \leq (1 + \delta)c$. Now suppose it holds for the cut $C(p_i, q_{i_j})$. Observe that by Lemma 5.2 we have

$$|C(p_i, q_{i_{j+1}})| \leq |C(p_i, q_{i_j})| + \delta c \leq c + (2j + 1)\delta + 2\delta c = c + (2(j + 1) + 1)\delta c,$$

where the last inequality holds by the induction hypothesis. This implies the induction claim.

To prove the last property, it suffices to construct a shorter chain for p_i using the near minimum cut $C(p_{i'}, p_{j'})$. Let q'_{i_j}, q'_{i_k} be the first, and the last vertices of the chain that are contained in $[p_{i'}, p_{j'})$. By claim's assumptions we have $k \geq j + 2$.

Let $P_i := \{q'_{i_0}, q'_{i_1}, \ldots, q'_{i_j}, p_{i'}, q'_{i_{k+1}}, \ldots, q'_{i_l}\}$. Since $k \geq j + 2$, the length of P_i is smaller than P'_i. Suppose $[q'_{i_{k+1}}, r'_{k+1}]$ is the cut that crosses $[p_i, q'_{i_k}]$. To prove P_i is indeed a chain we need to show that $[p_{i'}, p_{j'}]$ crosses $[p_i, q'_{i_j}]$, and $[q'_{i_{k+1}}, r'_{k+1}]$ crosses $[p_i, p_{i'}]$. The latter can be proved simply by noting that $p_{i'} \in (q'_{i_{k+1}}, q'_{i_k}]$.

It remains to prove $[p_{i'}, p_{j'}]$ crosses $[p_i, q'_{i_j}]$. By claim's assumption $[p_{i'}, p_{j'}]$ either crosses $[p_i, q_i]$, or is a subinterval of it. If $[p_{i'}, p_{j'}]$ crosses $[p_i, q_i]$, then we have $q'_{i_j} = q'_0$. Therefore, $[p_{i'}, p_{j'}]$ crosses $[p_i, q'_{i_j}]$. On the other hand, if $[p_{i'}, p_{j'}]$ is a subinterval of $[p_i, q_i]$, then $p_{j'} \in (p_{i'}, q_i]$. Since $q'_{i_j} \in [p_{i'}, p_{j'})$, $[p_{i'}, p_{j'}]$ must cross $[p_i, q'_{i_j}]$.

$\square$

Now we are ready to define the set A. We define $A := \{v_i : \ length(P_i) \leq \sigma\}$ to be the set of vertices v_i such that the length of the shortest chain P_i of p_i is at most σ. By property 2 of the above claim, if length of P_i is at most σ, then we have $|E(v_i, v_{i+1})| \geq (1 - 2\sigma\delta)\frac{c}{2}$. It remains to prove that $|A| \geq (1 - \frac{8}{\sigma-10})n$.

We say a vertex p_i is *bad* if $length(P_i) > \sigma$ (i.e., $v_i \notin A$). In the next lemma we show that the number of bad vertices is at most $\frac{8}{\sigma-10}n$. This completes the proof of Lemma 5.8. $\square$

Lemma 5.10. *The number of bad vertices is at most $\frac{8}{\sigma-10}n$.*

Proof. We prove this claim by a double counting argument. Consider a graph D, with n vertices one for each vertex of the polygon and b additional "bad" vertices one for each interval $[p_i, q_i]$ corresponding to a bad vertex p_i. We construct a directed acyclic graph (DAG) by adding directed edges from the bad vertices to the rest (we allow directed edges between the bad vertices too). We will show that the indegree of each vertex is at most 4, while the outdegree of each bad vertex is at least $\frac{\sigma-2}{2}$. The lemma follows by noting that the sum of the indegrees is equal to the sum of the outdegrees, thus:

$$4(n + b) \geq b\frac{\sigma - 2}{2} \Rightarrow b \leq \frac{8}{\sigma - 10}n.$$

The construction of the graph D is quite intricate and we will do it in several steps. We say a vertex p_j is a *potential child* of $[p_i, q_i]$ iff $p_j \in [p_{i+2}, q_i)$. We say an interval $[p_j, q_j]$ is a potential child of $[p_i, q_i]$ iff it is a subinterval of $[p_i, q_i]$, and both of its endpoints are potential children of $[p_i, q_i]$. The directed edges of D are from a bad vertex to a *subset* of its potential children. Since

the edges are directed only from an interval to the intervals/vertices inside it, D will be a directed acyclic graph.

We add the directed edges inductively in a bottom up manner. Consider an interval $[p_i, q_i]$ and suppose we have added the edges between the potential children of $[p_i, q_i]$, such that the outdegree of all intervals is at least $\frac{\sigma-2}{2}$.

Let D_i be the induced subgraph of the set of potential children of $[p_i, q_i]$. Trivially D_i is also a directed acyclic graph. We show that it contains at least $(\sigma - 2)/2$ sources (vertices of indegree 0). Then we connect $[p_i, q_i]$ to some specific set of the sources of D_i.

Let $P_i = q_{i_0}, q_{i_1}, \ldots, q_{i_l}$ be the shortest chain assigned to vertex p_i. Since p_i is a bad vertex we have $l \geq \sigma$. Let σ' be the largest odd integer smaller than σ (i.e., $\sigma' := \sigma - \mathbb{I}[\sigma \text{ is even}]$). Define the set of vertices $S_i := \{q_{i_3}, q_{i_5}, \ldots, q_{i_{\sigma'}}\}$ in D_i. Note that S_i contains all vertices of P_i with odd index except q_{i_1}; this is a technical requirement and will be useful later in the proof of Claim 5.12.

In the next claim we show each source vertex in D_i has a directed path to at most one of the vertices of S_i. This implies that D_i contains at least $|S_i| = (\sigma' - 1)/2 \geq \frac{\sigma-2}{2}$ sources since either the vertex in S_i is a source or there is a unique source connected to it.

Claim 5.11. *Any source of D_i is connected by directed paths to at most one of the vertices of S_i.*

Proof. Let $[p_j, q_j]$ be a potential child of $[p_i, q_i]$, connected by directed paths to two vertices $q_{i_k}, q_{i_{k'}} \in S_i$, where $k + 2 \leq k'$. We show that using $[p_j, q_j]$, we can obtain a shorter chain for p_i, which is a contradiction.

First note the transitivity: if x is a potential child of y, and y is a potential child of z, then x is also a potential child of z. Since each interval is only adjacent to its potential children, and this property is transitive, q_{i_k} and $q_{i_{k'}}$ are potential children of $[p_j, q_j]$. Therefore, $q_{i_k}, q_{i_{k'}} \in [p_{j+2}, q_j) \subset (p_j, q_j)$. Hence, all the vertices between them in P_i, and in particular $q_{i_{k+1}}$, are also contained in (p_j, q_j).

Since $[p_j, q_j]$ is a potential child of $[p_i, q_i]$, $[p_j, q_j]$ is a subinterval of $[p_i, q_i]$, and $p_j \in [p_{i+2}, q_i)$. Moreover, since $C(p_j, q_j)$ is a near minimum cut, and at least three vertices of P_i are included in (p_j, q_j), by part 3 of Claim 5.9 we may obtain a shorter chain for p_i which is a contradiction. $\square$

Now we are ready to define the $\frac{\sigma'-1}{2}$ directed edges from (p_i, q_i) to its potential children: for each vertex $q_{i_j} \in S_i$, we add an edge from $[p_i, q_i]$ to one of the sources (i.e., vertices with indegree 0) in D_i that has a directed path to q_{i_j}.

It remains to show that after the construction of D the indegree of each vertex is at most 4. It is worth noting that, indeed some of the vertices may have indegree more than 1. As an example, suppose $[p_i, q_i]$, and $[p_{i+1}, q_{i+1}]$ are two bad intervals (note that q_{i+1} does not necessarily equal to the vertex next to q_i in the polygon). The crucial fact is that $[p_{i+1}, q_{i+1}]$ is not a potential child of $[p_i, q_i]$, even if it is a subinterval of $[p_i, q_i]$. Now suppose $[p_{i+1}, q_{i+1}]$ has a directed edge to a source vertex $x \in V(D_i) \cap V(D_{i+1})$. Since $[p_{i+1}, q_{i+1}] \notin V(D_i)$, the directed edge is not included in D_i. Hence x might seem to be a source vertex in D_i even though it already has an incoming edge.

First we show that if two intervals are both adjacent to a vertex of D_i, then they can not be crossing:

Claim 5.12. *Assume that two bad intervals $[p_i, q_i]$ and $[p_j, q_j]$ are both directly adjacent to a vertex x. Then we either have p_i is adjacent to p_j in the polygon, or $p_j, q_j \in [p_i, q_i)$, or $p_i, q_i \in [p_j, q_j)$.*

Proof. Assume that p_i is not adjacent to p_j (otherwise we are done). Since each bad vertex has exactly one associated interval in D, we also have $p_i \neq p_j$. First of all, note that by claim's

20

assumption, x is a potential child of both of the intervals. Therefore, $[p_i, q_i]$ and $[p_j, q_j]$ have a non-empty intersection. Without loss of generality, assume that $p_j \in [p_i, q_i)$ (the other cases are equivalent). We need to show that $q_j \in [p_i, q_i)$. Suppose not; let $P_i = q_{i_0}, \ldots, q_{i_l}$ be the shortest chain assigned to p_i. We show that we may modify P_i and obtain a shorter chain for p_i.

Since $[p_i, q_i]$ is adjacent to x, by definition x is a source in D_i that has a directed path to one of the vertices of S_i (say q_{i_k}) in D_i (note that x may be equal to q_{i_k}). Since $[p_j, q_j]$ is also adjacent to x, by transitivity, q_{i_k} is a potential child of $[p_j, q_j]$. Moreover, since $q_j \notin [p_i, q_i)$, all of the vertices $q_{i_1}, q_{i_2}, \ldots, q_{i_k}$ are also potential children of $[p_j, q_j]$ (note that since we may have $q_j = q_{i_0} = q_i$, q_{i_0} is not necessarily a potential child of $[p_j, q_j]$). Therefore, we have $q_{i_1}, q_{i_2}, \ldots, q_{i_k} \in (p_j, q_j)$. Since by construction of S_i, $k \geq 3$, at least 3 vertices of the chain P_i is contained in (p_j, q_j).

Since $p_j \in [p_i, q_i)$, but p_j is not equal to p_i, nor adjacent to it, we have $p_j \in [p_{i+2}, q_i)$. Moreover, since $q_j \notin [p_i, q_i)$, $[p_j, q_j]$ crosses $[p_i, q_i]$. Therefore, since $[p_j, q_j]$ is a near minimum cut that contains three consecutive vertices of P_i, by part 3 of Claim 5.9 we may obtain a shorter chain for p_i, which is a contradiction. $\square$

Now we can show that the indegree of each vertex in D is at most 4:

Claim 5.13. *The indegree of each vertex of D is at most 4.*

Proof. We prove the claim by contradiction. Let $x \in V(D)$ be a vertex with indegree at least 5. We show that one of the 5 intervals adjacent to x is indeed a potential child of another one, and thus x was not a source vertex for at least one of the induced DAGs associated to one of these intervals, which is a contradiction.

First of all, since each bad vertex has exactly one associated interval in D, the 5 intervals must start at distinct vertices of the polygon. Therefore, among these 5 intervals we can find 3 intervals $[p_i, q_i], [p_j, q_j], [p_k, q_k]$ such that no two of p_i, p_j, p_k are adjacent in the polygon.

By Claim 5.12 these three intervals may not cross. Without loss of generality, let $p_j, q_j, p_k, q_k \in [p_i, q_i)$. Therefore, since p_j, p_k are not equal to p_i, nor adjacent to it, we have $p_j, p_k \in [p_{i+2}, q_i)$. Therefore, $[p_j, q_j], [p_k, q_k]$ are potential children of $[p_i, q_i]$ unless they are not a subinterval of $[p_i, q_i]$ (i.e. $q_j \in [p_i, p_j)$, and $q_k \in [p_i, p_k)$).

Suppose they are not a subinterval of $[p_i, q_i]$; similarly we can assume $p_k, q_k \in [p_j, q_j)$, and $p_k \in [p_{j+2}, q_j)$. But since $[p_j, q_j], [p_k, q_k]$ are not a subinterval of $[p_i, q_i]$, $[p_k, q_k]$ must be a subinterval of $[p_j, q_j]$. Therefore, $[p_k, q_k]$ is a potential child of $[p_j, q_j]$. This means that $[p_k, q_k] \in V(D_j)$ was adjacent to x, and hence x has an indegree 1 in D_j which contradicts with the fact that $[p_j, q_j]$ is adjacent to x. $\square$

This completes the proof of Lemma 5.10. $\square$

5.1 Cut Classes with Inside Atoms

In this subsection we show that inside atoms (vertices) of the polygon representation do not have a great impact on the number of pairs of vertices that are highly connected. First, we prove the crucial lemma 5.16 which shows that the inside atoms are loosely connected to each other. We use this lemma in Corollary 5.17 to extend the proof of Lemma 5.8 in the case where we have inside atoms as well. Then we will show that the number of inside atoms is at most an $O(\sqrt{\delta})$ fraction of outside atoms. Let V_{in} and V_{out} be the set of inside/outside vertices of $V(H)$ respectively. Since by Corollary 5.17 there are at least $m := (1 - \frac{8}{\sigma - 10})|V_{out}|$ pairs of outside vertices of H such that the

vertices in each pair are connected by at least $(1-2\sigma\delta)\frac{c}{2}$ number of edges, and by Corollary 5.19 we have $|V_{in}| \le O(\sigma\delta|V_{out}|)$, setting $\sigma := 2/\sqrt{\delta}$, we will obtain that there are at least $m := (1 - 20\sqrt{\delta})|V|$ pairs of vertices of H, such that the vertices in each pair are connected by at least $(1-4\sqrt{\delta})\frac{c}{2}$ edges.

We start by describing the properties of inside atoms. We use the results in [6] to distinguish between inside and outside vertices of H. We need to use the definition of k-cycles:

Definition 5.14 (Benczur [6, Definition 5.3.1]). *We say that k cut sides $C_i \subset V$, for $1 \le i \le k$, form a k-cycle if*

- *C_i crosses both C_{i-1} and C_{i+1};*

- *$C_i \cap C_j = \emptyset$ for $j \ne i - 1, i$ or $i + 1$; and*

- *$\bigcap_{1 \le i \le k} \overline{C_i} \ne \emptyset$.*

One of the main differences between the properties of the collections of minimum and near minimum cuts is that minimum cuts do not admit any k-cycle. Benczur in [6] showed that the collection $(1 + \delta)$ near minimum cuts does not have a $1/\delta$-cycle. In particular, $6/5$ near minimum cuts do not have any k-cycle for $k \le 5$. He also showed that the set of inside atoms of the polygon representation are those atoms a, for which there exists a vertex subset W and a collection $\{C \cap W : (C, \overline{C}) \in \mathcal{C}_i, a \nsubseteq C\}$ that form a k-cycle for $k \ge 3$. It is not hard to see that if the set of cuts $\mathcal{C}$ possesses a polygon representation, then we have $W = V$ for any cycle associated to an inside atom. We say that a k-cycle is for an inside atom a, if $a \in \bigcap_i \overline{C_i}$.

Here is the strategy for the rest of this section. We will prove in Lemma 5.16 that the minimum cut within any subset of inside atoms is quite small. We use that to prove Corollaries 5.17 and 5.19. But before that, we state the following technical lemma whose proof is deferred to the appendix.

Lemma 5.15. *Let $T \subset V_{in}$ be a subset of inside vertices of H. Then the size of the minimum cut $\min_{S \subset T} |E(S, T \setminus S)| \le 12\delta\frac{c}{2} = 6\delta c$.*

Now we are ready to show that most of the edges adjacent to the inside atoms go to the outside atoms:

Lemma 5.16. $|E(V_{in})| \le 6\delta|V_{in}|c$.

Proof. The proof strategy is similar to Lemma 5.4. Define $\mathcal{S}$ as a collection of sets which are a partitioning of V_{in}. We start from $\mathcal{S} = \{V_{in}\}$, and iteratively break one of the sets in $\mathcal{S}$ into two sets through its minimum cut. We keep doing this until $\mathcal{S}$ becomes all singletons. By lemma 5.15 the size of the minimum cut of any subset of inside vertices is no more than $12\delta\frac{c}{2}$. Therefore, the total number of the edges of H encountered in the process will be at most $(|V_{in}| - 1)12\delta\frac{c}{2}$. Since each edge between the inside vertices of H is seen exactly once in the process, the number of edges between the inside vertices of H is no more than $12\delta|V_{in}|\frac{c}{2}$. $\square$

With the following two corollaries we complete the proof of Lemma 5.6. First we show how to extend the proof of Lemma 5.8 to the polygon representation of the cut classes that contain inside atoms:

Corollary 5.17. *For any integer $\sigma \ge 12$, there exists at least $m = (1 - \frac{8}{\sigma-10})|V_{out}|$ pairs of vertices of H, $\{(v_1, u_1), (v_2, u_2), \ldots, (v_m, u_m)\}$ such that for each $1 \le i \le m$, $|E(v_i, u_i)| \ge (1 - 2\sigma\delta)\frac{c}{2}$, and each vertex $v \in V$ is contained in at most two such pairs.*

Proof. We essentially use the same proof strategy of Lemmas 5.8 and 5.10. The main difference is that because of the existence of inside vertices the cuts $C(p_i, p_j)$ considered in the proof is not well defined (recall that $C(p_i, p_j)$ is not necessarily a near minimum cut in H, thus it may not correspond to a representing diagonal of the polygon representation). In fact depending on the set of inside vertices contributed to different sides of the cut $C(p_i, p_j)$, we get cuts of different sizes. For an interval $[p_i, p_j]$, define $I(p_i, p_j) := \text{argmin}_{I \subseteq V_{in}} |d(\{v_i, \ldots, v_{j-1}\} \cup I)|$ to be a subset of inside vertices which together with the outside vertices in the interval $[p_i, p_j]$ makes the minimum possible cut, and let

$$C(p_i, p_j) := d(\{v_i, v_{i+1}, \ldots, v_{j-1}\} \cup I(p_i, p_j)),$$

be that cut. It is not hard to see that the whole argument would work assuming this new definition, and we can show that for most of vertices $|C(p_i, p_{i+2})| \leq c(1 + 2\sigma\delta)$. Unfortunately, $|C(p_i, p_{i+2})|$ being small does not immediately imply that $|E(v_i, v_{i+1})|$ is large, since the cut $C(p_i, p_{i+2})$ may contain inside vertices. In the rest of the proof we show that for any vertex p_i of the polygon representation, if $|C(p_i, p_{i+2})| \leq c(1 + 2\sigma\delta)$, then $|E(v_i, v_{i+1})| \geq (1 - 2\sigma\delta)\frac{c}{2}$.

We prove the statement by contradiction; suppose $|C(p_i, p_{i+2})| \leq c(1 + 2\sigma\delta)$, and $|E(v_i, v_{i+1})| < (1 - 2\sigma\delta)\frac{c}{2}$. Let $I := I(p_i, p_{i+2})$, and $S := I \cup \{v_i, v_{i+1}\}$ (note that $C(p_i, p_{i+2}) = E(S, \overline{S})$). Trivially $I \neq \emptyset$. We show that $|E(S)| \leq (1 - 2\sigma\delta + 20|I|\delta)\frac{c}{2}$. Then we use the fact that H is c-edge connected to argue that the degree of each vertex is at least c, thus the number of the edges adjacent to vertices of S is at least $c|S|$. This implies that $|C(p_i, p_{i+2})| \gg c(1 + 2\sigma\delta)$, unless $I = \emptyset$ which is a contradiction.

First we show that the $|E(I, \{v_i, v_{i+1}\})| \leq 4\delta|I|c$.

Claim 5.18. *For any inside vertex $v_{in} \in V_{in}$ and $v_{out} \in V_{out}$ we have $|E(v_{in}, v_{out}))| \leq 2\delta c$.*

Proof. Let $C = \{C_1, C_2, \ldots, C_k\}$ be a k-cycle for v_{in}. Since v_{out} is an outside vertex, C is not a cycle for v_{out}, and thus $v_{out} \in C_i$ for some $1 \leq i \leq k$. Since C_i crosses C_{i-1} and C_{i+1}, we have $|E(C_i, C_{i-1})| \geq (1 - \delta)\frac{c}{2}$ and $|E(C_i, C_{i+1})| \geq (1 - \delta)\frac{c}{2}$. Therefore,

$$|E(v_{in}, v_{out})| \leq |E(v_{in}, C_i)| \leq c(1 + \delta) - c(1 - \delta) = 2\delta c.$$

$\square$

By the above claim and Lemma 5.16 we have

$$|E(S)| = |E(I, v_i)| + |E(I, v_{i+1})| + |E(I)| + |E(v_i, v_{i+1})| \leq 4\delta|I|c + 6\delta|I|c + (1 - 2\sigma\delta)\frac{c}{2}$$

Since $|d(S)| = |C(p_i, p_{i+2})| \leq c(1 + 2\sigma\delta)$ we get

$$c(|I| + 2) - c(1 - 2\sigma\delta + 20|I|\delta) \leq |d(S)| \leq c(1 + 2\sigma\delta) \Rightarrow |I| \leq 20\delta < 1,$$

where the last inequality holds by the fact that $\delta < 1/20$. Therefore $I = \emptyset$ and we get a contradiction.

$\square$

It remains to prove an upperbound on the number of inside vertices:

Corollary 5.19. *The number of inside vertices is no more than $|V_{in}| \leq \frac{1}{1-27\delta}|V_{out}|\left[(2\sigma + 3)\delta + \frac{8}{\sigma-10}\right]$.*

Proof. We use a double counting argument. Let E_{in} be the set of edges between the inside vertices, E_{out} be the set of edges between the outside vertices and E_{in-out} be the set edges from the inside to outside vertices. By Lemma 5.16 we have $|E_{in}| \leq 6\delta|V_{in}|c$. Since the degree of each inside vertex is at least c, we have

$$|E_{in-out}| \geq c(1 - 12\delta)|V_{in}|. \tag{1}$$

Let $\sigma \geq 12$ be an integer (it will be fixed later), by lemma 5.17, there are $m := (1 - \frac{8}{\alpha-10})|V_{out}|$ pairs of outside vertices such that the vertices in each pair are connected by at least $(1 - 2\sigma\delta)\frac{c}{2}$ edges. We have

$$|E_{out}| \geq \frac{c}{2}(1 - 2\sigma\delta)(1 - \frac{8}{\sigma - 10})|V_{out}|. \tag{2}$$

Finally, by lemma 5.4 we have

$$|V|(1 + 3\delta)\frac{c}{2} \geq |E_{in}| + |E_{out}| + |E_{in-out}|. \tag{3}$$

By combining equations (1),(2), and (3) we obtain:

$$[|V_{in}| + |V_{out}|](1 + 3\delta)\frac{c}{2} \geq c(1 - 12\delta)|V_{in}| + \frac{c}{2}(1 - 2\sigma\delta)(1 - \frac{8}{\sigma - 10})|V_{out}|$$

Therefore

$$|V_{in}| \leq \frac{1}{1 - 27\delta}|V_{out}|\left[(2\sigma + 3)\delta + \frac{8}{\sigma - 10}\right]$$

$\square$

Now we may complete the proof of Lemma 5.6:

Proof of Lemma 5.6. By Corollary 5.17 there are $m := (1 - \frac{8}{\sigma-10})|V_{out}|$ pairs of vertices of H that are connected by at least $(1 - 2\sigma\delta)\frac{c}{2}$ edges (for any integer $\sigma \geq 12$). Also by Corollary 5.19, we have $|V_{in}| \leq \frac{|V|}{1-27\delta}\left[(2\sigma + 3)\delta + \frac{8}{\sigma-10}\right]$. Therefore, there is at least

$$(1 - \frac{8}{\sigma - 10})|V_{out}| \geq \left(1 - \frac{8}{\sigma - 10} - \frac{1}{1 - 27\delta}\left[(2\sigma + 3)\delta + \frac{8}{\sigma - 10}\right]\right)|V|$$

pairs of vertices of H such that each pair is connected by at least $(1 - 2\sigma\delta)\frac{c}{2}$ edges. Since $\delta < \frac{1}{100}$, setting $\sigma = \sqrt{4/\delta}$ and rounding it down to the nearest integer implies the Lemma. $\square$

This completes the proof of Theorem 4.9.

5.2 Cyclic Cut Classes

In this subsection we prove a much stronger statement (cf. Theorem 4.9) on small cut classes. For a small cut class C_i, let $\kappa(C_i) := 10|\phi(C_i)| \times \delta$.

We say a cut class is **cyclic** if it is not the root of $\Gamma(C)$, and

$$20|\phi(C_i)|\kappa(C_i) = 200|\phi(C_i)|^2 < 1$$

(note that τ is picked to ensure that each small cut class is cyclic). The following proposition shows nice properties of the cyclic cut classes.

Proposition 5.20. *Let C be the collection of $(1+\delta)$ near minimum cuts of fractional graph G, and let C_i be a cut class of C. If $\kappa(C_i) := 10|\phi(C_i)| \times \delta < 1$, and $\delta < 1/24$, then:*

1. *C_i does not contain any inside atoms,*

2. *For any $a \in \phi(C_i)$, the fractional degree $x(d(a)) \leq 2(1 + 2\delta(|\phi(C_i)| - 2)) \leq 2(1 + \kappa(C_i))$,*

3. *For any pair of adjacent outside atoms a, b of the polygon representation of C_i we have $x(E(a,b)) \geq 1 - 2\delta(|\phi(C_i)| - 3) \geq 1 - \kappa(C_i)$.*

Similar to the rest of the proofs in this section, we just look at the graph $H(\phi(C_i))$, and we prove the above proposition for such a graph when $10|\phi(C_i)|\delta < 1$. This is proved in the following Lemma:

Lemma 5.21. *Let $c \geq 1$ be an integer, $\delta < 1/100$, and $H = (V, E)$ be an unweighted c-edge connected graph such that for any non-trivial cut $(S, \overline{S})$, there exists a $(1+\delta)$ near minimum cut that crosses $(S, \overline{S})$. If the cross graph of $(1+\delta)$ near minimum cuts of H is connected, and $4|V|\delta < 1$, then:*

1. *H does not have any inside vertices in the polygon representation,*

2. *For any vertex v, we have $|d(v)| \leq c(1 + 2\delta(|V| - 2))$,*

3. *For any adjacent pair of vertices v_i, v_{i+1} in the polygon representation, we have $|E(v_i, v_{i+1})| \geq \frac{c}{2}(1 - 2\delta(|V| - 3))$.*

Proof. First we show that $V_{in} = \emptyset$. The proof simply follows from Lemma 5.16. Suppose $V_{in} \neq \emptyset$, by Lemma 5.16, there exists an inside vertex $v \in V_{in}$ which has at most $12\delta c$ edges to other inside atoms. Since $|d(v)| \geq c$, and $\delta < 1/24$, we have $|E(v, V_{out})| \geq c - 12\delta c \geq \frac{c}{2}$. Finally, by Claim 5.18, for any $v' \in V_{out}$ we have $|E(v, v')| \leq 2\delta c$. But

$$|E(v, V_{out}) = \sum_{v' \in V_{out}} |E(v, v')| \leq 2\delta|V_{out}|c \leq 2\delta|V|c < \frac{c}{2},$$

which is a contradiction. Note that the last inequality holds by Lemma's assumption.

In the rest of the proof we assume $V_{in} = \emptyset$, thus $V = V_{out}$. The proof of the other two statements follows from the chain construction argument of Lemma 5.8. Recall that $p_1, \ldots, p_n$ are the vertices of the polygon representation, and outside atoms are placed on the sides of the polygon such that v_i is placed on the edge (p_i, p_{i+1}). First we prove the third statement of the lemma; we show for any i, $|E(v_i, v_{i+1}| \geq \frac{c}{2}(1 - 2\delta(|V| - 3))$. It turns out that the longest possible chain for p_i is $P_i := p_{i-2}, p_{i-3}, \ldots, p_{i+3}, p_{i+2}$. (This follows from the fact that each vertex appears in a chain at most once. Since by Claim 5.9 any vertex v_i has at least one chain, the length of the shortest chain for v_i is at most $|V| - 3$. Therefore, by the second part of Claim 5.9 we have $|E(v_i, v_{i+1})| \geq \frac{c}{2}(1 - 2\delta(|V| - 3))$.

It remains to prove the second statement. For that we need to perturb the definition of a chain and require that the chain of v_i ends at $q_{i_l} = p_{i+1}$ instead of $q_{i_l} = p_{i+2}$. Let P_i be a chain for p_i. Observe that by inserting p_{i+1} at the end of P_i we obtain a feasible chain, say P_i', that ends at p_{i+1}. Also by following the lines of the proof of the second statement of Claim 5.9, we can show that if $length(P_i) = l$, then the size of the cut $|d(v_i)| \leq c(1 + 2l\delta)$. In other words, the degree of v_i is no more than $c(1 + 2l\delta)$. Also using the proof in the previous paragraph we get $length(P_i') \leq |V| - 2$. Therefore, $|d(v_i)| \leq c(1 + 2\delta(|V| - 2))$. $\qquad\square$

6 Existence of Good Edges

In this section, we prove the three types of edges defined in subsection 4 are good. Before that, we state some of the properties of random spanning trees and their corresponding generating functions. The results stated in the next subsection are not new. But we include them for the sake of completeness.

6.1 Random Spanning Trees and Strongly Rayleigh Measures

Let us start by a beautiful result of Burton and Pemantle that derives the exact probability that a set of edges appear in a random spanning tree as a determinant. Suppose G is an electrical network in which the conductance of every edge e is λ_e. Recall that λ_e was selected such that $\mathbb{P}[T \sim \mu] \propto \prod_{e \in T} \lambda_e$. With an abuse of notation, in this section we use T (instead of F) to denote the sampled spanning tree from the distribution μ.

Let e and f be two edges of G. Define $i^e(f)$ to be the current that flows across f when a unit current is imposed between the endpoints of e. It is known that $i^e(e) = z_e$ is the probability that e is chosen in the λ-uniform spanning tree. Burton and Pemantle [11] proved the following, known as Transfer-Current Theorem which also gives the exact value of correlation between any two edges in the graph.

Theorem 6.1 (Burton, Pemantle [11])**.** *For any distinct edges $e_1, \ldots, e_k \in G$,*

$$\mathbb{P}[e_1, \ldots, e_k \in T] = det[i^{e_i}(e_j)]_{1 \leq i, j \leq k}.$$

In particular, for any two edges e and f:

$$\mathbb{P}[e, f \in T] - \mathbb{P}[e \in T]\mathbb{P}[f \in T] = -i^e(f)i^f(e)$$

The next set of results are the main tools in proving the goodness of the trivial edges and are mostly restated from a recent work of Borcea, Branden and Ligget [8] on strongly Rayleigh probability measures. Strongly Rayleigh measures include determinantal measures (in a form similar to the above formula) and in particular uniform and λ-uniform random spanning tree measures. They also enjoy all the virtues of negative dependence and negative association.

Let E be the ground set of elements with $m = |E|$. For an element $e \in E$, let X_e be the indicator random variable for element e, and for $S \subseteq E$, let $X_S = \sum_{e \in S} X_e$. Let $\mathcal{P}_m$ be the set of all multi-affine polynomial in m variables $f(t_e : e \in E)$ with non-negative coefficients such that $f(1, 1, \ldots, 1) = 1$. There is a one-one correspondence between the measures on 2^E, and $\mathcal{P}_m$: if μ is a measure we may form its generating polynomial, namely $f(t) = \sum_{S \subseteq E} \mu(S)t^S$, where $t^S = \prod_{e \in S} t_e$.

A polynomial $f \in \mathbb{C}[t_e]_{e \in E}$ is called *stable* if $f(t_e : e \in E) \neq 0$ whenever $Im(t_e) > 0$ for all $e \in E$. A stable polynomial with all real coefficients is called *real stable*. For example, this simply implies that a polynomial in one variable is real stable iff all its roots are real. A measure μ on 2^E is called *strongly Rayleigh* if its generating function is real stable.

First we describe some operations (Projection, Conditioning and Truncation) that maintain the strongly Rayleigh property.

Definition 6.2 (Projection)**.** *Given $S \subset E$ and μ a measure on 2^E. The projection of μ onto 2^S is the measure μ' obtained from μ by restricting the samples to the subsets of S, i.e.:*

$$\forall A \subseteq S : \mu'(A) := \sum_{B \subseteq E : B \cap S = A} \mu(B)$$

Borcea et al. showed that any projection of a strongly Rayleigh measures is still a strongly Rayleigh measure. For example, if μ is a uniform measure on the spanning trees of $G = (V, E)$, and $S = (C, \overline{C}) \subset E$ is the set of edges in the cut $(C, \overline{C})$, the projection of μ on 2^S is a strongly Rayleigh measure.

Definition 6.3 (Conditioning). *Let μ be a measure on 2^E, fix some $e \in E$. The measure obtained from μ by conditioning on $X_e = 0$ is the measure μ' on $2^{E \setminus \{e\}}$ defined as follows:*

$$\forall A \subseteq E \setminus \{e\} : \mu'(A) := \frac{\mu(A)}{\sum_{B \subseteq E \setminus \{e\}} \mu(B)}$$

Similarly, we can define the measure obtained by conditioning on $X_e = 1$.

For example, if μ is a uniform measure on the spanning trees of $G = (V, E)$, the measure obtained by conditioning on $X_e = 0$ for some $e \in E$ is still a uniform measure on spanning trees that does not contain e, and it is still a strongly Rayleigh measure. Similarly, we can condition on the set of spanning trees that contain all of the edges in a set S, and none of the edges in S'. As a generalization, let $W \subset V$ be a set of vertices. Observe that the measure obtained by conditioning on $\sum_{e \in E(W)} X_e = |W| - 1$ (i.e. having an spanning tree inside W), is still a uniform measure on spanning trees.

Borcea et al. proved something stronger, they showed if we condition on $\sum_e X_e = p$, for some fixed number p, still we get a strongly Rayleigh measure. First we need to define the truncation of a measure:

Definition 6.4 (Truncation). *Let μ be a probability measure on 2^E, and $1 \le p \le q \le |E|$. The **truncation** of μ to $[p, q]$ is the conditional measure*

$$\mu_{p,q} := (\mu \mid p \le X_E \le q)$$

Borcea et al. proved that if $q - p \le 1$, the truncation of any strongly Rayleigh measure is still strongly Rayleigh:

Theorem 6.5 ([8, Corollary 4.18]). *Suppose that μ is a strongly Rayleigh probability measure on 2^E and that $0 \le p \le q \le |E|$ with $q - p \le 1$. Then $\mu_{p,q}$ is strongly Rayleigh.*

For example, let μ be a uniform measure on the spanning trees of $G = (V, E)$, and $S \subset E$; a truncation of the projection of μ on 2^S is strongly Rayleigh. Let μ^1 be the projection of μ on S. For any $1 \le p \le q \le |S|$ such that $q - p \le 1$, $\mu_{p,q}^1$ is a strongly Rayleigh measure. Moreover, since any spanning tree sampled from μ has exactly $|V| - 1 = n - 1$ edges, conditioning on $X_S = p$ is equivalent to $X_{\overline{S}} = n - 1 - p$. Therefore, conditioned on $p \le X_{\overline{S}} \le q$, μ^1 is a strongly Rayleigh measure too.

Next we describe some properties of the strongly Rayleigh measures that are essential in our proofs. We start with the negative association.

Definition 6.6 (Negative Association). *A measure μ on 2^E is called negatively associated or NA if*

$$\mathbb{E}_\mu[F] \mathbb{E}_\mu[G] = \int F d\mu \int G d\mu \ge \int FG d\mu = \mathbb{E}_\mu[FG]$$

for any increasing functions F, G on 2^E that depend on disjoint sets of elements.

Feder and Mihail [19] proved that uniform measures on balanced matroids (and in particular on spanning trees) have negative association. Borcea et al. in [8] proved that the strongly Rayleigh measure satisfy the strongest form of negative association known as *CNA+* which we do not describe here.

Theorem 6.7 ([8]). *Strongly Rayleigh measures are negatively associated*

We may associate an increasing function to any increasing event, and then use the properties of negative association:

Definition 6.8 (Increasing Events and Functions). *An* **increasing event** $\mathcal{A}$ *on* 2^E *is a collection of subsets of E that is closed upwards under containment, i.e. if $A \in \mathcal{A}$ and $A \subseteq B \subseteq E$, then $B \in \mathcal{A}$.*

For any increasing event $\mathcal{A}$, the function $f : 2^E \to \{0, 1\}$, defined as $f(A) := \mathbb{I}[A \in \mathcal{A}]$ is an increasing function. The following fact is a simple application of negative association.

Fact 6.9. *Let μ be a uniform measure on spanning trees of $G = (V, E)$. For any $S \subset E$, and $p \in \mathbb{R}$ we have*

1. $\forall e' \in E \setminus S : \mathbb{E}_\mu \left[X_{e'} \middle| X_S \geq p \right] \leq \mathbb{E}_\mu \left[X_{e'} \right]$

2. $\forall e' \in E \setminus S : \mathbb{E}_\mu \left[X_{e'} \middle| X_S \leq p \right] \geq \mathbb{E}_\mu \left[X_{e'} \right]$

The following corollary is a simple consequence of this:

Corollary 6.10. *Let μ be a uniform measure on spanning trees of a graph $G = (V, E)$, $S \subset E$, $s = \mathbb{E}_\mu [X_S]$, and $r(S)$ be the rank of S in the graphical matroid on G. For any set $S' \subseteq \overline{S}$, we have $\mathbb{E}_\mu [X_{S'} | X_S = 0] \leq \mathbb{E}_\mu [X_{S'}] + s$, and $\mathbb{E}_\mu [X_{S'} | X_S = r(S)] \geq \mathbb{E}_\mu [X_{S'}] - r(S) + s$.*

Proof. First of all since μ is a measure on spanning trees, and each spanning tree has $n - 1 = |V| - 1$ vertices, we have $\mathbb{E} \left[X_{\overline{S}} | X_S = 0 \right] = \mathbb{E} \left[X_{\overline{S}} \right] + s$, and $\mathbb{E} \left[X_{\overline{S}} | X_S = r(S) \right] = \mathbb{E} \left[X_{\overline{S}} \right] - r(S) + s$. The rest of the proof simply uses the negative association. Since any spanning tree selects at least zero, and at most $r(S)$ edges from S, the events $X_S = 0$, and $X_S = r(S)$ are downward, and upward closed event respectively. Hence, by negative association the probability of the edges in $\overline{S}$ can only increase/decrease respectively. Therefore, we have $\mathbb{E}[X_{S'} | X_S = 0] \leq \mathbb{E}[X_{S'}] + s$, and $\mathbb{E}[X_{S'} | X_S = r(S)] \geq \mathbb{E}[X_{S'}] - r(S) + s$. $\qquad \square$

The next property is ultra log-concavity (ULC) of the rank function of strongly Rayleigh measures. Recall that since there is negative correlation between the elements of a strongly Rayleigh measures, we may apply any standard concentration bounds like Chernoff Bounds, and say the $\sum_{e \in E} X_e$ is highly concentrated around its expectation. ULC in a sense is a generalization of that, it implies that the probability density of the rank sequence is a log-concave function (Note that not all measures with negative correlation have ULC rank functions):

Definition 6.11 (Ultra Log Concavity [8, Definition 2.8]). *A real sequence $\{a_k\}_{k=0}^m$ is log-concave if $a_k^2 \geq a_{k-1} a_{k+1}$, $1 \leq k \leq m - 1$, and it is said to have no internal zeros if the indices of its non-zero terms form an interval (of non-negative integers). We say that a non-negative sequence $\{a_k\}_{k=0}^m$ is*

- **LC** *if it is log-concave with no internal zeros;*

28

- **SLC** *(strongly log-concave) if the sequence* $\{k!a_k\}_{k=0}^m$ *is LC;*

- **ULC** *(ultra log-concave) if the sequence* $\{a_k/\binom{m}{k}\}_{k=0}^m$ *is LC.*

Clearly, ULC $\Rightarrow$ SLC $\Rightarrow$ LC. *If μ is a measure on 2^E, then $\mu(X_E = k)_{k=0}^{|E|}$ is called the* **rank sequence** *of μ. A measure is said to be ULC, SLC or LC if its rank sequence is ULC, SLC or LC, respectively.*

Let μ be a strongly Rayleigh measure, and $f(t)$ be the generating polynomial of μ. The *diagonal specialization* of μ obtained by pretending $f(t)$ as a univariate polynomial (i.e., considering $f(t, t, \ldots, t)$). This polynomial is indeed the generating polynomial of the rank sequence of μ. It simply follows that if $f(t_e : e \in E)$ is a stable polynomial then so is its diagonal specialization [8]. Since a univariate polynomial with real coefficients is stable iff all of its roots are real, $f(t)$ is a polynomial with real roots.

Therefore, the coefficients of $f(t)$ are corresponding to the probability density function of the convolution of a set of independent Bernoulli random variables [35]. In other words, they give the probability distribution of the number of successes in m independent trials for a sequence of success probabilities $p_1, \ldots, p_m$:

Fact 6.12. *[8, 35] The rank sequence of a strongly Rayleigh measure is the probability distribution of the number of successes in m independent trials for some sequence of success probabilities $p_1, \ldots, p_m$.*

The distribution of the number of successes of n independent trials is well studied in the literature [14, 27, 22, 38, 35]. Dorrach [14] proved that these distributions are unimodal, and the mode differs from the mean by less than 1 (recall that the mode is the value at which its probability mass function takes its maximum value). Moreover, according to Newton's inequality [25], these distributions are Ultra Log-concave. Hence,

Theorem 6.13 ([25, 14, 8]). *The rank sequence of any strongly Rayleigh measure is ULC, unimodal, and its mode differs from the mean by less than 1.*

In general, let $\mathcal{D}_m(p)$ be the set of all distributions of the sum of m independent trials where the probability of success in the ith trial is $0 \le p_i \le 1$ for $i = 1, \ldots, m$, and the expected number of successes is p. Let X be the number of successes. It is well known that the maximum of the variance of X is attained when $p_1 = p_2 = \ldots = p_m = p/m$ (cf. e.g. [27]). For a given m, here we are interested in lower bounding the probability that the number of successes is some fixed number r, over all distributions $\mathcal{D}_m(p)$ for $m \ge 1$, i.e.,

$$B(p, r) := \min_{m, \mu \in \mathcal{D}_m(p)} \mathbb{P}_\mu[X = r]. \tag{4}$$

The function $B(p, r)$ will be used in several lemmas in this section. As an example, we can lower bound the probability that exactly two edges are sampled from a $(1 + \delta)$ near minimum cut of G, by choosing $p = x(E(S, \overline{S}))$, and $r = 2$ (i.e., $\mathbb{P}\left[T \cap E(S, \overline{S}) = 2\right] \ge B(x(E(S, \overline{S})), 2)$).

First, observe that if $|r - p| \ge 1$, then the distribution with $m := \lceil p \rceil$ trials such that $\lfloor p \rfloor$ of p_i's are equal to 1 implies $B(p, r) = 0$. Therefore, here we focus on the cases where $|r - p| < 1$. Hoeffding in [27, Corollary 2.1] shows that the distribution that minimizes $\mathbb{P}[X = r]$ is the one where all p_i's take only one of the three different values, only one of which is different from 0 and 1. Since here we are minimizing over all distributions with any arbitrary number of trials $m \ge 0$, we can further

assume that none of the p_i's are 0. Let $\mu \in \mathcal{D}_m(p)$ be the optimum distribution, and suppose for $k \leq p, r$, $r - k$ of the p_i's in μ are equal to 1, We can deduce that the rest of the p_i's must be equal to $\frac{p-(r-k)}{m-(r-k)}$. Putting these together we get:

$$B(p,r) = \min_{m,k: k \leq p, r} \binom{m}{k} \left(\frac{p - (r-k)}{m} \right)^k \left(1 - \frac{p - (r-k)}{m} \right)^{m-k} \tag{5}$$

Therefore it is sufficient to find the optimum values of m, k. This is done in the following theorem:

Lemma 6.14. *For $r - 1 < p \leq r$, the optimum distribution is obtained by letting $m \to \infty$, and $k = r$, or 1:*

$$r - 1 < p \leq r: \quad B(p,r) = \min \left(e^{-(p-r+1)}(p - r + 1), \frac{e^{-p}p^r}{r!} \right). \tag{6}$$

On the other hand, if $r \leq p < r + 1$, the optimum distribution is obtained by letting $k = r$, and $m = r + 1$, or $m = \infty$:

$$r \leq p < r + 1: \quad B(p,r) = \min \left((r+1) \left(\frac{p}{r+1} \right)^r \left(1 - \frac{p}{r+1} \right), \frac{e^{-p}p^r}{r!} \right) \tag{7}$$

The proof of the above Lemma is left for the final version of this paper.

Remark 6.15. *Note that using simple Chernoff bounds, log-concavity, it follows that for constant r, $B(p,r)$ remains constant. The above lemma only characterizes the optimal lower bound.*

The last useful property is the stochastically dominance property on truncations of strongly Rayleigh measures.

Definition 6.16 ([8, Definition 2.14]). *Let μ, ν be two measures defined on 2^E. We say μ **stochastically dominates** ν ($\nu \preceq \mu$) if for any increasing event $\mathcal{A}$ on 2^E, we have $\mu(\mathcal{A}) \geq \nu(\mathcal{A})$.*

Borcea et al. showed that a truncation of strongly Rayleigh measures is stochastically dominated by a truncation of a larger value:

Theorem 6.17 ([8, Theorem 4.19]). *Let μ be a strongly Rayleigh probability measure on 2^E, and let $1 \leq k \leq |E|$. If $\mu(\{S : |S| = k - 1\})\mu(\{S : |S| = k\}) \neq 0$, then $\mu_{k-1} \preceq \mu_k$.*

As an example, let μ be the uniform measure on spanning trees of $G = (V, E)$, and $A \subset S \subset E$. Let μ' be the projection of μ on 2^S. Since μ' is strongly Rayleigh, we have $\mu'_k \preceq \mu'_{k+1}$, for any integer $k \geq 0$, where μ'_k, μ'_{k+1} are well defined. Therefore, we have

$$\mathbb{P}_{\mu'_{k+1}} [X_A \geq k'] \geq \mathbb{P}_{\mu'_k} [X_A \geq k']$$

for $k' \in \mathbb{R}$.

The rest of this section is organized as follows: in subsection 6.2 we show that certain trivial edges are good. In subsection 6.3 we show that any inside edge of a small non trivial cut class is good. Finally, in subsection 6.4 we prove an unbalanced thread of length 10 is assigned good edges of a constant fraction.

6.2 Trivial Good Edges

In this subsection we show that edges of fraction bounded away from $1/2$, that are included only in trivial cuts are good. Recall that $d(u) := \{(u, v) \in E\}$ is the set of edges with u as one endpoint. Let $d_w(u) := \{(u, v) : v \neq w\}$, be the set of edges incident to u, other than the edge (u, w).

Proposition 6.18. *Let u and v be two vertices of G of fractional degree 2, $e = (u, v)$, and $\epsilon < 1/100$. If $x_e < \frac{1}{2} - \epsilon$ or $x_e > \frac{1}{2} + \epsilon$, then $\mathbb{P}_{T \sim \mu}[\mathcal{E}(e)] \geq \frac{\epsilon}{100}$.*

Proof. Let $X := |T \cap d_v(u)|$ and $Y := |T \cap d_u(v)|$ be the random variables indicating the number of edges of $d_v(u)$ and $d_u(v)$ that are sampled in $T \sim \mu$, and $Z := |T \cap \{e\}|$ be the indicator random variable for e. If $x_e < 1/2 - \epsilon$, then from negative association (see Fact 6.9) we have $\mathbb{E}_{T \sim \mu}[X + Y | Z = 0] \geq \mathbb{E}_{T \sim \mu}[X + Y] \geq 3 + 2\epsilon$ and from Corollary 6.10, we have $\mathbb{E}_{T \sim \mu}[X + Y | Z = 0] \leq 4$. Moreover, for $r = 4, 3 + 2\epsilon \leq p \leq 4$ Lemma 6.14 implies that

$$\mathbb{P}_{T \sim \mu}[X + Y = 4 | Z = 0] \geq B(\mathbb{E}[X + Y | Z = 0], 4) \geq 2\epsilon e^{-2\epsilon} \geq \epsilon,$$

and therefore, from Corollary 6.20, we obtain that $\mathbb{P}_{T \sim \mu}[\mathcal{E}(e)] = \frac{\epsilon}{100}$.

If $x_e > 1/2 + \epsilon$ the statement can be proved similarly by conditioning on $Z = 1$. From Fact 6.9, we obtain that $\mathbb{E}_{T \sim \mu}[X + Y | Z = 1] \leq \mathbb{E}_{T \sim \mu}[X + Y] \leq 3 - 2\epsilon$ and Corollary 6.10 implies that we have $\mathbb{E}_{T \sim \mu}[X + Y | Z = 1] \geq 2$. Therefore, Lemma 6.14 implies that

$$\mathbb{P}_{T \sim \mu}[X + Y = 2 | Z = 1] \geq B(\mathbb{E}[X + Y | Z = 1], 2) \geq 2\epsilon \left(\frac{3 - 2\epsilon}{3} \right)^2 \geq \epsilon,$$

and applying Corollary 6.22 we obtain that $\mathbb{P}_{T \sim \mu}[\mathcal{E}(e)] = \frac{\epsilon}{100}$. □

In the following lemma we show that if $x(u, v) \leq 1/2$, and $\mathbb{P}_\mu[|d_v(u)| + |d_u(v)| = 4] \geq \epsilon$, then with some constant probability they both have an even degree. The proof uses the LC properties of Rayleigh measures crucially.

Lemma 6.19. *Let u and v be two vertices of G, such that $d(u) \cap d(v) = \emptyset$ (i.e., u and v does not share any edge). Define $X := |T \cap d(u)|$ and $Y := |T \cap d(v)|$. If the following conditions are satisfied then $\mathbb{P}_{T \sim \mu}[X = 2, Y = 2] \geq \frac{\epsilon \alpha}{4} \min(\frac{1}{2e}, \frac{\beta - 1}{e^{\beta - 1}})$:*

$$\epsilon \leq \mathbb{P}_{T \sim \mu}[X + Y = 4] \tag{8}$$

$$\alpha \leq \mathbb{P}_{T \sim \mu}[X \leq 2], \ \mathbb{P}_{T \sim \mu}[Y \leq 2] \tag{9}$$

$$1 < \beta \leq \mathbb{E}_{T \sim \mu}[X], \ \mathbb{E}_{T \sim \mu}[Y] \tag{10}$$

Proof. By equation (8), we have:

$$\mathbb{P}_{T \sim \mu}[X = 2, Y = 2] = \mathbb{P}_{T \sim \mu}[X = 2, Y = 2 | X + Y = 4]\mathbb{P}_{T \sim \mu}[X + Y = 4] \geq \epsilon \mathbb{P}_{T \sim \mu}[X = 2, Y = 2 | X + Y = 4]$$

Thus we need to show that $\mathbb{P}_{T \sim \mu}[X = 2, Y = 2 | X + Y = 4] \geq \alpha/4 \min(\frac{1}{2e}, \frac{\beta - 1}{e^{\beta - 1}})$. Let $\gamma := \alpha/4 \min(\frac{1}{2e}, \frac{\beta - 1}{e^{\beta - 1}})$. Theorem 6.13 implies that,

$$\mathbb{P}_{T \sim \mu}[X = 2, Y = 2 | X + Y = 4]^2 \geq \mathbb{P}_{T \sim \mu}[X = 3, Y = 1 | X + Y = 4]\mathbb{P}_{T \sim \mu}[X = 1, Y = 3 | X + Y = 4]$$

Note that in general both of the terms in the RHS could be zero (e.g. when we have $X = 2, Y = 2$ with high probability. Thus we prove $\mathbb{P}_{T \sim \mu}[X \geq 2 | X + Y = 4] \geq 2\gamma$ (it can be proved similarly that $\mathbb{P}_{T \sim \mu}[Y \geq 2 | X + Y = 4] \geq 2\gamma$). Note that this essentially implies $\mathbb{P}_{T \sim \mu}[X = 2, Y = 2 | X + Y = 4] \geq \gamma$.

31

Equivalently we can show that $\mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y = 4] \geq 2\gamma$. Since $X \geq 2$ and $Y \leq 2$ are an upward (resp. downward) event, Theorem 6.17 implies that

$$\mathbb{P}_{T\sim\mu}[X \geq 2 | X + Y = 4] \quad \geq \quad \mathbb{P}_{T\sim\mu}[X \geq 2 | X + Y = 3] \tag{11}$$
$$\mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y = 4] \quad \geq \quad \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y = 5] \geq \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y = 6] \geq \ldots$$

Note that inequality (11) is valid once the event $X + Y = 3$ is well defined. For this moment suppose this is the case, as we will see throughout the proof the other case is much simpler. These inequalities imply the following:

$$\mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y = 4] \quad \geq \quad \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4] \tag{12}$$

By equations (11), (12) to prove the lemma it is sufficient to show that

$$\mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4] + \mathbb{P}_{T\sim\mu}[X \geq 2 | X + Y = 3] \geq 4\gamma. \tag{13}$$

Let us consider the event $Y \leq 2$, we have

$$\begin{aligned}
\mathbb{P}_{T\sim\mu}[Y \leq 2] \quad &= \quad \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4]\,\mathbb{P}_{T\sim\mu}[X + Y \geq 4] + \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \leq 3]\,\mathbb{P}_{T\sim\mu}[X + Y \leq 3] \\
&\leq \quad \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4] + \mathbb{P}_{T\sim\mu}[X + Y \leq 3] \tag{14} \\
&= \quad \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4] + \mathbb{P}_{T\sim\mu}[X = 2, Y = 1] + \mathbb{P}_{T\sim\mu}[X = 1, Y \leq 2] \\
&\leq \quad \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4] + \mathbb{P}_{T\sim\mu}[X \geq 2 | X + Y = 3] + \mathbb{P}_{T\sim\mu}[X = 1, Y \leq 2]. \tag{15}
\end{aligned}$$

Note that since X and Y are indicator for all of the edges adjacent to u and v, we always have $X \geq 1$ and $Y \geq 1$, thus $X+Y \geq 2$. Therefore, if the event $X+Y = 3$ is not defined (i.e. $\mathbb{P}_{T\sim\mu}[X + Y \leq 3] = 0$), by equation (14) we obtain $\alpha \leq \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4]$ and we are done.

From Fact 6.9 and using equation (10) we have

$$\mathbb{E}_{T\sim\mu}[X | Y \leq 2] \geq \mathbb{E}_{T\sim\mu}[X] \geq \beta > 1.$$

Since $Y \geq 1$, theorem 6.5 implies that the measure $\{\mu : Y \leq 2\}$ is strongly Rayleigh. Therefore, we can apply Theorem 6.13 or Lemma 6.14 to upperbound $\mathbb{P}_{T\sim\mu}[X = 1 | Y \leq 2]$.

If $\beta \geq 2$, by Theorem 6.13 the mode of the distribution $\{X | Y \leq 2\}$ is at least 2, thus $\mathbb{P}_{T\sim\mu}[X = 1 | Y \leq 2] \leq \frac{1}{2}$. On the other hand, if $\beta \leq 2$, by Lemma 6.14

$$\mathbb{P}_{T\sim\mu}[X = 2 | Y \leq 2] \geq B(\beta, 2) = \min\{e^{-\beta+1}(\beta - 1), \frac{e^{-\beta}\beta^2}{2}\} \geq \min\{\frac{\beta - 1}{e^{\beta-1}}, \frac{1}{2e}\},$$

and $\mathbb{P}_{T\sim\mu}[X = 1 | Y \leq 2] \leq 1 - \min(\frac{\beta-1}{e^{\beta-1}}, \frac{1}{2e})$. Putting them together we obtain that

$$\mathbb{P}_{T\sim\mu}[X = 1, Y \leq 2] = \mathbb{P}_{T\sim\mu}[X = 1 | Y \leq 2]\,\mathbb{P}_{T\sim\mu}[Y \leq 2] \leq \left(1 - \min(\frac{1}{2e}, \frac{\beta - 1}{e^{\beta-1}})\right)\mathbb{P}_{T\sim\mu}[Y \leq 2]$$

Putting this with equation (15) we obtain:

$$\min(\frac{1}{2e}, \frac{\beta - 1}{e^{\beta-1}})\mathbb{P}_{T\sim\mu}[Y \leq 2] \leq \mathbb{P}_{T\sim\mu}[Y \leq 2 | X + Y \geq 4] + \mathbb{P}_{T\sim\mu}[X \geq 2 | X + Y = 3].$$

Finally using equation (9) we obtain equation (13). $\qquad\square$

Corollary 6.20. *Let u and v be two vertices of G of fractional degree 2. Let $X := |T \cap d_v(u)|$, $Y := |T \cap d_u(v)|$, and $Z := |T \cap \{(u,v)\}|$. If $x(u,v) \leq \frac{1}{2} + \frac{1}{5}$, and $\mathbb{P}_{T \sim \mu}[X + Y = 4 | Z = 0] \geq \epsilon$, then $\mathbb{P}[X = 2, \ Y = 2, \ Z = 0] \geq \epsilon/100$.*

Proof. Let $\mu_1 = \{\mu | Z = 0\}$ be the measure obtained from μ conditioned on $Z = 0$. We show that μ_1 satisfies all of the conditions of Lemma 6.19. Since $\mathbb{P}_{\mu_1}[X + Y = 4] \geq \epsilon$, the first condition is satisfied. Since $\mathbb{E}_{T \sim \mu}[X] = \mathbb{E}_{T \sim \mu}[Y] = 2 - x(u,v)$, Fact 6.9 implies that

$$1.5 - \frac{1}{5} \ \leq \ \mathbb{E}_{\mu_1}[X] \leq 2$$

$$1.5 - \frac{1}{5} \ \leq \ \mathbb{E}_{\mu_1}[Y] \leq 2$$

Now it is straightforward to see that $\alpha \geq \frac{1}{3}$ using Markov's inequality, and $\beta \geq 1.5 - \frac{1}{5}$, this implies corollary. $\qquad \square$

An identical argument to Lemma 6.19 gives the following lemma and its corollary. The proofs are deferred to the appendix.

Lemma 6.21. *Let A and B be two sets of edges such that $A \cap B = \emptyset$. Define $X := |T \cap A|$, $Y := |T \cap B|$. If the following conditions are satisfied then $\mathbb{P}_{T \sim \mu}[X = 1, Y = 1] \geq \frac{\epsilon \alpha}{4} \min(1/2, \beta e^{-\beta})$:*

$$\epsilon \ \leq \ \mathbb{P}_{T \sim \mu}[X + Y = 2] \tag{16}$$
$$\alpha \ \leq \ \mathbb{P}_{T \sim \mu}[X \leq 1], \ \mathbb{P}_{T \sim \mu}[Y \leq 1] \tag{17}$$
$$\beta \ \leq \ \mathbb{E}_{T \sim \mu}[X], \ \mathbb{E}_{T \sim \mu}[Y] \tag{18}$$

Corollary 6.22. *Let u and v be two vertices of G of fractional degree 2, and $e = (u,v)$. Let $X := |T \cap d_v(u)|$, $Y := |T \cap d_u(v)|$, and $Z := |T \cap \{e\}|$. If $x(u,v) \geq \frac{1}{2} - \frac{1}{5}$, and $\mathbb{P}_{T \sim \mu}[X + Y = 2 | Z = 1] \geq \epsilon$, then $\mathbb{P}[X = 1, \ Y = 1, \ Z = 1] \geq \frac{\epsilon}{100}$.*

In the following lemma we show that if $x(u,v) \simeq 1/2$ and $x(v,w) \simeq 1/2$ then with a constant probability either u and v will have an even degree, or v and w will have an even degree and therefore one of the edges (u,v) or (v,w) will be even with constant probability.

Lemma 6.23. *Let u, v, w be three vertices of the fractional graph G, and suppose $\frac{1}{2} - \epsilon \leq x(u,v), x(v,w) \leq \frac{1}{2} + \epsilon$, and $\epsilon \leq \frac{1}{8000}$. Then, either we have $\mathbb{P}_{T \sim \mu}[\mathcal{E}(u,v)] \geq \frac{\epsilon}{100}$ or $\mathbb{P}_{T \sim \mu}[\mathcal{E}(v,w)] \geq \frac{\epsilon}{100}$.*

Proof. Let $X_e(X_f)$ be the indicator random variable for the edge $e = (u,v)$ (resp. $f = (v,w)$). Also let $W := |T \cap d_v(u)|, Y := |T \cap d_u(v) \cap d_w(v)|, Z := |T \cap d_v(w)|$ be the random variables indicating the number of edges of $d_v(u), d_u(v) \cap d_w(v), d_v(w)$ sampled in T (e.g. note that $\mathbb{E}[W] \simeq \mathbb{E}[Z] \simeq 1.5$, and $\mathbb{E}[Y] \simeq 1$).

If X_e is highly negatively correlated with $W + Y + X_f$, i.e., if

$$\mathbb{E}\left[W + Y + X_f | X_e = 0\right] > \mathbb{E}\left[W + Y + X_f\right] + 4\epsilon \geq 3 + 2\epsilon$$

then similar to the proof of Proposition 6.18 we get $\mathbb{P}\left[W = 2, Y + X_f = 2 | X_e = 0\right] \geq \frac{\epsilon}{100}$ therefore e is good and we are done. Similarly, if

$$\mathbb{E}\left[Y + Z + X_e | X_f = 0\right] > \mathbb{E}[Y + Z + X_e] + 4\epsilon \geq 3 + 2\epsilon$$

33

then f is good and we are done.

Otherwise we have that conditioning on $X_e = 0/1$ or $X_f = 0/1$ does not change the probability of other edges adjacent to their endpoints. In particular, we have

$$\mathbb{E}\left[W + Y + X_f\right] \leq \mathbb{E}\left[W + Y + X_f | X_e = 0\right] \quad \leq \quad \mathbb{E}\left[W + Y + X_f\right] + 4\epsilon \qquad (19)$$

$$\mathbb{E}\left[Y + Z + X_e\right] \leq \mathbb{E}\left[Y + Z + X_e | X_f = 0\right] \quad \leq \quad \mathbb{E}\left[Y + Z + X_e\right] + 4\epsilon \qquad (20)$$

Observe that above equation also imply that and therefore,

$$\mathbb{E}\left[X_f | X_e = 0\right] \quad \leq \quad \mathbb{E}\left[X_f\right] + 4\epsilon \qquad (21)$$

$$\mathbb{E}\left[X_f | X_e = 1\right] \quad \geq \quad \mathbb{E}\left[X_f\right] - 4\epsilon. \qquad (22)$$

Here, for the sake of brevity we ignored the error term of $O(\epsilon^2)$ in (22). We first show that if X_e is not highly negatively correlated with Z, or if X_f is not highly correlated with W then also we are done.

Claim 6.24. *We have the following.*

1. *If $\mathbb{E}_{T \sim \mu}\left[Z | X_e = 0\right] \leq \mathbb{E}_{T \sim \mu}\left[Z\right] + \frac{1}{5}$, then $\mathbb{P}\left[\mathcal{E}(f)\right] \geq \mathbb{P}\left[Y = 1, Z = 1, X_f = 1, X_e = 0\right] \geq \frac{1}{5000}$ and f is good.*

2. *If $\mathbb{E}_{T \sim \mu}\left[W | X_f = 0\right] \leq \mathbb{E}_{T \sim \mu}\left[W\right] + \frac{1}{5}$, then $\mathbb{P}\left[\mathcal{E}(e)\right] \geq \mathbb{P}\left[Y = 1, W = 1, X_e = 1, X_f = 0\right] \geq \frac{1}{5000}$ and e is good.*

Proof. We prove the first statement; the second statement follows identically. Let $\mu_1 = \{\mu | X_e = 0\}$ be the measure obtained from μ conditioned on $X_e = 0$, and let $\mu_2 = \{\mu | X_e = 0, X_f = 1\}$. By the conditional probability axioms we have:

$$\mathbb{P}_{T \sim \mu}\left[Y = 1, Z = 1, X_e = 0, X_f = 1\right] \quad = \quad \mathbb{P}_{T \sim \mu}\left[Y = 1, Z = 1 | X_e = 0, X_f = 1\right] \mathbb{P}_{T \sim \mu}\left[X_f = 1 | X_e = 0\right] \mathbb{P}_{T \sim \mu}\left[X_e = 0\right]$$

$$\geq \quad \frac{1}{5}\mathbb{P}_{T \sim \mu_2}\left[Y = 1, Z = 1\right].$$

Thus it is sufficient to show that $\mathbb{P}_{T \sim \mu_2}\left[Y = 1, Z = 1\right] \geq 1/1000$. By the claim's assumption and equation (19) we have

$$\frac{1}{2} - 7\epsilon \quad \leq \quad \mathbb{E}_{T \sim \mu_2}\left[Y\right] \leq 1 + 6\epsilon$$

$$1 - 6\epsilon \quad \leq \quad \mathbb{E}_{T \sim \mu_2}\left[Z\right] \leq 1.7 + \epsilon.$$

Therefore, using Lemma 6.14, and $\epsilon < 1/1000$ we have

$$\mathbb{P}_{T \sim \mu_2}\left[Y + Z = 2\right] \geq B(\mathbb{E}_{T \sim \mu_2}\left[Y + Z\right], 2) \geq \frac{1}{5}.$$

Using Markov's inequality, we have $\mathbb{P}_{T \sim \mu_2}\left[Y \leq 1\right], \mathbb{P}_{T \sim \mu_2}\left[Z \leq 1\right] \geq \frac{1}{10}$. Hence, applying Lemma 6.21, for $\alpha^* = \frac{1}{10}, \beta^* = \frac{1}{2} - 7\epsilon$ and $\epsilon^* = \frac{1}{5}$, we obtain that $\mathbb{P}_{T \sim \mu_2}\left[Y = 1, Z = 1\right] \geq \frac{1}{1000}$ and f is good. $\qquad \square$

Thus suppose X_e is highly negatively correlated with Z and X_f is highly negatively correlated with W. We show that this implies a large negative correlation between X_e and X_f, which contradicts with equations (22) and (21).

First note that for a binary random variable B, and a random variable X, we can write

$$Cov(B, X) = \mathbb{E}[BX] - \mathbb{E}[B]\mathbb{E}[X] = \mathbb{P}[B](\mathbb{E}[X|B=1] - \mathbb{E}[X]).$$

Therefore, using the above claim we get:

$$Cov(X_e, Z) = \mathbb{P}[X_e](\mathbb{E}[Z|X_e=1] - \mathbb{E}[Z]) \leq -\frac{1}{10} \tag{23}$$

$$Cov(X_f, W) = \mathbb{P}[X_f](\mathbb{E}[W|X_f=1] - \mathbb{E}[W]) \leq -\frac{1}{10}, \tag{24}$$

where the lower order terms of $O(\epsilon)$ are ignored. In the rest of the proof we show the above equations imply that $Cov(X_e, X_f) \leq -4\epsilon$. But using equation (22) we have:

$$Cov(X_e, X_f) = \mathbb{P}[X_e](\mathbb{P}[X_f|X_e=1] - \mathbb{P}[X_f]) \geq -2\epsilon,$$

which is a contradiction.

It remains to compute $Cov(X_e, X_f)$.

Theorem 6.1 implies that we have $Cov(X_e, X_f) = -i^e(f)i^f(e)$ where $i^e(f)$ is the current through edge f when a unit current is imposed between the endpoints of e . In the next claim we show that $i^e(f) \geq -Cov(X_e, Z)/4$. Similarly, we have $i^f(e) \geq -Cov(X_f, W)/4$. Therefore, using equations (23) and (24), and $\epsilon < 1/8000$ we obtain:

$$Cov(X_e, X_f) \leq -\frac{-Cov(X_e, Z)}{4}\frac{-Cov(X_f, W)}{4} \leq -1/1600 \leq -4\epsilon.$$

Claim 6.25. *We have $i^e(f) \geq -Cov(X_e, Z)/4$.*

Proof. Suppose we impose a unit current between the endpoints of e (i.e. insert a unit current at u and extract it from v). First we show $\sum_{g \in d_v(w)} i^e(g) \geq -Cov(X_e, Z)$. Note that $i^e(g)$ is the absolute value of the current in edge g; thus it is always non-negative. Then we show that at least a quarter of the currents in edges adjacent to w belongs to the edge f. Let Z_g be the indicator random variable for edges $g \in d_v(w)$, Then:

$$Cov(X_e, Z) = \sum_{g \in d_v(w)} Cov(X_e, Z_g) = \sum_{g \in d_v(w)} -i^e(g)i^g(e) \geq - \sum_{g \in d_v(w)} i^e(g),$$

where the last inequality holds by the fact that $i^g(e) \leq 1$ for any two edges e and g.

Let $I_w = \frac{1}{2}\sum_{g \in d(w)} i^e(g)$ be the total current that comes into the vertex w, this current finally will be extracted at v. In the next claim we show that at least $I_w/2$ of it goes through edge f. The reason simply follows from the fact that $i^f(f) \simeq 1/2$; If we impose a unit current between the endpoints of f, roughly half of it crosses f. Therefore, when a current I_w is transferred from w to v, at least half of it should go through f

Since $\sum_{g \in d(w)} i^e(g) = 2I_w$, we get:

$$i^e(f) \geq \frac{I_w}{2} \geq \frac{1}{4}\sum_{g \in d(w)} i^e(g) \geq \frac{-Cov(X_e, Z)}{4}.$$

$\square$

Claim 6.26. *Suppose we insert a unit current at u and extract it from v. If I_w of it enters w, then we have $i^e(f) \geq i^f(f)I_w$*

Proof. We construct a unit current flow θ' from w to v in a network with smaller conductances $\lambda' \leq \lambda$, such that $\theta'(f) = i^e(f)/I_w$. Then by Thompson's Principle (cf. e.g. [29, Chapter 2]), θ' must consume more energy than i^f. This implies that the voltage difference of v and w must be larger in θ', thus $i^e(f)/I_w = \theta'(f) \geq i^f(f)$.

Let $P_{w,v}$ be the set of paths from w to v, and for any $P \in P_{w,v}$ let θ_P be the amount of current that goes through P when sending a unit current from u to v. By claim's assumption $\sum_{P \in P_{w,v}} \theta_P = I_w$. Also for any flow i, let $i(g)$ be the amount of flow that goes through an edge g. Note that $\theta(g) \leq i^e(g)$ for all edges g, but $\theta(f) = i^e(f)$.

We construct another electrical network with a different assignment of conductances in which θ is the electrical current from w to v. It is sufficient to decrease the conductance of each edge g such that the difference of the voltages of its endpoints remains the same i.e., define $\lambda'_g = \lambda g * \frac{\theta(g)}{i^e(g)}$. Now let $\theta' := \frac{1}{I_w}\theta$ be the unit current from w to v in the new network.

Since we only decreased the conductance of the edges, by Thompson's principle we have $Energy(\theta') \geq Energy(i^f)$. On the other hand, since θ', i^f are unit current flows from w to v we have:

$$Energy(\theta') = \lambda'_f \theta'(f) = \lambda_f \frac{i^e(f)}{I_w},$$
$$Energy(i^f) = \lambda_f i^f(f),$$

It follows that $i^e(f) \geq i^f(f)I_w$. $\qquad\square$

This completes the proof of the Lemma. $\qquad\square$

6.3 Inside Good Edges

In this subsection we show the following two lemmas which will lower bound the number of good edges assigned to cyclic cut classes. See subsection 5.2 For the definition of cyclic cut classes.

Lemma 6.27. *If $e = (u, v)$ be an inside edge of a cyclic cut class C_i then $\mathbb{P}_{T \sim \mu}[\mathcal{E}(e)] \geq \frac{1}{2000}$.*

Since each cut class has a large fraction of inside or trivial edges if its degree in $\Gamma(C)$ is small, we get the following bound on the good edges assigned to the cut class.

Lemma 6.28. *Let C_i be a cyclic cut class and the degree of C_i in $\Gamma(C)$ is k, then the sum of the fraction of good edges assigned to C_i is least $(1 - \kappa(C_i)) \left\lceil \frac{|\phi(C_i)| - 2k}{2} \right\rceil$.*

Suppose that we have an edge e which is an inside edge of a cyclic cut class. Although e may occur in many ($\Omega(\tau^2)$) near min-cuts, we still manage to argue that $\mathcal{E}(e)$ occurs with a constant probability. This is possible with the help of the following probabilistic arguments.

Let us start with a simple averaging argument that proves for a set of vertices $U \subset V$ of the graph $G = (V, E, x)$, if $(U, \overline{U})$ is a near minimum cut, and $|U| \leq |\overline{U}|$, then there is a constant probability that the sampled tree is also a spanning tree inside U (see Corollary 6.31).

Definition 6.29. *For $S \subseteq E$ of a graph $G = (V, E)$, let $r(S)$ be the rank of the graphical matroid defined on S. In other words, $r(S) = |V| - 1 - \#comp(S)$, where $\#comp(S)$ is the number of connected components of the subgraph (V, S).*

Recall that in the λ-uniform measure on spanning trees of $G = (V, E, x)$, for any edge $e \in E$, we have $\mathbb{P}[e \in T] = z_e = (1 - \frac{1}{n})x_e$. In the following, we show that if $z(S)$ is close to $r(S)$ for $S \subset E$, then $|T \cap S| = r(S)$ with high probability:

Proposition 6.30. *If for a set S, $z(S) > r(S) - 1$, then*

$$\mathbb{P}_{T \sim \mu}[|T \cap S| = r(S)] \geq 1 + z(S) - r(S),$$

where $z(S) := \sum_{e \in S} z_e$.

Proof. The proof is simply followed by the matroidal property of spanning trees. Indeed for any spanning tree in G we have $|T \cap S| \leq r(S)$. Let p be the probability that $|T \cap S| = r(S)$. We have:

$$
\begin{aligned}
z(S) = \mathbb{E}_{T \sim \mu}[|T \cap S|] &\leq p \cdot r(S) + (1 - p) \cdot (r(S) - 1) \\
&= r(S) + p - 1.
\end{aligned}
$$

Therefore $p \geq 1 + z(S) - r(S)$. $\qquad \square$

Suppose $U \subset V$ is small (say $|U| = O(1)$), and $(U, \overline{U})$ is a near minimum cut of G. The next corollary shows that with high probability the random spanning tree T chosen by the algorithm is also a spanning tree inside U. Even if $|U| \simeq n/2$, this probability is still roughly $1/2$ when δ is small enough.

Corollary 6.31. *If $(U, \overline{U})$ is a $(1 + \delta)$ near minimum cut with respect to x, and $|U|/n + \delta < 1$, then*

$$\mathbb{P}_{T \sim \mu}[|T \cap E(U)| = |U| - 1] \geq 1 - \frac{|U|}{n} - \delta$$

Proof. Since the fractional degree of each vertex with respect to x is 2, and $x(d(U)) \leq 2 + 2\delta$, we have

$$x(E(U)) \geq \frac{1}{2}(2|U| - 2 - 2\delta) = |U| - 1 - \delta.$$

Hence

$$z(E(U)) \geq (1 - 1/n)(|U| - 1 - \delta) \geq |U| - \frac{|U|}{n} - 1 - \delta.$$

Therefore, the statement follows from Proposition 6.30, by setting $S := E(U)$, and noting that $r(E(U)) \leq |U| - 1$. $\qquad \square$

Now, we are ready to prove the main result of the section. We will show that with constant probability, all the near minimum cuts of any cyclic (small) cut class are even. This is done by showing that T is a Hamiltonian cycle in $G(\phi(C_i))$ with constant probability. Lemma 6.27 and Lemma 6.28 now follow simply and are proven after the proof of the Lemma 6.32.

Lemma 6.32. *If C_i is cyclic then all of the near minimum cuts in C_i will have an even number of edges in the tree T, with probability at least $1/2000$.*

Proof. Since $\kappa(C_i) \leq 1$, by proposition 5.20, C_i does not have any inside atoms. Let $\phi(C_i) = \{a_1, a_2, \ldots, a_k\}$ and assume that they are placed around the polygon in the order of their labels. Let $E(G(\phi(C_i)))$ be the set of edges with endpoints in different atoms in C_i. Let $\mathcal{A}$ be the event that T contains exactly k edges from $E(G(\phi(C_i)))$ with the property that $|T \cap E(a_i, a_{i+1})| = 1$ for all adjacent pair of atoms. In other words, $\mathcal{A}$ is the event T is a Hamiltonian cycle in $G(\phi(C_i))$. Since the near minimum cuts in C_i are just the representing diagonals of the polygon, if $\mathcal{A}$ occurs, then T contains exactly two edges in each of the cuts in C_i. Thus it is sufficient to show that $\mathbb{P}[\mathcal{A}] = \frac{1}{2000}$.

Let a_1 be the father-connector of C_i. Observe that a_1 contains at least $\frac{n}{2}$ vertices. Let $F := \bigcup_{i=2}^{k} E(a_i)$ be the set of edges inside all small atoms, $I_1 := \bigcup_{2 \leq i < k} E(a_i, a_{i+1})$ be the set of edges between the consecutive small atoms, and $I_2 := E(a_k, a_1) \cup E(a_1, a_2)$ be the set of edges from a_1 to its adjacent atoms in the cyclic ordering. Finally, let $I_3 := E(G(\phi(C_i))) \setminus I_1 \setminus I_2$ be the rest of the edges.

We compute the probability of $\mathcal{A}$ in the following steps: first we condition on $|T \cap I_3| = 0$, and $|T \cap F| = r(F)$. The latter imply that T contains a spanning tree inside each of the atoms $a_2, \ldots, a_k$. Conditioned on $|T \cap F| = r(F)$, T can select at most one edge between each consecutive pair of small atoms. Next we condition on $|T \cap I_1| = k - 2$, which implies that T is a Hamiltonian path from a_2 to a_k in $G(\phi(C_i))$. In the last step we contract the whole Hamiltonian path as a single vertex v^*, and we show that with constant probability, T contains exactly one edge from each of $E(a_1, a_2)$ and $E(a_k, a_1)$.

First by Lemma 5.20 (3), we have

$$x(I_1) + x(I_2) \geq k(1 - \kappa(C_i)) \geq k - \frac{1}{20},$$

where the last inequality follows since C_i is cyclic and therefore $|\phi(C_i)|\kappa(C_i) \leq 1/20$. Since by Corollary 4.15, $x(E(G(\phi(C_i)))) \leq k(1 + 3\delta)$, we have

$$z(I_3) \leq x(I_3) \leq k(1 + 3\delta) - x(I_1) - x(I_2) \leq 3k\delta + \frac{1}{20} \leq \frac{1}{10}.$$

Therefore by Markov's inequality,

$$\mathbb{P}_{T \sim \mu}\left[|T \cap I_3| = 0\right] \geq 9/10. \tag{25}$$

Let $\mu_1 := \{\mu \mid |T \cap I_3| = 0\}$. This conditioning can only increase the probability of other edges. Now let us compute the probability that $T \sim \mu_1$, contains a spanning tree in all of the small atoms. First note that

$$\begin{aligned}
z(F) = (1 - \frac{1}{n})x(F) &\geq (1 - \frac{1}{n})\left(\sum_{i=2}^{k}[|a_i| - 1 - \kappa(C_i)]\right) &\tag{26} \\
&\geq \sum_{i=2}^{k}|a_i| - \frac{\sum_{i=2}^{k}|a_i|}{n} - (k-1)(1 + \kappa(C_i)) \\
&\geq \sum_{i=2}^{k}(|a_i| - 1) - \frac{1}{2} - \frac{1}{20}, &\tag{27}
\end{aligned}$$

where equation (26) follows from second part of Proposition 5.20, and equation (27) follows from the fact that $|a_1| \geq n/2$. Since $r(F) \leq \sum_{i=2}^{k}(|a_i| - 1)$, by Proposition 6.30, with probability at least

38

$1/2 - 1/20$, T contains an spanning tree inside each atom $a_2, \ldots, a_k$:

$$\mathbb{P}_{T \sim \mu_1} \left[|T \cap F| = r(F) \right] = \mathbb{P}_{T \sim \mu_1} \left[\bigwedge_{i=2}^{k} |T \cap E(a_i)| = |a_i| - 1 \right] \geq \frac{9}{20}. \tag{28}$$

Let $\mu_2 := \{\mu_1 \mid \forall 2 \leq i \leq k : |T \cap E(a_i)| = |a_i| - 1\}$. Fact 6.9 implies that this may decrease the probability of subset of edges not containing an edge from $I_3 \cup F$ by at most $1/2 - 1/20$. Hence we have

$$\mathbb{E}_{T \sim \mu_2} \left[|T \cap I_1| \right] \geq z(I_1) - \frac{1}{2} - \frac{1}{20} \geq k - 2 - \frac{3}{5}.$$

Since any spanning tree $T \sim \mu_2$ can have at most one edge between any consecutive pair of small atoms, $|T \cap I_1| = k - 2$ implies that T contains a Hamiltonian path through the small atoms. From Markov's inequality, this happens with probability at least $2/5$:

$$\mathbb{P}_{T \sim \mu_2} \left[|T \cap I_1| = k - 2 \right] = \mathbb{P}_{T \sim \mu_2} \left[\bigwedge_{i=2}^{k-1} |T \cap E(a_i, a_{i+1})| = 1 \right] \geq \frac{2}{5}. \tag{29}$$

Finally, let $\mu_3 := \{\mu_2 \mid |T \cap I_1| = k - 2\}$ therefore, under μ_3 probability of any subset of edges not containing an edge in $I_1 \cup F \cup I_3$ will decrease by at most $\frac{3}{5}$ using Fact 6.9. It remains to find the probability that $|T \cap E(a_k, a_1)| = 1$ and $|T \cap E(a_1, a_2)| = 1$. Since any tree $T \sim \mu_3$, contains a spanning tree on $V \setminus a_1$, we can contract these vertices to a new vertex v^*. The two cuts (a_k, a_1) and (a_1, a_2) then correspond to two set of edges incident to v^* (these are the only non-zero edges adjacent to v^*). Observe that

$$1 + \frac{1}{10} \geq \mathbb{E}_{T \sim \mu_3} \left[|T \cap E(a_k, a_1)| \right] \geq 1 - \kappa(C_i) - \frac{3}{5} \geq \frac{3}{10}$$

$$1 + \frac{1}{10} \geq \mathbb{E}_{T \sim \mu_3} \left[|T \cap E(a_2, a_1)| \right] \geq 1 - \kappa(C_i) - \frac{3}{5} \geq \frac{3}{10}$$

$$2 + \frac{1}{10} \geq \mathbb{E}_{T \sim \mu_3} \left[|T \cap E(a_k, a_1)| + |T \cap E(a_1, a_2)| \right] \geq 2 - 2\kappa(C_i) - \frac{3}{5} \geq \frac{13}{10}.$$

From Lemma 6.14 we obtain

$$\mathbb{P}_{T \sim \mu_3} \left[|T \cap E(a_k, a_1)| + |T \cap E(a_1, a_2)| = 2 \right] \geq B(\mathbb{E}_{T \sim \mu_3} \left[|T \cap E(a_k, a_1)| + |T \cap E(a_1, a_2)| \right], 2) \geq \frac{1}{5}.$$

From Lemma 6.21, setting $A = E(a_1, a_2)$ and $B = E(a_1, a_k)$, $\alpha = \frac{1}{2} - \frac{1}{20}$, $\beta = \frac{3}{10}$, $\epsilon = 1/5$, we obtain that T will contain an edge in each of the cuts (a_1, a_2) and (a_k, a_1) with probability at least $1/300$. Therefore by equations (25), (28), and (29):

$$
\begin{aligned}
\mathbb{P}_{T \sim \mu} \left[\mathcal{A} \right] &= \mathbb{P}_{T \sim \mu} \left[\mathcal{A} \mid |T \cap I_3| = 0 \right] \mathbb{P}_{T \sim \mu} \left[|T \cap I_3| = 0 \right] \\
&\geq \frac{9}{10} \mathbb{P}_{T \sim \mu_1} \left[\mathcal{A} \mid |T \cap F| = r(F) \right] \mathbb{P}_{T \sim \mu_1} \left[|T \cap F| = r(F) \right] \\
&\geq \frac{9}{10} \frac{9}{20} \mathbb{P}_{T \sim \mu_2} \left[\mathcal{A} \mid |T \cap I_1| = k - 2 \right] \mathbb{P}_{T \sim \mu_2} \left[|T \cap I_1| = k - 2 \right] \\
&\geq \frac{9}{10} \frac{9}{20} \frac{2}{5} \mathbb{P}_{T \sim \mu_3} \left[|T \cap E(a_k, a_1)| = 1, |T \cap E(a_1, a_2)| = 1 \right] \geq \frac{1}{2000}.
\end{aligned}
$$

$\square$

Proof of Lemma 6.27. Since (u, v) is an inside edge of C_i, u and v are singleton atoms of C_i and the only near minimum cuts containing e are diagonals of C_i and the trivial cuts $(\{u\}, \overline{\{u\}})$ and $(\{v\}, \overline{\{v\}})$. Lemma 6.32 implies that with probability $1/2000$, T is a Hamiltonian cycle when each of the atoms of C_i are contracted. Thus the degree of u and v is even and also each of the diagonal cut of C_i contains exactly two edges of T. $\qquad\square$

Proof of Lemma 6.28. Recall that an atom is connecting if it coincides with an atom of another cut class in a vertex of $K(C)$. Since the number of connecting atoms is at most the degree of C_i, C_i has at most k connecting atoms. Since $\kappa(C_i) < 1$, by proposition 5.20 it does not contain any inside atoms. Let $\phi(C_i) = \{a_1, a_2, \ldots, a_m\}$, and assume that they are placed around the polygon in the order of their labels. Let $(a_{i_1}, a_{i_1+1}), \ldots, (a_{i_{m'}}, a_{i_{m'}+1})$, be m' disjoint consecutive pair of atoms, such that none of the a_{i_j}'s are connecting. Observe that we can choose $m' \geq \lceil \frac{m-2k}{2} \rceil$. We show that each such pair corresponds to good edges of fraction at least $1 - \kappa(C_i)$ assigned to C_i.

Consider a pair (a_{i_j}, a_{i_j+1}), if they both are singletons (say $a_{i_j} = \{u\}$ and $a_{i_j+1} = \{v\}$) then by proposition 5.20 the inside edge (u, v) has fraction at least $1 - \kappa(C_i)$ and by Definition 4.23 it is assigned to C_i. Since C_i is cyclic, by Lemma 6.27 the inside edge (u, v) is a good edge and we are done.

Otherwise, without loss of generality, suppose a_{i_j} is not a singleton. Since a_{i_j} is not a connecting atom the edges between the vertices inside a_{i_j} are only included in trivial minimum cuts, and they are all assigned to C_i by Definition 4.23. By proposition 5.20, $x(d(a_{i_j})) \leq 2(1 + \kappa(C_i))$. Thus if $a_{i_j} = \{u, v\}$ we have $x(u, v) \geq 1 - \kappa(C_i)$, and by Proposition 4.17 it is a good edge. Otherwise there are at least $3 - (1 + \kappa(C_i))$ edges between the vertices of a_{i_j} and at least half of them are good by Proposition 4.17. $\qquad\square$

Corollary 6.33. *Let C_i be a cyclic cut class which is a leaf of $\Gamma(C)$. Then C_i is assigned good edges of fraction at least $1 - \kappa(C_i)$.*

Proof. If $|\phi(C_i)| > 2$, then we are done by Lemma 6.28. Otherwise, suppose $\phi(C_i) = \{a_1, a_2\}$, and suppose a_1 is the father-connector of C_i. Since C_i is a non-trivial cut class, a_2 is not a singleton. Also since C_i is a leaf, a_2 is not a connecting atom. Therefore, the edges between the vertices of a_2 are trivial and at least $1 - \kappa(C_i)$ of them are good and assigned to C_i as in proof of Lemma 6.28. $\qquad\square$

6.4 Threads and Good Edges

In this subsection we prove Proposition 4.25. Let $P \in \pi_{\epsilon^*}$, be an unbalanced thread with parameter $\epsilon^* < 1/25$. We show that it will be assigned good edges of fraction at least $\epsilon^*/4$. If P is assigned trivial or inside edges of total fraction at least $\epsilon^*/4$ we are done. Therefore, in the proof we assume the sum of the fraction of trivial/inside good edges assigned to P is less then $\epsilon^*/4$ (it can be zero). It turns out that this will eliminate many possible cases. First we show that if P contains a cut class C_i with $|\phi(C_i)| \neq 2, 4$, then it will be assigned trivial or inside good edges of fraction at least $3/4$.

By Lemma 6.28, any cut class with $|\phi(C_i)| > 4$ of degree 2, will be assigned good edges of fraction at least $1 - \kappa(C_i) > 3/4$. Since all of the small cut classes in P has degree 2 in $\Gamma(C)$, if P contains a cut class of more than 4 atoms, it will be assigned good edges of fraction $3/4$. Note that by definition 4.12 we always have $|\phi(C_i)| \neq 3$. Since P is assigned less than $\frac{\epsilon^*}{4} < \frac{1}{100}$ trivial/inside good edges, all of the cut classes in P must have either 2 or 4 atoms.

We prove the theorem by Lemmas 6.36 and 6.38. In lemma 6.36 we show that if P contains 6 consecutive cut classes of 2 atoms then it will be assigned good edges of fraction $\frac{\epsilon^*}{4}$. In Lemma

6.38 we show that if P contains 3 consecutive cut classes such that the third one has 4 atoms then it will be assigned good edges of fraction $\frac{\epsilon^*}{4}$. Putting them together it is straightforward that if P has at least 8 cut classes, then at least one of the two cases occurs.

Before proving the lemmas, let us describe another classes of trivial edges that is assigned to P. This special case would make the proofs of the two lemmas much simpler. Let us define some notations first: Let C_i, C_{i+1} be two consecutive cut classes in P. We use the notation $C_i \rightarrow C_{i+1}$ to describe C_{i+1} is the child of C_i. Note that since the degree of each cut class in P is 2, C_i has a unique child in $\Gamma(C)$. Moreover, it has exactly 2 connecting atoms. One of them is the father-connector that connects C_i to its father (see Definition 4.21), the other, call it **child-connector**, connects C_i to its only child (C_{i+1}).

Claim 6.34. *Let $P \in \pi_0$, and $C_1 \rightarrow C_2 \in P$ be cut classes with at most 4 atoms. Let $a \in \phi(C_1)$ be the child connector of C_1, and $b \in \phi(C_2)$ be the father-connector of C_2. If $|a \setminus \bar{b}| > 2$, then P is assigned good edges of fraction $\frac{3}{8}$. Otherwise, if $|a \setminus \bar{b}| = 2$ and the two vertices are connected by an edge of fraction at least ϵ, then P is assigned good edges of fraction at least ϵ.*

Proof. Since C_2 is the child of C_1, by definition 4.12 we have $a \cup b = V(G)$, thus $\bar{b} \subseteq a$. Let $S := a \setminus \bar{b}$. Since the edges between the vertices of S are not contained in C_1 nor C_2, they are only contained in trivial cuts. Also, by definition 4.23 these edges are assigned to C_1 (and thus to P). Therefore, it is sufficient to show that they are good and they have a large fraction.

Suppose $|S| \geq 3$. Since C_1, C_2 have at most 4 atoms, by Lemma 5.20 part 2, the fractional degree of their atoms are at most $2(1+2\delta)$. Hence, we have $x(a), x(b) \leq 2(1+2\delta)$. Since each vertex of S has degree 2, and $|S| \geq 3$, there is at least $1 - 3\delta$ edges between the vertices of S, and by proposition 4.17 at least half of these trivial edges are good. Therefore C_1 will be assigned good edges of fraction at least 3/8.

Now suppose $S = \{u, v\}$, and $x(u, v) > 0$. If $x(u, v)$ is bounded away from 1/2, then by Proposition 4.17 it is a good edge. Otherwise, since $(a, \bar{a})$ and $(b, \bar{b})$ are $(1 + 2\delta)$ near minimum cuts, by Lemma G.1 (u, v) is a good edge. Therefore, if $x(u, v) \geq \epsilon$, then C_1 will be assigned good edges of fraction at least ϵ. $\square$

Corollary 6.35. *Let $P \in \pi(0)$, $C_1 \rightarrow C_2 \in P$, a is the child-connector of C_1, b is the father connector of C_2, and $S = a \setminus \bar{b}$. If P is assigned trivial/inside good edges of fraction less than $\frac{\epsilon^*}{4}$, then $|S| \leq 2$. In addition, if we also have $S = \{u\}$, then $x(u, \bar{a}), x(u, \bar{b}) \geq 1 - 4\delta$. Otherwise, if we have $S = \{u, v\}$, then*

$$x(u, v) < \frac{\epsilon^*}{4}$$

$$x(u, \bar{a}), x(u, \bar{b}), x(v, \bar{a}), x(v, \bar{b}) > 1 - 4\delta - \frac{\epsilon^*}{4}$$

Proof. Since P is assigned trivial/inside good edges of fraction less than $\frac{\epsilon^*}{4}$, all cut classes in P have at most 4 atoms; thus by claim 6.34, we certainly have $|S| \leq 2$ and if $S = \{u, v\}$, then $x(u, v) < \frac{\epsilon^*}{4}$. Since $x(\bar{a}) \leq 2(1+2\delta)$, and $x(\bar{b}) \leq 2(1+2\delta)$, if $S = \{u\}$, then by algebraic calculations we observe that u must have edges of fraction at least $1 - 4\delta$ to both $\bar{a}, \bar{b}$; Otherwise, if $S = \{u, v\}$, since $x(u, v) \leq \frac{\epsilon^*}{4}$, both of them must have edges of fraction at least $1 - 4\delta - \frac{\epsilon^*}{4}$ to $\bar{a}, \bar{b}$. $\square$

In the next lemma we show that if P contains 6 consecutive cut classes of 2 atoms then it will be assigned good edges of fraction $\frac{\epsilon^*}{4}$.

Lemma 6.36. *Let $P \in \pi_\epsilon^*$, $C_1 \to C_2 \to \ldots \to C_6 \in P$, be cut classes of 2 atoms, with child-connectors $a_1, \ldots, a_6$ respectively. If $\epsilon^* > 60\delta$, then P is assigned good edges of fraction at least $\frac{\epsilon^*}{4}$, that are even with probability at least $\frac{\epsilon^{*3}}{100000}$.*

First we show that if there is $v \in a_i \setminus a_{i+1}$, and $u \in a_{i+1} \setminus a_{i+2}$, where $x(u, v) > 0$ then $e = (u, v)$ is a good edge.

Lemma 6.37. *Let $P \in \pi_\epsilon^*$, $C_1 \to C_2 \to C_3 \in P$ be cut classes each with 2 atoms, with child-connectors a_1, a_2, a_3 respectively. Let $u \in a_2 \setminus a_3$, $v \in a_1 \setminus a_2$, and $x(u, v) > 0$. If $\epsilon^* > 60\delta$, and P is assigned less than $\frac{\epsilon^*}{4}$ fraction of trivial/inside good edges, then*

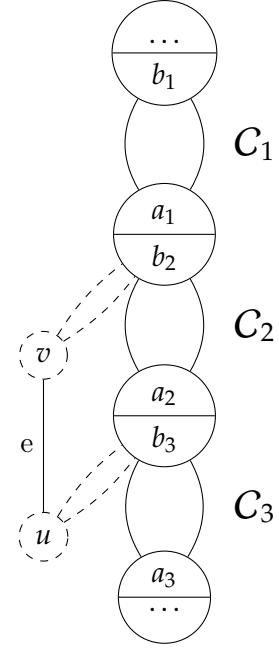$$\mathbb{P}_{T \sim \mu}[\mathcal{E}(u, v)] = \frac{\epsilon^{*3}}{100000}.$$

Figure 5: Cut classes C_1, C_2, C_3 as described in Lemma 6.37. Dashed edges represent deleted trivial cut classes. Connecting atoms are represented in the same circle.

Proof. We show that edge $e = (u, v)$ will be even with a constant probability. As shown in Figure 5 e is contained in 3 near minimum cuts: 2 trivial degree cut, and the near minimum cut defined by cut class C_2, $(a_2, \overline{a_2})$. We show that with a constant probability all of these 3 cuts will have even number of edges in the sampled tree $T \sim \mu$.

We show that with constant probability T is a spanning tree when restricted to a_2. Conditioned on this event, random variables which are functions of edges inside a_2 are independent of functions of edges outside a_2. We first show that u has degree one inside a_2 with constant probability. Observe that rest of the two min cuts do not have any edges in a_2. We then show that the event that degree of v and the size of cut $(a_2, \overline{a_2})$ are both two, and the degree of u restricted to $\overline{a_2}$ is one occur with constant probability. By the independence of the two events, all three cuts are even with constant probability completing the proof of the lemma.

Define $X := |T \cap E(a_2)|$. Since $(a_2, \overline{a_2})$ is a $(1 + \delta)$ near minimum cut, by corollary 6.31, $\mathbb{P}_{T \sim \mu}[X = |a_2| - 1] \geq 1 - \frac{|a_2|}{n} - \delta$. Since a_2 is not the father-connector of C_2, we get $|a_2| \leq n/2$. Define $\mu_1 := \{\mu \mid X = |a_2| - 1\}$. We get:

$$\mathbb{P}_\mu[\mathcal{E}(e)] = \mathbb{P}_{\mu_1}[\mathcal{E}(e)]\,\mathbb{P}_\mu[X = |a_2| - 1] \geq (\frac{1}{2} - \delta)\mathbb{P}_{\mu_1}[\mathcal{E}(e)]$$

Thus, it is sufficient to show that $\mathbb{P}_{T \sim \mu_1}[\mathcal{E}(e)] = \Omega(1)$. We show the following two inequalities

$$\mathbb{P}_{T \sim \mu_1}[|T \cap E(u, a_2)| = 1] = \Omega(1) \tag{30}$$
$$\mathbb{P}_{T \sim \mu_1}[T \cap E(a_2, \overline{a_2}) = 2, deg_T(v) = 2, |T \cap E(u, \overline{a_2})| = 1] = \Omega(1) \tag{31}$$

Since the above two events are independent under μ_1, we are done.

By Corollary 6.35 we have $\frac{9}{10} \leq x(E(u, a_2)) \leq \frac{11}{10}$. From Lemma 6.10 and Fact 6.9 we have

$$\frac{9}{10} \leq \mathbb{E}_{T \sim \mu_1}[|T \cap E(u, a_2)|] \leq \frac{16}{10} \implies \mathbb{P}_{T \sim \mu_1}[|T \cap E(u, a_2)| = 1] \geq B(\frac{16}{10}, 1) \geq \frac{3}{10}. \tag{32}$$

where the last inequality follows from Lemma 6.14. This proves inequality (30).

Now we prove inequality (31). Since restricted to a_2, T is a spanning tree, we contract a_2 to get a single vertex u'. In Lemma G.2, set $u := u'$, $v := v$, $N_1(u) = E(u, \overline{a_2})$, $N_2(u) = E(a_2 \setminus u, \overline{a_2})$ and $\epsilon = \frac{\epsilon^*}{3}$. Observe that the Lemma G.2 implies inequality (31). We now verify the conditions of the lemma. By Corollary 6.35, we have

$$
\begin{aligned}
1 - 4\delta - \frac{\epsilon^*}{4} &\leq\ x(v, a_2), x(v, \overline{a_2}), x(u, \overline{a_2}) \leq 1 + 4\delta + \frac{\epsilon^*}{4} \\
1 - 4\delta - \frac{\epsilon^*}{4} &\leq\ x(a_2 \setminus u, \overline{a_2}) \leq 1 + 6\delta + \frac{\epsilon^*}{4} \\
2 &\leq\ x(u, \overline{a_2}) + x(a_2 \setminus u, \overline{a_2}) \leq 2(1 + \delta)
\end{aligned}
$$

Since $z(e) \simeq x(e)$ the expectation under μ of the above values are nearly equal to their x-values. Since $P \in \pi_\epsilon^*$ we have $|\overline{a_2}| \geq n(\frac{1}{2} + \epsilon^*)$, and $|a_2| \leq n(\frac{1}{2} - \epsilon^*)$. Therefore, by negative association, conditioning on $X = |a_2| - 1$ may decrease the probability of edges outside of a_2 no more than $\frac{1}{2} - \epsilon^* + \delta$. Thus the LHS of the above equations can be decreased by at most $\frac{1}{2} - \epsilon^* + \delta$ and the conditions of the Lemma G.2 are satisfied. Hence,

$$
\mathbb{P}_{T \sim \mu_1}\left[T \cap E(a_2, \overline{a_2}) = 2, deg_T(v) = 2, |T \cap E(u, \overline{a_2})| = 1\right] \geq \frac{\epsilon^{*3}}{10000},
$$

$$
\Rightarrow\ \mathbb{P}_\mu\left[\mathcal{E}(e)\right] \geq (\frac{1}{2} - \delta)\frac{3}{10}\frac{\epsilon^{*3}}{10000} \geq \frac{\epsilon^{*3}}{100000}
$$

$\square$

Now we are ready to prove Lemma 6.36:

Proof of Lemma 6.36. Let $s_i := a_i \setminus a_{i+1}$, for $1 \leq i \leq 5$. By corollary 6.35 we have $|s_i| \leq 2$ for all $i \leq 5$. First note that if $|s_i| = 2$, for some i, then there is edge of fraction at least $1/3$ between a pair of vertices in s_i and s_{i+1}; thus by Lemma 6.37 P will be assigned good edges of fraction at least $\frac{1}{3} \geq \frac{\epsilon^*}{4}$.

So suppose all of s_i's contain only one vertex, say $s_i = \{u_i\}$ (see Figure 6). By lemma 6.37 if two consecutive vertices are connected by an edge of fraction $\frac{1}{100} \geq \frac{\epsilon^*}{4}$, then that edge will be good and we are done. Thus, suppose $\forall 1 \leq i \leq 5 : x(u_i, u_{i+1}) \leq \frac{1}{100}$. We show that in this case $x(u_2, u_4) \geq \frac{19}{20}$, and it will be even with a constant probability. Since this edge is assigned to C_4, P is assigned good edges of fraction at least $\frac{\epsilon^*}{4}$. Edge (u_2, u_4) is included in 4 near minimum cuts: 2 degree cuts, and the cuts corresponding to the cut classes C_3 and C_4 (i.e., $(a_3, \overline{a_3})$ and $(a_4, \overline{a_4})$); thus we need to show all these 4 cuts are even with a constant probability. We prove that the 3 edges $(u_1, u_3), (u_2, u_4), (u_3, u_5)$ all have a large fraction, then by union bound all of them will be sampled in spanning tree $T \sim \mu$, and thus the 2 cuts $(a_3, \overline{a_3})$ and $(a_4, \overline{a_4})$ are even. Finally, we use Lemma 6.21 to show that the degree of u_2 and u_4 is even with a constant probability.

By Corollary 6.35 we have

$$
\begin{aligned}
1 - 4\delta - \frac{2\epsilon^*}{4} &\leq\ x(u_3, \overline{a_2}), x(u_2, \overline{a_1}), \\
1 - 4\delta - \frac{\epsilon^*}{4} &\leq\ x(u_1, a_2)
\end{aligned}
$$

43

Hence $x(u_1, u_3) \geq 1 - 6\delta - \frac{3\epsilon^*}{4} \geq \frac{19}{20}$. Similarly, it can be shown that $x(u_3, u_5), x(u_2, u_4) \geq \frac{19}{20}$. Let $e_1 = (u_1, u_3), e_2 = (u_2, u_4), e_3 = (u_3, u_5)$. Define

$$X := |T \cap \{(a_3, \overline{a_3}) \cup (a_4, \overline{a_4})\} \setminus \{e_1, e_2, e_3\}|,$$

and let E_1, E_2, E_3 be the indicator random variables for the edges e_1, e_2, e_3 respectively.

Let $\mu_1 := \{\mu | X = 0\}$. Since

$$\mathbb{E}_{T \sim \mu}[X] \leq 4(1 + \delta) - x_{e_1} - 2x_{e_2} - x_{e_3} \leq \frac{1}{5} + 4\delta,$$

by Lemma 6.10, this can only increase the probability of other edges by at most $\frac{1}{5} + 4\delta$. Now let $\mu_2 = \{\mu_1 | E_1 = 1, E_2 = 1, E_3 = 1\}$. By Fact 6.9 we have:

$$\mathbb{E}_{T \sim \mu_1}[E_1 + E_2 + E_3] \geq \mathbb{E}_{T \sim \mu}[E_1 + E_2 + E_3] \geq 3 - \frac{3}{20}.$$

Therefore,

$$
\begin{aligned}
\mathbb{P}_{T \sim \mu}[\mathcal{E}(e_2)] &= \mathbb{P}_{T \sim \mu_1}[\mathcal{E}(e_2)]\, \mathbb{P}_{T \sim \mu}[X = 0] \\
&\geq \frac{4}{5}\mathbb{P}_{T \sim \mu_2}[\mathcal{E}(e_2)]\, \mathbb{P}_{T \sim \mu_1}[E_1 = 1, E_2 = 1, E_3 = 1] \\
&\geq \frac{16}{25}\mathbb{P}_{T \sim \mu_2}[\mathcal{E}(e_2)]
\end{aligned}
$$

Figure 6: Cut classes $C_1, \ldots, C_6$ as described in Lemma 6.36

Thus it is sufficient to show $\mathbb{P}_{T \sim \mu_2}[\mathcal{E}(e_2)]$ is a constant. In any tree $T \sim \mu_2$, there are exactly 2 edges in the cuts $(a_3, \overline{a_3})$ and $(a_4, \overline{a_4})$. It remains to show that u_2 and u_4 also have an even degree. Since e_2 is sampled in $T \sim \mu_2$, we need to show $\mathbb{P}_{T \sim \mu_2}[|T \cap d_{u_4}(u_2)| = 1, |T \cap d_{u_2}(u_4)| = 1] = \Omega(1)$. By Corollary 6.10, the expected number of edges sampled from $d_{u_4}(u_2)$ or $d_{u_2}(u_4)$ in $T \sim \mu_2$ is between $\frac{3}{5}$ and $\frac{7}{5}$. This implies that

$$\frac{8}{5} \leq \mathbb{E}_{T \sim \mu_2}[|T \cap d_{u_4}(u_2)| + |T \cap d_{u_2}(u_4)|] \leq \frac{12}{5}$$

$$\Rightarrow \mathbb{P}_{T \sim \mu_2}[|T \cap d_{u_4}(u_2)| + |T \cap d_{u_2}(u_4)| = 2] \geq \frac{1}{4},$$

where the last inequality follows from Lemma 6.14. Thus by applying Lemma 6.21, setting $A := d_{u_4}(u_2), B := d_{u_2}(u_4), \epsilon := \frac{1}{4}, \alpha := \frac{3}{10}, \beta := \frac{3}{5}$, we obtain that

$$\mathbb{P}_{T \sim \mu_2}[\mathcal{E}(e_2)] = \mathbb{P}_{T \sim \mu_2}[|T \cap d_{u_4}(u_2)| = 1, |T \cap d_{u_2}(u_4)| = 1] = \frac{1}{200}.$$

Therefore $\mathbb{P}_{\mu}[\mathcal{E}(e_2)] \geq \frac{1}{500}$. $\qquad\qquad\square$

It remains to consider the cases where P contains cut classes with 4 atoms. Let $C_i \in P$, be a cut class with 4 atoms. Using the proof of Lemma 6.28, observe that if C_i has a non-connecting non-singleton atom, or a consecutive pair of singletons, it will be assigned good edges of fraction
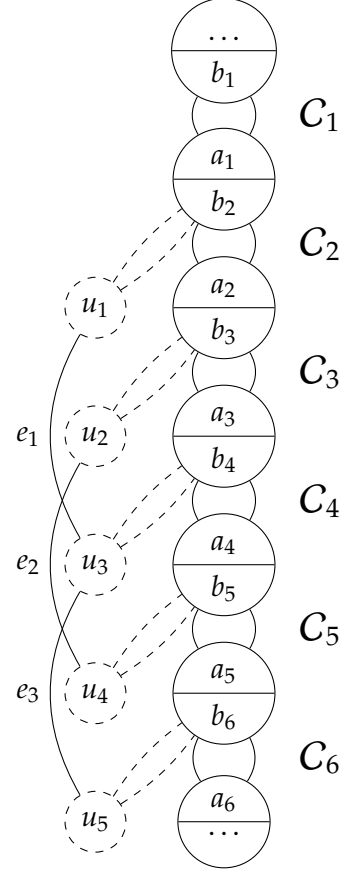
44

at least 3/4. In particular, suppose $\phi(C_i) = \{a_1, a_2, a_3, a_4\}$ and the atoms placed in the polygon representation according to their labels, and a_1 is the father-connector. Then C_i will be assigned good edges of fraction 3/4, unless a_2 and a_4 are (non-connecting) singletons and a_3 is the child-connector. It is immediate that each of the edges $x(a_i, a_{i+1}) \geq 1 - 2\delta$. We call such a cut class **exceptional 4-cut class**.

Lemma 6.38. *Let* $C_1 \to C_2 \to C_3 \in P$, *be 3 cut classes with child-connectors* a_1, a_2, a_3, *father-connectors* b_1, b_2, b_3, *such that* C_3 *has 4 atoms. If all of classes with 4 atoms in P are exceptional 4-cut class,* $\epsilon^* > 216\delta$, *and* $\delta < \frac{1}{1000}$, *then P is assigned good edges of fraction at least* $\frac{\epsilon^*}{4}$, *that are even with probability at least* $\frac{\epsilon^{*3}}{100000}$.

First we show that if $|a_2 \setminus a_3| > 2$ (or equivalently $b_3 \neq \overline{a_2}$) then we are done.

Lemma 6.39. *Let* $\epsilon^* > 216\delta$, *and* $\delta < 1/2000$, $P \in \pi_{\epsilon^*}$, $C_1 \to C_2 \in P$, *be 2 cut classes with child connectors* a_1, a_2, *father connectors* b_1, b_2 *such that* C_2 *is an exceptional 4-cut class with u and v the two singletons of* C_2. *Moreover, let* $1 \leq |a_1 \setminus a_2 \setminus \{u, v\}| \leq 2$. *For any* $w \in a_1 \setminus a_2 \setminus \{u, v\}$, *if* $x(u, w) > 0$ *then event that* $\mathcal{E}(u, w) \cap |T \cap E(b_2, \overline{b_2})| = 2$ *occurs with probability at least* $\frac{\epsilon^{*3}}{100000}$.

Proof. From Corollary 6.35, we have $x(w, \{u, v\}) \geq 1 - 8\delta - \frac{\epsilon^*}{4}$. Let $e = (u, v)$. It is sufficient to show that with a constant probability $T \sim \mu$ selects a Hamiltonian cycle on the atoms of C_2, and w has an even degree. Let $\mathcal{A}$ denote this event.

We use the proof strategy of Lemma 6.37. First we condition on T be an spanning tree on a_2 and contains exactly one edge from each of $E(u, a_2)$, and $E(v, a_2)$. Then we contract the vertices in $\{u, v, a_2\}$ and get a new vertex u'. Finally, we use Lemma G.2 to show that u' and w will have an even degree in the new measure and also that $|T \cap E(u, b_2)| = 1$ and $|T \cap E(v, b_2)| = 1$ with constant probability.

Let $X := |T \cap ((u, v) \cup E(a_2, b_2))|$, $Y := |T \cap E(a_2)|$, and $Z := |T \cap (E(u, a_2) \cup E(v, a_2))|$. Observe that

$$0 \leq \quad \mathbb{E}_{T \sim \mu}[X] \quad \leq 10\delta,$$

$$|a_2| - 1 - \frac{1}{2} - 4\delta \leq \quad \mathbb{E}_{T \sim \mu}[Y] \quad \leq |a_2| - 1,$$

$$2 - 4\delta \leq \quad \mathbb{E}_{T \sim \mu}[Z] \quad \leq 2,$$

where the second equation holds by the fact that $|a_2| \leq \frac{n}{2}$. Define $\mu_1 := \{\mu | X = 0, Y = |a_2| - 1\}$. By Corollary 6.10, conditioning on $X = 0, Y = |a_2| - 1$ can only decrease $\mathbb{E}_{T \sim \mu}[Z]$ by $\frac{1}{2} + 4\delta$. Observe that under measure μ_1, Z can take a maximum value of two. Thus $Z = 2$ is an upward event and when $Z = 2$, we must have $|T \cap E(u, a_2)| = |T \cap E(v, a_2)| = 1$. Let $\mu_2 := \{\mu_1 | Z = 2\}$. We have

$$\mathbb{P}_{T \sim \mu}[\mathcal{A}] = \mathbb{P}_{T \sim \mu_1}[\mathcal{A}] \, \mathbb{P}_{T \sim \mu}[X = 0, Y = |a_2| - 1]$$

$$\geq (\frac{1}{2} - 14\delta)\mathbb{P}_{T \sim \mu_2}[\mathcal{A}] \, \mathbb{P}_{T \sim \mu_1}[Z = 2]$$

$$\geq \frac{1}{3}(\frac{1}{2} - 18\delta)\mathbb{P}_{T \sim \mu_2}[\mathcal{A}]$$

It suffices to show $\mathbb{P}_{T \sim \mu_2}[\mathcal{A}] = \Omega(1)$. Observe that any tree $T \sim \mu_2$ contains a spanning tree inside $\{u, v, a_2\}$. Let us contract these vertices, and let u' be the new contracted vertex. If the event

45

$\{deg(u') = 2 \cap |T \cap E(u,b_2)| = 1\}$ occurs for $T \sim \mu_2$, then we get a Hamiltonian cycle inside C_2. Therefore, it is sufficient to show that:

$$\mathbb{P}_{T \sim \mu_2}[\mathcal{A}] \geq \mathbb{P}_{T \sim \mu_2}[deg(w) = 2, deg(u') = 2, |T \cap (u,b_2)| = 1] = \Omega(1).$$

Similar to Lemma 6.37, we use Lemma G.2. We set $u := u'$, $v := w$, $N_1(u) := E(u,b_2)$, $N_2(u) := E(v,b_2)$, $\epsilon := \frac{\epsilon^*}{3}$. We now check the conditions of the Lemma G.2. Observe, under μ_2, $\mathbb{E}_{T \sim \mu_2}[|T \cap E(a_2,b_2)|] = 0$. By corollary 6.35 we have

$$1 - 8\delta - \frac{\epsilon^*}{4} \leq x(w,u'), x(w,b_2), x(u,b_2), x(v,b_2) \leq 1 + 8\delta + \frac{\epsilon^*}{4}$$
$$2 - 4\delta \leq x(u,b_2) + x(v,b_2) \leq 2 + 4\delta$$

The expectation of the above values are equal to their x-values (within $O(\frac{1}{n})$ error). Since $P \in \pi_{\epsilon^*}$ we have $|\overline{b_2}| \geq n(\frac{1}{2} + \epsilon^*)$, and $|a_2| \leq n(\frac{1}{2} - \epsilon^*)$. Therefore, by Corollary 6.10, conditioning on $Y = |a_2| - 1$ may decrease the probability of edges outside of a_2 no more than $\frac{1}{2} - \epsilon^* + 4\delta$. Also conditioning on $Z = 2$ may decrease it by at most 4δ. Thus the LHS of the above equations can be decreased by at most $\frac{1}{2} - \epsilon^* + 8\delta$. Moreover, the RHS can be increased by at most 10δ and since $\epsilon^* \geq 216\delta$, the conditions of Lemma G.2 are satisfied. Therefore,

$$\mathbb{P}_{T \sim \mu_2}[deg(w) = 2, deg(u') = 2, |T \cap E(u,b_2)| = 1] \geq \frac{\epsilon^{*3}}{10000}$$
$$\Rightarrow \mathbb{P}_\mu[\mathcal{A}] \geq \frac{1}{3}(\frac{1}{2} - 18\delta)\frac{\epsilon^{*3}}{10000} \geq \frac{\epsilon^{*3}}{100000}.$$

$\square$

The lemma implies the following Corollary.

Corollary 6.40. *Let $P \in \pi_0$, $C_1 \to C_2 \to C_3$ be 3 cut classes with child-connectors a_1, a_2, a_3, and father-connectors b_1, b_2, b_3. If C_3 is an exceptional 4-cut class, C_2 is a 2-cut class, $|a_2 \setminus a_3| = 2$, $1 \leq |a_1 \setminus a_2| \leq 2$, and $\epsilon^* > 216\delta, \delta < \frac{1}{1000}$, then P is assigned good edges of fraction at least $\frac{\epsilon^*}{4}$.*

Proof. Observe that $3 \leq |a_1 \setminus a_3| \leq 4$. Since C_2 contains two atoms and $|a_2 \setminus a_3| = 2$, the cuts (a_2, b_2) and $(b_3, \overline{b_3})$ are the same. Therefore, if $|T \cap E(b_3, \overline{b_3})|$ is even, $|T \cap E(a_2, b_2)|$ is also even. Thus, we can ignore C_2 and assume $C_1 \to C_3$ and apply Lemma 6.39. Therefore, the edges connecting the singleton of C_3 to vertices in $a_1 \setminus a_2$ are good. $\square$

Now suppose $|a_2 \setminus a_3| = 2$ (or equivalently $b_3 = \overline{a_2}$). In the next lemma we show that if C_2 also is an exceptional 4-cut class then we are done.

Lemma 6.41. *Let $P \in \pi_0$, $C_1 \to C_2$, be exceptional 4-cut classes with child connectors a_1, a_2 and father-connectors b_1, b_2. If $|a_1 \setminus a_2| = 2$, $\epsilon^* > 50\delta$, and $\delta < \frac{1}{10000}$, then T is a Hamiltonian cycle in $G(\phi(C_1))$ and $G(\phi(C_2))$ with a probability $\frac{1}{4000}$, thus P is assigned good edges of fraction at least $3/2$.*
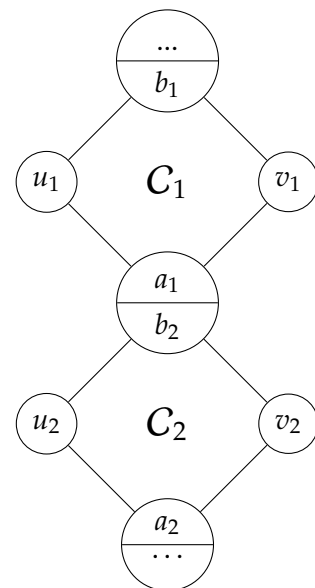

Figure 7: Exceptional 4 Cut classes C_1, C_2 described in Lemma 6.41

Proof. Let u_1, v_1 be the singletons of C_1, and u_2, v_2 be the two singletons of C_2 (see Figure 7). Since $|a_1 \setminus a_2| = 2$, we have $a_1 = a_2 \cup \{u_2, v_2\}$, and $a_1 = \overline{b_2}$. Hence, $x(u_1, a_1) = x(u_1, \overline{b_2})$ and $x(v_1, a_1) = x(v_1, \overline{b_2})$. Thus:

$$1 - 2\delta \leq x(u_1, \overline{b_2}), x(v_1, \overline{b_2}) \leq 1 + 2\delta$$
$$1 - 2\delta \leq x(u_2, b_2), x(v_2, b_2) \leq 1 + 2\delta.$$

Since $x(d(b_2)) \leq 2(1 + 2\delta)$, we obtain $x(\{u_2, v_2\}, \{u_1, v_1\}) \geq 2 - 8\delta$. Let $\mathcal{A}$ be the event that T is a Hamiltonian cycle in both $G(\phi(C_1))$ and $G(\phi(C_2))$. We show that $\mathcal{A}$ occurs with probability at least $\frac{1}{4000}$, thus all edges in the cut $(\{u_2, v_2\}, \{u_1, v_1\}$ are good.

The proof is basically a generalization of Lemma 6.32. Let $X := |T \cap (E(a_1, b_1) \cup (u_1, v_1))\}|$ and $Y := |T \cap E(a_2, b_2) \cup (u_2, v_2)\}|$, be the indicator random variables for the edges between the non-consecutive atoms of C_1 and C_2. Observe that $\mathbb{E}_\mu[X], \mathbb{E}_\mu[Y] \leq 8\delta$. Let $\mathcal{A}_2$ be the event that T contains a Hamiltonian path through atoms of $\{u_2\}, a_2, \{v_2\}$ and $|T \cap E(a_2)| = |a_2| - 1$ and $Y = 0$. The proof is structured as follows: first we condition on $X = 0$, then on $\mathcal{A}_2$. Since u_1, v_1 are singletons in C_1, and T contains a spanning tree in $a_2 \cup \{u_2, v_2\}$, T must have exactly two edges in the cut $(\{u_2, v_2\}, \{u_1, v_1\})$, one (wlog) (u_2, u_1), and the other (v_2, v_1). Finally, we show that exactly one edge is sampled from each of $E(u_1, b_1)$ and $E(v_1, b_1)$ and we are done.

Let $\mu_1 : \{\mu| X = 0\}$. Since $8\delta < \frac{1}{1000}$, following the proof of Lemma 6.32, $\mathcal{A}_2$ occurs with probability $\frac{9}{10} \frac{9}{20} \frac{2}{5}$. In particular, let $\mu_2 := \{\mu_1 | \mathcal{A}_2\}$, Since any tree $T \sim \mu_2$ contains a spanning tree in $\overline{b_2} = a_2 \cup \{u_2, v_2\}$ we may contract all these vertices as a new vertex u'. Since $X = 0$, T does not have any edge in cut (u', b_1), thus u' is only connected to u_1 and v_1. Therefore $\mathbb{P}_{T \sim \mu_2}[deg(u') \leq 2] = 1$. Moreover, u' can have at most one edge to each of the vertices u_1, v_1. Hence the event $deg(u') = 2$ is an upward closed event, which implies T has exactly one edge in all of the 4 edge sets $E(u_2, b_2), E(v_2, b_2), E(u_1, a_1), E(v_1, a_1)$.

Define $\mu_3 := \{\mu_2 | deg(u') = 2\}$. Since,

$$\frac{3}{2} - 12\delta \leq \mathbb{E}_{T \sim \mu_2}[deg(u')] \leq 2.$$

Therefore, by Markov's inequality $\mathbb{P}_{T \sim \mu_2}[deg(u') = 2] = \Omega(1)$. It remains to show that

$$\mathbb{P}_{T \sim \mu_3}[|T \cap E(u_1, b_1)| = 1, |T \cap E(v_1, b_1)| = 1]$$

is a constant. By Corollary 6.10, the probability of the edges in $E(G) \setminus E(a_1) \cup \{u_1, v_1\})$ can only be decreased by $\frac{1}{2} + 12\delta$ under the measure μ_2. Now we use Lemma 6.21 to show that the event $|T \cap E(u_1, b_1)| = 1, |T \cap E(v_1, b_1)| = 1$ also occurs with a constant probability, and we have spanning tree that has a Hamiltonian cycle through the atoms of C_1 and C_2 with a constant probability. Set $A := E(u_1, b_1)$, $B := E(v_1, b_1)$, $\alpha = \frac{1}{3}$, $\beta = \frac{1}{2} - 12\delta$, $\epsilon = \frac{1}{5}$ from Lemma 6.14. This implies that

$$\mathbb{P}_{T \sim \mu_3}[|T \cap E(u_1, b_1)| = 1, |T \cap E(v_1, b_1)| = 1] \geq \frac{1}{300} \Rightarrow \mathbb{P}_\mu[\mathcal{A}] \geq \frac{1}{4000}$$

$\square$

The lemma implies the following Corollary.

Corollary 6.42. *Let $P \in \pi_0$, $C_1 \to C_2 \to C_3$ where C_1 and C_3 are exceptional 4-cut classes with child connectors a_1, a_3 and father-connectors b_1, b_3 and C_2 is a 2-cut class with child-connector a_2. If $|a_1 \setminus a_3| = 2$ and $\epsilon^* > 50\delta$, and $\delta < \frac{1}{10000}$, the P is assigned even edges of fraction at least $\frac{1}{3}$.*

Proof. Observe that $|a_1 \setminus a_3| = 2$ implies that $a_2 = a_1$. We ignore C_2 and assume $C_1 \to C_3$ and apply Lemma 6.41. Since T is a Hamiltonian cycle on C_1 and C_3 with constant probability, the edges connecting the singleton atoms of C_1 and C_3 are good. □

Now we are ready to prove Lemma 6.38:

Proof of Lemma 6.38. If $|a_2 \setminus a_3| > 4$ then we are done since the thread is assigned at least $\frac{1}{3}$ trivial good edges. If $3 \le |a_2 \setminus a_3| \le 4$ then applying Lemma 6.39 we are done. Hence, suppose $|a_2 \setminus a_3| = 2$. If C_2 has four atoms then we are done by Lemma 6.41. Thus suppose that C_2 has two atoms. Now we do a case analysis depending on the value of $|a_1 \setminus a_2|$.

If $|a_1 \setminus a_2| = 0$. Since $C_1 \ne C_2$, we have C_1 is an exceptional 4-cut class and $|a_1 \setminus a_3| = 2$. Applying Corollary 6.42, we are done.

Now suppose $1 \le |a_1 \setminus a_2| \le 2$. Now applying Corollary 6.40, we are done.

□

References

[1] N. Aggarwal, N. Garg, and S. Gupta. A 4/3-approximation for tsp on cubic 3-edge-connected graphs. *CoRR*, abs/1101.5586, 2011.

[2] S. Arora. Polynomial time approximation schemes for euclidean traveling salesman and other geometric problems. In *Journal of the ACM*, pages 2–11, 1996.

[3] S. Arora, M. Grigni, D. Karger, P. Klein, and A. Woloszyn. A polynomial-time approximation scheme for weighted planar graph tsp. In *Proceedings of the ninth annual ACM-SIAM symposium on Discrete algorithms*, pages 33–41. Society for Industrial and Applied Mathematics, 1998.

[4] A. Asadpour, M. X. Goemans, A. Madry, S. Oveis Gharan, and A. Saberi. O(logn/loglogn) approximation to the asymmetric traveling salesman problem. *Proceeding of 21st ACM Symposium on Discrete Algorithms*, 2010.

[5] A. A. Benczúr. A representation of cuts within 6/5 times the edge connectivity with applications. In *FOCS*, pages 92–102, 1995.

[6] A. A. Benczúr. *Cut structures and randomized algorithms in edge-connectivity problems.* PhD thesis, 1997. Supervisor-Goemans, Michel X.

[7] A. A. Benczúr and M. X. Goemans. Deformable polygon representation and near-mincuts. *Building Bridges: Between Mathematics and Computer Science, M. Groetschel and G.O.H. Katona, Eds., Bolyai Society Mathematical Studies*, 19:103–135, 2008.

[8] J. Borcea, P. Branden, and T. M. Liggett. Negative dependence and the geometry of polynomials. *Journal of American Mathematical Society*, 22:521–567, 2009.

[9] S. Boyd, R. Sitters, S. van der Ster, and L. Stougie. Tsp on cubic and subcubic graphs. In *to appear in fifteenth Conference on Integer Programming and Combinatorial Optimization*, 2011.

[10] S. C. Boyd and W. R. Pulleyblank. Optimizing over the subtour polytope of the travelling salesman problem. *Math. Program.*, 49:163–187, 1991.

[11] R. Burton and R. Pemantle. Local characteristics, entropy and limit theorems for spanning trees and domino tilings via transfer-impedances. *Ann. Probab.*, 21(math.PR/0404048):1329–1371, 1993.

[12] R. D. Carr and S. Vempala. On the held-karp relaxation for the asymmetric and symmetric traveling salesman problems. *Math. Program.*, 100(3):569–587, 2004.

[13] N. Christofides. Worst case analysis of a new heuristic for the traveling salesman problem. Report 388, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh, PA, 1976.

[14] J. N. Darroch. On the distribution of the number of successes in independent trials. *Ann. Math. Stat.*, 36:1317–1321, 1964.

[15] E. D. Demaine, M. Hajiaghayi, and K. ichi Kawarabayashi. Contraction decomposition in h-minor-free graphs and algorithmic applications. In *to appear in fourty-third ACM Symposium on Theory of Computing*, 2011.

[16] E. D. Demaine, M. Hajiaghayi, and B. Mohar. Approximation algorithms via contraction decomposition. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 278–287, Philadelphia, PA, USA, 2007. Society for Industrial and Applied Mathematics.

[17] E. Dinits, A. Karzanov, and M. Lomonosov. On the structure of a family of minimal weighted cuts in graphs. *Studies in Discrete Mathematics (in Russian), ed. A.A. Fridman, 290-306, Nauka (Moskva)*, 1976.

[18] J. Edmonds and E. Johnson. Matching: A well-solved class of integer linear programs. In e. R. Guy et al., editor, *Proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications*, pages 82–92, 1970.

[19] T. Feder and M. Mihail. Balanced matroids. In *Proceedings of the twenty-fourth annual ACM symposium on Theory of Computing*, pages 26–38, New York, NY, USA, 1992. ACM.

[20] T. Fleiner and A. Frank. A quick proof for the cactus representation of mincuts. Technical Report QP-2009-03, Egerváry Research Group, Budapest, 2009. `www.cs.elte.hu/egres`.

[21] D. Gamarnik, M. Lewenstein, and M. Sviridenko. An improved upper bound for the tsp in cubic 3-edge-connected graphs. *Operations Research Letters*, 33(5):467–474, 2005.

[22] L. J. Gleser. On the distribution of the number of successes in independent trials. *Ann. Probab.*, 3(1):182–188, 1975.

[23] M. X. Goemans. Minimum bounded degree spanning trees. In *FOCS*, pages 273–282, 2006.

[24] M. Grigni, E. Koutsoupias, and C. Papadimitriou. An approximation scheme for planar graph tsp. In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, page 640. IEEE Computer Society, 1995.

[25] G. H. Hardy, J. E. Littlewood, and G. Polya. *Inequalities*. Cambridge Univ. Press, 1952.

[26] M. Held and R. Karp. The traveling salesman problem and minimum spanning trees. *Operations Research*, 18:1138–1162, 1970.

[27] W. Hoeffding. On the distribution of the number of successes in independent trials. *Ann. Math. Statist.*, 27:713–721, 1956.

[28] P. N. Klein. A linear-time approximation scheme for planar weighted tsp. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 647–657, Washington, DC, USA, 2005. IEEE Computer Society.

[29] R. Lyons and Y. Peres. *Probability on Trees and Networks*. Cambridge University Press, 2010.

[30] C. L. Monma, B. S. Munson, and W. R. Pulleyblank. Minimum-weight two-connected spanning networks. *Math. Program.*, 46:153–171, 1990.

[31] H. Nagamochi, K. Nishimura, and T. Ibaraki. Computing all small cuts in an undirected network. *SIAM Journal on Discrete Mathematics*, 10:469–481, 1994.

[32] C. H. Papadimitriou and S. Vempala. On the approximability of the traveling salesman problem. *Combinatorica*, 26(1):101–120, 2006.

[33] C. H. Papadimitriou and M. Yannakakis. The traveling salesman problem with distances one and two. *Mathematics of Operations Research*, 18:1–11, February 1993.

[34] R. Pemantle. Towards a theory of negative dependence. *J. Math. Phys*, 41:1371–1390, 1999.

[35] J. Pitman. Probabilistic bounds on the coefficients of polynomials with only real zeros. *J. Comb. Theory Ser. A*, 77:279–303, February 1997.

[36] A. Schrijver. *Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer, 2003.

[37] D. B. Shmoys and D. P. Williamson. Analyzing the held-karp tsp bound: a monotonicity property with application. *Information Processing Letters*, 35:281–285, September 1990.

[38] Y. H. Wang. On the number of successes in independent trials. *Statistica Sinica*, 3:295–312, 1993.

[39] L. A. Wolsey. Heuristic analysis, linear programming and branch and bound. In *Combinatorial Optimization II*, volume 13 of *Mathematical Programming Studies*, pages 121–134. Springer Berlin Heidelberg, 1980.

A   Christofides Algorithm and Subtour Elimination LP

In this section we prove the following lemma for completeness.

Lemma A.1. *[39, 37] Let H be the Eulerian subgraph returned by the Christofides algorithm. Then $c(H) \le \frac{3}{2}c(x)$ where x is the optimal solution to Held-Karp linear program.*

Proof. Observe that $H = F \cup J$ where F is the minimum spanning tree and J is the minimum cost matching on the odd-degree vertices of F. The lemma follows from the following two claims.

Claim A.2. $c(F) \le c(x)$.

Proof. Since $(1 - \frac{1}{n})x$ is in the spanning tree polyhedron, we have $c(F) \le (1 - \frac{1}{n})c(x) \le c(x)$. □

Claim A.3. *Let T be any set of even number of vertices and let J be the minimum cost T-join on T. Then $c(J) \le c(x)$.*

Proof. It follows from Edmonds and Johnson [18] that the integrality gap of the T-join polytope is 1. Therefore, the minimum cost T-join on T is the optimum solution to to LP_{T-join} (see Figure 1).

Since $x/2$ is a feasible solution to the LP_{T-join} since $x(d(S))/2 \ge 1$ for each set $S \subset V$. Thus $c(J) \le c(x)/2$. □

This completes the proof of the lemma. □

B   Proof of Theorem 1.1

In this section, we prove Theorem 1.1 using Theorem 3.1.

Case 1: x has at least $\epsilon_1 n$ good edges.

Lemma B.1. *The expected cost of F is at most $c(x)$.*

Proof. Since $\mathbb{P}[e \in F] = z_e \le x_e$, we have $\mathbb{E}[c(F)] = \sum_{e \in E^*} c(e)\mathbb{P}[e \in F] \le c(x)$. □

Hence, we need to bound the cost of the T-join. The rest of the argument depends on which case of Theorem 3.1 holds.

Lemma B.2. *Let x be a fractional solution of $(LP_{subtour})$, $E^* \subset E$ be the set of good edges. If there are a lot of good edges, that is if $x(E^*) \ge \epsilon_1 n$, then the expected cost of the smallest Eulerian tour is at most $3/2 - \frac{\epsilon_1 \delta \rho}{4(1+\delta)}$.*

Proof. We provide a fractional solution to the (LP_{T-join}) (Figure 1 to make it Eulerian. For any edge $e \in E$ if e is contained in at least one odd $(1 + \delta)$ near minimum cut $(S, \bar{S})$, set $y_e = x_e/2$,

otherwise set $y_e = x_e/2(1 + \delta)$). Observe that a cut $(S, \overline{S})$ is odd in F iff $|S \cap T|$ is odd. Therefore, y is indeed a fractional solution of (LP_{T-join}). Now, to bound the cost of y in Step 3 note:

$$
\begin{aligned}
\mathbb{E}\left[c(y)\right] &\leq \frac{c(x)}{2} - \sum_{e \in E} x_e c(e) \mathbb{P}\left[e \notin \text{ odd near minimum cuts}\right] \left(\frac{1}{2} - \frac{1}{2(1 + \delta)}\right) \\
&\leq \frac{c(x)}{2} - \frac{\delta}{2(1 + \delta)} \sum_{e \in E^*} x_e \rho \\
&\leq c(x)(\frac{1}{2} - \frac{\epsilon_1 \delta \rho}{4(1 + \delta)}).
\end{aligned}
$$

The second inequality holds because $c(e) \geq 1$ for all $e \in E$, and the last one because $c(x) \leq 2n$. Since the T-join polytope is integral [18], the minimum cost integral T-join costs at most $c(y)$. By adding the edges of minimum T-join to F we obtain an Eulerian tour of expected total weight at most $c(x)(\frac{3}{2} - \frac{\epsilon_1 \delta \rho}{4(1+\delta)})$. $\qquad\square$

The above argument bounds the cost of the tour in expectation. By sampling a tree $\log n$ times and choosing the best solution, one can obtain an Eulerian tour of cost at most $c(x)(\frac{3}{2} - \frac{\epsilon_1 \delta \rho}{4(1+\delta)})$ with high probability.

Case 2: x is nearly integral. In this case, we bound the cost of the tree F and T-join J together and prove the following lemma. The construction of the fractional T-join in the lemma is similar to a construction by Monma, Munson and Pulleyblank [30].

Lemma B.3. *Let x be a fractional solution of $(LP_{subtour})$. If x contains at least $(1 - \epsilon_2)n$ edges of fraction greater than $1 - \gamma$, then the tour computed in Algorithm 2, step 5 is at most $c(x)(\frac{4}{3} + 2\epsilon_2 + 4\gamma)$.*

Proof. Let $I' = \{e \mid x_e > 1 - \gamma\}$ be the set of nearly integral edges, and let F' be the minimum cost spanning graph that contains I'. Since G_0 is connected, I' can be augmented into a connected graph using only edges of cost 1. Hence, we have $c(F') = c(I') + |F' \setminus I'| \leq \frac{\sum_{e \in I'} c(e) x_e}{1 - \gamma} + |F' \setminus I'|$.

Recall that F is a minimum cost spanning subgraph of F'. Because of the constraints of LP and since $\gamma < 1/3$, it is easy to see that I' consists of disjoint cycles and paths and the length of each cycle is at least $\frac{1}{\gamma}$. Therefore, F will have at least $n(1 - \epsilon_2)(1 - \gamma)$ edges from I'. Therefore, $|F \setminus I'| \leq n(\epsilon_2 + \gamma)$. Let us set $I = I' \cap F$.

Let T denote the set of odd vertices in F. Again, we bound the cost of T-join by constructing a fractional solution to the LP_{T-join}, and then invoking the integrality of the T-join polytope. Let

$$
\begin{cases}
y_e = \frac{x_e}{3(1-\gamma)} & \text{if } e \in I \\
y_e = 1 & \text{if } e \in F \setminus I \\
y_e = x_e & \text{otherwise.}
\end{cases}
$$

We first show that y is feasible for LP_{T-join}. Let $(U, \overline{U})$ be any cut which has an odd number of vertices of T in U (equivalently, a cut that has an odd number of edges of F). If there exists an $e \in (F \setminus S) \cap E(U, \overline{U})$, then $y(d(U)) \geq y_e \geq 1$ and the constraint is satisfied. Otherwise, we have $E(U, \overline{U}) \cap F \subseteq S$. Therefore since $(U, \overline{U})$ has an odd number of edges F, and $I \subset F$, $(U, \overline{U})$ must

contain an odd number of edges of I. By the values assigned to the edges in y, we have

$$\begin{cases} y(d(U)) \geq x(d(U) \setminus I) \geq 1 & \text{if } |I \cap E(U, \overline{U})| = 1 \\ y(d(U)) \geq y(d(U) \cap I) \geq 3\frac{1}{3(1-\gamma)}(1-\gamma) = 1 & \text{if } |I \cap E(U, \overline{U})| \geq 3 \end{cases}$$

thus y is a feasible solution of (LP_{T-join}).

Now we bound the cost of the final Eulerian subgraph which will be at most $c(F) + c(y)$

$$\begin{aligned} c(F) + c(y) & \leq \frac{c(x(I))}{1-\gamma} + c(F \setminus I) + c(F \setminus I) + c(x(E \setminus F)) + \frac{c(x(I))}{3(1-\gamma)} \\ & \leq \frac{4c(x(I))}{3(1-\gamma)} + n(2\epsilon_2 + 2\gamma) + c(x(E \setminus I)) \\ & \leq \frac{4c(x)}{3(1-\gamma)} + c(x)(2\epsilon_2 + 2\gamma) \leq c(x)(\frac{4}{3} + 2\epsilon_2 + 4\gamma). \end{aligned}$$

The one to the last inequality follows from the fact that $c(x) \geq n$, and the last inequality follows from $\gamma < 1/3$. □

C Proofs from Section 4

First we prove the following simple claim which is useful in later proofs.

Lemma C.1. *Any collection of cuts satisfies the following statements:*

1. *Among any set $S = \{C_{i_1}, C_{i_2}, \ldots, C_{i_l}\}$ of cut classes, we can find a set M of atoms such that the atoms in M are pairwise disjoint, and*

$$|M| \geq -2(l-1) + \sum_{j=1}^{l} |\phi(C_{i_j})|.$$

2. *If C is the set of $(1 + \delta)$ near minimum cuts of $G = (V, E, x)$, and $\delta < 1/10$, then any edge e contained in a cut class C_i is contained in at most $|\phi(C_i)|^2$ cuts of C_i.*

Proof. The first statement can be proved by considering the properties of the cactus defined on cut classes of C. Let's consider the tree $\Gamma(C)$ (see Definition 4.21) on the cycles of the cactus (cut classes) and let's make one of the classes of S to be the root of the tree. Starting from the root let's traverse the tree (e.g. using Depth First Search), and keep the invariant that M always contain a set of atoms that are pairwise disjoint. At the first step we add all of the atoms of the root to M. At the time of visiting a cut class $C_{i_j} \in S$, it is not hard to see that all except one of the atoms of $\phi(C_{i_j})$ are a subset of an atom of M. Therefore, we can add all except one of its atoms to M and remove their superset from M. Since each time we add $|\phi(C_i)| - 1$ atoms and we delete one of the previous atoms, after $l - 1$ steps we have $|M| \geq -2(l-1) + \sum_{C_{i_j}} |\phi(C_i)|$.

Nagamochi et al. [31] show that the number of $4/3$ near minimum cuts of any graph is at most $\binom{n}{2}$. Therefore, the second statement can be proved simply by applying this result to the graph $G(\phi(C_i))$. □

Proof of **Lemma 4.6.**   We show that a constant fraction of edges in G are not incident to any of the atoms of $L(\tau)$. By Corollary 4.15, for any cut class C_i, the total fraction of edges in $G(\phi(C_i))$ is at most $|\phi(C_i)|(1 + \beta\delta)$, where $\beta := 3$. Thus the total fraction of edges that are contained in at least one of the large cut classes is no more than

$$|L(\tau)|(1 + \beta\delta) < n(1 - \epsilon)(1 + \beta\delta) \leq n - n(\epsilon - \beta\delta).$$

Therefore, since the total sum of the fraction of edges in G is n (i.e. $x \in LP_{subtour}$) we have $x(E_S) \geq (\epsilon - \beta\delta)n$. □

Proof of **Lemma 4.4.**   We show that for any $\delta < \frac{1}{100}$, if $L(\tau) \geq (1 - \epsilon)n$, then x contains at least $n(1 - \alpha\sqrt{\delta} - 5\epsilon - \frac{12}{\tau-2}) \geq n(1 - \alpha\sqrt{\delta} - 17\epsilon)$ edges of fraction $1 - \alpha'\sqrt{\delta}$, where $\alpha := 20, \alpha' := 4$. Observe that for any cut class C_i, a near integral edge in $G(\phi(C_i))$ incident to two singleton atoms is corresponding to a near integral edge in G. On the other hand, we show that most of the atoms in $L(\tau)$ are singletons. Therefore, most of the near integral edges in $G(\phi(C_i))$ are corresponding to the near integral edges in G. Since by Corollary 4.15, for any cut class C_i, $G(\phi(C_i))$ contains are at least $|\phi(C_i)|(1 - \alpha\sqrt{\delta})$ edges of fraction at least $1 - \alpha'\sqrt{\delta}$, there exists a large number of edges of fraction $1 - \alpha'\sqrt{\delta}$ between the vertices of G.

Let L be the number of large cut classes. By property (1) of Lemma C.1, $|L(\tau)| \leq n + 2L$. But then we have $L\tau \leq n + 2L$ and therefore $L \leq \frac{n}{\tau-2}$. By the first property in Lemma C.1, we can find at least $|L(\tau)| - 2L \geq n(1 - \epsilon - \frac{2}{\tau-2})$ atoms in $L(\tau)$ that are mutually disjoint. But these atoms define a partition of the ground set V, therefore at least $n(1 - 2\epsilon - \frac{4}{\tau-2})$ of them are singletons. Therefore, the number of non-singleton atoms of $L(\tau)$ is at most $n(2\epsilon + \frac{6}{\tau-2})$.

On the other hand, by Corollary 4.15, there are

$$\sum_{C_i : |\phi(C_i)| \geq \tau} |\phi(C_i)|(1 - \alpha\sqrt{\delta}) = |L(\tau)|(1 - \alpha\sqrt{\delta}) \geq n(1 - \epsilon - \alpha\sqrt{\delta})$$

edges of fraction $1 - \alpha'\sqrt{\delta}$ in graphs $G(\phi(C_i))$ for large cut classes C_i. Hence, at least $n(1 - \alpha\sqrt{\delta} - 5\epsilon - \frac{12}{\tau-2})$ of these edges are incident only to singleton atoms. But edges adjacent to two singletons are corresponding to actual edges of G. We conclude that there are $n(1 - \alpha\sqrt{\delta} - 5\epsilon - \frac{12}{\tau-2})$ edges of fraction $1 - \alpha'\sqrt{\delta}$ in G. Then lemma follows from the assumption $\epsilon > \frac{1}{\tau-2}$. □

D   Proof of Theorem 4.26

For simplicity, in the rest of the argument we will not consider root as a special case. Indeed if root is a large cut class, then similar to other large cut classes, we ignore the edges contained in C_r, and we do not assign any good edge to it. Otherwise, if it a small cut class, only a constant number of edges ($O(\tau)$) are included in it, thus removing the good edges contained in the root will not have a big effect.

We prove this theorem in several steps. Suppose the sum of fraction of edges in E_{S_T} is very small (say $o(n)$). First we prove that this implies $x(E_{S_N}) = \Omega(n)$. Then we show that $\Gamma(C)$ contains $\Omega(n)$ nodes. Finally, we prove that the tree can not have many leaves, thus it should contain many long paths (or threads).

Lemma D.1. *If $\epsilon > \frac{32}{\tau}$, $x(E_S) \geq \epsilon n$, and $x(E_{S_T}) < \frac{\epsilon n}{20}$, then $x(E_{S_N}) \geq \frac{\epsilon n}{5}$.*

Proof. We prove by contradiction, suppose $x(E_{S_N}) < \frac{\epsilon n}{5}$. Define $F := E_S \setminus E_{S_T} \setminus E_{S_N}$. By assumption $x(F) \geq \frac{3\epsilon n}{4}$. Let us describe the properties of the edges in F. For an edge $e \in F$, since $e \in E_S$ but $e \notin E_{S_N}$, e must only contained in trivial near minimum cuts. Moreover, since $e \notin E_{S_T}$, e is not a good edge. Therefore, by Proposition 4.17, we have $x_e = \frac{1}{2} \pm \frac{1}{8000}$.

Note that if the graph originally does not have any almost half edges (edges e for which x_e is close to 1/2), we get $F = \emptyset$, and we reach to a contradiction. Since by proposition 4.17, from any adjacent pair of half edges, at least one of them is good, the edges in F are not adjacent to each other. Let $e = (u, v) \in F$, since there is not any near minimum cut (other than the trivial cuts) containing e, u and v must be contained in the same atom in all of the non-trivial cut classes. Hence, we can contract e (and all of the edges in F), without changing the structure of tree hierarchy.

Let U be the contracted set of vertices. Recall that for any vertex $u \in U$, $d(u)$ is the set of edges adjacent to u. Since u is a contraction of an almost half edge, we have $x(d(u)) = 3 \pm \frac{1}{4000}$

Let $d'(u) \subseteq d(u)$ be the set of edges adjacent to u that are contained in at least one large cut class. We have

$$d(u) \setminus d'(u) \subseteq E_S \setminus F. \tag{33}$$

Define $W := \{u \in U : x(d'(u)) \geq 2.5 - \frac{1}{4000}\}$. We drop this term of $\frac{1}{4000}$ in the rest of the calculations. This affects the constant very slightly. Then

$$\frac{\epsilon n}{4} \geq x(E_S \setminus F) \geq \frac{1}{2} \sum_{u \in U} (3 - x(d'(u))) \geq \sum_{u \in U \setminus W} \frac{1}{4} = \frac{|U \setminus W|}{4}, \tag{34}$$

where the second inequality holds by equation (33), and noting that each edge $e \in E_S \setminus F$ can be counted at most twice in the RHS. Since $x(F) \geq \frac{3\epsilon n}{4}$, we get $|U| = |F| \geq \frac{6\epsilon n}{4}$. Using equation (34) we get $|W| \geq \frac{\epsilon n}{2}$. We show that this implies the existence of large degree atoms in large cut classes and we reach to a contradiction with Corollary 4.15.

For a set S of atoms let $D(S) := \sum_{a \in S} x(d(a))$ be sum of fraction of edges adjacent to the atoms in S. To reach a contradiction, we use a double counting argument for the value $D(L(\tau))$. Since by Corollary 4.15, $D(\phi(C_i)) = x(E(G(\phi(C_i)))) \leq 2(1 + 3\delta)|\phi(C_i)|$ for any cut class C_i, we have

$$D(L(\tau)) \leq 2(1 + 3\delta)|L(\tau)| \tag{35}$$

On the other hand, we show that there is a set $S^* \subset L(\tau)$ such that $D(S^*) \geq \frac{5\epsilon n}{4}$, while $|S^*| \leq 9D(S^*)/20$. Since the size of the minimum cut of G is 2, for any $a \in S$, we have $x(d(a)) \geq 2$. Therefore

$$D(L(\tau)) = D(S^*) + D(L(\tau) \setminus S^*) \geq D(S^*) + 2(|L(\tau)| - |S^*|) \geq \frac{D(S^*)}{10} + 2|L(\tau)| \geq 2|L(\tau)| + \frac{\epsilon n}{8}.$$

This is however a contradiction with equation 35, since $|L(\tau)| \leq n$, and $\epsilon > 48\delta$.

It remains to show the existence of S^*. Consider the set of large cut classes $C_1, \ldots, C_l$, by Lemma C.1, part 1, there is a subset $S_1 \subset L(\tau)$ such that atoms in S_1 are pairwise disjoint, and $|S_1| \geq |L(\tau)| - 2l$.

Let $S_2 := \{a \in S_1 : \exists w \in W, w \in a\}$ be the set of atoms in S_1, each contain at least one vertex of W. Set $S^* = S_2 \cup (L(\tau) \setminus S_1)$. Then

$$|S^*| = |S_2| + |L(\tau) \setminus S_1| \leq |W| + 2l \leq |W| + \frac{2n}{\tau} \leq \frac{9|W|}{8},$$

where the last inequality holds by the fact that $\tau \geq 32/\epsilon$.

Let $w \in W$, observe that if w is contained in an atom of $a \in S_1$, then a is the smallest part in $L(\tau)$ that contains w, thus $d'(w) \subseteq d(a)$. Hence, if an atom $a \in S_2$ contains k vertices of W we have $x(d(a)) \geq 2.5k$. If w is not contained in any atoms of S_1, then it is contained in some atoms of $L(\tau) \setminus S_1$, thus the edges in $d'(w)$ will be adjacent to those atoms. Consequently,

$$D(S^*) \geq \sum_{w \in W} x(d'(w)) \geq 2.5|W|.$$

Hence $D(S^*) \geq \frac{5\epsilon n}{4}$, while $|S^*| \leq 9D(S^*)/20$.

$\square$

Lemma D.2. *Let C be the collection of $(1 + \delta)$ near minimum cuts, with small non trivial cut classes $C_1, \ldots, C_s$. If $x(E_{S_I}) + x(E_{S_T}) \leq \frac{\epsilon n}{8}$, and $x(E_{S_N}) \geq \epsilon n$, then $\Gamma(C)$ contains at least $\frac{\epsilon n}{6}$ nodes.*

Proof. From Corollary 4.15, it is straightforward to prove that the number of small cut classes is at least $\Omega(\frac{\epsilon n}{\tau})$. We prove a much stronger bound here. Since $x(E_{S_N}) \geq \epsilon n$, the sum of the fraction edges inside small cut classes is at least ϵn. By Corollary 4.15:

$$\sum_{i=1}^{s} |\phi(C_i)| \geq \epsilon n. \tag{36}$$

We show that each small cut class should have a large fraction of good edges, unless it has a high degree in $\Gamma(C)$. The latter implies the sum of the degrees of small cut classes (and thus the number of edges of $\Gamma(C)$ is $\Omega(\epsilon n)$. Since $\Gamma(C)$ is a tree, it must have at least that many nodes.

For each $1 \leq i \leq s$, let d_i be the degree of small cut class C_i in $\Gamma(C)$, and let y_i be the sum of the fraction of good edges in E_{S_I} or E_{S_T} assigned to C_i. Let C_i be a small cut class; by lemma 6.28 we have $y_i \geq \frac{3}{4} \left\lceil \frac{|\phi(C_i)| - 2d_i}{2} \right\rceil$. Summing up over all small cut classes we get:

$$\frac{\epsilon n}{8} \geq x(E_{S_I}) + x(E_{S_T}) \geq \sum_{i=1}^{s} y_i \geq \sum_{i=1}^{s} \frac{3}{4} \left\lceil \frac{|\phi(C_i)| - 2d_i}{2} \right\rceil \geq \frac{3}{8} \sum_{i=1}^{s} |\phi(C_i)| - \frac{3}{4} \sum_{i=1}^{s} d_i \geq \frac{3\epsilon n}{8} - \frac{3}{4} \sum_{i=1}^{s} d_i,$$

where the first inequality holds by lemmas assumption, and the last inequality holds by equation (36). Therefore, $\sum_{i=1}^{s} d_i \geq \frac{\epsilon n}{3}$. Thus $\Gamma(C)$ contains at least $\frac{\epsilon n}{6}$ edges, and at least this many nodes. In other words, we have at least $\frac{\epsilon n}{6}$ cut classes.

$\square$

Lemma D.3. *If $\Gamma(C)$ contains at least ϵn nodes, $\epsilon > \frac{820}{\tau}$ and $x(E_{S_I}) + x(E_{S_T}) \leq \frac{\epsilon n}{64}$, then $|\pi_{\frac{\epsilon n}{80}}| \geq \frac{\epsilon n}{40}$.*

Proof. We show that $\Gamma(C)$ must contain small number of leaves. This implies that it should contain many nodes of degree 2. Let L be the number of leaves of $\Gamma(C)$. By corollary 6.33, any small cut class which is a leaf is assigned good edges of fraction at least 3/4. Since $x(E_{S_I}) + x(E_{S_T}) \leq \frac{\epsilon n}{64}$, the number of leaves which are a small cut class is no more than $\frac{\epsilon n}{48}$. Also the number of large cut classes which are a leaf is no more than n/τ since $L(\tau) \leq n$. Hence:

$$L \leq \frac{\epsilon n}{48} + \frac{n}{\tau} \leq \frac{\epsilon n}{41}$$

Since in any tree, the number of nodes of degree more than 2 is at most the number of leaves, we get the number of small cut class of degree strictly more than 2 is at most L.

56

Let $\Gamma'(C)$ be the tree obtained by contracting each thread (or equivalently, each small cut class of degree 2 except the root). The number of nodes of $\Gamma'(C)$ is no more than $2L + n/\tau \le \frac{\epsilon n}{20}$. Since each edge of $\Gamma'(C)$ is corresponding to a thread in $\Gamma(C)$, and $\Gamma(C)$ contains at least ϵn vertices, it must have at least $\frac{\epsilon n}{20}$, disjoint threads of length 10. Note that if the father-connector atom of a cut class contains at least m vertices, the father-connector atom of its descendants contain at least m vertices too. Therefore, by Lemma F.2 at least half of these threads are unbalanced with parameter $\frac{\epsilon n}{80}$. □

This finishes the proof of Theorem 4.26.

E Proofs from Section 5

We prove a crucial lemma which shows that the minimum cut of any subset of the inside vertices of H is very small.

Proof of Lemma 5.15.   We assume that $|T| > 1$, otherwise the lemma is trivial. We prove by contradiction, suppose for any non-empty set $S \subset T$, we have $|E(S, T \setminus S)| > 6\delta c$. The idea is to find a k-cycle (say C), such that at least one (but no more than a constant number) of its cut sides crosses T. Then we may argue that at least one of those cut sides (say $(C_j, \overline{C_j})$) should have a large number of edges ($\Omega(\delta c)$) to vertices in T that are not included in any of the cut sides of C. We reach to a contradiction by showing that $(C_j, \overline{C_j})$ can not be a near minimum cut using the fact that a cut side should have a large number of edges ($(1 - \delta)\frac{c}{2}$) to its adjacent cut sides in C.

Since $(T, \overline{T})$ is a non-trivial cut, there exists at least one near minimum cut that crosses $(T, \overline{T})$. Among the set of near minimum cuts of H that crosses the set T, let $(D, \overline{D})$ be the cut that maximizes $|D \cap T|$. Note that since $(D, \overline{D})$ crosses T, we have $\overline{D} \cap T \ne \emptyset$. Let $v^* \in \overline{D} \cap T$ be an inside vertex. Since v^* is an inside vertex, there exists a k-cycle C for it. Let $\{C_1, C_2, \ldots, C_k\}$ be the cut sides of C. Recall that by definition none of the cut sides contain v^*. In the next claim we show that there is another cycle for v^* that contains a set with properties similar to D as one of the cut sides, or exactly two adjacent cut sides that cross D.

Claim E.1. *Let $(D, \overline{D})$ be a near minimum cut that crosses T and maximizes $|D \cap T|$, and let $v^* \in T \setminus D$. Then, there is a cycle C for the vertex v^* which either contains C_i such that $|C_i \cap T| = |D \cap T|$, or has at most two other cut sides which cross D.*

Proof. Let $C = \{C_1, \ldots, C_k\}$ be the shortest cycle for v^* as described above which maximizes $\max_i |C_i \cap D \cap T|$. Observe that if this maximum equals $|D \cap T|$, then we are done. Hence, assume that $C_i \cap D \cap T \subsetneq D \cap T$ for all $1 \le i \le k$. Let t denote the number of cut sides C_i whose intersection with D is non-empty. Since each near minimum cut corresponds to a diagonal in the polygon representation, and $D \nsubseteq C_i$ for any i, we must have $t > 1$. Moreover, the set of cut sides intersecting D must occur consecutively in the cycle C. Let C_x, C_y be the first, and the last cut sides that are crossing D, respectively.

If $t \ge 3$, then consider the cycle $C' = \{C_1, \ldots, C_x, D, C_y, C_{y+1}, \ldots, C_k\}$. It is easy to verify that it is a valid cycle for v^*. Moreover either it is shorter or improves $\max_i |C_i \cap D \cap T|$.

Else $t = 2$, then we cannot have $C_i \subset D$ for any i. Therefore, both cut sides intersecting with D (i.e., C_x, C_y) cross D. □

Let $O := \bigcap_j \overline{C_j}$ be the set of inside vertices that are not included in any of the cut-sides of C (note that we have $v^* \in O$ and thus $O \ne \emptyset$).

In the next claim we show that there exists a cut side C_j such that $|E(C_j, O)| > 2\delta c$.

Claim E.2. *Consider the cycle $C = \{C_1, \ldots, C_k\}$ be a k-cycle for v^* as given by Claim E.1. Then C contains a cut side C_j such that $|E(C_j, O)| > 2\delta c$.*

Proof. In the first case, let C_2 denote the cut side such that $|C_2 \cap T| = |D \cap T|$ and in the second case let C_1 and C_2 denote the sides which intersect D. We show that only the cut sides C_1, C_2, C_3 may cross T. Then since $C_2 \cap T \neq \emptyset$, and $v^* \notin C_1 \cup C_2 \cup C_3$, we have $C_1 \cup C_2 \cup C_3$ crosses T. Therefore, by lemma's assumption we have $|E(C_1 \cup C_2 \cup C_3, T \setminus (C_1 \cup C_2 \cup C_3))| > 6\delta c$. But since all of the vertices of T are in one of the four sets C_1, C_2, C_3 or O, we have $O = T \setminus (C_1 \cup C_2 \cup C_3)$, which implies that $|E(C_1 \cup C_2 \cup C_3, O)| > 6\delta c$. Therefore, we obtain that $|E(C_j, O)| > 2\delta c$ for some $1 \leq j \leq 3$, which completes the proof.

It remains to show that only the first three cut-sides may cross T. First consider the case that $|C_2 \cap T| = |D \cap T|$. Then for all near minimum cuts $(S, \overline{S})$ crossing T, we have $|\overline{C_2} \cap T| \leq |S \cap T| \leq |C_2 \cap T|$. Therefore, for all $3 < j$, we either have $C_j \cap T = \emptyset$ (i.e. $(C_j, \overline{C_j})$ does not cross T), or $|C_j \cap T| \geq |\overline{C_2} \cap T|$. If the former occurs we are fine. So suppose the latter occurs for some $j > 3$. Since C is a k-cycle, and $j > 3$, we have $C_j \cap C_2 = \emptyset$. Since $(C_j, \overline{C_j})$ crosses T we get

$$|C_2 \cap T| + |C_j \cap T| \geq |C_2 \cap T| + |\overline{C_2} \cap T| = |T|.$$

But we know that v^* is not in any of them, thus we must have $C_2 \cap T \cap C_j \neq \emptyset$, which is a contradiction. Therefore none of the cut sides C_j crosses T for $j > 3$.

In the other case, let the cut sides intersecting D be C_1 and C_2. Following the argument in the previous proof, we claim that $C_j \cap T = \emptyset$ for all $j \neq 1, 2$. Suppose that is not the case and there is a $j > 2$ such that $C_j \cap T \neq \emptyset$. Since we have $D \cap C_j = \emptyset$, and $|\overline{D} \cap T| \leq |C_j \cap T| \leq |D \cap T|$ as before there must be a common element in C_j and D which is a contradiction. Thus there exists a cut side C_j for $1 \leq j \leq 2$ such that $|E(C_j, O)| > \frac{6\delta c}{2} > 2\delta c$. $\qquad\square$

By Lemma 5.3, we have $|E(C_j, C_{j-1} \setminus C_j)| \geq (1 - \delta)\frac{c}{2}$ and $|E(C_j, C_{j+1} \setminus C_j)| \geq (1 - \delta)\frac{c}{2}$. Therefore, $|d(C_j)| > c(1 - \delta) + 2\delta c = c(1 + \delta)$ which is a contradiction. $\qquad\square$

F   On the Properties of Tree Hierarchies

In this section we explain some basic properties of tree hierarchies.

Lemma F.1. *Let C be the collection of $(1 + \delta)$ near minimum cuts of fractional graph G, with cut classes $C_1, \ldots, C_l$. In all except one of the cut classes there is an atom which is a subset of at least $n/2$ of the vertices of G.*

Proof. The proof follows from the fact that the union of two connecting atoms is V. Therefore, at least one of them must have at least $n/2$ vertices.

Let C_r be a cut class such that all of its atoms have less than $n/2$ vertices. Then any cut class C_i which is adjacent to C_r in $\Gamma(C)$ contains an atom with $n/2$ vertices (see Definition 4.21). In fact since C_i is adjacent to C_r there is an atom $b \in \phi(C_i)$ that coincide with an atom $a \in \phi(C_r)$ in a vertex of $K(C)$. Since all atoms of C_r have less than $n/2$ vertices, b must have at least $n/2$ vertices.

Similarly, it can be shown that all cut classes C_j that are adjacent to C_i (and in general all cut classes at distance 2 from C_r) must contain an atom with $n/2$ vertices. In particular, suppose an

atom $c \in \phi(C_j)$ coincides with an atom $b' \in \phi(C_i)$. Since C_j is at distance 2 from C_r we must have $b' \neq b$. Moreover, since $|b| \geq n/2$, the rest of the atoms of C_i has at most $n/2$ vertices. Therefore $|b'| \leq n/2$, thus $|c| \geq n/2$. By using a simple induction on the set of cut classes at distance k from C_r it can be seen that all of the cut classes other than C_r contain an atom with at least $n/2$ vertices. □

Lemma F.2. *Let C be the collection of $(1 + \delta)$ near minimum cuts of fractional graph G. Let $\Gamma(C)$ be the rooted tree defined in 4.21. If $l < \frac{n}{6}$, then for any set of cut classes $C_1, \ldots, C_{4l}$, the father-connector atom of at least $2l$ of them has at least $\frac{n}{2} + l$ vertices.*

Proof. We prove that for any $l < \frac{n}{6}$ in all except at most $2l$ cut classes, the father-connector contains at least $\frac{n}{2} + l$ vertices. Let $f(C_i)$ be the father-connector of the cut class C_i. Let $S := \{C_i : |f(C_i)| < \frac{n}{2} + l\}$. We first show that the cut classes in S form an (undirected) path in $\Gamma(C)$. Since $C_r \in S$, the undirected path is made of two directed path that starts from the root. Therefore there is a directed path that starts at root of length at least $|S|/2$. In Claim F.5 we show that the father-connector of the cut classes at distance k from the root has at least $\frac{n}{2} + k$ vertices. Since father-connectors of cut classes in S have less than $\frac{n}{2} + l$ vertices, we must have $|S| \leq 2l$.

Since the father-connectors of the descendants of a cut class C_i in $\Gamma(C)$ can only have more vertices than $f(C_i)$, the set of cut classes in S forms a connected subgraph of $\Gamma(C)$. In other words, for any $C_i \in S$, all of the ancestors of C_i must be in S too. Thus we need to show that this subgraph is a path.

In the next claim we show that it is not possible that 3 atoms of 3 different cut class in S coincide in a vertex of $K(C)$. Using this in Claim F.4 we show that each cut class in S is adjacent to at most 2 other cut classes of S.

Claim F.3. *Let C_1, C_2, C_3 be 3 cut classes such that there is $a \in \phi(C_1), b \in \phi(C_2), c \in \phi(C_3)$ such that a, b, c coincides in a vertex of $K(C)$. If $l < \frac{n}{6}$, then at least one of the three cut classes is not in S.*

Proof. The proof follows from the fact that the union of atoms that coincide in a vertex $K(C)$ is V. Hence we have $a \cup b = a \cup c = b \cup c = V$. Wlog suppose $|a| \geq |b|$. Since $|a \cup b| = n$ we have

$$|a| \geq |a \cap b| + \frac{n - |a \cap b|}{2} = \frac{n + |a \cap b|}{2}.$$

Therefore, if $|a \cap b| \geq 2l$ we get $|a| \geq \frac{n}{2} + l$ and we get $C_1 \notin S$. Thus suppose $|a \cap b| < 2l$. Since $|a \cup b| = n$, we have

$$|a \setminus b| + |b \setminus a| = n - |a \cap b| \geq n - 2l > \frac{n}{2} + l,$$

where the last inequality follows from the fact that $l < \frac{n}{6}$. On the other hand, since $a \cup c = V$, and $a \cup b = V$, we have $b \setminus a \subseteq c$. Similarly, $a \setminus b \subseteq c$. Hence,

$$|c| \geq |a \setminus b| + |b \setminus a| > \frac{n}{2} + l.$$

Therefore $c = f(C_3)$ and $C_3 \notin S$. □

In the next claim we show that the cut classes in S must form a path in $\Gamma(C)$:

Claim F.4. *If $l < \frac{n}{6}$, then any cut class $C_1 \in S$ can be connected to at most 2 other cut classes of S in the tree $\Gamma(C)$.*

Proof. Suppose $C_1 \in S$ is adjacent to 3 cut classes C_2, C_3, C_4 in $\Gamma(C)$. We show that the father connector of at least one of them must have more than $\frac{n}{2} + l$ vertices. Let $a_2, a_3, a_4 \in \phi(C_1)$ be the atoms that coincide with $b_2 \in \phi(C_2), b_3 \in \phi(C_3), b_4 \in \phi(C_4)$. By Claim F.3 we must have a_2, a_3, a_4 are three different atoms in $\phi(C_i)$. Since $a_2 \cup a_3 \cup a_4 \subseteq V$, we have $|a_2 \cup a_3 \cup a_4| \leq n$. Since the three atoms are disjoint wlog we have $|a_2| \leq \frac{n}{3} < \frac{n}{2} - l$. Since b_2 coincide with a_2 we have $a_2 \cup b_2 = V$, thus $|b_2| \geq \frac{n}{2} + l$. This implies that $b_2 = f(C_2)$ and $C_2 \notin S$. $\square$

Therefore, S forms a path in $\Gamma(C)$. Hence there is a directed path of the cut classes in S of length $\frac{|S|}{2}$ that starts at C_r. The following claim implies that the father-connector of a cut class at distance k from the root must have at least $\frac{n}{2} + k$ vertices. This implies that $|S| \leq 2l$:

Claim F.5. *Let C_2 be a child of C_1, and C_3 be a child of C_2 which is not adjacent to C_1. Then we either have $|f(C_2)| \geq |f(C_1)| + 1$, or $|f(C_3)| \geq |f(C_1)| + 2$.*

Proof. The proof follows from a simple case analysis. Let $a \in \phi(C_1)$ be the atom coincides with $f(C_2)$. First suppose $|\phi(C_1)| > 2$; since we do not have any cut class with 3 atoms, we must have $|\phi(C_1)| \geq 4$. Since the atoms form a partitioning of the vertices, we have $|a| \leq |f(C_1)| - 2$. Finally, $a \cup f(C_2) = V$, thus we must have $|f(C_2)| \geq |f(C_1)| + 2$.

Now suppose $|\phi(C_1)| = 2$. If $|\phi(C_2)| = 2$, then since $C_1 \neq C_2$ we must have $|f(C_2)| \geq |f(C_1)| + 1$ and we are done. It remains the case where $|\phi(C_1)| = 2$ and $|\phi(C_2)| > 2$ (note that in this case we can have $|f(C_1)| = |f(C_2)|$). By the argument in the previous paragraph in this case we have $|f(C_3)| \geq |f(C_2)| + 2$. Therefore, $|f(C_1)| \geq |f(C_3)| + 2$ and we are done. $\square$

$\square$

G   Proofs from Section 6.2

Proof of Lemma 6.21.   The proof strategy is similar to Lemma 6.19. By equation (16), we have:

$$\mathbb{P}_{T \sim \mu}[X = 1, Y = 1] = \mathbb{P}_{T \sim \mu}[X = 1, Y = 1 | X + Y = 2]\mathbb{P}_{T \sim \mu}[X + Y = 2] \geq \epsilon\mathbb{P}_{T \sim \mu}[X = 1, Y = 1 | X + Y = 2]$$

Thus we need to show that $\mathbb{P}_{T \sim \mu}[X = 1, Y = 1 | X + Y = 2] \geq \alpha/4 \min(1/2, \beta e^{-\beta})$. Let $\gamma := \alpha/4 \min(1/2, \beta e^{-\beta})$. By ULC theorem we have:

$$\mathbb{P}_{T \sim \mu}[X = 1, Y = 1 | X + Y = 2]^2 \geq \mathbb{P}_{T \sim \mu}[X = 2, Y = 0 | X + Y = 2]\,\mathbb{P}_{T \sim \mu}[X = 0, Y = 2 | X + Y = 2]$$

Note that in general both of the terms in the RHS could be zero (e.g. when we have $X = 1, Y = 1$ with probability 1). Thus we prove $\mathbb{P}_{T \sim \mu}[X \geq 1 | X + Y = 2] \geq 2\gamma$ (and similarly $\mathbb{P}_{T \sim \mu}[Y \geq 1 | X + Y = 2] \geq 2\gamma$). This is equivalent to $\mathbb{P}_{T \sim \mu}[Y \leq 1 | X + Y = 2] \geq 2\gamma$. Since $X \geq 1$ and $Y \leq 1$ are an upward (resp. downward) event, by Theorem 6.17 we have

$$\mathbb{P}_{T \sim \mu}[X \geq 1 | X + Y = 2] \quad \geq \quad \mathbb{P}_{T \sim \mu}[X \geq 1 | X + Y = 1] \tag{37}$$
$$\mathbb{P}_{T \sim \mu}[Y \leq 1 | X + Y = 2] \quad \geq \quad \mathbb{P}_{T \sim \mu}[Y \leq 1 | X + Y = 3] \geq \mathbb{P}_{T \sim \mu}[Y \leq 1 | X + Y = 4] \geq \ldots$$

Note that inequality (37) is valid once the event $X + Y = 1$ is well defined. For this moment suppose this is the case, as we will see throughout the proof the other case is much simpler. The second inequality implies the following:

$$\mathbb{P}_{T \sim \mu}[Y \leq 1 | X + Y = 2] \geq \mathbb{P}_{T \sim \mu}[Y \leq 1 | X + Y \geq 2] \tag{38}$$

By equations (37), (38) to prove the lemma it is sufficient to show that

$$\mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]+\mathbb{P}_{T\sim\mu}\left[X\geq 1|X+Y=1\right]\geq 4\gamma. \tag{39}$$

Let us consider the event $Y\leq 1$, we have

$$
\begin{aligned}
\mathbb{P}_{T\sim\mu}\left[Y\leq 1\right] &= \mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]\mathbb{P}_{T\sim\mu}\left[X+Y\geq 2\right]+\mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\leq 1\right]\mathbb{P}_{T\sim\mu}\left[X+Y\leq 1\right]\\
&\leq \mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]+\mathbb{P}_{T\sim\mu}\left[X+Y\leq 1\right] \tag{40}\\
&= \mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]+\mathbb{P}_{T\sim\mu}\left[X=1,Y=0\right]+\mathbb{P}_{T\sim\mu}\left[X=0,Y\leq 1\right]\\
&\leq \mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]+\mathbb{P}_{T\sim\mu}\left[X\geq 1|X+Y=1\right]+\mathbb{P}_{T\sim\mu}\left[X=0,Y\leq 1\right]. \tag{41}
\end{aligned}
$$

Therefore, if the event $X+Y\leq 1$ is not defined (i.e. $\mathbb{P}_{T\sim\mu}\left[X+Y\leq 1\right]=0$), by equation (40) we obtain $\alpha\leq\mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]$ and we are done.

By Fact 6.9, and using equation (18) we have

$$\mathbb{E}_{T\sim\mu}\left[X|Y\leq 1\right]\geq\mathbb{E}_{T\sim\mu}\left[X\right]\geq\beta.$$

Since by Theorem 6.17, the measure $\{\mu:Y\leq 1\}$ is strongly Rayeligh, we can apply Theorem 6.13 and Lemma 6.14 to upper bound $\mathbb{P}_{T\sim\mu}\left[X=0|Y\leq 1\right]$.

If $\beta\geq 1$, then by theorem 6.13 the mode of the distribution $\{\mu:Y\leq 1\}$ is at least 1, thus $\mathbb{P}_{T\sim\mu}\left[X=0|Y\leq 1\right]\leq\frac{1}{2}$. On the other hand, if $\beta\leq 1$, by Lemma 6.14

$$\mathbb{P}_{T\sim\mu}\left[X=1|Y\leq 1\right]\geq B(\beta,1)=\beta e^{-\beta},$$

and $\mathbb{P}_{X=0|Y\leq 1}\left[\leq\right]1-\beta e^{-\beta}$. Putting them together we obtain

$$\mathbb{P}_{T\sim\mu}\left[X=0,Y\leq 1\right]=\mathbb{P}_{T\sim\mu}\left[X=0|Y\leq 1\right]\mathbb{P}_{T\sim\mu}\left[Y\leq 1\right]\leq\left(1-\min(\frac{1}{2},\beta e^{-\beta})\right)\mathbb{P}_{T\sim\mu}\left[Y\leq 1\right]$$

Putting this with equation (41) we obtain:

$$\min(\frac{1}{2},\beta e^{-\beta})\mathbb{P}_{T\sim\mu}\left[Y\leq 1\right]\leq\mathbb{P}_{T\sim\mu}\left[Y\leq 1|X+Y\geq 2\right]+\mathbb{P}_{T\sim\mu}\left[X\geq 1|X+Y=1\right].$$

Finally using equation (17) we obtain equation (39). $\qquad\square$

Proof of Corollary 6.22. Let $\mu_1=\{\mu|Z=1\}$ be the measure obtained from μ conditioned on $Z=1$. We show that μ_1 satisfies all of the conditions of Lemma 6.21. Since $\mathbb{P}_{\mu_1}\left[X+Y=2\right]\geq\epsilon$, the first condition is satisfied. Since $\mathbb{E}_{T\sim\mu}\left[X\right]=\mathbb{E}_{T\sim\mu}\left[Y\right]=1-x(u,v)$, By the negative association between the edges we have:

$$
\begin{aligned}
1 &\leq \mathbb{E}_{\mu}\left[X\right]\leq 1.5+\frac{1}{5}\\
1 &\leq \mathbb{E}_{\mu}\left[Y\right]\leq 1.5+\frac{1}{5}
\end{aligned}
$$

Now it is straightforward to see that $\alpha\geq\frac{3}{20}$, and $\beta\geq 1$, this implies corollary. $\qquad\square$

G.1 Proofs from Section 6.4

Lemma G.1. *Let $u, v \in V$ be two vertices of $G = (V, E, x)$ and suppose that there exists a set $S_1 \subset V \setminus \{u, v\}$, and $S_2 := V \setminus \{u, v\} \setminus S_1$ such that both $(S_1, \overline{S_1})$, and $(S_2, \overline{S_2})$ are $(1+\delta')$-near minimum cuts, for $\delta' < 1/1000$. If (u, v) is a trivial edge and $|x(u, v) - \frac{1}{2}| \leq \frac{1}{200}$, then $\mathbb{P}_{T \sim \mu}[\mathcal{E}(u, v)] \geq 1/2000$.*

Proof. Let $D := E(\{u, v\}, \overline{\{u, v\}})$, be the set of edges separating u and v from the rest of the graph. By Corollary 6.22, we have $\mathbb{P}_{T \sim \mu}[\mathcal{E}(u, v)]$ is a constant, unless the size of the cut $|T \cap D| = 3$ with high probability. Here we show that this can not be the case, and indeed there is a constant probability that $|T \cap D| = 2$.

Observe that by lemma's assumption we should have

$$x(D \cap E(\{u, v\}, S_1)) = 1.5 \pm \frac{1}{100},$$

$$x(D \cap E(\{u, v\}, S_2)) = 1.5 \pm \frac{1}{100}.$$

Let $X := |T \cap D \cap E(\{u, v\}, S_1)|, Y := |T \cap D \cap E(\{u, v\}, S_2)|$, and $Z := |T \cap (u, v)|$.

We show that with a constant probability T contains the edge (u, v) and a spanning tree inside S_1 (or S_2). This implies that $\mathbb{P}_{T \sim \mu}[X + Y = 2|Z = 1] \geq 1/20$ which finishes the proof using Corollary 6.22.

Let $\mu_1 = \{\mu|Z = 1\}$ be the measure obtained from μ conditioned on $Z = 1$. By negative association this can only decrease the probability of other edges. Let $F := E(S_1) \cup E(S_2)$. From Fact 6.9, we have:

$$
\begin{aligned}
\mathbb{E}_{T \sim \mu_1}[|T \cap F|] &\geq z(F) - 0.5 - \frac{1}{200} \\
&\geq (1 - \frac{1}{n})(|S_1 \cup S_2| - 2 - 2\delta') - 0.5 - \frac{1}{200} \\
&\geq |S_1| + |S_2| - 3.5 - 2\delta' - \frac{1}{200}
\end{aligned}
$$

Dropping the $\frac{1}{200}$ term and δ' terms for simplicity, we either have $\mathbb{E}_{T \sim \mu_1}[|T \cap E(S_1)|] \geq |S_1| - 1.75$, or $\mathbb{E}_{T \sim \mu_1}[|T \cap E(S_2)|] \geq |S_2| - 1.75$. Wlog suppose the former happens. By Proposition 6.30 we have $\mathbb{P}_{T \sim \mu_1}[|T \cap S_1| = |S_1| - 1] \geq 0.25$.

Let $\mu_2 = \{\mu_1 \mid |T \cap S_1| = |S_1| - 1\}$. Again by Fact 6.9 this can only decrease the probability of the edges in D. Hence, $\mathbb{E}_{T \sim \mu_2}[Y] \leq 1.5$. Since the sets S_1 and $\{u, v\}$ act as a single vertex in the trees sampled from μ_2, any such tree can have at most one edge from $D \cap E(\{u, v\}, S_1)$, thus $\mathbb{E}_{T \sim \mu_2}[X] \leq 1$. Therefore,

$$1.5 \leq \mathbb{E}_{T \sim \mu_2}[X + Y] \leq 2.5 \Rightarrow \mathbb{P}_{T \sim \mu_2}[X + Y = 2] \geq \frac{1}{5},$$

where the RHS follows from Lemma 6.14. Hence,

$$\mathbb{P}_{\mu}[X + Y = 2|Z = 1] = \mathbb{P}_{T \sim \mu_1}\left[X + Y = 2 \mid |T \cap S_1| = |S_1| - 1\right] \mathbb{P}_{T \sim \mu_1}[|T \cap S_1| = |S_1| - 1] \geq \frac{1}{20}.$$

Hence, by applying Corollary 6.22, we get $\mathbb{P}[\mathcal{E}(u, v)] \geq \frac{1}{2000}$. $\qquad \square$

Lemma G.2. *Let u, v be two vertices in G and let $d(u) = d_1(u) \cup d_2(u)$ be a partitioning of edges adjacent to u (different parts of an edge can be divided too). Also let $X := |T \cap d_1(u)|, Y := |T \cap d_2(u)|, Z := |T \cap d_u(v)|$. Also let W_e be the indicator for the edge $e = (u, v)$. If the following conditions are satisfied for $\frac{1}{10} > \epsilon > 0$, then $\mathbb{P}_{T \sim \mu}[Z = 1, W_e = 1, X = 1, Y = 1] = \frac{\epsilon^3}{1000}$:*

$$\frac{1}{2} + \epsilon \leq \mathbb{E}_{T \sim \mu}[Z], \mathbb{E}_{T \sim \mu}[W_e] \leq 1 + \frac{\epsilon}{2} \tag{42}$$

$$\mathbb{E}_{T \sim \mu}[X + Y - W_e] \leq 1 + \frac{\epsilon}{2} \tag{43}$$

$$\mathbb{E}_{T \sim \mu}[X], \mathbb{E}_{T \sim \mu}[Y] \leq 1 + \frac{\epsilon}{2} \tag{44}$$

$$\frac{3}{2} + \epsilon \leq \mathbb{E}_{T \sim \mu}[X + Y] \tag{45}$$

$$\frac{5}{2} + \epsilon \leq \mathbb{E}_{T \sim \mu}[X + Y + Z] \tag{46}$$

Proof. Let us define $X_1 := |T \cap d_1(u) \cap \{e\}|$, $Y_1 := |T \cap d_2(u) \cap \{e\}|$, and $X_2 = X - X_1, Y_2 = Y - Y_1$. Wlog suppose

$$\mathbb{E}_{T \sim \mu}[Y_2] \geq \frac{1}{2} \mathbb{E}_{T \sim \mu}[X_2 + Y_2]. \tag{47}$$

We prove by conditioning on $X_1 = 1$, and then on $X_2 = 0$. First note that

$$\mathbb{E}_{T \sim \mu}[X_1] + 1 + \frac{\epsilon}{2} \geq \mathbb{E}_{T \sim \mu}[X_1] + \mathbb{E}_{T \sim \mu}[Y_1 + Y_2] \geq \frac{1}{2}(\mathbb{E}_{T \sim \mu}[X + Y] + \mathbb{E}_{T \sim \mu}[X_1 + Y_1]) \geq 1 + \epsilon,$$

where the first inequality holds by condition (44), and the last one holds by condition (42) and (45). Let $\mu_1 := \{\mu | X_1 = 1\}$ be the measure obtained from μ conditioned on $X_1 = 1$. Since we always have $E = X_1 + Y_1 \leq 1$, we get $\mathbb{E}_{T \sim \mu_1}[Y_1] = 0$. Therefore,

$$\mathbb{P}_{T \sim \mu}[Z = 1, E = 1, X = 1, Y = 1] \geq \mathbb{P}_{T \sim \mu_1}[X_2 = 0, Y_2 = 1, Z = 1] \mathbb{P}_{T \sim \mu}[X_1 = 1]$$

$$\geq \frac{\epsilon}{2} \mathbb{P}_{T \sim \mu_1}[X_2 = 0, Y_2 = 1, Z = 1] \tag{48}$$

It is sufficient to show that $\mathbb{P}_{T \sim \mu_1}[X_2 = 0, Y_2 = 1, Z = 1]$ is a constant. By Fact 6.9, we have

$$\mathbb{E}_{T \sim \mu_1}[X_2] \leq \mathbb{E}_{T \sim \mu}[X_2] \leq \frac{1}{2} \mathbb{E}_{T \sim \mu}[X_2 + Y_2] \leq \frac{1}{2} + \frac{\epsilon}{4}.$$

Let $\mu_2 := \{\mu_1 | X_2 = 0\}$. We have

$$\mathbb{P}_{T \sim \mu_1}[X_2 = 0, Y_2 = 1, Z = 1] = \mathbb{P}_{T \sim \mu_2}[Y_2 = 1, Z = 1] \mathbb{P}_{T \sim \mu_1}[X_2 = 0] \geq \frac{1}{3} \mathbb{P}_{T \sim \mu_2}[Y_2 = 1, Z = 1]. \tag{49}$$

Hence, we only need to show that $\mathbb{P}_{T \sim \mu_2}[Y_2 = 1, Z = 1]$ is a constant. We use Lemma 6.21. First note that conditioning on $X_1 = 1$ may decrease $\mathbb{E}_{T \sim \mu}[X_2 + Y_2 + Z]$ by at most $1 - \mathbb{E}_{T \sim \mu}[X_1 + Y_1]$ using the fact that conditioned on $X_1 + Y_1 = 1$, the random variables X_1 and Z are independent. Hence using condition (42) we get

$$\mathbb{E}_{T \sim \mu_2}[Z] \geq \mathbb{E}_{T \sim \mu_1}[Z] \geq \mathbb{E}_{T \sim \mu}[Z] - (1 - \mathbb{E}_{T \sim \mu}[X_1 + Y_1]) \geq 2\epsilon.$$

$$\mathbb{E}_{T \sim \mu_2}[Y_2] \geq \mathbb{E}_{T \sim \mu}[Y_2] - (1 - \mathbb{E}_{T \sim \mu}[X_1 + Y_1]) \geq \frac{1}{2}(\mathbb{E}_{T \sim \mu}[X + Y] + \mathbb{E}_{T \sim \mu}[X_1 + Y_1]) - 1 \geq \epsilon.$$

Also,
$$\mathbb{E}_{T \sim \mu_2}[Y_2 + Z] \geq \mathbb{E}_{T \sim \mu}[X + Y + Z] - \frac{1}{2}\mathbb{E}_{T \sim \mu}[X_2 + Y_2] - 1 \geq 1 + \frac{\epsilon}{2},$$

where the last inequality holds by conditions (43) and (46). On the other hand, by negative association conditioning on $X_2 = 0$ may increase $\mathbb{E}_{T \sim \mu}[Y_2 + Z]$ by at most $\frac{1}{2} + \frac{\epsilon}{4}$. Therefore, using equations (44) and (42) we get

$$\mathbb{E}_{T \sim \mu_2}[Z], \mathbb{E}_{T \sim \mu_2}[Y_2] \leq 1.5 + \epsilon$$
$$\mathbb{E}_{T \sim \mu_2}[Y_2 + Z] \leq 2.5 + \epsilon.$$

Hence, by Lemma 6.14, we have

$$\mathbb{P}[Y_2 + Z = 2] \geq B(\mathbb{E}_{\mu_2}[Y_2 + Z], 2) \geq B(1 + \frac{\epsilon}{2}, 2) \geq \frac{\epsilon}{4}.$$

By Markov's inequality, $\mathbb{P}_{\mu_2}[Y_2 \leq 1], \mathbb{P}_{\mu_2}[Z \leq 1] = \frac{1}{5}$. Therefore, from Lemma 6.21 setting $\epsilon := \frac{\epsilon}{4}, \alpha = \frac{1}{5}, \beta = \epsilon$ we obtain that $\mathbb{P}_{T \sim \mu_2}[Y_2 = 1, Z = 1] = \frac{\epsilon^2}{160}$. Putting this together with equations (48) and (49) we conclude that: $\mathbb{P}_{\mu}[Z = 1, E = 1, X = 1, Y = 1] \geq \frac{\epsilon^3}{1000}$. $\square$

H Constants

In this section, we summarize the discussion about constants.

1. We will fix δ at the end to be very small.

2. We set $\tau = \frac{1}{20\sqrt{\delta}}$ from Proposition 4.19.

3. To set ϵ in $L(\tau)$ to be more or less than $(1 - \epsilon)n$, Lemma 4.4 needs $\epsilon > \frac{1}{\tau - 2}$ and Theorem 4.26 needs $\epsilon > \frac{4920}{\tau}$. We set $\epsilon = \frac{5000}{\tau} = 10^5\sqrt{\delta}$.

4. Now we obtain ϵ_2 in the algorithm to be at least $20\sqrt{\delta} + 17\epsilon = (17 \cdot 10^5 + 20)\sqrt{\delta}$ from Lemma 4.4. We set it to be $2 \cdot 10^6\sqrt{\delta}$.

5. We set $\rho := \epsilon^3 10^{-16}$ from Proposition 4.25 which gives the smallest probability for an edge to be even.

6. ϵ_1 in Theorem 1.1 is set by Theorem 4.26. We have $\epsilon_1 = \min\{\frac{\epsilon}{3840}, \frac{\epsilon^2}{6 \cdot 10^7}\} = \frac{\epsilon^2}{6 \cdot 10^7} \simeq 3000\delta$.

7. Now we compute δ. We also need $2\epsilon_2 \leq 0.1$ for Lemma B.3. This implies $\delta \leq 6.25 \cdot 10^{-16}$. We set $\delta = 6.25 \cdot 10^{-16}$.

8. Finally, ϵ_0 the improvement of the algorithm can be computed by Lemma B.2 and set to $\epsilon_0 := \frac{\epsilon_1 \delta \rho}{4(1 + \delta)}$.

Hence, finally we have

1. $\delta = 6.25 \cdot 10^{-16}$.

2. $\tau = 2 \cdot 10^6$.

3. $\epsilon = 2.5 \cdot 10^{-3}$.

4. $\rho = 1.5 \cdot 10^{-24}$.

5. $\epsilon_2 = 0.05$.

6. $\epsilon_1 = 18.75 \cdot 10^{-13}$.

7. $\epsilon_0 = 4 \cdot 10^{-52}$.