

Finding Structure in Entropy: Improved Approximation Algorithms for TSP and other Graph Problems

Nathan Klein

A dissertation

submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Anna Karlin, Chair

Shayan Oveis Gharan, Chair

Cynthia Vinzant

©Copyright 2023
Nathan Klein

University of Washington

Abstract

Finding Structure in Entropy: Improved Approximation Algorithms for TSP and other Graph Problems

Nathan Klein

Co-Chairs of the Supervisory Committee:

Anna Karlin and Shayan Oveis Gharan

Department of Computer Science

This dissertation demonstrates that there is an approximation algorithm for the metric traveling salesperson problem (TSP) with approximation ratio below $3/2$. This represents the first improvement in nearly half a century, answering a long-standing open problem in combinatorial optimization.

The algorithm we analyze, a variant of Christofides' $3/2$ approximation from the 1970s, exploits a distribution over spanning trees that has as much entropy as possible subject to obeying certain marginal constraints. A key component of this work is to show that despite the inherent unpredictability of such a distribution, the trees it produces nevertheless exhibit surprisingly robust structural properties. To show these properties, we use that the generating polynomials of these distributions have a zero-free region in the complex plane, allowing us to employ a suite of tools coming from work on the geometry of polynomials. As a byproduct of our analysis, we prove several new statements that sharply characterize the behavior of such distributions.

We also discuss several other results in network design, including a lower bound for this algorithm, an optimal rounding algorithm for a special case of TSP, and improved algorithms for the k -edge-connected multi-subgraph problem and the laminar thin tree problem.

Acknowledgements

Thanks to my amazing advisors Anna Karlin and Shayan Oveis Gharan, I am one of the lucky few for which graduate school was mostly just *fun*. You showed me beautiful mathematics, taught me how to give talks and do research, and pushed me to work hard, but all with an exceptionally light touch which never made being your student feel like a job. I don't know how you did it – I often feel like you tricked me into getting a PhD. I am immensely grateful to both of you.

I have been fortunate to have many teachers and mentors over my life who have gone above and beyond to guide and nurture me: Lynn Legler, Alan Tung, Luca Matone, Sam Reiner, Sibylle Johner, Amir Eldan, Gregory Fulkerson, Nicholas Kraft, Kami Vaniea, Jonathan Katz, Justin Kolb, Kira Goldner, Tom Wexler, and Alexa Sharp. I also thank Robert Bosch, Lola Thompson, Elizabeth Wilmer, Chris Marx, Cynthia Vinzant, Thomas Rothvoss, and James Lee for being exceptional teachers: you all taught me how to think and how to do mathematics.

And thank you to my collaborators, from which I have learned that the greatest part of research is time spent working with others: Fateme Abbasi, Dorna Abdolazimi, Jannis Blauth, Jarek Byrka, Vincent Cohen-Addad, Cody Freitag, Ellis Hershkowitz, Billy Jin, Kevin Kim, Janardhan Kulkarni, Jonathan Leake, Kasper Lindberg, Martin Nägele, Neil Olver, Sherry Sarkar, Jakub Tarnawski, David Williamson, Rico Zenklusen, and Xinzhi Zhang. I am also very grateful to the UW theory group for creating a fun and welcoming community, and for my friends in Seattle and elsewhere for supporting me and giving me a life beyond math.

I thank NSF for generously supporting me through the Graduate Research Fellowship Program, and the University of Washington for its support through teaching. I am also grateful to the Hausdorff Institute, the Institute of Advanced Study, and Microsoft Research for their support.

Finally, thank you to my parents, Bill and Jane, my brother Alex, and my partner Flora. I could not have done this without you, or really much of anything at all.

Contents

1	Introduction	8
1.1	Approximation Algorithms	9
1.2	High Level Overview of Techniques	10
1.3	Other Results on TSP	14
1.3.1	Lower Bounds for the Max Entropy Algorithm	14
1.3.2	Optimal Results for a Special Case of TSP	15
1.4	Summary of Results on TSP and Corollaries	16
1.5	Other Results on Network Design Problems	17
1.5.1	Designing k -Edge-Connected Graphs	17
1.5.2	The Thin Tree Conjecture and the Laminar Crossing Spanning Tree Problem	18
2	Essential Background	19
2.1	Polyhedra of Interest	19
2.1.1	Notation	19
2.1.2	The Subtour LP	19
2.1.3	The Spanning Tree Polytope	20
2.2	The O -Join Polyhedron and the Integrality Gap of P_{Sub}	21
2.3	The Maximum Entropy Algorithm for TSP	22
2.3.1	λ -uniform Spanning Trees	22
2.3.2	The Algorithm	23
2.4	Generating Polynomials and λ -uniform Spanning Trees	25
2.5	Real Stable Polynomials	26
2.6	Strongly Rayleigh Distributions	28
2.7	Additional Properties of SR Distributions	32
2.8	Staying in the λ -Uniform Distribution	33
2.9	The Structure of (Near) Minimum Cuts	36
2.9.1	Notation	37
2.9.2	The Cactus Representation	38
2.9.3	Near Minimum Cuts	38
2.9.4	The Polygon Representation	39
3	Building Up: Degree Cuts and The Half Integral Case	42
3.1	The Degree Cut Case	42
3.1.1	The Construction of s in the Half Integral Degree Cut Case	43
3.1.2	Proof of Lemma 3.1	45
3.1.3	The Construction of s in the General Degree Cut Case	46
3.2	The Half Integral Case	47
3.2.1	Critical Sets	48
3.3	Overview of Analysis	52
3.4	Probabilistic lemmas	54
3.5	Proof of Main Theorem	56
3.6	Summary of special cases	60

4	Overview	61
4.1	Three New Techniques	61
4.1.1	Polygon Structure for Near Minimum Cuts Crossed on one Side.	61
4.1.2	Generalized Gurvits' Lemma	62
4.1.3	Conditioning while Preserving Marginals	63
4.2	Overview of Proof	63
4.2.1	Ideas underlying proof of Theorem 4.6	65
4.2.2	Proof ideas for Theorem 4.7	68
5	Polygons and the Hierarchy of Near Minimum Cuts	76
5.1	Notation	76
5.2	Cuts Crossed on Both Sides	76
5.3	Proof of the Main Technical Theorem, Theorem 12.1	79
5.4	Structure of Polygons of Cuts Crossed on One Side	80
5.5	Happy Polygons	85
5.6	Hierarchy of Cuts and Proof of Theorem 5.6	89
5.7	Hierarchy Notation	92
6	Probabilistic statements	94
6.1	Properties of Bernoulli-Sum Random Variables	94
6.2	Random Spanning Trees	97
6.3	Gurvits' Machinery and Generalizations	99
6.4	Max Flow	102
6.5	Good Edges	106
6.6	2-1-1 and 2-2-2 Good Edges	111
7	Matching	115
8	Reduction and payment	119
8.1	Increase for Good Top Edges	122
8.2	Increase for Bottom Edges	128
8.2.1	Case 1: \hat{S} is a degree cut	129
8.2.2	Case 2: S and its parent \hat{S} are both polygon cuts	131
9	Introduction to the Integrality Gap Result and a Proof Overview	134
9.1	New techniques and contributions	134
9.2	Proof Overview	136
9.3	Overview of proof of Theorem 9.7 – no inside atoms	139
9.4	Satisfying the above desiderata	141
9.4.1	Defining bad events for each polygon point	143
9.4.2	Defining the slack increase sets for bad events	143
9.5	Extending to polygons with inside atoms	145

10 Polygon Representation: Redux	148
10.1 Our polygon notation and the root	149
10.2 Properties of inside atoms	151
10.3 New properties of polygon representations	151
10.3.1 Almost diagonal cuts and the chain lemma	152
10.4 Another structural property of inside atoms	154
11 Using the Polygon Representation for Cuts Crossed on Both Sides	156
11.1 Notation and a preliminary lemma	156
11.2 Main theorem	157
11.3 All cuts are satisfied	159
11.4 Every cut is mapped to a constant number of bad events	160
12 Putting everything together for the integrality gap	162
12.1 Constructing the slack vector s	163
12.2 Proof of Theorem 1.2 using Theorem 12.1	164
13 Derandomizing the Max Entropy Algorithm	167
13.1 High level proof overview	167
13.2 Computing probabilities	169
13.2.1 Notation	169
13.2.2 Matrix tree theorem	169
13.3 Computing parities in a simple case	170
13.4 A deterministic algorithm in the degree cut case	170
13.5 A deterministic algorithm in the general case	172
14 Computation for $c(s^*)$	174
14.1 Polygon representation preprocessing	175
14.1.1 Computation for cuts crossed on both sides	175
14.2 Preprocessing for cuts crossed on one side	177
14.3 Computation of $\mathbb{E}[c(s^*) \mid \text{Set}]$	179
15 Computation for $c(s)$	179
15.1 Hierarchy definition and computation	179
15.2 Edge bundles, A, B, C degree partition, and edge classification	181
15.3 Max Flow	184
15.4 Matching	185
15.5 Reductions	186
15.6 Increases	187
15.6.1 Increases for bottom edges	187
15.6.2 Increases for top edges	189
15.7 Computation of $\mathbb{E}[c(s) \mid \text{Set}]$	189
16 A Lower Bound for Max Entropy: k-Donuts	191
16.1 The Max Entropy Algorithm on the k -Donut	191
16.2 Analyzing the Performance of Max Entropy on the k -Donut	193

17	4/3 for Cycle Cut Instances	196
17.1	Technical Overview	198
17.2	Preliminaries	199
17.2.1	The structure of minimum cuts	199
17.3	Cycle cut instances and the hierarchy of critical tight sets	200
17.4	Structure of cycle cuts	202
17.5	Proof of Theorem 17.1	205
17.6	The Markov Chains	207
17.7	Fixed Point and Algorithm	213
17.8	Characterizing the Feasible Region	214
17.9	Conclusion and Open Questions	217
18	An Improved Approximation Algorithm for k-ECSM	218
18.1	Proof Overview	219
18.2	Preliminaries	221
18.2.1	Background	221
18.3	Warm-up: a Simple Algorithm with a $1 + O(\sqrt{\frac{\ln k}{k}})$ -Approximation Ratio.	223
18.4	Improved Algorithm and Proof of Main Theorem	225
18.5	Conclusion	231
19	Thin Trees for Laminar Families	232
19.1	Introduction	232
19.2	Preliminaries and Results	235
19.3	Notation	235
19.4	Polyhedral Background	235
19.5	Our Results	236
19.6	Proof Overview	237
19.7	Reduction to \mathcal{L} -aligned Points	238
19.8	Laminar thin trees for \mathcal{L} -aligned points via iterative relaxation	241
19.9	Conclusion	247
A	Preliminary Background	257
A.1	Graphs, Walks, and Circuits	257
A.2	Metric Completion and Equivalent Formulations of Metric TSP	258
A.3	The Double Tree Algorithm	259
A.4	The Christofides-Serdyukov Algorithm	259
A.5	Integrality Gap	260
B	Proofs from Section 6	263
C	Cuts crossed on one side	270
C.1	Cuts crossed on one side	270
C.2	Notation and results that we re-use	270
C.3	Main theorem for cuts crossed on one side	272
D	Proof of Theorem 12.1	274

1 Introduction

One of the flagship optimization problems in computer science is the metric traveling salesperson problem (TSP). Metric TSP is the following: given a list of cities and their pairwise symmetric distances, find the shortest route that visits each city at least once and returns to the starting point. Equivalently, given a weighted graph, we wish to find the shortest closed walk that visits every vertex. In the following, we will call such closed walks TSP tours.

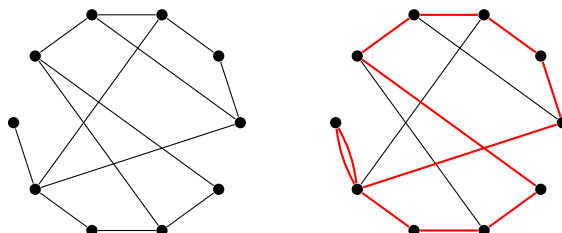


Figure 1: On the left, a graph in which all edge weights are equal to 1. On the right is a minimum cost TSP tour. It has cost 11.

Metric TSP has been of practical use for as long as traveling salespeople existed. See, for example, a note on designing practical TSP tours from 1832 discussed in [App+07]. And it is of immense use today, as TSP models a large class of planning and decision problems. Many companies use TSP to route and schedule their vehicles and airplanes; researchers at the NIH use TSP solvers to construct maps related to genome sequencing; semi-conductor companies use it to design chips and optimize drilling routes on circuitboards. Typically, these organizations solve TSP to obtain a solution to the standard problem using approximation algorithms and then modify it to satisfy additional application-specific constraints.

For theorists, though, TSP’s real utility comes from its ability to inspire exciting new mathematics. It has played the role of mathematical muse: it sparked the use of linear programming relaxations to model and solve non-convex optimization problems [DFJ54], was one of the first problems proved NP-Complete [Kar72], and was one of the first problems for which an approximation algorithm was developed [Chr76]. It also inspired local search techniques such as simulated annealing [KGV83] and the field of average case analysis [BHH59]. To give you a sense of the impact of these ideas, these five works collectively have roughly 80,000 citations and each created research areas that are active to this day.

It is difficult to say why exactly TSP encourages so much new mathematics. One possibility is that it is simply more studied than many other problems, perhaps because of its ubiquity in real life: popping into the theorist’s head when they run errands or go on a road trip. However, I think there is also something special about the mathematical structure of TSP: there is a *global* constraint on the set of edges (that the output is connected) as well as a *local* constraint (that every vertex has even degree), and – unlike many other problems – these local constraints cannot be fixed locally.¹ This throws a wrench in algorithm design and makes many methods unusable, prompting frustration and wonder in just the right ratio to hook the unsuspecting researcher. As Christos Papadimitriou once said, “TSP isn’t a problem, it’s an addiction.”

¹If I have a connected set of edges, and a vertex is odd, my impulse might be to simply add an arbitrary edge to that vertex to make it even. However, that edge might make another vertex odd. So, this local constraint also deals with the global structure of the graph.

1.1 Approximation Algorithms

Optimization problems like TSP abound in the world around us. Unfortunately, many of those routinely seen in applications are NP-Hard. In other words, unless $P=NP$, finding exact answers to these questions takes exponential time. Therefore, we must often accept finding *approximate* answers. Enter approximation algorithms:

For a minimization problem, we say an algorithm \mathcal{A} is a (randomized) α approximation algorithm if on every instance \mathcal{I} , the (expected) cost of $\mathcal{A}(\mathcal{I})$ is at most α times the cost of the optimal solution for \mathcal{I} .

In addition to providing worst-case guarantees and good approximate solutions quickly, developing approximation algorithms helps us understand the core difficulties of solving NP-Hard problems, illuminates their structure, and contributes to the algorithmic toolbox. Approximation algorithms are the main focus of this thesis.

The golden standard in approximation algorithms is to find an algorithm with α very close to 1. In certain special cases, it is even possible to find an approximation algorithm with $\alpha = 1 + \epsilon$ for any $\epsilon > 0$ and running time polynomial in n and $1/\epsilon$. For metric TSP, however, we know that any polynomial time algorithm must obey $\alpha \geq \frac{123}{122}$ unless $P=NP$ [KLS15]. However, we are currently far from finding an algorithm with a guarantee so close to 1.

In the 1970s, Christofides and Serdyukov independently found the same beautiful $3/2$ approximation algorithm for metric TSP [Chr76; Ser78]. The algorithm, often taught in a first algorithms course, is to find a minimum cost spanning tree T of the input graph and then add the minimum cost matching on the odd degree vertices of T . The resulting graph is Eulerian and therefore has an Eulerian tour, which is a TSP tour as it visits every vertex and returns to the starting point.²

Due to the algorithm's simplicity, some researchers believed it would quickly be improved upon. However that was not the case, and improving upon it remained open despite significant work, e.g. [Wol80; SW90; BP91; Goe95a; CV00; GLS05; BEM10; BC11a; SWZ12; HNR17; HN19; KKO20]. As stated by Bill Cook [Coo12]:

Christofides' algorithm first appeared in a Carnegie Mellon University research report in 1976. At the time, the result seemed easy enough. Thirty years later, with no improvements in sight, it no longer seems so simple. Indeed, it is a pressing problem to find a polynomial-time approximation algorithm with α less than 1.5, capable of handling all metric instances.

Case in point, it is at the top of a list of ten open problems in approximation algorithms in Shmoys and Williamson's textbook on the subject from 2011 [WS11].

The main result of this dissertation is to finally address this problem, albeit with an astronomically small improvement. In particular, I discuss joint work with my advisors Anna Karlin and Shayan Oveis Gharan in which we show the following theorem:

Theorem 1.1. *For some absolute constant $\epsilon > 10^{-36}$, there is a deterministic $\frac{3}{2} - \epsilon$ approximation algorithm for metric TSP.*

²If desired, this tour can be shortcut to a Hamiltonian cycle by assuming that the costs form a metric (as is the case in metric TSP).

As discussed in [Section 1.4](#) and summarized in the below [Table 1](#), this also improves the best known approximation algorithm for a range of other problems such as Path TSP and 2-ECSM.

Problem	Input	Approximability
Path TSP	An arbitrary symmetric metric, but find a path instead of a cycle	$1.5 - 10^{-36}$ approximation (this work with [TVZ20])
Prize-Collecting TSP	An arbitrary symmetric metric. Vertices are given a value $\pi(v)$, and the tour can skip vertices at a cost of $\pi(v)$ per vertex skipped	1.599 [BN23; BKN23]
Many Visits TSP	An arbitrary symmetric metric, and vertices have requirements $r(v)$ for the number of visits	$1.5 - 10^{-36}$ approximation (this work with [PS23])
Capacitated Vehicle Routing	An arbitrary symmetric metric, and vertices have demands between 0 and 1. Design a set of tours visiting all nodes such that each contains a fixed depot vertex and no tour has demand more than 1	$(1.5 - 10^{-36} + 2)(1 - \frac{1}{3000})$ (this work with [Fri+22; BV22])
2-Edge-Connected Multi-Subgraph	Given an arbitrary symmetric metric, buy the cheapest set of edges that 2-edge-connects the graph	$1.5 - 10^{-36}$ approximation (this work with [GB93b])

Table 1: Approximability of some TSP variants and related problems. While the approximability of Prize-Collecting TSP is not currently related to the TSP constant, over the years many works have used approximation algorithms for TSP as a subroutine, e.g. [\[Arc+11; Goe09\]](#).

1.2 High Level Overview of Techniques

The reason Christofides’ algorithm is no better than a $3/2$ approximation (see [Appendix A](#) for a proof that it is no worse than a $3/2$ approximation) is because it can be forced to choose a spanning tree with an expensive matching, as shown in [Fig. 2](#).

In other words, it is easy to design a graph where the minimum spanning tree has an expensive matching. However, intuitively, one might imagine that it is not so easy to design a graph in which “many” spanning trees simultaneously have expensive matchings. This observation is a small piece of intuition for the merits of the algorithm we study, first proposed by Oveis Gharan, Saberi, and Singh [\[Asa+10\]](#) in 2010. The algorithm is very similar to Christofides’ algorithm. It only differs in the first step: instead of picking a minimum spanning tree, it picks a *random* spanning tree from a well-chosen distribution of large support.

To find this distribution, we first solve the natural linear programming relaxation for TSP. This LP, first formulated by Dantzig, Fulkerson and Johnson [\[DFJ54\]](#), is often known as the subtour elimination LP or the Held-Karp relaxation (see also [\[HK70\]](#)). As input, we assume we get our TSP instance in the form of a weighted graph $G = (V, E)$ with metric completion c (see [Appendix A](#) for

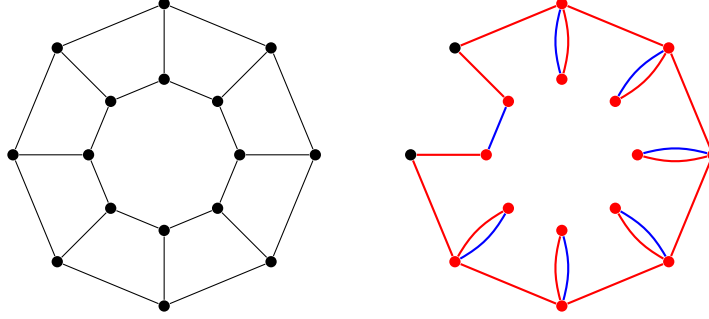


Figure 2: A tight example for Christofides' algorithm. Let $c_e = 1$ for all edges. In red is a minimum spanning tree, and marked in red are the odd vertices. In blue is a minimum cost matching on the odd vertices. We can force the algorithm to pick the red edges by making their costs just slightly cheaper than the remaining edges. The cost of the tour produced by the algorithm here is $\frac{3}{2}n - 2$, which is a $3/2$ approximation as $n \rightarrow \infty$.

various equivalent formulations of TSP and the definition of metric completion). Where for $F \subseteq E$ we let $x(F) = \sum_{e \in F} x_e$, the subtour LP finds the cheapest point $x \in P_{\text{Sub}}$, i.e. the one minimizing $\sum_{u,v} c_{\{u,v\}} x_{\{u,v\}}$. Where $\delta(S) := \{\{u,v\} \in E : u \in S, v \notin S\}$, P_{Sub} is defined as follows:

Subtour Polytope

$$P_{\text{Sub}} := \begin{cases} x(\delta(S)) \geq 2 & \forall S \subsetneq V \\ x(\delta(v)) = 2 & \forall v \in V \\ x_{\{u,v\}} \geq 0 & \forall u, v \in V \end{cases} \quad (1)$$

Notice that if we replace the constraint $x_{\{u,v\}} \geq 0$ with $x_{\{u,v\}} \in \{0,1\}$, this polytope describes the set of Hamiltonian tours. Thus it is a natural (and valid) relaxation for the problem, meaning that the cost of any optimal solution to this LP is at lower bound on the cost of the optimal TSP tour. By solving³ this LP we obtain a fractional solution x which we will use as a starting point for constructing our distribution.

It is well known that any point $x \in P_{\text{Sub}}$ can be written as a convex combination of 1-trees (spanning trees plus an edge) [Wol80; SW90]. Instead of finding an arbitrary convex combination, this algorithm finds the distribution μ over 1-trees such that $\mathbb{P}_{T \sim \mu}[e \in T] = x_e$ for all edges E , and *has as much entropy as possible* among all such distributions. This is why the algorithm is often called the “max entropy” algorithm for TSP. As hinted at above, one way to account for the improvement given by the max entropy algorithm is to argue that it is able to avoid such adversarial examples via its inherent randomness.

In order to prove this formally, we will need to have a deep understanding of the properties of the max entropy algorithm. It turns out that max entropy distributions over spanning trees are equivalent to so-called λ -uniform spanning tree distributions, in which we have a vector $\lambda \in \mathbb{R}_{\geq 0}^E$

³Note this LP has exponentially many constraints, one for each set $S \subsetneq V$. However there is an efficient separation oracle for the problem in the form of any global minimum cut algorithm, and it can be solved in strongly polynomial time (see [Sch03], pp. 984-986).

and the probability of a tree T is proportional to $\prod_{e \in T} \lambda_e$. This allows us to state the algorithm as follows:

Max Entropy Algorithm

1. Solve the subtour LP to obtain $x \in P_{\text{Sub}}$.(2)
2. Choose an arbitrary root r . Then there is a λ -uniform distribution μ_λ over the graph $G' = (V \setminus \{r\}, E \setminus \delta(r))$ such that $\mathbb{P}_{T \sim \mu_\lambda} [e \in T] = x_e$ (see [Section 2](#) for more details, and note that the equality is met with possibly exponentially small error).
3. Sample $T \sim \mu_\lambda$ and find the minimum cost matching M on the odd vertices of T .
Output $T \uplus M$.

We give a sample point $x \in P_{\text{Sub}}$ and its resulting distribution μ_λ in [Fig. 3](#) and [Fig. 4](#).

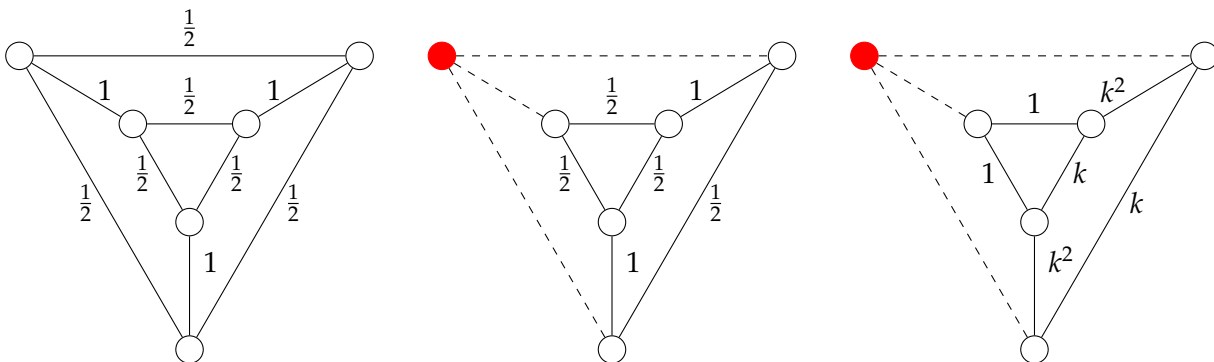


Figure 3: On the left is a point in P_{Sub} . In the middle, the root (chosen arbitrarily) is marked in red. The remaining nodes and non-dashed edges form a point $z \in P_{\text{ST}}$. On the right is a λ -uniform distribution with marginals approximately equal to z , where k is chosen to be a positive number tending to ∞ . (The ϵ error in [Theorem 2.6](#) goes to 0 as $k \rightarrow \infty$.)

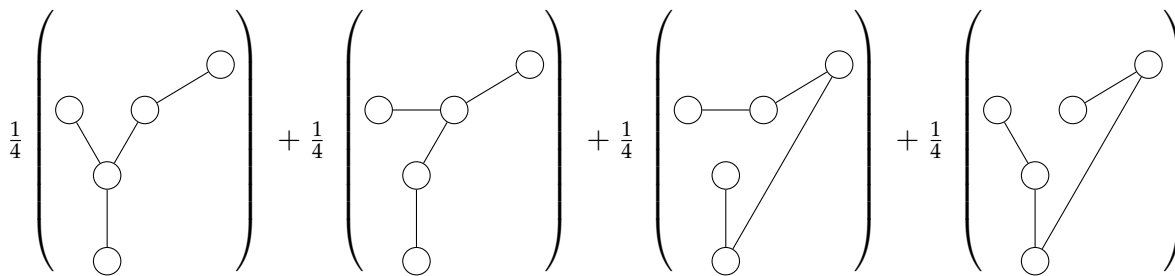


Figure 4: The trees and their probabilities in the above λ -uniform distribution. We show only the trees T for which $\lim_{k \rightarrow \infty} \mathbb{P}[T] > 0$ and give their probabilities in the limit. Note that after sampling, two edges from the red vertex above are added at random (sampled proportional to x_e).

λ -uniform distributions have been studied for some time, with their roots going back to the work of Kirchhoff and Borchardt in the mid-1800s. We will study λ -uniform distributions in

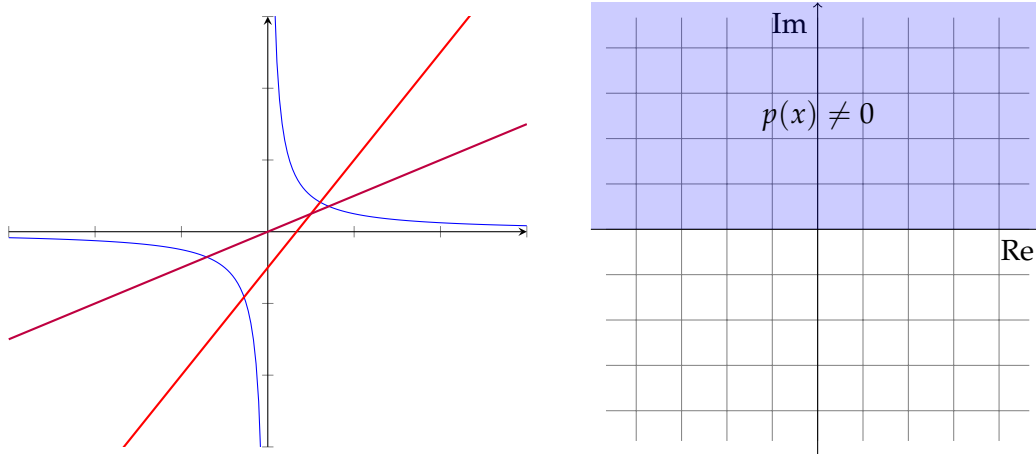


Figure 5: A distribution is strongly Rayleigh if its generating polynomial is real stable. One way to characterize a multivariate real stable polynomial is that its hyperbolicity cone contains the positive orthant, or equivalently that every restriction of the polynomial to a line is real rooted. $y = 1/x$ is real stable, as shown by the above plot: in blue are the real roots of $y = 1/x$, and every line with positive slope hits two real zeros and thus is real rooted. An equivalent definition of real stability is that the polynomial has no roots in the upper half of the complex plane, i.e. if all variables have positive imaginary value then the polynomial cannot vanish.

their own right, building on this long history. However, we will crucially use a more modern view of λ -uniform distributions as part of a broader class of *strongly Rayleigh (SR) distributions*, introduced by Borcea, Brändén, and Liggett [BBL09]. These distributions are characterized by the location of the roots of their generating polynomials (see Fig. 5 for a brief explanation). This allows us to study their properties using tools from the geometry of polynomials, linking operations from the two worlds: conditioning, for example, is now kin to differentiation, truncation with homogenization, projection with specialization, and so on.

By exploiting this link, probabilists have proved powerful properties of strongly Rayleigh distributions. In our proof we exploit a range of these properties and find some new ones as well. One elementary fact that is a consequence of this link is that every vertex has even degree in the sampled tree T from the algorithm above with probability at least $\frac{1}{2}(1 - e^{-2}) \approx 0.43$. This alone should give the reader some hope that max entropy is better than Christofides: in the bad example above, the matching is expensive because every vertex is odd. This fact demonstrates that such a thing does not occur in a typical max entropy tree.

However, properties of strongly Rayleigh distributions alone are not enough, and neither is having many even vertices.⁴ After using properties of λ -uniform and SR distributions to understand our spanning tree distribution, we turn to *the structure of the near-minimum cuts of graphs*. The cost of the matching is closely related to these structures, as we will explain in future sections. In particular, we prove new properties of the polygon representation [Ben95; BG08] and show that cut structures of a particular type lead to polygons with additional properties.

⁴For example, even if there are only *two* odd vertices for a given tree, if they have distance $OPT/2$ we get a $3/2$ approximation. Indeed, even the property that every vertex is even independently with probability $1 - o(1)$ would not even be enough to beat $3/2$ on the cycle graph. One would need to combine this property with other observations.

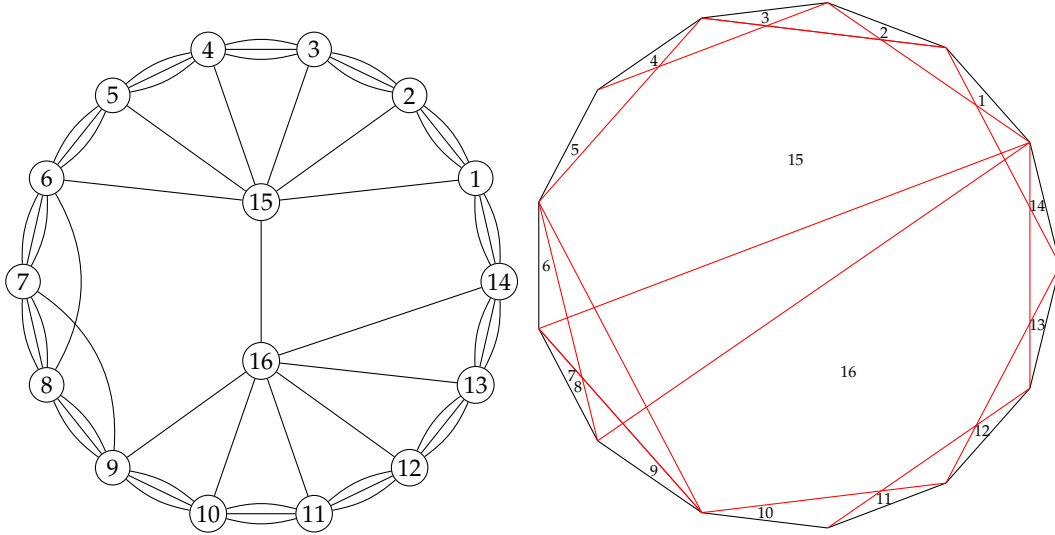


Figure 6: Consider the graph on the left, with minimum cut 7 (or, consider setting $x_e = \frac{2}{7}$ for all edges and a min cut of 2 as in the support graph of a solution to P_{Sub}). On the right is the *polygon representation* of the connected component of all proper cuts with at most 8 edges (or $x(\delta(S)) \leq 2 + 1/7$). We prove many new properties of this representation for use in our analysis.

These two pieces are the mathematical core of our proof. We then use these to carefully construct an analysis which shows that the matching is slightly cheaper than the one guaranteed by Christofides in expectation.

Finally, we show that the algorithm can be derandomized using the method of conditional expectation and the matrix tree theorem, which allows us to evaluate the generating polynomial of the max entropy distribution at any point in the complex plane. This derandomization shows that while randomness is a useful tool for analysis, it is not fundamental to the improvement.

1.3 Other Results on TSP

1.3.1 Lower Bounds for the Max Entropy Algorithm

We now know that max entropy out-performs Christofides' algorithm by at least a small constant. While perhaps the most pressing problem is to figure out a significantly improved analysis for max entropy ($\frac{3}{2} - 10^{-36}$ is certainly not the true approximation ratio of the algorithm), it is also worthwhile to determine *lower* bounds for the algorithm to better understand how far we might expect to push the analysis.

It has long been conjectured that the natural linear programming relaxation for TSP has an integrality gap of $4/3$, as we will discuss later – this is informally known as simply “the $4/3$ rd conjecture.” Therefore a natural target for max entropy is $4/3$. I will discuss joint work with Billy Jin and David Williamson in which we show for the first time that there are examples where the Eulerian tour returned by max entropy has cost strictly greater than $4/3$ times the cost of the LP. In particular, we show that a variant of the “ k -donut” instance introduced by Boyd and Sebő [BS21] gives a lower bound of $11/8$ for the Eulerian tour returned by max entropy. Interestingly,

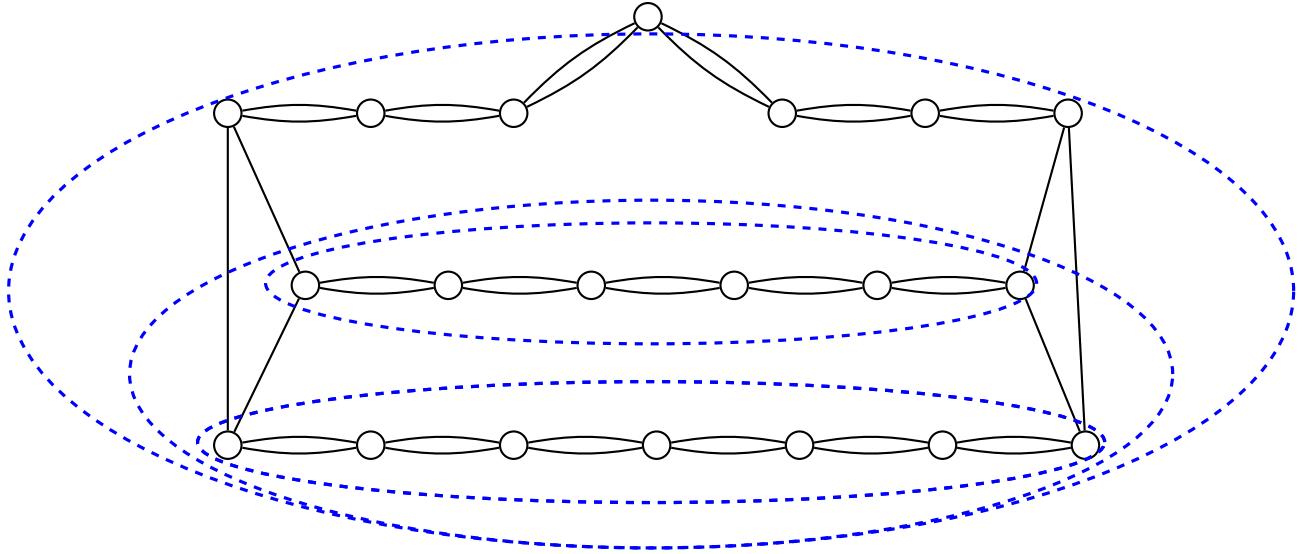


Figure 7: An example of a half integral cycle cut instance, in which every edge represents LP value $1/2$. This is also the canonical integrality gap example for the Subtour LP. The cuts present in the hierarchy are shown in blue.

our lower bound example is an instance of graphic TSP, and empirically, max entropy performs extremely well on graphic instances (see [GW17]). This highlights the often large gap between worst case and average case performance.

This work shows that despite its promise, the max entropy algorithm is unlikely to resolve the $4/3$ conjecture. However there is still a slim possibility that it does: we do not bound the cost of the tour produced after shortcutting. There may be a way to exploit this in the proof, although at the moment all known algorithms for TSP of which we are aware only look at the cost of the Eulerian tour.

1.3.2 Optimal Results for a Special Case of TSP

Instead of trying to bring down the approximation ratio from $3/2 - \epsilon$ to something slightly smaller, one in pursuit of proving the $4/3$ rd conjecture can ask: what is the largest class of instances for which we can find a $4/3$ approximation? In joint work with Billy Jin and David Williamson, we found a $4/3$ approximation for a class of instances we call “half integral cycle cuts,” and proved that the integrality gap of this class is at most $4/3$. This class includes the two known examples achieving the integrality gap of the LP relaxation for TSP, thus our result is tight.

Our results work by constructing a hierarchy of cuts similar to what is done in the analysis of max entropy, taking the form of a laminar family $\mathcal{L} \subseteq 2^V$. We then perform a top-down induction on \mathcal{L} : for each set in \mathcal{L} , we define a set of “patterns” of edges incident on it such that the set has even degree. For each pattern, we give a distribution of edges connecting the chain of child nodes in the cycle cut, which induces a distribution of patterns on each child. Crucially, we then show that there is a *feasible region* R of distributions over patterns, such that if the distribution of patterns on the parent node belongs to R , then the induced distribution on patterns on each child node also belong to R . The algorithmic idea is illustrated in Fig. 8.

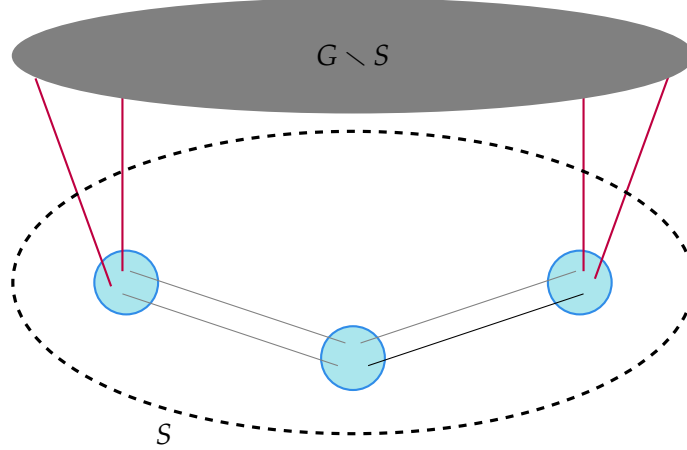


Figure 8: S is an example of a cut in \mathcal{L} with three children. In gray is the rest of the graph with S contracted. In our recursive algorithm, we are given a distribution of Eulerian tours over G/S , so in particular we know the distribution on the red edges. We then extend it to G with the children of S in \mathcal{L} contracted by picking a distribution over the black edges.

To handle the many possible configurations of the red edges, we designate four distinct states that they may be in and analyze the behavior of a Markov chain showing transitions between them. Our proof then boils down to finding transitions which have a good stationary distribution.

Interestingly, the lower bound example from above is in this class of instances. Thus, the algorithm in this work outperforms max entropy (without shortcutting) on this example, achieving $4/3$ instead of $11/8$. This suggests two interesting directions: (i) further develop this algorithm to work on larger classes of instances, and (ii) use ideas from this algorithm to modify the max entropy algorithm, so that it performs better on this class of examples.

1.4 Summary of Results on TSP and Corollaries

The majority of this thesis is devoted to proving the following theorem:

Theorem 1.2. *Let $x \in P_{\text{Sub}}$. Sample a tree T from the max entropy distribution μ_λ as in [Algorithm 1](#), and add a minimum cost matching M on the odd vertices of T . Then,*

$$\mathbb{E}[c(M)] \leq \left(\frac{1}{2} - \epsilon\right) c(x),$$

where $\epsilon > 10^{-36}$. Therefore, the max entropy algorithm produces a solution of expected cost at most $(\frac{3}{2} - \epsilon)c(x)$.

Implication for Path TSP. In recent exciting work, Traub, Vygen, Zenklusen [[TVZ20](#)] showed that an α -approximation algorithm for metric TSP can be used as a black box to get a $\alpha(1 + \epsilon)$ approximation algorithm for Path TSP. With the above theorem, this implies that there is a $3/2 - \epsilon$ approximation algorithm for Path TSP (for $\epsilon > 10^{-36}$). On the other hand, it is a folklore result that the integrality gap of the natural LP relaxation of Path TSP is at least $3/2$. Therefore, an interesting consequence of the above theorem is that although the best possible approximation

factors of the two problems are the same (up to polynomial reductions), the natural LP relaxation of metric TSP has a strictly smaller integrality gap.

Implication for 2-ECSM. In the 2-edge-connected multi-subgraph problem, or 2-ECSM for short, we are given a weighted graph G and we want to find a minimum cost 2-edge-connected spanning subgraph, where an edge can be chosen multiple times. The classical Christofides-Serdyukov algorithm gives a $3/2$ -approximation for 2-ECSM and despite significant attempts [CR98; BFS16; SV14; Boy+20] improved algorithms were designed only for special cases of the problem. Since in [GB93a] it is shown that LP (2) is a valid relaxation for 2-ECSM, we also obtain a randomized LP-relative $\frac{3}{2} - \epsilon$ approximation for the 2-edge-connected multi-subgraph problem.

We later show that the algorithm can be derandomized, giving the following theorem:

Theorem 1.3. *The max entropy algorithm can be derandomized using the method of conditional expectation. Therefore, there is a deterministic $\frac{3}{2} - 10^{-36}$ approximation algorithm for TSP, 2-ECSM, and Path TSP.*

As mentioned, we also show a *lower bound* on the performance of the algorithm.

Theorem 1.4. *There is an infinite family of instances of graphic TSP for which the performance of the max entropy algorithm is 1.375 as the size of the instances go to infinity.*

A major open question is to determine the true worst-case performance of the max entropy algorithm. We currently only know that it is in the range $[1.375, \frac{3}{2} - 10^{-36}]$.⁵

1.5 Other Results on Network Design Problems

I will also discuss new progress on two classic problems in network design: the k -edge-connected spanning multi-subgraph problem and the thin tree conjecture.

1.5.1 Designing k -Edge-Connected Graphs

In an instance of the minimum k -edge connected spanning subgraph problem, or k -ECSS, we are given an (undirected) graph $G = (V, E)$ with $n := |V|$ vertices and a cost function $c : E \rightarrow \mathbb{R}_{\geq 0}$, and we want to choose a minimum cost set of edges $F \subseteq E$ such that the subgraph (V, F) is k -edge connected. The k -edge-connected *multi*-subgraph problem, k -ECSM, is a close variant of k -ECSS in which we want to choose a k -edge-connected *multi*-subgraph of G of minimum cost, i.e., we can choose an edge $e \in E$ multiple times. Note that without loss of generality we can assume the cost function c in k -ECSM is a metric, i.e., for any three vertices $x, y, z \in V$, we have $c(x, z) \leq c(x, y) + c(y, z)$.

Around four decades ago, Fredrickson and Jájá [FJ81; FJ82] designed a 2-approximation algorithm for k -ECSS and a $3/2$ -approximation algorithm for k -ECSM. The latter essentially follows from a reduction to the traveling salesperson problem (TSP). Since then, despite a number of papers on the problem [CT00; KR96; Kar99; Gab05; GG08; Gab+09; Pri11; LOS12], progress was made only in cases where the underlying graph is unweighted or $k \gg \log n$. Most notably, Gabow, Goemans, Tardos and Williamson [Gab+09] showed that if the graph G is unweighted then k -ECSS

⁵Although as remarked, if you consider the version of the algorithm with shortcutting, it may be possible to go below 1.375.

and k -ECSM admit $1 + 2/k$ approximation algorithms, i.e., as $k \rightarrow \infty$ the approximation factor approaches 1.

Motivated by [Gab+09], Pritchard posed the following conjecture:

Conjecture 1.5 ([Pri11]). *The k -ECSM problem admits a $1 + O(1)/k$ approximation algorithm.*

In other words, if true, the above conjecture implies that the $3/2$ -classical factor can be substantially improved for large k , and moreover that it is possible to design an approximation algorithm whose factor gets arbitrarily close to 1 as $k \rightarrow \infty$. With Anna Karlin, Shayan Oveis Gharan, and Xinzhi Zhang we proved a weaker version of the above conjecture.

Theorem 1.6. *There is a polynomial time randomized algorithm for (weighted) k -ECSM with approximation factor (at most) $1 + \frac{5.06}{\sqrt{k}}$.*

This gives the first approximation algorithm with ratio tending to 1 as $k \rightarrow \infty$.

1.5.2 The Thin Tree Conjecture and the Laminar Crossing Spanning Tree Problem

A spanning tree T of a graph G is called α -thin if the number of edges of T crossing any given cut of G is at most an α fraction of the total number of edges. In 2004, Goddyn [God04] made the following conjecture: there exists a function $f : \mathbb{Z}_+ \rightarrow [0, 1]$ with $\lim_{k \rightarrow \infty} f(k)/k = 0$ such that every k -edge-connected graph G has an $f(k)$ -thin spanning tree. This has become known as the *thin tree conjecture*, and it remains open despite substantial efforts.

A natural strengthening of the conjecture, which we will refer to as the *strong thin tree conjecture* makes the same claim, but for $f(k) = C/k$ for some constant C . This conjecture is found explicitly in [Asa+17] and is the best that one could hope for up to constant factors; clearly no k -edge-connected graph has an α -thin tree for any $\alpha < 1/k$. In a different direction, there is also an algorithmic question one can ask: if a thin tree always exists, can we *find* one in polynomial time?

The thin tree conjecture has some surprising implications. It implies the *weak 3-flow conjecture* of Jaeger [Jae84]. This has since been resolved, by Thomassen [Tho12], however this would provide an alternate proof. Another application lies in the asymmetric traveling salesman problem (ATSP). As shown by Asadpour, Goemans, Madry, Oveis Gharan and Saberi [Asa+17; OS11], if the constructive form of the strong thin tree conjecture is true, it would yield an $O(1)$ -approximation algorithm to ATSP. This has since been resolved by Svensson, Tarnawski and Végħ [STV20] using completely different methods. Nonetheless, a new algorithm stemming from thin trees would be of significant interest.

In joint work with Neil Olver, we focus on a special case of the strong thin tree conjecture. We show that we can obtain a strong thin tree if we only need it to be thin with respect to an arbitrary *laminar* family of cuts. Previously results were only known if we want the tree to be thin with respect to the vertex set [Goe06; SL15] or a chain of constraints [OZ18]. This work generalizes both of these works in the context of the thin tree problem.

Our work also gives the first constant factor approximation to the laminar crossing spanning tree problem. Here, we are given a weighted graph and a laminar family of cuts \mathcal{L} with bounds b_S for each $S \in \mathcal{L}$, and we suppose a tree T exists which has at most b_S edges in $\delta(S)$ for every $S \in \mathcal{L}$. We show it is possible to find a tree with cost at most $O(1) \cdot c(T)$ and at most $O(1) \cdot b_S$ edges in every cut $S \in \mathcal{L}$. Previously, it was not even known how to achieve $O(1) \cdot b_S$ violation on the cut constraints and an arbitrary violation in cost.

2 Essential Background

This section will introduce essential background for the main result of this dissertation, a deterministic $\frac{3}{2} - \epsilon$ approximation for metric TSP for some constant $\epsilon > 10^{-36}$. It will:

- Introduce the subtour polytope, the spanning tree polytope, and the O -Join polytope and show that the integrality gap is at most 1.5 using the analysis of Wolsey [Wol80] and Shmoys and Williamson [SW90].
- Introduce the max entropy algorithm of Oveis Gharan, Saberi, and Singh [OSS11], the main focus of this work.
- Introduce real stable polynomials and strongly Rayleigh distributions which we will use to prove some essential properties of the max entropy algorithm.
- Introduce the structure of (near) minimum cuts and give some intuition about why their structure is useful in the proof.
- Give a brief overview on the current state of knowledge of the max entropy algorithm and some intuition for the proof.

In other words, here we give a much more in-depth introduction to our framework for studying metric TSP. For the reader unfamiliar with basic terminology and background on TSP, such as the definition of a closed walk, the notion of shortcutting an Eulerian tour to a Hamiltonian cycle, the double tree algorithm, and the proof that Christofides' algorithm is a $3/2$ approximation, see [Appendix A](#).

Before starting, the reader may want to review [Table 2](#). This shows a number of TSP variants and their approximability, placing metric TSP in its larger context.

2.1 Polyhedra of Interest

Here we describe three important polyhedra for studying TSP.

2.1.1 Notation

For a graph $G = (V, E)$ and a set $S \subseteq V$, we write

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

For a vector $x : E \rightarrow \mathbb{R}$, and a set $F \subseteq E$, we write $x(F) = \sum_{e \in F} x_e$. We write $x|_F$ to denote the restriction of x to the edges in F such that $x|_F \in \mathbb{R}^F$.

2.1.2 The Subtour LP

As mentioned in the introduction, the following linear program, first formulated by Dantzig, Fulkerson and Johnson [DFJ54], is extremely well-studied and an important tool for solving TSP instances. It is often known as the subtour elimination polytope or the Held-Karp LP relaxation

Problem	Input	Approximability
(Asymmetric) TSP	A (directed) weighted graph. Find the min cost (directed) Hamiltonian cycle.	No α approximation for any α [Kar72]
Asymmetric Metric TSP	An arbitrary asymmetric metric	$22 + \epsilon$ approximation [STV20; TV22]
Metric TSP	An arbitrary symmetric metric	$1.5 - 10^{-36}$ approximation (this work)
Graphic Metric TSP	The metric completion of an unweighted graph	1.4 approximation [OSS11; MS16; SV12]
Euclidean TSP	A Euclidean metric in \mathbb{R}^d	$1 + \epsilon$ approximation for any $\epsilon > 0$ if d is fixed [Aro96; Mit99]
Bounded Genus TSP	The metric completion of a bounded genus graph	$1 + \epsilon$ approximation for any $\epsilon > 0$ [DHM07] (following work on planar graphs [GKP95; Aro+98; Kle05])

Table 2: Fundamental classes of TSP instances and their approximability.

(see also [HK70]). As input, we assume we get a weighted graph $G = (V, E)$ with metric completion c . We now want to minimize $\sum_{u,v} c_{\{u,v\}} x_{\{u,v\}}$ subject to $x \in P_{\text{Sub}}$, defined as follows:

Subtour Polytope

$$P_{\text{Sub}} := \begin{cases} x(\delta(S)) \geq 2 & \forall S \subsetneq V \\ x(\delta(v)) = 2 & \forall v \in V \\ x_{\{u,v\}} \geq 0 & \forall u, v \in V \end{cases} \quad (2)$$

This LP has exponentially many constraints, one for each set $S \subsetneq V$. However there is an efficient separation oracle for the problem in the form of any global minimum cut algorithm, and it can be solved in strongly polynomial time (see [Sch03], pp. 984-986).

2.1.3 The Spanning Tree Polytope

For any graph $G = (V, E)$, Edmonds [Edm70] gave the following description for the convex hull of spanning tree of a graph G , known as the *spanning tree polytope*.

Spanning Tree Polytope

$$P_{\text{ST}} := \begin{cases} z(E) = n - 1 \\ z(E(S)) \leq |S| - 1 \quad \forall S \subsetneq V \\ z_e \geq 0 \quad \forall e \in E \end{cases} \quad (3)$$

Edmonds [Edm70] proved that the extreme point solutions of this polytope are the characteristic vectors of the spanning trees of G . This implies that the integrality gap of P_{ST} is 1.

First we note the following:

Fact 2.1. *Let $x \in P_{\text{Sub}}$. Then for all sets $S \subsetneq V$, we have $x(E(S)) \leq |S| - 1$.*

Proof. We have

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

Where we use that $x(\delta(S)) \geq 2$ for all $S \subsetneq V$. □

Therefore the only constraint broken by $x \in P_{\text{Sub}}$ is the constraint $x(E) = n - 1$. We get the following corollary:

Corollary 2.2. *Let $x \in P_{\text{Sub}}$. Then $\frac{n-1}{n}x \in P_{\text{ST}}$.*

A similar useful fact is as follows:

Fact 2.3. *Let $x \in P_{\text{Sub}}$ and $S \subsetneq V$ such that $x(\delta(S)) = 2$. Then, if P_{ST} is the spanning tree polytope on the graph $G' = (S, E(S))$, we have $x|_{E(S)} \in P_{\text{ST}}$.*

Proof. Let $z = x|_{E(S)}$. We have

$$\sum_{e \in E(S)} z_e = \frac{2|S| - x(\delta(S))}{2} = \frac{2|S| - 2}{2} = |S| - 1.$$

Thus the first constraint holds. The remaining constraints hold due to [Fact 2.1](#). □

This follows from [Fact 2.1](#).

Corollary 2.4. *Given a graph $G = (V, E)$, let $x \in P_{\text{Sub}}$ and suppose e is an edge with $x_e = 1$. Define x' such that $x'_f = x_f$ for all $f \neq e$ and $x'_e = 0$. Then $x' \in P_{\text{ST}}$.*

2.2 The O -Join Polyhedron and the Integrality Gap of P_{Sub}

The following characterization of the O -Join polytope is due to Edmonds and Johnson [EJ73]: for any graph $G = (V, E)$, cost function $c : E \rightarrow \mathbb{R}_+$, and a set $O \subseteq V$ with an even number of vertices, the minimum weight of an O -join equals the optimum value of the following integral linear program: minimize $c(y)$ subject to $y \in P_{\text{OJ}}$, where P_{OJ} is given by the following.

O-Join Polyhedron

$$P_{\text{OJ}} = \begin{cases} y(\delta(S)) \geq 1 & \forall S \subseteq V, |S \cap O| \text{ odd} \\ y_e \geq 0 & \forall e \in E \end{cases} \quad (4)$$

In other words, they proved that the integrality gap of P_{OJ} is 1. The following useful fact is immediate from the fact that $x(\delta(S)) \geq 2$ for all $S \subseteq V$.

Fact 2.5. *Let $x \in P_{\text{Sub}}$. Then for any $O \subseteq V$ (with $|O|$ even), $x/2 \in P_{\text{OJ}}$.*

We combine the above facts to show that the integrality gap of P_{Sub} is at most $3/2$, following the analysis of Wolsey [Wol80] and Shmoys and Williamson [SW90].

To do so, we show that Christofides' algorithm produces a solution of cost at most $\frac{3}{2}c(x)$ for any solution x to the subtour LP.

First, notice that cost of a minimum spanning tree T is at most $c(x)$. This is due to [Corollary 2.2](#). Since we can write $\frac{n-1}{n}x$ as a convex combination of spanning trees, the cheapest one must cost at most $c(x)$. It remains to show that the cost of the matching on T is at most $\frac{1}{2}c(x)$. However this follows from the fact that the integrality gap of the O -Join polyhedron is 1 and [Fact 2.5](#).

2.3 The Maximum Entropy Algorithm for TSP

Given the above background, we now introduce the max entropy algorithm.

2.3.1 λ -uniform Spanning Trees

The algorithm relies upon create a distribution of the following type.

λ -Uniform Spanning Trees

For a vector $\lambda : E \rightarrow \mathbb{R}_{\geq 0}$, a λ -uniform distribution μ_λ over spanning trees of a graph $G = (V, E)$ is a distribution where for every spanning tree $T \subseteq E$, $\mathbb{P}_{\mu_\lambda}[T] = \frac{\prod_{e \in T} \lambda_e}{\sum_{T'} \prod_{e \in T'} \lambda_e}$.

[Asa+10] showed that for any point $z \in P_{\text{ST}}$, one can find a λ -uniform distribution over trees with marginals z (up to negligible error) in polynomial time. In particular, they proved the following theorem:

Theorem 2.6 ([Asa+10, Theorem 5.2]). *Let z be a point in the spanning tree polytope (85) of a graph $G = (V, E)$. For any $\epsilon > 0$, a vector $\lambda : E \rightarrow \mathbb{R}_{\geq 0}$ can be found in time polynomial in n , $-\log \min_{e \in E} z_e$ and $\log(1/\epsilon)$ such that the corresponding λ -uniform spanning tree distribution μ_λ satisfies*

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

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

[Fig. 9](#) is an example output of the above theorem for the point z in the spanning tree polytope where the topmost edge has $z_e = \frac{1}{2}$ and the remaining edges have $z_f = \frac{3}{4}$. Notice that in this case $\epsilon = 0$, i.e. the λ -uniform distribution has marginals exactly z .

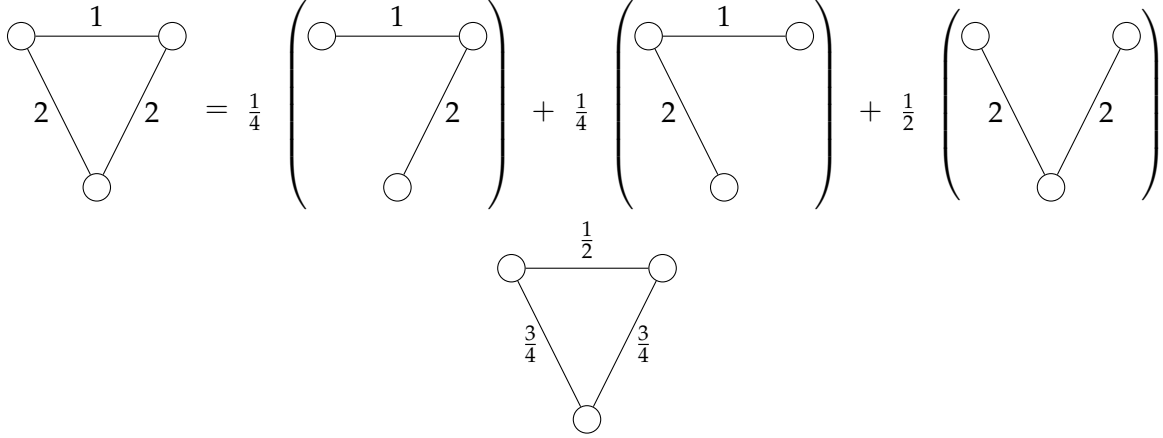


Figure 9: Top is an example of a λ -uniform spanning tree distribution. The normalizing term is $\sum_T \prod_{e \in T} \lambda_e = 8$, so the probability of each tree is $\frac{1}{8} \prod_{e \in T} \lambda_e$. The bottom picture represents the marginal vector of this distribution, i.e. it represents the probability of each edge. Thus, this λ -uniform distribution is a potential output of [Theorem 2.6](#) for the point $z \in P_{\text{ST}}$ above.

This computation can be done using the multiplicative weight update algorithm [[Asa+10](#)] or by applying interior point methods [[SV19](#)] or the ellipsoid method [[Asa+10](#)]. (We note that the multiplicative weight update method can only guarantee $\epsilon < 1/\text{poly}(n)$ in polynomial time.) We will sometimes call such a distribution the *maximum entropy* distribution because a λ -uniform distribution has maximal entropy over all distributions with marginals x .⁶

This theorem suggests the following algorithm for TSP. Instead of finding a minimum spanning tree as in Christofides, select a *random* spanning tree from the distribution μ_λ . Then add the minimum cost matching as before.

2.3.2 The Algorithm

Our starting point is to solve the subtour LP (2) to obtain a solution x^0 . We then pick an arbitrary node, u , split it into two nodes u_0, v_0 and set $x_{(u_0, v_0)} = 1, c(u_0, v_0) = 0$ and we assign half of every edge incident to u to u_0 and the other half to v_0 . This allows us to assume without loss of generality that x^0 has an edge $e_0 = (u_0, v_0)$ such that $x_{e_0} = 1, c(e_0) = 0$.

Using the above [Theorem 2.6](#), we find a vector λ such that for every edge $e \in E(G \setminus \{r\})$, $\mathbb{P}_{\mu_\lambda}[e \in T] = x_e(1 \pm \epsilon)$, for some $\epsilon < 2^{-n}$. We then sample a tree $T \sim \mu_\lambda$ and add the minimum cost matching on the odd degree vertices of T .

The guiding intuition for this algorithm is that a random spanning tree should have a cheaper matching on average than the minimum spanning tree. This is what we are able to prove. However to do so we need to understand properties of μ_λ . The next few sections are dedicated to giving some essential background for studying μ_λ .

We will use the following definition:

Definition 2.7 (Satisfied cuts). *For a set $S \subseteq V$ such that $u_0, v_0 \notin S$ and a spanning tree $T \subseteq E$ we say a vector $y : E \rightarrow \mathbb{R}_{\geq 0}$ satisfies S if one of the following holds:*

⁶Although, note that since we do not always find a distribution which preserves marginals exactly (as one does not necessarily exist), for precision we generally refer to the distribution used by the algorithm as λ -uniform instead.

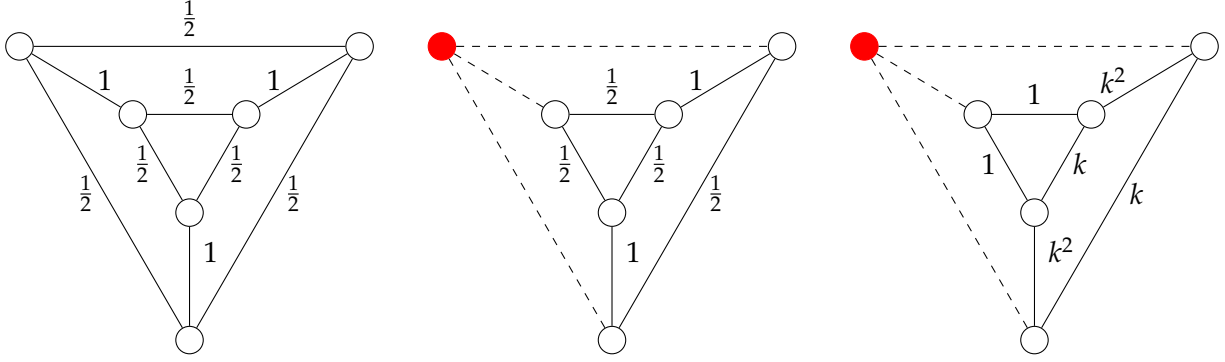


Figure 10: On the left is a point in P_{Sub} . In the middle, the root (chosen arbitrarily) is marked in red. The remaining nodes and non-dashed edges form a point $z \in P_{\text{ST}}$. On the right is a λ -uniform distribution with marginals approximately equal to z that could be the output of [Theorem 2.6](#), where k is chosen to be a large positive number. The ϵ error in [Theorem 2.6](#) goes to 0 as $k \rightarrow \infty$.

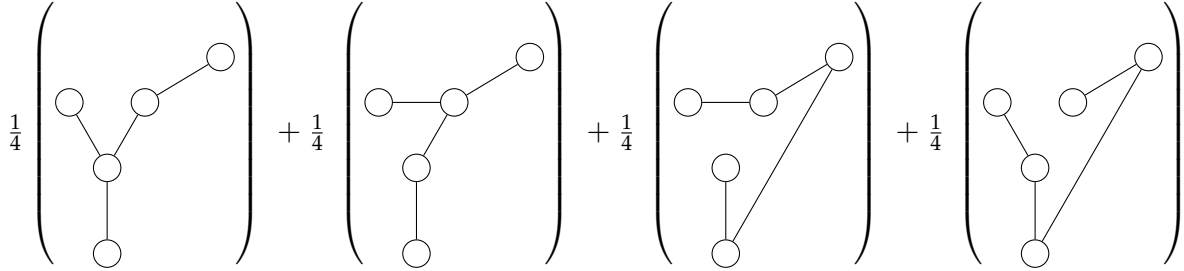


Figure 11: The trees and their probabilities in the above λ -uniform distribution. We show only the trees T for which $\lim_{k \rightarrow \infty} \mathbb{P}[T] > 0$ and give their probabilities in the limit. Note that after sampling, two edges from the red vertex above are added at random (sampled proportional to x_e).

Algorithm 1 Max Entropy Algorithm (slight modification of [[OSS11](#)])

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

- $\delta(S)_T$ is even, or
- $y(\delta(S)) \geq 1$.

In other words, a cut is satisfied if the constraint corresponding to S is satisfied in P_{OJ} (where O is the odd vertices of T) by y or the constraint does not exist.

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

2.4 Generating Polynomials and λ -uniform Spanning Trees

Given a vector $z \in \mathbb{R}^E$ and a set $S \subseteq E$, let $z^S := \prod_{e \in S} z_e$. Let $\mu : \{0, 1\}^E \rightarrow \mathbb{R}$ be a probability distribution over subsets of E . The generating polynomial $g_\mu \in \mathbb{R}[\{z_e\}_{e \in E}]$ of μ is defined as follows:

$$g_\mu(z) := \sum_S \mu(S) z^S.$$

In general, this polynomial may have exponentially many terms and no compact representation. However, if μ is the uniform distribution over spanning trees on a graph G (equivalently, the 1-uniform distribution), then Kirchhoff's matrix tree theorem tells us that g_μ has a very succinct representation, which we prove here for completeness.

Theorem 2.8 (Matrix tree theorem). *For a graph $G = (V, E)$ let $g_{\mathcal{T}}$ be the generating polynomial of the 1-uniform spanning tree distribution on G . Pick a root $r \in V$ arbitrarily. Then,*

$$g_{\mathcal{T}}(\{z_e\}_{e \in E}) = \det\left(\sum_{e \in E} z_e L_e\right),$$

where for $e = \{u, v\}$ with $u, v \neq r$ we let $L_e \in \mathbb{R}^{V \setminus \{r\}} = (\mathbf{1}_u - \mathbf{1}_v)(\mathbf{1}_u - \mathbf{1}_v)^T$ be the Laplacian of e . For $e = \{u, r\}$ we let $L_e = \mathbf{1}_u \mathbf{1}_u^T$.

Proof. Let $b_e = (\mathbf{1}_u - \mathbf{1}_v)$ for $e = \{u, v\}$ with $u, v \neq r$ and $b_e = \mathbf{1}_u$ for $e = \{r, u\}$. Let $B \in \mathbb{R}^{(n-1) \times m}$ be the matrix with columns b_e for all $e \in E$ (where $|E| = m$). Then, by Cauchy-Binet, where Z is the diagonal matrix with entries z_e ,

$$\det\left(\sum_{e \in E} z_e L_e\right) = \det(BZB^T) = \sum_{S \in \binom{[m]}{n-1}} \det((BZ)_{[n-1], S} B_{S, [n-1]}^T) = \sum_{S \in \binom{[m]}{n-1}} z^S \det(B_{[n-1], S})^2$$

If the edges in S contain a cycle, then $\det(B_{[n-1], S}) = 0$ as the set of edges in the cycle are linearly dependent. So it remains to show that the determinant is ± 1 if S forms a spanning tree. To see this, notice that since the tree contains a leaf, there must be a row of the matrix with exactly one non-zero "leaf" entry. We can place this entry in the bottom right of the matrix, and recurse (considering the leaf as deleted in the tree and the matrix), leaving us with an upper triangular matrix. This implies $\det(B_{[n-1], S}) = \pm 1$ as desired. \square

We will often use the slightly more convenient form

$$g_{\mathcal{T}}(\{z_e\}_{e \in E}) = \det\left(\sum_{e \in E} z_e L_e + 11^T/n\right) \quad (5)$$

where $L_e = (\mathbf{1}_u - \mathbf{1}_v)^T$ for all $e = \{u, v\}$. We leave the proof of the equivalence as an exercise.

This immediately generalizes to any λ -uniform distribution μ_λ . We can simply write (using the above, and assuming λ is normalized such that $\sum_{T \in \mathcal{T}} \lambda^T = 1$):

$$g_{\mu_\lambda}(z) = \sum_{T \in \mathcal{T}} \lambda^T z^T = g_{\mathcal{T}}(\{\lambda_e z_e\}_{e \in E}) = \frac{1}{n} \det\left(\sum_{e \in E} z_e \lambda_e L_e + 11^T/n\right)$$

Another consequence of this fact which we will rely on later is that g_{μ_λ} can be computed at any point in \mathbb{C} in polynomial time.

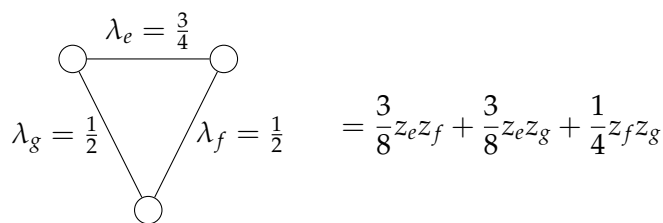


Figure 12: A normalized λ -uniform distribution and its generating polynomial.

For example, for the λ -uniform distribution from Fig. 12, one can check:

$$\frac{1}{3} \det \begin{pmatrix} \frac{3}{4}z_e + \frac{1}{2}z_g + \frac{1}{3} & -\frac{3}{4}z_e + \frac{1}{3} & -\frac{1}{2}z_g + \frac{1}{3} \\ -\frac{3}{4}z_e + \frac{1}{3} & \frac{3}{4}z_e + \frac{1}{2}z_f + \frac{1}{3} & -\frac{1}{2}z_f + \frac{1}{3} \\ -\frac{1}{2}z_g + \frac{1}{3} & -\frac{1}{2}z_f + \frac{1}{3} & \frac{1}{2}z_f + \frac{1}{2}z_g + \frac{1}{3} \end{pmatrix} = \frac{3}{8}z_e z_f + \frac{3}{8}z_e z_g + \frac{1}{4}z_f z_g$$

2.5 Real Stable Polynomials

For a field \mathbb{F} , let $\mathbb{F}[z_1, \dots, z_n]$ be the ring of polynomials with coefficients in \mathbb{F} in the indeterminates z_1, \dots, z_n .

Real Stability

Let $\mathcal{H} = \{z : \Im(z) > 0\}$ be the upper half of the complex plane. We say a polynomial $p \in \mathbb{C}[z_1, \dots, z_n]$ is **stable** if $z \in \mathcal{H}^n$ implies $p(z) \neq 0$, i.e. it has no zeros in \mathcal{H}^n .^a

We say a polynomial p is **real stable** if it is stable and $p \in \mathbb{R}[z_1, \dots, z_n]$.

^aNote we do not treat the zero polynomial as stable.

For example, $z_1 z_2 + 1$ is not stable since $z_1 = z_2 = i$ is a root in \mathcal{H}^2 . However, $z_1 z_2 - 1$ is (real) stable, since for any root in \mathcal{H}^n we have z_1, z_2 not equal to 0 so $z_2 = \frac{1}{z_1}$. However $\Im(z_1)$ and $\Im(\frac{1}{z_1})$ cannot both be positive.

We will often work with the following equivalent definition of stability.

Fact 2.9. A polynomial $p \in \mathbb{C}[z_1, \dots, z_n]$ is (real) stable if and only if for every $a \in \mathbb{R}_{>0}^n$ and every $b \in \mathbb{R}^n$, the univariate polynomial $p(ax + b)$ is (real) stable.

Proof. This follows from the identity $\mathcal{H}^n = \{ax + b \mid a \in \mathbb{R}_{>0}^n, b \in \mathbb{R}^n, x \in \mathcal{H}\}$. \square

Recall the following for univariate polynomials with real coefficients.

Fact 2.10. If $p \in \mathbb{R}[x]$, then if x is a root of p so is \bar{x} .

Proof. Let $p = \sum a_i x^i \in \mathbb{R}[x]$ and $x \in \mathbb{C}$ for which $p(x) = 0$. Then,

$$0 = \overline{p(x)} = \overline{\sum a_i x^i} = \sum \bar{a}_i \bar{x}^i = \sum a_i \bar{x}^i = p(\bar{x}).$$

More abstractly, this is because complex conjugation is an automorphism of \mathbb{C} which fixes \mathbb{R} . \square

Note that the above is *not* true for univariate polynomials with complex coefficients, for example $z + i$ has just a single root $z = -i$. This is because complex conjugation does not fix \mathbb{C} . An immediate consequence of this is the following:

Corollary 2.11. *A univariate polynomial $p \in \mathbb{R}[x]$ is real stable if and only if it is real rooted.*

Thus, by Fact 2.9, a polynomial $p \in \mathbb{R}[z_1, \dots, z_n]$ is real stable if and only if for every $a \in \mathbb{R}_{>0}^n$ and every $b \in \mathbb{R}^n$, the univariate polynomial $p(ax + b)$ is real rooted.

We also record the following fact about univariate real rooted polynomials that we will give as a consequence a key lemma in future sections:

Lemma 2.12 ([Edr53]). *Let $p \in \mathbb{R}_{\geq 0}[x]$ be real rooted, and suppose $p(1) = 1$. Then where $p(x) = a_0 + a_1x + a_2x^2 + \dots + a_dx^d$ there exists independent Bernoullis B_1, \dots, B_d such that $\forall i, a_i = \mathbb{P}\left[\sum_{j=1}^d B_j = i\right]$.*

Proof. Let $\lambda_1, \dots, \lambda_d \in \mathbb{R}$ be the negation of the (not necessarily distinct) roots of p . Note $\forall i, \lambda_i \geq 0$ since p has non-negative coefficients. Then $p(x) = C \prod_{i=1}^d (x + \lambda_i)$ for some $C > 0$, and since $p(1) = 1$, we have $C = \frac{1}{\prod_{i=1}^d (1 + \lambda_i)}$. Therefore

$$p(x) = \prod_{i=1}^d \frac{x + \lambda_i}{1 + \lambda_i} = \prod_{i=1}^d \left(\frac{x}{1 + \lambda_i} + \left(1 - \frac{1}{1 + \lambda_i}\right) \right)$$

However this is simply $\prod_{i=1}^d (p_i x + (1 - p_i))$ for Bernoullis with success probabilities $p_i = \frac{1}{1 + \lambda_i}$, as desired. \square

The following is an important class of real stable polynomials:

Lemma 2.13. *Suppose $p(z_1, \dots, z_n) = \det(\sum_{i=1}^n z_i P_i + S)$ for positive semidefinite matrices $P_1, \dots, P_n \in \mathbb{R}^{n \times n}$ and a symmetric matrix $S \in \mathbb{R}^{n \times n}$. Then, p is real stable.*

Proof. Consider any $a \in \mathbb{R}_{>0}^n, b \in \mathbb{R}^n$ and the polynomial

$$q(x) = \det \left(\sum_{i=1}^n (a_i x + b_i) P_i + S \right) = \det \left(x \sum_{i=1}^n a_i P_i + \sum_{i=1}^n b_i P_i + S \right).$$

By the above fact it is enough to prove that q is real rooted.

First assume that each P_i is positive definite. Then $D = \sum_{i=1}^n a_i P_i$ is positive definite since $a \in \mathbb{R}_{>0}^n$, so it has a symmetric square root $D^{1/2}$ and symmetric negative square root $D^{-1/2}$. Therefore we can rewrite this polynomial as

$$\det(D^{1/2}) \det \left(xI + D^{-1/2} \left(\sum_{i=1}^n b_i P_i + S \right) D^{-1/2} \right) \det(D^{1/2})$$

$\det(D^{1/2})$ is a constant, so this is real rooted if $\det(xI + D^{-1/2}(\sum_{i=1}^n b_i P_i + S)D^{-1/2})$ is real rooted. However this is the characteristic polynomial of the matrix $Q = D^{-1/2}(\sum_{i=1}^n b_i P_i + S)D^{-1/2}$, so its roots are the eigenvalues of Q . Finally we observe that Q is symmetric since all P_i and S are symmetric, so its eigenvalues are all real, implying that this polynomial is real rooted as desired.

If the P_i are not positive definite, this follows from taking the limit of $P_i + \epsilon I$ as $\epsilon \rightarrow 0$ and applying Hurwitz's theorem, summarized in the below corollary. \square

Corollary 2.14 (Of Hurwitz's theorem (1.3.8 in [RS05])). *A polynomial obtained as the limit of a convergent sequence of stable polynomials is stable.*

Real Stability of g_{μ_λ}

It follows from [Lemma 2.13](#) that the generating polynomial g_{μ_λ} of a λ -uniform spanning tree distribution is real stable, since by [Theorem 13.2](#) we can write it as a constant times $\det(\sum_{e \in E} z_e \lambda_e L_e + 11^T/n)$ (and $\lambda_e L_e$ are all positive semidefinite matrices and $11^T/n$ is symmetric).

Said again, the zeros of the polynomial g_{μ_λ} for any graph G and any $\lambda \in \mathbb{R}_{\geq 0}$ avoid the upper half of the complex plane, i.e. if $\Im(z_e) > 0$ for all $e \in E$, then $g_{\mu_\lambda}(z) \neq 0$.

Real stable polynomials are closed under many natural operations, some of which we list below:

- **External field.** $p(z_1, \dots, z_n) \rightarrow p(\lambda_1 z_1, \dots, \lambda_n z_n)$ for $\lambda \in \mathbb{R}_{\geq 0}$.
- **Specialization.** $p(z_1, \dots, z_n) \rightarrow p(z_1, \dots, z_{i-1}, a, z_{i+1}, \dots, z_n)$ for $a \in \mathbb{R}$ and $i \in [n]$.
- **Diagonalization.** $p(z_1, \dots, z_n) \rightarrow p(z_i, \dots, z_i)$ for any $i \in [n]$.
- **Differentiation.** $p(z_1, \dots, z_n) \rightarrow \frac{\partial}{\partial z_i} p(z_1, \dots, z_n)$ for any $i \in [n]$.

We leave these as an exercise. Note differentiation will need the Gauss-Lucas theorem: that the roots of p' for $p \in \mathbb{R}[x]$ lie in the convex hull of the roots of p .

2.6 Strongly Rayleigh Distributions

We can now introduce a class of probability distributions key to this thesis. Let \mathcal{B}_E be the set of all probability measures on 2^E .

Strongly Rayleigh distributions

We say $\mu \in \mathcal{B}_E$ is a strongly Rayleigh distribution, defined in [BBL09], if g_μ is real stable.^a The above shows that every λ -uniform distribution over spanning trees is strongly Rayleigh.

^aNotice that we only consider SR distributions on 2^E . Thus, the generating polynomial of any SR distribution is multiaffine.

There are a number of useful closure properties of strongly Rayleigh (SR) distributions. A simple one is as follows:

Fact 2.15. *If μ, μ' are strongly Rayleigh distributions then so is $\mu \times \mu'$.*

Proof. g_μ and $g_{\mu'}$ are both real stable. But $g_{\mu \times \mu'} = g_\mu g_{\mu'}$. Thus the zeros of $g_{\mu \times \mu'}$ are the union of the zeros of g_μ and $g_{\mu'}$. \square

We summarize a number of important closure operations here:

Closure Operations of SR Distributions. SR distributions are closed under the following operations. Here we are given a SR distribution $\mu \in \mathcal{B}_E$ and we are sampling a set $S \sim \mu$.

- **Projection.** For any $F \subseteq E$, the projection of μ onto F is SR and is defined as the measure μ_F where for any $A \subseteq F$,

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

- **Conditioning.** For any $e \in E$, $\mu_{|e \in S}$ and $\mu_{|e \notin S}$ are SR, i.e. we can condition on an element to be in or out of the sample.
- **Truncation.** For any integer $k \geq 0$ and $\mu \in \mathcal{B}_E$, the truncation of μ to k , μ_k defined as μ conditioned on $|S| = k$, is SR.⁷
- **Product.** As in the above fact, for any two SR distributions $\mu_E \in \mathcal{B}_E, \mu_F \in \mathcal{B}_F$ the product measure $\mu_{E \times F}$ is SR. In particular this is the measure where for any $A \subseteq E, B \subseteq F$, $\mu_{E \times F}(A \cup B) = \mu_E(A)\mu_F(B)$.

These can be derived from the closure properties of stable polynomials listed in the previous subsection. For example, conditioning an element e in is equivalent to differentiating with respect to z_e , and conditioning an element out is equivalent to specializing $z_e = 0$.

Rank Sequence. The *rank sequence* of a distribution μ is the sequence

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

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

$$\tilde{g}_\mu(z) := g_\mu(z, z, \dots, z).$$

Observe that $\tilde{g}(\cdot)$ is the generating polynomial of the rank sequence of μ . By the above closure properties of real stable polynomials, if μ is SR then \tilde{g}_μ is real rooted. Thus, the following is an immediate corollary of [Lemma 2.12](#):

Corollary 2.16. *Let μ be a strongly Rayleigh distribution. Then the rank sequence r_0, r_1, \dots, r_m of μ follows the law of a sum of independent Bernoullis, i.e. there exists Bernoullis B_1, \dots, B_m such that $r_i = \mathbb{P}[\sum_{j=1}^m B_j = i]$.*

Bernoullis are particularly easy to analyze, especially when one applies the following theorem of Hoeffding:

Theorem 2.17 ([\[Hoe56, Corollary 2.1\]](#)). *Let $g : \{0, 1, \dots, n\} \rightarrow \mathbb{R}$ and $0 \leq q \leq n$ for some integer $n \geq 0$. Let B_1, \dots, B_n be n independent Bernoulli random variables with success probabilities p_1, \dots, p_n , where $\sum_{i=1}^n p_i = q$ that minimizes (or maximizes)*

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

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

⁷Note that SR distributions are not necessarily closed under truncation of a proper subset, i.e., if we require exactly k elements from $F \subsetneq E$. We can of course project and then truncate, however we then lose the distribution on the remaining elements.

We can now describe one of the most important aspects of SR distributions for our purposes:

Sum of Bernoullis and Even Vertices

We call a random variable with expectation q which has the same law as a sum of independent Bernoullis as $BS(q)$. By the closure of projection and the above corollary, we get the following as immediately from [Corollary 2.16](#), which is a crucial window into the behavior of the max entropy algorithm.

Corollary 2.18. *Let $F \subseteq E$ and let μ_λ be the λ -uniform distribution from [Algorithm 1](#) with marginals x . Then, the random variable $|F \cap T|$ is a $BS(x(F))$.*

It is intuitive that an important property of μ_λ should be that *vertices have even degree in T with constant probability.*^a The above corollary along with [Theorem 2.17](#) lets us prove this with ease, detailed shortly:

Lemma 2.19. *Let $T \sim \mu_\lambda$ where the marginals of x are a solution to (2) as in [Algorithm 1](#). Then for all vertices v , the probability that v has even degree in T is at least **0.43**.*

More generally, if $S \subseteq V$ and $2 \leq x(\delta(S)) < 2.2$, then

$$\mathbb{P}_{T \sim \mu_\lambda} [|\delta(S) \cap T| \text{ even}] \geq \frac{1}{2}(1 - e^{-2(x(\delta(S))-1)}) \geq 0.43.$$

^aIn particular, its behavior should differ from Christofides as detailed in [Fig. 83](#).

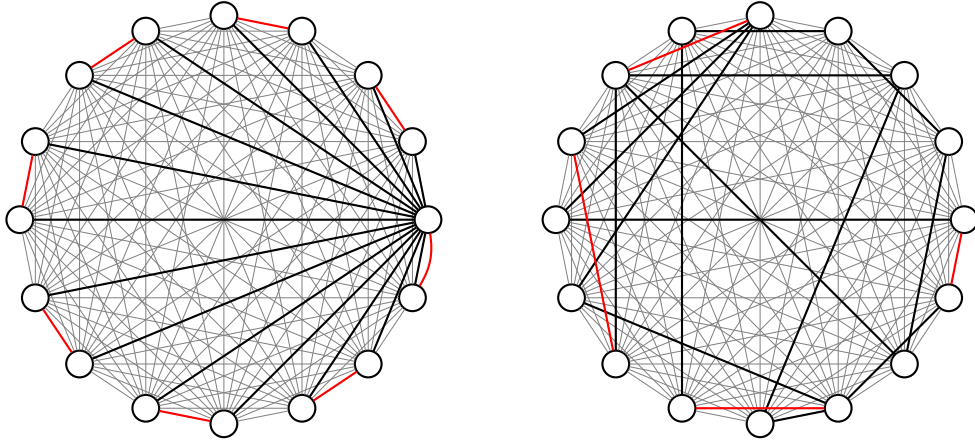


Figure 13: Trees are shown in black and matchings in red. Assume that all edges have cost 1 in this complete graph. While Christofides' algorithm (left) may pick a minimum spanning tree with many odd vertices, causing the matching on the odd vertices to be expensive (here 8 edges), a tree sampled from a λ -uniform or max entropy distribution (right) will usually have a cheaper matching (here only 4 edges). In this simple example, this is because in the tree drawn from the max-entropy distribution will have many vertices of even degree by [Lemma 2.19](#).

To prove the above lemma, we first need the following simple fact, which demonstrates the benefits of working with Bernoullis.

Fact 2.20. Let B_1, \dots, B_n be independent Bernoulli random variables each with expectation $0 \leq p \leq 1$. Then

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

Proof. Note that

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

Summing them up we get,

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

as desired. □

This now lets us prove the following:

Lemma 2.21. Given a $BS(q)$ random variable with $0 < q \leq 1.2$, then

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

Proof. We apply Hoeffding's theorem [Theorem 2.17](#) to the function indicating if $BS(q)$ is even. First, if $q \leq 1$, then by Hoeffding's theorem we can write $BS(q)$ as sum of n Bernoullis with success probability $p = q/n$ for some $n \geq 1$. If $n = 1$, then the statement obviously holds. Otherwise, by the previous fact, we have (for some n),

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

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

So, now assume $q > 1$. Write $BS(q)$ as the sum of n Bernoullis, each with success probabilities 1 or p . First assume we have no Bernoullis with success probability 1. Then, either we only have two non-zero Bernoullis with success probability $q/2$ in which case

$$\mathbb{P} [BS(q) \text{ even}] = (q/2)^2 + (1 - q/2)^2 \leq 0.6^2 + 0.4^2 = \frac{1}{2}$$

and we are done. Otherwise, $n \geq 3$ so $p \leq 1/2$ and similar to the previous case we get $\mathbb{P} [BS(q) \text{ even}] \leq \frac{1}{2}(1 + e^{-2q})$.

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

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

where we used that $p \leq 0.5$ (since $q \leq 1.2$). □

Now to obtain [Lemma 2.19](#), we observe that (i) by [Corollary 2.18](#), $|\delta(S) \cap T|$ for any cut is a $BS(x(\delta(S)))$; (ii) since any cut has at least 1 edge in T (with probability 1) there is a Bernoulli with success probability 1 in $BS(x(\delta(S)))$, so we can model the parity of $|\delta(S) \cap T|$ as $1 + BS(x(\delta(S)) - 1)$; and finally (iii) we apply [Lemma 2.21](#) so $BS(x(\delta(S)) - 1)$.

Another beautiful consequence of the above Bernoulli fact (see [[HLP52](#); [Dar64](#); [BBL09](#)]) is that the rank sequence of any strongly Rayleigh measure is *ultra log concave* (see below for the definition), unimodal, and its mode differs from the mean by less than 1.

Definition 2.22 (Log-concavity [[BBL09](#), Definition 2.8]). *A real sequence $\{a_k\}_{k=0}^m$ is log-concave if $a_k^2 \geq a_{k-1} \cdot a_{k+1}$ for all $1 \leq k \leq m - 1$, and it is said to have no internal zeros if the indices of its non-zero terms form an interval (of non-negative integers). It is ultra log concave if $\left(\frac{a_k}{\binom{m}{k}}\right)^2 \geq \frac{a_{k-1}}{\binom{m}{k-1}} \cdot \frac{a_{k+1}}{\binom{m}{k+1}}$.*

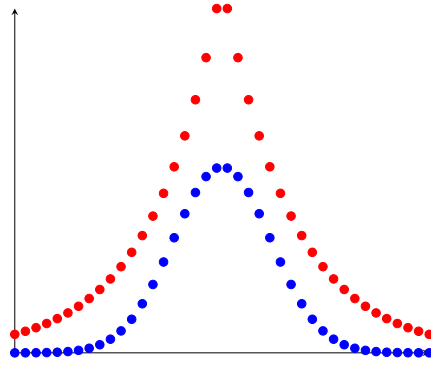


Figure 14: Two log concave sequences.

2.7 Additional Properties of SR Distributions

In this section we provide some additional properties of SR distributions. On a first reading one could skip this subsection.

Negative Dependence Properties. An *upward event*, \mathcal{A} , on 2^E is a collection of subsets of E that is closed under upward containment, i.e. if $A \in \mathcal{A}$ and $A \subseteq B \subseteq E$, then $B \in \mathcal{A}$. Similarly, a *downward event* is closed under downward containment. An *increasing function* $f : 2^E \rightarrow \mathbb{R}$, is a function where for any $A \subseteq B \subseteq E$, we have $f(A) \leq f(B)$. We also say $f : 2^E \rightarrow \mathbb{R}$ is a *decreasing function* if $-f$ is an increasing function. So, an indicator of an upward event is an increasing function. For example, if E is the set of edges of a graph G , then the existence of a Hamiltonian cycle is an increasing function, and the 3-colorability of G is a decreasing function.

Definition 2.23 (Negative Association). *A measure $\mu \in \mathcal{B}_E$ is negatively associated if for any increasing functions $f, g : 2^E \rightarrow \mathbb{R}$, that depend on disjoint sets of edges,*

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

It is shown in [[BBL09](#); [FM92](#)] that strongly Rayleigh measures are negatively associated.

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

$$\begin{aligned} \sum_B \rho(A, B) &= \mu(A), \forall A \in 2^E, \\ \sum_A \rho(A, B) &= \nu(B), \forall B \in 2^E, \end{aligned}$$

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

Theorem 2.24 ([BBL09]). *If μ is strongly Rayleigh and μ_k, μ_{k+1} are well-defined, then $\mu_k \preceq \mu_{k+1}$.*

Note that in the above particular case the coupling ρ satisfies the following: For any $A, B \subseteq E$ where $\rho(A, B) > 0$, $B \supseteq A$ and $|B \setminus A| = 1$, i.e., B has exactly one more element.

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

Fact 2.25 ([BBL09, Theorems 4.8, 4.19]). *Let μ be any SR distribution on E , then for any $F \subseteq E$, and any integer k*

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

The following fact is a direct consequence of the above, see e.g. Corollary 6.10 of [OSS11].

Fact 2.26. *Let μ be a homogenous SR distribution on E . Then,*

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

$$\mathbb{E}_\mu [B_T | A_T = 0] \leq \mathbb{E}_\mu [B_T] + \mathbb{E}_\mu [A_T] \tag{6}$$

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

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

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

2.8 Staying in the λ -Uniform Distribution

Since λ -uniform spanning tree distributions are special classes of SR distributions, if we perform any of the above operations on a λ -uniform spanning tree distribution μ we get another SR distribution. However the distributions are no longer necessarily λ -uniform. Below we list two special operations that also preserve the λ -uniform property (with possibly a different λ). The second operation is especially important to our proof.

Closure Operations of λ -uniform Spanning Tree Distributions For $G = (V, E)$, a spanning tree distribution $\mu \in \mathcal{B}_E$, and $T \sim \mu$, we have:

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

Note that λ -uniform spanning tree distributions are not necessarily closed under truncation and projection.⁸ We remark that SR measures are also closed under an analogue of tree conditioning, i.e., for a set $F \subseteq E$, let $k = \max_{S \in \text{supp } \mu} |S \cap F|$. Then, $\{\mu \mid |S \cap F| = k\}$ is SR. But if μ is a spanning tree distribution we get an extra *independence* property which is a property we heavily rely on throughout the entire work. Thus a general proof strategy we follow in this work to study a particular event is to start with μ_λ , apply several tree conditioning operations, and only then apply the stronger closure properties of general SR distributions. The independence property is as follows:

Fact 2.27. For a graph $G = (V, E)$, and a vector $\lambda(G) : E \rightarrow \mathbb{R}_{\geq 0}$, let $\mu_{\lambda(G)}$ be the corresponding λ -uniform spanning tree distribution. Then for any $S \subsetneq V$,

$$\{\mu_{\lambda(G)} \mid |E(S) \cap T| = |S| - 1\} = \{\mu_{\lambda(G)} \mid S \text{ is a subtree}\} = \mu_{\lambda(G[S])} \times \mu_{\lambda(G/S)}.$$

Proof. Intuitively, this holds because in the max entropy distribution (recall a λ -uniform distribution maximizes entropy subject to matching the marginals of x), conditioned on S being a tree, any tree chosen inside S can be composed with any tree chosen on G/S to obtain a spanning tree on G . So, to maximize the entropy these trees should be chosen independently. More formally for any $T_1 \in G[S]$ and $T_2 \in G/S$,

$$\begin{aligned} \mathbb{P}[T = T_1 \cup T_2 \mid S \text{ is a tree}] &= \frac{\lambda^{T_1} \lambda^{T_2}}{\sum_{T'_1 \in G[S], T'_2 \in G/S} \lambda^{T'_1} \lambda^{T'_2}} \\ &= \frac{\lambda^{T_1}}{\sum_{T'_1 \in G[S]} \lambda^{T'_1}} \cdot \frac{\lambda^{T_2}}{\sum_{T'_2 \in G/S} \lambda^{T'_2}} \\ &= \mathbb{P}_{T'_1 \sim G[S]} [T'_1 = T_1] \mathbb{P}_{T'_2 \sim G/S} [T'_2 = T_2], \end{aligned}$$

giving independence. □

In the following, to denote that $|E(S) \cap T| = |S| - 1$ we simply say “ S is a tree.” This situation is actually quite common due to the following lemma, which is very crucial to our proof:

Lemma 2.28. Let $G = (V, E, x)$, and let μ be any distribution over spanning trees with marginals x . For any ϵ -near min cut $S \subseteq V$ (such that none of the endpoints of $e_0 = (u_0, v_0)$ are in S), we have

$$\mathbb{P}_{T \sim \mu} [S \text{ is a subtree of } T] = \mathbb{P}_{T \sim \mu} [|T \cap E(S)| = |S| - 1] \geq 1 - \epsilon/2.$$

⁸For example, if we project to $E(S)$ for some $S \subseteq V$ and then truncate to μ_k for some $k < |S| - 1$, we have an SR distribution on $E(S)$ but we are not sampling a spanning tree anymore.

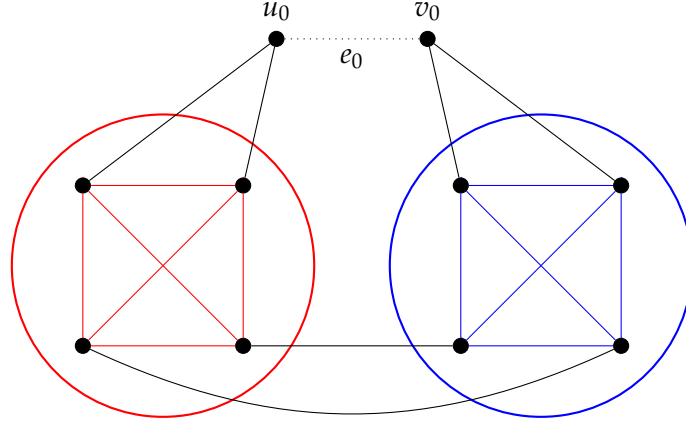


Figure 15: Suppose this is a solution to P_{Sub} where every displayed edge has $x_e = 1/2$. Then by [Fact 2.27](#) and [Lemma 2.28](#), (i) the sampled tree is always a subtree inside the red set and inside the blue set, and (ii) the edges in red are sampled independent of all other edges in T (and similarly for blue).

Proof. First, observe that

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

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

Let $p_S = \mathbb{P}_{T \sim \mu} [S \text{ is a subtree of } T]$. Then, we must have

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

Therefore, $p_S \geq 1 - \epsilon/2$. □

Near minimum cuts and independence

As an immediate corollary of [Fact 2.27](#) and [Lemma 2.28](#), we obtain the following.

Let $S \subseteq V$ be a near minimum cut, i.e. $x(\delta(S)) \leq 2 + \eta$. Then, with probability at least $1 - \eta/2$, **the distribution μ_λ decomposes into $\mu_\lambda = \mu_{\lambda(G[S])} \times \mu_{\lambda(G/S)}$.**

This is an incredibly powerful fact and says that (near) minimum cuts essentially break our tree distribution into a collection of independent pieces. See [Fig. 15](#) for an example. We will use this fact over and over again in our proofs.

We also record a useful fact here:

Corollary 2.29. *Let $A, B \subseteq V$ be disjoint sets such that $A, B, A \cup B$ are $\epsilon_A, \epsilon_B, \epsilon_{A \cup B}$ -near minimum cuts w.r.t., x respectively, where none of them contain endpoints of e_0 . Then for any distribution μ of spanning trees on E with marginals x ,*

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

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

2.9 The Structure of (Near) Minimum Cuts

Recall [Fact 2.5](#), which claims that if $x \in P_{\text{Sub}}$ then $x/2 \in P_{\text{OJ}}$ for any $O \subseteq V$ with $|O|$ even. This proves that the max entropy algorithm is a $3/2$ approximation.

Our proof strategy will show that there is a random variable $y : \mathcal{T} \rightarrow \mathbb{R}^E$ such that: (i) $y \in P_{\text{OJ}}$ with probability 1, where O is the set of odd vertices in T , and (ii) $\mathbb{E}[c(y)] \leq (\frac{1}{2} - \epsilon)c(x)$ for some $\epsilon > 0$. This would complete the proof as it upper bounds the expected cost of the matching by slightly less than half of the cost of the LP. To do this, we will show the following:

Slack vector and slack space

For some very small $\eta > 0$ there exists a slight perturbation of $x/2$ we call a “slack vector” $s : \mathcal{T} \rightarrow \mathbb{R}^E$ such that:

- (i) Given a tree T , $s(\delta(S)) \geq 0$ with probability 1 for every $S \subseteq V$ with $|\delta_T(S)|$ odd and $x(\delta(S)) \leq 2 + \eta$.
- (ii) $\mathbb{E}[s_e] \leq -\epsilon x_e$ for all $e \in E$ for some $\epsilon > 0$,
- (iii) $s_e \geq -\frac{1}{2+\eta}x_e$ for all $e \in E$ with probability 1.

We will call the space which $x/2 + s$ lives the “slack space.”

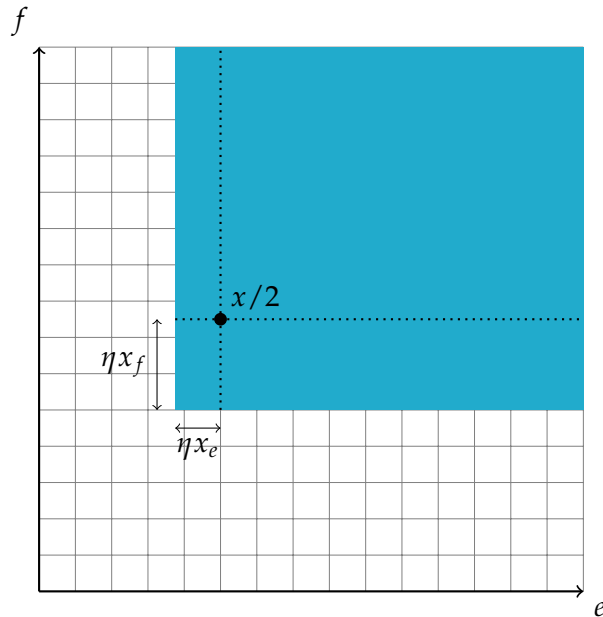


Figure 16: The space in which $x/2 + s$ lives (with probability 1) projected onto e and f . We will sometimes call this the slack space.

The existence of such a vector immediately implies a $\frac{3}{2} - \epsilon$ approximation for TSP. First observe that $y = x/2 + s$ has the property $y \in P_{\text{OJ}}$ with probability 1. The reason is as follows: for every cut $S \subseteq V$, if $x(\delta(S)) \geq 2 + \eta$, then $y(\delta(S)) \geq (2 + \eta)(\frac{1}{2 + \eta}) \geq 1$ by condition (iii). Therefore, the cuts that are not η -near minimum are never violated in P_{OJ} , regardless of their parity in the sampled tree T . However, condition (i) implies that all the near minimum cuts are satisfied. So, to finish the proof one simply uses (ii). By the integrality of the O -Join polyhedron, the expected cost of the matching is at most the expected cost of the solution $x/2 + s$ we have designed, which itself has expected cost at most $(\frac{1}{2} - \epsilon)c(x)$.

The reason we design a slack vector of this type, which only deals with the near minimum cuts, is that it reduces the space of constraints we have to worry about from *every* cut to a much smaller set. To handle the near minimum cuts, we will need to intimately understand their structure. In this section, as a warmup we introduce the structure of *minimum* cuts (where $\eta = 0$) and briefly introduce polygons, a tool for understanding the structure of near minimum cuts.

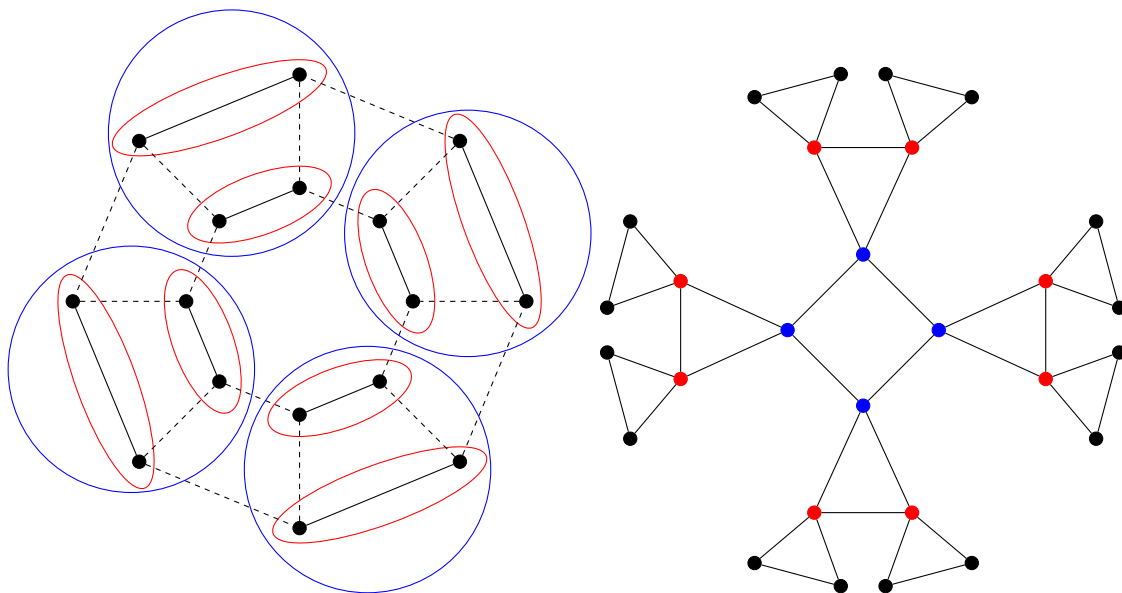


Figure 17: Here is an example support graph and its cactus representation which is a variant of the k -donut introduced by Boyd and Sebő [BS21]. The dotted edges have $x_e = 1/2$ and the solid edges $x_e = 1$. One can check that every minimum cut of the graph on the left is represented in the cactus as a cut containing exactly two edges. In particular the mapping ϕ from Theorem 2.34 should be clear: each of the black vertices in the left is mapped to a black vertex on the right. We will return to this example later: it shows a lower bound of $11/8$ for the performance of the max entropy algorithm as the number of nodes goes to infinity.

2.9.1 Notation

For a set of edges $A \subseteq E$ and (a tree) $T \subseteq E$, we write

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

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

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

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

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

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

2.9.2 The Cactus Representation

Definition 2.30. Consider a graph $G = (V, E)$ with min-cuts of value k .

- Any set $S \subseteq V$ such that $|\delta(S)| = k$ (i.e., its boundary is a min-cut) is called a **tight set**.
- A cut (S, \bar{S}) is **proper** if $|S| \geq 2$ and $|\bar{S}| \geq 2$.
- Two sets S and S' **cross** if all of $S \setminus S', S' \setminus S, S \cap S'$ and $V \setminus (S \cup S')$ are non-empty.

To discuss the cactus representation, we will rely on a number of basic facts about min-cuts. For proofs, see [FF09]. Suppose G is a k -edge connected graph, i.e. $|\delta(S)| \geq k$ for all $S \subseteq V$.

Fact 2.31. If two tight sets S and S' cross, then each of $S \setminus S', S' \setminus S, S \cap S'$ and $\overline{S \cup S'}$ are tight. Moreover, there are no edges from $S \setminus S'$ to $S' \setminus S$, and there are no edges from $S \cap S'$ to $\overline{S \cup S'}$.

Therefore, if two distinct tight sets S and S' cross each other, then $\delta(S) \cap \delta(S') = \emptyset$.

The following fact is especially useful to us, since the support graph of x is 4-edge-connected.

Fact 2.32. Suppose that every proper mincut is crossed by some other proper mincut. Then k is even and G is a cycle, with $k/2$ parallel edges between each adjacent pair of vertices.

Definition 2.33 (Cactus Graph). A loopless and 2-edge connected graph $C = (U, F)$ is a cactus if each edge belongs to exactly one cycle.

Theorem 2.34 (Cactus Representation [DKL76]). Let $G = (V, E)$ be a loopless graph with min-cut size $k \geq 1$. There is a cactus $C = (U, F)$ and a mapping $\phi : V \rightarrow U$ such that the 2-element cuts of C are in one to one correspondence with the min-cuts of G . Equivalently, S is a tight set of G if and only if $\phi(S)$ is a tight set of C .

2.9.3 Near Minimum Cuts

The previous subsection demonstrates that minimum cuts have a very simple structure. Unfortunately, near minimum cuts get quite a bit more complicated. Here are some essential facts about them. One can see that the lemmas are much more powerful when $\epsilon = 0$.

Lemma 2.35 ([OSS11]). For $G = (V, E, x)$, let $A, B \subsetneq V$ be two crossing ϵ_A, ϵ_B near min cuts respectively. Then, $A \cap B, A \cup B, A \setminus B, B \setminus A$ are $\epsilon_A + \epsilon_B$ near min cuts.

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

$$x(\delta(A \cap B)) + x(\delta(A \cup B)) \leq x(\delta(A)) + x(\delta(B)) \leq 4 + \epsilon_A + \epsilon_B.$$

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

The following lemma is proved in [Ben97]:

Lemma 2.36 ([Ben97, Lem 5.3.5]). *For $G = (V, E, x)$, let $A, B \subsetneq V$ be two crossing ϵ -near minimum cuts. Then,*

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

Lemma 2.37. *For $G = (V, E, x)$, let $A, B \subsetneq V$ be two ϵ near min cuts such that $A \subsetneq B$. Then*

$$x(\delta(A) \cap \delta(B)) = x(E(A, \overline{B})) \leq 1 + \epsilon, \text{ and}$$

$$x(\delta(A) \setminus \delta(B)) \geq 1 - \epsilon/2.$$

Proof. Notice

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

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

Summing these up, we get

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

Since $B \setminus A$ is non-empty, $x(\delta(B \setminus A)) \geq 2$, which implies the first inequality. To see the second one, let $C = B \setminus A$ and note

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

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

2.9.4 The Polygon Representation

Near min cuts can be represented by what is known as a polygon, which we will use to study the η -near minimum cuts of a fractionally 2-edge-connected graph. All of the statements are generalizable to the $(1 + \eta)\alpha$ near minimum cuts of an α -edge-connected graph (by rescaling) for $\eta \leq 6/5$.

Definition 2.38 (Connected Component of Crossing Cuts). *Given the set of η -near min cuts of a graph $G = (V, E)$, construct a graph where two cuts are connected by an edge if they cross. Partition this graph into maximal connected components. In the following, we will consider maximal connected components \mathcal{C} of crossing cuts and simply call them connected components. We say a connected component is a singleton if it has exactly one cut and a non-singleton otherwise.*

For a connected component \mathcal{C} , let $\{a_i\}_{i \geq 0}$ be the coarsest partition of vertices V such that for any $C \in \mathcal{C}$, either $a_i \subseteq C$ or $a_i \subseteq \overline{C}$. Each set a_i is called an atom of \mathcal{C} and we write $\mathcal{A}(\mathcal{C})$ to denote the set of all atoms.

Note for any atom $a_i \in \mathcal{A}(\mathcal{C})$ which is an η -near min cut, (a_i, \bar{a}_i) is a singleton component, and is not crossed by any η -near min cut. Therefore $(a_i, \bar{a}_i) \notin \mathcal{C}$.

We can now represent any cut in $S \in \mathcal{C}$ either by the set of vertices it contains or as a subset of $\mathcal{A}(\mathcal{C})$. In the following, we will often identify an atom with the set of vertices that it represents⁹.

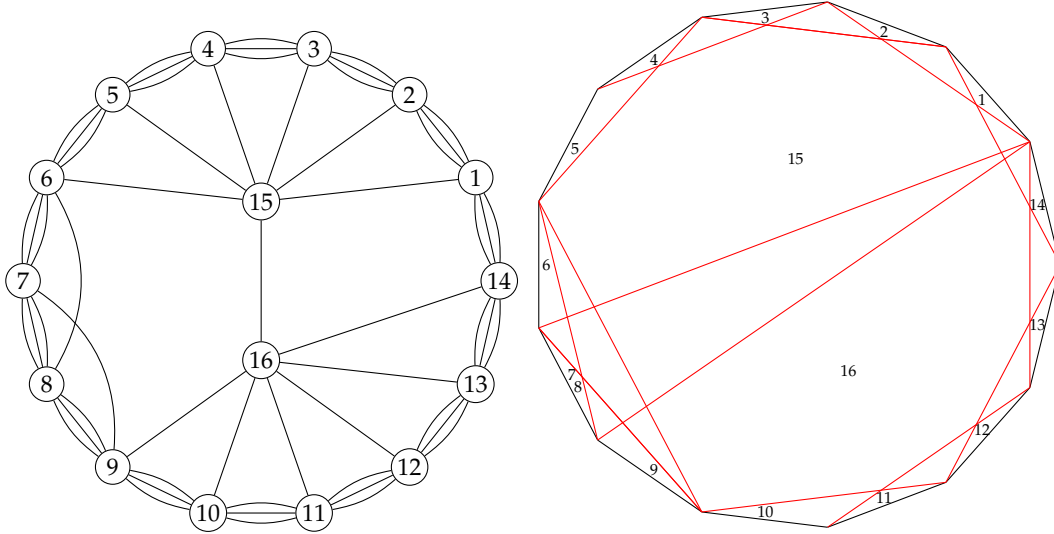


Figure 18: Consider the graph on the left, with minimum cut 7 (or, consider setting $x_e = \frac{2}{7}$ for all edges and a min cut of 2 as in the support graph of a solution to P_{Sub}). On the right is the polygon representation of the connected component of all proper cuts with at most 8 edges (or $x(\delta(S)) \leq 2 + 1/7$). This component consists of all proper near minimum cuts of the graph excluding the cut $\{7,8\}$, which is in its own connected component of size 1. As in the below definition, 15 and 16 are inside atoms, the others are outside atoms. Note $\{7,8\}$ is a single atom.

To study these systems, we will utilize the polygon representation of near minimum cuts defined in [Ben97] and then extended in [BG08]. Their work implies that any connected component \mathcal{C} of crossing η -near minimum cuts has a polygon representation with the following properties, so long as $\eta \leq \frac{2}{5}$:

⁹For example, it will be convenient to write cuts as subsets of atoms. In this case the cut is the union of the vertices in those atoms.

Polygon Representation

1. A polygon representation is a convex regular polygon with a collection of *representing diagonals*. All polygon edges and diagonals are drawn using straight lines in the plane. The diagonals partition the polygon into *cells*.
2. Each atom of \mathcal{C} is mapped to a cell of the polygon. If one of these cells is bounded by some portion of the polygon boundary it is *non-empty* and we call its atom an *outside atom*. We call the atoms of all other non-empty cells *inside atoms*. Note that some cells may not contain any atom. WLOG label the outside atoms a_0, \dots, a_{m-1} in counterclockwise order, and label the inside atoms arbitrarily. We also label points of the polygon p_0, \dots, p_{m-1} such that outside atom a_i is on the side (p_i, p_{i+1}) and a_0 is on the side (p_{m-1}, p_0) . (In future sections we will refer to the special atom called the root, and if it is an outside atom WLOG we will label a_0 as the root.)
3. No cell has more than one incident outer polygon edge.
4. Each representing diagonal defines a cut such that each side of the cut is given by the union of the atoms on each side. Furthermore, the collection of cuts given by these diagonals is exactly \mathcal{C} .

The following fact follows immediately from the above discussion:

Fact 2.39. *Any cut $S \in \mathcal{C}$ (represented by a diagonal of P) must have at least two outside atoms.*

Definition 2.40 (Outside atoms). *For a polygon P and a set S of atoms of P , we write $O_P(S)$ to denote the set of outside atoms of P in S ; we drop the subscript when P is clear from context. We also write $O(P)$ (or $O(\mathcal{A}(\mathcal{C}))$ where \mathcal{C} is the connected component of P) to denote the set of all outside atoms of P .*

Note that, given $S \in \mathcal{C}$, since S may be identified with a set of atoms, $O(S)$ is also well defined.

3 Building Up: Degree Cuts and The Half Integral Case

We will start with two special cases in terms of the structure of $x \in P_{\text{Sub}}$ to slowly build up to the fully general case. The first, called the “degree cut case” is not too difficult to beat $3/2$ on and provides a good introduction to our proof strategy. And despite its simplicity, it is an intriguing open question to design an algorithm (or analysis for max entropy) with a significantly improved constant in this simple case.

In the second case, we assume that every edge $e \in E$ has $x_e \in \{0, \frac{1}{2}, 1\}$. This assumption on x is known as the *half integral* case and has been studied extensively in its own right. The proofs in this section closely follow [KKO20], in which we obtained a 1.49993 approximation for the half integral case prior to the improvement in the general case. The best known approximation for this case is now 1.4983 [Gup+22], an approach which combines max entropy with a matroid-based rounding algorithm introduced in [HN19].

The half integral case (building on the degree cut case) is an especially useful window into the proof in the general case. As we highlight at the end of the section (Section 3.6), while the analysis ideas we present may at first sound specialized to the half integral case, essentially all of them will be present in some form in the general case. More than that, these ideas really appear in their “idealized” form in the half integral case. So, we highly recommend fully understanding this section before moving on, as then in future sections a reader can see a complicated statement and parse it more quickly by comparing it to its cleaner analog in the half integral case.¹⁰

3.1 The Degree Cut Case

Here we assume that there exists an absolute constant $\eta > 0$ such that for every proper cut $S \subseteq V$ we have $x(\delta(S)) \geq 2 + \eta$. We allow one exception for the cut containing u_0, v_0 , which by definition of the algorithm has $x(\delta(S)) = 2$. In this section we will deal with this by contracting u_0 and v_0 to a single vertex r we call the root (this is achievable since we assume $c(\{u_0, v_0\}) = 0$). Note that $\delta_T(r) = 2$ with probability 1.

As will be the case for the remainder of the paper, “all” we need to do is show how to construct the slack vector $s : \mathcal{T} \rightarrow \mathbb{R}^E$ with the three properties detailed in Section 2.9. Since this is the first time we explicitly construct such a vector, we remind the reader of the three necessary properties:

- (i) Given a tree T , $s(\delta(S)) \geq 0$ with probability 1 for every $S \subseteq V$ with $|\delta_T(S)|$ odd and $x(\delta(S)) \leq 2 + \eta$.
- (ii) $\mathbb{E}[s_e] \leq -\epsilon x_e$ for all $e \in E$ for some $\epsilon > 0$,
- (iii) $s_e \geq -\frac{1}{2+\eta} x_e$ for all $e \in E$ with probability 1.

The reason the degree cut case is such a massive simplifying assumption is that condition (i) instead simply reads that for all *vertices* v with $\delta_T(v)$ odd we require $s(\delta(v)) \geq 0$.

In the half integral case, as mentioned we have $\eta = 1$. This is because by splitting every edge with $x_e = 1$ into two edges with $x_e = 1/2$, we obtain a 4-regular 4-edge connected graph. Such a graph is Eulerian, so every cut has an even number of edges. Therefore in the degree cut case, every non-vertex cut S has $x(\delta(S)) \geq 3$.

¹⁰Of course, many completely new techniques are required in the general case so this trick does not always work.

In this section, we say that an edge $e = (u, v)$ is **good** if $\mathbb{P}_{T \sim \mu} [u, v \text{ both even in } T] \geq p$, where we say a vertex v is even in a tree T if $\delta(v)_T$ is even and we define $p = 1/27$. We will first assume the following probabilistic lemma and obtain a $\frac{3}{2} - \epsilon$ approximation in the degree cut case. We will then prove it in the following subsection.

Lemma 3.1 (Good Edges). *Let v be a vertex. Then, if G_v is the set of good edges adjacent to v , $x(G_v) \geq 1/2$.*

We note that not all edges are good, and call these edges “bad.” Indeed, the variant of the k -donut has a large number of bad edges. In the below picture (where we omit the root for a cleaner picture), *every* edge connecting the two cycles is bad!

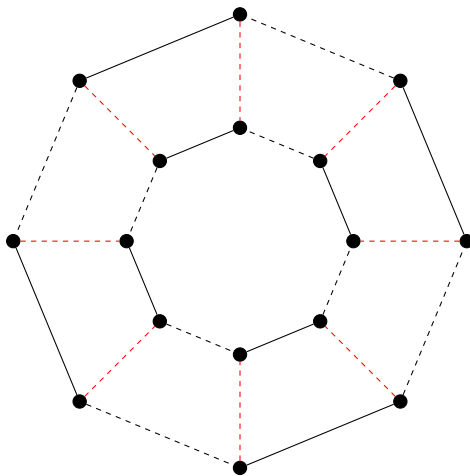


Figure 19: In the k -donut variant that is a running example in this dissertation, every red edge (with the exception of those adjacent to the root, not shown here) has the property that its two vertices never have the same parity in the tree at the same time. In this section, we call such edges “bad.” Note in future sections the definition of bad depends on the structure of near minimum cuts.

3.1.1 The Construction of s in the Half Integral Degree Cut Case

The vector m will consist of the convex combination of two feasible points in the $O(T)$ -Join polyhedron, g and b (where g is for “good” edges and b is for “bad” edges).

For a tree T and an edge $e = (u, v)$ we let:

$$g_e = \begin{cases} \frac{1}{6} & \text{If } u \text{ and } v \text{ are both even in } T \\ \frac{1}{4} & \text{Otherwise} \end{cases}$$

Lemma 3.2. g is in P_{OJ} with probability 1.

Proof. First, consider any cut consisting of a single vertex v (or its complement). If v is odd, we need to ensure that $g(\delta(v)) \geq 1$. If v is odd, then $g_e = 1/4$ for all $e \in \delta(v)$, so this follows from the fact that the support graph is 4-regular.

Now consider any cut S with $2 \leq |S| \leq n - 2$. We now argue that $g(\delta(S)) \geq 1$ with probability 1. This follows from the fact that:

$$g(\delta(S)) \geq \frac{1}{6}|\delta(S)| \geq \frac{1}{6} \cdot 6 = 1,$$

where we use that every cut S with $2 \leq |S| \leq n - 2$ is not minimal so it has at least 6 edges. \square

We now design our second vector b . For a tree T and an edge $e = (u, v)$ we let:

$$b_e = \begin{cases} \frac{1}{2} & \text{If } e \text{ is good} \\ \frac{1}{6} & \text{If } e \text{ is bad} \end{cases}$$

Lemma 3.3. b is in P_{OJ} .

Proof. For any non-vertex cut, similar to above, the $O(T)$ -Join constraint is easily satisfied. For a vertex cut v , by [Lemma 3.1](#) there is at least one good edge adjacent to every vertex. Therefore, $b(v) \geq \frac{1}{2} + \frac{3}{6} = 1$. \square

Definition 3.4 (Matching vector m in the degree cut case). Let $m = \alpha b + (1 - \alpha)g$, for some $0 < \alpha < 1$ we choose in the next subsection. Since b and g are both in the $O(T)$ -Join polyhedron, so is m .

Lemma 3.5. For any good edge e , $\mathbb{E}[g_e] \leq \frac{1}{4} - \frac{p}{12}$.

Proof. Let $p_e = \mathbb{P}_{T \sim \mu}[u, v \text{ even}]$. We can compute:

$$\mathbb{E}[g_e] = \frac{p_e}{6} + \frac{1 - p_e}{4} \leq \frac{p}{6} + \frac{1 - p}{4} = \frac{1}{4} - \frac{p}{12},$$

as desired. \square

Therefore, for any good edge e ,

$$\mathbb{E}[m_e] \leq \frac{1}{2}\alpha + (1 - \alpha) \left(\frac{1}{4} - \frac{p}{12} \right)$$

For any bad edge e , we have

$$\mathbb{E}[m_e] \leq \frac{1}{6}\alpha + \frac{1}{4}(1 - \alpha)$$

To make the two equal, we set $\alpha = \frac{p}{4+p}$. Therefore,

$$\mathbb{E}[m_e] \leq \frac{1}{4} - \frac{p}{48 + 4p}$$

for all edges e , giving a 1.4985 approximation for $p = 1/27$.

3.1.2 Proof of Lemma 3.1

We now prove the lemma from above which claims there is at least one good edge adjacent to every vertex for $p = 1/27$.

To do so we first slightly strengthen Lemma 2.19 in the half integral case.

Lemma 3.6. *Let $S \subseteq V$ with $x(\delta(S)) = 2$ in the half integral case. Then $\mathbb{P}[\delta_T(S) \text{ even}] \geq 13/27$.*

Proof. By Corollary 2.16 $\delta_T(S)$ is a BS(2) for at most four Bernoullis with success probabilities $p_1 \geq p_2 \geq p_3 \geq p_4$. We now apply Hoeffding (Theorem 2.17) to the function indicating if $\delta_T(S)$ is even. There are four possible cases for the worst case setting of the remaining three Bernoullis:

- $p_1 = p_2 = 1$, in which case the probability it is even is 1.
- $p_1 = 1$ and $p_2 = p_3 = 1/2$. In this case the probability it is even is $1/2$.
- $p_1 = 1$ and $p_2 = p_3 = p_4 = 1/3$. In this case, the probability it is even is $4/9 + 1/27 = 13/27$.
- $p_1 = p_2 = p_3 = p_4 = 1/2$. Then the probability it is even is $1/2$.

Note one can also exclude the last case by observing that since S is a cut in the graph, $\delta_T(S) \geq 1$ with probability 1, implying $p_1 = 1$. One then applies Hoeffding to the remaining three Bernoullis and the function indicating their sum is odd. \square

Lemma 3.7. *For every vertex v in the half integral case, there is an edge $e = \{u, v\}$ such that $\mathbb{P}[u, v \text{ even}] \geq 1/27$.*

Proof. Let μ' be the conditional measure where $\delta_T(v) = 1$. Observe that in μ' , first we sample a tree in $G \setminus \{v\}$, and then we *independently* add an edge in $\delta(v)$. As conditioning $\delta_T(v) = 1$ is equivalent to conditioning $\delta_T(v) \leq 1$, by Fact 2.25, $\mathbb{P}_{\mu'}[e \in T] \leq 1/2$ for all edges $e \in \delta(v)$. Thus, for at least one edge $e \in \delta(v)$, we have $1/4 \leq \mathbb{P}_{\mu'}[e \in T] \leq 1/2$. In the special case that $\mathbb{P}[\delta_T(v) = 1] = 0$, we just let e be an arbitrary edge in $\delta(v)$. Let $e = \{u, v\}$.

It remains to prove that $\mathbb{P}_{\mu}[\delta(u), \delta(v) \text{ even}] > 1/27$.

$$\begin{aligned} \mathbb{P}[\delta(u), \delta(v) \text{ even}] &= 1 - \mathbb{P}[\delta(u) \text{ or } \delta(v) \text{ odd}] \\ &= 1 - \mathbb{P}[\delta(u) \text{ odd}] - \mathbb{P}[\delta(v) \text{ odd}] + \mathbb{P}[\delta(u), \delta(v) \text{ odd}] \\ &\geq 13/27 - \mathbb{P}[\delta(v) \text{ odd}] + \mathbb{P}[\delta_T(v) = 1 \wedge \delta(u) \text{ odd}], \end{aligned}$$

by Lemma 3.30. First, note if $\mathbb{P}[\delta_T(v) = 1] = 0$, then we get that $\mathbb{P}[\delta_T(v) = 2] = 1$ (since $\mathbb{P}[\delta_T(v) \geq 1] = 1$, and $\mathbb{E}[\delta_T(v)] = 2$). So, the RHS is $13/27$ and we are done.

Otherwise,

$$\begin{aligned} \mathbb{P}[\delta_T(v) = 1 \wedge \delta_T(u) \text{ odd}] &= \mathbb{P}[\delta_T(u) \text{ odd} \mid \delta_T(v) = 1] \cdot \mathbb{P}[\delta_T(v) = 1] \\ &\geq \frac{1}{4} \cdot \mathbb{P}[\delta_T(v) = 1]. \end{aligned}$$

The inequality is because edge e can make $\delta(u)$ odd by being in/out of the tree and that has probability at least $1/4$. Therefore,

$$\begin{aligned} \mathbb{P}[\delta(u), \delta(v) \text{ even}] &\geq 13/27 - \mathbb{P}[\delta_T(v) \text{ odd}] + \frac{1}{4} \mathbb{P}[\delta_T(v) = 1] \\ &= 13/27 - \frac{3}{4} \mathbb{P}[\delta_T(v) = 1] - \mathbb{P}[\delta_T(v) = 3]. \end{aligned}$$

Finally, by [Fact 2.26](#), the RHS attains its minimum value when $\delta_T(v)$ is sum of 4 Bernoullis with success probabilities $1, 1/3, 1/3, 1/3$. \square

3.1.3 The Construction of s in the General Degree Cut Case

We simply generalize the above to the degree cut case when η is an arbitrary constant and the edges are not necessarily half integral. As before the vector m will consist of the convex combination of two feasible points in the $O(T)$ -Join polyhedron, g and b (where g is for “good” edges and b is for “bad” edges).

Here we will rely on a probabilistic lemma which is slightly more involved to prove. It is an immediate corollary of a statement proved later in this dissertation:

Corollary 3.8 (Corollary of [Theorem 6.21](#)). *Let v be a vertex. Then, if G_v is the set of good edges adjacent to v , $x(G_v) \geq 1$.*

For a tree T and an edge $e = (u, v)$ we let:

$$g_e = \begin{cases} \frac{1}{2+\eta}x_e & \text{If } u \text{ and } v \text{ are both even in } T \\ \frac{1}{2}x_e & \text{Otherwise} \end{cases}$$

Lemma 3.9. g is in P_{OJ} .

Proof. First, consider any cut consisting of a single vertex v (or its complement). If v is odd, we need to ensure that $g(\delta(v)) \geq 1$. If v is odd, then $g_e = x_e/2$ for all $e \in \delta(v)$, so this follows from the fact that $x(\delta(v)) = 2$.

Now consider any cut S with $2 \leq |S| \leq n - 2$. We now argue that $g(\delta(S)) \geq 1$ with probability 1. This follows from the fact that:

$$g(\delta(S)) \geq \frac{1}{2+\eta}x(\delta(S)) \geq \frac{1}{2+\eta}(2+\eta) = 1,$$

where we use that every cut S with $2 \leq |S| \leq n - 2$ has $x(\delta(S)) \geq 2 + \eta$. \square

We now design our second vector b . For a tree T and an edge $e = (u, v)$ we let:

$$b_e = \begin{cases} \frac{1+\eta}{2+\eta}x_e & \text{If } e \text{ is good} \\ \frac{1}{2+\eta}x_e & \text{If } e \text{ is bad} \end{cases}$$

Lemma 3.10. b is in P_{OJ} .

Proof. For any non-vertex cut, similar to above, the $O(T)$ -Join constraint is easily satisfied. For a vertex cut v , by [Corollary 3.8](#) the x weight of the set of good edges adjacent to v is at least 1. Therefore, $b(v) \geq \frac{1+\eta}{2+\eta} + \frac{1}{2+\eta} = 1$. \square

Definition 3.11 (Matching vector m in the degree cut case). *Let $m = \alpha b + (1 - \alpha)g$, for some $0 < \alpha < 1$ we choose in the next subsection. Since b and g are both in the $O(T)$ -Join polyhedron, so is m .*

Lemma 3.12. *For any edge e , $\mathbb{E}[m_e] \leq (\frac{1}{2} - \frac{\eta}{9})x_e$.*

Proof. Let $p_e = \mathbb{P}_{T \sim \mu} [u, v \text{ even}]$. We can compute:

$$\mathbb{E} [g_e] = \left(\frac{p_e}{2+\eta} + \frac{1-p_e}{2} \right) x_e \leq \left(\frac{p}{2+\eta} + \frac{1-p}{2} \right) x_e = \left(\frac{1}{2} - \frac{\eta p}{4+2\eta} \right) x_e,$$

as desired. \square

Therefore, for any good edge e ,

$$\mathbb{E} [m_e] \leq \left(\alpha \left(\frac{1+\eta}{2+\eta} \right) + (1-\alpha) \left(\frac{1}{2} - \frac{\eta p}{4+2\eta} \right) \right) x_e$$

For any bad edge e , we have

$$\mathbb{E} [m_e] \leq \left(\frac{\alpha}{2+\eta} + \frac{1-\alpha}{2} \right) x_e$$

To make the two equal, we set $\alpha = \frac{p}{2+p}$. Therefore,

$$\mathbb{E} [m_e] \leq \left(\frac{p/(2+p)}{2+\eta} + \frac{1-p/(2+p)}{2} \right) x_e < \left(\frac{1}{2} - \frac{p\eta}{9} \right) x_e$$

for all edges e . Since η, p are absolute constants, this is at most $(\frac{1}{2} - \epsilon)x_e$ for some absolute constant $\epsilon > 0$. Therefore in the degree cut case, the expected cost of the max entropy algorithm is at most $(\frac{3}{2} - \epsilon)c(x)$.

3.2 The Half Integral Case

We now build up significantly and obtain a better-than-3/2 approximation for the half integral case with no additional assumptions. Here we assume that there is an optimal LP solution with $x_e \in \{0, \frac{1}{2}, 1\}$ for all edges $e \in E$. In particular, we show:

Theorem 3.13. *Let x be a half-integral solution to (2), i.e. $x_e \in \{0, \frac{1}{2}, 1\}$ for all edges $e \in E$. Then, max entropy algorithm produces a solution of expected cost at most $1.49993 \cdot c(x)$.*

Assumption 3.14. *Throughout the paper, we assume that we are given a feasible half-integral solution of the P_{Sub} , that is, for each $e = \{u, v\}$, $x_e \in \{0, \frac{1}{2}, 1\}$.*

Remark 3.15. We will often talk about the support graph $G = (V, E)$ of x , replacing any edge of value 1 with two parallel edges. Therefore the number of edges crossing any minimum cut is 4 (corresponding to fractional value 2), and the graph is Eulerian. **Henceforth, any reference to the graph G refers to this support graph.**

One of the very nice things about the half integral case is that there are no cuts S with $2 < x(\delta(S)) < 3$, as every cut in this 4-regular support graph has an even number of edges (and at least 4). This allows us to define a completely equivalent algorithm that adds a lot of intuition for the proof. The equivalence is an immediate consequence of [Fact 2.27](#) and [Fact 2.1](#): for every minimum cut S , $E(S)$ is in the spanning tree polytope, thus μ decomposes into $\mu_{E(S)} \times \mu_{G/S}$.

Thus we get the immediate consequence:

Algorithm 2 Algorithm for half-integral TSP

- 1: Given a half-integral solution x of the subtour LP, with an edge e^+ with $x_{e^+} = 1$.
 - 2: Let G be the support graph of x .
 - 3: Set $T = \emptyset$ ▷ T will be a 1-tree
 - 4: **while** there exists a proper tight set of G that is not crossed (by a tight set) **do**
 - 5: Let S be a minimal such set such that $e^+ \notin E(S)$ ▷ Note such a set always exists, as S, \bar{S} are both proper tight sets, so one does not have e^+ . In [Fact 3.17](#) we show that $e^+ \notin \delta(S)$.
 - 6: Compute the maximum entropy distribution μ of $E(S)$
 - 7: Sample a tree from μ and add its edges to T
 - 8: Set $G = G/S$ ▷ Note we never contract e^+ .
 - 9: **end while**
 - 10: Randomly sample a cycle from G (including e^+) and add it to T ▷ In [Fact 3.18](#) we show G itself is a cycle
 - 11: Compute the minimum O-Join on the odd nodes of T . Shortcut it and output the resulting Hamiltonian cycle.
-

Fact 3.16. *Any tree chosen from a max-entropy distribution corresponding to a proper tight set S which is not crossed is independent of all other edges of T that we choose in different iterations of the while loop in our algorithm.*

Our algorithm can be viewed as essentially constructing a cactus representation of the min-cuts. More precisely, the critical cuts of our algorithm (defined below) are in one to one correspondence with the cycles of the cactus.

3.2.1 Critical Sets

In this section we will more fully explore the interaction between our algorithm and the structure of the cactus representation of minimum cuts. From now on, assume G is the 4-regular support graph of the half-integral LP solution x and that we have executed our algorithm on G .

Fact 3.17. *In step 5 of the algorithm, we have $e^+ \notin \delta(S)$.*

Proof. Say $e^+ \in \delta(S)$, and let $e^+ = \{u, v\}$. Then, since $x_{e^+} = 1$, $\{u, v\}$ is a tight set. It also crosses S (as S is a proper set). That is a contradiction. □

Fact 3.18. *In step 10 of the algorithm the remaining graph G is a cycle of length at least 3 such that there are exactly two parallel edges between each pair of consecutive vertices.*

Proof. Let G be the graph which remains after the while loop in the algorithm terminates. By the algorithm, $e^+ = \{u, v\}$ is not contracted yet. G has at least 3 vertices, as otherwise in the last of the while we contracted a set S where $e^+ \in \delta(S)$ which contradicts [Fact 3.17](#). If G has 3 vertices then it must be a cycle. Otherwise, $\{u, v\}$ is a proper tight set in G , and it must be crossed. In this case by [Fact 17.4](#) G is a cycle of length at least 4. □

A tight set S selected in step 5 of the algorithm is called a **critical set** and the corresponding cut $\delta(S) := E(S, \bar{S})$ is called a **critical cut**. Vertices of G are degenerate critical sets.

There is a natural *hierarchy of critical sets* associated with the execution of the algorithm. The leaves of the hierarchy are vertices of the original graph. If S and S' are critical sets such that S or a contracted version of S is a vertex in S' , then S is a child of S' (respectively S' is the parent of S). If \tilde{S} is an ancestor of S in the hierarchy of critical sets, then we say that \tilde{S} is a **higher** critical set than S (resp. S is a **lower** critical set than \tilde{S}). For example, in Fig. 21, critical set F is the parent of and is higher in the hierarchy than critical sets A, B and C .

The root of the hierarchy is the graph G once we get to step 10 of the algorithm.

Definition 3.19 (Going higher). *An edge e in $\delta(S)$ goes higher if the lowest critical set S' such that $S \subsetneq S'$ satisfies $e \in \delta(S')$.*

Note that by Fact 3.16 any edge going higher is independent of all edges which do not.

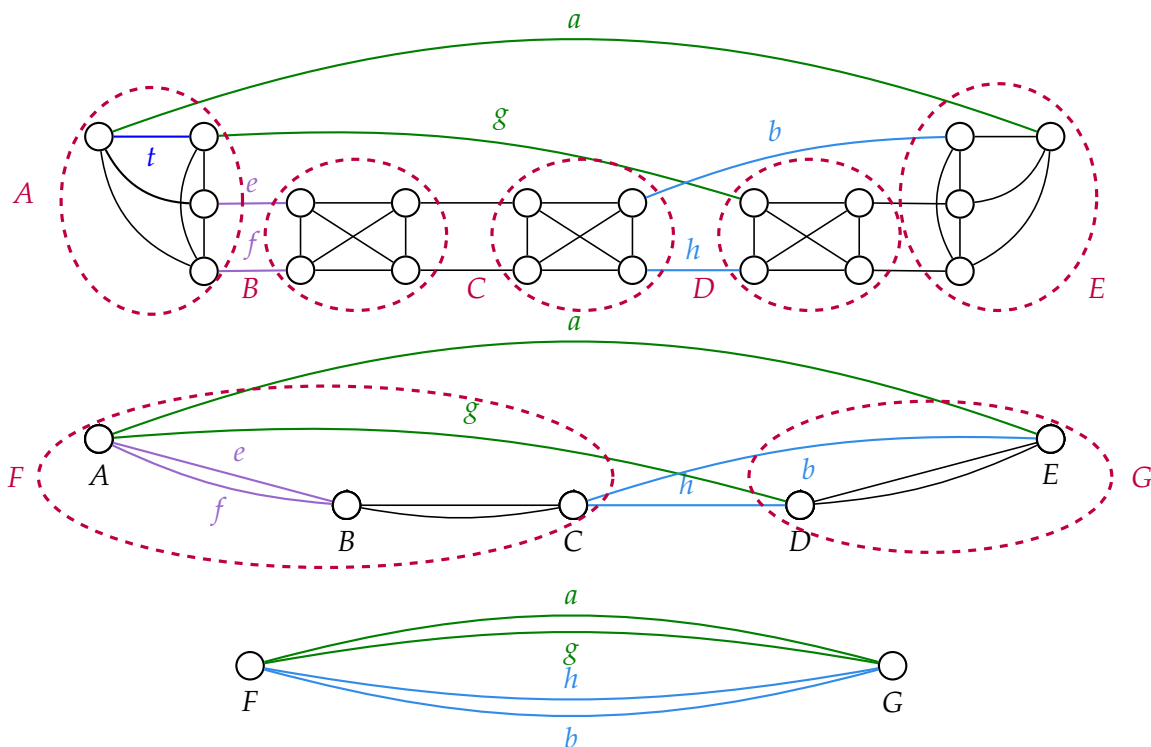


Figure 20: Example execution on a half integral graph. In the first figure, we visualize five tree operations in parallel, which we may do since all these tight sets have size 4 and are not crossed by other tight sets. Similarly in the second figure we do two operations in parallel. In the final step, a cycle is chosen by picking two edges at random. A, B, C, D, E are all “degree cuts” whereas F and G are both “cycle cuts.” t is an example top edge (as are all edges picked in the first graph). a, g and b, h are cycle partners with respect to the cut F . e, f are companions.

Structure of critical cuts: Consider a critical set S chosen in step 5 in the algorithm. We will abuse notation and, at any time during the execution of the algorithm, refer to G with vertex set V as the graph remaining at that time, after contraction of all trees that have been sampled before S

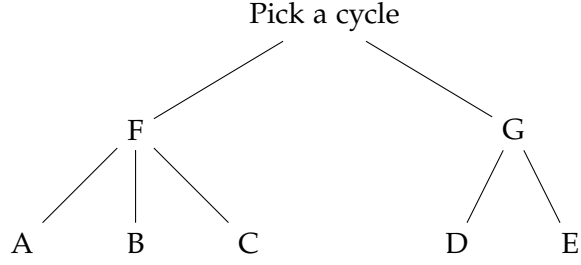


Figure 21: An execution tree on this graph.

is considered. Consider the graph $G' := G/V \setminus S$ and let w be the contracted vertex representing $V \setminus S$. There are two possibilities for the structure of G' :

- Case 1: There are no proper min-cuts inside S . In this case, we call $\delta(S)$ a **degree cut**. In Fig. 20, A, B, C, D, E are all degree cuts.¹¹
- Case 2: There is a proper min-cut (S_0, \bar{S}_0) such that $S_0 \subsetneq S$. In this case, it (and every other proper min-cut inside S) is crossed by some other min-cut (or would be more minimal than S).

It follows that in G' , every proper mincut is crossed by some other proper mincut and therefore, by Fact 17.4, the graph is a cycle with two edges between each pair of adjacent vertices in the cycle. In this case, we call $\delta(S)$ a *cycle cut*. For example, F and G in Fig. 20 are cycle cuts.

We divide the 4 edges from w into two pairs, such that each pair share an endpoint inside S . We call each such pair *cycle partners* with respect to $\delta(S)$. Every other pair of edges between two adjacent vertices in the cycle are called *companions*.

For example, in Fig. 20, $\delta(F)$ is a cycle cut and a, g and b, h are cycle partners with respect to $\delta(F)$. e and f are companions.

Cycle cuts correspond to cycles of length 3 or more in the cactus.

Note that every edge has at most one companion but possibly many partners depending on the underlying cactus.

Definition 3.20 (Highest critical cuts). *For a vertex u and an edge $e = \{u, v\}$, let $S_{u,e}$ be the highest critical set S such that $u \in S$ and $v \notin S$, and let S_e be the lowest critical set such that both $S_{u,e}$ and $S_{v,e}$ are (contracted) nodes in S_e . Then $\delta(S_{u,e})$ and $\delta(S_{v,e})$ are the highest critical cuts containing e . If the edge e is clear from context, we may drop e in the notation $S_{u,e}$.*

Definition 3.21 (Bottom Edge and Top Edges). *For an edge e , if S_e is a cycle cut, we say that e is a bottom edge and otherwise it is a top edge.*

For example, in Fig. 20, e, f, a, g, b, h are bottom edges (among the labeled edges) and t is a top edge. The following fact is immediate:

Fact 3.22. *Companion bottom edges e, f are in or out of T independently of every other edge of T .*

¹¹These cuts correspond to cycles of length two in the cactus.

Min-cuts containing a particular edge: The set of min-cuts an edge $e = (u, v)$ is on are the following:

- (a) all critical degree cuts $\delta(S)$ such that $e \in \delta(S)$. (This includes the cuts $(u, V \setminus u)$ and $(v, V \setminus v)$.)
- (b) For any set S such that $\delta(S)$ is a critical cycle cut, and e is either in S or on $\delta(S)$, every cut of the cycle that includes the edge e is a min-cut e is on.

It is easy to see that each of the above is a min-cut. To see that there are no others, it suffices to observe by induction that whenever a set S is contracted, we have accounted for all min-cuts in which nodes inside S are partitioned between the two sides of the cut.

Other facts: We end this part by recording the following basic facts about structure of min cuts, and we will use them throughout our proofs.

Fact 3.23. *Suppose that S is a critical set. If some (contracted) vertex $v \in S$ has two edges to $w := V \setminus S$, then S is a cycle cut.*

Proof. This is immediate if $|S| = 2$, so suppose that $|S| > 2$. Then v has two edges to w , which has two edges to $S \setminus v$ which has two edges to w . Since $S \setminus v$ is therefore a proper min-cut but was not selected in step 5, it must be crossed by some other set, which, by the earlier discussion of the structure of critical cuts, means that $\delta(S)$ is a cycle cut. \square

Fact 3.24. *Suppose that S and S' are two distinct tight sets. Then $|\delta(S) \cap \delta(S')| \leq 2$.*

Proof. By contradiction. Suppose that S and S' are both proper min-cuts and have $\delta(S) \cap \delta(S') \geq 3$. Then by Fact 17.3, they do not cross. Therefore it must be that, say, $S \subset S'$. But in this case, if $\delta(S) \cap \delta(S') \geq 3$, since $\delta(S)$ and $\delta(S')$ are both min-cuts, there can only be one edge from S to $S' \setminus S$ and at most one edge from $S' \setminus S$ to $V \setminus (S \cup S')$ which contradicts $\delta(S' \setminus S) \geq 4$. \square

Fact 3.25. *Suppose that S and S' are two critical sets such that $S \subset S'$. Then if $\delta(S) \cap \delta(S') = 2$, then S' is a cycle cut.*

Proof. Once S is contracted, it has two edges to $V \setminus S'$, and therefore by Fact 3.23 is a cycle cut. \square

Fact 3.26. *Suppose that $S \subset S'$ are two critical cycle cuts. Then any two edges are cycle partners on at most one of these (cycle) cuts.*

Proof. Suppose not. Then there is a pair of edges e and f that are cycle partners on both. Suppose that g and h are the other pair of cycle partners on $\delta(S)$ and that their endpoint inside S is node u . Then $(S' \setminus S) \cup u$ is a min-cut that crosses S , which is a contradiction to the selection of S . [Essentially this means that in fact there is a larger cycle here.] \square

Fact 3.27. *Say S is a critical set and exactly two edges of $\delta(S)$ are bottom edges that do not go higher. Then the other two edges of $\delta(S)$ must go higher.*

Proof. Say $\delta(S) = \{a, b, c, d\}$ and suppose a, b are bottom edges that do not go higher. Say S' is the parent of S in the hierarchy of critical cuts. This implies that $\delta(S')$ is a cycle cut. So, a, b are companions in this cycle. This implies that either c, d are also companions or they are cycle partners in $\delta(S')$. \square

Lemma 3.28. Let $S \subseteq E$ with $|S| = 3$. Furthermore, assume that $\mathbb{P}[\delta_T(S) \geq 1] = 1$. Then, $\mathbb{P}[\delta_T(S) = 1] \geq \frac{1}{2}$ and $\mathbb{P}[\delta_T(S) = 2] \geq \frac{3}{8}$.

Proof. By [Corollary 2.18](#), we can write the rank sequence of $\delta_T(S)$ as a sum of 3 independent Bernoullis B_1, B_2, B_3 , and since $\mathbb{P}[\delta_T(S) \geq 1] = 1$ we know that for one Bernoulli $p = 1$. Without loss of generality let $p_1 = 1$. Then by [Theorem 2.17](#) we know that $\mathbb{P}[\delta_T(S) = 1]$ and $\mathbb{P}[\delta_T(S) = 2]$ are minimized when $p_2 = p_3 = \frac{1}{4}$ or $p_2 = \frac{1}{2}$ and $p_3 = 0$. Therefore:

$$\mathbb{P}[\delta_T(S) = 1] \geq \min \left\{ \left(\frac{3}{4} \right)^2, \frac{1}{2} \right\} = \frac{1}{2}$$

$$\mathbb{P}[|S \cap T| = 2] \geq \min \left\{ 2 \left(\frac{1}{4} \right) \left(\frac{3}{4} \right), \frac{1}{2} \right\} = \frac{3}{8}$$

□

The following two lemmas are proved using a similar analysis.

Lemma 3.29. Let $S \subseteq E$ with $|S| = 2$. Let $\frac{1}{2} \leq \mathbb{E}[|S \cap T|] \leq \frac{3}{2}$. Then $\mathbb{P}[\delta_T(S) = 1] \geq \frac{3}{8}$.

This lemma was proved in the previous subsection (see [Lemma 3.30](#), but we repeat it here.

Lemma 3.30. For a min-cut S , $\mathbb{P}[\delta_T(S) \text{ even}] \geq 13/27$.

Lemma 3.31. Let $S_1, S_2 \subseteq E$ with $|S_1 \cap S_2| = \emptyset$. Let $|S_1| = |S_2| = 2$, or equivalently $\mathbb{E}[\delta_T(S_1)] = \mathbb{E}[\delta_T(S_2)] = 1$. Then $\mathbb{P}[\delta_T(S_1) = 1 \wedge \delta_T(S_2) = 1] \geq \frac{3}{16}$.

Proof. Let $S_1 = \{e, f\}$. Then condition on $e \in T$: this occurs with probability $\frac{1}{2}$. By [Fact 2.25](#) we have

$$\mathbb{E}[|f \cap T| \mid e \in T] \leq \frac{1}{2}$$

Then condition on $f \notin T$. Given the above, this happens with probability at least $\frac{1}{2}$. Similarly consider the event $e \notin T$ and $f \in T$. One of these occurs with probability $\frac{1}{2}$. Therefore, in either event we have:

$$\frac{1}{2} \leq \mathbb{E}[|S_2 \cap T|] \leq \frac{3}{2}$$

And now by [Lemma 3.29](#) both events occur simultaneously with probability at least $\frac{3}{16}$. □

3.3 Overview of Analysis

As already mentioned, our algorithm consists of two steps: sample a 1-tree T , and then construct an optimal O -join for the odd degree vertices in the 1-tree.

Given a feasible LP solution x , the choice $y_e = x_e/2$ for each edge $e \in E$ (which gives $y_e := 1/4$ in the half integral case), yields an O -join solution of total cost at most $OPT/2$. However, this is essentially Christofides' algorithm and guarantees only a $3/2$ approximation.

The key to improving on this is the observation that constraint (4) in the O -join LP is not binding if the intersection of the cut $\delta(S)$ with the tree is even.

Definition 3.32 (Even cuts). A cut $\delta(S)$ is *even* in T (or simply “even” when T is understood) if $\delta_T(S)$ is even.

Thus, for every edge e with the property that every min-cut that e is on is even, we can reduce y_e to $1/6$, since every non-min-cut has at least 6 edges, and therefore this guarantees that constraint (4) remains satisfied everywhere. This is the gist of the approach taken in [OSS11].

Suppose that there are multiple “good” edges e with the property that every min-cut they are on is even, say with probability at least p (over the randomness in the selection of T). Then for those outcomes T in which e has this property, we could set $y_e := 1/6$ and satisfy the O -join constraints. This would save us $\frac{1}{12}c(e)$ on every such edge e (the reduction from $1/4$ to $1/6$) and thereby guarantee a reduction in the cost of the O -join solution of $\sum_e \text{“good”} \frac{p}{12}c(e)$.

Unfortunately, in general, it is not possible to argue that every min-cut an edge is on is even simultaneously in T with constant probability. So, we will use a careful charging scheme.

Definition 3.33 (Last Cuts). *For an edge e , the last cuts of e are the only (two) min-cuts containing e and edges going higher in the graph right before contracting S_e .*

Observe that the last cuts of a top edge are critical cuts, but the last cuts of bottom edges are not critical.

Definition 3.34 (Even at Last). *For an edge e we say e is even at last if the two last cuts of e are even.*

Equivalently, if e is a bottom edge, we say e is even at last if all the min cuts containing e on the cycle defined by the graph consisting of S_e with $V \setminus S_e$ contracted are even. Otherwise, if $e = \{u, v\}$ is a top edge, then it is even at last if the critical cuts S_u, S_v are even simultaneously.

Fact 3.35. *If a bottom edge e is even at last, then all (bottom) edges f where $S_f = S_e$ are even at last.*

Proof. Since $\delta(S_e)$ is a cycle cut, the edges inside S_e form a path, and thus, exactly one edge between each pair of (possibly contracted) vertices inside S_e is selected as part of the tree on S_e chosen in step 6 of the algorithm. If e is even at last, we must have exactly one of each pair of cycle partners on $\delta(S)$ is in T ; therefore, every pair of adjacent nodes in the cycle have one edge connecting them in the tree. So, all cuts on the cycle have exactly two edges in T . This implies every bottom edge of this cycle is even at last. \square

Remark 3.36. By [Fact 17.4](#), the companion of every bottom edge e has exactly the same pair of last cuts as e .

Definition 3.37 (Good edges). *An edge $e = (u, v)$ is **good** if it is even at last with probability at least p for some constant $p > 0$.*

Instead of proving that all min-cuts that a single edge is on are even, we will instead prove that every minimum cut contains at least one good edge. Each good edge e will then be responsible for its last two cuts. This will allow edges to be reduced when they are even at last, as all cuts lower in the hierarchy are handled by other edges.

Theorem 3.38. *There is a universal constant $p \geq 1/27$ such that every every min-cut has at least one good edge.*

We already proved this theorem for vertices in [Lemma 3.1](#). Proving it for min-cuts is identical, thus we omit the proof here.

As we hinted at, we will reduce the value of y_e to $1/6$ whenever an edge e is even at last. However, since e may also be on many other lower min-cuts, if we reduce y_e , the solution may

not be feasible ((4) may be violated) as the lower min-cuts may be odd. To handle any lower min-cut C that e is on, we show that, conditioned on e being even at last, the probability that C is also even is at least q for some $q \geq \Omega(1)$. Therefore, we only need to worry about the lower cuts with probability $1 - q$ each. In the bad event that a lower cut C is odd, we will need to fix the solution to guarantee that (4) still holds: our approach is to split the deficit introduced in the O -Join constraint for C among the good edges that do not go higher (see Definition 3.19). We then simply show that in expectation each edge gains. This part of the proof heavily exploits the properties of cactus representation of the min-cuts that we discussed above.

We note that Theorem 3.38 on its own is not enough to run our charging argument; so, we need a slightly stronger version. In particular, in some cuts we may need to have two or three good edges.

3.4 Probabilistic lemmas

In this section, we present three probabilistic lemmas which show that in every min-cut there is at least one good edge, (and in some there are even more). This immediately proves Theorem 3.38.

Note the last cycle that we choose in step 10 of the algorithm has all edges even at last so we don't need to address it in this section. Furthermore, by Fact 3.17, e^+ does not belong to any critical cut.

Lemma 3.39 (Bottom edge lemma). *Suppose that $e = (u, v)$ is a bottom edge. Then e is good (where $p \geq 3/16$).*

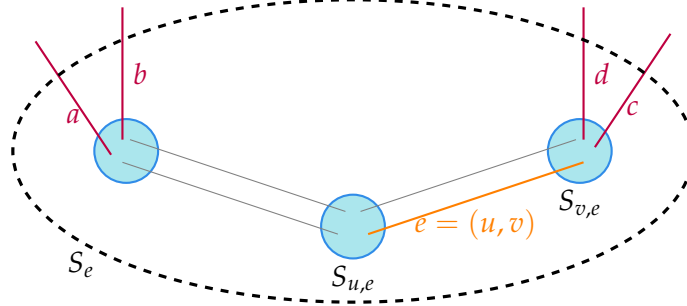


Figure 22: Illustration of edges in Bottom Edge Lemma.

Proof. If e is a bottom edge then $S_e (= S_{(u,v)})$ is a cycle cut. By construction, when a tree on S_e is selected in step 8 of the algorithm, exactly one edge is chosen between every pair of adjacent nodes in $E(S_e)$. So it suffices to consider the edges in $\delta(S_e)$. These divide up into two pairs of cycle partners connecting S_e to $V \setminus S_e$, say $\{a, b\}$ and $\{c, d\}$. (See Fig. 22.) Then by Lemma 3.31, setting $S_1 := \{a, b\}$ and $S_2 := \{c, d\}$, we have $\mathbb{P}[\delta_T(S_1) = 1 \text{ and } \delta_T(S_2) = 1] \geq 3/16$. \square

Lemma 3.40 (Top edge lemma). *In a critical cut $\delta(S)$ with one edge that goes higher, of the remaining three edges in the cut, at least two are good with $p \geq \frac{1}{16}$.*

Proof. First, suppose that e is the edge that goes higher from $\delta(S)$, and f, g and h are the other edges in $\delta(S)$.

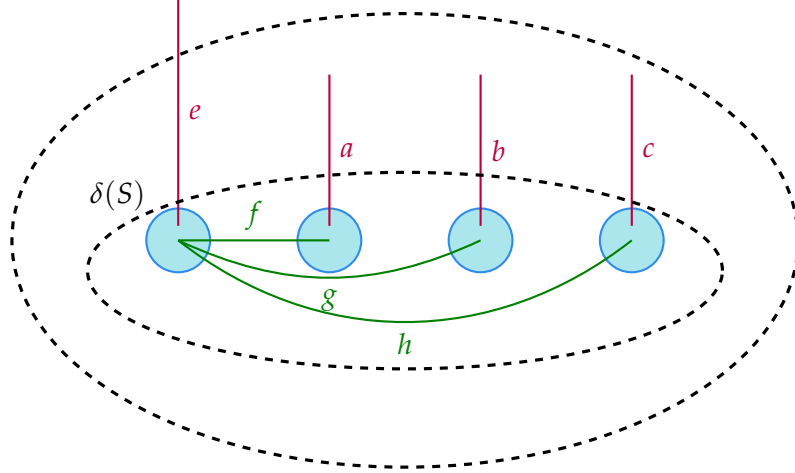


Figure 23: Illustration of top lemma. The figure shows the case where all three of S_f , S_g and S_h have an edge that goes higher (a, b, c respectively). It may be that some number of these nodes do not have an edge going higher.

If the other endpoint of one of these three edges, say S_h , has no edge that goes higher then h is good. (See left side of Fig. 23.) To see this, observe that we can condition on $\delta(S_h)$ being even which by Lemma 3.30 has probability at least $13/27$. Given this event, $|\{f, g, h\} \cap T|$ is either even or odd. In either case, the event $e \in T$ is an independent event that occurs with probability $1/2$. Therefore,

$$\mathbb{P}[|\delta(S_h) \cap T| \text{ even}] \cdot \mathbb{P}[e \text{ makes } \delta(S) \text{ even} \mid |\delta(S_h) \cap T| \text{ even}] \geq \frac{13}{27} \cdot \frac{1}{2}.$$

Therefore, at least two of f, g and h are good if at least two of S_f, S_g and S_h do not have an edge that goes higher.

Consider next the case that, say, S_f and S_g both have an edge that goes higher, but S_h doesn't. As before, h is good. We claim that one of f and g is also good. To see this, since $e \in T$ is independent of $f, g, h \in T$, we observe using Lemma 3.28 that

$$\mathbb{P}[e \in T \text{ and } |\{f, g, h\} \cap T| = 1] \geq \frac{1}{2} \cdot \frac{1}{2}.$$

Next, note that for one of a and b , say a ,

$$\mathbb{P}[a \in T \mid e \in T] \geq \frac{1}{4}.$$

Therefore, as before, regardless of the even/odd status of $X \cup f$ after conditioning e in and f, g, h to 1, the cut can be fixed by a , meaning we have $p \geq \frac{1}{16}$.

Finally, if all three of S_f, S_g and S_h have an edge that goes higher, then again, condition e in and $|\{f, g, h\} \cap T| = 1$. Then Among a, b and c , for two of them, say a and b , their probability of being in T given that $e \in T$ is at least $1/4$. Therefore, each can fix their corresponding cut ($\delta(S_f)$ for a and $\delta(S_g)$ for b) and both f and g are good. \square

As mentioned, we proved this lemma for vertices in [Lemma 3.1](#), and it can be easily extended for any critical cut S and the fact that every bottom edge is good.

Lemma 3.41. *For every critical cut S , there is an edge $e \in \delta(S)$ such that $\mathbb{P}[e \text{ even at last}] \geq 1/27$.*

3.5 Proof of Main Theorem

Recall that every good edge is even on top with probability at least p . The following statement is the main technical result of this section.

Lemma 3.42. *There is a (random) feasible O -join solution such that for every good edge e ,*

$$\mathbb{E}[y_e] \leq 1/4 - p/240,$$

and for every bad edge $y_e = 1/4$ with probability 1.

Before, proving the above statement we use it to prove [Theorem 3.13](#).

Proof of Theorem 3.13. Consider the trivial O -join solution y' where $y'_e = 1/2$ if e is good and $y'_e = 1/6$ otherwise. Note that this is a valid O -join by [Theorem 3.38](#). Now, define $z = \alpha y + (1 - \alpha)y'$ for some α that we choose later. It follows that for any good edge e ,

$$\mathbb{E}[z_e] \leq \alpha \left(\frac{1}{4} - \frac{p}{240} \right) + (1 - \alpha) \frac{1}{2},$$

and for a bad edge f :

$$\mathbb{E}[z_f] = \alpha \frac{1}{4} + (1 - \alpha) \frac{1}{6}$$

So, for $p = \frac{1}{27}$ and $\alpha = \frac{2160}{2161}$ we obtain $\mathbb{E}[z_e] \leq 0.249962$. Since any edge e is chosen in T with probability $1/2$ (up to a 2^{-n} error), we pay at most $1/2 + \mathbb{E}[z_e]$ for any edge e whereas x pays $1/2$. Therefore, we get a $0.749962/0.5$ approximation algorithm. \square

So, in the rest of this section we prove [Lemma 3.42](#).

O -join construction for good edges: For each good edge e , define B_e to be an independent Bernoulli random variable which is 1 with probability p/p_e , where p is the *lower bound* on the probability that *any* good edge is even on top, and p_e is the *actual* probability that e is even on top. If e, f are bottom edge companions, then we let $B_f = B_e$ (with probability 1). Note that this still makes selection of e, f independent of B_e and any other edge of the graph.

We then construct an O -join solution for each 1-tree T using the following three step process:

1. Initialize $y_e := 1/4$ for each edge $e \in E$.
2. Next, if e is even at last in T and $B_e = 1$, reduce y_e by $r_e(T)$ where:

$$r_e(T) := \begin{cases} \beta & \text{if } e \text{ is a bottom edge.} \\ \tau_2 & \text{if } e = \{u, v\} \text{ is a good top edge and there are exactly 2 good top edges} \\ & \text{in both } \delta(S_u) \text{ and } \delta(S_v) \text{ that do not go higher.} \\ \tau_3 & \text{if } e \text{ is a good top edge that does not meet the previous criteria.} \end{cases}$$

β, τ_2, τ_3 are parameters we will set later. For now, we just assume $\tau_3 \leq \tau_2 \leq \beta \leq 1/12$. When $r_e(T) > 0$, we say that e is **reduced**.

3. On each cut C that is odd, let $\Delta(C) := \sum_{e \in C} r_e(T)$ be the amount by which edges on that cut were reduced in step 2, and let G_C be the set of good edges on C such that C is one of their last cuts. Now, for an edge $e \in G_C$, let C' (and C) be the last cuts of e . Then increase y_e by $\max \left\{ \frac{\Delta(C)}{|G_C|}, \frac{\Delta(C')}{|G_{C'}|} \right\}$. Notice that in this case, since C is one of e 's top cuts, e is not even on top in T and therefore is not reduced in step 2.

By construction, this is a valid O -join solution since on every min-cut we began with at least 4 edges crossing every cut with $y_e = 1/4$ and then guaranteed that every reduction on an odd cut in step 2 was compensated for by a matching increase on that cut in step 3. (The ultimate gain of course will come from the fact that many cuts will be even and hence, there will not need to be an increase.) All non-min-cuts have at least 6 edges on them, each of value at least $1/6$ (after reduction) and are therefore satisfied in (4).

In the rest of the proof, it is enough to show that for any good edge e , $\mathbb{E}[y_e] \leq 1/4 - p/240$. We complete the proof using the following two lemmas.

Lemma 3.43. *If $\beta \geq 5\tau_2/4$, then for any good top edge $e = \{u, v\}$,*

$$\mathbb{E}[y_e] \leq \frac{1}{4} - p \min\{\tau_2 - \beta/2, \tau_3 - 5\beta/12\}.$$

Proof. Let $C := \delta(S_u)$ and $C' := \delta(S_v)$. Recall that, since e is a top edge, C and C' are critical cuts. Let H_C (resp. $H_{C'}$) be the good edges in C (resp. C') that go higher. Note that $|H_C|$ and $|H_{C'}|$ is either 0 or 1 by [Fact 3.25](#). We consider 3 cases:

Case (i): $|G_C| = |G_{C'}| = 2$. In the worst case, there is an edge, say $f \in H_C$ and an edge $g \in H_{C'}$. If f is a bottom edge, then

$$\mathbb{E}[r_f(T)] = \beta \cdot p \quad \text{and} \quad \mathbb{P}[C \text{ is odd} | f \text{ reduced}] = 1/2$$

since

$$\mathbb{P}[\text{parity of } |f \cap T| = \text{parity of } |(\delta(S_u) \setminus f) \cap T|] = 1/2,$$

by [Fact 3.22](#). Therefore, the expected reduction in step 2 on C is at most

$$\mathbb{E}[r_f(T)] \cdot \frac{1}{|G_C|} \cdot \mathbb{P}[C \text{ is odd} | f \text{ reduced}] = \beta p \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{\beta \cdot p}{4}.$$

On the other hand, if f is a top edge, then the expected reduction in step 2 on C is at most

$$\mathbb{E}[r_f(T)] \cdot \frac{1}{|G_C|} \cdot \mathbb{P}[C \text{ is odd} | f \text{ reduced}] \leq \tau_2 p \cdot \frac{1}{2} \cdot \frac{5}{8} \leq \frac{\beta p}{4},$$

where we use [Lemma 3.28](#) and the fact that edge f is independent of the rest of edges in C to infer that $\mathbb{P}[C \text{ is odd} | f \text{ reduced}] \leq 5/8$. The same reasoning applies to g , so we get

$$\mathbb{E}[y_e] \leq \frac{1}{4} - \tau_2 p + 2 \cdot \frac{\beta p}{4}.$$

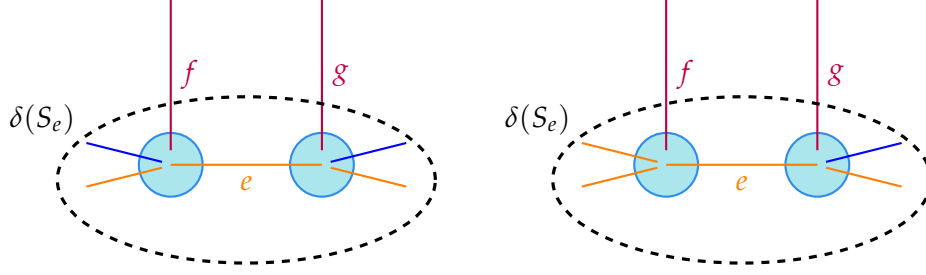


Figure 24: Cases (i) is on the left and case (ii) is on the right. Orange edges are good. On the left, y_e is reduced by τ_2 with probability p and pays half the burden on both sides; on the right y_e is reduced by τ_3 with probability p and pays half the burden on one side and one third of the burden on the other.

Case (ii): $|G_C| \geq 3$ or $|G_{C'}| \geq 3$ (or both). Again, in the worst case, there is an edge, say $f \in H_C$ and an edge $g \in H_{C'}$. By the same reasoning as above, $\mathbb{E}[y_e]$ is largest if f and g are bottom edges. In this case, the same calculation as above gives,

$$\mathbb{E}[r_f(T)] \cdot \frac{1}{|G_C|} \cdot \mathbb{P}[C \text{ is odd} | f \text{ reduced}] \leq \beta p \cdot \frac{1}{|G_C|} \cdot \frac{1}{2},$$

and similarly for g , so

$$\mathbb{E}[y_e] \leq \frac{1}{4} - \tau_3 p + \frac{\beta p}{2} \left(\frac{1}{2} + \frac{1}{3} \right).$$

Case (iii): $|H_C| + |H_{C'}| \leq 1$. In the worst case, $|H_C| = 1$ and $|H_{C'}| = 0$ and $f \in H_C$ is a bottom edge. Note that in this case we may have $G_{C'} = \{e\}$. But the advantage is that we never increase y_e to fix C' since $H_{C'} = \emptyset$. Then,

$$\mathbb{E}[y_e] = \frac{1}{4} - \tau_3 p + \beta p \cdot \frac{1}{|G_C|} \cdot \frac{1}{2} \leq 1/4 - \tau_3 p + \frac{\beta p}{4}.$$

Note that if case (iii) does not happen, then by [Lemma 3.40](#) we have $|G_C|, |G_{C'}| \geq 2$. So, either (i) or (ii) will happen. \square

Lemma 3.44. *If $3\tau_3 \leq 2\tau_2$, then for any (good) bottom edge e ,*

$$\mathbb{E}[y_e] \leq 1/4 - p \min\{\beta/4, 3\beta/4 - \tau_2, \beta - 4\tau_2/3\}.$$

Proof. Say f is the companion of e . Let $S = S_e$ and S' be the parent of S in the hierarchy of critical cuts. Say the last cuts of e (and f) are $C = \{e, f, a, b\}$ and $C' = \{e, f, g, h\}$. In other words a, b are partners and g, h are partners. Note that $|G_C| = |G_{C'}| = 2$ because all edges $\{a, b, g, h\}$ go to the higher critical cut S_e .

Case (i): a and g go higher than S . We have $a, g \in \delta(S')$. So, by [Fact 3.25](#), S' is also a cycle cut. This means that b and h are companions and a and g are cycle partner pairs on $\delta(S')$. (See [Fig. 25](#)). Edge e has to increase to fix the cuts C, C' whenever a, b, g or h are decreased and the corresponding cut is odd. The expected increase in y_e due to reductions on a, b, g, h divides into two types.

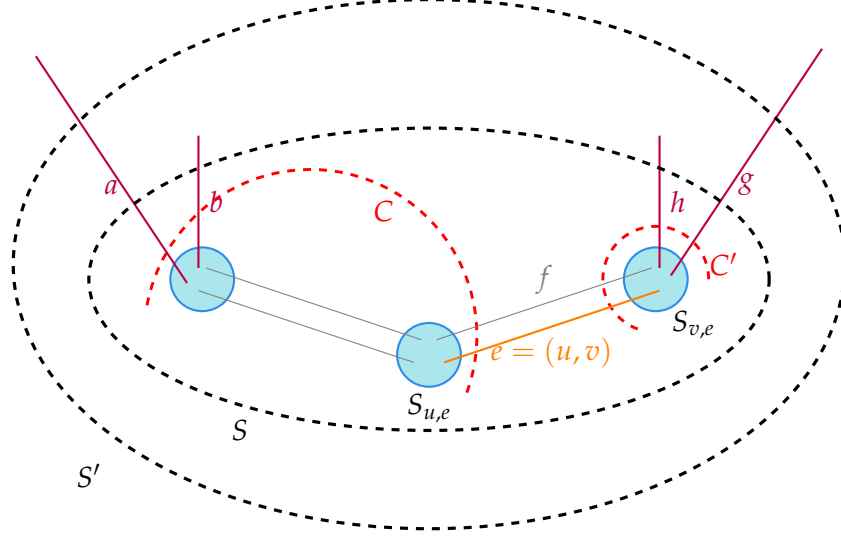


Figure 25: Case (i) of Lemma 3.44.

We start with b, h : By Fact 3.22 and the fact that $B_b = B_h$, we know that b and h are always reduced at the same time. Furthermore, conditioned on b (and h) being reduced, we have

$$\text{parity of } |T \cap C| = \text{parity of } |T \cap C'|.$$

This is simply because b (and h) are reduced only when they are even at last which implies $|T \cap \{a, g\}| = 1$. So, we can fix the reduction of b, h simultaneously when we increase e (or f). In other words, it is enough to only take into account the expected increase of y_e due to b , i.e.,

$$\mathbb{E}[r_b(T)] \cdot \frac{1}{|G_C|} \cdot \mathbb{P}[|C \cap T| \text{ is odd} | b \text{ reduced}] = \beta p \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{\beta p}{4}.$$

Now, we calculate the expected increase due to a, g : We compute the charge due to a and the same will hold for g . Again, by Fact 3.22, b and h are chosen independently of a, g , i.e., $\mathbb{P}[|T \cap C| \text{ odd} | a \text{ reduced}] \leq 1/2$. Therefore, the expected increase due to a is

$$\mathbb{E}[r_a(T)] \cdot \frac{1}{|G_C|} \cdot \mathbb{P}[|C \cap T| \text{ odd} | a \text{ reduced}] \leq \beta p \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{\beta p}{4}$$

using $\tau_3, \tau_2 \leq \beta$. Therefore, altogether,

$$\mathbb{E}[y_e] \leq \frac{1}{4} - \beta p + 3 \cdot \frac{\beta p}{4}.$$

Case (ii): Only one edge, say a , goes higher than S . In this case, by the same reasoning as above, we have

$$\mathbb{E}[r_a(T)] \cdot \frac{1}{|G_C|} \cdot \mathbb{P}[|C \cap T| \text{ is odd} | a \text{ reduced}] \leq \beta p \cdot \frac{1}{2} \cdot \frac{1}{2},$$

since b is independent of a and it can be chosen to correct the parity.

For the remaining three edges $\{b, g, h\}$, by [Lemma 3.40](#) either two or three of b, g and h are good. If two are good, then each has an expected reduction of at most $\tau_2 p$ or three are good and each has an expected reduction of at most $\tau_3 p$.

Therefore, altogether,

$$\mathbb{E}[y_e] \leq \frac{1}{4} - \beta p + \frac{\beta p}{4} + \frac{p}{2} \cdot \max\{2\tau_2, 3\tau_3\} \leq \frac{1}{4} - \frac{3\beta p}{4} + p\tau_2,$$

by the assumption of the lemma.

Case (iii): a, g are companions and b, h are companions (on the next critical cut). This case follows the same analysis as case (i) but gains because in this case a, g are also reduced simultaneously.

Case (iv): No edge goes higher than S and all $\{a, b, g, h\}$ are top edges. Some number of these edges are good; if more than two are good we pay $4\tau_3$ (at most) with probability p and otherwise we pay $2\tau_2$ (at most) with probability p . Then:

$$\mathbb{E}[y_e] \leq \frac{1}{4} - \beta p + \frac{p}{|G_C|} \max\{2\tau_2, 4\tau_3\} \leq \frac{1}{4} - \beta p + \frac{4\tau_2 p}{3}.$$

To finish the proof we just need to argue that we exhausted all cases. By [Fact 3.24](#), among $\{a, b, g, h\}$ at most two go higher. By [Fact 3.26](#), from each pair of cycle partners, i.e., $\{a, b\}$ or $\{g, h\}$, at most one goes higher. Therefore, if case (i) does not happen, we have at most one that goes higher. If (i), (ii) do not happen, then no edge goes higher. So, by [Fact 3.27](#) either all four edges in $\{a, b, g, h\}$ are bottom edges, i.e., case (iii), or none are bottom edges, i.e., case (iv). \square

To finish the proof of [Lemma 3.42](#), let $\beta = 1/12$, $\tau_2 = 7/120$ and $\tau_3 = 7/180$ chosen to satisfy $\tau_3 \leq \tau_2 \leq \beta$, $\beta \geq 5\tau_2/4$ and $3\tau_3 \leq 2\tau_2$. Plugging in these numbers into [Lemma 3.43](#) and [Lemma 3.44](#) we obtain that $\mathbb{E}[y_e] \leq 1/4 - p/240$ for any good edge e as desired.

3.6 Summary of special cases

This completes the proof for the half integral case. The key technical ideas to take away from the degree cut case and the half integral case are as follows:

- The edges that end in different levels of the hierarchy of minimum cuts *are independent*.
- Many edges in every critical cut are *good*, i.e. their top two cuts are even simultaneously with constant probability.
- We can define a *matching* between edges that end at level i and edges that end at level higher than i in the hierarchy.
- Bottom edges are difficult to define this matching for since there are fewer edges ending at level i per edge going higher. However, they are very well behaved: in the half integral case, one of each pair of cycle partners is chosen *independently of all other events*.

While all of these notions may appear specialized to the half integral case, they are in fact all robust enough to be generalized and each one appears in some form in what follows. For this reason, we again highlight that understanding this section on special cases is an excellent window into the proof in the general case.

4 Overview

In this section we will first describe three of the most important new theorems we proved that allowed us to give an improvement in the general case, and then give a proof overview. This section will focus on [KKO21], which shows that the cost of the max entropy algorithm is at most $(\frac{3}{2} - \epsilon) \cdot c(OPT)$ instead of $(\frac{3}{2} - \epsilon) \cdot c(x)$ where $x \in P_{\text{Sub}}$ (as is done in [KKO22]). In a later section we discuss the tools necessary to obtain the stronger claim. The essential difference is here we use the optimal cycle in our analysis, which dramatically simplifies a key part of our argument. Thus the theorem we outline in this section is as follows:

Theorem 4.1. *The max entropy is a randomized $\frac{3}{2} - \epsilon$ approximation for metric TSP for some $\epsilon > 10^{-36}$.*

While this section is an overview, we do include some technically important material, in particular a proof of [Theorem 4.1](#) using the two main technical theorems in later sections.

4.1 Three New Techniques

4.1.1 Polygon Structure for Near Minimum Cuts Crossed on one Side.

It should be clear from the previous section that understanding the structure of the set of near minimum cuts is key. We already know that if the only minimum cuts are the vertices, we can obtain a better-than-3/2 algorithm. So we need to study what other structures can arise.

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

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

There have been several works studying the structure of edges between these atoms and the structure of cuts in a connected component of cuts \mathcal{C} w.r.t. the a_i 's. The *cactus structure* (see [DKL76]) shows that if $\eta = 0$, then we can arrange the a_i 's of a connected component around a cycle, say a_1, \dots, a_m (after renaming), such that $x(E(a_i, a_{i+1})) = 1$ for all i .

Benczúr and Goemans [Ben95; BG08] studied the case when $\eta \leq 6/5$ and introduced the notion of *polygon representation*, in which case atoms can be placed on the sides of an equilateral polygon and some atoms placed inside the polygon, such that every cut in \mathcal{C} can be represented by a diagonal of this polygon. Later, [OSS11] studied the structure of edges of G in this polygon when $\eta < 1/100$.

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

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

Then, we have (i) For any atom a_i , $x(\delta(a_i)) \leq 2 + O(\eta)$ and (ii) For any pair of atoms a_i, a_{i+1} , $x(E(a_i, a_{i+1})) \geq 1 - \Omega(\eta)$ (see [Theorem 5.9](#) for details).

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

4.1.2 Generalized Gurvits' Lemma

Given a real stable polynomial $p \in \mathbb{R}_{\geq 0}[z_1, \dots, z_n]$ (with non-negative coefficients), Gurvits proved the following inequality [Gur06; Gur08]

$$e^{-n} \inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n} \leq \partial_{z_1} \dots \partial_{z_n} p|_{z=0} \leq \inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}. \quad (7)$$

He defined $\inf_{z>0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}$ as the *capacity* of the polynomial p at the all ones vector, $\mathbf{1}$.

As a consequence, one can prove the following theorem about [strongly Rayleigh](#) (SR) distributions by demonstrating that the capacity of the generating polynomial of such a distribution is at least 1.

Theorem 4.2. *Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be SR and A_1, \dots, A_m be random variables corresponding to the number of elements sampled in m disjoint subsets of $[n]$ such that $\mathbb{E}[A_i] = n_i$ for all i . If $n_i = 1$ for all $1 \leq i \leq n$, then $\mathbb{P}[\forall i, A_i = 1] \geq e^{-m}$.*

One can ask what happens if the vector (n_1, \dots, n_m) in the above theorem is not equal but close to $\mathbf{1}$. To answer this, we prove a generalization of the above statement, although with a significantly weaker constant. Roughly speaking, we show that as long as $\sum_{i=1}^m |n_i - 1| < 1 - \epsilon$ then $\mathbb{P}[\forall i, A_i = 1] \geq f(\epsilon, m)$ where $f(\epsilon, m)$ has no dependence on n , the number of underlying elements in the support of μ .

Theorem 4.3 (Informal version of [Proposition 6.8](#)). *Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be SR and let A_1, \dots, A_m be random variables corresponding to the number of elements sampled in m disjoint subsets of $[n]$. Suppose that there are integers n_1, \dots, n_m such that for any set $S \subseteq [m]$, $\mathbb{P}[\sum_{i \in S} A_i = \sum_{i \in S} n_i] \geq \epsilon$. Then,*

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

If one only cares about bounding the probability by a constant, the above statement is even stronger than [Theorem 4.2](#) as we only require $\mathbb{P}[\sum_{i \in S} A_i = \sum_{i \in S} n_i]$ to be bounded away from 0 for any set $S \subseteq [m]$ and we don't need a bound on the expectation. Our proof of the above theorem has doubly exponential dependence on ϵ . We leave it an open problem to find the optimum dependency on ϵ . Furthermore, our proof of the above theorem is probabilistic in nature; we expect that an algebraic proof based on the theory of real stable polynomials will provide a significantly improved lower bound.

In an independent work, Gurvits and Leake [GL21] proved a variant of the above theorem with a much better dependence on ϵ and m for a homogeneous strong Rayleigh distribution. This was subsequently updated to a version of the above theorem that does not require homogeneity. While the constant they obtain is much better, the statement is still slightly weaker in that it deals with expectations and not probabilities. Thus it gives constant bounds for a strictly smaller set of events. We reproduce it here, stated in the case in which we want all m sets to equal 1 for simplicity (the theorem in [GL21] deals with a general vector κ in which we want $A_i = \kappa_i$ for all i).

Theorem 4.4 (Corollary 8.1 in [GL21]). *Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a strongly Rayleigh distribution and let A_1, \dots, A_m be random variables corresponding to the number of elements sampled in m disjoint subsets of $[n]$. Define $\beta_i = \mathbb{E}[A_i]$. Then if $\|\beta - \mathbf{1}\|_1 \leq 1 - \epsilon$ then*

$$\mathbb{P}[A_i = 1, \forall i] \geq \left(\frac{\epsilon}{e}\right)^m$$

This is proved by lower bounding the capacity of the generating polynomial of such an SR distribution μ by ϵ^m and applying Gurvits' bound (7). It is interesting to study how to further bridge the gap between the two theorems: is it possible to bound the capacity for every regime in which [Theorem 4.3](#) gives us a constant?

4.1.3 Conditioning while Preserving Marginals

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

Let $A, B \subseteq [n]$ be two disjoint sets such that $\mathbb{E} [A_T], \mathbb{E} [B_T] \approx 1$. It follows from [Theorem 4.3](#) that $\mathbb{P} [A_T = B_T = 1] \geq \Omega(1)$. Here, however, we are interested in a stronger event; let $\nu = \mu |_{A_T = B_T = 1}$ and let $y_i = \mathbb{P}_{T \sim \nu} [i \in T]$. It turns out that the y vector can be very different from the x vector, in particular, for some i 's we can have $|y_i - x_i|$ bounded away from 0. We show that there is an event of non-negligible probability that is a subset of $A_T = B_T = 1$ under which the marginals of elements in A, B are almost preserved.

Theorem 4.5 (Informal version of [Proposition 6.13](#)). *Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a SR distribution and let $A, B \subseteq [n]$ be two disjoint subsets such that $\mathbb{E} [A_T], \mathbb{E} [B_T] \approx 1$. For any $\alpha \ll 1$ there is an event $\mathcal{E}_{A,B}$ such that $\mathbb{P} [\mathcal{E}_{A,B}] \geq \Omega(\alpha^2)$ and*

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

We remark that the quadratic lower bound on α is necessary in the above theorem for a sufficiently small $\alpha > 0$. The above theorem can be seen as a generalization of [Theorem 4.2](#) in the special case of two sets.

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

4.2 Overview of Proof

As alluded to earlier, the crux of the proof of [Theorem 4.1](#) is to show that the expected cost of the minimum cost matching on the odd degree vertices of the sampled tree is at most $OPT(1/2 - \epsilon)$. We do this by showing the existence of a cheap feasible O -join solution to (4).

First, recall that if we only wanted to get an O -join solution of value at most $OPT/2$, to satisfy all cuts, it is enough to set $y_e := x_e/2$ for each edge [[Wol80](#)]. To do better, we want to take advantage of the fact that we only need to satisfy a constraint in the O -join for S when $\delta(S)_T$ is odd. Here, we are aided by the fact that the sampled tree is likely to have many even cuts because it is drawn from a [strongly Rayleigh distribution](#).

If an edge e is exclusively on even cuts then y_e can be reduced below $x_e/2$. This, more or less, was the approach in [[OSS11](#)] for graphic TSP, where it was shown that a constant fraction of LP edges will be exclusively on *even* near min cuts with constant probability. The difficulty in implementing this approach in the metric case comes from the fact that a high cost edge can be on many cuts and it may be exceedingly unlikely that *all* of these cuts will be even simultaneously.

Overall, our approach to addressing this is to start with $y_e := x_e/2$ and then modify it with a random¹² slack vector $s : E \rightarrow \mathbb{R}$: When certain special (few) cuts that e is on are even we let $s_e = -x_e\beta$ (for a carefully chosen constant $\beta > 0$); for other cuts that contain e , whenever they are odd, we will increase the slack of other edges on that cut to satisfy them. The bulk of our effort is to show that we can do this while guaranteeing that $\mathbb{E}[s_e] < -\epsilon\beta x_e$ for some $\epsilon > 0$.

By carefully choosing β smaller than η , we do not need to worry about the reduction breaking a constraint for any cut S such that $x(\delta(S)) > 2(1 + \eta)$. In particular, if we choose $\beta \approx \eta/4$, any such cut is always satisfied, even if every edge in $\delta(S)$ is decreased and no edge is increased.

Let OPT be the optimum TSP tour, i.e., a Hamiltonian cycle, with set of edges E^* ; throughout the paper, we write e^* to denote an edge in E^* . To bound the expected cost of the O -join for a random spanning tree $T \sim \mu_\lambda$, we also construct a random slack vector $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ such that $(x + \text{OPT})/4 + s + s^*$ is a feasible for Eq. (4) with probability 1. In Section 4.2.1 we explain how to use s^* to satisfy all but a linear number of near mincuts.

Theorem 4.6 (Main Technical Theorem). *Let x^0 be a solution of (2) with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E . Let $z := (x + \text{OPT})/2$, $\eta \leq 10^{-12}$, $\beta > 0$, and let μ be the max-entropy distribution with marginals x . Also, let E^* denote the support of OPT . There are two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$), , such that*

- i) For each edge $e \in E$, $s_e \geq -x_e\beta$.
- ii) For each η -near-min-cut S of z , if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.
- iii) For every OPT edge e^* , $\mathbb{E}[s_{e^*}^*] \leq 218\eta\beta$ and for every LP edge $e \neq e_0$, $\mathbb{E}[s_e] \leq -\frac{1}{3}x_e\epsilon_P\beta$ for $\epsilon_P = 3.12 \cdot 10^{-16}$ (defined in (38)).

In the next subsection, we explain the main ideas needed to prove this technical theorem. But first, we show how our main theorem follows readily from Theorem 4.6.

Proof of Theorem 4.1. Let x^0 be an extreme point solution of the 2, with support E_0 and let x be x^0 restricted to E . By Fact 2.3 x is in the spanning tree polytope. For $\mu = \mu_{\lambda^*}$ the max entropy distribution with marginals x and $\beta > 0$ a parameter we choose below, let s, s^* be as defined in Theorem 4.6. We will define $y : E_0 \rightarrow \mathbb{R}_{\geq 0}$ and $y^* : E^* \rightarrow \mathbb{R}_{\geq 0}$. Let

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

we also let $y_{e^*}^* = 1/4 + s_{e^*}^*$ for any edge $e^* \in E^*$. We will show that $y + y^*$ is a feasible solution¹³ to (4). First, observe that for any S where $e_0 \in \delta(S)$, we have $y(\delta(S)) + y^*(\delta(S)) \geq 1$. Otherwise, we assume $u_0, v_0 \notin S$. If S is an η -near min cut w.r.t., z and $\delta(S)_T$ is odd, then by property (ii) of Theorem 4.6, we have

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

¹²where the randomness comes from the random sampling of the tree

¹³Recall that we merely need to prove the *existence* of a cheap O -join solution. The actual optimal O -join solution can be found in polynomial time.

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

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

where in the first inequality we used property (i) of [Theorem 4.6](#) which says that $s_e \geq x_e \beta$ with probability 1 for all LP edges and that $s_{e^*}^* \geq 0$ with probability 1. In the second inequality we used that $z = (x + OPT)/2$, so, since $OPT \geq 2$ across any cut, $x(\delta(S)) \leq 2(z(\delta(S)) - 1)$. Finally, if we choose

$$\beta = \eta/4.1 \tag{8}$$

then the righthand side is at least 1, so $y + y^*$ is a feasible O -join solution.

Finally, using $c(e_0) = 0$ and part (iii) of [Theorem 4.6](#),

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

choosing η such that

$$218\eta = \frac{1}{6}\epsilon_P \tag{9}$$

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

Now, we are ready to bound the approximation factor of our algorithm. First, since x^0 is an extreme point solution of the [2](#), $\min_{e \in E_0} x_e^0 \geq \frac{1}{n!}$. So, by [Theorem 2.6](#), in polynomial time we can find $\lambda : E \rightarrow \mathbb{R}_{\geq 0}$ such that for any $e \in E$, $\mathbb{P}_{\mu_\lambda}[e] \leq x_e(1 + \delta)$ for some δ that we fix later. It follows that

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

By stability of maximum entropy distributions (see [\[SV19, Thm 4\]](#) and references therein), we have that $\|\mu - \mu_\lambda\|_1 \leq O(n^4\delta) =: q$. Therefore, for some $\delta \ll n^{-4}$ we get $\|\mu - \mu_\lambda\|_1 = q \leq \frac{\epsilon_P\beta}{100}$. That means that

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

where we used that for any spanning tree the cost of the minimum cost matching on odd degree vertices is at most $OPT/2$. Finally, since $\mathbb{E}_{T \sim \mu_\lambda}[c(T)] \leq OPT(1 + \delta)$, $\epsilon_P = 3.12 \cdot 10^{-16}$, and $\beta = \eta/4.1 = \epsilon_P/5362.8$ (from [\(62\)](#)) we get a $3/2 - 3 \cdot 10^{-36}$ approximation algorithm for TSP. \square

4.2.1 Ideas underlying proof of [Theorem 4.6](#)

The first step of the proof is to show that it suffices to construct a slack vector s for a ‘‘cactus-like’’ structure of near min-cuts that we call a *hierarchy*. Informally, a hierarchy \mathcal{H} is a laminar family of mincuts¹⁴, consisting of two types of cuts: *triangle cuts* and *degree cuts*. A triangle S is the union of two min-cuts X and Y in \mathcal{H} such that $x(E(X, Y)) = 1$. See [Fig. 52](#) for an example of a hierarchy with three triangles.

¹⁴This is really a family of near-min-cuts, but for the purpose of this overview, assume $\eta = 0$

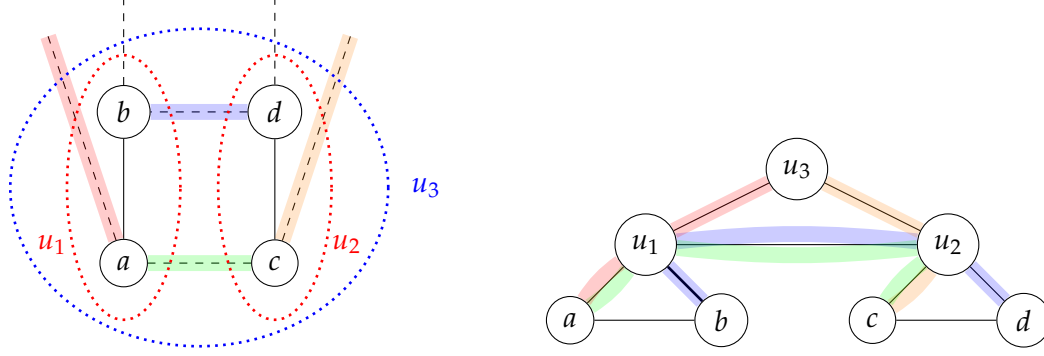


Figure 26: An example of part of a hierarchy with three triangles. The graph on the left shows part of a feasible LP solution where dashed (and sometimes colored) edges have fraction $1/2$ and solid edges have fraction 1 . The dotted ellipses on the left show the min-cuts u_1, u_2, u_3 in the graph. (Each vertex is also a min-cut). On the right is a representation of the corresponding hierarchy. Triangle u_1 corresponds to the cut $\{a, b\}$, u_2 corresponds to $\{c, d\}$ and u_3 corresponds to $\{a, b, c, d\}$. Note that, for example, the edge (a, c) , represented in green, is in $\delta(u_1)$, $\delta(u_3)$, and inside u_3 . For triangle u_1 , we have $A = \delta(a) \setminus (a, b)$ and $B = \delta(b) \setminus (b, d)$.

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

Theorem 4.7 (Main Payment Theorem (informal)). *Let $G = (V, E, x)$ for LP solution x and let μ be the max-entropy distribution with marginals x and $\beta > 0$. Given a hierarchy \mathcal{H} , there is a slack vector $s : E \rightarrow \mathbb{R}$ such that*

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

In the following subsection, we discuss how to prove this theorem. Here we explain at a high level how to define the hierarchy and reduce [Theorem 4.6](#) to this theorem. The details are in [Section 5](#).

First, observe that, given [Theorem 4.7](#), cuts in \mathcal{H} will automatically satisfy (ii) of [Theorem 4.6](#). The approach we take to satisfying all other cuts is to introduce additional slack, the vector s^* , on OPT edges.

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

To ensure that any cut S that is *crossed on both sides* is satisfied, we first observe that S is odd with probability $O(\eta)$. To see this, let S_L and S_R be the cuts crossing S on the left and right with minimum intersection with S and consider the two (bad) events $\{E(S \cap S_L, S_L \setminus S)_T \neq 1\}$ and $\{E(S \cap S_R, S_R \setminus S)_T \neq 1\}$. Recall that if A, B and $A \cup B$ are all near-min-cuts, then $\mathbb{P}[E(A, B)_T \neq 1] = O(\eta)$ (see [Corollary 2.29](#)). Applying this fact to the two aforementioned bad events implies that each of them has probability $O(\eta)$. Therefore, we will let the two *OPT* edges in $\delta(S)$ be responsible for these two events, i.e., we will increase the slack s^* on these two *OPT* edges by $O(\eta)$ when the respective bad events happens. This gives $\mathbb{E}[s^*(e^*)] = O(\eta^2)$ for each *OPT* edge e^* . As we will see, this simple step will reduce the number of near-min-cuts of z that we need to worry about satisfying to $O(n)$.

Next, we consider the set of near-min-cuts of z that are crossed on at most one side. Partition these into maximal connected components of crossing cuts. Each such component corresponds to an interval along the *OPT* cycle and, by definition, these intervals form a laminar family.

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

Theorem 4.8 (Polygons look like cycles (Informal version of [Theorem 5.9](#))). *Given a connected component \mathcal{C} of near-min-cuts of z that are crossed on one side, consider the coarsest partition of vertices of the *OPT* cycle into a sequence a_1, \dots, a_{m-1} of sets called atoms (together with a_0 which is the set of vertices not contained in any cut of \mathcal{C}). Then*

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

The main observation used to prove [Theorem 4.8](#) is that the cuts in \mathcal{C} crossed on one side can be partitioned into two laminar families \mathcal{L} and \mathcal{R} , where \mathcal{L} (resp. \mathcal{R}) is the set of cuts crossed on the left (resp. right). This immediately implies that $|\mathcal{C}|$ is linear in m . Since cuts in \mathcal{L} cannot cross each other (and similarly for \mathcal{R}), the proof boils down to understanding the interaction between \mathcal{L} and \mathcal{R} .

The approximations in [Theorem 4.8](#) are correct up to $O(\eta)$. Using additional slack in *OPT*, at the cost of an additional $O(\eta^2)$ for edge, we can treat these approximate equations as if they are exact. Observe that if $x(E(a_i, a_{i+1})) = 1$, and $x(\delta(a_i)) = x(\delta(a_{i+1})) = 2$ for $1 \leq i \leq m - 2$, then with probability 1, $E(a_i, a_{i+1})_T = 1$. Therefore, any cut in \mathcal{C} which doesn't include a_1 or a_{m-1} is even with probability 1. The cuts in \mathcal{C} that contain a_1 are even precisely¹⁵ when $E(a_0, a_1)_T$ is odd and similarly the cuts in \mathcal{C} that contain a_{m-1} are even when $E(a_0, a_{m-1})_T$ is odd. These observations are what allow us to imagine that each polygon is a triangle, i.e., assume $m = 3$. (Note that often it is convenient to look at the event in which $E(a_0, a_1)_T = 1$ and $E(a_0, a_{m-1})_T = 1$ since this is a simple criteria which implies that all cuts in \mathcal{C} are even.)

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

¹⁵Roughly, this corresponds to the definition of the polygon being left-happy.

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

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

We also observe that to prove [Theorem 4.1](#), we crucially used that $28\eta \ll \epsilon$. This forces us to take η very small, which is why we get only a “very slightly” improved approximation algorithm for TSP. Furthermore, since we use OPT edges in our construction, we don’t get a new upper bound on the integrality gap. We leave it as an open problem to find a reduction to the “cactus” case that doesn’t involve using a slack vector for OPT (or a completely different approach).

4.2.2 Proof ideas for [Theorem 4.7](#)

We now address the problem of constructing a good slack vector s for a hierarchy of degree cuts and triangle cuts. For each LP edge f , consider the lowest cut in the hierarchy, that contains both endpoints of f . We call this cut $p(f)$. If $p(f)$ is a degree cut, then we call f a *top edge* and otherwise, it is a *bottom edge*¹⁶. We will see that bottom edges are easier to deal with, so we start by discussing the slack vector s for top edges.

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

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

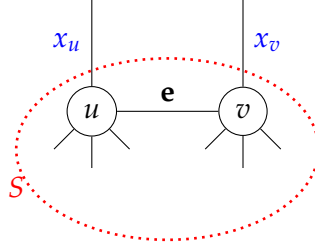
Now consider the situation from the perspective of the degree cut u (where $p(u) = S$) and consider any incident edge bundle in S , e.g., $\mathbf{e} = (u, v)$. Either its top cuts are both even and $s_{\mathbf{e}} := -\eta x_{\mathbf{e}}$, or they aren’t even, because, for example, $\delta(u)_T$ is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

Example 4.9 (Simple case). Fix a top edge bundle $\mathbf{e} = (u, v)$ with $p(\mathbf{e}) = S$. Let $x_u := x(\delta^\uparrow(u))$ and let $x_v := x(\delta^\uparrow(v))$. Suppose we have constructed a (fractional) *matching* between edges whose top two cuts are children of S in \mathcal{H} and the edges in $\delta(S)$, and this matching satisfies the following three conditions: (a) $\mathbf{e} = (u, v) \in S$ is matched (only) to edges going higher from its top two cuts (i.e., to edges in $\delta^\uparrow(u)$ and $\delta^\uparrow(v)$), (b) \mathbf{e} is matched to an $m_{\mathbf{e},u}$ fraction of every edge in $\delta^\uparrow(u)$ and to an $m_{\mathbf{e},v}$ fraction of each edge in $\delta^\uparrow(v)$, where

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

¹⁶For example, in [Fig. 52](#), $p(a, c) = u_3$, and (a, c) is a bottom edge.

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



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

$$\begin{aligned} \mathbb{E}[s_e] = & -\mathbb{P}[e \text{ reduced}] \eta x_e + m_{e,u} \sum_{g \in \delta^{\uparrow}(u)} \mathbb{P}[\delta(u)_T \text{ odd} | g \text{ reduced}] \mathbb{P}[g \text{ reduced}] \eta \frac{x_g}{x(\delta^{\uparrow}(u))} \\ & + m_{e,v} \sum_{g \in \delta^{\uparrow}(v)} \mathbb{P}[\delta(v)_T \text{ odd} | g \text{ reduced}] \mathbb{P}[g \text{ reduced}] \eta \frac{x_g}{x(\delta^{\uparrow}(v))} \end{aligned} \quad (10)$$

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

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

Unfortunately, we do not currently have a good handle on the parity of $\delta^{\uparrow}(u)_T$ conditioned on g reduced. However, we can use the following simple but crucial property: Since $x(\delta(S)) = 2$, by [Lemma 2.28](#), T consists of two independent trees, one on S and one on $V \setminus S$, each with the corresponding marginals of x . Therefore, we can write

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

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

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

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

The rest of the analysis of this special case follows from (a) the fact that our construction will guarantee that for *all* edges g , the probability that g is reduced is *exactly* p , i.e., it is the same for all edges, and (b) the fact that $m_{e,u}x_u + m_{e,v}x_v = x_e$. Plugging these facts back into (10), gives

$$\begin{aligned} \mathbb{E}[s_e] & \leq -p\eta x_e + m_{e,u}(1 - \epsilon)p\eta + m_{e,v}(1 - \epsilon)p\eta \\ & \leq -p\eta x_e + (1 - \epsilon)p\eta x_e = -\epsilon p\eta x_e. \end{aligned} \quad (11)$$

If we could prove (11) for *every* edge f in the support of x , that would complete the proof that the expected cost of the min O -join for a random spanning tree $T \sim \mu$ is at most $(1/2 - \epsilon)OPT$.

Remark: Throughout this paper, we repeatedly use a mild generalization of the above "independent trees fact": that if S is a cut with $x(\delta(S)) \leq 2 + \epsilon$, then S_T is very likely to be a tree. Conditioned on this fact, marginals inside S and outside S are nearly preserved and the trees inside S and outside S are sampled independently (see [Lemma 2.28](#)).

Ideal reduction: In the example, we were able to show that $\mathbb{P}[\delta(u)_T \text{ odd} \mid g \text{ reduced}]$ was bounded away from 1 for every edge $g \in \delta^\uparrow(u)$, and this is how we proved that the expected reduction for each edge was greater than the expected increase on each edge, yielding negative expected slack.

This motivates the following definition: A reduction for an edge g is k -ideal if, conditioned on g reduced, every cut S that is in the top k levels of cuts containing g is odd with probability that is bounded away from 1.

Moving away from an idealized setting: In [Example 4.9](#), we oversimplified in four ways:

- (a) We assumed that it would be possible to show that each top edge is *good*. That is, that its top two cuts are even *simultaneously* with constant probability.
- (b) We considered only top edge bundles (i.e., edges whose top cuts were inside a degree cut).
- (c) We assumed that $x_u, x_v \in [\epsilon, 1 - \epsilon]$.
- (d) We assumed the existence of a nice matching between edges whose top two cuts were children of S and the edges in $\delta(S)$.

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

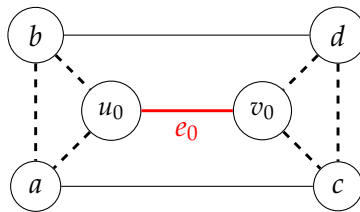


Figure 27: An Example with Bad Edges. A feasible solution of the 2 is shown; dashed edges have fraction $1/2$ and solid edges have fraction 1. Writing $E = E_0 \setminus \{e_0\}$ as a maximum entropy distribution μ we get the following: Edges $(a, b), (c, d)$ must be completely negatively correlated (and independent of all other edges). So, $(b, u_0), (a, u_0)$ are also completely negatively correlated. This implies (a, b) is a bad edge.

Bad edges. Consider first (a). Unfortunately, it is not the case that all top edges are good. Indeed, some are *bad*. However, it turns out that bad edges are rare in the following senses: First, for an edge to be bad, it must be a half edge, where we say that an edge e is a half edge if $x_e \in 1/2 \pm \epsilon_{1/2}$ for a suitably chosen constant $\epsilon_{1/2}$. Second, of any two half edge bundles sharing a common endpoint in the hierarchy, at least one is good. For example, in [Fig. 27](#), (a, u_0) and (b, u_0) are good half-edge bundles. We advise the reader to ignore half edges in the first reading of the paper. Correspondingly, we note that our proofs would be much simpler if half-edge bundles never

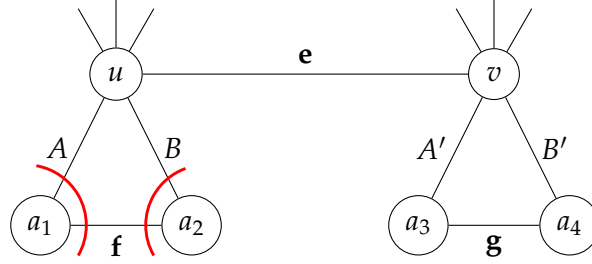


Figure 28: In this representation of the cut hierarchy (as in Fig. 52), for the triangle u corresponding to the cut $\delta(a_1 \cup a_2)$, when A_T and B_T are odd, all 3 cuts ($\delta(a_1)_T, \delta(a_2)_T$ and $\delta(a_1 \cup a_2)_T = \delta(u)_T$) are odd (since \mathbf{f}_T is always 1). (Recall also that the edges in the bundle \mathbf{e} must have one endpoint in $\{a_1 \cup a_2\}$ and one endpoint in $\{a_3 \cup a_4\}$, as was the case, e.g., for the edge (a, c) in Fig. 52.)

showed up in the hierarchy. It may not be a coincidence that half edges are hard to deal with, as it is conjectured that TSP instances with half-integral LP solutions are the hardest to round [SWZ12; SWZ13].

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

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

We deal with the second problem by proving Lemma 7.2, which guarantees a matching between *good* edge bundles $\mathbf{e} = (u, v)$ and fractions $m_{\mathbf{e},u}, m_{\mathbf{e},v}$ of edges in $\delta^\uparrow(u), \delta^\uparrow(v)$ such that, roughly, $m_{\mathbf{e},u} + m_{\mathbf{e},v} = (1 + O(\epsilon_{1/2}))x_{\mathbf{e}}$.

Dealing with triangles. Turning to (b), consider a triangle cut S , for example $\delta(a_1 \cup a_2)$ in Fig. 28. Recall that in a triangle, we can assume that there is an edge of fractional value 1 connecting a_1 and a_2 in the tree, and this is why we defined the cut to be happy when A_T and B_T are odd: this guarantees that all 3 cuts defined by the triangle ($\delta(a_1), \delta(a_2), \delta(a_1 \cup a_2)$) are even.

Now suppose that $\mathbf{e} = (u, v)$ is a top edge bundle, where u and v are both triangles, as shown in Fig. 28. Then we'd like to reduce $s_{\mathbf{e}}$ when both cuts u and v are happy. But this would require more than simply both cuts being even. This would require *all* of A_T, B_T, A'_T, B'_T to be odd. Note that if, for whatever reason, \mathbf{e} is reduced only when $\delta(u)_T$ and $\delta(v)_T$ are both even, then it could be, for example, that this only happens when A_T and B_T are both even. In this case, both $\delta(a_1)_T$ and $\delta(a_2)_T$ will be odd with probability 1 (recalling that $\mathbf{f}_T = 1$), which would then necessitate an increase in $s_{\mathbf{f}}$ whenever \mathbf{e} is reduced. In other words, the reduction will not even be 1-ideal.

It turns out to be easier for us to get a 1-ideal reduction rule for \mathbf{e} as follows: Say that \mathbf{e} is 2-1-1 happy with respect to u if $\delta(u)_T$ is even and both A'_T, B'_T are odd. We reduce \mathbf{e} with probability $p/2$

when it is 2-1-1 happy with respect to u and with probability $p/2$ when it is 2-1-1 happy with respect to v . This means that when \mathbf{e} is reduced, half of the time no increase in $s_{\mathbf{f}}$ is needed since u is happy. Similarly for v .

The 2-1-1 criterion for reduction introduces a new kind of bad edge: a half edge that is good, but not 2-1-1 good. We are able to show that non-half-edge bundles are 2-1-1 good (Lemmas 6.28 and 6.29), and that if there are two half edges which are both in A or are both in B , then at least one of them is 2-1-1 good (Lemma 6.30). Finally, we show that if there are two half edges, where one is in A and the other is in B , and neither is 2-1-1 good, then we can apply a different reduction criterion that we call 2-2-2 good. When the latter applies, we are guaranteed to decrease both of the half edge bundles simultaneously. All together, the various considerations discussed in this paragraph force us to come up with a relatively more complicated set of rules under which we reduce $s_{\mathbf{e}}$ for a top edge bundle \mathbf{e} whose children are triangle cuts. Section 6 focuses on developing the relevant probabilistic statements.

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

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

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

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

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

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

Coming to our aid is the second key idea, already discussed in Section 4.1.3. We reduce bottom edges only when $A_T = B_T = 1$ and the marginals of edges in A, B are approximately preserved (conditioned on $A_T = B_T = 1$). This allows us to get much stronger upper bounds on the probability that a lower cut a bottom edge is on is odd, given that the bottom edge is reduced, and enables us to show that bottom edge reduction is ∞ -ideal.

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

Example 4.10. [Bottom-bottom case] To see how preserving marginals helps us handle the interaction between bottom edges at consecutive levels, consider a triangle cut $a'_1 = \{a_1, a_2\}$ whose parent cut $\hat{S} = \{a'_1, a'_2\}$ is also a triangle cut (as shown in Fig. 29). Let's analyze $\mathbb{E}[s_{\mathbf{f}}]$ where $\mathbf{f} = (a_1, a_2)$. Observe first that $A^{\rightarrow} \cup B^{\rightarrow}$ is a bottom edge bundle in the triangle \hat{S} and all edges in

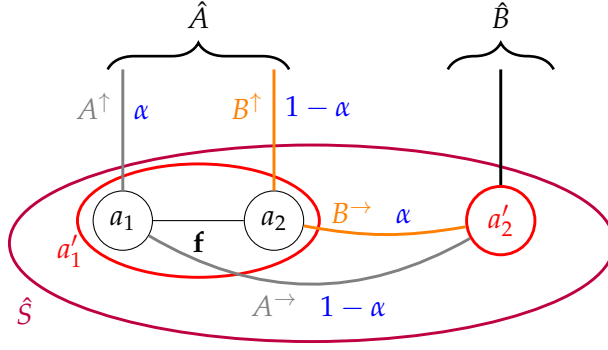


Figure 29: Setting of Example 4.10. Note that the set $A = \delta(a_1) \cap \delta(a'_1)$ decomposes into two sets of edges, A^\uparrow , those that are also in $\delta(S)$, and the rest, which we call A^\rightarrow . Similarly for B .

this bundle are reduced simultaneously when $\hat{A}_T = \hat{B}_T = 1$ and marginals of all edges in $\hat{A} \cup \hat{B}$ are approximately preserved. (For the purposes of this overview, we'll assume they are preserved exactly). Furthermore, since the tree inside \hat{S} is picked independently of the tree on G/\hat{S} (using Lemma 2.28 and assuming $\epsilon = 0$ for this overview), exactly one edge in $A^\rightarrow \cup B^\rightarrow$ is selected independently of the reduction event $\hat{A}_T = \hat{B}_T = 1$. Let $x(A^\uparrow) = \alpha$. Then since $A = A^\uparrow \cup A^\rightarrow$ and $x(A) = 1$, we have $x(A^\rightarrow) = 1 - \alpha$. Moreover, since $\hat{A} = A^\uparrow \cup B^\uparrow$ and $x(\hat{A}) = 1$, we also have $x(B^\uparrow) = 1 - \alpha$ and $x(B^\rightarrow) = \alpha$.

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

$$\mathbb{P}[a'_1 \text{ happy} \mid A^\rightarrow \cup B^\rightarrow \text{ reduced}] = \mathbb{P}[A_T = B_T = 1 \mid A^\rightarrow \cup B^\rightarrow \text{ reduced}] = \alpha^2 + (1 - \alpha)^2.$$

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

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

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

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

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

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

Dealing with x_u close to 1. ¹⁷ Now, suppose that $\mathbf{e} = (u, v)$ is a top edge bundle with $x_u := x(\delta^\uparrow(u))$ is close to 1. Then, the analysis in [Example 4.9](#), bounding $r := \mathbb{P}[\delta(u)_T \text{ odd} | g \text{ reduced}]$ away from 1 for an edge $g \in \delta^\uparrow(u)$ doesn't hold. To address this, we consider two cases: The first case, is that the edges in $\delta^\uparrow(u)$ break up into many groups that end at different levels in the hierarchy. In this case, we can analyze r separately for the edges that end at any given level, taking advantage of the independence between the trees chosen at different levels of the hierarchy.

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

Now, we define this more complex 2-1-1 reduction rule: Consider a top edge $\mathbf{f} = (u', v') \in \delta^\rightarrow(u')$. So far, we only considered the following reduction rule for \mathbf{f} : If both u', v' are degree cuts, \mathbf{f} reduces when they are both even in the tree; otherwise if say u' is a triangle cut, \mathbf{f} reduces when it is 2-1-1 good w.r.t., u' (and similarly for v'). But clearly these rules ignore the pseudo triangle. The simplest adjustment is, if u' is a pseudo triangle with partition $(u, u' \setminus u)$, to require \mathbf{f} to reduce when $A_T = B_T = 1$ and v' is happy. However, as stated, it is not clear that the sets A and B are well-defined. For example, u' could be an actual triangle or there could be multiple ways to see u' as a pseudo triangle only one of which is $(u, u' \setminus u)$. Our solution is to find the *smallest* disjoint pair of cuts $a, b \subset u'$ in the hierarchy such that $x(\delta(a) \cap \delta(u')), x(\delta(b) \cap \delta(u')) \geq 1 - \epsilon_{1/1}$, where $\epsilon_{1/1}$ is a fixed universal constant, and then let $A = \delta(a) \cap \delta(u'), B = \delta(b) \cap \delta(u')$ and $C = \delta(u') \setminus A \setminus B$ (see [Fig. 30](#) for an example). Then, we say \mathbf{f} is 2-1-1 happy w.r.t., u' if $A_T = B_T = 1$ and $C_T = 0$.

A few observations are in order:

- Since u is a candidate for, say a , it must be that a is a descendent of u in the hierarchy (or equal to u). In addition, b cannot simultaneously be in u , since $a \cap b = \emptyset$ and $x(\delta(u) \cap \delta(u')) \leq 1$ by [Lemma 2.37](#). So, when \mathbf{f} is 2-1-1 happy w.r.t. u' we get $(\delta(u) \cap \delta(u'))_T = 1$.
- If $u' = (X, Y)$ is an actual triangle cut, then we must have $a \subseteq X, b \subseteq Y$. So, when \mathbf{f} is 2-1-1 happy w.r.t. u' , we know that u' is a happy triangle, i.e., $(\delta(X) \cap \delta(u'))_T = 1$ and $(\delta(Y) \cap \delta(u'))_T = 1$.

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

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

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

Dealing with x_u close to 0 and the matching. We already discussed how the matching is modified to handle the existence of bad edges. We now observe that we can handle the case $x_u \approx 0$ by further modifying the matching. The key observation is that in this case, $x(\delta^\rightarrow(u)) \gg x(\delta^\uparrow(u))$. Roughly speaking, this enables us to find a matching in which each edge in $\delta^\rightarrow(u)$ has to increase about half as much as would normally be expected to fix the cut of u . This eliminates the need to prove a nontrivial bound on $\mathbb{P}[\delta(u)_T \text{ odd} | g \text{ reduced}]$. The details of the matching are in [Section 7](#).

¹⁷Some portions of this discussion might be easier to understand after reading the rest of the paper.

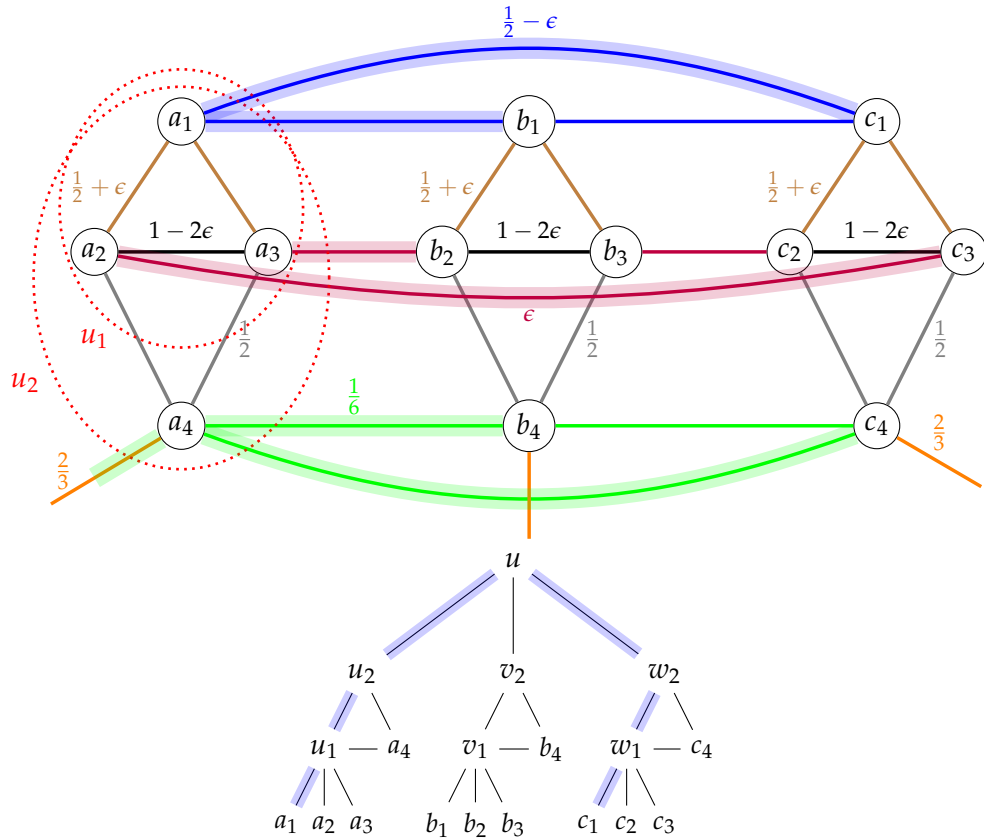


Figure 30: Part of the hierarchy of the graph is shown on top. Edges of the same color have the same fraction and $\epsilon \gg \eta$ is a small constant. u_1 corresponds to the degree cut $\{a_1, a_2, a_3\}$, u_2 corresponds to the triangle cut $\{u_1, a_4\}$ and u corresponds to the degree cut containing all of the vertices shown. Observe that edges in $\delta^\uparrow(a_1)$ are top edges in the degree cut u . If $\epsilon < \frac{1}{2}\epsilon_{1/1}$ then the (A, B, C) -degree partitioning of edges in $\delta(u_2)$ is as follows: $A = \delta(a_1) \cap \delta(u_2)$ are the blue highlighted edges each of fractional value $1/2 - \epsilon$, $B = \delta(a_4) \cap \delta(u_2)$ are the green highlighted edges of total fractional value 1, and C are the red highlighted edges each of fractional value ϵ . The cuts that contain edge (a_1, c_1) are highlighted in the hierarchy at the bottom.

5 Polygons and the Hierarchy of Near Minimum Cuts

This section, as detailed in [Section 4](#) above, demonstrates that **we can focus our attention on a laminar family of cuts** instead of every near minimum cut in the graph. This is an important step in the analysis as it gives us a clean combinatorial object to analyze using properties of SR distributions.

This section is modular up to [Section 5.6](#), in which the results of this section are combined with the main result from the remainder of the paper to demonstrate the main technical theorem, [Theorem 4.6](#). Thus readers wishing to skip this part may decide to only read [Section 5.6](#) before moving on and assuming that the set of near minimum cuts in the LP solution is laminar.

5.1 Notation

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

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

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

Definition 5.1 (Crossed on the Left/Right, Crossed on Both Sides). *For two crossing near minimum cuts S, S' , we say S crosses S' on the left if the leftmost endpoint of S on the optimal cycle is to the left of the leftmost endpoint of S' . Otherwise, we say S crosses S' on the right.*

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

5.2 Cuts Crossed on Both Sides

The following theorem is the main result of this section:

Theorem 5.2. *Given OPT TSP tour with set of edges E^* , and a feasible LP solution x^0 of (2) with support $E_0 = E \cup \{e_0\}$ and let x be x^0 restricted to E . For any distribution μ of spanning trees with marginals x and $\beta > 0$, if $\eta < 1/100$, then there is a random vector $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ (the randomness in s^* depends exclusively on $T \sim \mu$) such that*

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

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

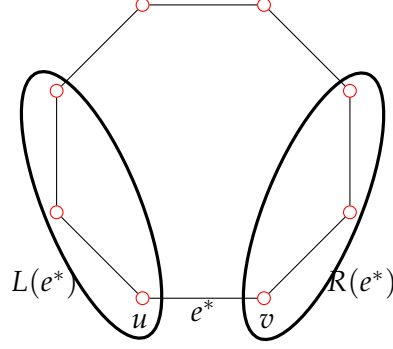


Figure 31: L and R for an OPT edge e^* .

v and not u which is crossed on both sides. (Note that $L(e^*), R(e^*)$ do not necessarily exist). For example, see Fig. 31.

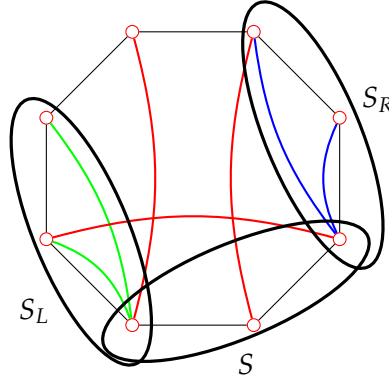


Figure 32: S is crossed on the left by S_L and on the right by S_R . In green are edges in $\delta(S)_L$, in blue edges in $\delta(S)_R$, and in red are edges in $\delta(S)_O$.

Definition 5.3. For a near minimum cut S that is crossed on both sides let S_L be the near minimum cut crossing S on the left which minimizes the intersection with S , and similarly for S_R ; if there are multiple sets crossing S on the left with the same minimum intersection, choose the smallest one to be S_L (and similar do for S_R).

We partition $\delta(S)$ into three sets $\delta(S)_L, \delta(S)_R$ and $\delta(S)_O$ as in Fig. 32 such that

$$\begin{aligned}\delta(S)_L &= E(S \cap S_L, S_L \setminus S) \\ \delta(S)_R &= E(S \cap S_R, S_R \setminus S) \\ \delta(S)_O &= \delta(S) \setminus (\delta(S)_L \cup \delta(S)_R)\end{aligned}$$

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

$$|T \cap \delta(L(e^*))_R| = 1, |T \cap \delta(R(e^*))_L| = 1, T \cap \delta(L(e^*))_O = \emptyset, \text{ and } T \cap \delta(R(e^*))_O = \emptyset. \quad (12)$$

In the proof of [Theorem 5.2](#) we will increase an OPT edge e^* whenever $\mathcal{I}_2(e^*)$ occurs.

Lemma 5.4. *For any OPT edge e^* , $\mathbb{P}[\mathcal{I}_2(e^*)] \leq 18\eta$.*

Proof. Fix e^* . To simplify notation we abbreviate $L(e^*), R(e^*)$ to L, R . Since L is crossed on both sides, L_L, L_R are well defined. Since by [Lemma 2.35](#) $L_L \cap L, L_L \setminus L$ are 4η -near min cuts and L is 2η -near mincut with respect to x , by [Corollary 2.29](#), $\mathbb{P}[|T \cap \delta(L)_L| = 1] \geq 1 - 5\eta$. Similarly, $\mathbb{P}[|T \cap \delta(R)_L| = 1] \geq 1 - 5\eta$. On the other hand, since L, L_L, L_R are 2η -near min cuts, by [Lemma 2.36](#), $x(E(L \cap L_R, L_R)), x(E(L \cap L_L, L_L)) \geq 1 - \eta$. Therefore

$$x(\delta(L)_O) \leq 2 + 2\eta - x(E(L \cap L_R, L_R)) - x(E(L \cap L_L, L_L)) \leq 4\eta.$$

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

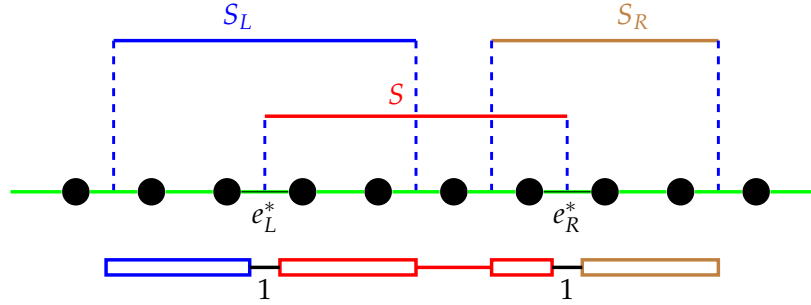


Figure 33: Setting of [Lemma 11.3](#). Here we zoom in on a portion of the optimal cycle and assume the root is not shown. If $\mathcal{I}_2(e_L^*)$ does not occur then $E(S \cap S_L, S_L \setminus S)_T = 1$.

Lemma 5.5. *Let S be a cut which is crossed on both sides and let e_L^*, e_R^* be the OPT edges on its interval where e_L^* is the edge further clockwise. Then, if $\delta(S)_T \neq 2$, at least one of $\mathcal{I}_2(e_L^*), \mathcal{I}_2(e_R^*)$ occurs.*

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

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

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

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

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

Proof of Theorem 5.2. For any OPT edge e^* whenever $\mathcal{I}_2(e^*)$ occurs, define $s_{e^*}^* = 2.02\beta$. Then, by Lemma 11.5, $\mathbb{E}[s_{e^*}] \leq 18 \cdot 2.02\beta$ and for any 2η -near min cut S (w.r.t., x) that is crossed on both sides if $\delta(S)_T$ is odd, then at least one of $\mathcal{I}_2(e_L^*), \mathcal{I}_w(e_R^*)$ occurs, so

$$s(\delta(S)) + s^*(\delta(S)) \geq -x(\delta(S))\beta + s_{e_L^*}^* + s_{e_R^*}^* \geq -(2 + 2\eta)\beta + 2.02\beta \geq 0$$

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

5.3 Proof of the Main Technical Theorem, Theorem 12.1

The following theorem is the main result of this section.

Theorem 5.6. *Let x^0 be a feasible solution of the 2 with support $E_0 = E \cup \{e_0\}$ and x be x^0 restricted to E . Let μ be the max entropy distribution with marginals x . For $\eta \leq 10^{-12}$, $\beta > 0$, there is a set $E_g \subset E \setminus \delta(\{u_0, v_0\})$ of good edges and two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$) such that*

- (i) *For each edge $e \in E_g$, $s_e \geq -x_e\beta$ and for any $e \in E \setminus E_g$, $s_e = 0$.*
- (ii) *For each η -near-min-cut S w.r.t. z , if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.*
- (iii) *We have $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ for all edges $e \in E_g$ and $\mathbb{E}[s_{e^*}^*] \leq 218\eta\beta$ for all OPT edges $e^* \in E^*$. for ϵ_P defined in (38).*
- (iv) *For every η -near minimum cut S of z crossed on (at most) one side such that $S \neq V \setminus \{u_0, v_0\}$, $x(\delta(S) \cap E_g) \geq 3/4$.*

Before proving this theorem we use it to prove the main technical theorem as given in the overview.

Theorem 4.6 (Main Technical Theorem). *Let x^0 be a solution of (2) with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E . Let $z := (x + \text{OPT})/2$, $\eta \leq 10^{-12}$, $\beta > 0$, and let μ be the max-entropy distribution with marginals x . Also, let E^* denote the support of OPT. There are two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$), such that*

- i) *For each edge $e \in E$, $s_e \geq -x_e\beta$.*
- ii) *For each η -near-min-cut S of z , if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.*
- iii) *For every OPT edge e^* , $\mathbb{E}[s_{e^*}^*] \leq 218\eta\beta$ and for every LP edge $e \neq e_0$, $\mathbb{E}[s_e] \leq -\frac{1}{3}x_e\epsilon_P\beta$ for $\epsilon_P = 3.12 \cdot 10^{-16}$ (defined in (38)).*

Proof of Theorem 4.6. Let E_g be the good edges defined in Theorem 5.6 and let $E_b := E \setminus E_g$ be the set of bad edges; in particular, note all edges in $\delta(\{u_0, v_0\})$ are bad edges. We define a new vector $\tilde{s} : E \cup \{e_0\} \rightarrow \mathbb{R}$ as follows:

$$\tilde{s}(e) \leftarrow \begin{cases} \infty & \text{if } e = e_0 \\ -x_e(4\beta/5)(1 - 2\eta) & \text{if } e \in E_b, \\ x_e(4\beta/3) & \text{otherwise.} \end{cases} \quad (13)$$

Let \tilde{s}^* be the vector s^* from [Theorem 5.2](#). We claim that for any η -near minimum cut S such that $\delta(S)_T$ is odd, we have

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

To check this note by (iv) of [Theorem 5.6](#) for every set $S \neq V \setminus \{u_0, v_0\}$ crossed on at most one side, we have $x(E_g \cap \delta(S)) \geq \frac{3}{4}$, so

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \geq \tilde{s}(\delta(S)) = \frac{4\beta}{3}x(E_g \cap \delta(S)) - \frac{4\beta}{5}(1 - 2\eta)x(E_b \cap \delta(S)) \geq 0. \quad (14)$$

For $S = V \setminus \{u_0, v_0\}$, we have $\delta(S)_T = \delta(u_0)_T + \delta(v_0)_T = 2$ with probability 1, so condition ii) is satisfied for these cuts as well. Finally, consider cuts S which are crossed on both sides. By [Theorem 5.2](#),

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \geq 0 \quad (15)$$

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

Now, we are ready to define s, s^* . Let \hat{s}, \hat{s}^* be the s, s^* of [Theorem 5.6](#) respectively. Define $s = \gamma\tilde{s} + (1 - \gamma)\hat{s}$ and similarly define $s^* = \gamma\tilde{s}^* + (1 - \gamma)\hat{s}^*$ for some γ that we choose later. We prove all three conclusions for s, s^* . (i) follows by (i) of [Theorem 5.6](#) and [Eq. \(13\)](#). (ii) follows by (ii) of [Theorem 5.6](#) and [Eq. \(14\)](#) above. It remains to verify (iii). For any OPT edge e^* , $\mathbb{E}[s_{e^*}^*] \leq 218\eta\beta$ by (iii) of [Theorem 5.6](#) and the construction of \tilde{s}^* . On the other hand, by (iii) of [Theorem 5.6](#) and [Eq. \(13\)](#),

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

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

5.4 Structure of Polygons of Cuts Crossed on One Side

Definition 5.7 (Connected Component of Crossing Cuts). *Given a family of cuts crossed on at most one side, construct a graph where two cuts are connected by an edge if they cross. Partition this graph into maximal connected components. We call a path in this graph, a path of crossing cuts.*

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

Definition 5.8 (Polygon). *For a connected component \mathcal{C} of crossing near min cuts that are crossed on one side, let a_0, \dots, a_{m-1} be the coarsest partition of the vertices V , such that for all $0 \leq i \leq m - 1$ and for any $A \in \mathcal{C}$ either $a_i \subseteq A$ or $a_i \cap A = \emptyset$. These are called atoms. We assume a_0 is the atom that contains the special edge e_0 , and we call it the root. Note that for any $A \in \mathcal{C}$, $a_0 \cap A = \emptyset$.*

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

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

See Fig. 34 for a visual example.

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

Theorem 5.9 (Polygon Structure). For $\epsilon_\eta \geq 14\eta$ and any polygon of cuts crossed on one side with atoms $a_0 \dots a_{m-1}$ (where a_0 is the root) the following holds:

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

The interpretation of this theorem is that the structure of a polygon of cuts crossed on one side converges to the structure of an actual integral cycle as $\eta \rightarrow 0$. The proof of the theorem follows from the lemmas in the rest of this subsection.

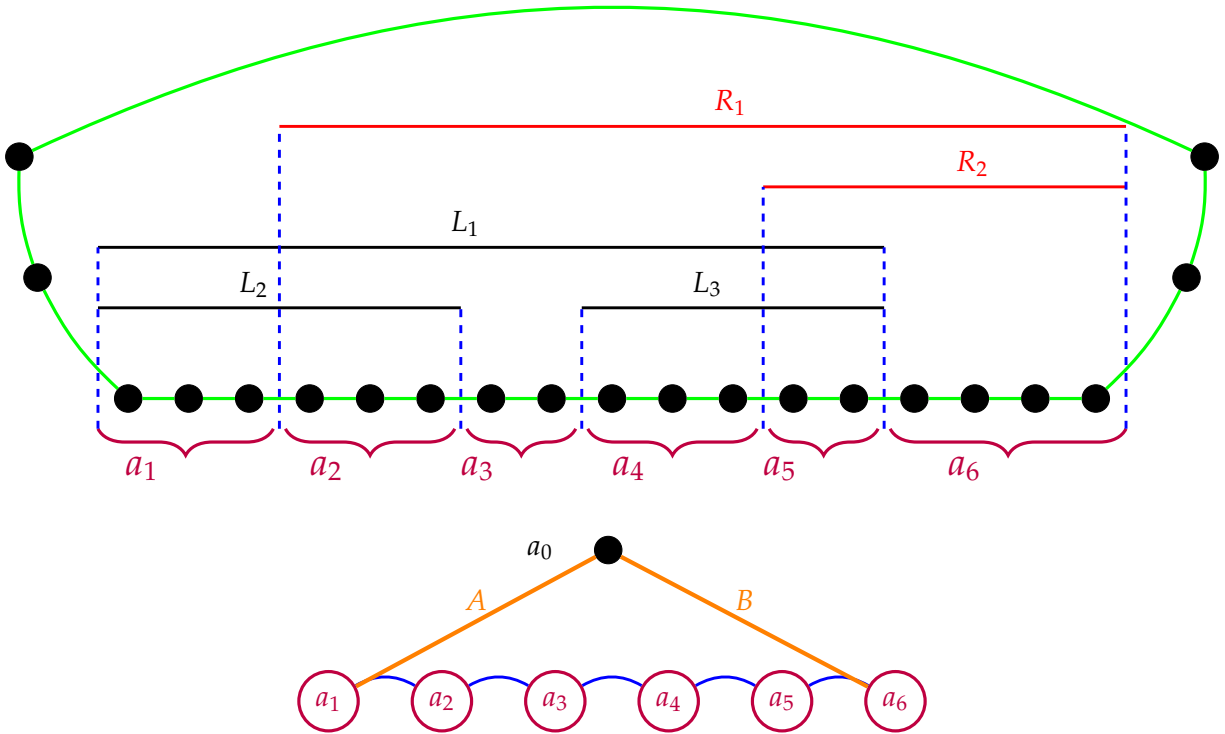


Figure 34: An example of a polygon with contracted atoms. In black are the cuts in the left polygon hierarchy, in red the cuts in the right polygon hierarchy. OPT edges around the cycle are shown in green. Here R_1 is an ancestor of R_2 , however it is not a strict ancestor of R_2 since they have the same right endpoint. L_1 is a strict ancestor and the strict parent of L_3 . By Theorem 5.9, every edge in the bottom picture represents a set of LP edges of total fraction at least $1 - \epsilon_\eta$.

Definition 5.10 (Left and Right Hierarchies). For a polygon u corresponding to a connected component \mathcal{C} of cuts crossed on one side, let \mathcal{L} (the left hierarchy) be the set of all cuts $A \in \mathcal{C}$ that are not crossed on the left. We call any cut in \mathcal{L} open on the left. Similarly, we let \mathcal{R} be the set of cuts that are open on the right. So, \mathcal{L}, \mathcal{R} is a partitioning of all cuts in \mathcal{C} .

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

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

See Fig. 34 for examples of sets and their parent/ancestor relationships.

Fact 5.11. If A, B are in the same hierarchy and they are not ancestors of each other, then $A \cap B = \emptyset$.

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

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

Lemma 5.12. For $A, B \in \mathcal{R}$ where B is a strict parent of A , there exists a cut $C \in \mathcal{L}$ that crosses both A, B . Similarly, if $A, B \in \mathcal{L}$ and B is a strict parent of A , there exists a cut $C \in \mathcal{R}$ that crosses A, B .

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

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

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

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

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

$$\ell(C_1) < \ell(C_i) \leq \ell(C_0)$$

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

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

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

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

Lemma 5.13. *Let $A, B \in \mathcal{R}$ such that $A \cap B = \emptyset$, i.e., they are not ancestors of each other. Then, they have a common ancestor, i.e., there exists a set $C \in \mathcal{R}$ such that $A, B \subseteq C$.*

Proof. WLOG assume $r(A) \leq \ell(B)$. Let C be the highest ancestor of A in the hierarchy, i.e., C has no ancestor. For the sake of contradiction suppose $B \cap C = \emptyset$ (otherwise, C is an ancestor of B and we are done). So, $r(C) \leq \ell(B)$. Consider the path of crossing cuts from C to B , say $C = C_0, \dots, C_k = B$.

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

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

Lemma 5.14. *If A is a cut in \mathcal{R} such that $r(A) < m$, then A has a strict ancestor. And, similarly, if $A \in \mathcal{L}$ satisfies $\ell(A) > 1$, then it has a strict ancestor.*

Proof. Fix a cut $A \in \mathcal{R}$. If there is a cut in $B \in \mathcal{R}$ such that $r(B) > r(A)$, then either B is a strict ancestor of A in which case we are done, or $A \cap B = \emptyset$, but then by [Lemma 5.13](#) A, B have a common ancestor C , and C must be a strict ancestor of A and we are done.

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

Corollary 5.15. *If $A \in \mathcal{C}$ has no strict ancestor, then $r(A) = m$ if $A \in \mathcal{R}$ and $\ell(A) = 1$ otherwise.*

Lemma 5.16 (Polygons are Near Minimum Cuts). $x(\delta(a_1 \cup \dots \cup a_{m-1})) \leq 2 + 4\eta$.

Proof. Let $A \in \mathcal{L}$ and $B \in \mathcal{R}$ be the unique cuts in the left/right hierarchy with no ancestors. Note that A and B are crossing (because there is a cut C that crosses A on the right, and B is an ancestor of C). Therefore, since A, B are both 2η near min cuts (with respect to x), by [Lemma 2.35](#), $A \cup B$ is a 4η near min cut. \square

Lemma 5.17 (Root Neighbors). $x(E(a_0, a_1)), x(E(a_0, a_{m-1})) \geq 1 - 2\eta$.

Proof. Here we prove $x(E(a_0, a_1)) \geq 1 - 2\eta$. One can prove $x(E(a_0, a_{m-1})) \geq 1 - 2\eta$ similarly. Let $A \in \mathcal{L}$ and $B \in \mathcal{R}$ be the unique cuts in the left/right hierarchy with no ancestors. First, observe that if $\ell(B) = 2$, then since A, B are crossing, by [Lemma 2.36](#) we have

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

as desired.

By definition of atoms, there exists a cut $C \in \mathcal{C}$ such that either $\ell(C) = 2$ or $r(C) = 2$; but if $r(C) = 2$ we must have $\ell(C) = 1$ in which case C cannot be crossed, so this does not happen. So,

we must have $\ell(C) = 2$. If $C \in \mathcal{R}$, then since C is a descendent of B , we must have $\ell(B) = 2$, and we are done by the previous paragraph.

Otherwise, suppose $C \in \mathcal{L}$. We claim that B crosses C . This is because, C is crossed on the right by some cut B' and B is an ancestor of B' , so $B \cap C \neq \emptyset$ and $C \not\subseteq B$ since $\ell(B) > 2$. Therefore, by [Lemma 2.35](#) $B \cup C$ is a 4η near min cut. Since A crosses $B \cup C$, by [Lemma 2.36](#) we have

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

as desired. \square

Lemma 5.18. *For any pair of atoms a_i, a_{i+1} where $1 \leq i \leq m - 2$ we have $x(\delta(\{a_i, a_{i+1}\})) \leq 2 + 12\eta$, so $x(E(a_i, a_{i+1})) \geq 1 - 6\eta$.*

Proof. We prove the following claim: There exists $j \leq i$ such that $x(\delta(\{a_j, \dots, a_{i+1}\})) \leq 2 + 6\eta$. Then, by a similar argument we can find $j' \geq i + 1$ such that $x(\delta(\{a_i, \dots, a_{j'}\})) \leq 2 + 6\eta$. By [Lemma 2.35](#) it follows that $x(\delta(\{a_i, a_{i+1}\})) \leq 2 + 12\eta$. Since $x(\delta(a_i)), x(\delta(a_{i+1})) \geq 2$, we have

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

But due to the bound on $x(\delta(\{a_i, a_{i+1}\}))$ we must have $x(E(a_i, a_{i+1})) \geq 1 - 6\eta$ as desired.

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

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

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

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

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

Lemma 5.19 (Atoms are Near Minimum Cuts). *For any $1 \leq i \leq m - 1$, we have $x(\delta(a_i)) \leq 2 + 14\eta$.*

Proof. By [Lemma 5.18](#), $x(\delta(\{a_i, a_{i+1}\})) \leq 2 + 12\eta$ (note that in the special case $i = m - 1$ we take the pair a_{i-1}, a_i). There must be a 2η -near minimum cut C (w.r.t., x) separating a_i from a_{i+1} . Then either $a_i = C \cap \{a_i, a_{i+1}\}$ or $a_i = \{a_i, a_{i+1}\} \setminus C$. In either case, we get $x(\delta(a_i)) \leq 2 + 14\eta$ by [Lemma 2.35](#). \square

5.5 Happy Polygons

Definition 5.20 (*A, B, C-Polygon Partition*). Let u be a polygon with atoms a_0, \dots, a_{m-1} with root a_0 where a_1, a_{m-1} are the atoms left and right of the root. The A, B, C -polygon partition of u is a partition of edges of $\delta(u)$ into sets $A = E(a_1, a_0)$ and $B = E(a_{m-1}, a_0)$, $C = \delta(u) \setminus A \setminus B$.

Note that by [Theorem 5.9](#), $x(A), x(B) \geq 1 - \epsilon_\eta$ and $x(C) \leq \epsilon_\eta$ where we set

$$\epsilon_\eta = 14\eta \tag{16}$$

as needed for [Theorem 5.9](#).

Definition 5.21 (*Leftmost and Rightmost cuts*). Let u be a polygon with atoms a_0, \dots, a_{m-1} and arcs labelled $1, \dots, m$ corresponding to a connected component \mathcal{C} of η -near minimum cuts (w.r.t., z). We call any cut $C \in \mathcal{C}$ with $\ell(C) = 1$ a leftmost cut of u and any cut $C \in \mathcal{C}$ with $r(C) = m$ a rightmost cut of u . We also call a_1 the leftmost atom of u (resp. a_{m-1} the rightmost atom).

Observe that by [Corollary 5.15](#), any cut that is not a leftmost or a rightmost cut has a strict ancestor.

Definition 5.22 (*Happy Polygon*). Let u be a polygon with polygon partition A, B, C . For a spanning tree T , we say that u is happy if

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

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

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

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

Definition 5.23 (*Relevant Cuts*). Given a polygon u corresponding to a connected component \mathcal{C} of cuts crossed on one side with atoms a_0, \dots, a_{m-1} , define a family of relevant cuts

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

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

Theorem 5.24 (*Happy Polygons and Cuts Crossed on One Side*). Let $G = (V, E, x)$ for x be an LP solution and $z = (x + \text{OPT})/2$. For a connected component \mathcal{C} of near minimum cuts of z , let u be the polygon with atoms a_0, a_1, \dots, a_{m-1} with polygon partition A, B, C . For μ an arbitrary distribution of spanning trees with marginals $x, \beta > 0$, there is a random vector $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ (as a function of $T \sim \mu$) such that for any vector $s : E \rightarrow \mathbb{R}$ where $s_e \geq -\beta x_e$ for all $e \in E$ the following holds:

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

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

Lemma 5.25 (Triangles as Degenerate Polygons). *Let $S = X \cup Y$ where X, Y, S are ϵ_η -near min cuts (w.r.t., x) and each of these sets is a contiguous interval around the OPT cycle. Then, viewing X as a_1 and Y as a_2 (and $a_0 = \overline{X \cup Y}$) the above theorem holds viewing S as a degenerate polygon.*

Proof. In this case $A = E(a_1, a_0), B = E(a_2, a_0), C = \emptyset$. For the OPT edge e^* between X, Y we define $\mathcal{I}_1(e^*)$ to be the event that at least one of $T \cap E(X), T \cap E(Y), T \cap E(S)$ is not a tree. Whenever this happens we define $s_{e^*}^* = 2.05 \cdot \beta$. If S is left-happy we need to show when $\delta(X)_T$ is odd, then $s(\delta(X)) + s^*(\delta(X)) \geq 0$. This is because when S is left-happy we have A_T is odd (and $C_T = 0$), so either $\mathcal{I}_1(e^*)$ does not happen and $\delta(X)_T$ is even, or it happens in which case $s(\delta(X)) + s^*(\delta(X)) \geq 0$ as $s(\delta(X)) \geq -(2 + 2\eta)\beta$ and $s_{e^*}^* = 2.05\beta$. Finally, observe that by [Corollary 2.29](#), $\mathbb{P}[\mathcal{I}_1(e^*)] \leq 3\epsilon_\eta$, so $\mathbb{E}[s_{e^*}^*] = 3\epsilon_\eta \cdot 2.05\beta \leq 87\eta\beta$ using $\eta < 1/100$ and ϵ_η as defined in [Eq. \(16\)](#). \square

Lemma 5.26. *For every cut $A \in \mathcal{C}$ that is not a leftmost or a rightmost cut, $\mathbb{P}[\delta(A)_T = 2] \geq 1 - 22\eta$.*

Proof. Assume $A \in \mathcal{R}$; the other case can be proven similarly. Let B be the strict parent of A . By [Lemma 5.12](#) there is a cut $C \in \mathcal{L}$ which crosses A, B on their left. It follows by [Lemma 2.35](#) that $C \setminus A, C \cap A$ are 4η near minimum cuts (w.r.t., x). So, by [Corollary 2.29](#), $\mathbb{P}[E(A \cap C, C \setminus A)_T = 1] \geq 1 - 5\eta$. On the other hand, $B \setminus (A \cup C)$ is a 6η near minimum cut and $A \setminus C, B \setminus C$ are 4η near min cuts (w.r.t., x). So, by [Corollary 2.29](#) $\mathbb{P}[E(A \setminus C, B \setminus (A \cup C))_T = 1] \geq 1 - 7\eta$.

Finally, by [Lemma 2.36](#), $x(E(A \cap C, C \setminus A)), x(E(A \setminus C, B \setminus (A \cup C))) \geq 1 - 3\eta$. Since A is a 2η near min cut (w.r.t., x), all remaining edges have fractional value at most 8η , so with probability $1 - 8\eta$, T does not choose any of them. Taking a union bound over all of these events, $\mathbb{P}[\delta(A)_T = 2] \geq 1 - 22\eta$. \square

Lemma 5.27. *For any atom $a_i \in \mathcal{C}'$ that is not the leftmost or the rightmost atom we have*

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

Proof. By [Lemma 5.18](#), $x(\delta(\{a_i, a_{i+1}\})) \leq 2 + 12\eta$, and by [Lemma 5.19](#), $x(\delta(a_{i+1})) \leq 2 + 14\eta$ (also recall by the assumption of lemma $x(\delta(a_i)) \leq 2 + 2\eta$, Therefore, by [Corollary 2.29](#),

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

where the second inequality holds similarly. Also, by [Lemma 5.18](#), $x(E(a_{i-1}, a_i)), x(E(a_i, a_{i+1})) \geq 1 - 6\eta$. Since $x(\delta(a_i)) \leq 2 + 2\eta$, $x(E(a_i, \overline{a_{i-1} \cup a_i \cup a_{i+1}})) \leq 14\eta$. So,

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

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

Let e_1^*, \dots, e_m^* be the OPT edges mapped to the arcs $1, \dots, m$ of the component \mathcal{C} respectively.

Lemma 5.28. *There is a mapping¹⁸ of cuts in \mathcal{C}' to OPT edges e_2^*, \dots, e_{m-1}^* such that each OPT edge has at most 4 cuts mapped to it, an OPT edge e^* is mapped to a cut S only if $e^* \in \delta(S)$, and every atom of the polygon in \mathcal{C}' gets mapped to two (not necessarily distinct) OPT edges.*

¹⁸Each cut will be mapped to one or two OPT edges.

Proof. Consider first the set of cuts in $\mathcal{C}'_{\mathcal{R}} := \mathcal{R} \cup \{a_i : 1 \leq i \leq m-1, z(\delta(a_i)) \leq 2 + \eta\}$ and similarly $\mathcal{C}'_{\mathcal{L}} := \mathcal{L} \cup \{a_i : 1 \leq i \leq m-1, z(\delta(a_i)) \leq 2 + \eta\}$. Observe that this is also a laminar family. Note that atoms are in both $\mathcal{C}'_{\mathcal{R}}$ and $\mathcal{C}'_{\mathcal{L}}$. We define a map from cuts in $\mathcal{C}'_{\mathcal{R}}$ to OPT edges such that every OPT edge e_2^*, \dots, e_{m-1}^* gets at most 2 cuts mapped to it. A similar argument works for cuts in $\mathcal{C}'_{\mathcal{L}}$.

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

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

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

Furthermore, we claim every cut $A \in \mathcal{C}'_{\mathcal{R}}$ gets mapped to at least one OPT edge. For the sake of contradiction let $A \in \mathcal{C}'_{\mathcal{R}}$ be a cut that is not mapped to any OPT edge. First note that a_1 is mapped to edge e_2^* (in both hierarchies) and a_{m-1} is mapped to edge e_{m-1}^* . Otherwise, if $A \in \mathcal{R}$, $\ell(A) \neq 1$. Furthermore, if $A \in \mathcal{R}$ and $r(A) = m$, then A is definitely the largest cut with left endpoint $\ell(A)$. So assume, $1 < \ell(A) < r(A) < m$. Let $B = \operatorname{argmax}_{B \in \mathcal{C}'_{\mathcal{R}}: \ell(B)=\ell(A)} |B|$ and let $C = \operatorname{argmax}_{B \in \mathcal{C}'_{\mathcal{R}}: r(B)=r(A)} |B|$. Since A is not mapped to any OPT edge but B, C are mapped by above definition, we must have $B, C \neq A$. But that implies $A \subsetneq B, C$. And this means B, C cross; but this is a contradiction with \mathcal{R} being a laminar family. \square

Definition 5.29 (Happy Cut). *We say a leftmost cut $L \in \mathcal{L}$ is happy if*

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

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

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

Lemma 5.30. *For every leftmost or rightmost cut A in u that is an η -near min cut w.r.t. z , $\mathbb{P}[A \text{ happy}] \geq 1 - 10\eta$, and for the leftmost atom a_1 (resp. rightmost atom a_{m-1}), if it is an η -near min cut then $\mathbb{P}[a_1 \text{ happy}] \geq 1 - 24\eta$ (resp. $\mathbb{P}[a_{m-1} \text{ happy}] \geq 1 - 24\eta$).*

Proof. Recall that if A is a η -near min cut w.r.t. z then it is a 2η -near min cut w.r.t. x . Also, recall for a cut $L \in \mathcal{L}$, L_R is the near minimum cut crossing L on the right that minimizes the intersection (see [Definition 5.3](#)). We prove this for the leftmost cuts and the leftmost atom; the other case can be proven similarly. Consider a cut $L \in \mathcal{L}$. Since by [Lemma 2.35](#) $L_R \cap L, L_R \setminus L$ are 4η near min cuts (w.r.t., x) and L_R is a 2η near min cut, by [Corollary 2.29](#), $\mathbb{P}[E(L_R \cap L, L_R \setminus L)_T = 1] \geq 1 - 5\eta$. On the other hand, by [Lemma 2.36](#), $x(E(L_R \cap L, L_R \setminus L)) \geq 1 - \eta$, and by [Lemma 5.17](#), $x(E(L, a_0)) \geq 1 - 2\eta$. It follows that

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

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

Now consider the atom a_1 , and suppose it is an η near min cut. By [Lemma 5.18](#), $x(\delta(\{a_1, a_2\})) \leq 2 + 12\eta$ and by [Lemma 5.19](#), $x(\delta(a_2)) \leq 2 + 14\eta$. Therefore, by [Corollary 2.29](#), $\mathbb{P}[E(a_1, a_2)_T = 1] \geq$

$1 - 14\eta$. On the other hand, by [Lemma 5.18](#), $x(E(a_1, a_2)) \geq 1 - 6\eta$ and by [Lemma 5.17](#), $x(E(a_1, a_0)) \geq 1 - 2\eta$. Therefore,

$$x(E(a_1, a_3 \cup \dots \cup a_{m-1})) \leq 2 + 2\eta - (1 - 6\eta) - (1 - 2\eta) \leq 10\eta.$$

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

Proof of [Theorem 5.24](#). Consider an OPT edge e_i^* for $1 < i < m$. For the at most four cuts mapped to e_i^* in [Lemma 5.28](#), we define the following three events:

- i) A leftmost cut assigned to e_i^* is not happy. (Equivalently, a leftmost cut $L \in \mathcal{L} \cap \mathcal{C}'$ with $r(L) = i$ is not happy.)
- ii) A rightmost cut assigned to e_i^* is not happy. (Equivalently, a rightmost cut $R \in \mathcal{R} \cap \mathcal{C}'$ with $l(R) = i$ is not happy.¹⁹)
- iii) A cut which is not leftmost or rightmost assigned to e_i^* is odd.

Observe that the cuts in (i) and (ii) are assigned to e_i^* in [Lemma 5.28](#). We say an atom a is singly-mapped to e_i^* if in the matching a is only mapped to e_i^* once, otherwise we say it is doubly-mapped to e_i^* .

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

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

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

$$s(\delta(S)) + s^*(\delta(S)) \geq -x(\delta(S))\beta + s^*(e_i^*) \geq -(2 + 2\eta)\beta + 2.05\beta \geq 0,$$

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

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

It remains to upper bound $\mathbb{E}[s^*(e_i^*)]$ for $1 < i < m$. By [Lemma 5.28](#) at most four cuts are mapped to e_i^* . Then, either there is an atom which is doubly-mapped to e_i^* or there is not.

¹⁹Note in the special case that $i = 2$, L in (i) will be the leftmost atom if it is a near min cut, and similarly in (ii) when $i = m - 1$, R will be the rightmost atom if it is a near min cut.

First suppose exactly one atom is doubly-mapped to e_i^* . Then there are at most three cuts mapped to e_i^* , including that atom. The probability of an event of type (i) or (ii) occurring for the leftmost or rightmost atom is at most $1 - 24\eta$ by [Lemma 5.30](#). Atoms which are not leftmost or rightmost are even with probability at least $1 - 42\eta$ by [Lemma 5.27](#). Therefore, in the worst case, the doubly-mapped atom is not leftmost or rightmost. For the remaining two cuts, leftmost and rightmost cuts are happy with probability at least $1 - 10\eta$ by [Lemma 5.30](#), and (non-atom) non leftmost/rightmost cuts are even with probability at least $1 - 22\eta$ by [Lemma 5.26](#). Therefore in the worst case the remaining two (non-atom) cuts mapped to e_i^* are not leftmost/rightmost. Therefore, if an atom is doubly-mapped to e_i^* ,

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

Note if two atoms are doubly-mapped to e_i^* ,

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

Otherwise, any atoms mapped to e_i^* are singly-mapped. In this case, if only an atom cut is odd/unhappy, we set $s^*(e_i^*) = 2.05\beta/2$. The probability of an event of type (i) or (ii) occurring for the leftmost or rightmost atom is at most $1 - 24\eta$ by [Lemma 5.30](#), so we can bound the contribution of this event to $\mathbb{E} [s^*(e_i^*)]$ by $24\eta \cdot 2.05\beta/2$. Atoms which are not leftmost or rightmost are even with probability at least $1 - 42\eta$ by [Lemma 5.27](#), and so we can bound their contribution by $42\eta \cdot 2.05\beta/2$. Therefore, in the worst case four non-leftmost/rightmost *non-atom* cuts are mapped to e_i^* , in which case,

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

as desired. □

5.6 Hierarchy of Cuts and Proof of [Theorem 5.6](#)

Definition 5.31 (Hierarchy). *For an LP solution x^0 with support $E_0 = E \cup \{e_0\}$ and x be x^0 restricted to E , a hierarchy \mathcal{H} is a laminar family of ϵ_η -near min cuts of $G = (V, E, x)$ with root $V \setminus \{u_0, v_0\}$, where every cut $S \in \mathcal{H}$ is either a polygon cut (including triangles) or a degree cut and $u_0, v_0 \notin S$. Furthermore, every cut S is a union of its children. For any (non-root) cut $S \in \mathcal{H}$, define the parent of S , $\mathfrak{p}(S)$, to be the smallest cut $S' \in \mathcal{H}$ such that $S \subsetneq S'$.*

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

- $A = E(\bar{S}, u_1), B = E(u_{m-1}, \bar{S})$ satisfy $x(A), x(B) \geq 1 - \epsilon_\eta$.
- For any $1 \leq i < m - 1$, $x(E(u_i, u_{i+1})) \geq 1 - \epsilon_\eta$.
- $C = \cup_{i=2}^{m-2} E(u_i, \bar{S})$ satisfies $x(C) \leq \epsilon_\eta$.

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

*We abuse notation, and for an (LP) edge $e = (u, v)$ that is not a neighbor of u_0, v_0 , let $\mathfrak{p}(e)$ denote the smallest²⁰ cut $S' \in \mathcal{H}$ such that $u, v \in S'$. We say edge e is a **bottom edge** if $\mathfrak{p}(e)$ is a polygon cut and we say it is a **top edge** if $\mathfrak{p}(e)$ is a degree cut.*

²⁰in the sense of the number of vertices that it contains

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

The following observation is immediate from the above definition.

Observation 5.32. *For any polygon cut $S \in \mathcal{H}$, and any cut $S' \in \mathcal{H}$ which is a descendant of S let $D = \delta(S') \cap \delta(S)$. If $D \neq \emptyset$, then exactly one of the following is true: $D \subseteq A$ or $D \subseteq B$ or $D \subseteq C$.*

We now introduce the “main payment theorem” which is the other key technical piece of the proof.

Theorem 5.33 (Main Payment Theorem). *For an LP solution x^0 and x be x^0 restricted to E and a hierarchy \mathcal{H} for some $\epsilon_\eta \leq 10^{-10}$ and any $\beta > 0$, the maximum entropy distribution μ with marginals x satisfies the following:*

- i) *There is a set of good edges $E_g \subseteq E \setminus \delta(\{u_0, v_0\})$ such that any bottom edge e is in E_g and for any (non-root) $S \in \mathcal{H}$ such that $\rho(S)$ is a degree cut, we have $x(E_g \cap \delta(S)) \geq 3/4$.*
- ii) *There is a random vector $s : E_g \rightarrow \mathbb{R}$ (as a function of $T \sim \mu$) such that for all e , $s_e \geq -x_e\beta$ (with probability 1), and*
- iii) *If a polygon cut u with polygon partition A, B, C is not left happy, then for any set $F \subseteq E$ with $\rho(e) = u$ for all $e \in F$ and $x(F) \geq 1 - \epsilon_\eta/2$, we have*

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

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

- iv) *For every cut $S \in \mathcal{H}$ such that $\rho(S)$ is not a polygon cut, if $\delta(S)_T$ is odd, then $s(\delta(S)) \geq 0$.*
- v) *For a good edge $e \in E_g$, $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ (see Eq. (38) for definition of ϵ_P).*

The above theorem is the main part of the paper in which we use that μ is a SR distribution. See Section 8 for the proof. We use this theorem to construct a random vector s such that essentially for all cuts $S \in \mathcal{H}$ in the hierarchy $z/2 + s$ is feasible; furthermore for a large fraction of “good” edges we have that $\mathbb{E}[s_e]$ is negative and bounded away from 0.

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

In the rest of this section we use the above theorem to prove Theorem 5.6. We start by explaining how to construct \mathcal{H} . Given the vector $z = (x + OPT)/2$ run the following procedure on the OPT cycle with the family of η -near minimum cuts of z that are crossed on at most one side:

For every connected component \mathcal{C} of η near minimum cuts (w.r.t., z) crossed on at most one side, if $|\mathcal{C}| = 1$ then add the unique cut in \mathcal{C} to the hierarchy. Otherwise, \mathcal{C} corresponds to a

polygon u with atoms a_0, \dots, a_{m-1} (for some $m > 3$). Add a_1, \dots, a_{m-1} ²¹ and $\cup_{i=1}^{m-1} a_i$ to \mathcal{H} . Since every vertex except u_0, v_0 has degree 2, they all appear in the hierarchy as singletons. Therefore, every set in the hierarchy is the union of its children. Note that since $z(\delta(\{u_0, v_0\})) = 2$, the root of the hierarchy is always $V \setminus \{u_0, v_0\}$.

Now, we name every cut in the hierarchy. For a cut S if there is a connected component of at least two cuts with union equal to S , then call S a polygon cut with the A, B, C partitioning as defined in [Definition C.4](#). If S is a cut with exactly two children X, Y in the hierarchy, then also call S a polygon cut²², $A = E(X, \bar{X} \setminus Y)$, $B = E(Y, \bar{Y} \setminus X)$ and $C = \emptyset$. Otherwise, call S a degree cut.

Fact 5.34. *The above procedure produces a valid hierarchy for $\epsilon_\eta \geq 14\eta$.*

Proof. First observe that whenever $|\mathcal{C}| = 1$ the unique cut in \mathcal{C} is a 2η near min cut (w.r.t, x) which is not crossed. For a polygon cut S in the hierarchy, by [Lemma 5.16](#), the set S is a ϵ_η near min cut w.r.t., x . If S is an atom of a polygon, then by [Lemma 5.19](#) S is a ϵ_η near min cut.

Now, it remains to show that for a polygon cut S we have a valid ordering u_1, \dots, u_k of cuts in $\mathcal{A}(S)$. If S is a non-triangle polygon cut, the u_1, \dots, u_k are exactly atoms of the polygon of S and $x(A), x(B) \geq 1 - \epsilon_\eta$ and $x(C) \leq \epsilon_\eta$ and $x(E(u_i, u_{i+1})) \geq 1 - \epsilon_\eta$ follow by [Theorem 5.9](#). For a triangle cut $S = X \cup Y$ because S, X, Y are ϵ_η -near min cuts (by the previous paragraph), we get $x(A), x(B) \geq 1 - \epsilon_\eta$ as desired, by [Lemma 2.37](#). Finally, since $x(\delta(X)), x(\delta(Y)) \geq 2$ we have $x(E(X, Y)) \geq 1 - \epsilon_\eta$. \square

The following observation is immediate:

Observation 5.35. *Each cut $S \in \mathcal{H}$ corresponds to a contiguous interval around OPT cycle. For a polygon u (or a triangle) with atoms a_0, \dots, a_{m-1} for $m \geq 3$ we say an OPT edge e^* is interior to u if $e^* \in E^*(a_i, a_{i+1})$ for some $1 \leq i \leq m - 2$. Any OPT edge e^* is interior to at most one polygon.*

Theorem 5.6. *Let x^0 be a feasible solution of the 2 with support $E_0 = E \cup \{e_0\}$ and x be x^0 restricted to E . Let μ be the max entropy distribution with marginals x . For $\eta \leq 10^{-12}$, $\beta > 0$, there is a set $E_g \subset E \setminus \delta(\{u_0, v_0\})$ of good edges and two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E^* \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$) such that*

- (i) *For each edge $e \in E_g$, $s_e \geq -x_e\beta$ and for any $e \in E \setminus E_g$, $s_e = 0$.*
- (ii) *For each η -near-min-cut S w.r.t. z , if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.*
- (iii) *We have $\mathbb{E}[s_e] \leq -\epsilon_p\beta x_e$ for all edges $e \in E_g$ and $\mathbb{E}[s_{e^*}^*] \leq 218\eta\beta$ for all OPT edges $e^* \in E^*$. for ϵ_p defined in (38).*
- (iv) *For every η -near minimum cut S of z crossed on (at most) one side such that $S \neq V \setminus \{u_0, v_0\}$, $x(\delta(S) \cap E_g) \geq 3/4$.*

Proof. For ϵ_η as in [Eq. \(16\)](#), let E_g, s be as defined in [Theorem D.2](#), and let $s_{e_0} = \infty$. Also, let s^* be the sum of the s^* vectors from [Theorem 5.2](#) and [Theorem 5.24](#). (i) follows (ii) of [Theorem D.2](#). $\mathbb{E}[s_{e^*}^*] \leq 218\eta\beta$ follows from [Theorem 5.2](#) and [Theorem 5.24](#) and the fact that every OPT edge

²¹Notice that an atom may already correspond to a connected component, in such a case we do not add it in this step.

²²Think about such set as a *degenerate* polygon with atoms $a_1 := X, a_2 := Y, a_0 := \bar{X} \cup \bar{Y}$. So, for the rest of this section we call them triangles and in later section we just think of them as polygon cuts.

is interior to at most one polygon. Also, $\mathbb{E}[s_e] \leq -\epsilon_p \beta x_e$ for edges $e \in E_g$ follows from (v) of [Theorem D.2](#).

Now, we verify (iv): For any (non-root) cut $S \in \mathcal{H}$ such that $p(S)$ is not a polygon cut $x(\delta(S) \cap E_g) \geq 3/4$ by (i) of [Theorem 5.33](#). The only remaining η -near minimum cuts are sets S which are either atoms or near minimum cuts in the component \mathcal{C} corresponding to a polygon u . So, by [Lemma 2.37](#), $x(\delta(S) \cap \delta(u)) \leq 1 + \epsilon_\eta$. By (i) of [Theorem D.2](#) all edges in $\delta(S) \setminus \delta(u)$ are in E_g . Therefore, $x(\delta(S) \cap E_g) \geq 1 - \epsilon_\eta \geq 3/4$.

It remains to verify (ii): We consider 4 groups of cuts:

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

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

Type 3: Near minimum cuts S crossed on both sides. Then, the inequality follows by [Theorem 5.2](#) and the fact that $s_e \geq -\beta x_e$ for all $e \in E$.

Type 4: Near minimum cuts S that are crossed on one side (and not in \mathcal{H}) or $S \in \mathcal{H}$ and $p(S)$ is a (non-triangle) polygon cut. In this case S must be an atom or a η -near minimum cut (w.r.t. z) in some polygon $u \in \mathcal{H}$. If S is not a leftmost cut/atom or a rightmost cut/atom, then the inequality follows by [Theorem 5.24](#). Otherwise, say S is a leftmost cut. If u is left-happy then by [Theorem 5.24](#) the inequality is satisfied. Otherwise, for $F = \delta(S) \setminus \delta(u)$, by [Lemma 2.37](#), we have $x(F) \geq 1 - \epsilon_\eta/2$. Therefore, by (iii) of [Theorem D.2](#) we have

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

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

Type 5: Near min cut S is the leftmost atom or the rightmost atom of a triangle u . This is similar to the previous case except we use [Lemma C.13](#) to argue that the inequality is satisfied when u is left happy. \square

5.7 Hierarchy Notation

In the rest of the paper we will not work with z , OPT edges, or the notion of polygons. So, practically, by [Definition 5.31](#), from now on, a reader can just think of every polygon as a triangle. In the rest of the paper we adopt the following notation.

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

Definition 5.36 (Edge Bundles, Top Edges, and Bottom Edges). *For every degree cut S and every pair of atoms $u, v \in \mathcal{A}(S)$, we define a **top edge bundle** $\mathbf{f} = (u, v)$ such that*

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

Note that in the above definition, u', v' are actual vertices of G .

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

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

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

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

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

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

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

6 Probabilistic statements

We first prove some useful properties of Bernoulli-Sum random variables and random spanning trees. We then have three important sections: [Section 6.3](#) provides a very general tool for proving lower bounds on the probability of certain events in any SR distribution, [Section 6.4](#) constructs and proves bounds on the events in which bottom edges can reduce, and finally [Section 6.5](#) and [Section 6.6](#) construct and prove bounds on the events in which top edges can reduce.

6.1 Properties of Bernoulli-Sum Random Variables

Lemma 6.1. *Let p_0, \dots, p_n be a log-concave sequence. If for some i , $\gamma p_i \geq p_{i+1}$ for some $\gamma < 1$, then,*

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

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

Proof. Since we have a log-concave sequence we can write

$$\frac{1}{\gamma} \leq \frac{p_i}{p_{i+1}} \leq \frac{p_{i+1}}{p_{i+2}} \leq \dots \quad (17)$$

Since all of the above ratios are at least $1/\gamma$, for all $l \geq 1$ we can write

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

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

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

□

Corollary 6.2. *Let X be a BS(q) random variable such that $\mathbb{P}[X = k] \geq 1 - \epsilon$ for some integer $k \geq 1$, $\epsilon < 1/10$. Then, $k(1 - \epsilon) \leq q \leq k(1 + \epsilon) + 3\epsilon$.*

Proof. The left inequality simply follows since $X \geq 0$. Since $\mathbb{P}[X = k+1] \leq \epsilon$, we can apply [Lemma 6.1](#) with $\gamma = \epsilon/(1 - \epsilon)$ to get

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

Therefore,

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

as desired. □

Fact 6.3. For integers $k < t$ and $k - 1 \leq p \leq k$,

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

Proof. We show that the LHS is a decreasing function of t . Since \ln is monotone, it is enough to show

$$\begin{aligned} 0 &\geq \partial_t \ln(\text{LHS}) = \partial_t \left(\sum_{i=1}^{k-1} \ln(1 - i/t) + (t - k) \ln(1 - p/t) \right) \\ &= \frac{1}{t^2} \sum_{i=1}^{k-1} \frac{1}{\frac{1}{i} - \frac{1}{t}} + \ln(1 - p/t) + \frac{(t - k)p}{t(t - p)} \end{aligned}$$

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

$$\begin{aligned} 0 &\geq -\frac{k-1}{t} - \ln(1 - \frac{k-1}{t}) + \ln(1 - p/t) + \frac{(t-k)p}{t(t-p)} + \frac{1}{t^2(\frac{1}{k-1} - \frac{1}{t})} \\ &= \ln \frac{t-p}{t-k+1} + \frac{p-k}{t-p} + \frac{1}{t} + \frac{k-1}{t(t-k+1)} \end{aligned}$$

Rearranging, it is equivalent to show

$$\ln(1 + \frac{p-k+1}{t-p}) \geq \frac{p-k}{t-p} + \frac{1}{t-k+1}$$

Since $p > k-1$, using Taylor series of \ln , to prove the above it is enough to show

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

This is equivalent to show

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

Finally the latter holds because $(t-k+1)(p-k+1) \leq (t-k+1) \leq 2(t-p)$ where we use $t \geq k+1$ and $p \leq k$. \square

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

Lemma 6.4. Let X be a Bernoulli sum $BS(p)$ for some n . For any integer $k \geq 0$ such that $k-1 < p < k+1$, the following holds true

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

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

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

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

where in the inequality we used [Fact 6.3](#) (also note if $n = k$ the inequality follows from Stirling's formula and that $p \geq k - 1$). If $k < p < k + 1$, then as above

$$\begin{aligned} \mathbb{P}[X = k] &= \prod_{i=1}^{k-1} (1 - i/n) \frac{p^k}{k!} (1 - p/n)^{n-p} (1 - p/n)^{p-k} \\ &\geq \prod_{i \geq k}^{k-1} (1 - i/n) \frac{p^k}{k!} (1 - p/n)^{n-k} (1 - p/n)^{p-k} \geq \text{Poi}(p, k) (1 - p/n)^{p-k}, \end{aligned}$$

where we used [Fact 6.3](#) in the last inequality. □

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

Lemma 6.5. *Let X be a Bernoulli sum $BS(p)$, where for some integer $k = \lceil p \rceil$, Then,*

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

where the minimum is over all non-negative integers $\ell \leq p$.

Proof. Suppose that X is a $BS(p)$ with n Bernoullis with probabilities p_1, \dots, p_n . If $p - 1 < k - 1 < p$, by [[Hoe56](#), Thm 4, (25)],

$$\mathbb{P}[X \leq k - 1] \leq \max_{0 \leq \ell < p} \sum_{i=0}^{k-1-\ell} \binom{n - \ell}{i} q^i (1 - q)^{n-\ell-i} \quad (18)$$

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

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

$$\max_{p_1 \dots p_n} \mathbb{P}[X \leq k - 1] \leq \max_{q_1 \dots q_m} \mathbb{P}[Y \leq k - 1]$$

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

Therefore, the upper bound in (18) obtained by taking the limit as n goes to infinity applies, from which it follows that

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

and therefore

$$\mathbb{P}[X \geq k] \geq \min_{0 \leq \ell < p} \text{Poi}(p - \ell, \geq k - \ell).$$

□

6.2 Random Spanning Trees

Lemma 6.6. *Let $G = (V, E, x)$, and let μ be a λ -uniform random spanning tree distribution with marginals x . For any edge $e = (u, v)$ and any vertex $w \neq u, v$ we have*

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

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

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

By [Theorem 2.24](#), $\mu_{in} \preceq \mu_{out}$ so there exists a coupling $\rho : 2^{E'} \times 2^{E'}$ between them such that for any T_{in}, T_{out} such that $\rho(T_{in}, T_{out}) > 0$, the tree T_{out} has exactly one more edge than T_{in} . Also, observe that T_{out} is always a spanning tree whereas $T_{in} \cup \{e\}$ is a spanning tree. The added edge (i.e., the edge in $T_{out} \setminus T_{in}$) is always along the unique path from u to v in T_{out} .

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

Say x_e is the marginal of e . Then,

$$\begin{aligned} \mathbb{E} [W_T] &= \mathbb{E} [W_T | e \notin T] (1 - x_e) + \mathbb{E} [W_T | e \in T] x_e \\ &= \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out}) W_0 (1 - x_e) + \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out}) W_i x_e \\ &= \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out}) ((1 - x_e) W_0 + x_e W_i), \end{aligned} \tag{19}$$

where we write W_i/W_0 instead of $W_{T_{in}}/W_{T_{out}}$

$$\begin{aligned} \mathbb{E} [W_T | e \notin T] &= \sum_{T_{in}, T_{out}} \rho(T_{in}, T_{out}) W_0 \\ &= \sum_{T_{in}, T_{out}: w \in P_{u,v}(T_{out})} \rho(T_{in}, T_{out}) W_0 + \sum_{T_{in}, T_{out}: w \notin P_{u,v}(T_{out})} \rho(T_{in}, T_{out}) W_0 \\ &\leq \sum_{T_{in}, T_{out}: w \in P_{u,v}(T_{out})} \rho(T_{in}, T_{out}) (x_e (W_i + 1) + (1 - x_e) W_0) \\ &\quad + \sum_{T_{in}, T_{out}: w \notin P_{u,v}(T_{out})} \rho(T_{in}, T_{out}) (x_e W_i + (1 - x_e) W_0) \\ &= \mathbb{E} [W_T] + \sum_{T_{in}, T_{out}: w \in P_{u,v}(T_{out})} \rho(T_{in}, T_{out}) x_e \\ &= \mathbb{E} [W_T] + \sum_{T_{out}: w \in P_{u,v}(T_{out})} \mu_{out}(T_{out}) x_e \\ &= \mathbb{E} [W_T] + \mathbb{P} [w \in P_{u,v} | e \text{ out}] \cdot \mathbb{P} [e \text{ in}] \end{aligned}$$

where in the inequality we used the following: When $w \notin P_{u,v}(T_{out})$ we have $W_i = W_0$ and when $w \in P_{u,v}(T_{out})$ we have $W_0 \leq W_i + 1$. Finally, in the third to last equality we used [\(19\)](#). \square

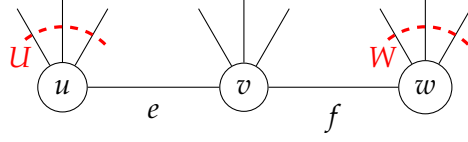


Figure 35: Setting of Lemma 6.7

Lemma 6.7. Let $G = (V, E, x)$, and let μ be a λ -uniform spanning tree distribution with marginals x . For any pair of edges $e = (u, v), f = (v, w)$ such that $|\mathbb{P}[e] - 1/2|, |\mathbb{P}[f] - 1/2| < \epsilon$ (see Fig. 35), if $\epsilon < 1/1000$, then

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

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

Proof. All probabilistic statements are with respect to ν so we drop the subscript. First, by Lemma 6.6, and negative association we can write,

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

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

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

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

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

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

So, we have

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

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

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

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

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

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

6.3 Gurvits' Machinery and Generalizations

The following is the main result of this subsection.

Proposition 6.8. *Given a SR distribution $\mu : 2^{[n]} \rightarrow \mathbb{R}_+$, let A_1, \dots, A_m be random variables corresponding to the number of elements sampled from m disjoint sets, and let integers $n_1, \dots, n_m \geq 0$ be such that for any $S \subseteq [m]$,*

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

it follows that,

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

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

We remark that in applications of the above statement, it is enough to know that for any set $S \subseteq [m]$, $\sum_{i \in S} n_i - 1 < \mathbb{E} [\sum_{i \in S} A_i] < \sum_{i \in S} n_i + 1$. Because, then by [Lemma 6.4](#) we can prove a lower bound on the probability that $\sum_{i \in S} A_i = \sum_{i \in S} n_i$.

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

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

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

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

$$\mathbb{P} [A_k = n_k | A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E}] \geq e^{2^{m-k+1}} \frac{1}{\max\{n_k, n_1 + \dots + n_{k-1}\} + 1} \quad (20)$$

By the following [Claim 6.9](#),

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

So, (20) simply follows by [Lemma 6.10](#). Now we prove this claim.

Claim 6.9. *Let $[k] := \{1, \dots, k\}$. For any $2 \leq k \leq m$, and any set $S \subsetneq [k]$,*

$$\begin{aligned} \mathbb{P} \left[\sum_{i \in S} A_i \geq \sum_{i \in S} n_i \mid A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E} \right] &\geq e^{2^{m-k+1}}, \\ \mathbb{P} \left[\sum_{i \in S} A_i \leq \sum_{i \in S} n_i \mid A_{k+1} = n_{k+1}, \dots, A_m = n_m, \mathcal{E} \right] &\geq e^{2^{m-k+1}} \end{aligned}$$

Proof. We prove by induction. First, notice for $k = m$ the statement holds just by lemma's assumption and [Lemma 6.11](#). Now, suppose the statement holds for $k + 1$. Now, fix a set $S \subsetneq [k]$. Let $\bar{S} = [k] \setminus S$. Define $A = \sum_{i \in S} A_i$ and $B = \sum_{i \in \bar{S}} A_i$, and similarly define n_A, n_B . By the induction hypothesis,

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

The same statement holds for events $A \geq n_A, B \leq n_B, B \geq n_B, A + B \geq n_A + n_B, A + B \leq n_A + n_B$. Let \mathcal{E}_{k+1} be the event $A_{k+2} = n_{k+2}, \dots, A_m = n_m, \mathcal{E}$. Note that conditioned on \mathcal{E}_{k+1} , $A + B = n_A + n_B$ if and only if $A_{k+1} = n_{k+1}$. By [Lemma 6.10](#), $\mathbb{P}[A + B = n_A + n_B | \mathcal{E}_{k+1}] > 0$. Therefore, by [Lemma 6.11](#),

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

as desired. \square

This finishes the proof of [Proposition 6.8](#) \square

Lemma 6.10. *Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a d -homogeneous SR distribution. If for an integer $0 \leq k \leq d$, $\mathbb{P}_{S \sim \mu}[|S| \geq k] \geq \epsilon$ and $\mathbb{P}_\mu[|S| \leq k] \geq \epsilon$. Then,*

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

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

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

Suppose s_i is the mode. It follows that there is $i \leq j \leq k-1$ such that $\frac{s_j}{s_{j+1}} \geq \left(\frac{s_i}{s_k}\right)^{1/(k-i)}$. So, by [Lemma 6.1](#),

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

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

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

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

Lemma 6.11. *Given a strongly Rayleigh distribution $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$, let A, B be two (nonnegative) random variables corresponding to the number of elements sampled from two disjoint sets such that $\mathbb{P}[A + B = n] > 0$ where $n = n_A + n_B$. Then,*

$$\mathbb{P}[A \geq n_A | A + B = n] = \mathbb{P}[B \leq n_B | A + B = n] \geq \mathbb{P}[A \geq n_A] \mathbb{P}[B \leq n_B], \quad (21)$$

$$\mathbb{P}[A \leq n_A | A + B = n] = \mathbb{P}[B \geq n_B | A + B = n] \geq \mathbb{P}[A \leq n_A] \mathbb{P}[B \geq n_B]. \quad (22)$$

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

$$\begin{aligned} & \mathbb{P}[A \leq n_A, A + B \geq n] + \mathbb{P}[B \geq n_B, A + B < n] \\ &= \mathbb{P}[B \geq n_B, A \leq n_A, A + B \geq n] + \mathbb{P}[A \leq n_A, B \geq n_B, A + B < n] \\ &= \mathbb{P}[B \geq n_B, A \leq n_A] \geq \mathbb{P}[B \geq n_B] \mathbb{P}[A \leq n_A] =: \alpha, \end{aligned}$$

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

$$\begin{aligned} \mathbb{P}[A \leq n_A | A + B = n] &\geq \mathbb{P}[A \leq n_A | A + B \geq n] \\ \mathbb{P}[B \geq n_B | A + B = n] &\geq \mathbb{P}[B \geq n_B | A + B < n] \end{aligned}$$

Note that in the special case that $A + B < n$ never happens, the lemma holds trivially. \square

Combining the previous two lemmas, we get

Corollary 6.12. *Let $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ be a SR distribution. Let A, B be two random variables corresponding to the number of elements sampled from two disjoint sets of elements such that $A \geq k_A$ with probability 1 and $B \geq k_B$ with probability 1. If $\mathbb{P}[A \geq n_A], \mathbb{P}[B \geq n_B] \geq \epsilon_1$ and $\mathbb{P}[A \leq n_A], \mathbb{P}[B \leq n_B] \geq \epsilon_2$, then, letting $n'_A = n_A - k_A, n'_B = n_B - k_B$,*

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

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

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

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

To get the first statement, we construct a new SR distribution from μ as follows. First, we symmetrize g_μ by setting all $x_a \in A$ to x and all $x_b \in B$ to y ; call the resulting polynomial q_μ . Then, notice $q'_\mu = q_\mu / (x^{k_A} y^{k_B})$ is real stable. Therefore, we can apply the above corollary to a distribution with generating polynomial q'_μ .²³

To get the second statement, notice that since the distribution of A is unimodal,

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

²³To be precise, we apply the above corollary to the polarization of q'_μ , where x, y are polarized by a disjoint set of variables of size equal to their maximum degree.

6.4 Max Flow

This proposition and the max flow event are crucially used in the analysis of the bottom-bottom case in the payment theorem ([Theorem D.2](#)). See [Example 4.10](#) and the preceding discussion for more high-level intuition. The main consequences of this section are [Corollary 6.17](#) and [Corollary 6.18](#).

Proposition 6.13. *Let $\mu : 2^E \rightarrow \mathbb{R}_{\geq 0}$ be a homogeneous SR distribution. For any $330\epsilon < \zeta < 0.002$ and disjoint sets $A, B \subseteq E$ such that $\mathbb{E}[A_T], \mathbb{E}[B_T] \in [1 - \epsilon, 1 + \epsilon]$ (where $T \sim \mu$) there is an event $\mathcal{E}_{A,B}(T)$ such that $\mathbb{P}[\mathcal{E}_{A,B}(T)] \geq 0.0246\zeta^2(1 - \zeta/2.1 - \epsilon)$ and it satisfies the following three properties.*

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

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

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

Lemma 6.14. *Let $\mu : 2^E \rightarrow \mathbb{R}_{\geq 0}$ be a homogeneous SR distribution. Let $A, B \subseteq E$ be two disjoint sets such that $\mathbb{E}[A_T], \mathbb{E}[B_T] \in [1 - \epsilon, 1 + \epsilon]$ (where $T \sim \mu$), $A' \subset A$ and $B' \subseteq B$ and $\mathbb{E}[A'_T \cup B'_T] \geq 1 + \alpha$ for some $\alpha > 100\epsilon$. If $\alpha < 0.001$, we have*

$$\mathbb{P}[A'_T = B'_T = A_T = B_T = 1] \geq 0.11\alpha^3.$$

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

$$\mathbb{E}_\nu[A'_T], \mathbb{E}_\nu[B'_T] \in [\alpha - \epsilon, 2 + 3\epsilon - \alpha]. \quad (23)$$

- **Case 1:** $\mathbb{E}_\nu[A'_T + B'_T] > 1.5$. Since $\mathbb{E}_\nu[A'_T + B'_T] \leq 2 + 2\epsilon$, by [Lemma 6.4](#), $\mathbb{P}_\nu[A'_T + B'_T = 2] \geq 0.25$. Furthermore,

$$\begin{aligned} \mathbb{P}_\nu[A'_T \geq 1], \mathbb{P}_\nu[B'_T \geq 1] &\geq 1 - e^{-(\alpha - \epsilon)} \geq 0.98\alpha && \text{(Lemma 6.5, } \alpha < 0.001) \\ \mathbb{P}_\nu[A'_T \leq 1], \mathbb{P}_\nu[B'_T \leq 1] &\geq \alpha/2 - 1.5\epsilon && \text{(Markov's Inequality)} \end{aligned}$$

Therefore, by [Corollary 6.12](#) and using $\alpha \leq 0.001$, $\mathbb{P}[A'_T = 1 | A'_T + B'_T = 2] \geq 0.45\alpha^2$. It follows that

$$\mathbb{P}[A_T = B_T = A'_T = B'_T = 1] \geq (0.98\alpha)\mathbb{P}_\nu[A'_T = B'_T = 1] \geq (0.98\alpha)0.25(0.45\alpha^2) \geq 0.11\alpha^3.$$

- **Case 2:** $\mathbb{E}[A'_T + B'_T] \leq 1.5$. Since $\mathbb{E}_\nu[A'_T + B'_T] \geq 1 + \alpha$, by [Lemma 6.4](#), $\mathbb{P}[A'_T + B'_T = 2] \geq \alpha e^{-\alpha} \geq 0.99\alpha$. But now $\mathbb{E}[A'_T], \mathbb{E}[B'_T] \leq 1.5$ and therefore by Markov's Inequality,

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

On the other hand, by [Lemma 6.5](#) (similar to case 1) $\mathbb{P}_\nu [A'_T \geq 1], \mathbb{P}_\nu [B'_T \geq 1] \geq 1 - e^{-\alpha+\epsilon} \geq 0.98\alpha$. It follows by [Corollary 6.12](#) that $\mathbb{P} [A'_T = 1 | A'_T + B'_T = 2] \geq 0.2\alpha$. Therefore,

$$\mathbb{P} [A_T = B_T = A'_T = B'_T = 1] \geq (0.98\alpha)\mathbb{P}_\nu [A'_T = B'_T = 1] \geq (0.98\alpha)(0.2\alpha)(0.99\alpha) \geq 0.11\alpha^3$$

as desired. \square

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

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

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

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

Proof of [Proposition 6.13](#). To prove the lemma, we construct an instance of the max-flow, min-cut problem. Consider the following graph with vertex set $\{s, A, B, t\}$. For any $e \in A, f \in B$ connect e to f with a directed edge of capacity $y_{e,f} = \mathbb{P} [e, f \in T | A_T = B_T = 1]$. For any $e \in E$, let $x_e := \mathbb{P} [e \in T]$. Connect s to $e \in A$ with an arc of capacity βx_e and similarly connect $f \in B$ to t with arc of capacity βx_f , where β is a parameter that we choose later. We claim that the min-cut of this graph is at least $\beta(1 - \epsilon - \zeta/2.1)$. Assuming this, we can prove the lemma as follows: let \mathbf{z} be the maximum flow, where $z_{e,f}$ is the flow on the edge from e to f . We define the event $\mathcal{E}_{A,B}(T) = \mathcal{E}(T)$ to be the union of events $z_{e,f}$. More precisely, conditioned on $A_T = B_T = 1$ the events $e, f \in T | A_T = B_T = 1$ are disjoint for different pairs $e \in A, f \in B$, so we know that we have a specific e, f in the tree T with probability $y_{e,f}$. And, of course, $\sum_{e \in A, f \in B} y_{e,f} = 1$. So, for $e \in A, f \in B$ we include a $z_{e,f}$ measure of trees, T , such that $A_T = B_T = 1, e, f \in T$. First, observe that

$$\mathbb{P} [\mathcal{E}] = \sum_{e \in A, f \in B} z_{e,f} \mathbb{P} [A_T = B_T = 1] \geq \beta(1 - \zeta/2.1 - \epsilon) \mathbb{P} [A_T = B_T = 1]. \quad (24)$$

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

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

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

$$\begin{aligned} \sum_{e \in A} |x_e - \mathbb{P} [e \in T | \mathcal{E}]| &= 2 \sum_{e \in A: z_e/z > x_e} \left(\frac{z_e}{z} - x_e \right) \\ &\leq 2 \sum_{e \in A: z_e/z > x_e} \left(\frac{\beta x_e}{\beta(1 - \zeta/2.1 - \epsilon)} - x_e \right) \\ &\leq 2 \cdot \sum_e x_e \frac{\zeta/2.1 + \epsilon}{1 - \zeta/2.1 - \epsilon} \leq 2 \frac{(1 + \epsilon)(\zeta/2.1 + \epsilon)}{1 - \zeta/2.1 - \epsilon} \leq \zeta. \end{aligned}$$

The first inequality uses that the max-flow is at least $\beta(1 - \zeta/2.1 - \epsilon)$ and that the incoming flow of e is at most βx_e , and the last inequality follows by $\zeta < 0.003$ and $\epsilon < \zeta/330$. (iii) follows by the same argument.

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

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

If $x(S_B) \geq x(S_A) - \zeta/2.1$, then

$$\text{cap}(S, \bar{S}) \geq \beta x(\bar{S}_A \cup S_B) \geq \beta(x(\bar{S}_A \cup S_A) - \zeta/2.1) \geq \beta(1 - \epsilon - \zeta/2.1),$$

and we are done. Otherwise, say $x(S_B) + \gamma = x(S_A)$, for some $\gamma > \zeta/2.1$. So,

$$x(\bar{S}_B) + x(S_A) = x(\bar{S}_B) + x(S_B) + \gamma \geq 1 - \epsilon + \gamma$$

So, by [Lemma 6.14](#) with $(\alpha = \gamma - \epsilon > \zeta/2.1 - \epsilon > 100\epsilon)$

$$\mathbb{P}[(S_A)_T = (\bar{S}_B)_T = 1 | A_T = B_T = 1] \geq \frac{\mathbb{P}[(S_A)_T = (\bar{S}_B)_T = A_T = B_T = 1]}{\mathbb{P}[A_T = B_T = 1]} \geq \frac{0.11(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1]}.$$

It follows that

$$\begin{aligned} \text{cap}(S, \bar{S}) &\geq \beta x(\bar{S}_A \cup S_B) + \frac{0.11(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1]} \\ &\geq \beta(x(\bar{S}_A \cup S_A) - \gamma) + \frac{0.11(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1]} \\ &\geq \beta(1 - \epsilon - \gamma) + \frac{0.11(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1]} \end{aligned}$$

To prove the lemma we just need to choose β such that RHS is at least $\beta(1 - \epsilon - \zeta/2.1)$. Or equivalently,

$$\frac{0.11(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1]} \geq \beta(\gamma - \zeta/2.1).$$

In other words, it is enough to choose $\beta \leq \frac{0.11(\gamma - \epsilon)^3}{\mathbb{P}[A_T = B_T = 1](\gamma - \zeta/2.1)}$. Since $\gamma > \zeta/2.1$ and $\zeta > 330\epsilon$, we have $\gamma - \epsilon \geq 0.473\zeta$. Therefore, we can set $\beta = \frac{0.11(0.473\zeta)^2}{\mathbb{P}[A_T = B_T = 1]}$. Finally, this plus [\(24\)](#) gives

$$\mathbb{P}[\mathcal{E}] \geq (1 - \zeta/2.1 - \epsilon)\beta\mathbb{P}[A_T = B_T = 1] = 0.11(0.473\zeta)^2(1 - \zeta/2.1 - \epsilon) \geq 0.0246\zeta^2(1 - \zeta/2.1 - \epsilon)$$

as desired. \square

Definition 6.15 (Max-flow Event). *For a polygon cut $S \in \mathcal{H}$ with polygon partition A, B, C , let ν be the max-entropy distribution conditioned on S is a tree and $C_T = 0$. By [Lemma 2.28](#), we can write $\nu : \nu_S \times \nu_{G/S}$, where ν_S is supported on trees in $E(S)$ and $\nu_{G/S}$ on trees in $E(G/S)$. For a sample $(T_S, T_{G/S}) \sim \nu_S \times \nu_{G/S}$, we say \mathcal{E}_S occurs if $\mathcal{E}_{A,B}(T_{G/S})$ occurs, where $\mathcal{E}_{A,B}(\cdot)$ is the event defined in [Proposition 6.13](#) for sets A, B and $\zeta = \epsilon_M := \frac{1}{4000}$ and $\epsilon = 2\epsilon_\eta$.*

Corollary 6.16. For a polygon cut $S \in \mathcal{H}$ with polygon partition A, B, C , we have,

i) $\mathbb{P}[\mathcal{E}_S] \geq 0.0245\epsilon_M^2$.

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

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

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

$$\mathbb{P}[\mathcal{E}_S] = \mathbb{P}_\nu[\mathcal{E}_S] \mathbb{P}[C_T = 0, S \text{ tree}] \geq 0.0246\epsilon_M^2(1 - \epsilon_M/2.1 - \epsilon)\mathbb{P}[C_T = 0, S \text{ tree}] \geq 0.0245\epsilon_M^2,$$

using $\epsilon = 2\epsilon_\eta$ and $\epsilon_M = 1/4000$, which proves (i).

Now, we prove (ii). By [Proposition 6.13](#), the marginals of edges in $\delta(S)$ are preserved up to a total variation distance of ϵ_M , so

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

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

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

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

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

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

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

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

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

Corollary 6.17. For $u \in \mathcal{H}$ and a polygon cut $S \in \mathcal{H}$ that is an ancestor of u ,

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

Proof. First, notice by [Observation 5.32](#), $\delta(u) \cap \delta(S)$ is either a subset of A , B , or C . Therefore, by (iii) of [Corollary 6.16](#) we can write $\delta(u)_T | \mathcal{E}_S$ as a $BS(q)$ for $q \in 2 \pm [0.001]$ (where we use that $\epsilon_M + 3\epsilon_\eta < 0.001$). Furthermore, since $\delta(u)_T \neq 0$ with probability 1, we can write this as a $1 + BS(q - 1)$. Therefore, by [Lemma 2.21](#),

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

as desired. \square

Corollary 6.18. For a polygon cut $u \in \mathcal{H}$ and a polygon cut $S \in \mathcal{H}$ that is an ancestor of u ,

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

and the same follows for right happy.

Proof. Let A, B, C be the polygon partition of u . Recall that for u to be left-happy, we need $C_T = 0$ and A_T odd. Similar to the previous statement, we can write $A_T | \mathcal{E}_S$ as a $BS(q_A)$ for $q_A \in 1 \pm [0.00026]$ (where we used that $\epsilon_M = 1/4000$ and $\epsilon_\eta \leq \epsilon_M/300$). Therefore, by [Lemma 2.21](#),

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

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

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

as desired. □

6.5 Good Edges

Definition 6.19 (Half Edges). We say an edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ in a degree cut $S \in \mathcal{H}$, i.e., $p(\mathbf{e}) = S$, is a half edge if $|x_{\mathbf{e}} - 1/2| \leq \epsilon_{1/2}$, where $\epsilon_{1/2}$ is defined in [Global constants](#).

Definition 6.20 (Good Edges). We say a top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ in a degree cut $S \in \mathcal{H}$ is (2-2) good, if one of the following holds:

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

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

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

Theorem 6.21. For $\epsilon_{1/2} \leq 0.0002$, $\epsilon_\eta \leq \epsilon_{1/2}^2$, a top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ is bad only if the following three conditions hold simultaneously:

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

The proof of this theorem follows from [Lemma 6.23](#) and [Lemma 6.24](#) below.

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

Lemma 6.22. Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle in a degree cut $S \in \mathcal{H}$. Let $A = \delta(u)_{-\mathbf{e}}$ and $B = \delta(v)_{-\mathbf{e}}$. If $\epsilon_{1/2} \leq 0.001$ and $\epsilon_\eta < \epsilon_{1/2}/100$, then

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

Proof. Throughout the proof all probabilistic statements are with respect to the measure μ conditioned on u, v trees. Let $p_{\leq 2} = \mathbb{P}[A_T + B_T \leq 2]$ and similarly define $p_{\geq 4}$. Observe that whenever $\delta(u)_T = \delta(v)_T = 2$, we must have $A_T + B_T \neq 3$. Since \mathbf{e} is 2-2 good, this event happens with probability at least $3\epsilon_{1/2}$, i.e.,

$$p_{\leq 2} + p_{\geq 4} \geq 3\epsilon_{1/2} \tag{25}$$

By [Lemma 6.4](#), using the fact that $p_0 = 0$, we get $p_{=3} \geq 1/4$.

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

$$3 + 2\epsilon_{1/2} \geq \mathbb{E}[A_T + B_T] \geq 4p_{\geq 4} + 2p_{=2} + 3(1 - p_{\geq 4} - p_{\leq 2}) = 3 + p_{\geq 4} - p_{=2} - 3p_{=1}.$$

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

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

Finally, since $\epsilon_{1/2} < 0.001$, plugging this upper bound on $p_{\geq 4}$ into [Eq. \(25\)](#) we get $p_{\leq 2} \geq 0.4\epsilon_{1/2}$.

Now, we show $p_{\geq 4} \geq 0.4\epsilon_{1/2}/2$. Assume $p_{\geq 4} < \epsilon_{1/2}/2$ (otherwise we are done). Since $p_{=3} \geq 1/4$ by [Lemma 6.1](#) with $\gamma \leq (\epsilon_{1/2}/2)/(1/4) = 2\epsilon_{1/2}$

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

Therefore,

$$3 - 2\epsilon_{1/2} - 2\epsilon_\eta \leq \mathbb{E}[A_T + B_T] \leq 2p_{\leq 2} + \frac{p_{\geq 4}}{1 - 2\epsilon_{1/2}}(4 + 3\epsilon_{1/2}) + 3(1 - p_{\leq 2} - p_{\geq 4})$$

So, $1.01p_{\geq 4} \geq p_{\leq 2} - 2.02\epsilon_{1/2}$ where we used $\epsilon_{1/2} \leq 0.001$ and $\epsilon_\eta < \epsilon_{1/2}/100$. Now, $p_{\geq 4} \geq 0.4\epsilon_{1/2}$ follows by [Eq. \(25\)](#). \square

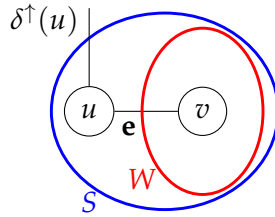


Figure 36: Setting of [Lemma 6.23](#)

Lemma 6.23. Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a half edge bundle in a degree cut $S \in \mathcal{H}$, and suppose $x(\delta^\uparrow(u)) \geq 1/2 + k\epsilon_{1/2}$. If $k \geq 9$, $\epsilon_{1/2} \leq 0.0002$, and $\epsilon_\eta \leq \epsilon_{1/2}^2$, then, \mathbf{e} is 2-2 good.

Proof. First, condition u, v, S to be trees. Let $W = S \setminus \{u\}$. Since S is a near mincut,

$$x(\delta(W)) = x(\delta(S)) + x(\delta(u)) - 2x(\delta^\dagger(u)) \leq 2(2 + \epsilon_\eta) - 2(1/2 + k\epsilon_{1/2}) = 3 - 2k\epsilon_{1/2} + 2\epsilon_\eta$$

So, by [Lemma 2.28](#), $\mathbb{P}[W \text{ is tree}] \geq 1/2 + k\epsilon_{1/2} - \epsilon_\eta - \epsilon_\eta$. Note that the extra $-\epsilon_\eta$ comes from the fact that conditioning u to be a tree can decrease marginals of edges in $E(W)$ by at most ϵ_η .

Let ν be the resulting measure, namely the measure obtained by first conditioning u, v, S to be trees and then W to be a tree. Note that ν is a strongly Rayleigh distribution on the set of edges in $E(W) \cup E(u, W) \cup E(G/S)$; this is because ν is a product of 3 SR distributions each supported on one of the aforementioned sets.

Let $X = \delta^\dagger(u)_T$ and $Y = \delta(v)_T - 1$. Observe that, under ν , $X = Y = 1$ iff $\delta(u)_T = \delta(v)_T = 2$. Furthermore, $Y \geq 0$ with probability 1, since v is connected to the rest of the graph. So, we just need to lower bound $\mathbb{P}_\nu[X = Y = 1]$. First, notice

$$\begin{aligned} \mathbb{E}_\nu[X] &\in [0.5 + k\epsilon_{1/2} - \epsilon_\eta, 1 + \epsilon_\eta] \\ \mathbb{E}_\nu[Y] &\in [0.5 + k\epsilon_{1/2} - 4\epsilon_\eta, 1.5 - k\epsilon_{1/2} + 3\epsilon_\eta] \end{aligned} \tag{26}$$

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

Note that using [Proposition 6.8](#), we can immediately argue that $\mathbb{P}_\nu[X = Y = 1] \geq \Omega(\epsilon_{1/2})$. We do the following more refined analysis to make sure that this probability is at least $6\epsilon_{1/2}$ (for $\epsilon_{1/2} \leq 0.0005$) and $k \geq 9$. Once we prove this, we obtain the lemma:

$$\mathbb{P}[\delta(u)_T = \delta(v)_T = 2 \mid u, v \text{ trees}] \geq \mathbb{P}[S, W \text{ trees} \mid u, v \text{ trees}] \mathbb{P}_\nu[X = Y = 1] \geq 0.5 \cdot 6\epsilon_{1/2}$$

Case 1: $\mathbb{P}_\nu[X + Y = 2] \geq 48\epsilon_{1/2}$. By [Lemma 6.5](#), $\mathbb{P}_\nu[X \geq 1], \mathbb{P}_\nu[Y \geq 1] \geq 1 - e^{-0.5}$. On the other hand, by [Theorem 2.17](#), $\mathbb{P}_\nu[X \leq 1], \mathbb{P}_\nu[Y \leq 1] \geq 7/16$. This is because if we have one Bernoulli of value 1, $\mathbb{P}_\nu[X \leq 1] \geq (1 - \frac{0.5}{n})^n$ is minimized at $n = 1$, whereas if we have no Bernoullis of value 1, $\mathbb{P}_\nu[X \leq 1] \geq (1 - \frac{1.5}{n})^n + 1.5(1 - \frac{1.5}{n})^{n-1}$ which is minimized at $n = 2$. Therefore, by [Corollary 6.12](#), $\mathbb{P}_\nu[X = 1 \mid X + Y = 2] \geq 0.1269$. Therefore, we get

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

Case 2: $\mathbb{P}_\nu[X + Y = 2] < 48\epsilon_{1/2} < 0.01$. By [Lemma 6.4](#), $\mathbb{P}_\nu[X + Y = 1] \geq 0.25$ (if $\mathbb{E}_\nu[X + Y] \geq 1.2$ then the assumption of this case obviously fails). So, since $\mathbb{P}_\nu[X + Y = 2] < 0.01$, by log concavity, $\mathbb{P}_\nu[X + Y = 3] \leq 0.01/25$. Furthermore, by [Lemma 6.1](#) (with $\gamma = 1/25, i = 1, k = 3$), $\mathbb{P}_\nu[X + Y > 2] < 0.0005$.

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

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

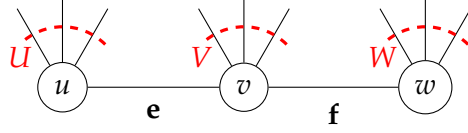


Figure 37: Setting of Lemma 6.24

Similarly, $\mathbb{P}_v[X \leq 1 | X + Y = 2] = \mathbb{P}_v[Y \geq 1 | X + Y = 2] \geq \mathbb{P}_v[Y \geq 1] - 0.0005 \geq 0.469$. Finally since the distribution of X conditioned on $X + Y = 2$ is the same as the number of successes in 2 independent Bernoulli trials, with probabilities, say, p_1 and p_2 , we can minimize $p_1(1 - p_2) + (1 - p_1)p_2$ subject to $1 - p_1p_2 \geq 0.469$ and $1 - (1 - p_1)(1 - p_2) \geq 0.469$. Solving this yields $\mathbb{P}_v[X = 1 | X + Y = 2] \geq 0.395$.

Lastly, observe that since by Eq. (26) $1.2 \geq \mathbb{E}_v[X + Y] \geq 1 + (2k - 1)\epsilon_{1/2}$, by Lemma 6.4 we can write

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

Therefore,

$$\mathbb{P}_v[X = Y = 1] = \mathbb{P}_v[X = 1 | X + Y = 2] \mathbb{P}_v[X + Y = 2] \geq 0.395(2k - 2)\epsilon_{1/2}$$

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

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

$$\mathbb{P}_v[X = 2] \leq \mathbb{P}_v[X + Y \geq 2] \leq 0.01 + 0.00042 \leq 0.0105$$

Also notice that $\mathbb{P}_v[X = 1] \geq 0.3$ by Lemma 6.4. Now, using Lemma 6.1 we can write, for $\gamma = 1/25$ and $i = 1$,

$$\mathbb{E}_v[X | X \geq 2] \mathbb{P}_v[X \geq 2] \leq 0.0224$$

Therefore, since X is integer valued,

$$\mathbb{P}_v[X \geq 1] \geq \mathbb{E}_v[X] - \mathbb{E}_v[X | X \geq 2] \mathbb{P}_v[X \geq 2] \geq \mathbb{E}_v[X] - 0.0224 \geq 0.47,$$

as desired. \square

Lemma 6.24. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be two half edge bundles in a degree cut $S \in \mathcal{H}$. If $\epsilon_{1/2} < 0.0005$ and $\epsilon_\eta \leq \epsilon_{1/2}^2$, then one of \mathbf{e} or \mathbf{f} is good.*

Proof. We use the following notation $V = \delta(v)_{-\mathbf{e}-\mathbf{f}}, U = \delta(u)_{-\mathbf{e}}, W = \delta(w)_{-\mathbf{f}}$ (see Fig. 37 for an illustration). For a set A of edges and an edge bundle \mathbf{e} we write $A_{+\mathbf{e}} = A \cup \{\mathbf{e}\}$. Furthermore, for a measure ν we write $\nu_{-\mathbf{e}}$ to denote ν conditioned on $\mathbf{e} \notin T$.

Condition u, v, w to be trees. This occurs with probability at least $1 - 3\epsilon_\eta$. Let ν be this measure. By Lemma 6.7, without loss of generality, we can assume

$$\mathbb{E}_\nu[W_T | \mathbf{e} \notin T] \leq \mathbb{E}_\nu[W_T] + 0.405. \quad (27)$$

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

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

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

Therefore, by [Lemma 6.4](#), $\mathbb{P}_{\nu_{-e}}[(V_{+f})_T + U_T = 4] \geq 0.029$, where we use the fact that $U_T \geq 1$ and $(V_{+f})_T \geq 1$ with probability 1 under ν_{-e} and apply this and the remaining calculations to $U_T - 1, (V_{+f})_T - 1$. In addition, we have

$$\begin{aligned} \mathbb{P}_{\nu_{-e}}[U_T \leq 2], \mathbb{P}_{\nu_{-e}}[(V_{+f})_T \leq 2] &\geq 0.499 && \text{(Markov Inequality)} \\ \mathbb{P}_{\nu_{-e}}[U_T \geq 2], \mathbb{P}_{\nu_{-e}}[(V_{+f})_T \geq 2] &\geq 0.39 && \text{(Lemma 6.5)} \end{aligned}$$

It follows by [Corollary 6.12](#) applied to $U_T - 1$ and $(V_{+f})_T - 1$ (with $\epsilon = 0.194$ and $p_m = 0.6$) that

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

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

Therefore,

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

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

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

$$\begin{aligned} \mathbb{E}_{\nu_{+f}}[(V_{+e})_T], \mathbb{E}_{\nu_{+f}}[W_T] &\in [1 - 2\epsilon_{1/2} - 3\epsilon_\eta, 1.5 + 2\epsilon_{1/2} + \epsilon_\eta] \\ \mathbb{P}_{\nu_{+f}}[(V_{+e})_T \leq 1], \mathbb{P}_{\nu_{+f}}[W_T \leq 1] &\geq 0.249 && \text{(Markov)} \\ \mathbb{P}_{\nu_{+f}}[(V_{+e})_T \geq 1], \mathbb{P}_{\nu_{+f}}[W_T \geq 1] &\geq 0.63 && \text{(Lemma 6.5)} \end{aligned}$$

So, by [Corollary 6.12](#) (with $\epsilon = 0.15, p_m = 0.7$), we get $\mathbb{P}_{\nu_{+f}}[W_T = 1 | (V_{+e})_T + W_T = 2] \geq 0.11$. On the other hand,

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

To derive the last inequality, we show $\mathbb{P}_{\nu_{+f-e}}[(V_{+e})_T + W_T = 2] \geq 0.0582$. This is because by negative association and [Eq. \(27\)](#)

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

So, since $(V_{+e})_T + W_T$ is always at least 1, so by [Theorem 2.17](#), in the worst case, $\mathbb{P}_{\nu_{+f-e}}[(V_{+e})_T + W_T = 2]$ is the probability that the sum of two Bernoullis with success probability $1.94/2$ is 1, which is 0.0582.

Therefore, similar to the previous case,

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

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

6.6 2-1-1 and 2-2-2 Good Edges

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

Definition 6.25 (A, B, C -Degree Partitioning). For $u \in \mathcal{H}$ and $\epsilon_{1/1}$ defined in [Global constants](#), we define a partitioning of edges in $\delta(u)$: Let $a, b \subsetneq u$ be minimal cuts in the hierarchy, i.e., $a, b \in \mathcal{H}$, such that $a \neq b$ and $x(\delta(a) \cap \delta(u)), x(\delta(b) \cap \delta(u)) \geq 1 - \epsilon_{1/1}$. Note that since the hierarchy is laminar, a, b cannot cross. Let $A = \delta(a) \cap \delta(u)$, $B = \delta(b) \cap \delta(u)$, $C = \delta(u) \setminus A \setminus B$.

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

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

See [Fig. 30](#) for an example. The following inequalities on A, B, C degree partitioning will be used in this section:

$$\begin{aligned} x(A), x(B) &\in [1 - \epsilon_{1,1}, 1 + \epsilon_\eta], \\ x(C) &\leq 2\epsilon_{1/1} + \epsilon_\eta. \end{aligned} \tag{28}$$

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

Definition 6.26 (2-1-1 Happy/Good). Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle. Let $A, B, C \subseteq \delta(u)$ be a Degree Partitioning of edges $\delta(u)$ as defined in [Definition 15.5](#). We say that \mathbf{e} is 2-1-1 happy with respect to u if the event

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

occurs.

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

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

Remark 6.27. Note we also use this A, B, C partitioning to help deal with the triangle cut case. In the special case that u is a polygon cut with A, B, C -polygon partitioning, let A', B', C' be the degree partitioning of $\delta(u)$. Then, by [Definition 5.31](#) we have $A' \subseteq A$, $B' \subseteq B$, $C' \subseteq C'$. Therefore, if an edge in $\delta(u)$ is reduced and is 2-1-1 happy with respect to u , the polygon u is also happy. See [the overview](#) for an example.

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

Lemma 6.28. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \leq 1/2 - \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$ then, \mathbf{e} is 2-1-1 happy with probability at least $0.005\epsilon_{1/2}^2$.*

Lemma 6.29. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \geq 1/2 + \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$, then, \mathbf{e} is 2-1-1 happy with respect to u with probability at least $0.006\epsilon_{1/2}^2$.*

Fix u in the hierarchy with degree partitioning A, B, C . The above two lemmas show that any edge bundle $\mathbf{e} \in \delta(u)$ which is not a half edge bundle is 2-1-1 good, so the difficult case is when the majority of $x(\delta^\rightarrow(u))$ comes from half edge bundles. In [Theorem 6.21](#) we showed that $\delta(u)$ can have at most one 2-2 bad edge. Oddly enough, one of the simplest cases of the reduction argument is when there is a bad edge in $\delta(u)$. This is because we never reduce bad edges, and therefore we never need to increase edges which are matched to them.²⁴ So, the main problem is good edges which are not 2-1-1 good. The following key statement, [Lemma 6.32](#), shows that these problematic edges are rare in the sense that there is at most one good half edge bundle in A (resp. B) which is not 2-1-1 good.

To prove this we need the following two lemmas. In the first one we show that if \mathbf{e}, \mathbf{f} are two half edge bundles which almost entirely land in A (or B), at least one of them is 2-1-1 good. In the second, we show that if a good half edge bundle does not entirely land in A (or B), then it is 2-1-1 good. This is the main tool we use to upper bound the expected increase of good top edges in [Section 8](#).

For a set of edges D , and an edge bundle \mathbf{e} , let $\mathbf{e}(D) := \mathbf{e} \cap D$. Note that $\mathbf{e}(D)$ is not really an edge bundle.

Lemma 6.30. *Let $\mathbf{e} = (\mathbf{v}, \mathbf{u})$ and $\mathbf{f} = (\mathbf{v}, \mathbf{w})$ be good half top edge bundles and let A, B, C be the [degree partitioning](#) of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \leq \epsilon_{1/2}$. Then, one of \mathbf{e}, \mathbf{f} is 2-1-1 happy with probability at least $0.005\epsilon_{1/2}^2$.*

Lemma 6.31. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle and let A, B, C be the [degree partitioning](#) of $\delta(u)$ (see [Fig. 87](#)). If $\epsilon_{1/2} \leq 0.001$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \geq \epsilon_{1/2}$, then*

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t } u] \geq 0.02\epsilon_{1/2}^2.$$

Lemma 6.32. *For a degree cut $S \in \mathcal{H}$, and $u \in \mathcal{A}(S)$, let A, B, C be the [degree partition](#) of u . Then, $A \cap \delta^\rightarrow(u) =: A^\rightarrow$ has fraction at most $1/2 + 4\epsilon_{1/2}$ of good edges that are not 2-1-1 good (w.r.t., u).*

Proof. Suppose by way of contradiction that there is a set $D \subseteq A^\rightarrow$ of good edges that are not 2-1-1 good w.r.t. u with $x(D) \geq \frac{1}{2} + 4\epsilon_{1/2}$. By [Lemma 6.28](#) and [Lemma 6.29](#), every edge in D is part of a half edge bundle.

There are at least two half edge bundles \mathbf{e}, \mathbf{f} such that $x(D \cap \mathbf{e}), x(D \cap \mathbf{f}) \geq \epsilon_{1/2}$, as there are at most four half edge bundles in $\delta^\rightarrow(u)$ (and using that for any half edge bundle \mathbf{e} , $x_{\mathbf{e}} \leq \frac{1}{2} + \epsilon_{1/2}$). Since $D \subseteq A^\rightarrow$, we have

$$x(A \cap \mathbf{e}), x(A \cap \mathbf{f}) \geq \epsilon_{1/2}.$$

²⁴The main problem with bad edges is that we cannot match them to edges going higher in the matching lemma [7.2](#). So, in order to prove the matching lemma we need to justify that there are not too many bad edges in any cut. Therefore we cannot simply “pretend” that one half edge bundle of $\delta(u)$ is bad.

Since $x(A \cap \mathbf{e}) \geq \epsilon_{1/2}$, if $x(B \cap \mathbf{e}) \geq \epsilon_{1/2}$ then, by [Lemma 6.31](#) \mathbf{e} is 2-1-1 good. But since every edge in D is not 2-1-1 good w.r.t u , we must have $x(B \cap \mathbf{e}) < \epsilon_{1/2}$. The same also holds for \mathbf{f} . Finally, since $x(B \cap \mathbf{e}) < \epsilon_{1/2}$ and $x(B \cap \mathbf{f}) < \epsilon_{1/2}$ by [Lemma 6.30](#) at least one of \mathbf{e}, \mathbf{f} is 2-1-1 good w.r.t u . This is a contradiction. \square

2-2-2 Good Edges. While [Lemma 6.32](#) is sufficient for bounding the increase of top edges, it is not sufficient for bottom edges. Fix a polygon u with partition A, B, C and suppose $\rho(u) = S$ is a degree cut (recall that by [Remark 6.27](#), the degree partitioning and polygon partitioning of u are essentially the same). Roughly speaking, a bottom edge $g \in E(u)$ is “matched” to all edges in $\delta(u)$, and needs to increase for edges $f \in A$ when f is reduced and A_T is even, and for edges $f \in B$ when f is reduced and B_T is even. Therefore, g is matched to essentially twice its fraction. If most of the edges in $\delta(u)$ are 2-1-1 good, this is sufficient to bound the expected increase of g because when such an edge is reduced and 2-1-1 happy with respect to u , g does not need to increase.

It turns out that the above lemmas are sufficient to bound the expected increase of $g \in E(u)$ except when $A \cap \delta(S) \approx B \cap \delta(S) \approx 1/2$ and $\mathbf{e} \approx A \cap \delta^\rightarrow(u)$ and $\mathbf{f} \approx B \cap \delta^\rightarrow(u)$ are both good edge bundles which are not 2-1-1 good. In this extreme case, we employ a new strategy. In [Lemma 6.34](#) below, we prove that the two edge bundles \mathbf{e}, \mathbf{f} are 2-2 happy *simultaneously* with a constant probability. We call such a pair 2-2-2 good. Later, in [Section 8](#), we use this to ensure that \mathbf{e} and \mathbf{f} are always reduced simultaneously. The point is that since \mathbf{e}, \mathbf{f} do not both come from A (or B), no cut inside u contains \mathbf{e} and \mathbf{f} . Therefore, g only needs to increase by the maximum of the decrease of \mathbf{e}, \mathbf{f} (not the sum), effectively saving a factor of 2.

Definition 6.33 (2-2-2 Happy/Good). *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be top half-edge bundles (with $\rho(\mathbf{e}) = \rho(\mathbf{f})$). We say \mathbf{e}, \mathbf{f} are 2-2-2 happy (with respect to v) if $\delta(u)_T = \delta(v)_T = \delta(w)_T = 2$ and u, v, w are all trees.*

We say \mathbf{e}, \mathbf{f} are 2-2-2 good with respect to v if $\mathbb{P}[\mathbf{e}, \mathbf{f} \text{ 2-2-2 happy}] \geq p$.

Lemma 6.34. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be two good top half edge bundles and let A, B, C be [degree partitioning](#) of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$. If \mathbf{e}, \mathbf{f} are not 2-1-1 good with respect to v , and $\epsilon_{1/2} \leq 0.0002$, then \mathbf{e}, \mathbf{f} are 2-2-2 happy with probability at least 0.01.*

The following theorem summarizes the above results in a compact form. This is the main result used in the analysis of the increase for bottom edges in [Section 8](#).

Theorem 6.35. *Let $v, S \in \mathcal{H}$ where $\rho(v) = S$, and let A, B, C be the [degree partitioning](#) of $\delta(v)$. For $p \geq 0.005\epsilon_{1/2}^2$, with $\epsilon_{1/2} \leq 0.0002$, $\epsilon_{1/1} \leq \epsilon_{1/2}/12$ and $\epsilon_\eta \leq \epsilon_{1/2}^2$, at least one of the following is true:*

- i) $\delta^\rightarrow(v)$ has at least $1/2 - \epsilon_{1/2}$ fraction of bad edges,
- ii) $\delta^\rightarrow(v)$ has at least $1/2 - \epsilon_{1/2} - \epsilon_\eta$ fraction of 2-1-1 good edges with respect to v .
- iii) There are two (top) half edge bundles $\mathbf{e}, \mathbf{f} \in \delta^\rightarrow(v)$ such that $x_{\mathbf{e}(B)} \leq \epsilon_{1/2}$, $x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$, and \mathbf{e}, \mathbf{f} are 2-2-2 good (with respect to v).

Proof. Suppose case (i) does not happen. Since every bad edge has fraction at least $1/2 - \epsilon_{1/2}$ this means that $\delta(v)$ has no bad edges. First, notice by [Lemma 6.28](#) and [Lemma 6.29](#) any non half-edge in $\delta^\rightarrow(v)$ is 2-1-1 good (with respect to v). (Recall we define $\delta^\rightarrow(v) = \delta(v) \setminus \delta(\rho(v))$, where $\rho(v)$ is the immediate parent of v in the hierarchy). If there is only one half edge in $\delta^\rightarrow(v)$,

then we have at least fraction $1 - \epsilon_\eta - (1/2 + \epsilon_{1/2})$ fraction of 2-1-1 good edges and we are done with case (ii). Otherwise, there are two good half edges $\mathbf{e}, \mathbf{f} \in \delta^{\rightarrow}(v)$.

First, by [Lemma 6.31](#) if $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \geq \epsilon_{1/2}$, then \mathbf{e} is 2-1-1 good (w.r.t., v) and we are done. Similarly, if $x_{\mathbf{f}(A)}, x_{\mathbf{f}(B)} \geq \epsilon_{1/2}$, then \mathbf{f} is good. So assume none of these happens.

Furthermore by [Lemma 6.30](#) if $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \leq \epsilon_{1/2}$ (or $x_{\mathbf{e}(A)}, x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$) then one of \mathbf{e}, \mathbf{f} is 2-1-1 good.

So, the only remaining case is when \mathbf{e}, \mathbf{f} are not 2-1-1 good and $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$. But in this case by [Lemma 6.34](#), \mathbf{e}, \mathbf{f} are 2-2-2 good; so (iii) holds. \square

7 Matching

The main result of this section is to construct a matching that we use in order to decide which edges will have positive slack to compensate for the negative slack of edges going higher. Refer to [Example 4.9](#) for a high-level motivation to construct a matching.

Definition 7.1 (ϵ_F fractional edge). For $z \geq 0$ we say that z is ϵ_F -fractional if $\epsilon_F \leq z \leq 1 - \epsilon_F$.

The following lemma is the main result of this section

Lemma 7.2 (Matching Lemma). For any $S \in \mathcal{H}$, $\epsilon_F \leq 1/10$, $\epsilon_B \geq 21\epsilon_{1/2}$, $\alpha \geq 2\epsilon_\eta$, $\epsilon_{1/2} \leq 0.0002$, there is a matching from good edges (see [Definition 6.20](#)) in $E^\rightarrow(S)$ to edges in $\delta(S)$ where every good edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ (where $u, v \in \mathcal{A}(S)$) is matched to a fraction $m_{\mathbf{e},u}$ of edges in $\delta^\uparrow(u)$ and a fraction $m_{\mathbf{e},v}$ of $\delta^\uparrow(v)$, and:

$$m_{\mathbf{e},u}F_u + m_{\mathbf{e},v}F_v \leq x_{\mathbf{e}}(1 + \alpha) \quad (29)$$

$$\sum_{e \in \delta^\rightarrow(u)} m_{\mathbf{e},u} = x(\delta^\uparrow(u))Z_u, \quad (30)$$

where for every atom $u \in \mathcal{A}(S)$, define

$$F_u = 1 - \epsilon_B \mathbb{I} \left\{ x(\delta^\uparrow(u)) \text{ is } \epsilon_F \text{ fractional} \right\}, \quad Z_u := \left(1 + \mathbb{I} \left\{ |\mathcal{A}(S)| \geq 4, x(\delta^\uparrow(u)) \leq \epsilon_F \right\} \right).$$

Roughly speaking, the intention of the above lemma is to match good edges in $E^\rightarrow(S)$ to a similar fraction of edges that go higher (such that an edge bundle e adjacent to atoms u, v is only matched to edges in $\delta^\uparrow(u), \delta^\uparrow(v)$). Since we never “reduce” bad edges in the proof of payment theorem ([Theorem 5.33](#)), we don’t use them in the matching. That inherently can cause a problem, as there could not be “enough” good edges in $E^\rightarrow(S)$ to saturate the edges going higher in the matching. The parameter F_u help us in this regard; in particular, it allows us to match some of the (good) edges in $E^\rightarrow(S)$ to more than their fraction in $\delta(S)$.

Next, we motivate the parameter Z_u . If $x(\delta^\uparrow(u)) \approx 0$, when those edges are reduced the conditional probability that $\delta(u)_T$ is even could be very close to 0. The parameter Z_u lets us match twice as many edges to $\delta^\uparrow(u)$; so there will be only half a burden to fix the parity of $\delta(u)_T$. See the discussion in [overview section](#) for more details.

Throughout this section we adopt the following notation: For a cut $S \in \mathcal{H}$ and a set $W \subseteq \mathcal{A}(S)$, we write

$$\begin{aligned} E(W, S \setminus W) &:= \cup_{u \in W, v \in \mathcal{A}(S) \setminus W} E(u, v), \\ \delta^\uparrow(W) &:= \cup_{u \in W} \delta^\uparrow(u) = \delta(W) \cap \delta(S), \\ \delta^\rightarrow(W) &:= \cup_{u \in W} \delta^\rightarrow(u). \end{aligned}$$

Note that in $\delta^\rightarrow(W) \not\subseteq \delta(W)$ since it includes edge bundles between atoms in W .

Before proving the main lemma we record the following facts.

Lemma 7.3. For any $S \in \mathcal{H}$ and $W \subseteq \mathcal{A}(S)$ (recall $\mathcal{A}(S)$ is the set of $u \in \mathcal{H}$ with $p(u) = S$), we have

$$x(\delta^\rightarrow(W)) \geq \frac{1}{2} \sum_{u \in W} x(\delta(u)) - \epsilon/2 \geq |W| - \epsilon/2.$$

Proof. We have

$$x(\delta^{\rightarrow}(W)) = \frac{1}{2} \left(\sum_{u \in W} (x(\delta(u)) + x(E(W, S \setminus W)) - x(\delta^{\uparrow}(W))) \right).$$

Since $x(\delta(S \setminus W)) \geq 2$ and $x(\delta(S)) \leq 2 + \epsilon$, we have:

$$(a) \ x(E(W, S \setminus W)) + x(\delta^{\uparrow}(S \setminus W)) \geq 2 \quad \text{and} \quad (b) \ x(\delta^{\uparrow}(W)) + x(\delta^{\uparrow}(S \setminus W)) \leq 2 + \epsilon.$$

Subtracting (b) from (a), we get

$$x(E(W, S \setminus W)) - x(\delta^{\uparrow}(W)) \geq -\epsilon,$$

which after substituting into the above equation, completes the proof of the first inequality in the lemma statement. The second inequality follows from the fact that $\delta(u) \geq 2$ for each atom u . \square

Lemma 7.4. *For $S \in \mathcal{H}$, if $|\mathcal{A}(S)| = 3$ then there are no bad edges in $E^{\rightarrow}(S)$.*

Proof. Suppose $\mathcal{A}(S) = \{u, v, w\}$ and $\mathbf{e} = (u, v)$ is a bad edge bundle. Then $|x_{\mathbf{e}} - \frac{1}{2}| \leq \epsilon_{1/2}$. In addition, by [Theorem 6.21](#), $x(\delta^{\uparrow}(u)), x(\delta^{\uparrow}(v)) \leq 1/2 + 9\epsilon_{1/2}$. Therefore,

$$x_{(\mathbf{u}, \mathbf{w})} = x(\delta(u)) - x_{\mathbf{e}} - x(\delta^{\uparrow}(u)) \geq 1 - 10\epsilon_{1/2}.$$

Similarly, $x_{(\mathbf{v}, \mathbf{w})} \geq 1 - 10\epsilon_{1/2}$. Finally, since $x(\delta(S)) \geq 2$, and $x(\delta^{\uparrow}(u)), x(\delta^{\uparrow}(v)) \leq 1/2 + 9\epsilon_{1/2}$, we must have $x(\delta^{\uparrow}(w)) \geq 1 - 18\epsilon_{1/2}$. But, this contradicts the assumption that $w \in \mathcal{H}$ must satisfy $x(\delta(w)) \leq 2 + \epsilon_{\eta}$. \square

Proof of Lemma 7.2. We will prove this by setting up a max-flow min-cut problem. Construct a graph with vertex set $\{s, X, Y, t\}$, where s, t are the source and sink. We identify X with the set of good edge bundles in $E^{\rightarrow}(S)$ and Y with the set of atoms in $\mathcal{A}(S)$. For every edge bundle $\mathbf{e} \in X$, add an arc from s to \mathbf{e} of capacity $c(s, \mathbf{e}) := (1 + \alpha)x_{\mathbf{e}}$. For every $u \in \mathcal{A}(S)$, there is an arc (u, t) with capacity

$$c(u, t) = x(\delta^{\uparrow}(u))F_uZ_u.$$

Finally, connect $\mathbf{e} = (u, v) \in X$ to nodes u and $v \in Y$ with a directed edge of infinite capacity, i.e., $c(\mathbf{e}, u) = c(\mathbf{e}, v) = \infty$. We will show below that there is a flow saturating t , i.e. there is a flow of value

$$c(t) := \sum_{u \in \mathcal{A}(S)} c(u, t) = \sum_{u \in \mathcal{A}(S)} x(\delta^{\uparrow}(u))F_uZ_u.$$

Suppose that in the corresponding max-flow, there is a flow of value $f_{\mathbf{e}, u}$ on the edge (\mathbf{e}, u) . Define

$$m_{\mathbf{e}, u} := \frac{f_{\mathbf{e}, u}}{F_u}.$$

Then (29) follows from the fact that the flow leaving \mathbf{e} is at most the capacity of the edge from s to \mathbf{e} , and (30) follows by conservation of flow on the node u (after cancelling out F_u from both sides).

We have left to show that for any s - t cut A, \bar{A} where $s \in A, t \in \bar{A}$ that the capacity of this cut is at least $c(t)$.

Claim 7.5. *If $A = \{s\}$, then capacity of (A, \bar{A}) is at least $c(t)$.*

Proof. First, note that

$$\begin{aligned}
c(t) &= \sum_{u \in \mathcal{A}(S)} x(\delta^\dagger(u)) F_u Z_u \leq \sum_{u \in \mathcal{A}(S)} x(\delta^\dagger(u)) Z_u \\
&\leq \mathbb{I}\{|\mathcal{A}(S)| \geq 4\} \cdot |\{u \in \mathcal{A}(S) : x(\delta^\dagger(u)) \leq \epsilon_F\}| \cdot \epsilon_F + x(\delta(S)) \\
&\leq 2 + \epsilon_\eta + \epsilon_F \mathbb{I}\{|\mathcal{A}(S)| \geq 4\} |\mathcal{A}(S)|
\end{aligned} \tag{31}$$

because $F_u \leq 1$ and $Z_u = 1 + \mathbb{I}\{|\mathcal{A}(S)| \geq 4, x(\delta^\dagger(u)) \leq \epsilon_F\}$.

Second, note that

$$x(E^\rightarrow(S)) = \frac{1}{2} \sum_{u \in \mathcal{A}(S)} (x(\delta(u)) - x(\delta^\dagger(u))) \geq \frac{2|\mathcal{A}(S)| - (2 + \epsilon_\eta)}{2} = |\mathcal{A}(S)| - 1 - \epsilon_\eta/2.$$

Therefore, if there are k bad edges in $E^\rightarrow(S)$, then

$$x_G \geq |\mathcal{A}(S)| - 1 - \epsilon_\eta/2 - k\left(\frac{1}{2} + \epsilon_{1/2}\right) \tag{32}$$

Case 1: $|\mathcal{A}(S)| = 3$. Then $Z_u = 1$ for all $u \in \mathcal{A}(S)$ and by [Lemma 7.4](#) all edges are good. So, by [Eq. \(32\)](#), $x(E^\rightarrow(S)) \geq 2 - \epsilon_\eta/2$. Thus, for $\alpha \geq 2\epsilon_\eta$ we have

$$c(s) = (1 + \alpha)x_G \geq (2 - \epsilon_\eta/2)(1 + \alpha) \geq 2 + \epsilon_\eta \stackrel{\text{Eq. (31)}}{\geq} c(t)$$

as desired.

Case 2: $|\mathcal{A}(S)| \geq 5$. By [Theorem 6.21](#) there is at most one bad half edge adjacent to every vertex. Therefore there are at most $|\mathcal{A}(S)|/2$ bad edges, so by [Eq. \(32\)](#),

$$(1 + \alpha)x_G \geq (1 + \alpha) \left(|\mathcal{A}(S)| - 1 - \epsilon_\eta/2 - \frac{1}{2}|\mathcal{A}(S)|\left(\frac{1}{2} + \epsilon_{1/2}\right) \right) \geq 2 + \epsilon_\eta + \epsilon_F |\mathcal{A}(S)| \stackrel{\text{Eq. (31)}}{\geq} c(t)$$

where the second to last inequality holds, using $\alpha \geq 2\epsilon_\eta$, $|\mathcal{A}(S)| \geq 5$, $\epsilon_{1/2} \leq 0.01$, and $\epsilon_F \leq 0.1$.

Case 3: $|\mathcal{A}(S)| = 4$, and we have 0 or 1 bad edges. Then by [Eq. \(32\)](#), $x_G \geq 2.5 - \epsilon_\eta/2 - \epsilon_{1/2}$, so by [Eq. \(31\)](#), $(1 + \alpha)x_G \geq 2 + \epsilon_\eta + 4\epsilon_F \geq c(t)$ for $\epsilon_F \leq 0.1$, $\alpha \geq 2\epsilon_\eta$, $\epsilon_{1/2} \leq 0.01$.

Case 4: $|\mathcal{A}(S)| = 4$, and there are 2 bad edges. Then they form a perfect matching inside S and for each $u \in \mathcal{A}(S)$, $x(\delta^\dagger(u)) \leq 1/2 + 9\epsilon_{1/2}$ (see [Theorem 6.21](#)).

Therefore it must also be the case that $x(\delta^\dagger(u)) \geq \epsilon_F$ for each $u \in \mathcal{A}(S)$. If not, there would have to be a node $u' \in \mathcal{A}(S)$ such that $x(\delta^\dagger(u')) \geq (2 - \epsilon_F)/3 > 1/2 + 9\epsilon_{1/2}$, which is a contradiction to u' having an incident bad edge. Thus, for each $u \in \mathcal{A}(S)$, $x(\delta^\dagger(u))$ is ϵ_F -fractional, i.e., $F_u = 1 - \epsilon_B$ and $Z_u = 1$ implying that $c(t) \leq (2 + \epsilon_\eta)(1 - \epsilon_B)$. Therefore, by [Eq. \(32\)](#),

$$c(s) = (1 + \alpha)x_G \geq (1 + \alpha)(2 - 2\epsilon_{1/2} - \epsilon_\eta/2),$$

and the rightmost quantity is at least $c(t)$ for $\epsilon_B \geq 2\epsilon_{1/2}$ and $\alpha \geq 2\epsilon_\eta$. \square

From now on, we assume that the min s-t cut $A \neq \{s\}$. In the following we will prove that for any set of atoms $W \subsetneq S$, we have:

$$c(s, \delta^\rightarrow(W)) = (1 + \alpha)x_G(\delta^\rightarrow(W)) \geq c(\delta^\dagger(W), t) \tag{33}$$

where for a set F of edges we write $x_G(F)$ to denote the total fractional value of good edges in F .

Let $A_X = A \cap X, A_Y = A \cap Y$ and so on. Assuming the above inequality, let us prove the lemma: First, for the set of edges A_X chosen from X , let Q be the set of endpoints of all edge bundles in A_X (in $\mathcal{A}(S)$).

Observe that we must choose all atoms in Q inside A_Y due to the infinite capacity arcs, i.e., $Q \subseteq A_Y$. Let $W = S \setminus Q$. Note that $W \neq S$. Then:

$$\begin{aligned} c(A, \bar{A}) &= c(A_Y, t) + c(s, \bar{A}_X) \\ &\geq c(\delta^\uparrow(Q), t) + c(s, \delta^\rightarrow(W)) \\ &= c(\delta^\uparrow(S), t) - c(\delta^\uparrow(W)) + c(s, \delta^\rightarrow(W)) \geq c(\delta^\uparrow(S), t), \end{aligned}$$

where the last inequality follows by (33).

Finally, we prove (33). Suppose atoms in W are adjacent to k bad edges. Then

$$x_G(\delta^\rightarrow(W)) = x(\delta^\rightarrow(W)) - x_B(\delta^\rightarrow(W))$$

which by Lemma 7.3 and the fact that each bad edge has fraction at most $1/2 + \epsilon_{1/2}$, is

$$\geq |W| - \epsilon_\eta/2 - k(1/2 + \epsilon_{1/2}). \quad (34)$$

To upper bound $c(\delta^\uparrow(W), t)$, we observe that for any $u \in \mathcal{A}(S)$,

$$c(u, t) \leq \begin{cases} x(\delta^\uparrow(u))Z_u \leq 1/5 & \text{if } x(\delta^\uparrow(u)) < \epsilon_F \\ (1/2 + 9\epsilon_{1/2})(1 - \epsilon_B) & \text{if } x(\delta^\uparrow(u)) > \epsilon_F \text{ and } u \text{ incident to bad edge} \\ 1 + \epsilon_\eta & \text{otherwise, using Lemma 2.37.} \end{cases}$$

Therefore, we can write,

$$c(\delta^\uparrow(W), t) \leq k(1/2 + 9\epsilon_{1/2})(1 - \epsilon_B) + (|W| - k)(1 + \epsilon_\eta).$$

Now, to prove (33), using (34), it is enough to choose α and ϵ_B such that,

$$(1 + \alpha) (|W| - \epsilon_\eta/2 - k(1/2 + \epsilon_{1/2})) \geq k(1/2 + 9\epsilon_{1/2})(1 - \epsilon_B) + (|W| - k)(1 + \epsilon_\eta),$$

or equivalently,

$$|W|(\alpha - \epsilon_\eta) \geq k(\alpha/2 + 10\epsilon_{1/2} + \alpha\epsilon_{1/2} - \epsilon_B/2 - 9\epsilon_B\epsilon_{1/2} - \epsilon_\eta) + \frac{\epsilon_\eta}{2}(1 + \alpha)$$

Since every atom is adjacent to at most one bad edge, $k \leq |W|$ and $|W| \geq 1$, the inequality follows using $\epsilon_B \geq 21\epsilon_{1/2}$ and $\alpha > 2\epsilon_\eta$ and $\epsilon_{1/2} \leq 0.0002$ and $\epsilon_\eta \leq \epsilon_{1/2}^2$. \square

8 Reduction and payment

In this section we prove the main payment theorem, [Theorem 5.33](#).

In [Section 6](#) we defined a number of happy events, such as 2-1-1 happy or 2-2-2 happy and showed that each of these events occurs with probability at least p . In this section, we will subsample these events to define a corresponding decrease event that occurs with probability *exactly*²⁵ p .

Reduction Events.

- **Bottom edges.** For each polygon cut $S \in \mathcal{H}$, let \mathcal{R}_S be the indicator of a uniformly random subset of measure p of the [max flow event](#) \mathcal{E}_S . Note that when $\mathcal{R}_S = 1$ then in particular we know that the polygon S is happy.
- **Top edges.** For a top edge bundle $\mathbf{e} = (u, v)$ define

$$\mathcal{H}_{\mathbf{e},u} = \begin{cases} 1 & \text{if } \mathbf{e} \text{ is 2-1-1 happy and good w.r.t. } u \\ 1 & \text{if } \mathbf{e} \text{ is 2-2 happy and good, but not 2-1-1 good with respect to } u \\ 0 & \text{otherwise.} \end{cases}$$

and let $\mathcal{H}_{\mathbf{e},v}$ be defined similarly. Since p is a lower bound on the probability a good edge is happy, we may now let $\mathcal{R}_{\mathbf{e},u}$ and $\mathcal{R}_{\mathbf{e},v}$ be indicators of subsets of measure p of $\mathcal{H}_{\mathbf{e},u}$ and $\mathcal{H}_{\mathbf{e},v}$ respectively (note $\mathcal{R}_{\mathbf{e},u}$ and $\mathcal{R}_{\mathbf{e},v}$ may overlap). In this way every top edge bundle $\mathbf{e} = (u, v)$ is associated with indicators $\mathcal{R}_{\mathbf{e},u}$ and $\mathcal{R}_{\mathbf{e},v}$. In the special case that u is in case 3 (and not case 1 or 2) of [Theorem 6.35](#), fix two half edge bundles \mathbf{e}, \mathbf{f} that are neighbors of u which satisfy the conditions of case 3. For these edges, by [Theorem 6.35](#), $\mathcal{H}_{\mathbf{e},u} \cap \mathcal{H}_{\mathbf{f},u}$ has measure at least p . This is because $\mathcal{H}_{\mathbf{e},u} \cap \mathcal{H}_{\mathbf{f},u}$ happens if and only if \mathbf{e}, \mathbf{f} are 2-2-2 happy with respect to u . Here, we choose $\mathcal{R}_{\mathbf{e},u}, \mathcal{R}_{\mathbf{f},u}$ to be the same subset of measure p of $\mathcal{H}_{\mathbf{e},u} \cap \mathcal{H}_{\mathbf{f},u}$.

Define $r : E \rightarrow \mathbb{R}_{\geq 0}$ as follows: For any (non-bundle) edge e ,

$$r_e = \begin{cases} \beta x_e \mathcal{R}_S & \text{if } p(e) = S \text{ for a polygon cut } S \in \mathcal{H} \\ \frac{1}{2} \tau x_e (\mathcal{R}_{\mathbf{f},u} + \mathcal{R}_{\mathbf{f},v}) & \text{if } e \in \mathbf{f} \text{ for a top edge bundle } \mathbf{f} = (u, v), \end{cases}$$

for β , the parameter of [Theorem 5.33](#) and τ as defined in [Global constants](#).

Increase Events Let \mathbf{E} be the set of edge bundles, i.e., top/bottom edge bundles. Now, we define the increase vector $I : \mathbf{E} \rightarrow \mathbb{R}_{\geq 0}$ as follows:

²⁵Suppose that under the distribution μ on spanning trees, some event \mathcal{D}' has probability $q \geq p$ and we seek to define an event $\mathcal{D} \subseteq \mathcal{D}'$ that has probability *exactly* p . To this end, one can copy every tree T in the support of μ , exactly $\lfloor \frac{kq}{p} \rfloor$ times for some integer $k > 0$ and whenever we sample T we choose a copy uniformly at random. So, to get a probability exactly p for an event, we say this event occurs if for a “feasible” tree T one of the first k copies are sampled. Now, as $k \rightarrow \infty$ the probability that \mathcal{D} occurs converges to p . Now, for a number of decreasing events, $\mathcal{D}_1, \mathcal{D}_2, \dots$, that occur with probabilities q_1, q_2, \dots (respectively), we just need to let k be the least common multiple of $p/q_1, p/q_2, \dots$ and follow the above procedure. Another method is to choose an independent Bernoulli with success probability p/q for any such event \mathcal{D} .

- **Bottom edges.** For each polygon $S \in \mathcal{H}$ (and corresponding bottom edge bundle) with polygon partition A, B, C , let $r(A) := \sum_{f \in A} r_f$, $r(B) := \sum_{f \in B} r_f$, and $r(C) := \sum_{f \in C} r_f$. Then set

$$I_S := (1 + \epsilon_\eta) \left(\max\{r(A) \cdot \mathbb{I}\{S \text{ not left happy}\}, r(B) \cdot \mathbb{I}\{S \text{ not right happy}\}\} + r(C) \mathbb{I}\{S \text{ not happy}\} \right). \quad (35)$$

- **Top edges.** For every degree cut $S \in \mathcal{H}$, invoke [Lemma 7.2](#) with

$$\alpha = 2\epsilon_\eta, \epsilon_B = 21\epsilon_{1/2}, \epsilon_F = 1/10 \quad (\text{Matching parameters})$$

and let $m_{\mathbf{e},u}$ be the resulting matching for every $u \in \mathcal{A}(S)$. For each top edge bundle $\mathbf{e} = (u, v)$, let

$$I_{\mathbf{e},u} := \sum_{g \in \delta^\dagger(u)} r_g \cdot \frac{m_{\mathbf{e},u}}{\sum_{f \in \delta^\dagger(u)} m_{f,u}} \mathbb{I}\{u \text{ is odd}\}, \quad (36)$$

and define $I_{\mathbf{e},v}$ analogously. Let $I_{\mathbf{e}} = I_{\mathbf{e},u} + I_{\mathbf{e},v}$.

The following theorem is the main technical result of this section.

Theorem 8.1. *For any good top edge bundle \mathbf{e} , $\mathbb{E}[I_{\mathbf{e}}] \leq (1 - \frac{\epsilon_{1/1}}{6}) p \tau x_{\mathbf{e}}$, and for any bottom edge bundle S , $\mathbb{E}[I_S] \leq 0.99994 \beta p$.*

Using this theorem, we can prove the desired theorem:

Theorem 5.33 (Main Payment Theorem). *For an LP solution x^0 and x be x^0 restricted to E and a hierarchy \mathcal{H} for some $\epsilon_\eta \leq 10^{-10}$ and any $\beta > 0$, the maximum entropy distribution μ with marginals x satisfies the following:*

- There is a set of good edges $E_g \subseteq E \setminus \delta(\{u_0, v_0\})$ such that any bottom edge e is in E_g and for any (non-root) $S \in \mathcal{H}$ such that $\mathfrak{p}(S)$ is a degree cut, we have $x(E_g \cap \delta(S)) \geq 3/4$.*
- There is a random vector $s : E_g \rightarrow \mathbb{R}$ (as a function of $T \sim \mu$) such that for all e , $s_e \geq -x_e \beta$ (with probability 1), and*
- If a polygon cut u with polygon partition A, B, C is not left happy, then for any set $F \subseteq E$ with $\mathfrak{p}(e) = u$ for all $e \in F$ and $x(F) \geq 1 - \epsilon_\eta/2$, we have*

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

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

- For every cut $S \in \mathcal{H}$ such that $\mathfrak{p}(S)$ is not a polygon cut, if $\delta(S)_T$ is odd, then $s(\delta(S)) \geq 0$.*
- For a good edge $e \in E_g$, $\mathbb{E}[s_e] \leq -\epsilon_p \beta x_e$ (see [Eq. \(38\)](#) for definition of ϵ_p).*

Proof of [Theorem 5.33](#). First, we set the constants:

$$\epsilon_{1/2} = 0.0002, \epsilon_{1/1} = \frac{\epsilon_{1/2}}{12}, p = 0.005 \epsilon_{1/2}^2, \epsilon_M = 0.00025, \tau = 0.571 \beta \quad (\text{Global constants})$$

Define E_g to be the set of bottom edges together with any edge e which is part of a good top edge bundle. Now, we verify (i): We show for any $S \in \mathcal{H}$ such that $p(S)$ is a degree cut, $x(E_g \cap \delta(S)) \geq 3/4$. First, by [Theorem 6.21](#), if $x(\delta^\uparrow(S)) \geq 1/2 + 9\epsilon_{1/2}$ then all edges in $\delta^\rightarrow(S)$ are good, so the claim follows because by [Lemma 2.37](#), $x(\delta^\rightarrow(S)) \geq 1 - \epsilon_\eta \geq 3/4$. Otherwise, $x(\delta^\uparrow(S)) \leq 1/2 + 9\epsilon_{1/2}$. Then, by [Theorem 6.21](#) there is at most one bad edge in $\delta^\rightarrow(S)$. Therefore, there is a fraction at least $x(\delta^\rightarrow(S)) - (1/2 + \epsilon_{1/2}) \geq 3/4$ of good edges in $\delta^\rightarrow(S)$.

For any edge $e \in E'$ define

$$s_e = -r_e + \begin{cases} I_{\mathbf{f}} \frac{x_e}{x_{\mathbf{f}}} & \text{if } e \in \mathbf{f} \text{ for a top edge bundle } \mathbf{f}, \\ I_S x_e & \text{if } p(e) = S \text{ for a polygon cut } S \in \mathcal{H}. \end{cases} \quad (37)$$

Now, we verify (ii): First, we observe that $s_e = 0$ (with probability 1) if e is part of a bad edge bundle since we defined reduction events only for good edges and $m_{e,u}$ is non-zero only for good edge bundles. Since $r_e \leq \beta x_e$ for bottom edges and $r_e \leq \tau x_e$ for top edges, and $\tau \leq \beta$, it follows that $s_e \geq -x_e \beta$ with probability 1.

Now, we verify (iii): Suppose a polygon cut u is not [left-happy](#). Since u is not happy we must have $\mathcal{R}_u = 0$ and $r_e = 0$ for any $e \in F$. Therefore,

$$\begin{aligned} s(A) + s(F) + s^-(C) &= s(A) + I_S x(F) + s^-(C) \\ &\geq -r(A) + (1 + \epsilon_\eta)(r(A) + r(C))(1 - \epsilon_\eta/2) - r(C) \geq 0. \end{aligned}$$

where we used that $x(F) \geq 1 - \epsilon_\eta/2$.

Now, we verify (iv): Let $S \in \mathcal{H}$, where $p(S)$ is a degree cut. If S is odd, then $r_e = 0$ for all edges $e \in \delta^\rightarrow(S)$; so by [Eq. \(36\)](#)

$$\begin{aligned} s(\delta(S)) &\geq - \sum_{g \in \delta^\uparrow(S)} r_g + \sum_{\mathbf{e} \in \delta^\rightarrow(S)} I_{\mathbf{e},S} \\ &= - \sum_{g \in \delta^\uparrow(S)} r_g + \sum_{\mathbf{e} \in \delta^\rightarrow(S)} \sum_{g \in \delta^\uparrow(S)} r_g \frac{m_{\mathbf{e},S}}{\sum_{\mathbf{f} \in \delta^\rightarrow(S)} m_{\mathbf{f},S}} = 0. \end{aligned}$$

Finally, we verify (v): Here, we use [Theorem 8.1](#). For a good top edge e that is part of a top edge bundle \mathbf{f} we have

$$\mathbb{E}[s_e] = -\mathbb{E}[r_e] + \mathbb{E}[I_{\mathbf{f}}] \frac{x_e}{x_{\mathbf{f}}} \leq -\tau p x_e + (1 - \frac{\epsilon_{1/1}}{6}) p \tau x_e = -\frac{\epsilon_{1/1}}{6} p \tau x_e.$$

On the other hand, for a bottom edge e with $p(e) = S$, then

$$\mathbb{E}[s_e] = -\mathbb{E}[r_e] + \mathbb{E}[I_S] x_e \leq -\beta p x_e + 0.99994 p \beta x_e \leq -0.00006 p \beta x_e.$$

Finally, we can let

$$\epsilon_p := \frac{\epsilon_{1/1}}{6} p \frac{\tau}{\beta} = \frac{\epsilon_{1/2}}{72} 0.005 \epsilon_{1/2}^2 0.571 \geq 0.000039 \epsilon_{1/2}^3 \geq 3.12 \cdot 10^{-16} \quad (38)$$

as desired. □

Name	Value	Set In	Explanation
$\epsilon_{1/2}$	0.0002	Global constants	Half edge threshold, Definition 6.19
$\epsilon_{1/1}$	$\frac{\epsilon_{1/2}}{12}$	Global constants	A, B, C partitioning threshold, Definition 15.5
p	$0.005\epsilon_{1/2}^2$	Global constants	Min prob. of happiness for a (2-*) good edge
ϵ_M	0.00025	Global constants	Marginal errors due to max flow, Definition 6.15
τ	0.571β	Global constants	Top edge decrease
ϵ_P	$\frac{\epsilon_{1/1}}{6} p \frac{\tau}{\beta}$	(38)	Expected decrease constant, Theorem 5.33
α	$2\epsilon_\eta$	Matching parameters	Parameter of Lemma 7.2
ϵ_B	$21\epsilon_{1/2}$	Matching parameters	Parameter of Lemma 7.2
ϵ_F	1/10	Matching parameters	Parameter of Lemma 7.2
ϵ_η	14η	(16)	Definition 5.31
η	$\frac{1}{1308}\epsilon_P$	(62)	Near min cut constant
β	$\eta/4.1$	(8)	Slack shift constant e.g. Theorems 5.2, 5.33 and C.12

Table 3: A table of all constants used in the paper.

In the rest of this section we prove [Theorem 8.1](#). Throughout the proof, we will repeatedly use the following facts proved in [Section 6](#): If a top edge $e = (u, v)$ that is part of a bundle \mathbf{f} is reduced (equivalently $\mathcal{H}_{\mathbf{f},u} = 1$ or $\mathcal{H}_{\mathbf{f},v} = 1$), then u and v are trees, which means that tree sampling inside u and v is independent of the reduction of e .

Note however, that conditioning on a near-min-cut or atom to be a tree increases marginals inside and reduces marginals outside as specified by [Lemma 2.28](#). Since for any $S \in \mathcal{H}$, $x(\delta(S)) \leq 2 + \epsilon_\eta$, the overall change is $\pm\epsilon_\eta/2$.

The proof of [Theorem 8.1](#) simply follows from [Lemma 8.2](#) and [Lemma 8.7](#) that we will prove in the following two sections.

8.1 Increase for Good Top Edges

The following lemma is the main result of this subsection.

Lemma 8.2 (Top Edge Increase). *Let $S \in \mathcal{H}$ be a degree cut and $\mathbf{e} = (u, v)$ a good edge bundle with $p(\mathbf{e}) = S$. If $\epsilon_{1/2} \leq 0.0002$, $\epsilon_{1/1} \leq \epsilon_{1/2}/12$ and $\epsilon_\eta \leq \frac{\epsilon_{1/1}}{100}$, $\epsilon_F = 1/10$ then*

$$\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] \leq p\tau x_{\mathbf{e}} \left(1 - \frac{\epsilon_{1/1}}{6}\right).$$

We will use the following technical lemma to prove the above lemma.

Lemma 8.3. *Let $S \in \mathcal{H}$ be a degree cut with an atom $u \in \mathcal{A}(S)$. If $x(\delta^\dagger(u)) > \epsilon_F$, $\epsilon_{1/2} \leq 0.0002$,*

$\epsilon_{1/1} \leq \epsilon_{1/2}/12$, $\epsilon_\eta \leq \frac{\epsilon_{1/1}}{100}$, then we have

$$\begin{aligned} & \sum_{\substack{g \in \delta^\dagger(u), \\ g \in \mathbf{f}=(u',v') \text{ good top}}} \frac{1}{2} \tau x_g \cdot (\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}] + \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},v'}]) \\ & + \sum_{g \in \delta^\dagger(u), \mathbf{p}(g)=S' \text{ polygon}} \beta x_g \cdot \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{S'}] \leq \tau(1 - \frac{\epsilon_{1/1}}{5})x(\delta^\dagger(u))F_u, \end{aligned} \quad (39)$$

where recall we set $F_u := 1 - \epsilon_B \mathbb{I}\{x(\delta^\dagger(u)) \text{ is } \epsilon_F \text{ fractional}\}$ in [Lemma 7.2](#), where $\epsilon_B := 21\epsilon_{1/2}$ and $\epsilon_F = 1/10$ as in [Matching parameters](#).

Proof of Lemma 8.2. By linearity of expectation and using [Eq. \(36\)](#):

$$\begin{aligned} \mathbb{E}[I_{\mathbf{e},u}] &= \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f} \in \delta^\rightarrow(u)} m_{\mathbf{f},u}} \mathbb{E} \left[\sum_{g \in \delta^\dagger(u)} r_g \cdot \mathbb{I}\{u \text{ is odd}\} \right] \\ &= \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f} \in \delta^\rightarrow(u)} m_{\mathbf{f},u}} \left(\sum_{\substack{g \in \delta^\dagger(u): \\ g \in \mathbf{f}=(u',v') \text{ good top}}} \frac{1}{2} \tau x_g (\mathbb{P}[\mathcal{R}_{\mathbf{f},u'}, \delta(u)_T \text{ odd}] + \mathbb{P}[\mathcal{R}_{\mathbf{f},v'}, \delta(u)_T \text{ odd}]) \right. \\ & \quad \left. + \sum_{g \in \delta^\dagger(u): \mathbf{p}(g)=S' \text{ polygon}} \beta x_g \mathbb{P}[\mathcal{R}_{S'}, \delta(u)_T \text{ odd}] \right) \end{aligned} \quad (40)$$

A similar equation holds for $\mathbb{E}[I_{\mathbf{e},v}]$.

The case where $x(\delta^\dagger(u)) \leq \epsilon_F$ or $x(\delta^\dagger(v)) \leq \epsilon_F$ is dealt with in [Lemma 8.6](#). So, consider the case where $x(\delta^\dagger(u)), x(\delta^\dagger(v)) > \epsilon_F$. Now recall that from [\(30\)](#),

$$\sum_{\mathbf{f} \in \delta^\rightarrow(u)} m_{\mathbf{f},u} = Z_u x(\delta^\dagger(u)) \quad (41)$$

where $Z_u = 1 + \mathbb{I}\{|S| \geq 4, x(\delta^\dagger(u)) \leq \epsilon_F\}$. In this case, $Z_u = Z_v = 1$.

Using $\mathbb{P}[\mathcal{R}_{\mathbf{f},u'}, \delta(u)_T \text{ odd}] = p \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f},u'}]$, and plugging [\(39\)](#) into [\(40\)](#) for u and v , we get (and using [Eq. \(41\)](#)):

$$\begin{aligned} \mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] &\leq p\tau(1 - \frac{\epsilon_{1/1}}{5}) \left(x(\delta^\dagger(u))F_u \frac{m_{\mathbf{e},u}}{x(\delta^\dagger(u))} + x(\delta^\dagger(v))F_v \frac{m_{\mathbf{e},v}}{x(\delta^\dagger(v))} \right) \\ &= p\tau(1 - \frac{\epsilon_{1/1}}{5})(F_u m_{\mathbf{e},u} + F_v m_{\mathbf{e},v}) \\ &\leq p\tau(1 - \frac{\epsilon_{1/1}}{5})(1 + 2\epsilon_\eta)x_{\mathbf{e}} < p\tau x_{\mathbf{e}}(1 - \frac{\epsilon_{1/1}}{6}). \end{aligned} \quad (42)$$

where on the final line we used [\(29\)](#) and $\epsilon_\eta < \frac{\epsilon_{1/1}}{100}$. \square

Proof of Lemma 8.3. Suppose that $S_i \in \mathcal{H}$ are the ancestors of S in the hierarchy (in order) such $S_1 = S$ and for each i , $S_{i+1} = \mathbf{p}(S_i)$. Let

$$\delta^{\geq i} := \delta(u) \cap \delta(S_i) \quad \text{and} \quad \delta^i := \delta(u) \cap \delta^\rightarrow(S_i).$$

Each group of edges δ^i is either entirely top edges or entirely bottom edges. First note that if $g \in \delta^i$ and g is a bottom edge, i.e., S_{i+1} is a polygon cut, then by [Corollary 6.17](#),

$$\mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{S_{i+1}}] = \mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{E}_{S_{i+1}}] \leq 0.5678$$

(see [Definition 6.15](#) and [Section 8](#) for definition of $\mathcal{E}_{S_{i+1}}, \mathcal{R}_{i+1}$) where in the equality we used that $\mathcal{R}_{S_{i+1}}$ is a uniformly random event chosen in $\mathcal{E}_{S_{i+1}}$. Therefore, to prove [Eq. \(39\)](#) it is enough to show

$$\begin{aligned} & \sum_{\substack{g \in \delta_{\text{good}}^\uparrow(u): \\ g \in \mathbf{f} = (u', v') \text{ top,}}} \frac{1}{2} \tau x_g (\mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] + \mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, v'}]) \\ & \leq \tau \left(\left(1 - \frac{\epsilon_{1/1}}{5}\right) F_u \left(x(\delta_{\text{good}}^\uparrow(u)) + x(\delta_{\text{bad}}^\uparrow(u)) \right) + 0.0014 x(\delta_\beta^\uparrow(u)) \right) \end{aligned} \quad (43)$$

where we write $\delta_\beta(u), \delta_{\text{good}}(u), \delta_{\text{bad}}(u)$ to denote the set of bottom edges, good top edges, and bad (top) edges in $\delta(u)$ respectively and we used that

$$\tau \left(1 - \frac{\epsilon_{1/1}}{5}\right) (1 - \epsilon_B) - 0.5678\beta \geq 0.0014\tau$$

since $\tau = 0.571\beta$, $\epsilon_{1/1} \leq \frac{\epsilon_{1/2}}{12}$, $\epsilon_{1/2} \leq 0.0002$, and $\epsilon_B = 21\epsilon_{1/2}$ as defined in [Matching parameters](#).

Since $h(\mathbf{f}) := \frac{1}{2} (\mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] + \mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, v'}]) \leq 1$ and $(1 - \frac{\epsilon_{1/1}}{5})F_u$ is nearly 1, in each of the following cases

$$x(\delta_\beta^\uparrow(u)) \geq \begin{cases} 0.003 & \text{when } F_u = 1 \\ \frac{4}{5}x(\delta^\uparrow(u)) & \text{when } F_u = 1 - \epsilon_B \end{cases} \quad \text{or} \quad x(\delta_{\text{bad}}^\uparrow(u)) \geq 0.006 \quad \text{when } F_u \geq 1 - \epsilon_B, \quad (44)$$

[\(43\)](#) holds. To see this, just plug in $\epsilon_{1/1} \leq \frac{\epsilon_{1/2}}{12}$, $\epsilon_{1/2} \leq 0.0002$, $\epsilon_B = 21\epsilon_{1/2}$, $\epsilon_\eta \leq 10^{-10}$, $x(\delta^\uparrow(u)) \leq 1 + \epsilon_\eta$ and any inequality from [\(44\)](#) into [\(43\)](#), using the upper bound $h(\mathbf{f}) = 1$.

Alternatively, for $\delta_{\text{top}}(u) = \delta_{\text{good}}(u) \cup \delta_{\text{bad}}(u)$ be the set of top edges in $\delta(u)$, if we can show the existence of a set $D \subseteq \delta_{\text{top}}^\uparrow(u)$ such that

$$x(D) \cdot \min_{\substack{g \in D: \\ g \in \mathbf{f} = (u', v') \text{ good}}} 1 - \frac{\mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] + \mathbb{P} [\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, v'}]}{2} \geq \left(\frac{\epsilon_{1/1}}{5} + 1 - F_u \right) x(\delta_{\text{top}}^\uparrow(u)), \quad (45)$$

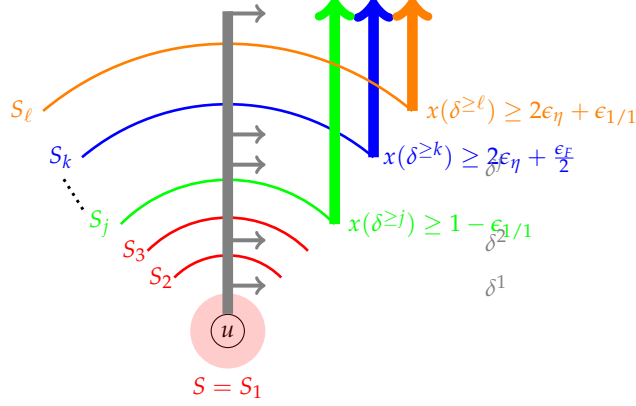
then, again, [\(43\)](#) holds.

In the rest of the proof, we will consider a number of cases and show that in each of them, either one of the inequalities in [\(44\)](#) or the inequality in [\(45\)](#) for some set D is true, which will imply the lemma.

First, let

$$\begin{aligned} j &= \max\{i : x(\delta^{\geq i}) \geq 1 - \epsilon_{1/1}\} \\ k &= \max\{i : x(\delta^{\geq i}) \geq 2\epsilon_\eta + \epsilon_F/2\}, \\ \ell &= \max\{i : x(\delta^{\geq i}) \geq 2\epsilon_\eta + \epsilon_{1/1}\} \end{aligned}$$

Just note $j \leq k \leq \ell$. Note that levels ℓ and k exist since $x(\delta^\uparrow(u)) \geq \epsilon_F$, whereas level j may not exist (if $x(\delta^\uparrow(u)) < 1 - \epsilon_{1/1}$). We consider three cases:



Case 1: $x(\delta^\uparrow(u)) \geq 1 - \epsilon_{1/1}$: Then j exists and S_j has a valid A, B, C degree partitioning (Definition 15.5) where $A = \delta(v) \cap \delta(S_j)$ such that either $u = v$ or v is a descendant of u in \mathcal{H} . Note that, $x(\delta(u) \cap \delta(S_j)) \geq 1 - \epsilon_{1/1}$, and by Definition 15.5, $B \cap \delta(u) = \emptyset$. In addition, in this case, $x(\delta^\uparrow(u))$ is not ϵ_F fractional (see Lemma 7.2), so $F_u = 1$.

Case 1a: $x(\delta^j) \geq 3/4$. If δ^j are bottom edges then (44) holds. So, suppose that δ^j is a set of top edges. By Lemma 6.32, at most $1/2 + 4\epsilon_{1/2}$ fraction of edges in $A \cap \delta^j$ are good but not 2-1-1 good (w.r.t., u). So, the rest of the edges in $A \cap \delta^j$ are either bad or 2-1-1 good. Since

$$x(A \cap \delta^j) \geq 3/4 - x(C) \geq 3/4 - 2\epsilon_{1/1} - \epsilon_\eta,$$

δ^j either has a mass of $\frac{1}{2}(1/4 - 2\epsilon_{1/1} - \epsilon_\eta - 4\epsilon_{1/2}) > 1/8 - 3\epsilon_{1/2}$ of bad edges or of 2-1-1 good edges.²⁶ The former case implies that (44) holds. In the latter case, by Claim 8.4 for any 2-1-1 good edge $g \in \delta^j$ with $g \in \mathbf{f} = (u', v')$ we have $\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] \leq 2\epsilon_\eta + \epsilon_{1/1}$; so (45) holds for D defined as the set of 2-1-1 good edges in δ^j .

Case 1b: $x(\delta^j) < 3/4$. If $x(\delta_\beta^\uparrow(u)) \geq 0.003$, then (44) holds. Otherwise, we apply Claim 8.5 with $\epsilon = \epsilon_{1/1}$ to all good top edge bundles $\mathbf{f} \in D = \delta^{\geq j+1} \setminus \delta^{\geq \ell+1}$ and we get that

$$\frac{1}{2}(\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] + \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, v'}]) \leq 1 - \epsilon_{1/1} + \epsilon_{1/1}^2.$$

Since $x(D) \geq 1 - \epsilon_{1/1} - 3/4 - 2\epsilon_\eta - \epsilon_{1/1} - 0.003 > 0.24$, (45) holds.

Case 2: $1 - \epsilon_F < x(\delta^\uparrow(u)) < 1 - \epsilon_{1/1}$. Again we have $F_u = 1$. So we can either show that $x(\delta_\beta^\uparrow(u)) \geq 0.003$ or take D to be the top edges in $\delta^\uparrow(u) \setminus \delta^{\geq \ell+1}$ and use Claim 8.5 with $\epsilon = \epsilon_{1/1}$. This will enable us to show that (45) holds as in the previous case.

Case 3: $\epsilon_F < x(\delta^\uparrow(u)) < 1 - \epsilon_F$: In this case $F_u = 1 - \epsilon_B$. If at least $4/5$ of the edges in $\delta^\uparrow(u)$ are bottom edges, then we are done by (44).

Otherwise, let $u' = \rho(u)$. For any top edge $e \in \delta^\uparrow(u)$ where $e \in \mathbf{f} = (u'', v'')$ we have

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u''}] \leq \mathbb{P}[u' \text{ tree} | \mathcal{R}_{\mathbf{f}, u''}] \mathbb{P}[\delta(u)_T \text{ odd} | u' \text{ tree}, \mathcal{R}_{\mathbf{f}, u''}] + \mathbb{P}[u' \text{ not tree} | \mathcal{R}_{\mathbf{f}, u''}]$$

²⁶We are using the fact that $\epsilon_{1/1} = \epsilon_{1/2}/12$ and that ϵ_η is tiny by comparison to these.

Using that $u' \subseteq u''$ is a tree under $|\mathcal{R}_{\mathbf{f}, u''}$ with probability at least $1 - \epsilon_\eta/2$, and applying [Claim 8.5](#) (to u and u') with $\epsilon = \epsilon_F$ we have $\mathbb{P}[\delta(u)_T \text{ odd} | u' \text{ tree}, \mathcal{R}_{\mathbf{f}, u''}] \leq 1 - \epsilon_F + \epsilon_F^2$ we get

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u''}] \leq 1 - \epsilon_F + \epsilon_F^2 + \epsilon_\eta/2.$$

Now, let D be all top edges in $\delta^\uparrow(u)$. Then, we apply [Eq. \(45\)](#) to this set of mass at least $x(\delta^\uparrow(u))/5$, and we are done, using that $(\epsilon_F - 2\epsilon_F^2)/5 \geq (\frac{\epsilon_{1/1}}{5} + \epsilon_B)$ which holds for $\epsilon_F \geq 1/10$, $\epsilon_B = 21\epsilon_{1/2}$, and $\epsilon_{1/2} \leq 0.0002$. \square

Claim 8.4. For $u \in \mathcal{H}$ and a top edge $e \in \mathbf{f} = (u', v')$ for some $u' \in \mathcal{H}$ that is an ancestor of u , if $x(\delta(u) \cap \delta(u')) \geq 1 - \epsilon_{1/1}$ and \mathbf{f} is 2-1-1 good, then

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] \leq 2\epsilon_\eta + \epsilon_{1/1}.$$

Proof. Let A, B, C be the [degree partitioning](#) of $\delta(u')$. By the assumption of the claim, without loss of generality, assume $A \subseteq \delta(u) \cap \delta(u')$. Furthermore, by definition, $B \cap \delta(u) = \emptyset$. This means that if $\mathcal{R}_{\mathbf{f}, u'} = 1$ then u' is a tree and $A_T = 1 = (\delta(u) \cap \delta(u'))_T$ (also using $C_T = 0$ and $B \cap \delta(u) = \emptyset$). Therefore,

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{\mathbf{f}, u'}] = \mathbb{P}[(\delta(u) \setminus \delta(u'))_T \text{ even} | \mathcal{R}_{\mathbf{f}, u'}].$$

To upper bound the RHS first observe that

$$\mathbb{E}[(\delta(u) \setminus \delta(u'))_T | \mathcal{R}_{\mathbf{f}, u'}] \leq \epsilon_\eta/2 + x(\delta(u) \setminus \delta(u')) \leq \epsilon_\eta/2 + x(\delta(u)) - x(A) < 1 + 2\epsilon_\eta + \epsilon_{1/1}.$$

Under the conditional measure $|\mathcal{R}_{\mathbf{f}, u'}$, u' is a tree, so u must be connected inside u' , i.e., $(\delta(u) \setminus \delta(u'))_T \geq 1$ with probability 1. Therefore,

$$\mathbb{P}[(\delta(u) \setminus \delta(u'))_T \text{ even} | \mathcal{R}_{\mathbf{f}, u'}] \leq \mathbb{P}[(\delta(u) \setminus \delta(u'))_T - 1 \neq 0 | \mathcal{R}_{\mathbf{f}, u'}] \leq 2\epsilon_\eta + \epsilon_{1/1}$$

as desired. \square

Claim 8.5. For $u, u' \in \mathcal{H}$ such that u' is an ancestor of u . Let $\nu = \nu_{u'} \times \nu_{G/u'}$ be the measure resulting from conditioning u' to be a tree. if $x(\delta(u) \cap \delta(u')) \in [\epsilon, 1 - \epsilon]$, then

$$\mathbb{P}_\nu[\delta(u) \text{ odd} | (\delta(u) \cap \delta(u'))_T] \leq 1 - \epsilon + \max\{2\epsilon_\eta, \epsilon^2\}. \quad (46)$$

In other words, for any integer $k \geq 0$, we have $\mathbb{P}_\nu[\delta(u) \text{ odd} | (\delta(u) \cap \delta(u'))_T = k] \leq 1 - \epsilon + \max\{2\epsilon_\eta, \epsilon^2\}$.

Proof. Let $D = \delta(u) \setminus \delta(u')$. By assumption, u' is a tree, so $D_T \geq 1$ with probability 1. Therefore, since we have no control over the parity of $(\delta(u) \cap \delta(u'))_T$

$$\mathbb{P}_\nu[\delta(u)_T \text{ even} | (\delta(u) \cap \delta(u'))_T] \geq \min\{\mathbb{P}[D_T - 1 \text{ odd} | u' \text{ tree}], \mathbb{P}[D_T - 1 = 0 | u' \text{ tree}]\}$$

where we removed the conditioning by taking the worst case over $(\delta(u) \cap \delta(u'))_T$ even, $(\delta(u) \cap \delta(u'))_T$ odd. First, observe by the assumption of the claim and that $x(\delta^\uparrow(u)) \leq 2 + \epsilon_\eta$ we have

$$\mathbb{E}[D_T - 1 | u' \text{ tree}] \in [\epsilon, 1 - \epsilon + 2\epsilon_\eta].$$

Furthermore, since we have a SR distribution on $G[u']$, $D_T - 1$ is a [Bernoulli sum](#) random variable. Therefore,

$$\mathbb{P}[D_T - 1 = 0 | u' \text{ tree}] \geq \epsilon - 2\epsilon_\eta$$

and by [Lemma 2.21](#)

$$\mathbb{P}[D_T - 1 \text{ odd} | u' \text{ tree}] \geq 1 - 1/2(1 + e^{-2\epsilon}) \geq \epsilon - \epsilon^2$$

as desired. \square

Lemma 8.6. Let $S \in \mathcal{H}$ be a degree cut and $\mathbf{e} = (u, v)$ a good edge bundle with $\rho(\mathbf{e}) = S$. If $x(\delta^\uparrow(u)) < \epsilon_F$, $\epsilon_{1/2} \leq 0.0002$, $\epsilon_{1/1} \leq \epsilon_{1/2}/10$, then,

$$\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] \leq p\tau x_{\mathbf{e}} \left(1 - \frac{\epsilon_{1/1}}{6}\right)$$

Proof. First notice, by [Corollary 6.17](#) for any bottom edge $g \in \delta^\uparrow(u)$ with $\rho(g) = S'$, we have

$$\mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{R}_{S'}] = \mathbb{P}[\delta(u)_T \text{ odd} | \mathcal{E}_{S'}] \leq 0.5678,$$

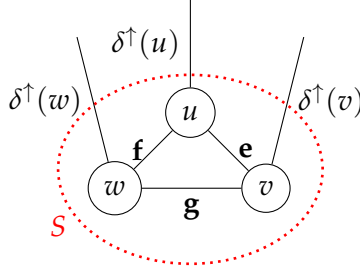
using $0.5678\beta \leq \tau$ and $F_u = 1$ (as $x(\delta^\uparrow(u)) \leq \epsilon_F$) we can write,

$$\mathbb{E}[I_{\mathbf{e},u}] \leq \sum_{h \in \delta^\uparrow(u)} x_h p\tau F_u \cdot \frac{m_{\mathbf{e},u}}{Z_u x(\delta^\uparrow(u))}. \quad (47)$$

Secondly, if $x(\delta^\uparrow(v)) \geq \epsilon_F$, applying [\(39\)](#) and [\(40\)](#) to $I_{\mathbf{e},v}$ and using $Z_v \geq 1$ we get

$$\mathbb{E}[I_{\mathbf{e},v}] \leq \frac{m_{\mathbf{e},v}}{\sum_{\mathbf{f} \in \delta^\rightarrow(v)} m_{\mathbf{f},v}} p\tau \left(1 - \frac{\epsilon_{1/1}}{5}\right) x(\delta^\uparrow(v)) F_v = m_{\mathbf{e},v} p\tau \left(1 - \frac{\epsilon_{1/1}}{5}\right) F_v \quad (48)$$

Case 1: $|\mathcal{A}(S)| = 3$, where $\mathcal{A}(S) = \{u, v, w\}$. Let $\mathbf{f} = (u, w)$, $\mathbf{g} = (v, w)$ (and of course $\mathbf{e} = (u, v)$). We will use the following facts below:



$$\begin{aligned} x_{\mathbf{e}} + x_{\mathbf{f}} &\geq 2 - \epsilon_F && (x(\delta(u)) \geq 2 \text{ and } x(\delta^\uparrow(u)) \leq \epsilon_F) \\ x(\delta^\uparrow(v)) + x(\delta^\uparrow(w)) &\geq 2 - \epsilon_F && (x(\delta(S)) \geq 2) \\ x_{\mathbf{f}}, x(\delta^\uparrow(w)) &\leq 1 + \epsilon_\eta, && \text{(Lemma 2.37)} \end{aligned}$$

so we have,

$$x_{\mathbf{e}}, x(\delta^\uparrow(v)) \geq 1 - \epsilon_F - \epsilon_\eta. \quad (49)$$

Now we bound $\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}]$. By [Eq. \(47\)](#) and [Eq. \(48\)](#) (which we may apply to $\mathbb{E}[I_{\mathbf{e},v}]$ since $x(\delta^\uparrow(v)) \geq \epsilon_F$),

$$\begin{aligned} \mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] &\leq \sum_{h \in \delta^\uparrow(u)} x_h p\tau F_u \cdot \frac{m_{\mathbf{e},u}}{Z_u x(\delta^\uparrow(u))} + p\tau \left(1 - \frac{\epsilon_{1/1}}{5}\right) F_v m_{\mathbf{e},v} \\ &= p\tau F_u m_{\mathbf{e},u} + p\tau \left(1 - \frac{\epsilon_{1/1}}{5}\right) F_v m_{\mathbf{e},v} && (Z_u = 1 \text{ as } |\mathcal{A}(S)| = 3) \\ &= p\tau (F_u m_{\mathbf{e},u} + F_v m_{\mathbf{e},v}) - \frac{\epsilon_{1/1}}{5} p\tau F_v m_{\mathbf{e},v} \\ &\leq p\tau (1 + 2\epsilon_\eta) x_{\mathbf{e}} - \frac{\epsilon_{1/1}}{5} p\tau F_v m_{\mathbf{e},v} \end{aligned} \quad (50)$$

where the final inequality follows from (29). To complete the proof, we lower bound $m_{\mathbf{e},v}$.

Using (30) for v and w , we can write,

$$\begin{aligned}
x(\delta^\dagger(v)) + x(\delta^\dagger(w)) &= m_{\mathbf{e},v} + m_{\mathbf{g},v} + m_{\mathbf{f},w} + m_{\mathbf{g},w} \\
&\leq m_{\mathbf{e},v} + \frac{(1+2\epsilon_\eta)}{(1-\epsilon_B)}(x_{\mathbf{f}} + x_{\mathbf{g}}) && \text{(using (29))} \\
&= m_{\mathbf{e},v} + \frac{(1+2\epsilon_\eta)}{(1-\epsilon_B)} \left(\sum_{a \in \mathcal{A}(S)} \frac{x(\delta(a))}{2} - \frac{x(\delta(S))}{2} - x_{\mathbf{e}} \right) \\
&\leq m_{\mathbf{e},v} + \frac{(1+2\epsilon_\eta)}{(1-\epsilon_B)}(2 + 3\epsilon_\eta - x_{\mathbf{e}})
\end{aligned}$$

and using the fact that $x(\delta^\dagger(v)) + x(\delta^\dagger(w)) \geq 2 - \epsilon_F$, we get

$$m_{\mathbf{e},v} \geq x_{\mathbf{e}} - \epsilon_F - 4\epsilon_B \geq (1 - 1.2\epsilon_F)x_{\mathbf{e}},$$

where the second inequality follows from (49) and $\epsilon_B = 21\epsilon_{1/2}$ and $\epsilon_\eta < \epsilon_{1/2}^2$ and $\epsilon_F \geq 1/10$. Plugging this back into (50) and using $F_v \geq 1 - \epsilon_B = 1 - 21\epsilon_{1/2}$ we get

$$\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] \leq p\tau x_{\mathbf{e}} \left(1 + 2\epsilon_\eta - \frac{\epsilon_{1/1}}{5}(1 - 1.2\epsilon_F)(1 - 21\epsilon_{1/2}) \right) \leq p\tau x_{\mathbf{e}} \left(1 - \frac{\epsilon_{1/1}}{6} \right)$$

as desired. In the last inequality we used $\epsilon_F \leq 1/10$ and $\epsilon_{1/2} \leq 0.0002$.

Case 2: $|S| \geq 4$. In this case, $Z_u = 2$. Therefore, by Eq. (47)

$$\mathbb{E}[I_{\mathbf{e},u}] \leq \sum_{e \in \delta^\dagger(u)} x_e p\tau F_u \frac{m_{\mathbf{e},u}}{Z_u x(\delta^\dagger(u))} = \frac{1}{2} p\tau F_u m_{\mathbf{e},u}.$$

If $x(\delta^\dagger(v)) < \epsilon_F$, we get the same inequality for $I_{\mathbf{e},v}$. Then,

$$\mathbb{E}[I_{\mathbf{e},u}] + \mathbb{E}[I_{\mathbf{e},v}] \leq \frac{1}{2} p\tau (F_u m_{\mathbf{e},u} + F_v m_{\mathbf{e},v}) \stackrel{(29)}{\leq} \frac{1}{2} p\tau x_{\mathbf{e}} (1 + 2\epsilon_\eta),$$

which is clearly sufficient for the lemma statement.

Otherwise, $x(\delta^\dagger(v)) \geq \epsilon_F$ in which case by (48) we get $\mathbb{E}[I_{\mathbf{e},v}] \leq m_{\mathbf{e},v} p\tau F_v (1 - \epsilon_{1/1}/5)$. We conclude the lemma similar to the previous case. \square

8.2 Increase for Bottom Edges

The following lemma is the main result of this subsection.

Lemma 8.7 (Bottom Edge Increase). *If $\epsilon_{1/2} \leq 0.0002$, $\epsilon_\eta \leq \epsilon_{1/2}^2$, for any polygon cut $S \in \mathcal{H}$,*

$$\mathbb{E}[I_S] \leq 0.99994\beta p.$$

Proof. For a set of edges $D \subseteq \delta(S)$ define the random variable.

$$\begin{aligned}
I_S(D) &:= (1 + \epsilon_\eta) (\max\{r(A \cap D) \mathbb{I}\{S \text{ not left happy}\}, r(B \cap D) \mathbb{I}\{S \text{ not right happy}\}\} \\
&\quad + r(C \cap D) \mathbb{I}\{S \text{ not happy}\}).
\end{aligned} \tag{51}$$

Note that by definition $I_S(\delta(S)) = I_S$ and for any two disjoint sets D_1, D_2 , $I_S(D_1 \cup D_2) \leq I_S(D_1) + I_S(D_2)$. Also, define $I_S^\uparrow = I_S(\delta^\uparrow(S))$ and $I_S^\rightarrow = I_S(\delta^\rightarrow(S))$.

First, we upper bound $\mathbb{E}[I_S^\uparrow]$. Let $f \in \delta^\uparrow(S)$ and suppose that f with $p(f) = S'$ is a bottom edge. Say we have $f \in A^\uparrow(S)$ ($f \in B^\uparrow(S)$ is similar). We write,

$$\begin{aligned} \mathbb{E}[I_S(f)] &= (1 + \epsilon_\eta)\beta x_f \mathbb{P}[\mathcal{R}_{S'}] \mathbb{P}[S \text{ not left happy} | \mathcal{R}_{S'}] \\ &\leq 0.568x_f p\beta \leq x_f p\tau \end{aligned}$$

where in the inequality we used [Corollary 6.18](#) and that

$$\mathbb{P}[S \text{ not left happy} | \mathcal{R}_S] = \mathbb{P}[S \text{ not left happy} | \mathcal{E}_S]$$

since \mathcal{R}_S is a uniformly random subset of \mathcal{E}_S . If $f \in C^\uparrow(S)$, we use the trivial guarantee $\mathbb{E}[I_S(f)] \leq (1 + \epsilon_\eta)x_f p\beta$.

On the other hand, if f is a top edge, then we use the trivial bound

$$\mathbb{E}[I_S(f)] \leq (1 + \epsilon_\eta)\tau p x_f. \quad (52)$$

Therefore,

$$\mathbb{E}[I_S^\uparrow] \leq (1 + \epsilon_\eta)\tau p x(\delta^\uparrow(S)) + (1 + \epsilon_\eta)\epsilon_\eta p\beta \leq (1 + \epsilon_\eta)(0.571)\beta p x(\delta^\uparrow(S)) + 2\epsilon_\eta p\beta \quad (53)$$

since $x(C) \leq \epsilon_\eta$.

Now, we consider three cases:

Case 1: $\hat{S} = p(S)$ is a degree cut. Combining (53) and [Lemma 8.8](#) below, we get

$$\mathbb{E}[I_S] \leq (1 + \epsilon_\eta)p(0.571)\beta(7/4 + 6\epsilon_{1/2} + \epsilon_\eta) + 2\epsilon_\eta p\beta \leq 0.99994\beta p$$

using $\epsilon_{1/2} \leq 0.0002$ and $\epsilon_\eta \leq \epsilon_{1/2}^2$.

Case 2: $\hat{S} = p(S)$ is a polygon cut with ordering u_1, \dots, u_k of $\mathcal{A}(\hat{S})$, $S = u_1$ or $S = u_k$ Then, by [Lemma 8.9](#) below,

$$\mathbb{E}[I_S] \leq (1 + \epsilon_\eta)\beta p(0.571x(\delta^\uparrow(S)) + 0.31) + 2\epsilon_\eta p\beta \leq 0.89\beta p$$

where we used $x(\delta^\uparrow(S)) \leq 1 + \epsilon_\eta$.

Case 3: $\hat{S} = p(S)$ is a polygon cut with ordering u_1, \dots, u_k of $\mathcal{A}(\hat{S})$, $S \neq u_1, u_k$ Then, by [Lemma 8.11](#) below

$$\mathbb{E}[I_S] \leq (1 + \epsilon_\eta)\beta p(0.571x(\delta^\uparrow(S)) + 0.85) + 2\epsilon_\eta p\beta \leq 0.86\beta p$$

where we use that $x(\delta^\uparrow(S)) \leq \epsilon_\eta$ since we have a [hierarchy](#). This concludes the proof. \square

8.2.1 Case 1: \hat{S} is a degree cut

Lemma 8.8. *Let $S \in \mathcal{H}$ be a polygon cut with parent \hat{S} which is a degree cut. Then*

$$\mathbb{E}[I_S^\rightarrow] \leq (1 + \epsilon_\eta)p\tau(x(\delta^\rightarrow(S)) - (1/4 - 6\epsilon_{1/2})).$$

Proof. Let A, B, C be the polygon partition of S . We will show that for a constant fraction of the edges in $\delta^{\rightarrow}(S)$, we can improve over the trivial bound in (52). To this end, consider the cases given by [Theorem 6.35](#).

Case 1: There is a bad half edge e in $\delta^{\rightarrow}(S)$. Since bad edges never decrease, no corresponding increase occurs, so by the trivial bound [Eq. \(52\)](#)

$$\mathbb{E} [I_S^{\rightarrow}] \leq (1 + \epsilon_\eta) p \tau (x(\delta^{\rightarrow}(S)) - (1/2 - \epsilon_{1/2})).$$

This concludes the proof.

Case 2: There is a set of 2-1-1 good edges (w.r.t., S) $D \subseteq \delta^{\rightarrow}(S)$, such that $x_D \geq 1/2 - \epsilon_{1/2} - \epsilon_\eta$. For any (top) edge $e \in \mathbf{f} = (S, u)$ such that $e \in D$, if $\mathcal{R}_{\mathbf{f}, S}$, then S is happy, that is $A_T = B_T = 1, C_T = 0$ by [Remark 6.27](#).

Therefore,

$$\begin{aligned} \mathbb{E} [I_S(D)] &\leq \sum_{e \in D: e \in \mathbf{f} = (S, u)} \frac{1 + \epsilon_\eta}{2} \tau x_e \mathbb{P} [S \text{ not happy} | \mathcal{R}_{\mathbf{f}, u}] \mathbb{P} [\mathcal{R}_{\mathbf{f}, u}] \\ &\leq \frac{1 + \epsilon_\eta}{2} p \tau x(D). \end{aligned}$$

Using the trivial inequality [Eq. \(52\)](#) for edges in $\delta^{\rightarrow}(S) \setminus D$ we get

$$\mathbb{E} [I_S^{\rightarrow}] \leq (1 + \epsilon_\eta) p \tau \left(\frac{x(D)}{2} + x(\delta^{\rightarrow}(S)) - x(D) \right) \leq (1 + \epsilon_\eta) p \tau (x(\delta^{\rightarrow}(S)) - (1/4 - \epsilon_{1/2}))$$

as desired. In the last inequality we used $x(D) \geq 1/2 - \epsilon_{1/2} - \epsilon_\eta$.

Case 3: Cases 1 and 2 do not hold. Therefore, by [Theorem 6.35](#) there are least two 2-2-2 good top half edge bundles. In this case, S has chosen a fixed pair of 2-2-2 good edges $\mathbf{e} = (S, v), \mathbf{f} = (S, w)$ in $\delta^{\rightarrow}(S)$ (as defined in [the reduction events](#)) such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$ and $\mathcal{R}_{\mathbf{e}, S} = \mathcal{R}_{\mathbf{f}, S}$ with probability 1. (Recall that $\mathbf{e}(A) = \mathbf{e} \cap A$.) Let $D = \mathbf{e}(A) \cup \mathbf{f}(B)$. In this case, \mathbf{e} and \mathbf{f} are reduced simultaneously by τ when they are 2-2-2 happy (w.r.t., S), i.e., when $\mathcal{R}_{\mathbf{e}, S} = \mathcal{R}_{\mathbf{f}, S} = 1$. In such a case we have $\delta(S)_T = \delta(v)_T = \delta(w)_T = 2$. Therefore,

$$\begin{aligned} \mathbb{E} [I_S(D)] &\leq (1 + \epsilon_\eta) \mathbb{E} [\max\{r(A \cap D), r(B \cap D)\}] \\ &\leq (1 + \epsilon_\eta) \frac{\tau}{2} \max\{x_{\mathbf{e}(A)}, x_{\mathbf{f}(B)}\} (\mathbb{P} [\mathcal{R}_{\mathbf{e}, S} \wedge \mathcal{R}_{\mathbf{f}, S}] + \mathbb{P} [\mathcal{R}_{\mathbf{e}, v}] + \mathbb{P} [\mathcal{R}_{\mathbf{f}, w}]) \\ &\leq (1 + \epsilon_\eta) \tau \frac{3p}{2} x(D) \left(\frac{1}{2} + 3\epsilon_{1/2} \right) = (1 + \epsilon_\eta) \tau p x(D) \left(\frac{3}{4} + 4.5\epsilon_{1/2} \right) \end{aligned}$$

where we used that $1/2 - 2\epsilon_{1/2} - x(C) \leq x_{\mathbf{e}(A)}, x_{\mathbf{f}(B)} \leq 1/2 + \epsilon_{1/2}$ and that $x(C) \leq \epsilon_\eta$. Using the trivial inequality [Eq. \(52\)](#) for edges in $\delta^{\rightarrow}(S) \setminus D$ we get

$$\begin{aligned} \mathbb{E} [I_S^{\rightarrow}] &\leq (1 + \epsilon_\eta) p \tau (x(D)(3/4 + 4.5\epsilon_{1/2}) + x(\delta^{\rightarrow}(S)) - x(D)) \\ &\leq (1 + \epsilon_\eta) p \tau (x(\delta^{\rightarrow}(S)) - (1/4 - 6\epsilon_{1/2})) \end{aligned}$$

where we used $x(D) \geq 1 - 4\epsilon_{1/2} - \epsilon_\eta$. □

8.2.2 Case 2: S and its parent \hat{S} are both polygon cuts

In this subsection we prove two lemmas: [Lemma 8.9](#), which bounds $\mathbb{E}[I_S^\rightarrow]$ when S is the leftmost or rightmost atom of \hat{S} , and [Lemma 8.11](#), which bounds this quantity when S is not leftmost or rightmost.

Lemma 8.9. *Let $S \in \mathcal{H}$ be a polygon cut with $p(S) = \hat{S}$ also a polygon cut. Let u_1, \dots, u_k be the ordering of cuts in $\mathcal{A}(\hat{S})$ (as defined in [Definition 5.31](#)). If $\epsilon_M \leq 0.001$, $\epsilon_\eta \leq \epsilon_M^2$, $S = u_1$ or $S = u_k$, then*

$$\mathbb{E}[I_S^\rightarrow] \leq 0.31\beta p.$$

Proof. Let S be the leftmost atom of \hat{S} and let A, B, C be the polygon partition of $\delta(S)$. First, note

$$\mathbb{E}[I_S^\rightarrow] \leq (1 + \epsilon_\eta) (\mathbb{E}[\max(r(A^\rightarrow), r(B^\rightarrow)) \cdot \mathbb{I}\{S \text{ not happy}\}] + \mathbb{E}[r(C^\rightarrow) \mathbb{I}\{S \text{ not happy}\}]). \quad (54)$$

where recall that $A^\rightarrow = A \cap \delta^\rightarrow(S)$. WLOG assume $x(A^\rightarrow) \geq x(B^\rightarrow)$. Then,

$$\mathbb{E}[\max\{r(A^\rightarrow), r(B^\rightarrow)\} \mathbb{I}\{S \text{ not happy}\}] = \beta p x(A^\rightarrow) \cdot \mathbb{P}[S \text{ not happy} | \mathcal{R}_{\hat{S}}]$$

By [Lemma 8.10](#) we have

$$\begin{aligned} x(A^\rightarrow) \cdot \mathbb{P}[S \text{ not happy} | \mathcal{R}_{\hat{S}}] &\leq x(A^\rightarrow) (1 - ((1 - x(A^\rightarrow))^2 + (x(A^\rightarrow))^2 - 2\epsilon_M - 17\epsilon_\eta)) \\ &\leq (2x(A^\rightarrow)^2 - 2x(A^\rightarrow)^3 + 2\epsilon_M x(A^\rightarrow) + 17\epsilon_\eta x(A^\rightarrow)) \\ &\leq (8/27 + 2\epsilon_M + 17\epsilon_\eta), \end{aligned}$$

where in the final inequality we used that the function $x \mapsto x^2(1 - x)$ is maximized at $x = 2/3$, and using $\epsilon_M \leq 0.001, \epsilon_\eta < \epsilon_M^2$.

Plugging this back into (54), and using $x(C) \leq \epsilon_\eta$, we get

$$\mathbb{E}[I_S^\rightarrow] \leq (1 + \epsilon_\eta) \beta p \left(\frac{8}{27} + 2\epsilon_M + 18\epsilon_\eta \right) \leq 0.31\beta p,$$

where the last inequality follows since $\epsilon_M \leq 0.001$ and $\epsilon_\eta < \epsilon_M^2$. \square

Lemma 8.10. *Let $S \in \mathcal{H}$ be a polygon cut with $p(S) = \hat{S}$ also a polygon cut. Let u_1, \dots, u_k be the ordering of cuts in $\mathcal{A}(\hat{S})$. If $S = u_1$, (or $S = u_k$) then*

$$\mathbb{P}[S \text{ happy} | \mathcal{R}_{\hat{S}}] \geq (1 - x(A^\rightarrow))^2 + (x(A^\rightarrow))^2 - 2\epsilon_M - 17\epsilon_\eta.$$

Proof. Let $A, B, C, \hat{A}, \hat{B}, \hat{C}$ be the polygon partition of S, \hat{S} respectively. Observe that since $S = u_1$, we have $\hat{A} = E(u_1, \hat{S}) = A^\uparrow \cup B^\uparrow \cup C^\uparrow$ and $\hat{B}, \hat{C} \cap (A \cup B \cup C) = \emptyset$. Conditioned on $\mathcal{R}_{\hat{S}}$, \hat{S} is a tree, and marginals of all edges in \hat{A} is changed by a total variation distance at most $\epsilon'_M := \epsilon_M + 2\epsilon_\eta$ from x (see [Corollary 6.16](#)) and they are independent of edges inside \hat{S} . The tree conditioning increases marginals inside by at most $\epsilon_\eta/2$. Since after the changes just described

$$\mathbb{E}[C_T] \leq x_C + \epsilon_\eta + \epsilon'_M \leq 4\epsilon_\eta + \epsilon_M,$$

it follows that $\mathbb{P}[C_T = 0 | \mathcal{R}_{\hat{S}}] \geq 1 - 4\epsilon_\eta - \epsilon_M$. So,

$$\mathbb{P}[S \text{ happy} | \mathcal{R}_{\hat{S}}] \geq (1 - 4\epsilon_\eta - \epsilon_M) \mathbb{P}[A_T = B_T = 1 | C_T = 0, \mathcal{R}_{\hat{S}}]. \quad (55)$$

Let ν be the conditional measure $C_T = 0$, $\mathcal{R}_{\hat{s}}$. We see that

$$\mathbb{P}_\nu [A_T = B_T = 1] = \mathbb{P}_\nu [A_T^\uparrow = 1, B_T^\uparrow = 0, A_T^\rightarrow = 0, B_T^\rightarrow = 1] + \mathbb{P}_\nu [A_T^\uparrow = 0, B_T^\uparrow = 1, A_T^\rightarrow = 1, B_T^\rightarrow = 0]$$

so using independence of $(\delta^\uparrow(S))_T$ and $(\delta^\rightarrow(S))_T$.

$$\begin{aligned} &= \mathbb{P}_\nu [A_T^\uparrow = 1, B_T^\uparrow = 0] \mathbb{P}_\nu [A_T^\rightarrow = 0, B_T^\rightarrow = 1] + \mathbb{P}_\nu [A_T^\uparrow = 0, B_T^\uparrow = 1] \mathbb{P}_\nu [A_T^\rightarrow = 1, B_T^\rightarrow = 0] \\ &\geq (x(A^\uparrow) - \epsilon'_M) \mathbb{P}_\nu [A_T^\rightarrow = 0, B_T^\rightarrow = 1] + (x(B^\uparrow) - \epsilon'_M) \mathbb{P}_\nu [A_T^\rightarrow = 1, B_T^\rightarrow = 0]. \end{aligned}$$

In the final inequality, we used the fact that conditioned on $\mathcal{R}_{\hat{s}}$, $\hat{A} = (A^\uparrow \cup B^\uparrow \cup C^\uparrow)_T = 1$ and marginals in A^\uparrow and B^\uparrow are approximately preserved. Now, we lower bound $\mathbb{P}_\nu [A_T^\rightarrow = 1, B_T^\rightarrow = 0]$.

Let ϵ_A, ϵ_B be such that

$$\mathbb{E}_\nu [A_T^\rightarrow] = \mathbb{P}_\nu [A_T^\rightarrow = 1, B_T^\rightarrow = 0] + \epsilon_A, \quad \mathbb{E}_\nu [B_T^\rightarrow] = \mathbb{P}_\nu [A_T^\rightarrow = 0, B_T^\rightarrow = 1] + \epsilon_B$$

First notice that $\mathbb{P}_\nu [A_T^\rightarrow + B_T^\rightarrow \geq 1] = 1$, and so $\mathbb{P}_\nu [A_T^\rightarrow + B_T^\rightarrow \geq 2] \leq \mathbb{E}_\nu [A_T^\rightarrow + B_T^\rightarrow] - 1$. So,

$$\begin{aligned} \epsilon_A + \epsilon_B &= \mathbb{E}_\nu [A_T^\rightarrow + B_T^\rightarrow] - \mathbb{P}_\nu [A_T^\rightarrow + B_T^\rightarrow = 1] = \mathbb{E}_\nu [A_T^\rightarrow + B_T^\rightarrow] - (1 - \mathbb{P}_\nu [A_T^\rightarrow + B_T^\rightarrow \geq 2]) \\ &\leq 2(\mathbb{E}_\nu [A_T^\rightarrow + B_T^\rightarrow] - 1) \leq 5\epsilon_\eta. \end{aligned}$$

To see the last inequality, first, by [Definition 5.31](#), $x(\delta^\uparrow(S)) \geq 1 - \epsilon_\eta$. Since $x(\delta(S)) \leq 2 + \epsilon_\eta$, we get that $x(\delta^\rightarrow(S)) \leq 1 + 2\epsilon_\eta$. Therefore,

$$\mathbb{E}_\nu [A_T^\rightarrow + B_T^\rightarrow] \leq \mathbb{E} [\delta^\rightarrow(S) \mid \mathcal{R}_{\hat{s}}] \leq x(\delta^\rightarrow(S)) + \epsilon_\eta/2 \leq 1 + 2.5\epsilon_\eta.$$

Therefore,

$$\begin{aligned} \mathbb{P}_\nu [A_T = B_T = 1] &\geq (x(A^\uparrow) - \epsilon'_M)(\mathbb{E}_\nu [B_T^\rightarrow] - \epsilon_B) + (x(B^\uparrow) - \epsilon'_M)(\mathbb{E}_\nu [A_T^\rightarrow] - \epsilon_A) \\ &\geq (x(A^\uparrow) - \epsilon'_M)(x(B^\rightarrow) - 5\epsilon_\eta) + (x(B^\uparrow) - \epsilon'_M)(x(A^\rightarrow) - 5\epsilon_\eta) \end{aligned}$$

where the second inequality uses that the tree conditioning and $C_T^\rightarrow = 0$ can only increase the marginals of edges in A^\rightarrow and B^\rightarrow . Simplify the above using $x(A^\uparrow) + x(A^\rightarrow) \geq 1 - \epsilon_\eta$, and similarly for B ,

$$\begin{aligned} &\mathbb{P}_\nu [A_T = B_T = 1] \\ &\geq (1 - x(A^\rightarrow) - \epsilon_\eta - \epsilon'_M)(x(B^\rightarrow) - 5\epsilon_\eta) + (1 - x(B^\rightarrow) - \epsilon_\eta - \epsilon'_M)(x(A^\rightarrow) - 5\epsilon_\eta) \end{aligned}$$

and since $x(A^\rightarrow) + x(B^\rightarrow) \geq 1 - 2\epsilon_\eta$ (because $x(A^\uparrow) + x(B^\uparrow) \leq 1 + \epsilon_\eta$ and $x_C \leq \epsilon_\eta$), this is

$$\begin{aligned} &\geq (1 - x(A^\rightarrow) - \epsilon_\eta - \epsilon'_M)(1 - x(A^\rightarrow) - 7\epsilon_\eta) + (x(A^\rightarrow) - 3\epsilon_\eta - \epsilon'_M)(x(A^\rightarrow) - 5\epsilon_\eta) \\ &\geq (1 - x(A^\rightarrow))^2 + (x(A^\rightarrow))^2 - \epsilon'_M - 8\epsilon_\eta. \end{aligned}$$

Plugging this into [Eq. \(55\)](#), we obtain

$$\begin{aligned} \mathbb{P} [A_T = B_T = 1, C_T = 0 \mid \mathcal{R}_{\hat{s}}] &\geq (1 - 2\epsilon_\eta - \epsilon'_M) \mathbb{P} [A_T = B_T = 1 \mid C_T = 0, \mathcal{R}_{\hat{s}}] \\ &\geq (1 - 2\epsilon_\eta - \epsilon'_M)((1 - x(A^\rightarrow))^2 + (x(A^\rightarrow))^2 - \epsilon'_M - 8\epsilon_\eta) \\ &\geq (1 - x(A^\rightarrow))^2 + (x(A^\rightarrow))^2 - 2\epsilon'_M - 10\epsilon_\eta, \end{aligned}$$

which noting $\epsilon'_M = \epsilon_M + 2\epsilon_\eta$ completes the proof of the lemma. \square

Lemma 8.11. *Let $S \in \mathcal{H}$ be a polygon cut with $\rho(S) = \hat{S}$ also a polygon cut with u_1, \dots, u_k be the ordering of cuts in $\mathcal{A}(\hat{S})$. If $S \neq u_1, u_k$, then*

$$\mathbb{E} [I_S^\rightarrow] \leq 0.85\beta p.$$

Proof. Let $S = u_i$ for some $2 \leq i \leq k-1$. Let A, B, C be the polygon partitioning of $\delta(u_i)$ and $\hat{A}, \hat{B}, \hat{C}$ be the polygon partitioning of \hat{S} . Since u_i is in the [hierarchy](#) $A^\uparrow \cup B^\uparrow \cup C^\uparrow \subseteq \hat{C}$. So, conditioned on $\mathcal{R}_{\hat{S}}$, $A_T^\uparrow = B_T^\uparrow = C_T^\uparrow = 0$.

Once again, let ν be the conditional measure $C_T = 0$, $\mathcal{R}_{\hat{S}}$. Similar to the previous case, we will lower-bound

$$\begin{aligned} \mathbb{P} [S \text{ happy} | \mathcal{R}_{\hat{S}}] &\geq (1 - 2\epsilon_\eta) \mathbb{P} [A_T^\rightarrow = 1, B_T^\rightarrow = 1, | C_T = 0, \mathcal{R}_{\hat{S}}] \\ &= (1 - 2\epsilon_\eta) \mathbb{P}_\nu [A_T^\rightarrow = 1 | A_T^\rightarrow + B_T^\rightarrow = 2] \mathbb{P}_\nu [A_T^\rightarrow + B_T^\rightarrow = 2] \end{aligned} \quad (56)$$

where we used $\mathbb{E} [C_T^\rightarrow | \mathcal{R}_{\hat{S}}] \leq 2\epsilon_\eta$ in the first inequality. So, it remains to lower-bound each of the two terms in the RHS.

We start with the first one. Since $x(A) \in [1 - \epsilon_\eta, 1 + \epsilon_\eta]$ and $x(A^\uparrow) \leq \epsilon_\eta$ we have

$$\mathbb{E}_\nu [A_T^\rightarrow] \in [1 - 2\epsilon_\eta, 1 + 3\epsilon_\eta].$$

The same bounds hold for $\mathbb{E}_\nu [x(B^\rightarrow)]$.

Therefore,

$$\begin{aligned} \mathbb{P}_\nu [A_T^\rightarrow \geq 1], \mathbb{P}_\nu [B_T^\rightarrow \geq 1] &\geq 1 - e^{-1+2\epsilon_\eta} && \text{(Lemma 6.5)} \\ \mathbb{P}_\nu [A_T^\rightarrow \leq 1], \mathbb{P}_\nu [B_T^\rightarrow \leq 1] &\geq 0.495 && \text{(Markov)} \end{aligned}$$

Therefore, by [Corollary 6.12](#) (with $\epsilon = 0.495(1 - e^{-1+2\epsilon_\eta}) \geq 0.31$) we have

$$\mathbb{P}_\nu [A_T^\rightarrow = 1 | A_T^\rightarrow + B_T^\rightarrow = 2] \geq 0.155.$$

By [Corollary 2.29](#), $\mathbb{P}_\nu [E(u_{i-1}, u_i)_T = 1] \geq 1 - 4\epsilon_\eta$. Similarly, $\mathbb{P}_\nu [E(u_i, u_{i+1})_T = 1] \geq 1 - 4\epsilon_\eta$. And,

$$\mathbb{P}_\nu [\delta^\rightarrow(u_i)_T - E(u_{i-1}, u_i)_T - E(u_i, u_{i+1})_T = 0] \geq 1 - 4\epsilon_\eta$$

So, by a union bound all of these events happen simultaneously and we get $\mathbb{P}_\nu [\delta^\rightarrow(u_i)_T = 2] \geq 1 - 12\epsilon_\eta$. Therefore,

$$\mathbb{P}_\nu [(A^\rightarrow)_T = (B^\rightarrow)_T = 1] \geq 0.155(1 - 12\epsilon_\eta) \geq 0.153.$$

Plugging this back into (56), we get

$$\mathbb{P} [S \text{ happy} | \mathcal{R}_{\hat{S}}] \geq 0.153(1 - 2\epsilon_\eta) \geq 0.152.$$

Plugging this in (54) we get

$$\begin{aligned} \mathbb{E} [I_S^\rightarrow] &\leq (1 + \epsilon_\eta)\beta p \mathbb{P} [S \text{ not happy} | \mathcal{R}_{\hat{S}}] (\max\{x(A^\rightarrow), x(B^\rightarrow)\} + x(C^\rightarrow)) \\ &\leq (1 + \epsilon_\eta)\beta p (1 - 0.152)(1 + \epsilon_\eta + \epsilon_\eta) \leq 0.85\beta p \end{aligned}$$

as desired. □

9 Introduction to the Integrality Gap Result and a Proof Overview

While the above sections describe how to obtain a $\frac{3}{2} - \epsilon$ approximation for TSP, the result is not with respect to the subtour LP. Thus, we have not yet established that the integrality gap of P_{Sub} is less than $\frac{3}{2}$ the purpose of the next few sections is to show that this is true.

9.1 New techniques and contributions

This result can be seen as a case study on how to reason about and deal with *near* minimum cuts. One can deduce from the classical cactus representation of a graph G [DKL76] (i) the structure of *all* min cuts of G and (ii) the structure of the edges of G in the sense that every edge $\{u, v\}$ maps to a unique *path* in the cactus between the images of u and v . Furthermore, such a path intersects every cycle of the cactus on at most one cactus edge. The theory has found many application from designing fast algorithms [Kar00; KP09] to the analysis of approximation algorithms for TSP [KKO20] and connectivity augmentation [BGA20a; CTZ21].

Two decades later, the theory of min cuts was extended to near min cuts in works of Benczúr and Goemans [Ben95; BG08] where they introduced the polygon representation which represents all cuts of a graph with at most $\frac{6}{5}k$ edges, where k is its edge connectivity. Although these works completely classify the structure of all near min cuts of a given graph G , they do not characterize the structure of the *edges* of G with respect to these cuts, which can be important in applications (for example, in many of the recent applications of min cuts, one also needs to exploit the structure of the edges in relation to the cactus). The structure on the edges turns out to be highly relevant in this work as well, and as a byproduct of our analysis we make progress towards classifying the way in which the edges of G relate to the structure of the polygon representation.

For motivation, consider a generic family of network design problems in which we want to construct a network such that every pair u, v of vertices has connectivity at least $c_{u,v}$. A natural approach is to write an LP relaxation to find a (minimum cost) vector $x : E \rightarrow \mathbb{R}_{\geq 0}$ such that for every cut S separating u and v , $x(\delta(S)) \geq c_{u,v}$. We can round this LP using independent rounding or a dependent rounding scheme such as sampling from max entropy distributions. Using classical concentration bounds one can show that if $x(\delta(S)) \gg c_{u,v}$ then with high probability the rounded solution has at least $c_{u,v}$ edges across this cut. So the main challenge is to “fix” near tight cuts, i.e., cuts where $x(\delta(S)) \approx c_{u,v}$. For an explicit instantiation of this scheme see [Kar+21]. A better understanding of the global structure of the family of near tight cuts has the potential to significantly simplify or even improve the approximation factor of such rounding algorithms. A classical technique to design algorithms for such network design problems is to apply uncrossing to extreme point solutions of the LP. One can view our contribution as an approximate uncrossing technique that deals with all near tight cuts (instead of just tight cuts) as we explain next.

An Approximate Uncrossing Technique. A fundamental technique in the field of approximation algorithms is the uncrossing technique²⁷ of Jain [Jai01]. Given a graph $G = (V, E)$, a weight vector $x : E \rightarrow \mathbb{R}_{\geq 0}$, and a function $f : V \rightarrow \mathbb{R}$, suppose that $x(\delta(S)) \geq f(S)$ for all $S \subseteq V$. Let \mathcal{N} be the family of sets S such that $x(\delta(S)) = f(S)$, i.e., the family of *tight* sets with respect to f . The uncrossing technique says that if f is (weakly) supermodular then we can refine \mathcal{N} to a laminar family of sets, \mathcal{H} , such that if all sets of \mathcal{H} are tight, then all sets of \mathcal{N} are tight as well. For a

²⁷See e.g. [LRS11] for a number of applications of this technique.

concrete example, suppose f is a constant function, say $f(S) = 2$ for all $\emptyset \subsetneq S \subsetneq V$. Then, sets of \mathcal{H} can be constructed using the cactus representation [DKL76] of cuts in \mathcal{N} . The significance of this method is that if x is a basic feasible solution to a LP with constraints $x(\delta(S)) \geq f(S)$ for all S , one can use this machinery to argue that the support of x has size $O(|V|)$.

Informally, we prove the following, which can be seen as an *approximate uncrossing technique*:

Theorem 9.1 (Informal). *Suppose we have a vector $x : E \rightarrow \mathbb{R}_{\geq 0}$ such that $x(\delta(S)) \geq f(S)$ for all S ; define \mathcal{N} to be sets S where $x(\delta(S)) \leq f(S)(1 + \epsilon)$ for some fixed $\epsilon > 0$. If $f(\cdot)$ is constant, say $f(S) = 2$ for all S , then there is a set $\mathcal{N}^* \subseteq \mathcal{N}$ and a collection of edge sets $F_1, \dots, F_m \subseteq E$ such that the following hold:*

- $|\mathcal{N}^*| = O(|V|)$, $m = O(|V|)$.
- $x(F_i) \geq 1 - \epsilon/2$ for all $1 \leq i \leq m$.
- Every edge e is in at most $O(1)$ of the F_i 's.
- For every set $S \in \mathcal{N} \setminus \mathcal{N}^*$ there exists $1 \leq i < j \leq m$ such that $F_i \cap F_j = \emptyset$ and $F_i \cup F_j \subseteq \delta(S)$ and for every $S \in \mathcal{N}^*$, there exists $1 \leq i \leq m$ such that $F_i \subseteq \delta(S)$.

In words, although we cannot simply refine \mathcal{N} to a linear number of sets, we can refine the edges in cuts of \mathcal{N} to a linear number of sets F_1, \dots, F_m such that we can essentially capture the edges of $\delta(S)$ for any $S \in \mathcal{N} \setminus \mathcal{N}^*$ by a pair of disjoint F_i 's. We give a slightly weaker condition for cuts in \mathcal{N}^* ; namely we only capture half of their edges by F_i 's.

Example 9.2. For a simple example of the above theorem, suppose $\epsilon = 0$, i.e. \mathcal{N} is the set of min cuts of a graph G . Furthermore, suppose that every proper cut in \mathcal{N} is **crossed** (recall that S is proper if $1 < |S| < |V| - 1$) and that \mathcal{N} has at least one proper cut. Then, one can use an uncrossing technique, namely that if $A, B \in \mathcal{N}$ then $A \cap B \in \mathcal{N}$, to prove that G must be cycle, namely we can order vertices of G , v_0, \dots, v_{n-1} such that $x_{\{v_i, v_{i+1 \bmod n}\}} = 1$. In such a case we let $\mathcal{N}^* = \emptyset$ and $F_i = E(v_i, v_{i+1 \bmod n})$.

Example 9.3. For a second example, suppose again $\epsilon = 0$ and \mathcal{N} is the set of mincuts of a graph G where \mathcal{N} forms a laminar family (no two cuts cross). It turns out that we cannot decompose edges of cuts of \mathcal{N} into a linear sized collection of sets where every edge appears only a constant number of times. The main reason is that some edges may appear in an unbounded number of cuts. In this case we let $\mathcal{N}^* = \mathcal{N}$ and for every $A \in \mathcal{N}$ (with immediate parent $B \in \mathcal{N}$ in the laminar family) we add a set $F_A = \delta(A) \setminus \delta(B)$ to our collection. It is straightforward to show, using the structure of min cuts, that $x(F_A) \geq 1$; furthermore, since the size of a laminar family is linear in V , this gives a valid decomposition in the sense of above theorem.

Lastly, if $\epsilon = 0$ and \mathcal{N} is the set of min cuts of an arbitrary graph, one can represent all min cuts of \mathcal{N} by a cactus [DKL76] which can be seen as a tree of cycles. In such a case, one can use a construction similar to [Example 9.2](#) for each cycle where instead of a vertex v_i we have a set $a_i \subseteq V$ and one similar to [Example 9.3](#) for the tree part of the cactus. For a concrete application of such a decomposition of min cuts see [KKO20].

One of the main challenges in dealing with near min cuts relative to min cuts is that if $x(\delta(A)), x(\delta(B)) \leq 2 + \epsilon$ then $x(\delta(A \cap B)) \leq 2 + 2\epsilon$. Therefore, if $\epsilon = 0$, then min cuts are closed under intersection, set difference and union, but this is no longer true when $\epsilon > 0$. So, to employ the classical uncrossing machinery one should be very careful to "uncross" only a constant number

of times (independent of ϵ) to make sure that every cut remains within $2 + O(\epsilon)$. This is the main reason that the polygon representation of near min cuts (see below) is more sophisticated, e.g., we can no longer argue $x(E(a_i, a_{i+1})) \approx 1$, see Fig. 42.

Although we don't study it here, we believe it may be worthwhile to find generalizations of Theorem 9.1 which hold for any (weakly) supermodular function.

Remark 9.4. We do not explicitly prove Theorem 9.1 here, as it is not used to prove Theorem 1.2. However it can be deduced from arguments in Section 11 and Appendix C.

Extensions to the Polygon Representation To obtain our uncrossing framework we prove new properties of the polygon representation. Given a graph $G = (V, E)$, let k be the edge-connectivity of G , i.e. the number of edges in a minimum cut of G . For $\epsilon > 0$, consider the set of $(1 + \epsilon)$ -near minimum cuts of G : cuts (S, \bar{S}) where $|E(S, \bar{S})| < (1 + \epsilon)k$. Benczúr [Ben95] and Benczúr, Goemans [BG08] proved that if $\epsilon \leq 1/5$ then the near minimum cuts of G admit a *polygon representation*. Namely, every connected component \mathcal{C} of crossing $(1 + \epsilon)$ near min cuts can be represented by the diagonals of a convex polygon. In this polygon, the vertices of G are partitioned into sets called *atoms*, and every atom is mapped to a cell of this polygon defined by the diagonals and the boundary of the polygon itself (see ?? for more details).

The polygon representation can be seen as a generalization of the well-known cactus representation [DKL76] of minimum cuts where a cycle of the cactus is replaced by a convex polygon. Unlike a cycle, some vertices/atoms map to the interior of the polygon, which are called "inside" atoms. The inside atoms at first look like a mystery and one can ask many questions about them such as how many can exist and what structures they can exhibit.

Here, we explain two lemmas we proved which might find further applications beyond TSP in the future. First, we give a necessary condition for a cell of a polygon to contain an inside atom:

Lemma 9.5 (Informal, see Lemma 10.27). *Consider a polygon P for a connected component \mathcal{C} of a family of $1 + \epsilon$ near min cuts for $\epsilon \leq 1/5$ (where representing diagonals correspond to cuts in \mathcal{C}). Any cell of P that has an inside atom must have at least $\Omega(1/\epsilon)$ many sides.*

This can be seen as a generalization of [BG08, Lem 22] to the case in which the cell is allowed to be adjacent to vertices of the polygon P .

Now, we explain our second extension: it follows from the cactus representation of minimum cuts that for a graph G and a min cut S one can partition the set of all min cuts that cross S into two groups $\mathcal{A} = \{A_1, \dots, A_k\}$ and $\mathcal{B} = \{B_1, \dots, B_l\}$ for some $k, l \geq 0$ such that $S \cap A_1 \subseteq S \cap A_2 \subseteq \dots \subseteq S \cap A_k$ and, similarly, $S \cap B_1 \subseteq \dots \subseteq S \cap B_l$. We prove a generalization of this fact for near min cuts:

Lemma 9.6 (Informal, see Lemma 10.26). *Consider the set of $1 + \epsilon$ near min cuts of a graph G for $\epsilon \leq 1/10$; for any such near min cut S , one can partition the $1 + \epsilon$ near min cuts crossing S into two groups $\mathcal{A} = \{A_1, \dots, A_k\}$ and $\mathcal{B} = \{B_1, \dots, B_l\}$ such that $S \cap A_1 \subseteq S \cap A_2 \subseteq \dots \subseteq S \cap A_k$ and similarly for cuts in \mathcal{B} .*

9.2 Proof Overview

Algorithm 1 consists of two steps: sampling a tree whose marginals match x (and hence has expected cost equal to $c(x)$), and then augmenting this with a minimum cost matching on the

odd degree vertices of the tree. The goal of this section (the content of [KKO22]) is to show that the expected cost of the minimum cost matching on the odd degree vertices of the sampled tree is at most $(1/2 - \epsilon)c(x)$ instead of $(1/2 - \epsilon)c(OPT)$. This is done by showing the existence of a cheap feasible O -join solution to (4). Note that as before we merely need to prove the existence of a cheap O -join solution. The actual optimal O -join solution can be found in polynomial time.

Recall that if we only wanted to get an O -join solution of value at most $c(x)/2$, to satisfy all cuts, it is enough to set $y_e := 0.5x_e$ for each edge²⁸ [Wol80]. If all of the near min cuts of x containing e are even, then we can reduce y_e strictly below $0.5x_e$. The difficulty in implementing this approach comes from the fact that a high cost edge can be on many near min cuts and it may be exceedingly unlikely that *all* of these cuts will be even simultaneously. Our overall approach is exactly as before: initialize $y_e := 0.5x_e$ and then modify it by adding to it a random²⁹ slack vector $s : E \rightarrow \mathbb{R}$: For each edge e , when certain special (few) η -near-mincuts that e is on are even in the tree, s_e is set to $-x_e\beta$ where $\beta \approx \eta/4$; for other cuts that contain e , whenever they are odd, the slack of *other* edges on that cut is increased to satisfy them (i.e., maintain feasibility of y for that cut). The bulk of the effort was to show that this can be done while guaranteeing that $\mathbb{E}[s_e] < -\epsilon x_e$ for some $\epsilon > 0$, and therefore $\mathbb{E}[y_e] = 0.5x_e + \mathbb{E}[s_e] < (0.5 - \epsilon)x_e$.

To help the reader understand both the big picture as well as the ideas and contribution of this section, it is useful to describe the approach taken in the previous section at a high level. Let \mathcal{N}_η be the set of all η -near min cuts of x . A key idea there was to partition \mathcal{N}_η into three types: a set of near min cuts \mathcal{H} that form a *hierarchy* (which is a laminar family of cuts), a set of cuts $\mathcal{N}_{\eta,1}$ that are "crossed on one side" and a set of cuts $\mathcal{N}_{\eta,2}$ that are "crossed on both sides." [KKO21] (the section above) showed that if we only need to satisfy the O -Join constraints coming from \mathcal{H} , then we can find such a vector s .

However, this vector s (which is negative in expectation) might "break" O -join constraints on cuts that are *not* in the hierarchy (i.e., cuts in $\mathcal{N}_{\eta,1}$ and $\mathcal{N}_{\eta,2}$). To resolve this, we showed how a negligible *increase* in the slack of certain edges (a slack component they called s^*) can be used to restore the feasibility of the O -join solution on *all* cuts, including those that are not in the hierarchy. See Section 12 for more on this.

Concretely, because the cuts in $\mathcal{N}_{\eta,2}$ have a rather complex structure, to simplify their handling, we changed the plan: Instead of starting with $y_e = 0.5x_e$, they started with $y_e = (x_e + OPT_e)/2$, where OPT_e is an indicator for edge e being in the optimal integral TSP solution. They then constructed slack vectors relative to the near min cuts of $(OPT + x)/2$. The advantage of doing so is that it guarantees that all near mincuts correspond to intervals of vertices along the optimal cycle, *greatly* simplifying the structure of the family of near min cuts under consideration. Slack on the edges in the optimal cycle was then used to handle the cuts in $\mathcal{N}_{\eta,1}$ and $\mathcal{N}_{\eta,2}$.

Unfortunately, this meant that the bound on the expected cost of the minimum cost matching from [KKO21] is at most $(1/2 - \epsilon)((c(x) + c(OPT))/2)$, which is insufficient to prove that the integrality gap of the LP is strictly below $3/2$.

In the present section, we return to the plan of initializing $y_e := 0.5x_e$ and then construct a slack vector for each edge with the desired properties. Our starting point is the *polygon decomposition* \mathcal{D}

²⁸This is because x satisfies $x(\delta(S)) \geq 2$ for all S , whereas y must satisfy $y(\delta(S)) \geq 1$ just for those cuts that have odd intersection with the tree T .

²⁹where the randomness comes from the random sampling of the tree

of the η -near min cuts of x [BG08]³⁰. As stated previously, a polygon³¹ is a connected component of crossing $2 + \eta$ near minimum cuts, where two cuts are connected if they cross each other. It turns out that the way the polygon representation \mathcal{D} is constructed, each cut in $\mathcal{N}_{\eta,2}$ is in exactly one polygon, and each edge on such a cut will have its slack increase in at most one polygon. Thus, cuts in $\mathcal{N}_{\eta,2}$ can be handled independently for each polygon.

The main result of this paper is to show how to handle the cuts in $\mathcal{N}_{\eta,2}$ (polygon by polygon) without resorting to the use of the OPT vector. Specifically, we prove the following:

Theorem 9.7 (Informal main theorem). *For any connected component C of \mathcal{N}_η (i.e. a polygon), let C_2 be the cuts in C that are crossed on both sides. For any $\alpha > 0$, there is a vector $s^* : E \rightarrow \mathbb{R}$ depending on T s.t.*

- (i) $\forall e \in E, s_e^* \geq 0$;
- (ii) $\mathbb{E}[s_e^*] = O(\eta \alpha x_e)$, where the expectation is over the choice of tree T .
- (iii) If $S \in C_2$ is a cut such that $\delta(S)_T$ is odd, then $s_T^*(\delta(S)) := \sum_{e \in \delta(S)} s_e^* \geq \alpha(1 - \eta)$.

Once cuts in $\mathcal{N}_{\eta,2}$ are handled, the remaining cut structure becomes significantly simpler in that the polygons start to look very much like cycles: they contain only "outside atoms"³² and the fractional mass $x(a_i, a_{i+1})$ between adjacent atoms is $1 \pm \Theta(\eta)$, as we have discussed in previous sections. This enables us, with minor modification to the way in which cuts crossed on one side are handled (see [Theorem C.12](#)), to adapt one of the main results on handling cuts crossed on one side.

Theorem 9.8 (Informal Theorem adapted from [Theorem C.12](#) and [Theorem D.2](#)). *Given a family $\mathcal{N}_{\eta, \leq 1}$ of near-min cuts containing **no** cuts crossed on both sides, for any $\beta > 0$, there is a vector $s : E \rightarrow \mathbb{R}$ depending on T such that*

- (i) $\forall e \in E, s_e \geq -\beta x_e$
- (ii) $\mathbb{E}[s_e] < -\epsilon \beta x_e$ for some absolute constant $\epsilon > 0$, independent of η , where the expectations are over the choice of T .
- (iii) If $S \in \mathcal{N}_{\eta, \leq 1}$ is a cut such that $\delta(S)_T$ is odd, then $s_T(\delta(S)) = \sum_{e \in \delta(S)} s_e \geq 0$.

Note that [Theorem D.2](#) (used to prove the above theorem) crucially relies on the fact that the tree is sampled from a max-entropy distribution, whereas [Theorem 9.7](#) does not.

Before we explain the ideas underlying the proof of [Theorem 9.7](#), we quickly show how by setting

$$y_e(T) := 0.5x_e + s_e^* + s_e \quad \forall e,$$

these two theorems together imply the main result of this paper.

First, we show that $\mathbb{E}[c(y)] \leq c(x)(0.5 - \epsilon)$. To see this, observe that [Theorem 9.7\(ii\)](#) together with [Theorem 9.8\(ii\)](#) imply that for every edge $e \in E$,

$$\mathbb{E}[y_e] = 0.5x_e + \mathbb{E}[s_e] + \mathbb{E}[s_e^*] \leq x_e(0.5 + O(\eta\alpha) - \epsilon\beta) \leq x_e(0.5 - \eta\epsilon'),$$

³⁰See ?? for a formal introduction to polygons. In particular, a reader unfamiliar with polygons will likely need to read ?? to understand this section, though we provide a very brief overview now and again in [Section 9.3](#).

³¹One difference between a cycle on m nodes and a polygon with m "outside atoms" is that in a cycle all of the $\binom{m}{2}$ simple cuts are min-cuts, whereas in a polygon only some of the $\binom{m}{2}$ simple cuts are η -near min cuts. Indeed a cycle is the "simplest" kind of polygon. Another major difference is that polygons may also have "inside" atoms. See ??.

³²See ??.

for α, β as chosen below and η sufficiently smaller than ϵ . Summing over all edges, this gives

$$\mathbb{E} [c(y)] \leq \left(\frac{1}{2} - \epsilon' \right) c(x).$$

Note that since s_e^* is always nonnegative, it does not help us in our quest to reduce y_e strictly below $0.5x_e$. That reduction comes only from s_e being negative. Indeed, the raison d'être of the slack vector s^* is to repair the feasibility of cuts which are odd in the tree but which have s_e negative on some edges in $\delta(S)$. This is why it is crucial that $\mathbb{E} [s_e]$ is much smaller than $-\mathbb{E} [s_e^*]$.

Next, we show that $y(T)$ is feasible for every tree T . For this, we need to consider three types of cuts:

Case 1: $\delta(S)_T$ is odd and $x(\delta(S)) > 2 + \eta$. Since $s_T^*(\delta(S)) \geq 0$ and $s_T(\delta(S)) \geq -\beta x(\delta(S))$, we have

$$y_T(\delta(S)) = 0.5x(\delta(S)) + s_T^*(\delta(S)) + s_T(\delta(S)) \geq (0.5 - \beta)x(\delta(S)) \geq (0.5 - \beta)(2 + \eta) \geq 1,$$

for $\beta \approx \eta/4$.

Case 2: $\delta(S)_T$ is odd, $S \in \mathcal{N}_{\eta, \leq 1}$. In this case $s_T(\delta(S)), s_T^*(\delta(S)) \geq 0$ so $y_T(\delta(S)) \geq 0.5x(\delta(S)) \geq 1$.

Case 3: $\delta(S)_T$ is odd, $x(\delta(S)) \leq 2 + \eta$, and $S \in \mathcal{N}_{\eta, 2}$. In this case, $s_T^*(\delta(S)) \geq \alpha(1 - \eta)$ and $s_T(\delta(S)) \geq -\beta x(\delta(S))$, so we have

$$y_T(\delta(S)) = 0.5x(\delta(S)) + s_T^*(\delta(S)) + s_T(\delta(S)) \geq (0.5 - \beta)x(\delta(S)) + \alpha(1 - \eta) \geq 1,$$

for $\alpha \approx 2\beta$ using $x(\delta(S)) \geq 2$.

9.3 Overview of proof of [Theorem 9.7](#) – no inside atoms

Given a connected component \mathcal{C} of cuts in \mathcal{N}_η , we can partition vertices of G into sets a_0, \dots, a_{m-1} (called atoms); this is the coarsest partition such that for each a_i , and each $(S, \bar{S}) \in \mathcal{C}$, we have $a_i \subseteq S$ or $a_i \subseteq \bar{S}$. One of these atoms, a_0 is the atom that contains u_0, v_0 . We call a_0 the *root*. In the following, we will often identify an atom with the set of vertices that it represents³³.

If $\eta = 0$, then [\[DKL76\]](#) shows that the structure of cuts in \mathcal{C} can be represented by a cycle; namely we can arrange these atoms around a cycle such that, perhaps after renaming, for any $0 \leq i \leq m - 1$, $x(E(a_i, a_{i+1 \bmod m})) = 1$ and cuts of \mathcal{C} are just the mincuts of this cycle.

As mentioned, [\[Ben95; BG08\]](#) studied the case when $0 < \eta \leq 2/5$ and introduced the notion of *polygon representation*, in which case atoms can be placed on the sides of an equilateral polygon P and some atoms placed inside the polygon, such that every cut in \mathcal{C} can be represented by a diagonal of this polygon. See [Figure 18](#).

In the rest of this section, we fix \mathcal{C} and we outline the ideas behind the proof of [Theorem 9.7](#) in the special case that the polygon P representing the connected component of cuts \mathcal{C} contains no inside atoms. This latter assumption simplifies the argument but still illustrates many of the main ideas.

³³For example, it will be convenient to write cuts as subsets of atoms. In this case the cut is the union of the vertices in those atoms.

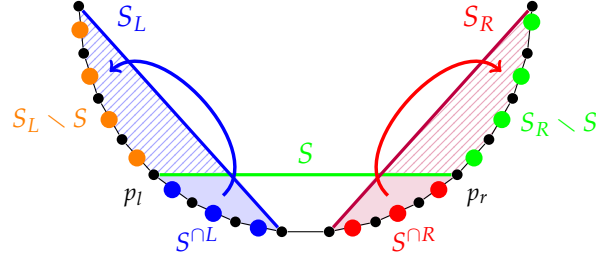


Figure 38: (Note that to simplify the pictures, we usually draw a polygon as a circle.) The figure shows a cut S that is crossed on both sides. The cut S consists of all atoms below the green line. The cut S_R is the cut crossing S on the right which minimizes the number of outside atoms in the intersection, i.e., it minimizes the number of red atoms. Similarly, S_L crosses S on the left and minimizes the number of blue atoms. While not shown in the picture, it's possible for the red and blue atoms to overlap. (S_R might cross S_L .) Edges between red atoms and green atoms are in $E^{\rightarrow}(S)$ and edges between blue atoms and orange atoms are in $E^{\leftarrow}(S)$. Edges in $E^{\circ}(S)$ are all remaining edges in $\delta(S)$. [Claim 9.9](#) shows that with probability $1 - O(\eta)$, in the randomly sampled tree T , there is exactly one (red,green) edge (i.e., $B^{\rightarrow}(S)$ does not occur) and exactly one (blue, orange) edge (i.e., $B^{\leftarrow}(S)$ does not occur) and those are the only edges in $\delta(S) \cap T$ (i.e., also $B^{\circ}(S)$ does not occur).

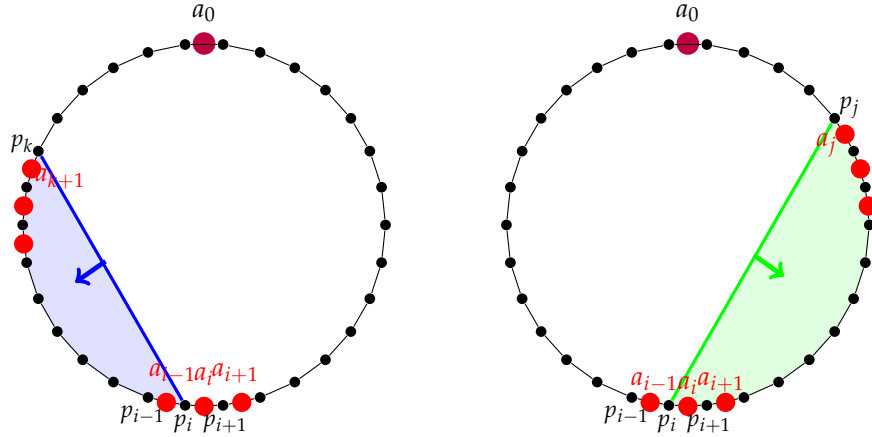


Figure 39: Of all cuts crossed on both sides, $L(p_i)$, the blue set, extends farthest to the left from p_i . Similarly, $R(p_i)$, the green set, is the one that extends farthest to the right from p_i . (For reference when we later include inside atoms: if the root is not an outside atom, $L(p_i)$ can wrap around past a_0 and there may be atoms in the interior of the blue region, aka inside atoms. However, the outside atoms of $L(p_i) \cup R(p_i)$ form a contiguous interval around the cycle and don't include all outside atoms. Even in the case in which the root is not an outside atom, the shaded region is the side of the diagonal which does not contain the root.)

We assume that the atoms of P are labelled counterclockwise from a_0 to a_{m-1} . We associate to each diagonal (defining a cut) the side which does not contain a_0 .³⁴ Thus, we will refer to a cut by the set of outside atoms it contains, say $[a_i, a_j]$, $i < j$. (This denotes the side of the diagonal containing the atoms a_i, a_{i+1}, \dots, a_j .) We equivalently refer to this cut by giving the left and right polygon points of its diagonal $[p_{i-1}, p_j]$.

As mentioned above, the raison d'être for the slack vector s^* that we construct here is to

³⁴The reason we do this is that it is crucial for subsequent arguments to be able to condition on near min cuts being trees using [Lemma 2.28](#), i.e., that for S a near min cut, $E(S) \cap T$ is very likely to be a tree. However, this lemma can only be used on sets which do not contain u_0, v_0 .

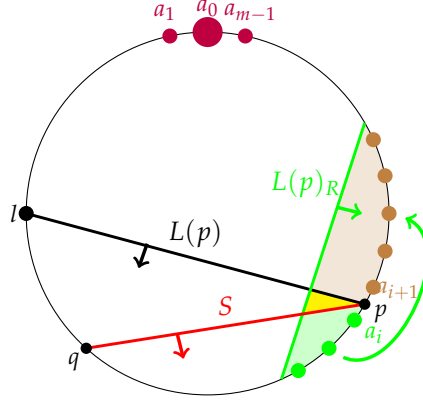


Figure 40: Since S is crossed on the right and any cut that crosses S on the right also crosses $L(p)$ on the right, the cut S_R , which contains the fewest number of atoms in S (green atoms), is the same as $L(p)_R$. The edges in $E^{\rightarrow}(L(p)) = E^{\rightarrow}(S)$ are those that go between green atoms and brown atoms. Note also that any edge in $\delta(S)$ with one endpoint to the right of p that is *not* in $E^{\rightarrow}(L(p))$ is in $E^{\circ}(L(p))$. (It can't be in $E^{\leftarrow}(L(p))$ since those edges have one endpoint to the left of l .) Note also that since the green + yellow region as well as the brown region are each the difference of two crossing η near min cuts, each is a 2η near min cut. So by ??, the fraction of edges with one endpoint in each of these regions is $1 - O(\eta)$. (To extend this to the case where polygon P may have inside atoms we show that there are no atoms in the yellow region see [Lemma 11.1](#), or [Lemma 10.27](#) for a more general statement.)

restore the feasibility of cuts S in \mathcal{C}_2 which are odd in the tree but which have s_e negative on some edges in $\delta(S)$. The high level approach in the proof is the following. Initialize $s_e^* := 0$ for all e . Now define a set of *bad events* whose occurrence signifies that some of these near min cuts are potentially in need of such a repair. These bad events should satisfy the follow desiderata:

- (a) Each bad event occurs with probability $O(\eta)$, where the probability is taken over the choice of tree T .
- (b) The occurrence of a bad event B in a tree T triggers a *slack increase* on an associated set of edges $E(B)$. Specifically, when B occurs, each edge $e \in E(B)$ has its slack s_e^* increased by αx_e .
- (c) Each edge e is in $E(B)$ for a *constant number of bad events* B . Combining (a) and (b), this implies that $\mathbb{E}[s_e^*] = O(\eta \alpha x_e)$ (condition 2 of [Theorem 9.7](#)).
- (d) Each η -near-min cut S is associated with a constant number of bad events $\mathcal{B}(S)$, such that when $\delta(S)_T$ is odd, at least one of the bad events $B \in \mathcal{B}(S)$ occurs. We will ensure that the edges in $E(B)$ (on which slack increases are triggered) are a subset of $\delta(S)$ of fractional value at least $\Omega(1)$. Therefore, if S is odd in the tree, $s_T^*(\delta(S)) \geq \alpha x(E(B)) \geq \Omega(\alpha)$ implying condition (iii) of [Theorem 9.7](#) (once the constant are chosen appropriately).

9.4 Satisfying the above desiderata

Consider any near min cut S in P , which is crossed on the left and on the right (see [Definition 10.14](#) for the definition of being crossed on left/right) . Let S_L and S_R be the cuts crossing S on the left and right *with minimum sized intersection* with S . See [Figure 38](#).

One of the very nice things about cuts crossed on both sides is the following:

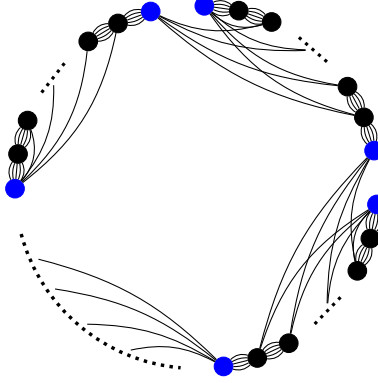


Figure 42: Suppose that there are exactly $1/\eta$ (black) vertices between any two vertices in the above figure, where each edge has fractional value η . Also, any two consecutive vertices (if both of them are not blue) have exactly $1/\eta$ parallel edges between them. Then, it is easy to check that the above graph is fractionally 2 edge connected. Furthermore, the set of $2 + O(\eta)$ -near minimum cuts comprises a single connected component and every vertex will become an (outside) atom of the corresponding polygon. This is because every diagonal which separates two blue vertices on both sides is a near min cut. In addition, every interval with $O(1)$ many consecutive vertices where at most one of vertex is blue is also a near mincut. In such a case, for every pair of adjacent blue vertices, we have $E(a_i, a_{i+1}) = \emptyset$.

9.4.1 Defining bad events for each polygon point

For a fixed polygon point p , let $L := L(p)$ be the set crossed on both sides that extends farthest clockwise from p and as above, let L_R be the cut that crosses it on the right with the minimum number of outside atoms in the intersection. Analogously define $R := R(p)$ and R_L . See Figure 39.

Now we consider two bad events:

$$\begin{aligned} B^\rightarrow(p) &= \mathbb{1}\{E^\rightarrow(L(p))_T \neq 1 \text{ or } E^\circ(L(p))_T \neq 0\} \\ B^\leftarrow(p) &= \mathbb{1}\{E^\leftarrow(R(p))_T \neq 1 \text{ or } E^\circ(R(p))_T \neq 0\}. \end{aligned} \quad (58)$$

For these events, we have the following two claims:

Claim 9.10. *For any near min cut $S = [p, q]$, $E^\rightarrow(L(q)) = E^\rightarrow(S)$ and $E^\leftarrow(R(p)) = E^\leftarrow(S)$. Moreover, $E^\circ(S) \subset E^\circ(L(p)) \cup E^\circ(R(p))$. See Figure 40. Therefore, if neither $B^\rightarrow(q)$ or $B^\leftarrow(p)$ occur, then $\delta(S)_T$ is even.*

In addition we have

Claim 9.11. *For any polygon point p , $\mathbb{P}[B^\rightarrow(p)], \mathbb{P}[B^\leftarrow(p)] = O(\eta)$.*

This follows arguments similar to those used in Claim 9.9, using that $x(E^\rightarrow(L(p))) \geq 1 - \eta/2$, and $x(E^\circ(L(p))) = O(\eta)$ (and similarly for $R(p)$).

These bad events satisfy the desiderata (a) and (d) (assuming we define $E(B)$ such that $x(E(B)) \in \Omega(1)$).

9.4.2 Defining the slack increase sets for bad events

It remains to determine the sets $E(B^\rightarrow(p)), E(B^\leftarrow(p))$ for which slack increases are triggered when the bad events occur. In particular, we will let $E(B^\rightarrow(p)) \subseteq E^\rightarrow(L(p))$ and $E(B^\leftarrow(p)) \subseteq E^\leftarrow(R(p))$ such that:

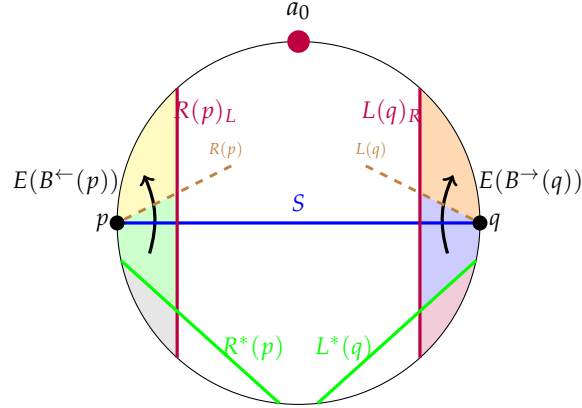


Figure 43: This figure illustrates the definitions in (59). The near min cut S contains all atoms below the blue line. $L(q)_R$ crosses $L(q)$ on the right and $R(p)_L$ crosses $R(p)$ on the left. $L(q)_R$ is the set of atoms in the blue + pink region on the right. $L^*(q)$ is the near min cut crossing $L(q)_R$ on the left that maximizes the number of outside atoms in the pink region. Similarly $R^*(p)$ is the near min cut crossing $R(p)_L$ on the right that maximizes the number of outside atoms in the grey region. The edges in $E(B^{\rightarrow}(q))$ are those edges from the blue region to the orange region and the edges in $E(B^{\leftarrow}(p))$ are those edges from the green region to the yellow region. Note that the figure is misleading in that sets that are shown as disjoint here may not in fact be disjoint.

(*) $x(E(B^{\rightarrow}(p))) \geq \Omega(1)$ and $x(E(B^{\leftarrow}(p))) \geq \Omega(1)$ (to guarantee (d)),

(**) All edges e are in at most a constant number of sets $E(B)$ (to guarantee (c)).

Assuming we can satisfy (*) and (**), we can set $s_e^* = \alpha x_e$ for all $e \in E(B)$ when B occurs to satisfy all four desiderata.

First try: The most natural choice is to simply let $E(B^{\rightarrow}(p)) = E^{\rightarrow}(L(p))$. Here, (*) obviously holds but unfortunately (**) fails. Indeed, there are examples (see Figure 41) for which there exist edges $e \in E(a_i, a_j)$ with $|j - i| = \Omega(m)$ that belong to $E^{\rightarrow}(L(p_k))$ for $\Omega(m)$ many values of $i \leq k \leq j$.

Second try: Let a_i be the atom immediately to the left of p and a_{i+1} the atom immediately to the right of p (i.e. $p = p_i$). Note that all edges with one endpoint in a_i and one in a_{i+1} are in $E^{\rightarrow}(L(p))$. Now if it was always the case that $x(E(a_i, a_{i+1})) \geq \gamma$ for some universal constant $\gamma > 0$, then, when one of these bad events occurs, say $B^{\rightarrow}(p)$, we could simply increase the slack of every edge e in $E(a_i, a_{i+1})$ by αx_e . This approach is analogous to the method employed in [KKO21] where slack was increased on OPT edges. One might have some hope that this is true since it holds with $\gamma = 1$ for the cactus representation of min cuts (i.e. when $\eta = 0$).

Unfortunately, as observed in [OSS11] there is a family of near minimum cuts such that the polygon representation has no inside atoms, yet $E(a_i, a_{i+1}) = \emptyset$ for some consecutive pairs of (outside) atoms (see Figure 42) (even though there are cuts whose diagonals end between those atoms). So, this method is doomed even if inside atoms are not present.

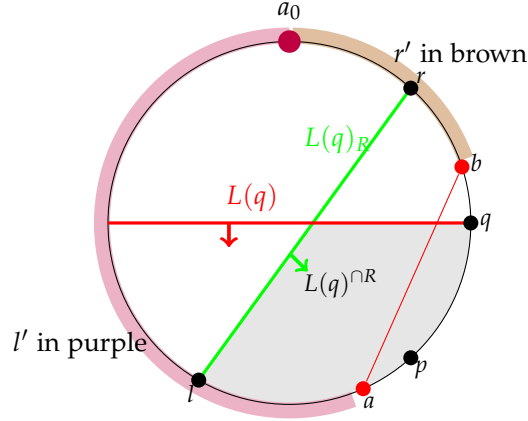


Figure 44: Setup for proof of [Claim 9.13](#): Let $L(q)_R = (l, r)$ and $L(p)_R = (l', r')$ (where l, r, l', r' are polygon points). The grey region is $L(q)^{\cap R} := L(q) \cap L(q)_R$. Note that neither $L(p)$ or $L(p)_R$ are shown in this figure, since our proof in fact will need to argue about how these cuts are situated relative to those shown. WLOG (as shown in the figure) p is to the left of q . Now, for contradiction, suppose that $e = \{a, b\} \in E(B^\rightarrow(p)) \cap E(B^\rightarrow(q))$. Then $a, b \in L(p)_R \cap L(q)_R$. So, in the above, since no cut contains a_0 , it must be that l' is to the left of a and r' is to the right of b .

Our method: The first try works if there are no “long” edges. So, to rectify that attempt we essentially “ignore” long edges (edges between distant atoms) in our charging argument and argue that they only contribute minimally to $E^\rightarrow(L(p))$ and $E^\leftarrow(R(p))$.

To this end, define $L(p)^{\cap R} := L(p) \cap L(p)_R$, and let $L^*(p)$ be the cut crossing $L(p)^{\cap R}$ on the left that *maximizes* the number of outside atoms in the intersection of $L^*(p)$ and $L(p)^{\cap R}$ (and similarly $R^*(p)$ to maximize the intersection with $R(p)^{\cap L}$ on the right). If $L^*(p)$ does not exist, i.e. no cut crosses $L(p)^{\cap R}$ on the left, set $L^*(p) = \emptyset$, and similarly for $R^*(p)$. See [Figure 43](#). We let:

$$\begin{aligned} E(B^\rightarrow(p)) &:= E(L(p)^{\cap R} \setminus L^*(p), L(p)_R \setminus L(p)^{\cap R}) \\ E(B^\leftarrow(p)) &:= E(R(p)^{\cap L} \setminus R^*(p), R(p)_L \setminus R(p)^{\cap L}). \end{aligned} \tag{59}$$

The following claim establishes (*) for [Eq. \(59\)](#). It can be proved using methods similar to [Claim 9.9](#); see [Figure 40](#).

Claim 9.12. For all polygon points p , $x(E(B^\rightarrow(p))), x(E(B^\leftarrow(p))) \geq 1 - O(\eta)$.

And finally, the following claim establishes (**):

Claim 9.13. For any edge e , we have $e \in E(B^\rightarrow(p))$ for at most one polygon point p and similarly $e \in E(B^\leftarrow(q))$ for at most one polygon point q .

The proof of [Claim 9.13](#) is more involved. For an outline of the arguments used, see [Figures 44, 45, 46, 47, 48](#).

9.5 Extending to polygons with inside atoms

For the general case, we still follow the same proof outline satisfying the four desiderata using the same definition of bad events given in [\(58\)](#) and the same definition of edges on which slack is increased in response to bad events given in [\(59\)](#).

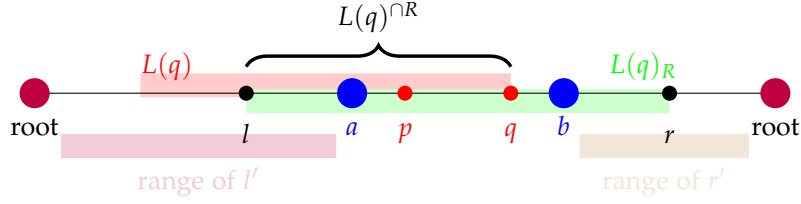


Figure 45: Proof of Claim 9.13 continued: Because we are dealing with outside atoms only and no cuts contain the root a_0 , we may as well visualize the polygon as a line (with wraparound) and each cut as an interval along the line. The above figure repeats Fig. 44 when viewed as a line segment.

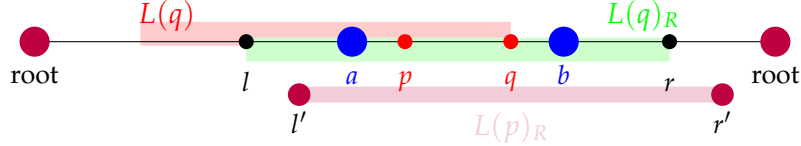


Figure 46: Claim: l' can not be to the right of l . If it is, as in the figure above, then $L(p)_R$ crosses $L(q)$ on the right and has a smaller intersection with $L(q)$. Contradiction to choice of $L(q)_R$!

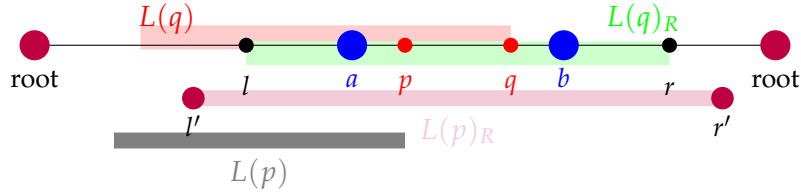


Figure 47: Claim: l' can not be to the left of l . If so, as in the figure above, then $L(q)_R$ crosses $L(p)$ on the right and has a smaller intersection with $L(p)$. Contradiction to choice of $L(p)_R$! Therefore $l' = l$.

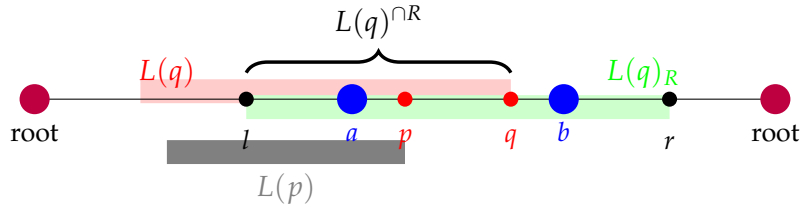


Figure 48: Since the left endpoint of $L(p)_R$ is l , the left endpoint of $L(p)$ is to the left of l . Therefore, $L(p)$ crosses $L(q)^{\cap R}$ on left and it is a candidate for $L^*(q)$. Therefore, $L(p) \cap L(q)^{\cap R} \subseteq L^*(q) \cap L(q)^{\cap R}$ and we have $a \in L^*(q) \cap L(q)^{\cap R}$. But then, $e \notin E(B^{\rightarrow}(q))$. Contradiction!.

We need to address the following challenges when generalizing the above proof:

- (1) Edges may have endpoints adjacent to inside atoms. The proof outline above crucially used that both endpoints of every edge were outside atoms.
- (2) Regions of the polygon we could previously assume were empty may now contain inside atoms.
- (3) While the sets are still defined such that they do not contain the root, the root (i.e., the atom containing $\{u_0, v_0\}$) is no longer necessarily an outside atom. Therefore we can have a sequence of cuts not containing the root wrapping around the polygon such that each cut

crosses the one before it. In this case, the notion of "left" and "right" becomes unclear. One can still define "left" and "right" synonymously with clockwise and counterclockwise, but we can no longer say that an outside atom is to the left (and not to the right of) another outside atom, nor can we collapse the diagonals of the polygon to intervals of a line.

To handle these complexities, we introduce additional structural properties of polygons with inside atoms. These are presented in [subsection 10.3](#).

10 Polygon Representation: Redux

We previously utilized the optimal Hamiltonian cycle to drastically simplify the polygon representation. In this section we gain a more in-depth understanding of the polygon representation to avoid this trick, thereby allowing us to bound the integrality gap of the subtour polytope.

We invite the reader to look at [Section 2.9](#) for a basic review of polygons. We present an example polygon here for reference, and provide a few additional facts not present in the earlier section.

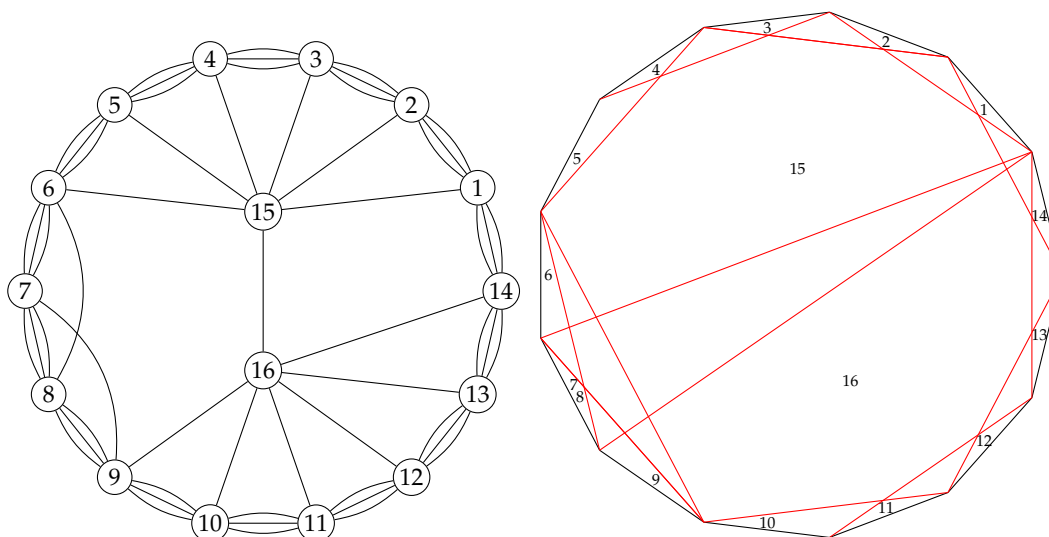


Figure 49: Consider the graph on the left, with minimum cut 7 (or, consider setting $x_e = \frac{2}{7}$ for all edges and a min cut of 2 as in the support graph of a solution to P_{Sub}). On the right is the polygon representation of the connected component of all proper cuts with at most 8 edges (or $x(\delta(S)) \leq 2 + 1/7$). This component consists of all proper near minimum cuts of the graph excluding the cut $\{7, 8\}$, which is in its own connected component of size 1. 15 and 16 are inside atoms, the others are outside atoms. Note $\{7, 8\}$ is a single atom.

The following observation follows from the fact that cuts correspond to straight diagonals in the plane and the polygon P is regular:

Observation 10.1 ([BG08, Prop 19]). *If $S, S' \in \mathcal{C}$ cross then $O(S)$ and $O(S')$ cross, and $O(S \cup S') \neq O(P)$.*

[Lemma 10.3](#) is used in the proof of [Lemma 10.20](#), and depends on the following:

Theorem 10.2 ([Ben97, Lem 4.1.7]). *Let $\mathcal{C}, \mathcal{C}'$ be two (distinct) connected components of crossing cuts for a family of cuts of $G = (V, E)$. Then, there exists an atom $a \in \mathcal{A}(\mathcal{C})$ and $a' \in \mathcal{A}(\mathcal{C}')$ such that $a \cup a' = V$.*

Lemma 10.3. *Consider the set of η -near minimum cuts (NMCs) of G and let \mathcal{C} be a connected component. Let $B \subset \mathcal{A}(\mathcal{C})$ be an η -NMC such that $1 < |B| < |\mathcal{A}(\mathcal{C})| - 1$. Then, $B \in \mathcal{C}$.*

Proof. For the sake of contradiction, suppose $B \notin \mathcal{C}$. Since B is an η -NMC, it is in some connected component of cuts, say $B \in \mathcal{C}'$, where $\mathcal{C} \neq \mathcal{C}'$. Then, by [Theorem 10.2](#) there exist atoms

$a \in \mathcal{A}(\mathcal{C}), a' \in \mathcal{A}(\mathcal{C}')$ such that $a \cup a' = V$. Observe that by the definition of atoms, each of a, a' is contained in either B or \overline{B} . Without loss of generality assume $a \subseteq B$ and $a' \subseteq \overline{B}$. But then, $a \neq B$ since $|B| > 1$. So, $a \cup a' \neq V$ which is a contradiction. \square

10.1 Our polygon notation and the root

All statements in the previous section do not depend on which side of each diagonal we consider. The ambiguity on the sides of the cut considered makes it difficult to define a consistent orientation on the polygon, for example to say whether a cut A, \overline{A} crosses B, \overline{B} “on the left” or “on the right.” Motivated by this, we identify every cut with the side that does not contain u_0, v_0 . This has the added benefit of allowing us to apply [Lemma 2.28](#) to every cut considered. For every polygon P , we call the unique atom r containing u_0, v_0 *the root*.

Recall that, for $\eta > 0$, we write $\mathcal{N}_\eta \subseteq 2^{V \setminus \{u_0, v_0\}}$ to denote the family of all η -NMCs of $G_{/e_0}$. If η is clear in context, we drop the subscript of \mathcal{N}_η . Throughout the paper, we will need to show that various sets $A, B \subseteq V \setminus \{u_0, v_0\}$ cross. Since $u_0, v_0 \in \overline{A \cup B}$ to verify that a pair of sets A and B cross, it suffices to check the three conditions in the following fact.

Fact 10.4. *For $A, B \subseteq V \setminus \{u_0, v_0\}$, A, B cross iff*

$$A \cap B, A \setminus B, B \setminus A \neq \emptyset.$$

As above, unless otherwise specified, we let \mathcal{C} be a connected component of cuts in \mathcal{N}_η with corresponding polygon P for $\eta \leq 1/10$. Again, call the *outside* atoms (of P) a_0, \dots, a_{m-1} , ordered counter-clockwise though these are not necessarily all the atoms in P . Note that the root r is not necessarily an outside atom, but if it is, it is the atom labelled a_0 .

We use the existence of the root to prove the following two facts. The first fact is a consequence of [Observation 10.1](#):

Fact 10.5. *Let $S, S' \in \mathcal{C}$. Then, $O(S \cup S') \neq O(\mathcal{A}(\mathcal{C}))$.*

Proof sketch. If S, S' are (the non-root) sides of two diagonals of a polygon and there is an atom r which is in neither of them, then there must be a polygon point which is in neither of them. \square

The following lemma is the main reason why we can treat each polygon separately in constructing the slack vector mentioned in [Section 9](#). In [Section 11](#), we are careful to only define positive slack on edges of a polygon that do not have an endpoint in its root.

Fact 10.6. *For any edge $e = \{u, v\}$, there is at most one polygon in which the endpoints of e lie in two different atoms which are not the root of their respective polygons.*

Proof. Suppose not, and let P, P' be two polygons in which u and v lie in different atoms that do not contain r . By [Theorem 10.2](#), there exists an atom $a \in P, a' \in P'$ such that $a \cup a' = V$. Since u, v lie in different atoms (and the atoms of a polygon partition V) in both P, P' it must be that (WLOG) $u \in a, v \in a'$. However, since $a \cup a' = V$, u_0, v_0 lies in a or a' , so either a is the root of P or a' is the root of P' , which is a contradiction. \square

We will use “left” synonymously with “clockwise” and “right” synonymously with “counter-clockwise.”

Definition 10.7 (Near min cut notation). We will interchangeably refer to a set $S \in \mathcal{N}_\eta$ by specifying the extreme outside atoms it contains or by specifying the polygon points defining its diagonal. For the former, if $S = (a_l, a_r)$, then a_l is the leftmost outside atom in S and a_r is the rightmost outside atom in S . For the latter, if $S = (p_l, p_r)$, then p_l is the polygon point immediately to the left of a_l and p_r is the polygon point immediately to the right of a_r .

Definition 10.8 ($L(p)$, $R(p)$). For a polygon point p_i , let $L(p_i)$ be the largest cut in \mathcal{N}_η containing a_i and not a_{i+1} which is crossed on both sides. Let $R(p_i)$ be the largest cut in \mathcal{N}_η containing a_{i+1} and not a_i which is crossed on both sides. (Note that $L(p_i), R(p_i)$ do not necessarily exist). See [Figure 39](#).

The following definitions make formal the notion of ‘‘crossed on one side’’ and ‘‘crossed on both sides’’ (introduced in [Section 9](#)) for polygons with inside atoms.

Definition 10.9 (Left, Right Crossing). Let $S, S' \in \mathcal{C}$ such that S' crosses S . For such a pair, we say S' crosses S on the left if the leftmost (clockwise-most) outside atom of $O(S' \cup S)$ is in S' . Otherwise, we say that S' crosses S on the right. Note that by [Observation 10.1](#), $O(S), O(S')$ cross.

Definition 10.10 (Crossed on one, both sides). We say a cut S is crossed on both sides if it is crossed by a cut (in \mathcal{C}) on the left and a cut (in \mathcal{C}) on the right and we say S is crossed on one side if it is crossed only on the left or only on the right.

Definition 10.11 ($\mathcal{N}_{\eta, \leq 1}, \mathcal{N}_{\eta, 1}, \mathcal{N}_{\eta, 2}$). Let $\mathcal{N}_{\eta, 2} \subseteq \mathcal{N}_\eta$ be the set of cuts which are crossed on both sides in their respective polygons. Let $\mathcal{N}_{\eta, 1} \subseteq \mathcal{N}_\eta$ be the set of cuts that are crossed on one side in their respective polygons and finally let $\mathcal{N}_{\eta, \leq 1} = \mathcal{N}_\eta \setminus \mathcal{N}_{\eta, 2}$ (i.e. the set of cuts which are crossed on one side or not crossed at all).

Here we give an alternate set-theoretic characterization of $\mathcal{N}_{\eta, 2}$.

Lemma 10.12. Let $C \in \mathcal{N}_\eta$. Then, $C \in \mathcal{N}_{\eta, 2}$ if and only if there exist two cuts $A, B \in \mathcal{N}_\eta$ which cross C such that $(A \setminus C) \cap (B \setminus C) = \emptyset$.

Proof. The only if follows from [Lemma 10.24](#). So, assume there exist two cuts $A, B \in \mathcal{N}_\eta$ which cross C such that $(A \setminus C) \cap (B \setminus C) = \emptyset$. We will show $C \in \mathcal{N}_{\eta, 2}$.

Since A and B both cross C , it must be that C, A, B are in the same connected component of cuts $\mathcal{C} \subseteq \mathcal{N}_\eta$ (i.e. including all cuts in $\mathcal{N}_{\eta, 2}$). Let P be the corresponding polygon.

Suppose by way of contradiction that A, B both crossed C on the left (if they both cross on the right the argument is similar). Then A, B must both contain the outside atom immediately to the left of the leftmost atom of C , which contradicts $(A \setminus C) \cap (B \setminus C) = \emptyset$. \square

Consequently, we can give an alternate characterization of $\mathcal{N}_{\eta, 1}$ as the sets which are crossed but are not in $\mathcal{N}_{\eta, 2}$. This will be relevant in [Appendix C](#).

Definition 10.13 (\mathcal{C}_2). For a connected component of cuts $\mathcal{C} \subseteq \mathcal{N}_\eta$, let $\mathcal{C}_2 = \mathcal{C} \cap \mathcal{N}_2$.

The following definition is quite important throughout our paper as it is used to specify the set of bad events we use to construct our slack vector (see [Section 9](#) for a gentle introduction):

Definition 10.14 (S_L, S_R). For $S \in \mathcal{C}_2$ let S_L be the near minimum cut crossing S on the left which minimizes $|O(S \cap S_L)|$. If there are multiple sets crossing S on the left with the same minimum intersection, choose the smallest one to be S_L . Similarly, let S_R be the near min cut crossing S on the right which minimizes $|O(S \cap S_R)|$, and again choose the smallest set to break ties. See [Figure 38](#).

10.2 Properties of inside atoms

Before proving some new properties of the polygon representation we recall some basic properties of inside atoms from [BG08]:

Definition 10.15 ([BG08, Definition 3]). *A family of sets $C_1, \dots, C_k \subseteq V$, for some $k \geq 3$, forms a k -cycle if*

- C_i crosses both C_{i-1} and C_{i+1} (we treat C_{k+1} as C_1 and C_0 as C_k);
- $C_i \cap C_j = \emptyset$ for $j \neq i-1, i$ or $i+1$; and
- $\bigcup_{1 \leq i \leq k} C_i \neq V$.
- If $k = 3$, we have the additional condition $(C_i \cap C_{i+1}) \not\subseteq C_{i-1}$ for $i \in \{1, 2, 3\}$.

Lemma 10.16 ([BG08, Lemma 22]). *Any k -cycle formed by cuts in a connected component \mathcal{C} of η -near min cuts satisfies $k \geq 2/\eta$. (Note if $\eta = 0$ then no k -cycle exists.)*

Therefore, for all $2/5$ -near min cuts, there are no cycles with length less than 5.

Lemma 10.17 ([BG08, Def 4]). *An atom $a \in \mathcal{A}(\mathcal{C})$ is an inside atom (in the representation defined above) if and only if there is a k -cycle $C_1, \dots, C_k \in \mathcal{C}$, such that $a \cap C_i = \emptyset$ for all $1 \leq i \leq k$.*

See [Figure 18](#) for an example of an inside atom and a cycle.

Fact 10.18. *Let C_1, \dots, C_k be a k -cycle for a connected component \mathcal{C} with polygon representation P and $k \geq 5$. For any adjacent pair of outside atoms $a, b \in O(\mathcal{A}(\mathcal{C}))$, there is a $1 \leq j \leq k$ such that $a, b \in C_j$.*

Proof. Since a is an outside atom, there is a cut C_i for some $1 \leq i \leq k$ such that $a \in C_i$ (otherwise a would be an inside atom). If $b \in C_i$ we are done. Otherwise, a is a rightmost or leftmost outside atom in C_i . But then, since C_i is crossed by C_{i-1} and C_{i+1} and $C_{i-1} \cap C_{i+1} = \emptyset$, it follows from [Observation 10.1](#) (and the fact that both C_{i-1}, C_{i+1} contain at least two outside atoms) that either $a, b \in C_{i-1}$ or $a, b \in C_{i+1}$. \square

10.3 New properties of polygon representations

The following lemmas build on [Ben95; BG08]. [Proposition 10.19](#) is a key property of polygons which [Lemma 10.20](#) extends:

Proposition 10.19 ([BG08, Proposition 20]). *For any connected component \mathcal{C} of η -near min cuts with $\eta \leq 2/5$ with polygon representation P , and any $S_1, S_2 \in \mathcal{C}$ with $S_1 \neq S_2$ we have $O_P(S_1) \neq O_P(S_2)$.*

Lemma 10.20. *Let P be the polygon representation for a connected component $|\mathcal{C}| > 1$ of η -NMCs of a (fractionally) 2-edge connected graph G with atom set $\mathcal{A}(\mathcal{C})$. If $A, B \subsetneq \mathcal{A}(\mathcal{C})$ are two $2/5$ -NMCs with $O_P(A) = O_P(B) \neq \emptyset$ and there is an atom $r \in \mathcal{A}(\mathcal{C})$ such that $r \notin A, B$, then $A = B$ ³⁵.*

³⁵As indicated earlier, r is called the root of the polygon P

Proof. Take the graph G and contract each atom of $\mathcal{A}(\mathcal{C})$ to a single node. Call the resulting graph G' . Clearly, G' is still 2-edge connected since G is 2-edge-connected and all cuts in \mathcal{C} are represented in G' . Now, consider the set of non-singleton $2/5$ -near-min-cuts of G' . This set has a unique connected component of crossing cuts because any new (non-singleton) cut $S \notin \mathcal{C}$ is crossed by a cut in \mathcal{C} . (Suppose not: then, no cut on the atoms of S crosses a cut on the atoms of \bar{S} , which contradicts that \mathcal{C} forms a connected component.) Call this component of cuts \mathcal{C}' and the corresponding polygon P' . It follows that $\mathcal{A}(\mathcal{C}) = \mathcal{A}(\mathcal{C}')$ (more precisely, a set $S \subseteq V$ is an atom in $\mathcal{A}(\mathcal{C})$ if and only if it is an atom in $\mathcal{A}(\mathcal{C}')$): no two atoms from P can be merged in P' because we have not deleted any cuts, and no atoms in P can be split in P' because we have contracted them. While some outside atoms of P may become inside atoms in P' , it follows by [Lemma 10.17](#) that any inside atom of P remains an inside atom in P' (as any k -cycle of \mathcal{C} is also a k -cycle of \mathcal{C}'). Therefore,

$$O_{P'}(A) = O_{P'}(B).$$

Therefore, if $A, B \in \mathcal{C}'$, by [Proposition 10.19](#), $A = B$. So it remains to show that $A, B \in \mathcal{C}'$. First, assume $2 \leq |A| \leq |\mathcal{A}(\mathcal{C}')| - 2$ and $2 \leq |B| \leq |\mathcal{A}(\mathcal{C}')| - 2$. Then, by [Lemma 10.3](#), $A, B \in \mathcal{C}'$ and we are done.

Now, we claim that $2 \leq |A| \leq |\mathcal{A}(\mathcal{C}')| - 2$ and $2 \leq |B| \leq |\mathcal{A}(\mathcal{C}')| - 2$, which by the above would complete the proof. For contradiction, assume $A \neq B$ and $|A| = 1$ or $|A| = |\mathcal{A}(\mathcal{C}')| - 1$. First assume $|A| = 1$. Since $B \neq A$, $O_P(A) = O_P(B) \neq \emptyset$, and every polygon has at least three outside atoms, $2 \leq |B| \leq |\mathcal{A}(\mathcal{C})| - 2$. By [Lemma 10.3](#), $B \in \mathcal{C}'$. Yet this implies that B has at least two outside atoms in P' (and therefore in P), which contradicts $O_{P'}(A) = O_{P'}(B)$. Otherwise, $|A| = |\mathcal{A}(\mathcal{C}')| - 1$. Using that $A \neq B$ and $r \notin A, B$, it follows that B has at most $|\mathcal{A}(\mathcal{C}')| - 2$ atoms. Again using that every polygon has at least three outside atoms, this implies $|B| \geq 2$. So similarly to above, we have $B \in \mathcal{C}'$. Therefore, $|O_{P'}(B)| \leq |O(P')| - 2$ which contradicts that $|A| = |\mathcal{A}(\mathcal{C}')| - 1$ using $O_{P'}(A) = O_{P'}(B)$. \square

We generalize [Observation 10.1](#) in the next lemma.

Lemma 10.21. *Let P be a polygon representation of a connected component \mathcal{C} of η NMCs of a (fractionally) 2-edge-connected graph G for some $\eta \leq 2/5$. For any $2/5$ NMCs $A, B \subseteq \mathcal{A}(\mathcal{C})$ with $O(A), O(B) \neq \emptyset$, if A, B cross, then $O(A), O(B)$ cross and $O(A \cup B) \neq O(\mathcal{A}(\mathcal{C}))$.*

Proof. Similar to the previous lemma, consider the graph G' arising from contracting all atoms of $\mathcal{A}(\mathcal{C})$ and let \mathcal{C}' be the (unique) connected component of non-singleton $2/5$ -near-min-cuts of G' with corresponding polygon P' . As before, $\mathcal{A}(\mathcal{C}') = \mathcal{A}(\mathcal{C})$ and $O(\mathcal{A}(\mathcal{C}')) \subseteq O(\mathcal{A}(\mathcal{C}))$.

Notice that since A, B cross (in P), each of them contains at least two atoms of P . Therefore, since A, B are $2/5$ NMCs and A, B are not singletons, we must have $A, B \in \mathcal{C}'$. Since A, B cross in P and $\mathcal{A}(\mathcal{C}') = \mathcal{A}(\mathcal{C})$, they also cross in P' . By [Observation 10.1](#), it follows that $O_{P'}(A)$ and $O_{P'}(B)$ cross. Recall that outside atoms of $\mathcal{A}(\mathcal{C})$ may become inside atoms of $\mathcal{A}(\mathcal{C}')$, but inside atoms of $\mathcal{A}(\mathcal{C})$ remain inside atoms in $\mathcal{A}(\mathcal{C}')$. So, $O_P(A)$ and $O_P(B)$ cross as well (in particular it also follows that $O(A \cup B) \neq O(\mathcal{A}(\mathcal{C}))$). \square

10.3.1 Almost diagonal cuts and the chain lemma

In some cases, we will need to refer to cuts which are generated by intersections of diagonals in \mathcal{C} . Such cuts are a subset of the following class:

Definition 10.22 (Almost Diagonal Cuts). Let \mathcal{C} be a connected component of cuts in \mathcal{N}_η . We say a set of atoms $S \subseteq \mathcal{A}(\mathcal{C}) \setminus \{r\}$ is an almost diagonal cut if:

1. S is a 2η -near min cut,
2. $\emptyset \neq O(S) \subsetneq O(\mathcal{A}(\mathcal{C}))$,
3. $O(S)$ forms a contiguous interval in the polygon.

Notice that by definition any cut in \mathcal{C} is an almost diagonal cut.

When we reference almost diagonal cuts in the rest of the paper, we will always assume $\eta \leq 1/5$. Notice that given any two $A, B \in \mathcal{N}_\eta$, $A \cap B, A \setminus B, B \setminus A$ are almost diagonal cuts.

The following fact implies that one can naturally define left/right crossing analogous to [Definition 10.14](#) for almost diagonal cuts. The following is a consequence of [Lemma 10.21](#):

Fact 10.23. Let A, B be two crossing almost diagonal cuts. Then, $O(A)$ and $O(B)$ cross. In addition, neither $O(A \cup B)$ nor $O(A \cap B)$ contain all outside atoms and each of $O(A \cup B)$ and $O(A \cap B)$ form a contiguous interval of outside atoms.

Lemma 10.24. For an almost diagonal cut S with leftmost atom a_l and rightmost atom a_r , let $L \in \mathcal{C}$ cross S on the left and $R \in \mathcal{C}$ cross S on the right. Then, if $\eta \leq 1/5$, $(L \setminus S) \cap (R \setminus S) = \emptyset$.

Proof. Suppose $(L \setminus S) \cap (R \setminus S) \neq \emptyset$. We will show that L, R, S form a 3-cycle ([Definition 10.15](#)) which cannot exist. To show this (using that the root atom $r \notin S \cup L \cup R$), it is enough to prove that all pairs cross and none of the three sets is a superset of the intersection of the two others. First, by assumption L and R cross S . L and R cross because $a_r \in R \setminus L, a_l \in L \setminus R$, so neither is a subset of the other, and because we assumed $(L \setminus S) \cap (R \setminus S) \neq \emptyset$ their intersection is nonempty.

In addition, we have $S \cap L \not\subseteq R$ because $a_l \in S \cap L$ but not in R . Similarly $S \cap R \not\subseteq L$. Finally, $L \cap R \not\subseteq S$ by the assumption $(L \setminus S) \cap (R \setminus S) \neq \emptyset$. Therefore S, L, R form a 3-cycle which is a contradiction as S, L, R are all $2/5$ -near min cuts. \square

A fundamental property of the cactus representation is that the set of min cuts A_1, \dots, A_k crossing a min cut S form two laminar families inside S . In other words, perhaps after renaming we may assume $A_1 \cap S \subseteq A_2 \cap S \subseteq \dots \subseteq A_j \cap S$ and $A_{j+1} \cap S \subseteq \dots \subseteq A_k \cap S$.

It is not immediately obvious that such a property extends to polygons because of the existence of inside atoms. Nonetheless, the following lemma demonstrates that this property is also true of near min cuts provided that η is small enough. It is an immediate corollary of [Lemma 10.26](#).

Lemma 10.25 (Chain Lemma). Let S be an almost diagonal cut where $O(S)$ contains all outside atoms from a to b . In addition, let $A_1, \dots, A_k \in \mathcal{C}$ be the collection of η -near-min cuts crossing S on the left. Then there is a permutation, $\pi : [k] \rightarrow [k]$ such that

$$S \cap S_L = S \cap A_{\pi(1)} \subseteq \dots \subseteq S \cap A_{\pi(k)}$$

i.e., their intersections with S form a chain. The same statements also hold for cuts crossing S on the right.

Lemma 10.26. Let S be a near diagonal cut with leftmost outside atom a and rightmost outside atom b . Furthermore, let $A = [a_1, a_2], B = [b_1, b_2] \in \mathcal{C}$ be cuts which cross S on the right. If $\eta \leq 1/10$ and in the interval of the outside atoms of $O(S)$, a_1 is to the left of b_1 , then $B \cap S \subseteq A \cap S$. In the special case that $a_1 = b_1$, we have $S \cap A = S \cap B$.

Proof. First assume $b_1 \neq a_1$. Since b_1 is to the right of a_1 , and A, B cross S on the right, $O(A \cap S \cap B) = O(S \cap B) \neq \emptyset$ (as both sets have all outside atoms between b_1 and b). Note that B crosses $A \cap S$. This is because B has an atom outside of S (as it crosses S itself), $a_1 \in A \cap S \setminus B$, and $b_1 \in A \cap S \cap B$. So by [Lemma 2.35](#) $A \cap S \cap B$ is a 4η near min cut. In addition, $S \cap B$ is a 3η near min cut since B crosses S . Therefore, since $4\eta \leq 2/5$, by [Lemma 10.20](#) $A \cap S \cap B = S \cap B$.

In the special case that $a_1 = b_1$, $O(A \cap S) = O(B \cap S) \neq \emptyset$ has all outside atoms between $a_1 = b_1$ and b . Since $A \cap S, B \cap S$ are 3η near min-cuts, by [Lemma 10.20](#), we must have $A \cap S = B \cap S$ as desired. \square

10.4 Another structural property of inside atoms

The following lemma is not explicitly used in the proof of the main theorem, but the statement may be useful to guide the reader's intuition (or in other settings where near min cuts arise). For example, it implies that the yellow region is empty in [Fig. 40](#).

Lemma 10.27. *Let $\eta \leq 2/5$ and let P be the polygon representation for a connected component $|\mathcal{C}| > 1$ of η -NMCs of a (fractionally) 2-edge-connected graph. Suppose H is the intersection of half-planes³⁶ $H_0, \dots, H_{\ell-1}$ corresponding to diagonals $D_0, \dots, D_{\ell-1}$ of P that has a positive area. If H does not contain any side of P (equivalently, it does not contain any outside atom) and $\ell < 1/\eta$ then H does not have any inside atoms.*

Proof. Define $C_0, \dots, C_{\ell-1} \subseteq V$ such that $C_i = \mathcal{A}(\mathcal{C}) \cap \overline{H}_i$, i.e. all atoms which are not in the halfplane H_i . Without loss of generality assume there are no two sets C_i, C_j such that $C_i \subseteq C_j$ ³⁷.

First observe that H is a 2-dimensional polytope and therefore without loss of generality we can assume each D_i defines a side of H and $D_0, \dots, D_{\ell-1}$ are ordered such that the vertices of the polytope $v_0, \dots, v_{\ell-1}$ are arranged cyclically counterclockwise where v_i is the intersection of D_i and D_{i+1} (where for the rest of the proof we take all indices mod ℓ). We will call a vertex *external* if it is a polygon point of P and *internal* otherwise.

We prove the claim by induction over r that if H has r external vertices and $\ell + r < 2/\eta$, then H is empty. Note that if $\ell < 1/\eta$, since $r \leq \ell$, we have $\ell + r < 2/\eta$ which proves the theorem.

First assume $r = 0$. Then, C_i crosses C_{i+1} for all i . By way of contradiction suppose H contains an (inside) atom a . Without loss of generality, assume that $H_0, \dots, H_{k-1} \subseteq H_0, \dots, H_{\ell-1}$ (perhaps after renaming) is the minimal set of half-planes that contain H and have no external vertices. We claim that there are no indices i, j and $j \neq i-1, i$, or $i+1$ such that $C_i \cap C_j \neq \emptyset$. Since $C_i \not\subseteq C_j$, $C_j \not\subseteq C_i$, $a \notin C_i, C_j$, it follows that C_i, C_j cross and therefore the diagonals D_i, D_j intersect in the interior of P . Now $D_0, \dots, D_i, D_j, \dots, D_{k-1}$ contains H and has no external vertices and thus contradicts the minimality of H_0, \dots, H_{k-1} .

Therefore by minimality, $C_i \cap C_j = \emptyset$ if $j \neq i-1, i$ or $i+1$. So, C_0, \dots, C_{k-1} is a k -cycle for a : since there are no external vertices, C_i crosses C_{i+1} for all i , and $\bigcup_{i=0}^{\ell-1} C_i \neq V$ since $a \notin C_i$ for all i . By [Lemma 10.16](#), $k \geq 2/\eta$, which is a contradiction, because $k \leq \ell < 1/\eta$.

Now, suppose the claim is true when the number of external vertices is at most r ; we will prove it holds when the number is $r+1$.

Again, by way of contradiction suppose there is an inside atom a in H . Then there is a k -cycle for a , $L_1, \dots, L_k \in \mathcal{C}$. Now pick an arbitrary external vertex $v_i = p_j$ of H for some

³⁶Technically, half-polygons.

³⁷This is because if $C_i \subseteq C_j$ then $H_j \cap P \subseteq H_i \cap P$ which means H_j is redundant in defining H .

j. Let $a_{j-1}, a_j \in O(\mathcal{A}(\mathcal{C}))$ be the adjacent outside atoms immediately to the clockwise and counterclockwise of p_j . Therefore, by [Fact 10.18](#), there exists some j such that $a_{j-1}, a_j \in L_j$. Let H_ℓ be the halfspace corresponding to the side of L_j containing a . Now consider the set of halfspaces $H_0, \dots, H_{\ell-1}, H_\ell$. Note that $H' = \bigcap_{i=0}^{\ell} H_i \subsetneq H$ because $v_i = p_j \notin H_\ell$. Therefore, H' has at least one fewer external vertex and no sides of the polygon. Since $\ell + 1 + (r - 1) \leq 2/\eta$, by the induction hypothesis, H' has no inside atoms which is a contradiction with the existence of a . \square

11 Using the Polygon Representation for Cuts Crossed on Both Sides

11.1 Notation and a preliminary lemma

We will use the same set of definitions for bad events and increase sets that we did in [Section 9](#) for polygons without inside atoms. For the benefit of the reader we repeat them here. Recalling the definitions from [Section 10.1](#), partition each set $\delta(S)$ into three sets $E^{\leftarrow}(S), E^{\rightarrow}(S)$ and $E^{\circ}(S)$ such that

$$\begin{aligned} E^{\leftarrow}(S) &= E(S \cap S_L, S_L \setminus S) \\ E^{\rightarrow}(S) &= E(S \cap S_R, S_R \setminus S) \\ E^{\circ}(S) &= \delta(S) \setminus (E^{\leftarrow}(S) \cup E^{\rightarrow}(S)) \end{aligned}$$

In addition we define the left and right bad events:

$$\begin{aligned} B^{\rightarrow}(p) &= \mathbb{1}\{|E^{\rightarrow}(L(p)) \cap T| \neq 1 \text{ or } |E^{\circ}(L(p)) \cap T| \neq 0\} \\ B^{\leftarrow}(p) &= \mathbb{1}\{|E^{\leftarrow}(R(p)) \cap T| \neq 1 \text{ or } |E^{\circ}(R(p)) \cap T| \neq 0\}. \end{aligned} \tag{60}$$

If $L(p)$ does not exist, simply assume the left bad event never occurs, and similarly if $R(p)$ does not exist assume the right bad event never occurs.

Define $L(p)^{\cap R} := L(p) \cap L(p)_R$, and let $L^*(p) \in \mathcal{C}$ be the cut crossing $L(p)^{\cap R}$ on the left that maximizes $|O(L^*(p) \cap L(p)^{\cap R})|$ (and similarly $R^*(p)$ to maximize the intersection with $O(R(p)^{\cap L})$ on the right). If $L^*(p)$ does not exist, i.e. no cut crosses $L(p)^{\cap R}$ on the left, set $L^*(p) = \emptyset$, and similarly for $R^*(p)$. We let:

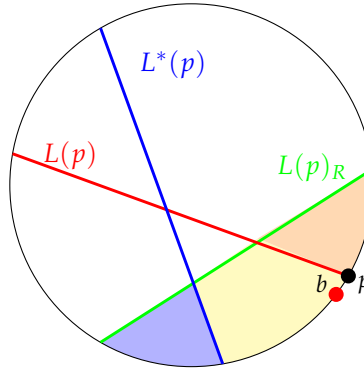


Figure 50: Recap of some basic definitions: $L(p)$ is the cut crossed on both sides with rightmost polygon point p (and contains all atoms below the red diagonal). $L(p)_R$ is the cut crossing $L(p)$ on the right that minimizes the number of outside atoms in $L(p)^{\cap R} = L(p) \cap L(p)_R$, i.e., in yellow + blue. Note that the cut $L(p)^{\cap R}$ contains all atoms in the yellow and blue regions (which may include inside atoms). Since this region is the set difference of two η near min cuts, it is a 2η near min cut. $L^*(p)$ is the cut crossing $L(p)^{\cap R}$ on the left that maximizes the number of outside atoms in the intersection, i.e., maximizes the number of outside atoms in the blue region. $E^{\rightarrow}(L(p))$ are the edges between atoms in the yellow region and atoms in the orange region. There is one edge in the tree that is in $E^{\rightarrow}(L(p))$ with probability $1 - O(\eta)$ and when this event does not occur, $B^{\rightarrow}(p)$ occurs. ($B^{\rightarrow}(p)$ also occurs if $|E^{\circ}(L(p)) \cap T| \neq 0$.)

$$\begin{aligned}
E(B^\rightarrow(p)) &:= E(L(p)^{\cap R} \setminus L^*(p), L(p)_R \setminus L(p)^{\cap R}) \\
E(B^\leftarrow(p)) &:= E(R(p)^{\cap L} \setminus R^*(p), R(p)_L \setminus R(p)^{\cap L}).
\end{aligned} \tag{61}$$

The following important lemma is the generalization of [Claim 9.10](#) from [Section 9](#) to the case in which there may be inside atoms. It uses that by [Lemma 10.20](#) many regions of the polygon do not contain inside atoms.

Lemma 11.1. *Let $A, B \in \mathcal{C}_2$ such that $A = (a_1, a_r)$ and $B = (a_2, a_r)$ share a rightmost polygon point p_r . Then, $A_R = B_R$ and $E^\rightarrow(A) = E^\rightarrow(B)$.*

Proof. WLOG assume $A \subseteq B$. First, we will prove that if a cut R crosses A on the right, it also crosses B on the right. So, let R be a set crossing A on the right. Then, since R contains a_r , $R \cap B \neq \emptyset$. Furthermore, R contains atom a_{r+1} , so $R \not\subseteq B$. Finally, B contains a_1 since $A \subseteq B$ yet R does not. Therefore, R crosses B on the right.

Therefore, by [Definition 10.14](#) $A_R = B_R$ since the set of cuts crossing B on the right is a superset of cuts crossing A on the right and any cut which crosses B but not A contains all atoms of A , so would have a larger intersection with $O(B)$. (If two sets have the same intersection with $O(A)$ we use the same tie-breaking rule for both A_R and B_R .)

Now we prove that $E^\rightarrow(A) = E^\rightarrow(B)$. Let $R = A_R = B_R$. It suffices to show that $R \cap A = R \cap B$ and $R \setminus A = R \setminus B$ because any edge $e \in E^\rightarrow(A)$ has one endpoint in $R \cap A$ and one in $R \setminus A$. To obtain $R \cap A = R \cap B$ notice that $O(R \cap A) = O(R \cap B) \neq \emptyset$ and by [Lemma 2.35](#) $R \cap A, R \cap B$ are 2η near minimum cuts (since R crosses both A and B), so by [Lemma 10.20](#), $R \cap A = R \cap B$. Similarly to obtain $R \setminus A = R \setminus B$ notice that $O(R \setminus A) = O(R \setminus B) \neq \emptyset$ and $R \setminus A, R \setminus B$ are 2η near min cuts so by [Lemma 10.20](#) we have $R \setminus A = R \setminus B$. \square

11.2 Main theorem

The following is the main technical result of the paper:

Theorem 11.2 (Main theorem). *Let x^0 be a feasible LP solution of [\(2\)](#) with support $E_0 = E \cup \{e_0\}$ and let x be x^0 restricted to E . For any distribution μ of spanning trees with marginals x , $0 < \eta \leq 1/10$ and $\alpha > 0$, there is a random vector $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (the randomness in s^* depends exclusively on $T \sim \mu$) such that*

- For any η -near minimum cut S which is crossed on both sides, if $\delta(S)_T$ is odd then $s^*(\delta(S)) \geq \alpha(1 - \eta)$;
- For any $e \in E$, $\mathbb{E}[s_e^*] \leq 18\alpha\eta x_e$.

Our slack vector for the above theorem will be exactly as in [Section 9](#). In particular, for every bad event B which occurs among those defined in [Eq. \(64\)](#), we will set $s^*(e) = \alpha x_e$ for all $e \in E(B)$, where $E(B)$ is defined as in [Eq. \(61\)](#). In order to extend the argument from [Section 9](#) to prove this theorem in the case in which the polygon contains inside atoms, we prove the following two lemmas from which the theorem follows easily:

Lemma 11.3 (All cuts are satisfied). *Let $S = (p_l, p_r)$ be a cut which is crossed on both sides. Then, if $\delta(S)_T \neq 2$, at least one of $B^\leftarrow(p_l), B^\rightarrow(p_r)$ occurs.*

Lemma 11.4 (Every edge is mapped to a constant number of bad events). *Let p, q be two polygon points such that $e = \{a, b\}$ and $a \in L(p) \cap L(q)$. Then, $e \notin E(B^\rightarrow(p)) \cap E(B^\rightarrow(q))$.*

Before proving these statements, we will show how they imply our main theorem. First we gives proofs for [Claim 9.11](#) and [Claim 9.12](#) (as the formal proofs were omitted in the overview):

Lemma 11.5. *For any polygon point p , $\mathbb{P}[B^\rightarrow(p)] \leq 4.5\eta$ and $\mathbb{P}[B^\leftarrow(p)] \leq 4.5\eta$.*

Proof. We will prove this for $B^\rightarrow(p)$, $B^\leftarrow(p)$ follows similarly. To simplify notation we abbreviate $L(p)$ to L . Since L is crossed on both sides, L_L, L_R are well defined. Since by [Lemma 2.35](#) $L_R \cap L, L_R \setminus L$ are 2η -near min cuts and L_R is an η -near mincut with respect to x , by [Corollary 2.29](#), $\mathbb{P}[E^\rightarrow(L)_T = 1] \geq 1 - 2.5\eta$.

On the other hand, since L, L_L, L_R are η -near min cuts, by [Lemma 2.36](#), $x(E^\rightarrow(L)), x(E^\leftarrow(L)) \geq 1 - \eta/2$. Therefore

$$x(E^\circ(L)) \leq 2 + \eta - x(E^\leftarrow(L)) - x(E^\rightarrow(L)) \leq 2\eta.$$

It follows that $\mathbb{P}[E^\circ(L)_T = 0] \geq 1 - 2\eta$. Finally, by the union bound, all events occur simultaneously with probability at least $1 - 4.5\eta$, which gives the lemma. \square

Lemma 11.6. *For any polygon point p , $x(E(B^\leftarrow(p))), x(E(B^\rightarrow(p))) \geq 1 - \eta$*

Proof. First, observe that $L^*(p)$ crosses $L(p)_R$. Notice (where we use \uplus to denote disjoint union):

$$L(p)^{\cap R} \setminus L^*(p) \uplus L(p)_R \setminus L(p) = L(p)_R \setminus L^*(p).$$

Therefore, by [Lemma 2.35](#) $L(p)^{\cap R} \setminus L^*(p) \uplus L(p)_R \setminus L(p)$ is a 2η near mincut. So, by ??, $x(E(B^\leftarrow(p))) \geq 1 - \eta$.

The proof for $x(E(B^\rightarrow(p)))$ is similar. \square

Proof of Theorem 11.2. Our slack vector is defined as follows. Initialize $s^*(e) = 0$ for all edges e . Then for each polygon P , for each polygon point $p \in P$, whenever $B^\leftarrow(p)$ occurs, let $s_e^* = \alpha x_e$ for each $e \in E(B^\leftarrow(p))$. Whenever $B^\rightarrow(p)$ occurs, let $s_e^* = \alpha x_e$ for each $e \in E(B^\rightarrow(p))$.

Now we show the first condition of the theorem. Let $S = [p, q] \in \mathcal{C}_2$ and suppose that $\delta(S)_T$ is odd. It appears in some polygon P . Then by [Lemma 11.3](#), either $B^\leftarrow(p)$ or $B^\rightarrow(q)$ has occurred. Assume the former, the other case is similar. In this event, we set $s^*(e) = \alpha x_e$ for all $e \in E(B^\leftarrow(p))$. However, using [Lemma 11.1](#), we have

$$E(B^\leftarrow(p)) \subseteq E^\leftarrow(p) = E^\leftarrow(S)$$

Therefore, by [Lemma 11.6](#) $s^*(S) \geq \alpha(1 - \eta)$ as desired.

Now we verify the second condition of the theorem. First note that by [Fact 10.6](#), for any edge e , there is at most one polygon P such that e does not have the root as one of its endpoints. Therefore, there is at most one polygon P for which s_e^* may be increased since if e is adjacent to the root, $e \notin E^\rightarrow(S), E^\leftarrow(S)$ for any set S in its connected component.

Now let $e = \{a, b\}$ for some polygon P such that $a, b \neq r$. We show that there are at most two polygon points p for which $e \in E(B^\rightarrow(p))$. Suppose otherwise and there are at least three such polygon points p . Since $E(B^\rightarrow(p)) \subseteq \delta(L(p))$ for each such point p , we have $e \in \delta(L(p))$, which implies that there are two polygon points p, q such that one of a or b , WLOG a , is in both $L(p)$ and $L(q)$ and $e \in E(B^\rightarrow(p)) \cap E(B^\rightarrow(q))$. However this contradicts [Lemma 11.4](#) since e is in at

most one such set. One can similarly show that there are at most two polygon points p such that $e \in E(B^\leftarrow(p))$. Therefore, any edge e is in $E(B)$ for at most four bad events B .

By [Lemma 11.5](#), each bad event occurs with probability at most 4.5η . Therefore, by the union bound:

$$\mathbb{E}[s^*(e)] \leq 4 \cdot 4.5\eta \alpha x_e = 18\alpha\eta x_e,$$

which gives the second condition and completes the proof. \square

11.3 All cuts are satisfied

In this section we first prove the following from which [Lemma 11.3](#) will easily follow:

Lemma 11.7. For $S = (p_l, p_r) \in \mathcal{C}_2$, $E^\circ(S) \subseteq E^\circ(L(p_r)) \cup E^\circ(R(p_l))$.

Note that if $S = (p_l, p_r) \in \mathcal{C}_2$ then $L(p_r)$ and $R(p_l)$ exist, because S is a candidate for both.

Before giving the proof of this we show how it implies [Lemma 11.3](#).

Lemma 11.3 (All cuts are satisfied). Let $S = (p_l, p_r)$ be a cut which is crossed on both sides. Then, if $\delta(S)_T \neq 2$, at least one of $B^\leftarrow(p_l), B^\rightarrow(p_r)$ occurs.

Proof. We prove by contradiction. Suppose none of $B^\leftarrow(p_l), B^\rightarrow(p_r)$ occur; we will show that this implies $\delta(S)_T = 2$.

Let $R = R(p_l)$. By [Lemma 11.1](#) we have $S_L = R_L$ and $E^\leftarrow(R) = E^\leftarrow(S)$. Similarly for $L = L(p_r)$ we have $E^\rightarrow(L) = E^\rightarrow(S)$.

Now, since $B^\leftarrow(p_l)$ has not occurred,

$$1 = E^\leftarrow(R)_T = E^\leftarrow(S)_T \text{ and } E^\circ(R)_T = 0$$

and since $B^\rightarrow(p_r)$ has not occurred,

$$1 = E^\rightarrow(L)_T = E^\rightarrow(S)_T \text{ and } E^\circ(L)_T = 0$$

So, to get $\delta(S)_T = 2$, it remains to show that $T \cap E^\circ(S) = \emptyset$. By [Lemma 11.7](#), we have $E^\circ(S) \subseteq E^\circ(L) \cup E^\circ(R)$, which gives the claim. \square

To prove [Lemma 11.7](#) we first need the following:

Corollary 11.8. For all sets $S \in \mathcal{C}_2$, we have $E^\leftarrow(S) \cap E^\rightarrow(S) = \emptyset$. Similarly, for all sets $A, B \in \mathcal{C}_2$ such that B crosses A on the right, $E^\leftarrow(A) \cap E^\rightarrow(B) = \emptyset$.

Proof. To see the first claim, suppose S_L crosses S on the left and S_R crosses S on the right, by [Lemma 10.24](#) $(S_L \setminus S) \cap (S_R \setminus S) = \emptyset$. Since every edge in $E^\leftarrow(S)$ has an endpoint in $S_L \setminus S$ and every edge in $E^\rightarrow(S)$ has an endpoint in $S_R \setminus S$, this proves the claim.

A similar argument using [Lemma 10.24](#) proves the second claim. \square

Now we can prove the main lemma:

Lemma 11.7. For $S = (p_l, p_r) \in \mathcal{C}_2$, $E^\circ(S) \subseteq E^\circ(L(p_r)) \cup E^\circ(R(p_l))$.

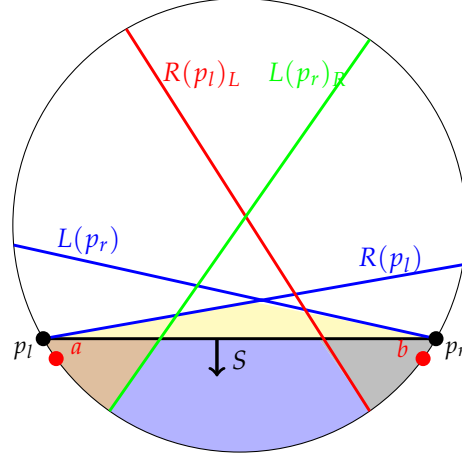


Figure 51: Here we have a set S which is crossed on both sides. We use in [Lemma 11.7](#) that $L(p_r) \cap R(p_l) = S$; in other words, the yellow region is empty.

Proof. For convenience, let $L = L(p_r)$ and $R = R(p_l)$. First suppose we had $L = S$ or $R = S$. In this case, by definition $E^\circ(S) = E^\circ(L)$ or $E^\circ(R)$, and we are done. So assume that $S \subsetneq L, R$. Therefore, L and R cross.

Now, notice that since $O(L \cap R) = O(S)$, by [Lemma 10.20](#), we have $L \cap R = S$. This implies $\delta(S) \subseteq \delta(L) \cup \delta(R)$. To see this, let e be an edge in $\delta(S)$. Then, it has an endpoint in both L and R . However, its other endpoint is in $\bar{S} = \bar{L} \cap \bar{R}$, and therefore cannot be in both L and R which implies it is in $\delta(L)$ or $\delta(R)$.

Now, by way of contradiction, suppose there exists an edge $e \in E^\circ(S)$ such that $e \notin E^\circ(L) \cup E^\circ(R)$. Since $E^\leftarrow(S) = E^\leftarrow(R)$, $E^\rightarrow(S) = E^\rightarrow(L)$, and $e \in \delta(L) \cup \delta(R)$, it must be that $e \in E^\leftarrow(L) \cup E^\rightarrow(R)$. Since L and R cross, by [Corollary 11.8](#) e is in exactly one of $E^\leftarrow(L), E^\rightarrow(R)$; assume that $e \in E^\leftarrow(L)$ but not in $E^\rightarrow(R)$, the other case is similar. Therefore $e \notin \delta(R)$, since it is not in $E^\leftarrow(R), E^\circ(R)$ or $E^\rightarrow(R)$.

However, L_L crosses L on the left and R crosses L on the right. Therefore, by [Lemma 10.24](#), we have $(L_L \setminus L) \cap (R \setminus L) = \emptyset$. However e has one endpoint in $L \cap R = S$, one in $R \setminus L$ (since $e \notin \delta(R), e \in \delta(L)$), and one in $L_L \setminus L$, which is a contradiction since all three sets are disjoint. \square

11.4 Every cut is mapped to a constant number of bad events

In this section we prove [Lemma 11.4](#).

Lemma 11.4 (Every edge is mapped to a constant number of bad events). *Let p, q be two polygon points such that $e = \{a, b\}$ and $a \in L(p) \cap L(q)$. Then, $e \notin E(B^\rightarrow(p)) \cap E(B^\rightarrow(q))$.*

Proof. First note that by [Fact 10.5](#), $O(L(p) \cup L(q)) \neq O(\mathcal{A}(\mathcal{C}))$, so it forms a contiguous interval. WLOG assume that q is the rightmost point in this interval.

Suppose by way of contradiction that $e \in E(B^\rightarrow(p)) \cap E(B^\rightarrow(q))$. In the below claim, we will show that $L(p)_R$ crosses $L(q)$ on the right. Now we will show that $L(p)$ crosses $L(q)^{\cap R}$ on the left, which would complete the proof. This is because $L(p)$ is a candidate for $L^*(q)$, and by [Lemma 10.26](#), $L(p) \cap L(q)^{\cap R} \subseteq L^*(q) \cap L(q)^{\cap R}$, which implies $a \in L^*(q)$ and therefore we could not have $e \in E(B^\rightarrow(q))$.

It remains to show that $L(p)$ crosses $L(q)^{\cap R}$ on the left. By assumption, $a \in L(p) \cap L(q)^{\cap R} \neq \emptyset$. $L(q)$'s rightmost atom is in $L(q)^{\cap R} \setminus L(p)$. So it remains to show that $L(p) \not\subseteq L(q)^{\cap R}$. By way of contradiction suppose $L(p) \subseteq L(q)^{\cap R} = L(q) \cap L(q)_R$. Since by the following claim, $L(p)_R$ crosses $L(q)$ on the right, $L(p)_R$ is a candidate for $L(q)_R$. We will show that $|O(L(p)_R \cap L(q))| < |O(L(q)_R \cap L(q))|$ which contradicts [Definition 10.14](#). Since $L(p)_R$ and $L(q)_R$ both cross $L(q)$ on the right, to prove the inequality it's enough to show that the leftmost outside atom of $L(q)_R$ is not in $L(p)_R$. However, this is immediate because $L(p)_R$ does not have the leftmost outside atom of $L(p)$, yet $L(p) \subseteq L(q)^{\cap R}$.

Claim 11.9. $L(p)_R$ crosses $L(q)$ on the right.

Proof. First we show that $L(p)_R$ crosses $L(q)$. Note $b \in L(p)_R \setminus L(q) \neq \emptyset$, and $a \in L(p)_R \cap L(q) \neq \emptyset$. So, $L(p)_R$ crosses $L(q)$ unless $L(q) \subseteq L(p)_R$. For contradiction, assume $L(q) \subseteq L(p)_R$.

Now we claim that $L(p)$ crosses $L(q)$. By assumption, $a \in L(p) \cap L(q)$. $L(q) \subseteq L(p)_R$ implies $L(p) \not\subseteq L(q)$. Since q is the rightmost point of the interval $O(L(p) \cup L(q))$, the rightmost atom of $L(q)$ is in $L(q) \setminus L(p)$, giving $L(q) \not\subseteq L(p)$. Therefore, $L(q)$ crosses $L(p)$ on the right.

Therefore, $L(q)$ is a candidate set for $L(p)_R$, but since $L(q) \subseteq L(p)_R$, we must have $L(q) = L(p)_R$. Yet $b \in L(p)_R \setminus L(q)$ which contradicts this.

Now we establish that $L(p)_R$ crosses $L(q)$ on the right. For contradiction, suppose it crosses on the left. Then, by [Lemma 10.24](#), we must have $(L(p)_R \setminus L(q)) \cap (L(q)_R \setminus L(q)) = \emptyset$, which contradicts the fact that b lies in both sets. □

□

12 Putting everything together for the integrality gap

In this section we use the following theorem to demonstrate [Theorem 1.2](#), bounding the integrality gap of the subtour polytope. While the proof of [Theorem 12.1](#) is non-trivial, using [Theorem 11.2](#) it follows from statements in [\[KKO21\]](#) and does not require any new ideas. For this reason, we sketch the proof in this section, leaving the formal proof to [Appendix D](#).

It turns out not to be useful to prove [Theorem 9.8](#) directly, but is an immediate corollary of the following theorem (we stated [Theorem 9.8](#) in the overview to improve readability and highlight the importance of [Theorem 11.2](#)).

Theorem 12.1 (Combination of [Theorem 9.8](#) and [Theorem 11.2](#)). *Let x^0 be a solution of LP (2) with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E . Let $\eta \leq 10^{-12}$, $\beta > 0$ and let μ be the max-entropy distribution with marginals x . Then there are two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$), such that*

- i) For each edge $e \in E$, $s_e \geq -x_e\beta$ (with probability 1).
- ii) For each $S \in \mathcal{N}_\eta$, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.
- iii) For every edge e , $\mathbb{E}[s_e^*] \leq 125\eta\beta x_e$ and $\mathbb{E}[s_e] \leq -\frac{1}{3}x_e\epsilon_P\beta$, where ϵ_P is defined in [Theorem D.2](#).

In [Section 12.2](#) we show how this theorem implies [Theorem 1.2](#), which is very similar to the proof of [Theorem 4.1](#) using [Theorem 4.6](#) (although with the crucial difference that OPT is not used). Now we sketch the ideas underlying the proof of [Theorem 12.1](#). To make this section as accessible as possible, we oversimplify and ignore the details of how parameters are set.

In the above theorem, the role of s is to generate gain over $3/2$. Roughly speaking, we follow the lead of [\[KKO21\]](#) and divide \mathcal{N}_η into three categories: cuts crossed on both sides, cuts crossed on one side, and the remainder, which form a laminar family \mathcal{H} defined below. We define an s^* vector to provide significant positive slack on each odd cut that is crossed; in particular, we start with the vector defined in [Theorem 11.2](#) and augment it to handle cuts crossed on one side. We will ensure that the expected cost of s^* is negligible³⁸. Now in \mathcal{H} , there are only a linear number of cuts and they have a simple structure (for example, most edges are only in a constant number of cuts of \mathcal{H}), so it is manageable to design a vector s which generates negative slack in expectation while still satisfying every cut in \mathcal{H} .

First, we explain how to augment s^* from [Theorem 11.2](#) to handle cuts crossed on one side. Observe that any polygon associated to a connected component in $\mathcal{N}_{\eta, \leq 1}$ contains no inside atoms. This follows from the fact that the existence of an inside atom is predicated on the existence of a k -cycle, which by its very definition contains cuts crossed on both sides. Thus, each connected component \mathcal{C} of $\mathcal{N}_{\eta, \leq 1}$ consists only of outside atoms, where a_0 is the root.

A key structure needed for the construction of the slack vector s is a laminar family of cuts \mathcal{H} that we call a *hierarchy*. This hierarchy \mathcal{H} includes the following set of cuts:

- The set of cuts in $\mathcal{N}_{\eta, \leq 1}$ that are not crossed by any other cut in $\mathcal{N}_{\eta, \leq 1}$;
- The cut consisting of the union of the non-root atoms $\{a_1, \dots, a_{m-1}\}$ of each connected component \mathcal{C} of $\mathcal{N}_{\eta, \leq 1}$, which (in this section) we call the *outer polygon cut* for \mathcal{C} , and

³⁸i.e., $\mathbb{E}[s_e^*] \leq 18\alpha\eta x_e$, which will ultimately be $O(\eta\beta x_e)$. In the end, this increase is dwarfed by a decrease in s_e of $\Omega(\beta x_e)$ since η is a minuscule constant.

- The atoms a_i , $1 \leq i \leq m - 1$ of each connected component \mathcal{C} of $\mathcal{N}_{\eta, \leq 1}$.

Notice that \mathcal{H} *excludes* some cuts in $\mathcal{N}_{\eta, \leq 1}$, namely all the near min cuts in any polygon P of $\mathcal{N}_{\eta, \leq 1}$ that are *not* outer polygon cuts. It also *includes* some cuts that are *not* in $\mathcal{N}_{\eta, \leq 1}$. For example, the outer polygon cut itself may not be an η near min cut, and there may be atoms in some polygon that are not η near min cuts. However, one of the consequences of the following theorem is that these extra cuts are ϵ_η near min cuts where $\epsilon_\eta = 7\eta$:

Theorem 12.2 (Structure of Polygons of $\mathcal{N}_{\eta, 1}$ (Theorem 4.9 from [KKO21])). *For $\epsilon_\eta \geq 7\eta$ and any polygon of η near min cuts \mathcal{C} crossed on one side with atoms $a_0 \dots a_{m-1}$ (where a_0 is the root) the following holds:*

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

[Theorem 12.2](#) shows that polygons of cuts crossed on one side nearly look like cycles. Now, if magically it was the case that $x(E(a_i, a_{i+1 \pmod{m}})) = 1$, and $x(\delta(a_i)) = 2$ for $1 \leq i \leq m - 1$, then with probability 1 (see [Corollary 2.29](#)), we would have $E(a_i, a_{i+1})_T = 1$ and we would be able to claim that:

- (i) any cut in \mathcal{C} which does not include either a_1 or a_{m-1} (and is therefore not in \mathcal{H}) is even in the tree with probability 1;
- (ii) The cuts in \mathcal{C} that contain a_1 but not a_{m-1} , i.e., the so-called "leftmost cuts" (also not represented in \mathcal{H}) are even precisely when $E(a_0, a_1)_T$ is odd and
- (iii) the cuts in \mathcal{C} that contain a_{m-1} but not a_1 i.e., the "rightmost cuts", are even when $E(a_0, a_{m-1})_T$ is odd.

[Theorem 12.2](#) can be used to show that this approximation is correct up to $O(\eta)$. In other words, we augment s_e^* as needed on each edge between adjacent non-root atoms in each connected component \mathcal{C} , at the cost of increasing $\mathbb{E}[s_e^*]$ by an additional (again negligible) $O(\eta\beta x_e)$. This allows us to pretend our magical thinking is correct. Thus, all of the η near min cuts in the polygon that are *not* represented in the hierarchy are satisfied so long as the outer polygon cut is *happy*, that is, $E(a_0, a_1)_T = E(a_0, a_{m-1})_T = 1$ and $E(a_0, \{a_2, \dots, a_{m-2}\})_T = 0$.

12.1 Constructing the slack vector s

Our main remaining task is to explain how to use the hierarchy \mathcal{H} to choose a slack vector s that has *negative* expected value, specifically, has $\mathbb{E}[s_e] = -\Omega(\beta x_e)$ for each edge, while ensuring that all O -Join constraints coming from \mathcal{H} are satisfied.

As mentioned in [Section 9](#), the approach taken is to set s_e to be negative (i.e. *reduce* it) when certain special cuts e is on are even in the tree and therefore induce no O -join constraint. Roughly speaking, it works as follows: For each LP edge f , consider the lowest cut S in the hierarchy that contains both endpoints of f . We call this cut $p(f)$ (for "parent of f "). Let $\mathbf{e} = \{u, v\}$ (where u and v are children of S in \mathcal{H} ; recall u, v are subsets of vertices) be the set of all edges $f = \{u', v'\}$ such that $u' \in u$ and $v' \in v$.

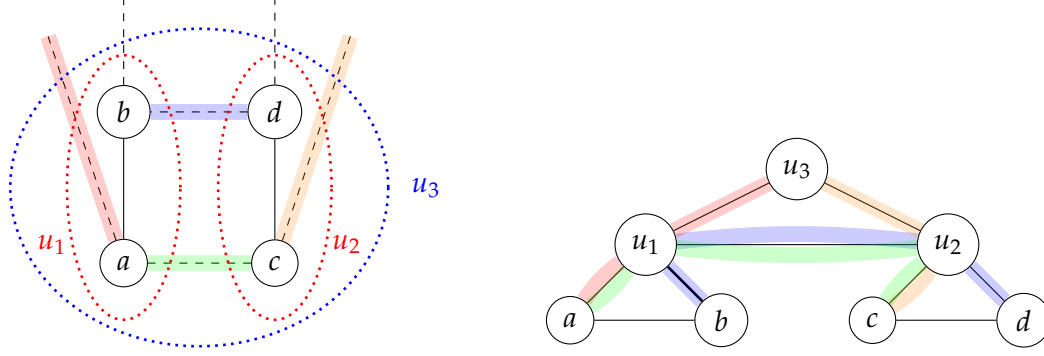


Figure 52: An example of part of a hierarchy with three "triangles". The graph on the left shows part of a feasible LP solution where dashed (and sometimes colored) edges have fraction $1/2$ and solid edges have fraction 1 . The dotted ellipses on the left show the min-cuts u_1, u_2, u_3 in the graph. (Each vertex is also a min-cut). On the right is a representation of the corresponding hierarchy. Triangle u_1 corresponds to the cut $\{a, b\}$, u_2 corresponds to $\{c, d\}$ and u_3 corresponds to $\{a, b, c, d\}$. Note that, for example, the edge $\{a, c\}$, represented in green, is in $\delta(u_1)$, $\delta(u_3)$, and inside u_3 . In this example, $\rho(\{a, b\}) = u_1$ and $\rho(\{a, c\}) = u_2$. Triangle u_1 is happy if $(\delta(a) \setminus \{a, b\})_T = (\delta(b) \setminus \{a, b\})_T = 1$.

Cuts in \mathcal{H} are separated into three types. If $S \in \mathcal{H}$ has at least three children and it is not an outer polygon cut, call it a *degree cut*. If it has exactly two children, call it a *triangle cut*. The remaining cuts, as defined above, are outer polygon cuts.

If $p(f)$ is a degree cut, then set $s_f := -0.57\beta x_f$ for all $f \in \mathbf{e}$ whenever the event that $\delta(u)_T$ and $\delta(v)_T$ are both even in the tree occurs. Call f "good" if this event occurs with constant probability. Furthermore, it is shown that every cut u with $p(u)$ a degree cut contains a $\Omega(1)$ fraction of good edges.

On the other hand, when $p(f)$ is an outer polygon cut or a "triangle" cut (see Fig. 52), set $s_f = -\beta x_f$ for all $f \in \mathbf{e}$ whenever $p(f)$ is happy (as defined above). Thus, when a polygon is happy, *all* edges e whose parent is that cut have their slack s_e reduced simultaneously. Moreover, the event that $p(f)$ is happy for a polygon cut (or triangle cut) occurs with constant probability.

However, regardless of the type of cut $p(f)$ is, setting s_f to a negative value can be problematic for the feasibility of other cuts lower down in the hierarchy that contain f . Therefore, when a cut S' lower down in the hierarchy such that $f \in \delta(S')$ is odd in the tree, the slack of *other* edges in $\delta(S')$ are increased to compensate for the reduction in s_f , (i.e., to maintain feasibility of y for the cut S').

The challenge is to do all of this in a way that still guarantees that overall $\mathbb{E}[s_e] < -\epsilon\beta x_e$, while simultaneously ensuring that for any cut $S \in \mathcal{H}$ if $\delta(S)_T$ is odd, $\sum_{e \in \delta(S)} s_e \geq 0$. Showing this is involved and requires careful probabilistic arguments that rely on the fact that the tree is sampled from a max-entropy distribution. We refer the reader to [KKO21] for the details.

12.2 Proof of Theorem 1.2 using Theorem 12.1

Theorem 12.1 (Combination of Theorem 9.8 and Theorem 11.2). *Let x^0 be a solution of LP (2) with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E . Let $\eta \leq 10^{-12}$, $\beta > 0$ and let μ be the max-entropy distribution with marginals x . Then there are two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$), such that*

i) For each edge $e \in E$, $s_e \geq -x_e\beta$ (with probability 1).

ii) For each $S \in \mathcal{N}_\eta$, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.

iii) For every edge e , $\mathbb{E}[s_e^*] \leq 125\eta\beta x_e$ and $\mathbb{E}[s_e] \leq -\frac{1}{3}x_e\epsilon_P\beta$, where ϵ_P is defined in [Theorem D.2](#).

Proof of [Theorem 1.2](#). Let x^0 be an extreme point solution of LP (2), with support E_0 and let x be x^0 restricted to E . By [Fact 2.3](#) x is in the spanning tree polytope. Let $\mu = \mu_{\lambda^*}$ be the max entropy distribution with marginals x , and let s, s^* be as defined in [Theorem 12.1](#). We will define $y : E_0 \rightarrow \mathbb{R}_{\geq 0}$ such that:

$$y_e = \begin{cases} x_e/2 + s_e + s_e^* & \text{if } e \in E \\ \infty & \text{if } e = e_0 \end{cases}$$

We will show that y is a feasible solution to (4). First, observe that for any S where $e_0 \in \delta(S)$, we have $y(\delta(S)) \geq 1$. Otherwise, we assume $u_0, v_0 \notin S$. If S is an η -near min cut and $\delta(S)_T$ is odd, then by property (ii) of [Theorem 12.1](#), we have

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

On the other hand, if S is not an η -near min cut, then

$$y(\delta(S)) \geq \left(\frac{1}{2} - \beta\right)x(\delta(S)) \geq \left(\frac{1}{2} - \beta\right) \cdot (2 + \eta) = 1 + \frac{\eta}{2} - 2\beta - \beta\eta$$

where in the first inequality we used property (i) of [Theorem 12.1](#) which gives $s_e \geq -x_e\beta$ with probability 1 along with the fact that s^* is non-negative. Therefore, choosing $\beta = \frac{\eta}{4+2\eta}$ ensures that y is a feasible O -join solution.

Finally, using $c(e_0) = 0$ and part (iii) of [Theorem 12.1](#),

$$\begin{aligned} \mathbb{E}[c(y)] &= c(x)/2 + \mathbb{E}[c(s)] + \mathbb{E}[c(s^*)] \\ &\leq c(x)/2 - \epsilon_P\beta \cdot \frac{1}{3}c(x) + 125\eta \cdot \beta \cdot c(x) \\ &\leq \left(1/2 - \frac{1}{6}\epsilon_P\beta\right) \cdot c(x) \end{aligned}$$

choosing η such that

$$125\eta = \frac{1}{6}\epsilon_P \tag{62}$$

Now, we are ready to bound the approximation factor of our algorithm. First, since x^0 is an extreme point solution of (2), $\min_{e \in E_0} x_e^0 \geq \frac{1}{n!}$. So, by [Theorem 2.6](#), in polynomial time³⁹ we can find $\lambda : E \rightarrow \mathbb{R}_{\geq 0}$ such that for any $e \in E$, $\mathbb{P}_{\mu_\lambda}[e] \leq x_e(1 + \delta)$ for some δ that we fix later. It follows that

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

³⁹Since the claim that the integrality gap is bounded below $3/2$ does not depend on the running time, it may appear that this step is unnecessary. However, we need to discuss the running time here because we are giving a stronger result that the max entropy algorithm returns a solution of expected cost at most $(3/2 - \epsilon)c(x)$ in polynomial time.

By stability of maximum entropy distributions (see [SV19, Thm 4] and references therein), we have that $\|\mu - \mu_\lambda\|_1 \leq O(n^4\delta) =: q$. Therefore, for some $\delta \ll n^{-4}$ we get $\|\mu - \mu_\lambda\|_1 = q \leq \frac{\epsilon_P \eta}{100}$. That means that

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

where we used that for any spanning tree the cost of the minimum cost matching on odd degree vertices is at most $c(x)/2$. Finally, since $\mathbb{E}_{T \sim \mu_\lambda} [c(T)] \leq c(x)(1 + \delta)$, $\epsilon_P = 3.12 \cdot 10^{-16}$, and $\eta = 4.16 \cdot 10^{-19}$ (from (62)) and $\beta = \eta/(4 + 2\eta)$, we get a $3/2 - 10^{-36}$ approximation algorithm (compared to $c(x)$). \square

13 Derandomizing the Max Entropy Algorithm

As we have discussed, the first result on the general case [KKO21] had two shortcomings. First, it did not show that the integrality gap of the subtour elimination polytope is bounded below $\frac{3}{2}$. Second, it was randomized, and the analysis in that work was by nature “non-constructive” in the sense that it used the optimal solution; thus it was not clear how to derandomize it using the method of conditional expectation. Other methods of derandomization seem at the moment out of reach and may require algorithmic breakthroughs. As we mentioned, the followup work [KKO22] remedied the first shortcoming by showing an improved integrality gap using the polygon representation. While it did not address the question of derandomization, a byproduct of that work is an analysis of the max entropy algorithm which is in principle polynomially-time computable as it avoids looking at the optimal solution. The purpose of this section is to show that this analysis can indeed be done in polynomial-time, from which the following can be deduced:

Theorem 13.1. *Let x be a solution to LP (2) for a TSP instance. For some absolute constant $\epsilon > 10^{-36}$, there is a deterministic algorithm (in particular, a derandomized version of max entropy) which outputs a TSP tour with cost at most $\frac{3}{2} - \epsilon$ times the cost of x .*

Thus, this section in some sense completes the exploratory program concerning whether the max entropy algorithm for TSP beats $3/2$ (initiated by [OSS11] in 2011), as now the above two weaknesses of [KKO21] have been addressed. Of course, much work remains in determining the true approximation factor of the algorithm; in this regard we are only at the tip of the iceberg.

As mentioned, this implies a $\frac{3}{2} - \epsilon$ approximation for 2-ECSM, and using the recent exciting work of Traub, Vygen, and Zenklusen reducing path TSP to TSP [TVZ20] this theorem also implies that there is a deterministic $\frac{3}{2} - \epsilon$ approximation algorithm for path TSP.

Now that we have completed the proof of the main result of the thesis, the remaining sections (including this one) will mostly stand alone from previous ones without significant integration.

13.1 High level proof overview

The high level strategy for derandomizing the max entropy algorithm is to use the method of conditional expectation on an objective function given by the analysis in [KKO22].

As discussed, the max entropy algorithm, similar to Christofides’ algorithm, first selects a spanning tree and then adds a minimum cost matching on the odd vertices of the tree. While Christofides selects a minimum cost spanning tree, here the spanning tree is sampled from a distribution. In particular, after solving the natural LP relaxation for the problem to obtain a fractional solution x , a tree is sampled from the distribution μ which has maximal entropy subject to the constraint $\mathbb{P}_{T \sim \mu} [e \in T] = x_e$ for all $e \in E$ (with possibly some exponentially small error in these constraints). [KKO21; KKO22] construct a so-called “slack” vector which is used to show the expected cost of the matching (over the randomness of the trees) is at most $\frac{1}{2} - \epsilon$ times the cost of an optimal solution to the LP. Given a solution x to LP (2) these works imply that there is a random vector m as a function of the tree $T \sim \mu$ such that:

- (1) The cost of the minimum cost matching on the odd vertices of tree T is at most $c(m)$ (with probability 1), and
- (2) $\mathbb{E}_{T \sim \mu} [c(m)] \leq (\frac{1}{2} - \epsilon)c(x)$.

Let $\mathcal{C} = \mathbb{E}_{T \sim \mu} [c(T) + c(m)]$. This will be the objective function to which we will apply the method of conditional expectation. Since the expected cost of the tree T is $c(x)$, as $\mathbb{P}_{T \sim \mu} [e \in T] = x_e$, by (2) \mathcal{C} is at most $(\frac{3}{2} - \epsilon)c(x)$. Since by (1) for a given tree T , $c(T) + c(m)$ is an upper bound on the cost of the output of the algorithm (with probability 1), this shows that the expected cost of the algorithm is bounded strictly below $3/2$.

Ideally, one would like μ to have polynomial sized support. Then one could simply check the cost of the output of the algorithm on every tree in the support, and the above would guarantee that some tree gives a better-than- $3/2$ approximation. However, the max entropy distribution can have exponential sized support, and it's not clear how to find a similarly behaved distribution with polynomial sized support.

Instead, let $\mathcal{T}_{\text{partial}}$ be the family of all **partial** settings of the edges of the graph to 0 or 1 where the edges set to 1 are acyclic. For $\text{Set} = \{X_{e_1}, \dots, X_{e_i}\} \in \mathcal{T}_{\text{partial}}$, and $1 \leq j \leq i$, we use X_{e_j} to indicate whether e_j is set to 1 or 0.

The method of conditional expectations is then used as follows: Process the edges in an arbitrary order e_1, \dots, e_m and for each edge e_i :

- (1) Assume we inductively have chosen a valid assignment $\text{Set} \in \mathcal{T}_{\text{partial}}$ to edges e_1, \dots, e_{i-1} .
- (2) Let $\text{Set}^+ = \text{Set} \cup \{X_{e_i} = 1\}$. Compute $\mathcal{C}^+ = \mathbb{E}_{T \sim \mu} [c(T) + c(m) \mid \text{Set}^+]$. Similarly, let $\text{Set}^- = \text{Set} \cup \{X_{e_i} = 0\}$ and compute $\mathcal{C}^- = \mathbb{E}_{T \sim \mu} [c(T) + c(m) \mid \text{Set}^-]$.
- (3) Let $\text{Set} \leftarrow \text{Set}^+$ or $\text{Set} \leftarrow \text{Set}^-$ depending on which quantity is smaller.

After a tree is obtained, add the minimum cost matching on the odd vertices of T . The resulting algorithm is shown in [Algorithm 4](#) (see [Algorithm 3](#) for its instantiation in a simple case).

As $\mathcal{C} \leq (\frac{3}{2} - \epsilon)c(x)$, this algorithm succeeds with probability 1. We only need to show it can be made to run in polynomial time. Since we can compute the expected cost of the tree conditioned on Set using linearity of expectation and the matrix tree theorem ([Theorem 13.2](#)), it remains to show that $\mathbb{E}_{T \sim \mu} [c(m) \mid \text{Set}]$ can be computed deterministically and efficiently for any $\text{Set} \in \mathcal{T}_{\text{partial}}$.

Key Contributions. The key contribution of this section is to show how to do this computation efficiently, which is based on two observations:

- (1) The first is that the vector m (whose cost upper bounds the cost of the minimum cost matching on the odd vertices of the tree) can be written as the (weighted) sum of indicators of events that depend on the sampled tree T , and each of these events happens only when a constant number of (not necessarily disjoint) sets of edges have certain parities or certain sizes.
- (2) The second is that the probability of any such event can be deterministically computed in polynomial time by evaluating the generating polynomial of all spanning trees at certain points in \mathbb{C}^E , see [Lemma 13.12](#).

Structure of the section. In [Section 13.2](#) we review the matrix tree theorem (in a slightly more compact form compared to as stated in [Theorem 13.2](#)) and show as a warmup how to compute the probability two (not necessarily disjoint) sets of edges both have an even number of edges in the sampled tree. In [Section 13.4](#), we then give a complete description and proof of a deterministic algorithm for the special “degree cut” case of TSP. [Section 13.4](#) is self-contained and thus directed

towards readers looking for more high-level intuition about the derandomization. In [Section 13.5](#) we show (2) from above and give the deterministic algorithm in the general case. The remainder then involves proving (1) for the general definition of m from [\[KKO21; KKO22\]](#).

13.2 Computing probabilities

The deterministic algorithm depends on the computation of various probabilities and conditional expectations. In this section (and additionally later in [Section 13.5](#)), we show to do these calculations efficiently.

13.2.1 Notation

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

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

13.2.2 Matrix tree theorem

Let $G = (V, E)$ with $|V| = n$. For $e = (u, v)$ we let $L_e = (\mathbf{1}_u - \mathbf{1}_v)(\mathbf{1}_u - \mathbf{1}_v)^T$ be the Laplacian of e . Recall Kirchhoff's matrix tree theorem:

Theorem 13.2 (Matrix tree theorem). *For a graph $G = (V, E)$ let $g_{\mathcal{T}} \in \mathbb{R}[z_{e_1}, \dots, z_{e_m}] = \sum_{T \in \mathcal{T}} z^T$ be the generating polynomial of the spanning trees of G .*

Then, we have

$$g_{\mathcal{T}}(\{z_e\}_{e \in E}) = \frac{1}{n} \det\left(\sum_{e \in E} z_e L_e + \mathbf{1}\mathbf{1}^T/n\right).$$

Given a vector $\lambda \in \mathbb{R}^{|E|}$ and a set $S \subseteq E$, let $\lambda^S := \prod_{i \in S} \lambda_i$. Recall that the λ -uniform distribution μ_λ is the probability distribution over spanning trees where the probability of every tree T is λ^T . Then the generating polynomial of μ_λ is

$$g_{\mu_\lambda}(z) = \sum_{T \in \mathcal{T}} \lambda^T z^T = g_{\mathcal{T}}(\{\lambda_e z_e\}_{e \in E}) = \frac{1}{n} \det\left(\sum_{e \in E} z_e \lambda_e L_e + \mathbf{1}\mathbf{1}^T/n\right)$$

and can be evaluated at any $z \in \mathbb{C}^E$ efficiently using a determinant computation.

Thus we can compute $\mathbb{P}_{T \sim \mu}[e \in T]$ by computing the sum of the probabilities of trees in the graph $G/\{e\}$, i.e. the graph with e contracted, as follows:

$$\mathbb{P}_{T \sim \mu}[e \in T] = 1 - \mathbb{P}_{T \sim \mu}[e \notin T] = 1 - \sum_{T \in \mathcal{T}: e \notin T} \lambda^T$$

where to compute the sum in the RHS we evaluate g_{μ_λ} at $z_e = 0, z_f = 1$ for all $f \neq e$. Thus,

Lemma 13.3. *Given a λ -uniform distribution μ_λ over spanning trees, for every edge e , we can compute $\mathbb{P}_{T \sim \mu_\lambda}[e \in T]$ in polynomial time.*

Given some $Set \in \mathcal{T}_{partial}$, we contract each edge e with $X_e = 1$ in Set and delete each edge e with $X_e = 0$ in Set . Let G' be the resulting graph with n' vertices, with corresponding $\lambda'_e \propto \lambda_e$ for all $e \in G'$ normalized such that $\sum_{T' \in G'} \lambda'^T = 1$.

Remark 13.4. A vector $\lambda \in \mathbb{R}^{|E|}$ is normalized by setting $\lambda'_e = \lambda_e / (\sum_T \lambda^T)^{1/n-1}$ i.e., $\lambda'_e = \lambda_e / g_{\mathcal{T}}(\{\lambda_e\}_{e \in E})^{1/n-1}$. Thus at the cost of another application of the matrix-tree theorem, we assume without loss of generality that we are always dealing with λ values that are normalized.

Putting the previous facts together, we obtain

Lemma 13.5. *Given a λ -uniform distribution μ_λ and some $Set \in \mathcal{T}_{partial}$, we can compute a vector λ' such that $\mu_{\lambda'} = \mu_{\lambda|Set}$.*

13.3 Computing parities in a simple case

Lemma 13.6. *Let $A, B \subseteq E$ and μ_λ be a λ -uniform distribution over spanning trees. Then, we can compute $\mathbb{P}_{T \sim \mu_\lambda} [A_T, B_T \text{ even}]$ in polynomial time.*

Proof. First observe that

$$\mathbb{I} \{A_T, B_T \text{ even}\} = \frac{1}{4} (1 + (-1)^{A_T} + (-1)^{B_T} + (-1)^{((A \setminus B) \cup (B \setminus A))_T})$$

One can easily check that if A_T and B_T are even, this is 1, and otherwise it is 0.

To compute $\mathbb{P}_{T \sim \mu_\lambda} [A \text{ and } B \text{ even in } T]$ it is enough to compute the expected value of this indicator. By linearity of expectation it is therefore enough to compute the expectation of $(-1)^{F_T}$ for any set $F \subseteq E$. We can do this using [Theorem 13.2](#). Setting $z_e^F = -1$ if $e \in F$ and $z_e^F = +1$ otherwise, we exactly have:

$$g_{\mu_\lambda}(z^F) = \sum_{T \in \mathcal{T}} (-1)^{F_T} \lambda^T = \mathbb{E}_{T \sim \mu_\lambda} \left[(-1)^{F_T} \right].$$

The lemma follows. □

Remark 13.7. We can use the same approach to compute $\mathbb{P}_{T \sim \mu_\lambda} [A_T \text{ odd}, B_T \text{ even}]$ or the probability that both are odd. All we need to do is to multiply $(-1)^{A_T}$ with a -1 if A_T needs to be odd (and similarly for B_T), and $(-1)^{((A \setminus B) \cup (B \setminus A))_T}$ with a -1 if we are looking for different parities in A_T, B_T .

Given some $Set \in \mathcal{T}_{partial}$, by [Lemma 13.5](#) we can compute $\mu_{\lambda'} = \mu_{\lambda|Set}$. Applying the above lemma to $\mu_{\lambda'}$, it follows (after appropriately updating the parities to account for edges set to 1 in Set):

Corollary 13.8. *Let $A, B \subseteq E$. We can compute $\mathbb{P}_{T \sim \mu} [A \text{ and } B \text{ even in } T \mid Set]$ in polynomial time.*

13.4 A deterministic algorithm in the degree cut case

As a warmup, in this section we show how to implement the deterministic algorithm for the so-called “degree cut case,” i.e., when for every set of vertices S with $2 \leq |S| \leq n - 2$ we have $x(\delta(S)) \geq 2 + \eta$ for some absolute constant $\eta > 0$. See [Algorithm 3](#).

Algorithm 3 A Deterministic Approximation Algorithm for Metric TSP in the Degree Cut Case

- 1: Given a solution x^0 of the LP (2), with an edge e_0 with $x_{e_0} = 1$.
 - 2: Let G be the support graph of x .
 - 3: Find a vector $\lambda : E \rightarrow \mathbb{R}_{\geq 0}$ such that for any $e \in E$, $\mathbb{P}_{T \sim \mu_\lambda} [e \in T] = x_e(1 \pm 2^{-n})$ (see ??).
 - 4: Initialize $Set := \emptyset$
 - 5: **while** there exists $e \neq e_0$ not set in Set **do**
 - 6: Let $Set^+ := Set \cup \{X_e = 1\}$ and let $Set^- := Set \cup \{X_e = 0\}$;
 - 7: **if** $\mathbb{E}_{T \sim \mu_\lambda} [c(T) + c(m) \mid Set^+] \leq \mathbb{E}_{T \sim \mu_\lambda} [c(T) + c(m) \mid Set^-]$ (m from Definition 13.9) **then**
 - 8: $Set := Set^+$;
 - 9: **else**
 - 10: $Set := Set^-$;
 - 11: **end if**
 - 12: **end while**
 - 13: Return $T = \{e : X_e = 1 \text{ in } Set\}$ together with min cost matching on odd degree vertices of T .
-

Construction of the matching vector. We will use the vector $m : \mathcal{T} \rightarrow \mathbb{R}^{|E|}$ for the degree cut case exactly as defined in Section 3.1.3. It will ensure that for a tree T , m is in the $O(T)$ -Join polyhedron where $O(T)$ is the set of odd vertices of T (we emphasize that m is a function of T). Therefore, $c(m)$ is an upper bound on the cost of the minimum cost matching on the odd vertices of T as desired.

Let $p = 2 \cdot 10^{-10}$ (note that we have not optimized this constant and in the degree cut case it can be greatly improved). Per usual, we say that an edge $e = (u, v)$ is **good** if $\mathbb{P}_{T \sim \mu} [u, v \text{ both even in } T] \geq p$, where we say a vertex v is even in a tree T if $\delta(v)_T$ is even. As before, the vector m will consist of the convex combination of two feasible points in the $O(T)$ -Join polyhedron, g and b (where g is for “good” edges and b is for “bad” edges).

For a tree T and an edge $e = (u, v)$ we let:

$$g_e = \begin{cases} \frac{1}{2+\eta} x_e & \text{If } u \text{ and } v \text{ are both even in } T \\ \frac{1}{2} x_e & \text{Otherwise} \end{cases} \quad b_e = \begin{cases} \frac{1+\eta}{2+\eta} x_e & \text{If } e \text{ is good} \\ \frac{1}{2+\eta} x_e & \text{If } e \text{ is bad} \end{cases}$$

Note b does not depend on T but g does. We proved in Section 3.1.3 that $g, b \in P_{OJ}$.

Definition 13.9 (Matching vector m in the degree cut case). *Let $m = \alpha b + (1 - \alpha)g$, for some $0 < \alpha < 1$ we choose shortly. Since b and g are both in the $O(T)$ -Join polyhedron, so is m .*

Setting $\alpha = \frac{p}{2+p}$, as in Section 3.1.3,

$$\mathbb{E}[m_e] \leq \left(\frac{p/(2+p)}{2+\eta} + \frac{1-p/(2+p)}{2} \right) x_e < \left(\frac{1}{2} - \frac{p\eta}{9} \right) x_e$$

for all edges e . Since η, p are absolute constants, this is at most $(\frac{1}{2} - \epsilon)x_e$ for some absolute constant $\epsilon > 0$. Therefore the randomized algorithm has expected cost at most $(\frac{3}{2} - \epsilon)c(x)$, which is enough to prove that Algorithm 3 deterministically finds a tree plus a matching whose cost is at most $(\frac{3}{2} - \epsilon)c(x)$. Thus the only remaining question is the computational complexity of Algorithm 3, which we address now.

Computing $\mathbb{E}[c(T) + c(m) \mid \text{Set}]$. Now that we have explained the construction of m , we observe that there is a simple deterministic algorithm to compute $\mathbb{E}[c(T) + c(m) \mid \text{Set}]$ in polynomial time.

First, compute $\mathbb{E}[c(T) \mid \text{Set}]$. By linearity of expectation it is enough to compute $\mathbb{P}[e \in T \mid \text{Set}]$ for all $e \in E$. To do this, we first apply [Lemma 13.5](#) to find λ' such that $\mu_{\lambda'} = \mu_{\lambda \mid \text{Set}}$ and then apply [Lemma 13.3](#).

Now to compute $\mathbb{E}[c(m) \mid \text{Set}]$, it suffices to compute $\mathbb{E}[m_e \mid \text{Set}]$ for any $\text{Set} \in \mathcal{T}_{\text{partial}}$, $\mathbb{P}[e \in T \mid \text{Set}]$ and any $e = (u, v)$. Given the definition of m , the only event depending on the tree is the event $\mathbb{P}[u, v \text{ even} \mid \text{Set}]$. This can be computed with [Corollary 13.8](#).

13.5 A deterministic algorithm in the general case

Algorithm 4 A Deterministic Approximation Algorithm for Metric TSP

- 1: Given a solution x^0 of the LP (2), with an edge e_0 with $x_{e_0} = 1$.
 - 2: Let G be the support graph of x .
 - 3: Find a vector $\lambda : E \rightarrow \mathbb{R}_{\geq 0}$ such that for any $e \in E$, $\mathbb{P}_{T \sim \mu_\lambda}[e \in T] = x_e(1 \pm 2^{-n})$
 - 4: Perform Preprocessing Steps 1, 2, 3, 4, 5, and 6
 - 5: Initialize $\text{Set} := \emptyset$.
 - 6: **while** there exists $e \neq e_0$ not set in Set **do**
 - 7: Let $\text{Set}^+ := \text{Set} \cup \{X_e = 1\}$ and let $\text{Set}^- := \text{Set} \cup \{X_e = 0\}$;
 - 8: Compute $S^+ = \mathbb{E}_{T \sim \mu_\lambda}[c(T) \mid \text{Set}^+] + \sum_{e \in E} \mathbb{E}_{c(s^*)}(e, \text{Set}^+) + \mathbb{E}_{c(s)}(e, \text{Set}^+)$.
 - 9: Compute $S^- = \mathbb{E}_{T \sim \mu_\lambda}[c(T) \mid \text{Set}^-] + \sum_{e \in E} \mathbb{E}_{c(s^*)}(e, \text{Set}^-) + \mathbb{E}_{c(s)}(e, \text{Set}^-)$.
 - 10: If $S^+ \leq S^-$, let $\text{Set} := \text{Set}^+$. Otherwise let $\text{Set} := \text{Set}^-$.
 - 11: **end while**
 - 12: Return $T = \{e : X_e = 1 \text{ in Set}\}$ together with min cost matching on odd degree vertices of T .
-

The matching vector m in the general case can be written as $s + s^* + \frac{1}{2}x$ where s, s^* are functions of the tree $T \sim \mu_\lambda$ and some independent Bernoullis \mathcal{B} . Roughly speaking, the (slack) vector $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ takes care of matching constraints for near minimum cuts that are crossed and the (slack) vector $s : E \rightarrow \mathbb{R}$ takes care of the constraints corresponding to cuts which are not crossed. Most importantly, the guarantee is that for a fixed tree T the expectation of $c(s) + c(s^*) + \frac{1}{2}c(x)$ over the Bernoullis is at least $c(M)$ where M is the minimum cost matching on the odd vertices of T . Furthermore, $\mathbb{E}[c(s) + c(s^*)] \leq -\epsilon c(x)$ which is the necessary bound to begin applying the method of conditional expectation in [Algorithm 4](#).

Remark 13.10. The definitions of s and s^* , the proof that $\mathbb{E}[c(s) + c(s^*)] \leq -\epsilon c(x)$, and the proof that $x/2 + \mathbb{E}[s + s^* \mid T]$ is in the $O(T)$ -join polyhedron come from [\[KKO21; KKO22\]](#). Here, we will review how to construct the random slack vectors s, s^* for a given spanning tree T and then explain how to efficiently compute $\mathbb{E}[c(s) + c(s^*) \mid \text{Set}]$ deterministically for any $\text{Set} \in \mathcal{T}_{\text{partial}}$.

Unfortunately, a reader who has not read the previous sections may not be able to understand the motivation behind the details of the construction of s, s^* . However, [Section 14](#) and [Section 15](#) are self-contained in the sense that a reader should be able to verify that $\mathbb{E}[c(s) + c(s^*) \mid \text{Set}]$ can be computed efficiently and deterministically.

Our theorem boils down to showing the following two lemmas:

Lemma 13.11. *For any $\text{Set} \in \mathcal{T}_{\text{partial}}$, there is a polynomial time deterministic algorithm that computes:*

- (1) $\mathbb{E}_{T \sim \mu_\lambda} [c(s^*) \mid \text{Set}]$ (shown in $\mathbb{E}_{c(s^*)}(e, \text{Set})$)
- (2) $\mathbb{E}_{T \sim \mu_\lambda} [c(s) \mid \text{Set}]$ (shown in $\mathbb{E}_{c(s)}(e, \text{Set})$)

The crux of proving the above lemma is to show that for a given edge e and any Set , each of $\mathbb{E}[s_e^* \mid \text{Set}]$ and $\mathbb{E}[s_e \mid \text{Set}]$ can be written as the (weighted) sum of indicators of events that depend on the sampled tree T , and each of these events happens only when a constant number of (not necessarily disjoint) sets of edges have certain parities or certain sizes. Technically speaking, these weighted sums are non-trivial for some of the events defined in [KKO21; KKO22]. Given that, the following is enough to prove Lemma 13.11, as it gives a deterministic algorithm to compute the probability that a collection of (not necessarily disjoint) sets of edges have certain parities or certain sizes.

(1) of Lemma 13.11 is proved in Section 14, and (2) in Section 15. The algorithm for each part requires a series of preprocessing steps and function definitions that we have marked with gray boxes. In each section, the final procedure to calculate the expected cost of the slack vector is given in a yellow box at the end of the corresponding section.

Lemma 13.12. *Given a probability distribution $\mu : 2^{[n]} \rightarrow \mathbb{R}_{\geq 0}$ and an oracle O that can evaluate $g_\mu(z_1, \dots, z_n)$ at any $z_1, \dots, z_n \in \mathbb{C}$. Let E_1, \dots, E_k be a collection of (not necessarily disjoint) subsets of $[n]$ and $(\sigma_1, \dots, \sigma_k) \in \mathbb{F}_{m_1} \times \dots \times \mathbb{F}_{m_k}$. Then, we can compute,*

$$\mathbb{P}_{T \sim \mu} [(E_i)_T = \sigma_i \pmod{m_i}, \forall 1 \leq i \leq k].$$

in $N := m_1 \dots m_k$ -many calls to the oracle.⁴⁰

Proof. For each of the sets E_i , define a variable x_i , and substitute $\prod_j x_j^{\mathbb{I}\{e \in E_j\}}$ for z_e into the polynomial g_μ and call the resulting polynomial g . Then

$$g(x_1, \dots, x_k) = \sum_{S \in \text{supp}(\mu)} \mathbb{P}[S] \prod_{i=1}^k x_i^{(E_i)_S}$$

Where recall $(E_i)_S = |E_i \cap S|$. Now, let $\omega_i := e^{\frac{2\pi\sqrt{-1}}{m_i}}$. We claim that

$$\begin{aligned} & \frac{1}{m_1 \dots m_k} \sum_{(e_1, \dots, e_k) \in \mathbb{F}_{r_1} \times \dots \times \mathbb{F}_{r_k}} \prod_{i=1}^k \omega_i^{-e_i \sigma_i} g(\omega_1^{e_1}, \dots, \omega_k^{e_k}) \\ &= \mathbb{P}_{S \sim \mu} [(E_i)_S \equiv \sigma_i \pmod{m_i}, \forall 1 \leq i \leq k] \end{aligned}$$

So the algorithm only needs to call the oracle N many times to compute the sum in the LHS.

⁴⁰Note that since we are dealing with irrational numbers, we will not be able to compute this probability exactly. However by doing all calculations with $\text{poly}(n, N)$ bits of precision we can ensure our estimate has exponentially small error which will suffice to get the bounds we need later.

To see this identity, notice that we can write the LHS as

$$\begin{aligned}
& \frac{1}{m_1 \cdots m_k} \sum_{(e_1, \dots, e_k) \in \mathbb{F}_{m_1} \times \cdots \times \mathbb{F}_{m_k}} \sum_{S \in \text{supp}(\mu)} \mathbb{P}[S] \prod_{i=1}^k \omega_i^{-e_i \sigma_i + e_i (E_i)_S} \\
&= \sum_{S \in \text{supp}(\mu)} \mathbb{P}[S] \prod_{i=1}^k \left(\frac{1}{m_i} \sum_{e_i \in \mathbb{F}_{m_i}} \omega_i^{((E_i)_S - \sigma_i) e_i} \right) \\
&= \sum_{S \in \text{supp}(\mu)} \mathbb{P}[S] \prod_{i=1}^k \mathbb{I}\{(E_i)_S - \sigma_i \equiv 0 \pmod{\sigma_i}\}
\end{aligned}$$

where the last equality uses that ω_i is the m_i 'th root of unity. The RHS is exactly equal to the probability that $(E_i)_S \equiv \sigma_i \pmod{m_i}$ for all i . \square

Remark 13.13. When we apply this lemma in this paper, we will always let k be a constant and $m_i \leq |V|$ for all i . Thus, it will always use a polynomial number of calls to an oracle evaluating the generating polynomial of a spanning tree distribution μ_λ . By [Theorem 13.2](#), for any $z \in \mathbb{C}^{|E|}$:

$$g_{\mu_\lambda}(\{z_e\}_{e \in E}) = \frac{1}{n} \det\left(\sum_{e \in E} \lambda_e z_e L_e + 11^T/n\right),$$

which can be computed in polynomial time.

Corollary 13.14. *Let μ_λ be a λ -uniform spanning tree distribution and let $\text{Set} \in \mathcal{T}_{\text{partial}}$. Then, let E_1, \dots, E_k be a collection of (not necessarily disjoint) subsets of $[n]$ and $(\sigma_1, \dots, \sigma_k) \in \mathbb{F}_{m_1} \times \cdots \times \mathbb{F}_{m_k}$. Then, we can compute,*

$$\mathbb{P}_{T \sim \mu_\lambda} [(E_i)_T = \sigma_i \pmod{m_i}, \forall 1 \leq i \leq k \mid \text{Set}].$$

in $N := m_1 \dots m_k$ -many calls to the oracle.

Proof. Construct a new graph G' by contracting all edges with $X_e = 1$ in Set and deleting all edges with $X_e = 0$. We then update all σ_i by subtracting the number of edges that are set to 1 in E_i by Set . Then we apply [Lemma 13.12](#) to the λ -uniform spanning tree distribution over G' with the updated $\vec{\sigma}$ and the same \vec{m} . \square

14 Computation for $c(s^*)$

We interleave the definitions of s and s^* with our method of computing the expected value of these vectors. While the definitions of s, s^* are essentially copied from the above, in some places we modify the notation and make the construction slightly more algorithmic to improve the presentation. To differentiate these two we put computations in boxes.

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

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

Definition 14.1 (Near Min Cut). *For $G = (V, E, x)$, we say a cut $S \subseteq V$ is an η -near min cut if $x(\delta(S)) < 2 + \eta$.⁴¹*

⁴¹Note this differs slightly from the notation in [[Ben95](#); [BG08](#)] in which an η near min cut is said to be within a $1 + \eta$ factor of the edge connectivity of the graph.

14.1 Polygon representation preprocessing

We recall some of the key definitions for constructing the vector s^* . We refer the reader to [Section 2.9](#) and [Section 5](#) for additional information about polygons.

Definition 14.2 (Left, Right Crossing). *Let $S, S' \in \mathcal{C}$ such that S' crosses S . For such a pair, we say S' crosses S on the left if the leftmost (clockwise-most) outside atom of $O(S' \cup S)$ is in S' . Otherwise, we say that S' crosses S on the right. Note that by [Observation 10.1](#), $O(S), O(S')$ cross.*

Definition 14.3 (Crossed on one, both sides). *We say a cut S is crossed on both sides if it is crossed by a cut (in \mathcal{C}) on the left and a cut (in \mathcal{C}) on the right and we say S is crossed on one side if it is crossed only on the left or only on the right.*

Definition 14.4 (Root node). *Recall $G_{/e_0}$ is the graph with e_0 contracted. Let $r \in V(G_{/e_0})$ be the result of contracting the nodes $\{u_0, v_0\}$. We will call r the **root node**.*

Definition 14.5 ($\mathcal{N}_\eta, \mathcal{N}_{\eta,0}, \mathcal{N}_{\eta,1}, \mathcal{N}_{\eta,2}, \mathcal{N}_{\eta,\leq 1}$). *Given an LP solution x , let $\mathcal{N}_\eta \subseteq 2^{V \setminus \{r\}}$ be the set of all η -near min cuts of x where we identify each cut with the side that does not contain the root node r .*

Let $\mathcal{N}_{\eta,0} \subseteq \mathcal{N}_\eta$ be the set of cuts that are not crossed. Let $\mathcal{N}_{\eta,1} \subseteq \mathcal{N}_\eta$ be the set of cuts that are crossed on one side in their respective polygons. Let $\mathcal{N}_{\eta,2} \subseteq \mathcal{N}_\eta$ be the set of cuts which are crossed on both sides in their respective polygons. Finally let $\mathcal{N}_{\eta,\leq 1} = \mathcal{N}_{\eta,0} \cup \mathcal{N}_{\eta,1}$.

Preprocessing Step 1: Compute the polygon representations

- Find all η -near min cuts of the support graph $G_{/e_0}$, which can be done in deterministic polynomial time (for example see [\[NNI94\]](#)).
- For each connected component of cuts \mathcal{C} , compute its polygon representation P . By [\[Ben95\]](#) this can be done in deterministic polynomial time.
- Given the collection of η -near min cuts and polygons, let r be the root node and compute $\mathcal{N}_\eta, \mathcal{N}_{\eta,0}, \mathcal{N}_{\eta,1}, \mathcal{N}_{\eta,2}$, and $\mathcal{N}_{\eta,\leq 1}$ (see [Definition 14.5](#)).

14.1.1 Computation for cuts crossed on both sides

Definition 14.6 (Internal). *We say an edge is internal to a polygon P (of a connected component of cuts) if its endpoints fall into two different atoms of P , both of which are not the root atom of P .*

Note by definition each edge is internal to at most one polygon P .

We iterate through each connected component of cuts \mathcal{C} in \mathcal{N}_η with polygon P and do as follows. First, we define and compute:

Definition 14.7 (S_L, S_R). *For each cut $S \in \mathcal{C}$ which is crossed on both sides, let S_L be the near minimum cut crossing S on the left which minimizes $|O(S \cap S_L)|$. If there are multiple sets crossing S on the left with the same minimum intersection, choose the smallest one to be S_L . Similarly, let S_R be the near min cut crossing S on the right which minimizes $|O(S \cap S_R)|$, and again choose the smallest set to break ties.*

For each cut $S \in \mathcal{C}$, we define:

$$\begin{aligned} E^{\leftarrow}(S) &= E(S \cap S_L, S_L \setminus S) \\ E^{\rightarrow}(S) &= E(S \cap S_R, S_R \setminus S) \\ E^{\circ}(S) &= \delta(S) \setminus (E^{\leftarrow}(S) \cup E^{\rightarrow}(S)) \end{aligned} \tag{63}$$

In addition we define the left and right bad events for each polygon point p .

$$\begin{aligned} B^{\rightarrow}(p) &= \mathbb{1}\{|E^{\rightarrow}(L(p)) \cap T| \neq 1 \text{ or } |E^{\circ}(L(p)) \cap T| \neq 0\} \\ B^{\leftarrow}(p) &= \mathbb{1}\{|E^{\leftarrow}(R(p)) \cap T| \neq 1 \text{ or } |E^{\circ}(R(p)) \cap T| \neq 0\}. \end{aligned} \tag{64}$$

If $L(p)$ does not exist, simply assume the left bad event never occurs, and similarly if $R(p)$ does not exist assume the right bad event never occurs.

Define $L(p)^{\cap R} := L(p) \cap L(p)_R$, and let $L^*(p) \in \mathcal{C}$ be the cut crossing $L(p)^{\cap R}$ on the left that maximizes $|O(L^*(p) \cap L(p)^{\cap R})|$ (and similarly $R^*(p)$ to maximize the intersection with $O(R(p)^{\cap L})$ on the right). If $L^*(p)$ does not exist, i.e. no cut crosses $L(p)^{\cap R}$ on the left, set $L^*(p) = \emptyset$, and similarly for $R^*(p)$. We let:

$$\begin{aligned} E(B^{\rightarrow}(p)) &:= E(L(p)^{\cap R} \setminus L^*(p), L(p)_R \setminus L(p)^{\cap R}) \\ E(B^{\leftarrow}(p)) &:= E(R(p)^{\cap L} \setminus R^*(p), R(p)_L \setminus R(p)^{\cap L}) \end{aligned} \tag{65}$$

Definition 14.8 (Increase event for cuts crossed on both sides). *For each edge e internal to polygon P , we define a random variable $\mathcal{I}_e : \mathcal{T} \rightarrow \{0, 1\}$ which indicates if there exists a p_i for which $e \in E(B^{\rightarrow}(p_i))$ and $B^{\rightarrow}(p_i)$ occurs or $e \in E(B^{\leftarrow}(p_i))$ and $B^{\leftarrow}(p_i)$ occurs.*

In this way, \mathcal{I}_e has been defined for every edge internal to some polygon. For all edges e which are not internal to any polygon, we simply let $\mathcal{I}_e = 0$ for every tree.

Lemma 14.9. *We can compute*

$$\mathbb{P}[\mathcal{I}_e \mid \text{Set}]$$

in polynomial time for any edge e .

Proof. If e is not internal to any polygon, $\mathbb{P}[\mathcal{I}_e \mid \text{Set}] = 0$ and we are done. Otherwise, it is internal to some polygon P . By Lemma 5.4 in [KKO22] (also see the proof of Theorem 5.2), there are at most two indices i in this polygon P for which $e \in E(B^{\rightarrow}(p_i))$ and at most two indices i for which $e \in E(B^{\leftarrow}(p_i))$. Therefore, we are interested in at most four events $B^{\rightarrow}(p_i)$ or $B^{\leftarrow}(p_i)$.

Using Corollary 13.14 it is straightforward to compute the probability that any collection of these (at most four) events occurs. For each event i (say, some $B^{\rightarrow}(p_i)$) we use the two sets $E^{\rightarrow}(L(p_i))$ and $E^{\circ}(L(p_i))$ and set their σ values to be 1 and 0 respectively and both of their m values to be $|V|$, and return 1 minus the computed probability.

Therefore, we can compute the probability that at least one event occurs, which is sufficient to prove the lemma. \square

Let $I_e : \mathcal{T}_{\text{partial}} \rightarrow \mathbb{R}_{\geq 0}$ be the function from the above lemma which given Set returns $\mathbb{P}[\mathcal{I}_e \mid \text{Set}]$.

Preprocessing Step 2: Compute polygon edge sets

For each polygon P with connected component of cuts \mathcal{C} :

- For each cut $S \in \mathcal{C}$, compute $E^{\leftarrow}(S), E^{\rightarrow}(S), E^{\circ}(S)$ (see Eq. (63)).
- For each polygon point p in P , compute $E(B^{\rightarrow}(p))$ and $E(B^{\leftarrow}(p))$ (see Eq. (65)).

Increase-Both-Sides(e, Set)

Given an edge e and $Set \in \mathcal{T}_{partial}$ compute $\mathbb{P}[\mathcal{I}_e \mid Set]$ using Lemma 14.9.

14.2 Preprocessing for cuts crossed on one side

Now partition the cuts in $\mathcal{N}_{\eta,1}$ into connected components. For each connected component of cuts \mathcal{C} and for each cut $C \in \mathcal{N}_{\eta,0}$ that can be written as the union of two other cuts $a_1, a_2 \in \mathcal{N}_{\eta,0}$ which are not crossed, let P be its (possibly degenerate)⁴² polygon P . Similar to above, each edge is internal (see Definition 14.6) to at most one such polygon P . By Lemma A.1 from [KKO22], P has no inside atoms. Label its outside atoms a_0, \dots, a_{m-1} , in counterclockwise order, where WLOG a_0 is the root atom.

We call a_1 the leftmost atom and a_{m-1} the rightmost atom. Finally, for $1 \leq i \leq m$ let $\mathcal{E}_i(P) = E(a_{i-1}, a_{i \pmod{m}})$ be the edges between atom a_{i-1} and $a_{i \pmod{m}}$ in P .

Now we define the following:

Definition 14.10 (A, B, C -Polygon Partition). *The A, B, C -polygon partition of a polygon P is a partition of edges of $\delta(a_0)$ into sets $A = \mathcal{E}_1(P)$, $B = \mathcal{E}_m(P)$, and $C = \delta(a_0) \setminus A \setminus B$.*

Definition 14.11 (Relevant Atoms and Relevant Cuts). *Define the family of relevant atoms of \mathcal{C} to be*

$$A = \{a_i : 1 \leq i \leq m-1, x(\delta(a_i)) \leq 2 + \eta\},$$

and define the relevant cuts to be

$$\mathcal{C}_+ = \mathcal{C} \cup A.$$

Definition 14.12 (Left and Right Hierarchies). *Let \mathcal{L} (the left hierarchy) be the set of all cuts $A \in \mathcal{C}$ that are not crossed on the left. Similarly, we let \mathcal{R} be the set of cuts that are not crossed on the right. In this way \mathcal{L}, \mathcal{R} partition all cuts in \mathcal{C} .*

Given this, define $\mathcal{C}_+^{\mathcal{R}} = \mathcal{R} \cup A$ and $\mathcal{C}_+^{\mathcal{L}} = \mathcal{L} \cup A$.

Definition 14.13 ($Map(\mathcal{E}_i(P))$). *We define a mapping from cuts in $\mathcal{C}_+^{\mathcal{R}}$ to the edges $E(a_1, a_2), \dots, E(a_{m-2}, a_{m-1})$. For any $2 \leq i \leq m-1$, we map*

$$\operatorname{argmax}_{A \in \mathcal{C}_+^{\mathcal{R}}: \ell(A)=i} |A| \text{ and } \operatorname{argmax}_{A \in \mathcal{C}_+^{\mathcal{R}}: r(A)=i} |A| \quad (66)$$

to $\mathcal{E}_{i-1}(P)$, where $\ell(A)$ is the index of the leftmost atom of A and $r(A)$ is the index of the rightmost atom of A . We then compute a similar mapping for $\mathcal{C}_+^{\mathcal{L}}$. For each edge group $\mathcal{E}_i(P)$ we record the set of cuts mapped to it by these two processes as a multiset $Map(\mathcal{E}_i(P))$ (since every atom is in both $\mathcal{C}_+^{\mathcal{R}}$ and $\mathcal{C}_+^{\mathcal{L}}$, some atoms may appear twice).

⁴²In the case that $C \in \mathcal{N}_{\eta,0}$, we simply let P be the three atoms $a_0 = G \setminus C$, a_1 , and a_2 .

We now introduce the following notion:

Definition 14.14 (Happy Cut). *We say a leftmost cut $L \in \mathcal{C}$ is happy if*

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

Similarly, the leftmost atom a_1 is happy if $E(a_1, \overline{a_0 \cup a_1})_T = 1$. Define rightmost cuts in \mathcal{C} or the rightmost atom in P to be happy in a similar manner.

We now define an “unhappy” event \mathcal{U}_C for each cut in $C \in \mathcal{C}_+$ such that

$$\mathcal{U}_C := \begin{cases} \mathbb{I}\{C \text{ is not happy}\} & \text{If } C \text{ is a leftmost or rightmost cut} \\ \mathbb{I}\{C \text{ is odd}\} & \text{If } C \text{ is not a leftmost or rightmost cut} \end{cases}$$

This allows us to define an increase random variable for each edge $e \in \mathcal{E}_i(P)$ for $1 \leq i \leq m$ called $\mathcal{I}'_e : \mathcal{T} \rightarrow \{0, \frac{1}{2}, 1\}$. In particular for $e \in \mathcal{E}_i(P)$ we let:

$$\mathcal{I}'_e := \min\{1, \sum_{C \in (\text{Map}(\mathcal{E}_i(P)) \setminus A)} \mathbb{I}\{\mathcal{U}_C\} + \frac{1}{2} \sum_{C \in (\text{Map}(\mathcal{E}_i(P)) \cap A)} \mathbb{I}\{\mathcal{U}_C\}\}$$

where notice that an atom may contribute twice to the sum since $\text{Map}(\mathcal{E}_i(P))$ may be a multiset.

In this way, every edge which is internal to some polygon of $\mathcal{N}_{\eta, \leq 1}$ constructed in this section has an associated random variable \mathcal{I}'_e . For every edge which is internal to no polygon constructed in this section, we say \mathcal{I}'_e never occurs.

Lemma 14.15. *We can compute*

$$\mathbb{P}[\mathcal{I}'_e \mid \text{Set}]$$

in polynomial time for any edge e internal to a polygon P of $\mathcal{N}_{\eta, 1}$.

Proof. If e is not internal to any polygon of $\mathcal{N}_{\eta, 1}$ then $\mathbb{P}[\mathcal{I}'_e \mid \text{Set}] = 0$ and we are done. Otherwise, it is internal to some polygon P with root a_0 . Since by [Eq. \(66\)](#), $|\text{Map}(\mathcal{E}_i(P))| \leq 4$, by linearity of expectation it is enough enough to compute $\mathbb{P}[\mathcal{U}_C]$ for (at most four) cuts in $\text{Map}(\mathcal{E}_i(P))$. We use [Corollary 13.14](#). Say C is a leftmost cut (it is similar if it is a rightmost cut). Then, compute $\mathbb{P}[\mathcal{U}_C] = 1 - \mathbb{P}[C \text{ is happy}]$; so it is enough to compute $\mathbb{P}[C \text{ is happy} \mid \text{Set}]$. We use [Corollary 13.14](#) with the set of edges $E(C, \overline{C \cup a_0})$ and with corresponding σ value of 1 and m value of $|V|$. If C is not a leftmost or a rightmost cut we use [Corollary 13.14](#) with the set $\delta(C)$, σ value of 1 and $m = 2$. \square

Preprocessing Step 3: Compute polygons of $\mathcal{N}_{\eta, 1}$ and the maps

Partition the cuts in $\mathcal{N}_{\eta, 1}$ into connected components. For each connected component of cuts \mathcal{C} compute its (possibly degenerate) polygon P . Now, for each polygon P corresponding to a connected component \mathcal{C} of cuts in $\mathcal{N}_{\eta, 1}$ with atoms a_0, \dots, a_{m-1} :

- Let $\mathcal{E}_i(P)$ be the set of edges between $a_{i-1}, a_{i \bmod m}$
- Let \mathcal{C}_+ be the set of relevant cuts as defined in [Definition 14.11](#).
- For each $i \in 0, \dots, m-1$ construct the multiset $\text{Map}(\mathcal{E}_i(P))$ of cuts mapped to $\mathcal{E}_i(P)$ in \mathcal{C}_+ .

Increase-One-Side(e, Set)

Given an edge e and $Set \in \mathcal{T}_{partial}$ compute $\mathbb{P}[I'_e \mid Set]$ using [Lemma 14.15](#).

14.3 Computation of $\mathbb{E}[c(s^*) \mid Set]$

Following Theorem 6.1 of [\[KKO22\]](#) we define $s_e^* : \mathcal{T} \rightarrow \mathbb{R}_{\geq 0}$:

$$s_e^* = (1 - \gamma) \frac{2 + \eta}{1 - \epsilon_\eta} \beta x_e (\mathcal{I}_e + \mathcal{I}'_e) + \gamma 2 \beta x_e \mathcal{I}_e, \quad (67)$$

where $\gamma = \frac{15}{32} \epsilon_P$. By the above two lemmas, we can compute $\mathbb{E}[c(s^*) \mid Set]$ in polynomial time. Thus the fact that the following function can be computed efficiently is the main result of this section:

$\mathbb{E}_{c(s^*)}(e, Set)$

Given an edge e and $Set \in \mathcal{T}_{partial}$, call functions [Increase-One-Side\(\$e, Set\$ \)](#) and [Increase-Both-Sides\(\$e, Set\$ \)](#) to compute $\mathbb{P}[I'_e \mid Set]$ and $\mathbb{P}[I_e \mid Set]$ respectively. Then use (67) to compute and return $\mathbb{E}[c(s_e^*) \mid Set]$.

This concludes the proof of (1) of [Lemma 13.11](#).

15 Computation for $c(s)$

Here we will compute some parameters which are fixed throughout the course of the algorithm. We also classify edges based on the probability of some events. In all computations we use the (unconditional) measure μ_λ .

We begin by setting the constants as in [Table 4](#).

15.1 Hierarchy definition and computation

Here we recall notation from [\[KKO22\]](#). The following is key to defining the slack vector.

Definition 15.1 (Hierarchy, [\[KKO22\]](#)). *For an LP solution x^0 with support $E_0 = E \cup \{e_0\}$ where x is x^0 restricted to E , a hierarchy $\mathcal{H} \subseteq \mathcal{N}_{\epsilon_\eta}$ is a laminar family with root $V \setminus \{u_0, v_0\}$, where every cut $S \in \mathcal{H}$ is called either a “near-cycle” cut or a degree cut. In the special case that S has exactly two children we call it a triangle cut. Furthermore, every cut S is the union of its children. For any (non-root) cut $S \in \mathcal{H}$, define the parent of S , $p(S)$, to be the smallest cut $S' \in \mathcal{H}$ such that $S \subsetneq S'$.*

For a cut $S \in \mathcal{H}$, let $\mathcal{A}(S) := \{a \in \mathcal{H} : p(a) = S\}$; we will call these the atoms of S . If S is called a “near-cycle” cut, then we can order cuts in $\mathcal{A}(S)$, a_1, \dots, a_{m-1} such that

- $x(E(\bar{S}, a_1)), x(E(a_{m-1}, \bar{S})) \geq 1 - \epsilon_\eta$.
- For any $1 \leq i < m - 1$, $x(E(a_i, a_{i+1})) \geq 1 - \epsilon_\eta$.
- $\cup_{i=2}^{m-2} E(a_i, \bar{S}) \leq \epsilon_\eta$.

Name	Value	Explanation
$\epsilon_{1/2}$	0.0002	Half edge threshold
$\epsilon_{1/1}$	$\frac{\epsilon_{1/2}}{12}$	A, B, C partitioning threshold, Definition 15.5
p	$0.005\epsilon_{1/2}^2$	Min prob. of happiness for a (2-*) good edge
ϵ_M	0.00025	Marginal errors due to max flow
τ	0.571β	Top edge decrease
ϵ_P	750η	Expected decrease constant
α	$2\epsilon_\eta$	Parameter of the matching
ϵ_B	$21\epsilon_{1/2}$	Parameter of the matching
ϵ_F	1/10	Parameter of the matching
ϵ_η	7η	Definition 15.1
η	$4.16 \cdot 10^{-19}$	Near min cut constant
β	$\frac{\eta}{4+2\eta}$	Slack shift constant

Table 4: A table of all constants used in the algorithm.

We abuse notation and for an edge $e = (u, v)$ that is not a neighbor of u_0, v_0 , we write $p(e)$ to denote the smallest⁴³ cut $S' \in \mathcal{H}$ such that $u, v \in S'$.

Definition 15.2 (A, B, C near-cycle partition, left-happy, right-happy, and happy). Let \mathcal{H} be a hierarchy and let $S \in \mathcal{H}$ be a near-cycle cut with cuts in $\mathcal{A}(S)$ ordered a_1, \dots, a_{m-1} . Then let $A = x(E(\bar{S}, a_1))$, $B = x(E(a_{m-1}, \bar{S}))$, and $C = \cup_{i=2}^{m-2} E(a_i, \bar{S})$. We call the sets A, B, C the **near-cycle partition** of $\delta(S)$.

We say S is **left-happy** when A_T is odd and $C_T = 0$, **right happy** when B_T is odd and $C_T = 0$, and **happy** when A_T, B_T are odd and $C_T = 0$.

By [Definition 15.1](#), we have $x(A), x(B) \geq 1 - \epsilon_\eta$ and $x(C) \leq \epsilon_\eta$.

Now we will define a hierarchy \mathcal{H} as the cuts in \mathcal{N}_η which are not crossed, plus some extra cuts in $\mathcal{H}_{\epsilon_\eta}$.

⁴³in the sense of the number of vertices that it contains

Preprocessing Step 4: Constructing the hierarchy

Let $\mathcal{N}_{\eta, \leq 1} \subseteq 2^{V \setminus \{u_0, v_0\}}$ be the set of cuts crossed on at most one side: this was computed in [Section 14](#). Now we construct \mathcal{H} as follows. For every connected component \mathcal{C} of $\mathcal{N}_{\eta, \leq 1}$, if $|\mathcal{C}| = 1$ then add the unique cut in \mathcal{C} to the hierarchy. Otherwise, \mathcal{C} corresponds to a connected component of cuts crossed on one side u with atoms a_0, \dots, a_{m-1} (for some $m > 3$). Add a_1, \dots, a_{m-1} ^a and $\cup_{i=1}^{m-1} a_i$ to \mathcal{H} . Note that since $x(\delta(\{u_0, v_0\})) = 2$, the root of the hierarchy is always $V \setminus \{u_0, v_0\}$.

Now, partition the cuts in \mathcal{H} into degree cuts and near-cycle cuts. For a cut $S \in \mathcal{H}$, if there is a connected component of at least two cuts with union equal to S , then call S a **near-cycle cut** and compute its near-cycle A, B, C partitioning as defined in [Definition 15.2](#). If S is a cut with exactly two children X, Y in the hierarchy, then also call S a near-cycle cut^b, with A, B, C partitioning $A = E(X, \overline{X} \setminus Y)$, $B = E(Y, \overline{Y} \setminus X)$ and $C = \emptyset$. Otherwise, call S a **degree cut**.

Finally, compute the A, B, C degree partitioning for all $S \in \mathcal{H}$ as described below in [Definition 15.5](#).

^aNotice that an atom may already correspond to a connected component, in such a case we do not add it in this step.

^bThink about such a set as a *degenerate* polygon with atoms $a_1 := X, a_2 := Y, a_0 := \overline{X \cup Y}$.

Remark 15.3. Since $|\mathcal{N}_{\eta, \leq 1}|$ has polynomial size in n this can be done in polynomial time.

Also note that since every vertex has degree 2, they all appear in the hierarchy as singletons. Therefore, every set in the hierarchy is the union of its children.

15.2 Edge bundles, A, B, C degree partition, and edge classification

Definition 15.4 (Edge Bundles, Top Edges, and Bottom Edges). *For every degree cut S and every pair of atoms $u, v \in \mathcal{A}(S)$, we define a **top edge bundle** $\mathbf{f} = (u, v)$ such that*

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

Note that in the above definition, u', v' are actual vertices of G .

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

Note in this way every edge e is in a unique edge bundle \mathbf{e} . We say e is a bottom edge if its edge bundle is a bottom edge bundle and otherwise e is a top edge.

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

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

$$\begin{aligned} \delta^\uparrow(u) &:= \delta(u) \cap \delta(S), \\ \delta^\rightarrow(u) &:= \delta(u) \setminus \delta(S), \\ E^\rightarrow(S) &:= \bigcup_{v \in \mathcal{H} : \mathfrak{p}(v) = S} \delta^\rightarrow(v). \end{aligned} \tag{68}$$

Variable	Name	Event
H_e	2-2 happy	$\mathbb{P}[u, v \text{ trees}, \delta(u)_T = \delta(v)_T = 2]$
$H_{e,u}$	2-1-1 happy w.r.t u	$\mathbb{P}[u, v \text{ trees}, A_T = B_T = 1, C_T = 0, \delta(v)_T = 2]$
$H_{\{e,f\}}$	2-2-2 happy w.r.t u with $f = (u, w)$	$\mathbb{P}[u, v, w \text{ trees}, \delta(u)_T = \delta(v)_T = \delta(w)_T = 2]$

Table 5: For a top edge bundle $e = (u, v)$ where A, B, C is the degree partitioning of u , we define the following “happy” events.

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

Now we compute the so-called A, B, C degree partitioning of each cut $S \in \mathcal{H}$ for which $p(S)$ is a degree cut. It can easily be implemented in polynomial time.

Definition 15.5 (A, B, C -Degree Partitioning). For $u \in \mathcal{H}$ and $\epsilon_{1/1}$ as in Table 4, we define a partitioning of edges in $\delta(u)$: Let $a, b \subsetneq u$ be minimal cuts in the hierarchy, i.e., $a, b \in \mathcal{H}$, such that $a \neq b$ and $x(\delta(a) \cap \delta(u)), x(\delta(b) \cap \delta(u)) \geq 1 - \epsilon_{1/1}$. Note that since the hierarchy is laminar, a, b cannot cross. Let $A = \delta(a) \cap \delta(u), B = \delta(b) \cap \delta(u), C = \delta(u) \setminus A \setminus B$.

If there is no cut $a \subsetneq u$ (in the hierarchy) such that $x(\delta(a) \cap \delta(u)) \geq 1 - \epsilon_{1/1}$, we just let A, B partition $\delta(u)$ such that $x(A), x(B) \in [1 - \epsilon_{1,1}, 1 + \epsilon_\eta]$, and set $C = \emptyset$. Note that this exists WLOG because we may split any edge into an arbitrary number of parallel copies.

If there is just one minimal cut $a \subsetneq u$ (in the hierarchy) with $x(\delta(a) \cap \delta(u)) \geq 1 - \epsilon_{1/1}$, i.e., b does not exist in the above definition, then we define $A = \delta(a) \cap \delta(u)$. Let $a' \in \mathcal{H}$ be the unique child of u such that $a \subseteq a'$, i.e., a is equal to a' or a descendant of a' . Then we define B to be an arbitrary subset of $\delta(u) \setminus \delta(a')$ such that $x(B) \in [1 - \epsilon_{1,1}, 1 + \epsilon_\eta]$. Finally let $C = \delta(u) \setminus (A \cup B)$. Note $C \supseteq \delta(a') \cap \delta(u) \setminus \delta(a)$.

Note we may have to divide a single edge e between the sets A, B, C to ensure such partitions exist.

Let p be as in Table 4. For a top edge bundles $e = (u, v)$, we say e is **2-2 happy**, or H_e occurs, if u, v are trees and $\delta(u)_T = \delta(v)_T = 2$. Recall that $u, v \in \mathcal{H}$ are sets of vertices.

$p_{2-2}(e, \text{Set})$

To compute $\mathbb{P}[e \text{ 2-2 happy}]$, use Corollary 13.14 with

$$E_1 = E(u), \quad E_2 = E(v), \quad E_3 = \delta(u), \quad E_4 = \delta(v),$$

$$\vec{\sigma} = (|u| - 1, |v| - 1, 2, 2), \quad \vec{m} = (|V|, |V|, |V|, |V|),$$

Definition 15.6 (Good and bad edges). A top edge e in edge bundle f is **good** (sometimes just “good”) if $p_{2-2}(f, \emptyset) \geq p$ and **bad** otherwise. We say every bottom edge is good, and edges in $\delta(\{u_0, v_0\})$ are bad (because they do not have both of their endpoints in the hierarchy).

- **2-1-1 happy w.r.t. u** : Let A, B, C be the A, B, C degree partition of u computed in the previous section. We say e is 2-1-1 happy w.r.t. u , or $H_{e,u}$ occurs, if u, v are trees, $A_T = B_T = 1, C_T = 0$, and $\delta(v)_T = 2$.

$p_{2-1-1}(\mathbf{e}, u, \text{Set})$

Let A, B, C be the degree partition of u . To compute $\mathbb{P}[\mathbf{e}$ 2-1-1 happy w.r.t u], use [Corollary 13.14](#) with $E_1 = E(u), E_2 = E(v), E_3 = A, E_4 = B, E_5 = C, E_6 = \delta(v)$, $\vec{\sigma} = (|u| - 1, |v| - 1, 1, 1, 0, 2)$, and $\vec{m} = (|V|, |V|, |V|, |V|, |V|, |V|)$.

- **2-2-2 happy (w.r.t. common endpoint u , with partner f):** We say the edge bundles $\mathbf{e} = (u, v)$ and $\mathbf{f} = (u, w)$ (where $p(u) = p(v) = p(w)$) are 2-2-2 happy w.r.t. u , or $H_{\{\mathbf{e}, \mathbf{f}\}}$ occurs, if u, v, w are trees and $\delta(u)_T = \delta(v)_T = \delta(w)_T = 2$.

$p_{2-2-2}(\mathbf{e}, \mathbf{f}, \text{Set})$

Assume \mathbf{e}, \mathbf{f} have a common endpoint u and v, w are the other endpoints of \mathbf{e}, \mathbf{f} . To compute $\mathbb{P}[\mathbf{e}$ 2-2-2 happy w.r.t u with \mathbf{f}], use [Corollary 13.14](#) with $E_1 = E(u), E_2 = E(v), E_3 = E(w), E_4 = \delta(u), E_5 = \delta(v), E_6 = \delta(w)$, $\vec{\sigma} = (|u| - 1, |v| - 1, |w| - 1, 2, 2, 2)$, and $\vec{m} = (|V|, |V|, |V|, |V|, |V|, |V|)$.

For each edge bundle $\mathbf{e} = (u, v)$, we define its type with respect to each endpoint as follows:

$\text{type}(u)$ returns $\text{type}(\mathbf{e}, u)$ for all edge bundles $\mathbf{e} \in \delta(u)$

For every edge bundle $\mathbf{e} \in \delta(u)$:

If $p_{2-1-1}(\mathbf{e}, u, \emptyset) \geq p$, then $\text{type}(\mathbf{e}, u) = 2-1-1$.

Let A, B, C be the degree partitioning of $\delta(u)$. Let $F \subseteq \delta(u)$ be the set of edge bundles adjacent to u with $\text{type}(\mathbf{e}, u) = 2-1-1$. Let $J \subseteq \delta(u)$ be the set of edges adjacent to u such that $p_{2-2}(\mathbf{e}, \emptyset) < p$. Now, if $x(F) \leq \frac{1}{2} - \epsilon_{1/2} - \epsilon_\eta$ and $x(J) \leq \frac{1}{2} - \epsilon_{1/2}$, then, by Theorem 5.28 of [KKO21], there exists two edges $\mathbf{e} = (u, v), \mathbf{f} = (u, w)$ such that: (i) $p_{2-2-2}(\mathbf{e}, \mathbf{f}, \emptyset) \geq p$, and (ii) $x(\mathbf{e} \cap B) \leq \epsilon_{1/2}, x(\mathbf{f} \cap A) \leq \epsilon_{1/2}$. Let $\text{type}(\mathbf{e}, u) = \text{type}(\mathbf{f}, u) = 2-2-2$.

For every edge bundle $\mathbf{e} \in \delta(u)$ such that $\text{type}(\mathbf{e}, u)$ is not set:

If $p_{2-2}(\mathbf{e}, \emptyset) \geq p$, then $\text{type}(\mathbf{e}, u) = 2-2$. Otherwise $\text{type}(\mathbf{e}, u) = \text{bad}$.

Finally, for each edge bundle $\mathbf{e} = (u, v)$, we define the following $\mathcal{R}_{\mathbf{e}, u}$ event with respect to each endpoint u :

- If $\text{type}(\mathbf{e}, u) = 2-1-1$, we define an independent Bernoulli $B_{\mathbf{e}, u}$ with success probability $p/p_{2-1-1}(\mathbf{e}, u, \emptyset)$ and we define

$$\mathcal{R}_{\mathbf{e}, u} := \mathbb{I} \{H_{\mathbf{e}, u} = B_{\mathbf{e}, u} = 1\}$$

We emphasize that this reduction indicator is purely a function of a tree T and an independent Bernoulli. The same will apply to all future reduction indicators.

- If $\text{type}(\mathbf{e}, u) = 2-2-2$, there exists an edge bundle \mathbf{f} such that $\text{type}(\mathbf{f}, u) = 2-2-2$. In this case, define an independent Bernoulli $B_{\{\mathbf{e}, \mathbf{f}\}}$ with success probability $p/p_{2-2-2}(\mathbf{e}, \mathbf{f}, \emptyset)$. Define

$$\mathcal{R}_{\mathbf{e}, u} := \mathcal{R}_{\mathbf{f}, u} := \mathcal{R}_{\{\mathbf{e}, \mathbf{f}\}} := \mathbb{I} \{H_{\{\mathbf{e}, \mathbf{f}\}} = B_{\{\mathbf{e}, \mathbf{f}\}} = 1\}$$

Note we use brackets to emphasize that $\mathcal{R}_{\{e,f\}}$ is the same event as $\mathcal{R}_{\{f,e\}}$.

- If $\text{type}(\mathbf{e}, u) = 2-2$. We define an independent Bernoulli B_e with success probability $p/p_{2-2}(\mathbf{e}, \emptyset)$ and we define

$$\mathcal{R}_{e,u} := \mathbb{I}\{H_e = B_e = 1\}.$$

- Otherwise $\text{type}(\mathbf{e}, u) = \text{bad}$. Define $\mathcal{R}_{e,u} = 0$.

The following allows us to compute the expected value of $\mathcal{R}_{e,u}$ for all edge bundles and each of their endpoints conditioned on Set . Note that $\mathbb{P}[\mathcal{R}_{e,u}] = p$ for all good edges. However, $\mathbb{P}[\mathcal{R}_{e,u} \mid Set]$ can be any number between 0 and the success probability of its Bernoulli (defined above).

$\mathbb{E}_{\mathcal{R}}(\mathbf{e}, u, Set)$

Call $\text{type}(u)$ to determine $\text{type}(\mathbf{e}, u)$. If $\text{type}(\mathbf{e}, u) = 2-1-1$, return

$$p_{2-1-1}(\mathbf{e}, u, Set) \cdot \left(\frac{p}{p_{2-1-1}(\mathbf{e}, u, \emptyset)} \right)$$

Otherwise if $\text{type}(\mathbf{e}, u) = \text{type}(\mathbf{f}, u) = 2-2-2$ for some edge bundle \mathbf{f} , return

$$p_{2-2-2}(\mathbf{e}, \mathbf{f}, Set) \cdot \left(\frac{p}{p_{2-2-2}(\mathbf{e}, \mathbf{f}, \emptyset)} \right)$$

Otherwise, if $\text{type}(\mathbf{e}, u) = 2-2$, return

$$p_{2-2}(\mathbf{e}, Set) \cdot \left(\frac{p}{p_{2-2}(\mathbf{e}, \emptyset)} \right)$$

Otherwise, $\text{type}(\mathbf{e}, u) = \text{bad}$. Return 0.

15.3 Max Flow

For each near-cycle cut $S \in \mathcal{H}$ with polygon partition A, B, C , we compute parameters $\alpha_{e,f}$ for all edges $e \in A, f \in B$ as well as p_S , which we define next. Let H_S be the event:

$$H_S := A_T = B_T = 1, C_T = 0, S \text{ is a tree.}$$

Preprocessing Step 5: Max Flow

For every near-cycle cut $S \in \mathcal{H}$ with polygon partition A, B, C , do the following.

Construct and solve an instance of the max-flow, min-cut problem. Consider the following graph with vertex set $\{s, A, B, t\}$. For any edge $e \in A, f \in B$ connect e to f with a directed edge of capacity $y_{e,f} = \mathbb{P}[e, f \in T \mid H_S] = \frac{\mathbb{P}[H_S \wedge e, f \in T]}{\mathbb{P}[H_S]}$. To compute the numerator (the denominator is similar), apply [Corollary 13.14](#) to

$$\begin{aligned} E_1 &= \{e\}, & E_2 &= \{f\}, & E_3 &= \delta(S) \setminus \{e, f\}, & E_4 &= E(S) \\ \vec{\sigma} &= (1, 1, 0, |S| - 1), & \vec{m} &= (2, 2, |V|, |V|) \end{aligned} \quad (69)$$

For any $e \in E$, let $x_e := \mathbb{P}[e \in T \mid C_T = 0, S \text{ is a tree}]$. Connect s to $e \in A$ with an arc of capacity qx_e and similarly connect $f \in B$ to t with arc of capacity qx_f , where $q = \frac{0.1\zeta^2/6^2}{\mathbb{P}[A_T=B_T=1 \mid C_T=0, S \text{ is a tree}]}$ and $\zeta = 1/4000$, computed using [Corollary 13.14](#). Then compute the maximum flow of this graph. Let \mathbf{z} be the maximum flow, where $z_{e,f}$ is the flow on the edge from e to f .

Now return

$$\alpha_{e,f} := \frac{z_{e,f}}{y_{e,f}}, \quad p_S := \sum_{e \in A, f \in B} \mathbb{P}[H_S] z_{e,f} \quad (70)$$

Note that it follows by Proposition 5.6 from [\[KKO21\]](#) that $p_S \geq p$. Define an independent Bernoulli B_S with success probability p/p_S as well as independent Bernoullis $B_{e,f}$ for all $e \in A, f \in B$ with success probability $\alpha_{e,f}$. For a tree T , we define the event:

$$\mathcal{R}_S := \bigcup_{e \in A, f \in B} \mathbb{I}\{A \cap T = \{e\}, B \cap T = \{f\}, H_S = B_S = B_{e,f} = 1\} \quad (71)$$

$\mathbb{E}_{\mathcal{R}}(S, \text{Set})$

To compute $\mathbb{E}[\mathcal{R}_S \mid \text{Set}]$, note by definition:

$$\begin{aligned} \mathbb{E}[\mathcal{R}_S \mid \text{Set}] &= \sum_{f \in A, g \in B} \mathbb{P}[A \cap T = \{f\}, B \cap T = \{g\}, H_S = B_S = B_{e,f} = 1 \mid \text{Set}] \\ &= (p/p_S) \alpha_{e,f} \sum_{f \in A, g \in B} \mathbb{P}[A \cap T = \{f\}, B \cap T = \{g\}, H_S \mid \text{Set}], \end{aligned} \quad (72)$$

Compute the inner probability using [Corollary 13.14](#) similarly to (69). Return the result.

15.4 Matching

Next we compute a matching from the good edges in $E^\rightarrow(S)$ to the edges in $\delta(S)$ for every degree cut $S \in \mathcal{H}$. The output of this procedure will be values $m_{\mathbf{e},u}$ which indicate that the good edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ (where $u, v \in \mathcal{A}(S)$) is matched to a fraction $m_{\mathbf{e},u}$ of edges in $\delta^\uparrow(u)$ and a fraction $m_{\mathbf{e},v}$ of $\delta^\uparrow(v)$. In the following, ϵ_F, ϵ_B , and α are set in [Table 4](#).

Preprocessing Step 6: Matching

For every $S \in \mathcal{H}$ which is a degree cut, do the following. For every $u \in \mathcal{A}(S)$, set:

$$F_u = 1 - \epsilon_B \mathbb{I} \left\{ \epsilon_F \leq x(\delta^\uparrow(u)) \leq 1 - \epsilon_F \right\}, \quad Z_u := \left(1 + \mathbb{I} \left\{ |\mathcal{A}(S)| \geq 4, x(\delta^\uparrow(u)) \leq \epsilon_F \right\} \right).$$

Set up and solve a polynomial size max-flow min-cut problem. Construct a graph with vertex set $\{s, X, Y, t\}$ with source s and sink t . We identify X with the set of good edge bundles in $E^\rightarrow(S)$ and Y with the set of atoms in $\mathcal{A}(S)$.

For every (good) edge bundle $\mathbf{e} \in X$, add an arc from s to \mathbf{e} of capacity $c(s, \mathbf{e}) := (1 + \alpha)x_{\mathbf{e}}$. For every $u \in \mathcal{A}(S)$, add an arc (u, t) with capacity $c(u, t) = x(\delta^\uparrow(u))F_u Z_u$. Finally, connect $\mathbf{e} = (u, v) \in X$ to each of $u, v \in Y$ with a directed edge of infinite capacity, i.e., $c(\mathbf{e}, u) = c(\mathbf{e}, v) = \infty$.

Let f be the max flow and return

$$m_{\mathbf{e}, u} := \frac{f_{\mathbf{e}, u}}{F_u}. \quad (73)$$

15.5 Reductions

In the following we compute the probability of events \mathcal{R} , corresponding to the probability of “decrease events” for every edge bundle. These then are used to compute values r_e for every edge, corresponding to the actual decrease amounts. For a set $F \subseteq E$, we let $r(F) = \sum_{e \in F} r_e$.

If e is a top edge (see [Definition 15.4](#)), then $e \in \mathbf{f}$ for some top edge bundle $\mathbf{f} = (u, v)$. Define

$$r_e = \frac{1}{2} \tau x_e (\mathbb{I} \{ \mathcal{R}_{\mathbf{f}, u} \} + \mathbb{I} \{ \mathcal{R}_{\mathbf{f}, v} \})$$

If \mathbf{e} is a bottom edge with near-cycle parent S , then define

$$r_e = \beta x_e \mathbb{I} \{ \mathcal{R}_S \}$$

Where τ, β are given in [Table 4](#).

$\mathbb{E}_r(e, Set)$

If e is a top edge in top edge bundle $\mathbf{f} = (u, v)$ return

$$\mathbb{E}[r_e \mid Set] = \frac{1}{2} \tau x_e (\mathbb{E}_{\mathcal{R}}(\mathbf{f}, u, Set) + \mathbb{E}_{\mathcal{R}}(\mathbf{f}, v, Set))$$

Otherwise, e is a bottom edge with near-cycle parent S . Return

$$\mathbb{E}[r_e \mid Set] = \beta x_e \mathbb{E}_{\mathcal{R}}(S, Set).$$

15.6 Increases

We now recall the definition of increase vectors in [KKO21] (over all edges) with the purpose of guaranteeing that every odd cut in \mathcal{H} is satisfied and then show how to compute its expectation. In the following subsection, the slack vectors are defined as the sum of the decrease vector and (a scaled version of) the increase vector.

15.6.1 Increases for bottom edges

Here we define the increase needed for bottom edges in each near-cycle cut S with near-cycle partition A, B, C . We let $I_S = I_S^\uparrow + I_S^\rightarrow$. We define:

$$I_S^\uparrow = (1 + \epsilon_\eta)(r(A^\uparrow)\mathbb{I}\{S \text{ not left happy}\} + r(B^\uparrow) \cdot \mathbb{I}\{S \text{ not right happy}\} + r(C)) \quad (74)$$

$\mathbb{E}_T(S, Set)$

Assume S is a near-cycle cut with near-cycle partition A, B, C . To compute the expected value of the first term in (74), note by linearity of expectation it suffices to compute the following for each $e \in A^\uparrow$.

- If e is a top edge in top edge bundle $\mathbf{f} = (u, v)$, compute:

$$\mathbb{E}[r_e \cdot \mathbb{I}\{S \text{ not left happy}\} \mid Set] = \frac{1}{2}x_e\tau\mathbb{E}[(\mathcal{R}_{\mathbf{f},u} + \mathcal{R}_{\mathbf{f},v}) \cdot \mathbb{I}\{S \text{ not left happy}\} \mid Set],$$

The expectation is equivalent to $\mathbb{E}_{\mathcal{R}}(\mathbf{f}, u, Set) - \mathbb{E}[\mathcal{R}_{\mathbf{f},u} \wedge \mathbb{I}\{S \text{ left happy}\} \mid Set]$ (plus the analogous quantity for v). To compute the second term, first recall the definition of left happy (see Definition 15.2): A_T is odd and $C_T = 0$. Now apply Corollary 13.14 using the necessary sets E_i and vectors $\vec{m}, \vec{\sigma}$ for $\mathcal{R}_{\mathbf{f},u}$ (as given in $\mathbb{E}_{\mathcal{R}}(\mathbf{f}, u, Set)$) and add two additional sets $E_A = A$ and $E_C = C$ and coordinates $\sigma_A = 1, \sigma_C = 0, m_A = 2, m_C = |V|$. (Note if \mathbf{f} is bad then the whole expectation is 0 and there is nothing to compute.)

- If instead e is a bottom edge in near-cycle cut \hat{S} , similarly compute:

$$\mathbb{E}[r_e \cdot \mathbb{I}\{S \text{ not left happy}\} \mid Set] = x_e\beta\mathbb{E}[(\mathcal{R}_{\hat{S}} \cdot \mathbb{I}\{S \text{ not left happy}\}) \mid Set].$$

Here to apply Corollary 13.14, sum over the events for $\mathcal{R}_{\hat{S}}$ as in (72) (used by $\mathbb{E}_{\mathcal{R}}(S, Set)$), to each one adding $E_A = A, E_C = C$ with $\sigma_A = 1, \sigma_C = 0, m_A = 2, m_C = |V|$ (similar to above).

Compute the remaining terms in the expectation of $\mathbb{E}[I_S^\uparrow \mid Set]$ analogously.

For an edge bundle \mathbf{e} and a set $A \subseteq E$ we use the shorthand $\mathbf{e}(A)$ to denote the set of edges in A and \mathbf{e} . We now define $I^\rightarrow(S)$ for a near-cycle cut S . There are three cases:

- **Case 1:** The parent \hat{S} of S is a polygon cut. Then define

$$I^\rightarrow(S) := (1 + \epsilon_\eta)\beta(\max\{x(A^\rightarrow), x(B^\rightarrow)\} + x(C^\rightarrow)) \cdot \mathbb{I}\{\mathcal{R}_{\hat{S}} = 1, S \text{ not happy}\}$$

- **Case 2:** The parent u of S is a degree cut with degree partition A', B', C' and has a pair $\mathbf{e} = (v, S), \mathbf{f} = (w, S)$ edges with $\text{type}(\mathbf{e}, S) = \text{type}(\mathbf{f}, S) = 2\text{-}2\text{-}2$. WLOG, let $x_{\mathbf{e}(B')}, x_{\mathbf{f}(A')} \leq \epsilon_{1/2}$. Then define:

$$I^{\rightarrow}(S) := (1 + \epsilon_{\eta}) \frac{\tau}{2} \max\{x_{\mathbf{e}(A')}, x_{\mathbf{f}(B')}\} (\mathbb{I}\{\mathcal{R}_{\{\mathbf{e}, \mathbf{f}\}}\} + \mathbb{I}\{\mathcal{R}_{\mathbf{e}, v}\} + \mathbb{I}\{\mathcal{R}_{\mathbf{f}, w}\}) \\ + (1 + \epsilon_{\eta}) \sum_{g \in \delta^{\rightarrow}(S) \setminus \mathbf{e}(A') \setminus \mathbf{f}(B')} r_g \cdot \mathbb{I}\{S \text{ not happy}\}.$$

- **Case 3:** Otherwise the parent u of S is a degree cut with no pair of 2-2-2 edges. Define:

$$I^{\rightarrow}(S) := (1 + \epsilon_{\eta}) \sum_{g \in \delta^{\rightarrow}(S)} \mathbb{E}[r_g \cdot \mathbb{I}\{S \text{ not happy}\} \mid \text{Set}].$$

$\mathbb{E}_{I^{\rightarrow}}(S, \text{Set})$

Assume S is a near-cycle cut. This function computes $\mathbb{E}[I^{\rightarrow}(S) \mid \text{Set}]$ for the above three cases as follows (using the notation from above):

- **Case 1:** Return

$$(1 + \epsilon_{\eta}) \beta (\max\{x(A^{\rightarrow}), x(B^{\rightarrow})\} + x(C^{\rightarrow})) \cdot \mathbb{E}[\mathcal{R}_{\hat{S}} \cdot \mathbb{I}\{S \text{ not happy}\} \mid \text{Set}]$$

Here to apply [Corollary 13.14](#) we sum over the events for $\mathcal{R}_{\hat{S}}$ as in (72) (used by $\mathbb{E}_{\mathcal{R}}(S, \text{Set})$), to each one adding $E_A = A, E_B = B, E_C = C$ and $\sigma_A = 1, \sigma_B = 1, \sigma_C = 0, m_A = 2, m_B = 2, m_C = |V|$.^a

- **Case 2:** Return

$$(1 + \epsilon_{\eta}) \frac{\tau}{2} \max\{x_{\mathbf{e}(A')}, x_{\mathbf{f}(B')}\} (\mathbb{P}[\mathcal{R}_{\{\mathbf{e}, \mathbf{f}\}} \mid \text{Set}] + \mathbb{P}[\mathcal{R}_{\mathbf{e}, v} \mid \text{Set}] + \mathbb{P}[\mathcal{R}_{\mathbf{f}, w} \mid \text{Set}]) \\ + (1 + \epsilon_{\eta}) \sum_{g \in \delta^{\rightarrow}(S) \setminus \mathbf{e}(A') \setminus \mathbf{f}(B')} \mathbb{E}[r_g \cdot \mathbb{I}\{S \text{ not happy}\} \mid \text{Set}].$$

Recall $\mathbb{E}[\mathcal{R}_{\{\mathbf{e}, \mathbf{f}\}} \mid \text{Set}]$ can be computed by $\mathbb{E}_{\mathcal{R}}(\mathbf{e}, S, \text{Set})$. To calculate $\mathbb{E}[r_g \cdot \mathbb{I}\{S \text{ not happy}\} \mid \text{Set}]$ we use techniques similar to $\mathbb{E}_{I^{\rightarrow}}(S, \text{Set})$.

- **Case 3:** Return

$$(1 + \epsilon_{\eta}) \sum_{g \in \delta^{\rightarrow}(S)} \mathbb{E}[r_g \cdot \mathbb{I}\{S \text{ not happy}\} \mid \text{Set}].$$

^aNote that the sets A, B, C written here are from the near-cycle partition of S , however the event $\mathcal{R}_{\hat{S}}$ uses the near-cycle partition of \hat{S} .

$\mathbb{E}_I(S, \text{Set})$

Return $\mathbb{E}_{I^{\rightarrow}}(S, \text{Set}) + \mathbb{E}_{I^{\leftarrow}}(S, \text{Set})$.

15.6.2 Increases for top edges

For each top edge bundle $\mathbf{e} = (u, v)$, using the values $m_{\mathbf{e},u}, m_{\mathbf{e},v}$ from [Section 15.4](#), define for a tree T :

$$I_{\mathbf{e},u} := \sum_{g \in \delta^\uparrow(u)} r_g \cdot \mathbb{I}\{\delta(u)_T \text{ is odd}\} \cdot \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f} \in \delta^\rightarrow(u)} m_{\mathbf{f},u}} \quad (75)$$

and define $I_{\mathbf{e},v}$ analogously. We then let

$$I_{\mathbf{e}} := I_{\mathbf{e},u} + I_{\mathbf{e},v}.$$

$\mathbb{E}_I(\mathbf{e}, \text{Set})$

Let $\mathbf{e} = (u, v)$. Compute:

$$\mathbb{E}[I_{\mathbf{e},u} \mid \text{Set}] := \sum_{g \in \delta^\uparrow(u)} \mathbb{E}[r_g \cdot \mathbb{I}\{\delta(u)_T \text{ is odd}\} \mid \text{Set}] \cdot \frac{m_{\mathbf{e},u}}{\sum_{\mathbf{f} \in \delta^\rightarrow(u)} m_{\mathbf{f},u}} \quad (76)$$

and compute $\mathbb{E}[I_{\mathbf{e},v} \mid \text{Set}]$ analogously, using the values $m_{\mathbf{e},u}$ computed previously in Step 6. To compute the above, we apply [Corollary 13.14](#) for sets E_i coming from $\mathbb{E}_r(g, \text{Set})$ and add an extra set $E_u = \delta(u)$ and coordinates $\sigma_u = 1, m_u = 2$.

Return

$$\mathbb{E}[I_{\mathbf{e}} \mid \text{Set}] := \mathbb{E}[I_{\mathbf{e},u} + I_{\mathbf{e},v} \mid \text{Set}].$$

15.7 Computation of $\mathbb{E}[c(s) \mid \text{Set}]$

First we define $s^{\mathcal{H}}$:

$$s_e^{\mathcal{H}} := -r_e + \begin{cases} I_{\mathbf{f}} \frac{x_e}{x_{\mathbf{f}}} & \text{if } e \in \mathbf{f} \text{ for a top edge bundle } \mathbf{f}, \\ I_S x_e & \text{if } p(e) = S \text{ for a polygon cut } S \in \mathcal{H}, \text{ i.e. } e \text{ is a bottom edge.} \end{cases} \quad (77)$$

Finally, we construct s . Note that $s_e^{\mathcal{H}} = 0$ with probability 1 for a bad edge bundle e . Therefore in [\[KKO22\]](#) a second slack vector was defined to allow bad edges to reduce. In particular, let E_g be the set of good edges and let $E_b := E \setminus E_g$ be the set of bad edges. Note all edges in $\delta(\{u_0, v_0\})$ are bad edges as they are not edge bundles in the hierarchy. Define the vector $s^{bad} : E \cup \{e_0\} \rightarrow \mathbb{R}$ as follows:

$$s_e^{bad} \leftarrow \begin{cases} \infty & \text{if } e = e_0 \\ -x_e(4\beta/5)(1 - 2\eta) & \text{if } e \in E_b, \\ x_e(4\beta/3) & \text{otherwise.} \end{cases} \quad (78)$$

Finally, where $\gamma = \frac{15}{32}\epsilon_P$, let $s = \gamma s^{bad} + (1 - \gamma)s^{\mathcal{H}}$. This is now exactly the vector s from Theorem 6.1 of [\[KKO22\]](#).

$\mathbb{E}_{c(s)}(e, Set)$

We have $\mathbb{E}[c(s_e) \mid Set] = \gamma c(s_e^{bad}) + (1 - \gamma) \mathbb{E}[s_e^{\mathcal{H}} \mid Set]$. $c(s_e^{bad})$ is a constant which can be computed by (78).

Thus, it is sufficient to compute $\mathbb{E}[c(s_e^{\mathcal{H}}) \mid Set]$. From (77), we just need $\mathbb{E}[r_e \mid Set]$, computed by $\mathbb{E}_r(e, Set)$, and $\mathbb{E}[I_f \mid Set]$ if $e \in \mathbf{f}$ is a top edge (computed by $\mathbb{E}_I(\mathbf{f}, Set)$) and $\mathbb{E}[I_S \mid Set]$ if it is a bottom edge with near-cycle parent S (computed by $\mathbb{E}_I(S, Set)$).

This concludes the proof of (2) of Lemma 13.11.

16 A Lower Bound for Max Entropy: k -Donuts

The main goal of this section is to lower bound the expected cost of the Eulerian tour output by the max entropy algorithm (although not the Hamiltonian tour if one uses shortcutting). We first describe the construction of a graphic k -donut instance, which will consist of $4k$ vertices. The cost function $c_{\{u,v\}}$ is given by the shortest path distance in the following graph.

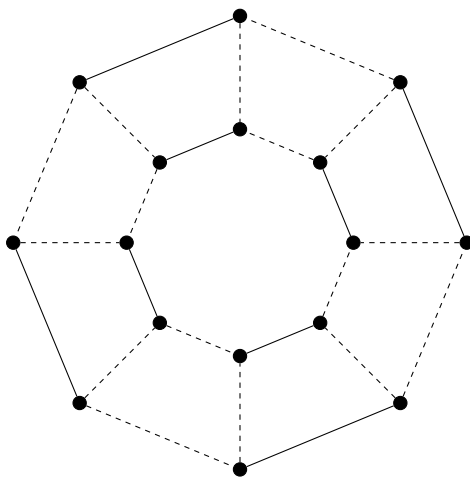


Figure 53: Our variant on the k -donut for $k = 4$, where k indicates the number of squares of dotted edges. There are $n = 4k$ vertices. The dotted edges have $x_e = \frac{1}{2}$ and the solid edges have $x_e = 1$ in the LP solution. All edges have cost 1, as this is a graphic instance.

Definition 16.1 (k -Donut Graph). For $k \in \mathbb{Z}_+$, $k \geq 3$, the k -donut is a 3-regular graph consisting of $2k$ "outer" vertices u_0, \dots, u_{2k-1} and $2k$ "inner" vertices v_0, \dots, v_{2k-1} . For each $0 \leq i \leq 2k - 1$, the graph has edges $\{u_i, u_{i+1 \pmod{2k}}\}$, $\{v_i, v_{i+1 \pmod{2k}}\}$, and $\{u_i, v_i\}$. See Fig. 53.

As noted by [BS21], there is a half-integral extreme point solution x of value $4k$ as follows, which we will work with throughout this note. Let $x_{\{u_i, v_i\}} = 1/2$ for all $0 \leq i \leq 2k - 1$, $x_{\{u_i, u_{i+1}\}} = x_{\{v_i, v_{i+1}\}} = 1/2$ for all odd i and $x_{\{u_i, u_{i+1 \pmod{2k}}\}} = x_{\{v_i, v_{i+1 \pmod{2k}}\}} = 1$ for all even i .⁴⁴ In the rest of the paper, we will say a set $S \subseteq V$ is *tight* if $x(\delta(S)) = 2$, and S is *proper* if $2 \leq |S| \leq |V| - 2$. For a set of edges M , we'll use $c(M) = \sum_{e \in M} c_e$.

16.1 The Max Entropy Algorithm on the k -Donut

We now describe the max entropy algorithm, and in particular discuss what it does when specialized to the k -donut. We will work with a description of the max entropy algorithm which is very similar to the one used for half-integral TSP in [KKO20]. In [KKO20], the authors show that without loss of generality there exists an edge e^+ with $x_{e^+} = 1$. To sample a 1-tree⁴⁵ T , their algorithm iteratively chooses a minimal proper tight set S not containing e^+ which is not crossed by any other tight set, picks a tree from the max entropy distribution on the induced graph $G[S]$,

⁴⁴By slightly perturbing the metric, one could ensure that x is the *only* optimal solution to the LP and thus the solution the max entropy algorithm works with. (Of course then the instance is no longer strictly graphic.)

⁴⁵A spanning tree plus an edge.

adds its edges to T , and contracts S . [KKO20] shows that if no such set remains, the graph is a cycle, possibly with multiple edges between (contracted) vertices. The algorithm then randomly samples a cycle and adds its edges to T . Finally the algorithm picks a minimum-cost perfect matching M on the odd vertices of T . Note that one can of course shortcut $T \uplus M$ to obtain a Hamiltonian cycle, however as discussed we will not consider the impact of this step. We also remark that this algorithm from [KKO20] is equivalent to the one used in [KKO21] as one lets the error measuring the difference between the marginals of the max entropy distribution and the subtour LP solution x go to 0 (see [KKO20; KKO21] for more details).

For ease of exposition, we work with the variant in which we do not use an edge e^+ and instead contract any minimal proper tight set which is not crossed. The two distributions over trees are essentially identical, perhaps with the exception of the edges adjacent to the vertices adjacent to e^+ . The performance of the two algorithms on graphic k -donuts can easily be seen to be the same as $k \rightarrow \infty$ since one can adjust the matching M with an additional cost of $O(1)$ to simulate any discrepancy between the two tree distributions.

Algorithm 5 Max Entropy Algorithm (Slight Variant of [KKO20])

- 1: Solve for an optimal solution x of the Subtour LP (2).
 - 2: Let G be the support graph of x .
 - 3: Set $T = \emptyset$. ▷ T will be a 1-tree
 - 4: **while** there exists a proper tight set of G that is not crossed by a proper tight set **do**
 - 5: Let S be a minimal such set.
 - 6: Compute the maximum entropy distribution μ of $E(S)$ with marginals $x_{|E(S)}$.
 - 7: Sample a tree from μ and add its edges to T .
 - 8: Set $G = G/S$.
 - 9: **end while**
 - 10: ▷ At this point G consists of a single cycle of length at least three, or two vertices with two fractional units of edges between them.
 - 11: **if** G consists of two vertices **then**
 - 12: Randomly sample two edges with replacement, choosing each edge each time with probability $x_e/2$.
 - 13: **else**
 - 14: Independently sample one edge between each adjacent pair, choosing each edge with probability x_e .
 - 15: **end if**
 - 16: Compute the minimum-cost perfect matching M on the odd vertices of T . Return $T \uplus M$.
-

The reason we use this description of the algorithm is that when specialized to k -donuts, Algorithm 5 is very simple and its behavior can be easily understood without using any non-trivial properties of the max entropy algorithm. It first adds the edges with $x_e = 1$ to the 1-tree. Then, it contracts the vertices $\{u_i, u_{i+1 \pmod{2k}}\}$ to a single vertex for all even i , and does the same for $\{v_i, v_{i+1 \pmod{2k}}\}$. After that, the the minimal proper tight sets consist of pairs of newly contracted vertices $\{u_i, u_{i+1 \pmod{2k}}\}, \{v_i, v_{i+1 \pmod{2k}}\}$ for even i . Since each of these pairs have two edges set to $1/2$ between them, the algorithm will simply choose one at random for each independently. After contracting these pairs the graph is a cycle. It follows that:

Claim 16.2. *On the k -donut, the max entropy algorithm will independently put exactly one edge among every pair $\{\{u_i, v_i\}, \{u_{i+1 \pmod{2k}}, v_{i+1 \pmod{2k}}\}\}$ in T for every odd i and exactly one edge among every pair*

$\{\{u_i, u_{i+1}\}, \{v_i, v_{i+1}\}\}$ in T for every even i .

We visualize these pairs in Figure 54.

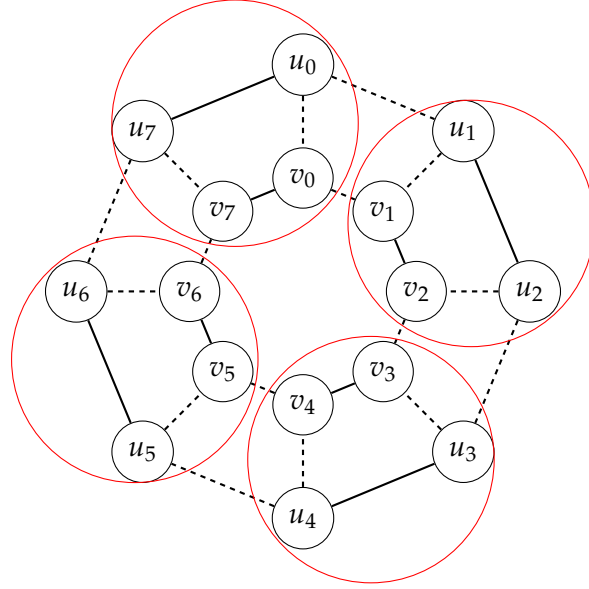


Figure 54: One edge among the pair of dotted edges inside each red cut will be chosen independently. Then one edge among each pair of dotted edges in the cycle resulting from contracting the red sets will be chosen independently.

The following claim is the only property we need in the remainder of the proof:

Claim 16.3. *For every pair of vertices (u_i, v_i) , $0 \leq i \leq 2k - 1$, exactly one of u_i or v_i will have odd degree in T , each with probability $\frac{1}{2}$. Let O_i indicate if u_i and u_{i+1} have the same parity. Then if $i \neq j$ and both are even or both are odd, O_i and O_j are pairwise independent.*

Proof. We will only do the case that both are odd as the other case is similar. To slightly simplify the notation we assume $i = 1$ perhaps after a cyclic shift of the indices.

Here the event O_1 depends only on the choice of the edges among the pairs $\{\{u_0, u_1\}, \{v_0, v_1\}\}$ and $\{\{u_2, u_3\}, \{v_2, v_3\}\}$. Similarly, O_j only depends on the independent choices among $\{\{u_{j-1}, u_j\}, \{v_{j-1}, v_j\}\}$ and $\{\{u_{j+1 \pmod{2k}}, u_{j+2 \pmod{2k}}\}, \{v_{j+1 \pmod{2k}}, v_{j+2 \pmod{2k}}\}\}$. The first choice for O_1 is independent of O_j if $j \neq 2k - 1$, and the second is independent of O_j if $j \neq 3$. Since $k \geq 3$ by definition of the k -donut, at most one of the independent choices is shared among the two events O_1, O_j . The proof follows by noticing that even after fixing one of the pairs, O_1 remains equally likely to be 0 or 1. \square

16.2 Analyzing the Performance of Max Entropy on the k -Donut

We now analyze the performance of the max entropy algorithm on graphic k -donuts. We first characterize the structure of the min-cost perfect matching on the odd vertices of T . We then use

this structure to show that in the limit as $k \rightarrow \infty$, the approximation ratio of the max entropy algorithm approaches 1.375 from below.

Claim 16.4. *Let T be any tree with the property that for every pair of vertices (u_i, v_i) for $0 \leq i \leq 2k - 1$, exactly one of u_i or v_i has odd degree in T . (This is [Claim 16.3](#)).*

Let o_0, \dots, o_{2k-1} indicate the odd vertices in T where o_i is the odd vertex in the pair (u_i, v_i) . Let M be a minimum-cost perfect matching on the odd vertices of T . Define:

$$M_1 = \{(o_0, o_1), (o_2, o_3), \dots, (o_{2k-2}, o_{2k-1})\}$$

$$M_2 = \{(o_{2k-1}, o_0), (o_1, o_2), \dots, (o_{2k-3}, o_{2k-2})\}$$

Then,

$$c(M) = \min\{c(M_1), c(M_2)\}.$$

Proof. We will show a transformation from M to a matching in which every odd vertex o_i is either matched to $o_{i-1 \pmod{2k}}$ or $o_{i+1 \pmod{2k}}$. This completes the proof, since then after fixing (o_0, o_1) or (o_{2k-1}, o_0) the rest of the matching is uniquely determined as M_1 or M_2 . During the process, we will ensure the cost of the matching never increases, and to ensure it terminates we will argue that the (non-negative) potential function $\sum_{e=(o_i, o_j) \in M} \min\{|i-j|, 2k-|i-j|\}$ decreases at every step. Note that this potential function is invariant under any renaming corresponding to a cyclic shift of the indices.

So, suppose M is not yet equal to M_1 or M_2 . Then there is some edge $(o_i, o_j) \in M$ such that $j \notin \{i-1, i+1 \pmod{2k}\}$. Without loss of generality (by switching the role of i and j if necessary), suppose $j \in \{i+2, i+3, \dots, i+k \pmod{2k}\}$. Possibly after a cyclic shift of the indices, we can further assume $i = 0$ and $2 \leq j \leq k$. Let o_l be the vertex that o_1 is matched to. We consider two cases depending on if $l \leq k+1$ or $l > k+1$.

Case 1: $l \leq k+1$. In this case, replace the edges $\{(o_0, o_j), (o_1, o_l)\}$ with $\{(o_0, o_1), (o_j, o_l)\}$. This decreases the potential function, as the edges previously contributed $j+l-1$ and now contribute $1+|j-l|$, which is a smaller quantity since $j, l \geq 2$. Moreover this does not increase the cost of the matching: We have $c_{\{o_0, o_1\}} \leq 2$ and $c_{\{o_j, o_l\}} \leq |j-l|+1$, so the two new edges cost at most $|j-l|+3$. On the other hand, the two old edges cost at least $c_{\{o_0, o_j\}} + c_{\{o_1, o_l\}} \geq j+l-1$, which is at least $|j-l|+3$ since $j, l \geq 2$.

Case 2: $l > k+1$. In this case, we replace the edges $\{(o_0, o_j), (o_1, o_l)\}$ with $\{(o_0, o_l), (o_1, o_j)\}$. This decreases the potential function, as the edges previously contributed $j+(2k-l+1)$ and now they contribute $(2k-l)+(j-1)$. Also, the edges previously cost at least $j+(2k-l+1)$, and now cost at most $(2k-l+1)+j$. Thus the cost of the matching did not increase. \square

We now analyze the approximation ratio of the max entropy algorithm.

Lemma 16.5. *If A is the output of the max entropy algorithm on the k -donut (i.e. $T \uplus M$), then*

$$\lim_{k \rightarrow \infty} \frac{\mathbb{E}[c(A)]}{c(\text{OPT})} = \lim_{k \rightarrow \infty} \frac{\mathbb{E}[c(A)]}{c(x)} = 1.375,$$

where $c(x)$ is the cost of the extreme point solution to the subtour LP.

Proof. We know that the LP value is $4k$. Since the k -donut is Hamiltonian, we also have that the optimal tour has length $4k$. On the other hand, $c(A) = c(T) + c(M)$, where T is the 1-tree and M is the matching. Note that the cost of the 1-tree is always $4k$. On the other hand, we know that $c(M) = \min\{c(M_1), c(M_2)\}$ from the previous claim. Thus, it suffices to reason about the cost of M_1 and M_2 . We know that for every i , $c_{\{o_i, o_{i+1} \pmod{2k}\}} = 2$ with probability $1/2$ and 1 otherwise, using [Claim 16.3](#). Thus, the expected cost of each edge in M_1 and M_2 is 1.5. Since each matching consists of k edges, by linearity of expectation, $\mathbb{E}[c(M_1)] = \mathbb{E}[c(M_2)] = 1.5k$. By Jensen's inequality, this implies $\mathbb{E}[c(M)] \leq 1.5k$. This immediately gives an upper bound on the approximation ratio of $\frac{4k+1.5k}{4k} = 1.375$. In the remainder we prove the lower bound.

For each i , construct a random variable X_i indicating if $c_{\{o_i, o_{i+1} \pmod{2k}\}} = 2$. By [Claim 16.3](#), these variables are pairwise independent. Thus, for M_1 , we have $\text{Var}(\sum_{i=0}^{k-1} X_{2i}) = \sum_{i=0}^{k-1} \text{Var}(X_{2i}) = k/4$, so $\sigma(\sum_{i=0}^{k-1} X_{2i}) = \sqrt{k}/2$.

Therefore, applying Chebyshev's inequality for M_1 ,

$$\mathbb{P}\left[c(M_1) \geq \left(\frac{3}{2} - \epsilon\right)k\right] = \mathbb{P}\left[\sum_{i=0}^{k-1} X_{2i} \geq \left(\frac{1}{2} - \epsilon\right)k\right] \geq 1 - \mathbb{P}\left[\left|\sum_{i=0}^{k-1} X_{2i} - \mu\right| \geq \epsilon k\right] \geq 1 - \frac{1}{4\epsilon^2 k}.$$

Choosing $\epsilon = k^{-1/4}$ and applying a union bound (the same bound applies to M_2), we obtain the chance that both matchings cost at least $\frac{3}{2}k - k^{3/4}$ occurs with probability at least $1 - \frac{1}{2\sqrt{k}}$. Even if the matching has cost 0 on the remaining instances, the expected cost of the matching is therefore at least $(1 - \frac{1}{2\sqrt{k}})(\frac{3}{2}k - k^{3/4}) \geq \frac{3}{2}k - 2k^{3/4}$. Since the cost of the 1-tree is always $4k$, we obtain an expected cost of $\frac{11}{2}k - 2k^{3/4}$ with a ratio of

$$\mathbb{E}[c(T \uplus M)] = \frac{\frac{11}{2}k - 2k^{3/4}}{OPT} = \frac{\frac{11}{2}k - 2k^{3/4}}{4k} \geq \frac{11}{8} - k^{-1/4},$$

which goes to $\frac{11}{8}$ as $k \rightarrow \infty$. □

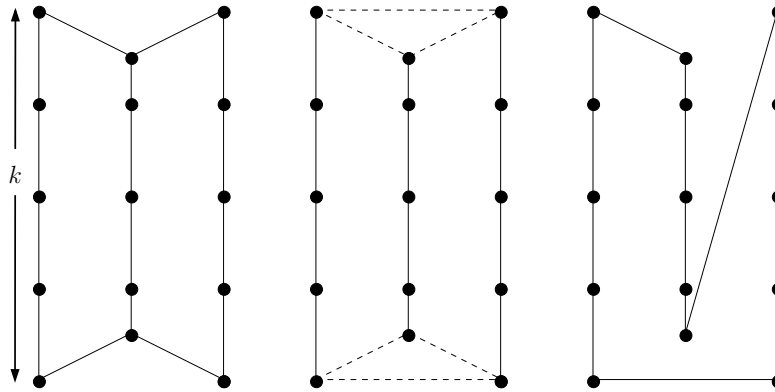


Figure 55: Illustration of a known worst-case example for the integrality gap for the symmetric TSP with triangle inequality. The figure on the left gives a graph, and costs c_{ij} are the shortest path distances in the graph. The figure in the center gives the LP solution, in which the dotted edges have value $1/2$, and the solid edges have value 1 . The figure on the right gives the optimal tour. The ratio of the cost of the optimal tour to the value of the LP solution tends to $4/3$ as k increases.

17 $4/3$ for Cycle Cut Instances

It is known that the integrality gap of the Subtour LP is at least $4/3$, due to a set of graph TSP instances shown in Figure 55, and another set of weighted instances due to Boyd and Sebó [BS21] known as k -donuts. These instances are half-integral instances. Schalekamp, Williamson, and van Zuylen [SWZ13] have conjectured that half-integral instances are the worst-case instances for the integrality gap. It has long been conjectured⁴⁶ that the integrality gap is exactly $4/3$, but until the work of Karlin et al. there had been no progress on that conjecture for several decades. Goemans [Goe95b] and Benoit and Boyd [BB08] give evidence that the $4/3$ conjecture is correct.

In the case of half-integral instances, some results are known. Mömke and Svensson [MS16] have shown a $4/3$ -approximation algorithm for half-integral graph TSP, also yielding an integrality gap of $4/3$ for such instances; because of the worst-case examples of Figure 55, their result is tight. Boyd and Carr [BC11b] give a $4/3$ -approximation algorithm (and an integrality gap of $4/3$) for a subclass of half-integer solutions they call triangle points (in which the half-integer edges form disjoint triangles); the examples of Figure 55 show that their result is tight also. Haddadan and Newman [HN19] prove interesting results for the half-integral case with symmetric and metric costs. Boyd and Sebó [BS21] give an upper bound of $10/7$ for a subclass of half-integral solutions they call square points (in which the half-integer edges form disjoint 4-cycles). In a paper released just prior to their general improvement, Karlin, Klein, and Oveis Gharan [KKO20] (KKO) gave a 1.49993 -approximation algorithm in the half-integral case; in particular, they show that given a half-integral solution, they can produce a tour of cost at most 1.49993 times the value of the corresponding objective function. Gupta, Lee, Li, Mucha, Newman, and Sarkar [Gup+22] improve this factor to 1.4983 .

With the improvements on the $3/2$ bound remaining very incremental for weighted instances

⁴⁶The first place that the authors are aware of a published statement of the conjecture is in a 1995 paper of Goemans [Goe95b], but the conjecture was in circulation earlier than that.

of the TSP, even in the half-integral case, we turn the question around and look for a large class of weighted half-integral instances for which we can prove that the $4/3$ conjecture is correct, preferably one containing the known worst-case instances.

To define our instances, we turn to some terminology from the half integral case. There, we used induction on a hierarchy of *critical tight* sets of the half-integral LP solution x . A set $S \subset V$ is *tight* if the corresponding LP constraint is met with equality; that is, $x(\delta(S)) = 2$. A set S is *critical* if it does not cross any other tight set; that is, for any other tight set T , either $S \cap T = \emptyset$ or $S \subseteq T$ or $T \subseteq S$. The critical tight sets then give rise to a natural tree-like hierarchy based on subset inclusion.

The algorithm constructs a tour on the hierarchy by sampling a random spanning tree on the child nodes for each critical tight set, starting with the minimal sets in the hierarchy and working bottom up. Following Christofides-Serdyukov, we then computed a minimum-cost T -join on the odd degree vertices of the resulting tree. In this algorithm, we differentiated between *cycle cuts* (in which the child nodes of a parent are linked by pairs of edges in a chain) and *degree cuts* (in which the child nodes of a parent form a 4-regular graph; more detail is given in subsequent sections).

Here, we will consider half-integral instances in which there are *only cycle cuts*, which we will refer to as half-integral cycle cut instances. Our contribution is to give a randomized $\frac{4}{3}$ -approximation algorithm for these instances; it generates a distribution over connected Eulerian subgraphs with expected cost at most $4/3$ the value of the LP solution. More precisely, we give a distribution over connected Eulerian subgraphs such that each edge e is used with expectation at most $\frac{4}{3}x_e$, which implies the result (note that edges are sometimes doubled in the Eulerian graph). It is not hard to show that both the bad examples in Figure 55 and the k -donut instances of Boyd and Sebő [BS21] are cycle cut instances (Boyd and Carr’s result for triangle points works for the examples of Figure 1, but not for k -donuts). Thus our bound of $4/3$ is tight and cannot be improved; furthermore, our result works for the known worst-case instances.

Our approach to the problem is novel and does not use the same Christofides-Serdyukov framework as employed above. Instead, we perform a top-down induction on the hierarchy of critical tight sets. For each set in the hierarchy, we define a set of “patterns” of edges incident on it such that the set has even degree.

For each pattern, we give a distribution of edges connecting the chain of child nodes in the cycle cut, which induces a distribution of patterns on each child. Crucially, we then show that there is a *feasible region* R of distributions over patterns, such that if the distribution of patterns on the parent node belongs to R , then the induced distribution on patterns on each child node also belong to R .

Our result leads to several interesting open questions. The first is whether it is possible to extend the $4/3$ -approximation algorithm to the general half-integral case. We believe it should be possible to improve modestly on the 1.4983-approximation of Gupta et al. [Gup+22] by combining our result for cycle cuts with some additional ideas. We do not elaborate on this potential improvement because both the improvement and the additional ideas are incremental relative to the ideas introduced in this paper. The second open question is whether our result extends to the case of cycle cuts for non-half-integral solutions. We believe this to be possible through a more refined understanding of the patterns that result from considering non-half-integral solutions. A third open question is whether we can unify our result and that of Boyd and Carr on triangle points. Triangle points need not be cycle cut instances, and it would be interesting to know of a single class of half-integral solutions that have an integrality gap of $4/3$ and which captures both

cycle cut instances and triangle points.

One major implication of our result is to focus attention on the half-integral degree cut case, in which every vertex has degree four, all edges have LP value $1/2$, and every non-trivial cut has at least six edges in it. While it is not clear whether a $\frac{4}{3}$ -approximation algorithm working on just these instances can be combined with our result for an overall $\frac{4}{3}$ -approximation algorithm for half-integral solutions, it is clear that understanding the degree cut case is a necessary next step to obtain any significant improvement in the approximation factor. We believe that giving a feasible region on the distribution of patterns as described above will be useful in obtaining an improved approximation that integrates both degree and cycle cuts.

17.1 Technical Overview

We now give a more in-depth overview of our algorithm and proof techniques.

Given a half-integral LP solution x , we construct a 4-regular 4-edge-connected multigraph G by including every edge with $x_e = 1/2$ once and every edge with $x_e = 1$ twice. Therefore, in G the tight sets S have $|\delta(S)| = 4$. For the remainder of this paper we will refer to this graph G instead of a half integral solution x .

Our strategy is to exhibit a convex combination of Eulerian tours that uses every edge at most $\frac{2}{3} = \frac{4}{3}x_e$ of the time. Using each edge of the graph at most $\frac{2}{3}$ of the time immediately implies that when we sample a tour from this distribution, its expected cost will be at most $\frac{4}{3}$ times the cost of x . This allows us to prove our main theorem.

Theorem 17.1. *There is a randomized $4/3$ -approximation algorithm for half-integral cycle cut instances of the TSP that produces an Eulerian tour with expected cost at most $\frac{4}{3} \sum_{e \in E} c_e x_e$.*

To construct our distribution of tours, we work on the cycle cut hierarchy from the top down. Each cut in the hierarchy is either a singleton vertex, or a cycle cut that contains two or more tight cuts inside. At the root of the hierarchy is a cycle cut S such that $V - S$ is a single vertex. For every cycle cut, its child cuts are linked together in a chain, each cut connected to the next by a pair of edges.

Our construction begins by specifying a distribution of *patterns* entering the root cycle cut, where a "pattern" refers to a given multiset of edges that enter the cycle cut. We then work down the hierarchy to determine the distribution of patterns entering every cycle cut. Inductively, consider a cut C in the hierarchy, and suppose we have already determined the distribution of patterns that enter C . We describe *rules* that dictate how to connect the child cuts in the chain inside of C , as a function of the pattern that enters C . This in turn determines the distribution of patterns entering each child of C .

The crux of the argument is to show a *feasible region* R of distributions over patterns, such that: 1) If the distribution of patterns on the parent node belongs to R , then the induced distribution on patterns on each child node also belong to R , and 2) the distributions in the region use each edge e at most $\frac{4}{3}x_e$ of the time. We are able to impose any distribution we want on the root, thus this is sufficient to give the result.

Conceptually key to our analysis is a visualization of the process through the lens of *Markov chains*. Each pattern corresponds to a state in the Markov chain. Given a distribution over patterns entering some cut C , applying the transition matrix gives the distribution of patterns entering the children of C . Using this language, the region R will satisfy the property that $PR \subseteq R$, where P is

the transition matrix of the Markov chain. In fact, there will turn out to be two Markov chains, depending on if the parity of the number of cuts inside C is even or odd; letting P_{even} and P_{odd} be the transition matrices of these two chains, our proof will show that $P_{\text{even}} R \subseteq R$ and $P_{\text{odd}} R \subseteq R$.

From the construction, one can easily sample an Eulerian tour in a top-down manner. To do so, we first choose an arbitrary distribution of patterns $\mathbf{p} \in R$ from the feasible region. Then, we sample from \mathbf{p} a pattern entering the topmost cycle cut. Next, supposing that we are at a cycle cut C and have determined the pattern entering it, we follow the (randomized) rules to sample a set of edges to connect the children of C . Applying this procedure all the way down the hierarchy gives an Eulerian tour. By design, this (random) tour satisfies the property that the distribution of edges entering each cut in the hierarchy belongs to R , and that each edge is used at most $\frac{4}{3}x_e$ of the time in expectation.

17.2 Preliminaries

Given a half-integral LP solution x , we construct a 4-regular 4-edge-connected multigraph $G = (V, E)$ by including a single copy of every edge e for which $x_e = \frac{1}{2}$ and two copies of every edge e for which $x_e = 1$. See Fig. 57 and Fig. 59 for examples.

17.2.1 The structure of minimum cuts

We state the following for general k -edge-connected multigraphs. In our setting, $k = 4$.

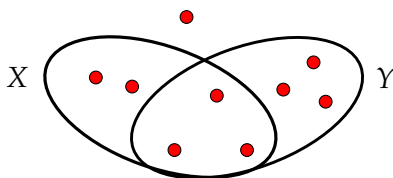


Figure 56: An example of two crossing sets.

Definition 17.2. For a k -edge-connected multigraph $G = (V, E)$, we say:

- Any set $S \subseteq V$ such that $|\delta(S)| = k$ (i.e., its boundary is a minimum cut) is a **tight set**.
- A set $S \subseteq V$ is **proper** if $2 \leq |S| \leq n - 2$ and a **singleton** if $|S| = 1$.
- Two sets $S, S' \subseteq V$ **cross** if all of $S \setminus S'$, $S' \setminus S$, $S \cap S'$, and $V \setminus (S \cup S') \neq \emptyset$ are non-empty (see Fig. 56).

The following are two standard facts about minimum cuts; for proofs see [FF09].

Fact 17.3. If two tight sets S and S' cross, then each of $S \setminus S'$, $S' \setminus S$, $S \cap S'$ and $\overline{S \cup S'}$ are tight. Moreover, there are no edges from $S \setminus S'$ to $S' \setminus S$, and there are no edges from $S \cap S'$ to $\overline{S \cup S'}$.

Fact 17.4. Let $G = (V, E)$ be a k -regular k -edge-connected graph. Suppose either $|V| = 3$ or G has at least one proper min cut, and every proper min cut is crossed by some other proper min cut. Then, k is even and G forms a cycle, with $k/2$ parallel edges between each adjacent pair of vertices.

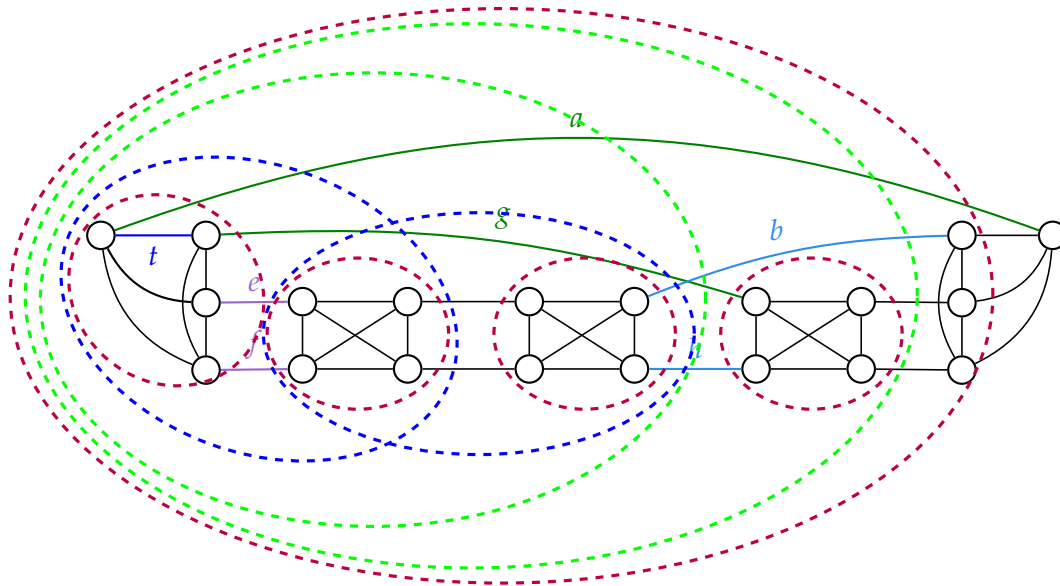


Figure 57: An example of a half integral instance, where the rightmost vertex is the root. The dotted lines circle (one side of) each minimum cut of the graph. The red cuts are "degree cuts" and the green cuts are the "cycle cuts." The blue cuts cross one another and therefore do not appear in the hierarchy of critical cuts.

17.3 Cycle cut instances and the hierarchy of critical tight sets

We first define our class of instances.

Definition 17.5 (Cycle cut instance). *We say a graph G is a **cycle cut instance** if every non-singleton tight set S can be written as the union of two tight sets $A, B \neq S$.*

Note that this definition includes complements of singletons, i.e. sets of size $n - 1$. As mentioned in the introduction this condition captures the two known integrality gap examples of the subtour LP. See Fig. 59 for a cycle cut instance. Due to the following fact, an equivalent definition of a cycle cut instance is that all non-singleton tight sets which are *not crossed* by any other tight set can be written as the union of two tight sets.

Fact 17.6. *In any graph G , every tight set which is crossed by another tight set can be written as the union of two tight sets.*

Proof. Let S be a tight set crossed by T . Then by Fact 17.3, $S \setminus T$ and $S \cap T$ are tight sets; the claim follows. \square

Thus, the condition in Definition 17.5 is met trivially by all crossed sets. So, it is enough to ensure all sets that are not crossed can be written as the union of two tight sets.

We now show a third equivalent definition of cycle cut instances matching what is described in the introduction. First, fix an arbitrary **root vertex** $r \in V$, and for all cuts we consider we will take the side which does not contain r .

Definition 17.7 (Critical cuts). *A critical cut is any tight set $S \subseteq V \setminus \{r\}$ which does not cross any other tight set.*

Definition 17.8 (Hierarchy of critical cuts, \mathcal{H}). Let $\mathcal{H} \subseteq 2^{V \setminus r}$ be the set of all critical cuts.

For an example, see Fig. 57. The hierarchy naturally gives rise to a parent-child relationship between sets as follows:

Definition 17.9 (Child, parent, $E^\rightarrow(S)$). Let $S \in \mathcal{H}$ such that $|S| \geq 2$. Call the maximal sets $C \in \mathcal{H}$ for which $C \subset S$ the children of S , and call S their parent. Finally, define $E^\rightarrow(S)$ to be the set of edges with endpoints in two different children of S .

Definition 17.10 (Cycle cut, degree cut). Let $S \in \mathcal{H}$ with $|S| \geq 2$. Then we call S a **cycle cut** if when $G \setminus S$ and all of the children of S are contracted, the resulting graph forms a cycle of length at least three with two parallel edges between each adjacent node. Otherwise, we call it a **degree cut**.

While this definition of a cycle cut may sound specialized, due to Fact 17.4, cycle cuts arise very naturally from collections of crossing min cuts. See Fig. 57 for a general example whose hierarchy of critical tight sets contains both degree cuts and cycle cuts.

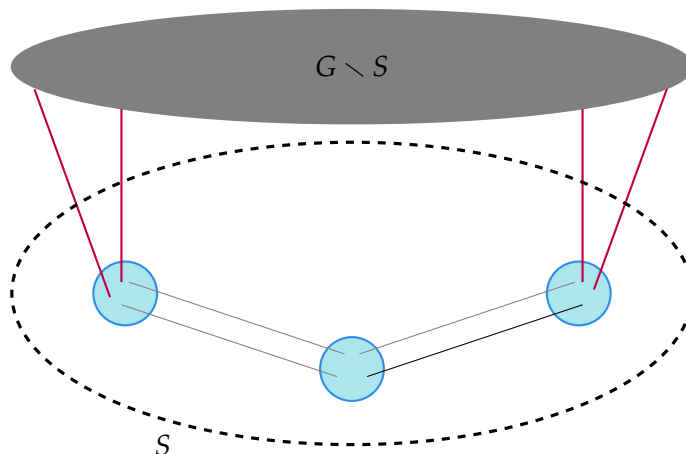


Figure 58: S is an example of a cycle cut with three children. In blue are contracted critical tight sets. In gray is the rest of the graph with S contracted. As in Fact 17.4, we can see that when $G \setminus S$ is contracted into a single vertex, the resulting graph is a cycle with 2 edges between each adjacent vertex. In our recursive proof of our main theorem in Section 17.5, we are given a distribution of Eulerian tours over G/S , so in particular on the red edges here, and will then extend it to G with the blue critical sets contracted it by picking a distribution over the black edges.

Fact 17.11. If G is a cycle cut instance, then for any choice of r , \mathcal{H} is composed only of cycle cuts (and singletons).

Proof. Let \mathcal{H} be the hierarchy of critical cuts for an arbitrary choice of r and let $S \in \mathcal{H}$ with $|S| \geq 2$ (note $\mathcal{H} \neq \emptyset$ since $V \setminus \{r\} \in \mathcal{H}$). We will show that S is a cycle cut.

Contract $G \setminus S$ and all of the children of S ; call the resulting graph G' . By definition of \mathcal{H} , G' contains no proper tight sets which are not crossed. If S contains a proper tight set which is crossed, then by Fact 17.4, it is a cycle cut and we are done.

Otherwise, G' contains no proper tight sets. Since G is a cycle cut instance and S is a tight set, there exist tight sets A, B such that $A \cup B = S$. Since the contracted children of S are not crossed

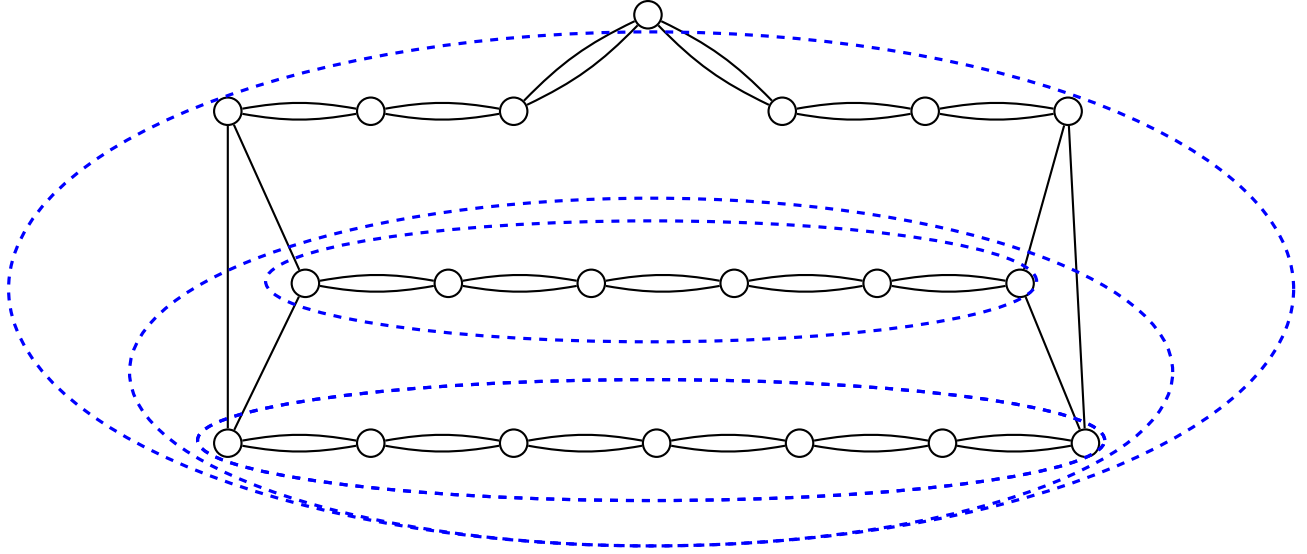


Figure 59: An example of a cycle cut instance, which is also the canonical integrality gap example for the Subtour LP. The non-singleton critical cycle cuts are shown in blue. The topmost vertex is the root.

and there are no proper tight sets in G' , A and B must be vertices of this graph. Therefore, G' must have exactly three vertices. Since it is 4-regular, it is a cycle of length three, and thus is a cycle cut. \square

Thus, in the remainder of the paper, we assume \mathcal{H} is a collection of cycle cuts. We remark that the following is also true:

Fact 17.12. *If for some choice of r , \mathcal{H} is composed only of cycle cuts, then G is a cycle cut instance.*

Proof. Fix any tight set S with $|S| \geq 2$. We will show it can be written as the union of two tight sets not equal to S . If it is crossed by another tight set, by Fact 17.6 we are done. If $r \notin S$, then S appears in \mathcal{H} . Then, the graph in which the children of S and $G \setminus S$ are contracted is a cycle with vertices a_0, \dots, a_k . Then S is the union of the vertices contained in tight sets $\{a_1, \dots, a_{k-1}\}$ and $\{a_k\}$.

Otherwise $r \in S$, and $\bar{S} \in \mathcal{H}$. Furthermore, since $|S| \geq 2$, $\bar{S} \neq S \setminus \{r\}$. Thus, \bar{S} has a parent S' . Consider the graph $a_0 = G \setminus S'$, a_1, \dots, a_k induced by contracting the children of S' and $G \setminus S'$. We have $\bar{S} = a_i$ for some $i \neq 0$. Then, S can be written as $(V \setminus \{a_i, a_{i-1}\}) \cup \{a_{i-1}\}$ as desired. \square

17.4 Structure of cycle cuts

Given $S \in \mathcal{H}$, let $a_0 = G \setminus S$ and let a_1, \dots, a_k be its children in \mathcal{H} (which are either vertices or cycle cuts). By Fact 17.4 a_0, \dots, a_k can be arranged into a cycle such that two edges go between each adjacent vertex. WLOG let a_1, \dots, a_k be in counterclockwise order starting from a_0 . We call a_1 the leftmost child of S and a_k the rightmost child.

Definition 17.13 (External and internal cycles cuts). *Let $S \in \mathcal{H}$ be a cut with parent S' . We call S external if in the ordering a_0, \dots, a_k of S' (as given above), $S = a_1$ or $S = a_k$. Otherwise, call S internal.*

For example, if the blue nodes in Fig. 58 are contracted cycle cuts, the left and right nodes are external, while the middle one is internal. Note that for an cycle cut S with parent S' , if S is external then $|\delta(S) \cap \delta(S')| = 2$, and if S is internal then $|\delta(S) \cap \delta(S')| = 0$.

Using the following simple fact, we will now describe our convention for drawing and describing cycle cuts:

Fact 17.14. *Let $A, B, C \in \mathcal{H}$ be three distinct critical cuts such that $A \subsetneq B$ and $B \cap C = \emptyset$ or $B \subseteq C$. Then $|\delta(A) \cap \delta(C)| \leq 1$.*

Proof. Suppose otherwise, and A shares two edges with C .

First, suppose $B \cap C = \emptyset$. Then, $A \cup C$ is a minimum cut, contradicting that B was a critical cut since it is crossed by $A \cup C$. Note $A \cup C$ crosses B since i) $B \setminus (A \cup C) \neq \emptyset$ since $A \subsetneq B$, ii) $(A \cup C) \setminus B \neq \emptyset$ since $C \cap B = \emptyset$, iii) $(A \cup C) \cap B \neq \emptyset$ since $A \subseteq B$ and finally iv) $(A \cup C) \cup B \neq V$ since neither contains the root.

Otherwise, suppose $B \subseteq C$. But then $B \setminus A$ has two edges to A and two edges to $C \setminus B$, implying that $C \setminus A$ was a minimum cut. This contradicts that B is a critical tight set, as it is crossed by $C \setminus A$. We can verify B crosses $C \setminus A$ as follows: i) $B \setminus (C \setminus A) \neq \emptyset$ as it contains $A \neq \emptyset$, ii) $(C \setminus A) \setminus B \neq \emptyset$ as $B \subsetneq C$ and $A \subseteq B$, iii) $(C \setminus A) \cap B \neq \emptyset$ since $B \setminus A \neq \emptyset$, and finally iv) the union is not everything since neither contains the root. \square

Definition 17.15 ($\delta^L(S), \delta^R(S)$). *Let $S \in \mathcal{H}$ be a cycle cut. We will define a partition of $\delta(S)$ into two sets $\delta^L(S), \delta^R(S)$ each consisting of two edges.*

If $S \neq V \setminus \{r\}$, then it has a parent S' . S' has children a_1, \dots, a_k such that $S = a_i$ for $i \neq 0$. Let $\delta^L(S) = \delta(S) \cap \delta(a_{i-1})$ and $\delta^R(S) = \delta(a_{i+1 \pmod{k+1}}) \cap \delta(S)$. In other words, we partition the edges of S into the two edges going to the left neighbor of S in the cycle defined by S' 's children and the two edges going to the right neighbor.

Otherwise $S = V \setminus \{r\}$. Then if a_1, \dots, a_k are the children of S , let $\delta^L(S)$ consist of an arbitrary edge from $\delta(S) \cap \delta(a_1)$ and an arbitrary edge from $\delta(S) \cap \delta(a_k)$. Let $\delta^R(S) = \delta(S) \setminus \delta^L(S)$.

By Fact 17.14 and the definition of $\delta^L(S), \delta^R(S)$ for $S = V \setminus \{r\}$, if S' is an external child of a cycle cut S , then $|\delta^L(S') \cap \delta(S)| = |\delta^R(S') \cap \delta(S)| = 1$. This allows us to adopt the following convention for drawing cycle cuts which we will call the **caterpillar drawing** of S : for an example, see Fig. 60. Formally, let $S \in \mathcal{H}$ be a cycle cut with children $a_1, \dots, a_k \in \mathcal{H}$. Arrange a_1, \dots, a_k in a horizontal line. First, expand the children of a_1 vertically (if it is not a singleton) such that the unique edge in $\delta^L(S) \cap \delta(a_1)$ is pointing up (if it is a singleton, simply draw this edge pointing up). Then, expand a_2, \dots, a_k one by one into their respective children (if they exist), placing the children vertically in increasing or decreasing order of their index so that the edges from a_i to a_{i+1} do not cross. If a_k is a singleton, arbitrarily choose which edge to draw pointing up. Otherwise, let a' be the topmost child of a_k . Draw the unique edge in $\delta(S) \cap \delta(a')$ pointing up.

There are two distinct types of cycle cuts:

Definition 17.16 (Straight and twisted cycle cuts). *Let $S \in \mathcal{H}$ be a cycle cut. caterpillar drawing of S . If $\delta^L(S)$ has both edges pointing up, then call it a straight cycle cut. Otherwise, call it a twisted cycle cut. See Fig. 60 for examples.*

In future sections, we abbreviate the caterpillar drawing by contracting the non-singleton children of S . We do so partially for cleaner pictures but also to emphasize that all the relevant

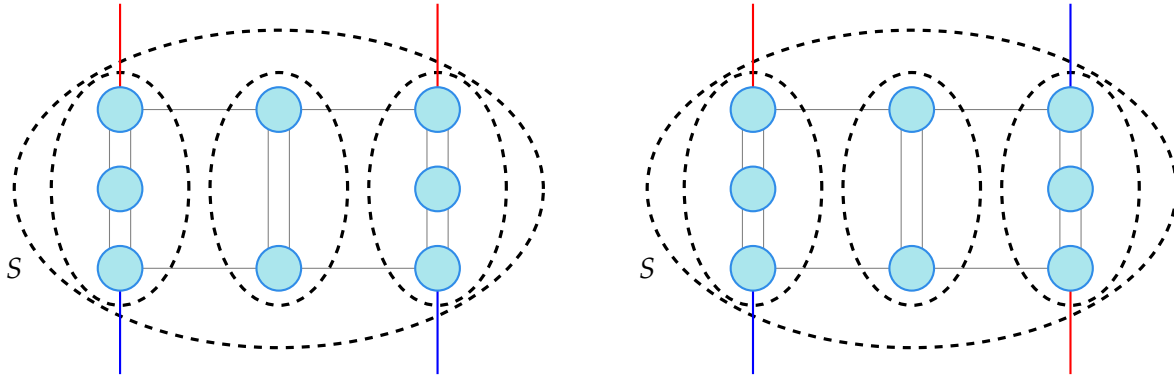


Figure 60: Two caterpillar drawing of two different cycle cuts S with three children. The red edges are in the $\delta^L(S)$ partition, and the blue edges are in the $\delta^R(S)$ partition. The left drawing is a straight cycle cut, and the right is a twisted cycle cut as per [Definition 17.16](#).

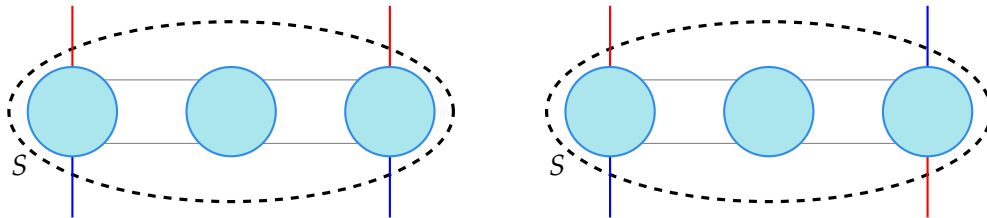


Figure 61: On the left is a shorthand caterpillar drawing for the straight cycle cut on the left in [Fig. 60](#), the style of picture we will use in future sections. To obtain this picture we contract the children of the cut. These are the shorthand drawings of [Fig. 60](#), so the left is a straight cycle cut and the right is twisted. We also label the edges as described below.

information used by our construction in the following section is contained in the abbreviated pictures.

To help the reader's understanding, we suggest looking at Fig. 59. The largest critical cut is arbitrarily chosen to be straight or twisted. The second largest critical cut is a straight cut. The smallest non-singleton critical cuts are arbitrarily chosen to be straight or twisted.

In the following, we will need to distinguish between the straight and twisted types of cuts as well as those with an even versus an odd number of children.

17.5 Proof of Theorem 17.1

We now present the proof of our main result, a $\frac{4}{3}$ -approximation for half-integral cycle-cut instances of the TSP. To prove Theorem 17.1, we construct a distribution of Eulerian tours such that every edge is used at most $\frac{2}{3}$ of the time. Since $x_e = \frac{1}{2}$ for every edge in the graph, this immediately implies that when we sample a tour from this distribution, its expected cost is at most $\frac{4}{3}$ times the value of the LP. We work on the cycle cut hierarchy from the top down, and inductively specify the distribution of edges that enter every cut.

Figure 61 depicts our convention for visualizing a cycle cut as described in Section 17.2. We say that a cycle cut is *even* if it contains an even number of children, and *odd* otherwise. Figure 63 illustrates the *patterns* we use, where "pattern" refers to a multiset of edges that enter a cycle cut. For each pattern entering a parent cycle cut, we give (randomized) rules which describe how to connect up its children – this induces a distribution of patterns for each child. We represent this process using a Markov chain with 4 states, illustrated in Figure 63. The figure shows the mapping from patterns to states; the transitions will come from the rules for connecting up the children, which we describe later. In the figure, each state contains two pictures, which represent the *parity* of the edges in the patterns that are mapped to the state. Specifically, an edge that is present is used exactly once, whereas an edge that is not present may be either unused or doubled. For example, Figure 64 illustrates all possible patterns that are captured by the top picture of state 1. Finally, we maintain the invariant that if a cycle cut is in a given state, then each of the two pictures are equally likely. (When we later give the rules for connecting up the children, we will ensure this invariant is preserved.) Thus, when we say a cycle cut is in a given state with probability p , this means the parity of the pattern entering it follows the top picture in the state with probability $\frac{p}{2}$, and the bottom picture with probability $\frac{p}{2}$. We will use the phrase "the distribution of patterns entering a cycle cut C is (p_1, p_2, p_3, p_4) " to mean that for all $i \in \{1, 2, 3, 4\}$, C is in state i with probability p_i .

To prove our main result, we give a *feasible region* R of distributions over the states of the Markov chain, and show that: 1) If the distribution of patterns entering a cycle cut C belongs to R , there is a way to connect up the children of C such that the distribution on each child also belongs to R , and 2) for each $\mathbf{p} \in R$, the corresponding rule for connecting the children of C uses each edge in $E^{\rightarrow}(C)$ at most $\frac{2}{3} = \frac{4}{3}x_e$ of the time in expectation. The feasible region is given in Definition 17.17. Since R is nonempty, 1) and 2) are sufficient to give the result since we can induce any distribution on the topmost cycle cut $V \setminus \{r\}$.

Definition 17.17 (The Feasible Region). *Let*

$$R = \left\{ (p_1, p_2, p_3, p_4) \in \mathbb{R}_+^4 : p_1 + p_2 + p_3 + p_4 = 1, p_1 + p_2 = \frac{2}{3}, p_2 + p_4 \geq \frac{1}{3} \right\}.$$

See Figure 62 for a visualization of R in a 2-dimensional space.

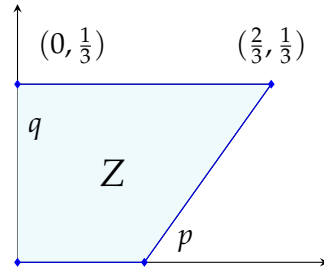


Figure 62: The feasible region of distributions is $R = \{(p, \frac{2}{3} - p, \frac{1}{3} - q, q) : (p, q) \in Z\}$, where Z is the polytope above.

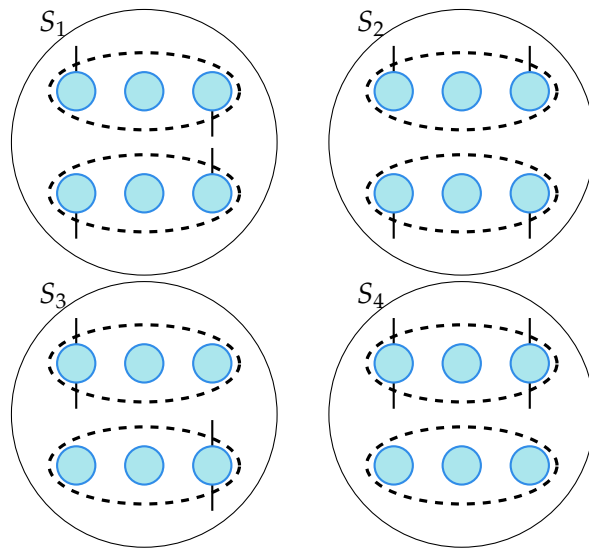


Figure 63: The patterns and how they map to states of a Markov chain. The states are unchanged regardless of the number of children: they are defined only with respect to which of the edges are used by the tour. Note that we ignore doubled edges.

To describe the transitions of the Markov chain, we give (randomized) rules that dictate, for a cycle cut C and a pattern entering it, how to connect up its children. These rules depend on whether C is even or odd. The final form of the Markov chains is illustrated in Figure 65.⁴⁷ The meaning of taking one transition is as follows. Suppose the distribution of patterns entering C is (p_1, p_2, p_3, p_4) , and suppose (q_1, q_2, q_3, q_4) is the resulting distribution after one transition of a Markov chain. What this means is that for each child of C , the distribution of patterns entering it will be **either** (q_1, q_2, q_3, q_4) or (q_2, q_1, q_3, q_4) depending on if the child is straight or twisted, respectively (see Definition 17.16 and Fig. 60). In particular, it can be shown that if (q_1, q_2, q_3, q_4) is the distribution induced on a child which is a straight cycle cut, then (q_2, q_1, q_3, q_4) would be the

⁴⁷In the figure, if there is a variable on an arc, it means that any transition probability in the range of that variable is possible. For example, in P_{even} , we can transition from S_2 to S_1 with probability z for any $z \in [0, 1]$; the transition from S_2 to S_3 then happens with probability $1 - z$.

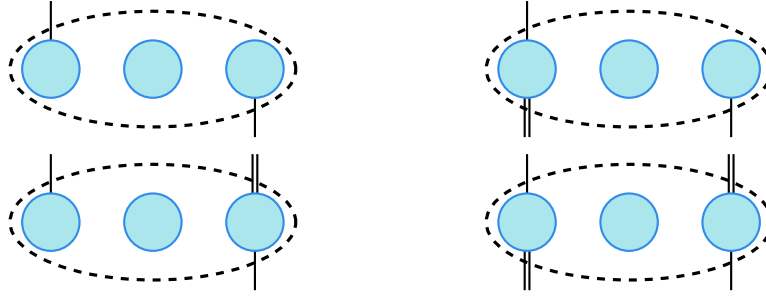


Figure 64: In our illustrations of the patterns entering a given cycle cut, any edge that is not present may either be unused or doubled. Therefore, all four of the given edge configurations are represented by the upper left most state, S_1 .

distribution induced on a child which is a twisted cycle cut. Thus, it is sufficient to check that: i) the distributions induced on straight children lie in the feasible region and ii) if (q_1, q_2, q_3, q_4) is a distribution induced on straight children, then (q_2, q_1, q_3, q_4) is also in the feasible region. This corresponds to the set of distributions induced on the children being symmetric under this transformation.⁴⁸

We ensure that in all cases, each edge in $E^{\rightarrow}(C)$ is used $\frac{1}{2}, \frac{1}{2}, 1, 1$ times in expectation if the pattern entering C belongs to state 1, 2, 3, 4, respectively. (We do not explicitly prove this, but it is straightforward to check when we give the rules from connecting the children.) Therefore, if (p_1, p_2, p_3, p_4) are the probabilities that we are in states 1, 2, 3, 4 respectively, then each edge in $E^{\rightarrow}(C)$ is used exactly

$$\frac{1}{2}p_1 + \frac{1}{2}p_2 + p_3 + p_4 = 1 - \frac{1}{2}(p_1 + p_2)$$

of the time in expectation. Thus, requiring that each edge be used at most $\frac{2}{3} = \frac{4}{3}x_e$ of the time is equivalent to requiring that $p_1 + p_2 \geq \frac{2}{3}$. Note that if $\mathbf{p} \in R$, then $p_1 + p_2 = \frac{2}{3}$ (i.e. each edge is used *exactly* $\frac{2}{3}$ of the time).

In Section 17.6, we illustrate the rules for connecting the children, and show why this leads to the Markov chains in Figure 65. Then in Section 17.7, we give a specific example of how to maintain feasible distributions on all the cuts in the hierarchy by choosing appropriate transition probabilities on the Markov chains in Figure 65. This already gives a $\frac{4}{3}$ -approximation algorithm for half-integral cycle cut TSP, which we describe in Algorithm 6. Finally in Section 17.8, we complete the picture by proving that R (as given in Definition 17.17), is the maximal feasible region of distributions achievable through these chains.

17.6 The Markov Chains

We describe the rules to connect the child cuts (given the edges entering the parent), which will allow us to transition according to the Markov chains depicted in Figure 65.⁴⁹

⁴⁸Note that the feasible region itself is *not* symmetric under this transformation. The distribution induced on the children is thus a symmetric subset of the feasible region.

⁴⁹In the figure, if there is a variable on an arc, it means that any transition probability in the range of that variable is possible. For example, in P_{even} , we can transition from S_2 to S_1 with probability z for any $z \in [0, 1]$; the transition from S_2 to S_3 then happens with probability $1 - z$.

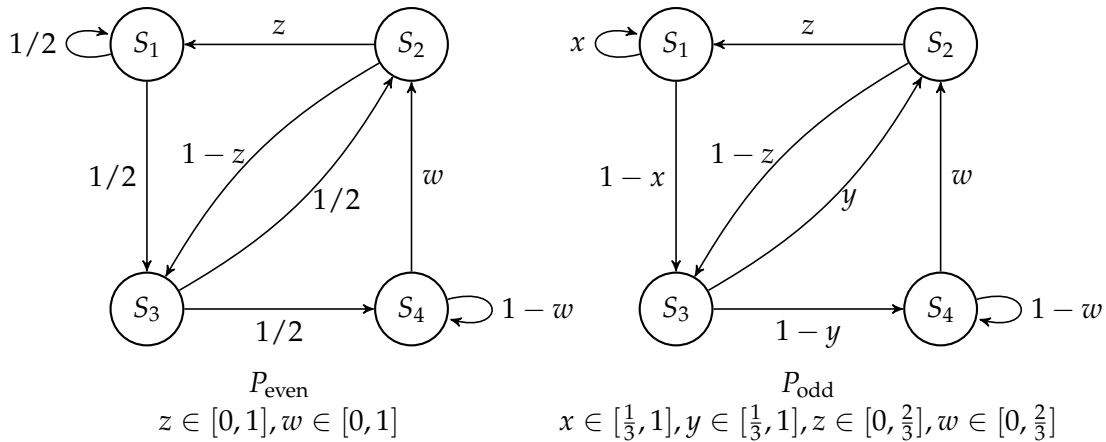


Figure 65: The variables on the arcs indicate that one can feasibly transition according any probability in the range. The ranges are given for the cases where the number of children is 2 (in the even case), and 3 (in the odd case), since these are the most restrictive. In general, when there are k children, the ranges are a superset of those given here, and depend on k .

Proposition 17.18. For any cycle cut $C \in \mathcal{H}$ and any distribution of patterns entering C , there is a way to connect its children so that the induced distribution on each child is given by 1) applying the corresponding Markov chain in Figure 65, and then 2) swapping the first two coordinates if the child is twisted.

Proof. To show that we can always feasibly transition according to the Markov chains in Figure 65, we give rules for connecting the children that result in these transitions. Consider a cycle cut C . We consider two cases, depending on whether C is even or odd.

Case 1: C is even. We consider the states one by one, and argue that the transitions depicted in Figure 65 are achievable.

- State 1.** The transitions out of state 1 are depicted in Figure 66. For each pair of edges inside the cycle cut, we pick one out of the two uniformly at random. This has the effect of transitioning each child to state 3 with probability $\frac{1}{2}$, and to state 1 with probability $\frac{1}{2}$.

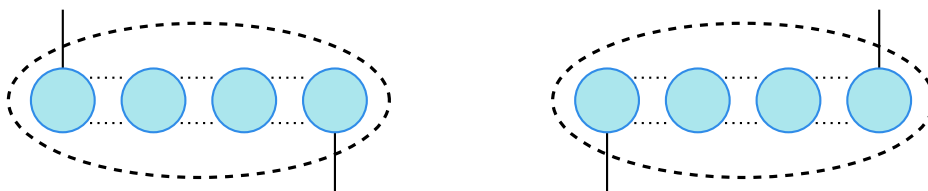


Figure 66: Transition for state 1 in the even case.

- State 2.** The way we transition out of this state is depicted in Figure 67. With probability $\alpha \in [0, 1]$, we alternate taking the top and bottom edges of each pair of edges such that each child transitions back to state 1 with probability 1. (Note that this rule *maximizes* the probability of transitioning back to state 1.) Otherwise, with probability $1 - \alpha$, we make all the children transition to state 3, by always picking the top edge of each pair or the bottom edge of each pair. Thus, the net transition probabilities out of state 2 can be made to be $(\alpha, 0, (1 - \alpha), 0)$, for any $\alpha \in [0, 1]$.

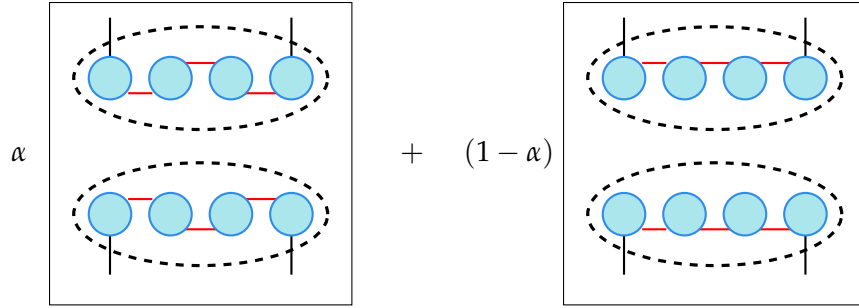


Figure 67: Transition for state 2 in the even case.

3. **State 3.** The transitions out of this state are depicted in [Figure 68](#). To connect up the children in this state, we consider each pair of edges in $E^{\rightarrow}(C)$ independently. Let e, f be such a pair. Then with probability $\frac{1}{2}$, we use both e and f (one copy each), as is illustrated by the solid orange edges. Otherwise, we either double e or double f , with equal probability, shown using the dotted black edges. The net effect is that each child transitions to state 2 with probability $\frac{1}{2}$, and to state 4 with probability $\frac{1}{2}$.

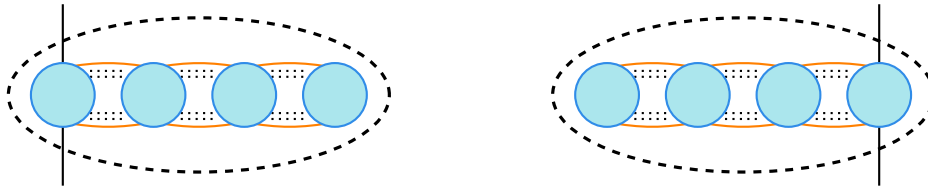


Figure 68: Transition for state 3 in the even case.

4. **State 4.** The transitions out of this state are depicted in [Figure 69](#). With probability α , we make all children transition to state 2 with probability 1. To do this, first suppose C has all 4 single edges entering it (the top picture in the left box). In this case, we consider the pairs of edges in $E^{\rightarrow}(C)$ from left to right, and alternate 1) doubling one of the two edges with equal probability (shown by the dotted black edges), and 2) using both edges (shown by the solid black edges). Because C is even, the rightmost pair of edges ends up falling in case 1) of the alternating rule, and so all children transition to state 2. The case where all the edges entering C are used an even number of times (the bottom picture in the left box) is quite similar, except we begin the alternating rule by using both edges.

On the other hand, with probability $1 - \alpha$, we transition back to state 4 with probability 1. This is accomplished by using each pair of edges in the top case of state 4, and by doubling one edge from each pair uniformly at random in the bottom case of state 4. The net transition probabilities are then $(0, \alpha, 0, 1 - \alpha)$, where α can be any number from 0 to 1.

This shows that the general form of the even chains, as depicted in [Figure 65](#), are achievable.⁵⁰

⁵⁰Actually, note that slightly more general transitions out of states 2 and 3 are possible as a function of k , the number of children. For example, one can show (similarly to the odd case) there are rules for connecting the children that allow the transition from state 3 to state 2 to be any value in the range $[\frac{1}{k}, \frac{k-1}{k}]$. However, the transitions are most restricted when $k = 2$, which results in the Markov chains we presented in [Figure 65](#). (e.g. Note that $\frac{1}{k} = \frac{k-1}{k} = \frac{1}{2}$ if $k = 2$.)

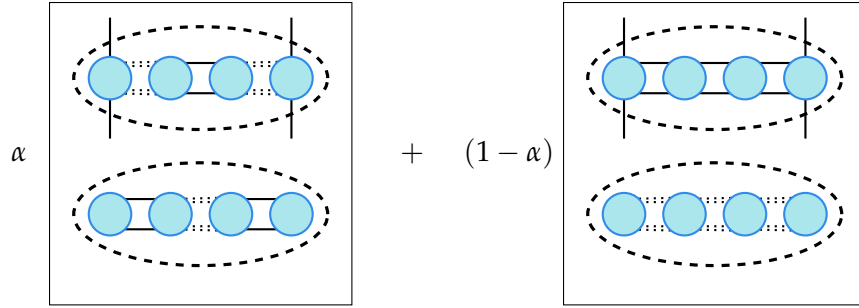


Figure 69: Transition for state 4 in the even case.

Case 2: C is odd. To show that we can feasibly transition according to the Markov chain P_{odd} in Figure 65, we consider each state one by one.

1. **State 1.** The way we transition out of this state is depicted in Figure 70. With probability $\alpha \in [0, 1]$, we alternate taking the top and bottom edges of each pair of edges such that each child transitions back to state 1 with probability 1. (Note that this rule *maximizes* the probability of transitioning back to state 1.) Otherwise, with probability $1 - \alpha$, we choose one of the children uniformly at random to transition to state 1 (the rightmost child in the figure), and make all other children transition to state 3. Note that once we have chosen which child to transition to state 1, there is a unique choice of edges that makes that child transition to state 1 and all other children transition to state 3. The net effect is that the children transition to state 1 with probability $\frac{1}{k}$, and to state 3 with probability $1 - \frac{1}{k}$, where k is the number of children. (This rule *minimizes* the probability of transitioning back to state 1.) Thus, the net transition probabilities out of state 1 are $(\alpha + \frac{1-\alpha}{k}, 0, (1-\alpha) \cdot \frac{k-1}{k}, 0)$. As α ranges from 0 to 1, the probability of transitioning back to state 1 ranges from $\frac{1}{k}$ to 1. Since this range is most restricted when $k = 3$, we conclude that it is always feasible to transition out of state 1 according to the probabilities $(x, 0, 1 - x, 0)$, for any $x \in [\frac{1}{3}, 1]$.

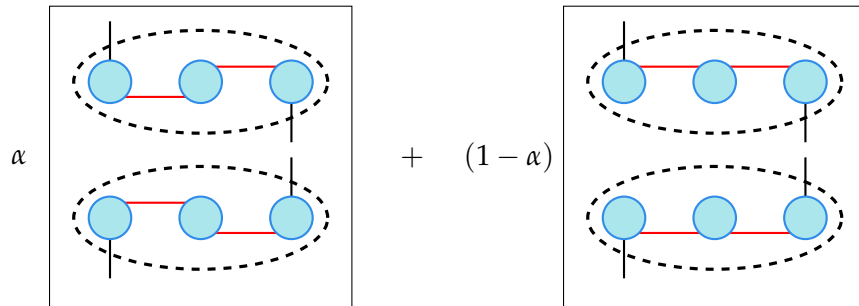


Figure 70: Transitions out of state 1 in the P_{odd} chain. On the left, each child transitions back to state 1. On the right, we pick one of the children uniformly at random to transition back to state 1 (we visualize this to be the rightmost child in the picture), and the remaining children transition to state 3.

2. **State 2.** The way we transition out of state 2 is depicted in Figure 71. With probability α , the net transition probabilities out of state 2 are $(\frac{k-1}{k}, 0, \frac{1}{k}, 0)$, where k is the number of children.

This is accomplished by choosing one of the children uniformly at random to transition to state 3, and the having the remaining children transition to state 1. In more detail, suppose the children are a_1, \dots, a_k from left to right, and suppose we chose a_i to transition to state 3. In the case where the two edges enter C from the top (the top left picture in Figure 71), we go through a_1, a_2, \dots, a_i , alternatingly using the bottom edge from a_1 to a_2 , the top edge from a_2 to a_3 , and so on, until we reach a_i . We then go through a_k, a_{k-1}, \dots, a_i , using the bottom edge from a_k to a_{k-1} , then the top edge from a_{k-1} to a_{k-2} , and so on, until we reach a_i . Since C is odd, a_i will end up either having two edges incident to it from the top or two edges incident to it from the bottom. Thus, a_i transitions to state 3, and all the children except for a_i transition to state 1. The case where the two edges enter C from the bottom (the bottom left picture in Figure 71) is the mirror image of this. Finally, since the child a_i which transitions to state 3 is chosen uniformly at random, the net transition probabilities are $(\frac{k-1}{k}, 0, \frac{1}{k}, 0)$.

On the other hand, with probability $1 - \alpha$ we always transition to state 3, by either always taking the top edge of each pair or the bottom edge of each pair, depending on if the two edges incident to C enter from the top or the bottom, respectively. The overall transition probabilities are therefore $(\alpha \cdot \frac{k-1}{k}, 0, \frac{\alpha}{k} + (1 - \alpha), 0)$. Since the range of transition probabilities is most constrained when $k = 3$; we conclude it is feasible to transition out of state 2 with probabilities $(z, 0, 1 - z, 0)$ for any $z \in [0, \frac{2}{3}]$.

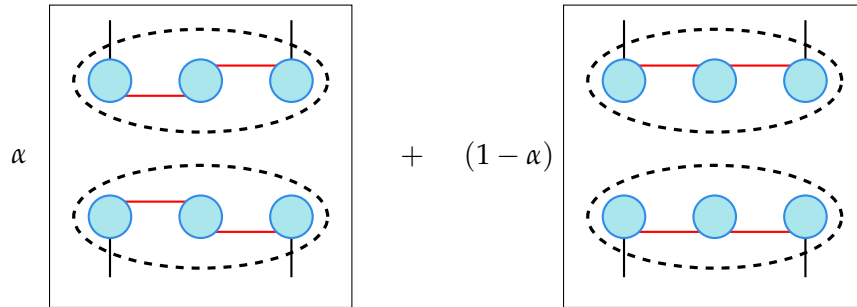


Figure 71: Transitions out of state 2 in the P_{odd} chain. On the left, we pick one of the children uniformly at random to transition to state 3 (we visualize this to be the rightmost child in the picture), and the remaining children transition to state 1. On the right, each child transitions to state 3.

3. **State 3.** The transition out of state 3 is depicted in Figure 72. With probability α , every child transitions to state 2. To do this, we start at the child with two edges entering it (i.e. the leftmost child in the top left picture of Figure 72, and the rightmost child in the bottom left picture of Figure 72), and for each pair of edges in $E^{\rightarrow}(C)$, we alternate 1) doubling one of the two edges with equal probability (shown by the dotted black edges), and 2) using both edges once (shown by the solid black edges).

On the other hand, with probability $1 - \alpha$, we choose one child uniformly at random to transition to state 2, and the remaining children transition to state 4. The way we accomplish this is as follows: Suppose the children are a_1, \dots, a_k from left to right, and suppose we chose a_i to transition to state 2. Then in the case where the two edges enter C from the left (the top right picture in Figure 72), we use both edges in each pair going from a_1, a_2, \dots , all the way

to a_i . Then from a_i to a_k , we double one edge uniformly at random between each pair. The case where the two edges enter C from the left (the bottom right picture in Figure 72) is the mirror image of this; we double one edge uniformly at random between each pair between a_1 and a_i , and then use both edges in each pair going from a_i to a_k . Since the child which transitions to state 2 is chosen uniformly at random, the transition probabilities are then $(0, \frac{1}{k}, 0, \frac{k-1}{k})$. Taking the convex combination of this with the earlier rule which transitions to state 2 deterministically, the net transition probabilities are then $(0, \alpha + \frac{1-\alpha}{k}, 0, (1-\alpha) \cdot \frac{k-1}{k})$. As α ranges from 0 to 1, the transition probability to state 2 ranges from $\frac{1}{k}$ to 1. Since this range is most restricted when $k = 3$, we conclude that it is always feasible to transition with probabilities $(0, y, 0, 1-y)$, for any $y \in [\frac{1}{3}, 1]$.

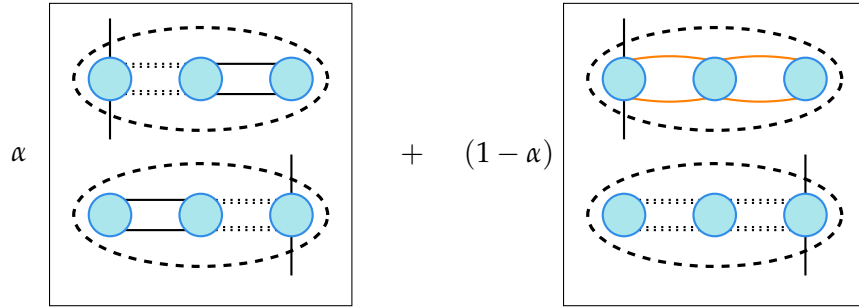


Figure 72: Transitions out of state 3 in the general P_{odd} chain. On the left, every child transitions to state 2. On the right, we pick one child uniformly at random to transition to state 2 (we visualize this to be the right child in the figure), and the remaining children transition to state 4.

4. **State 4.** The transition out of state 4 is depicted in Figure 73. With probability α , we choose one child uniformly at random to transition to state 4, and make the other children transition to state 2. The way we accomplish this is as follows: Suppose the children are a_1, \dots, a_k from left to right, and suppose a_i is chosen to transition to state 4. Then in the case where C has 4 edges entering it (the top left picture in Figure 73), we go through a_1, a_2, \dots, a_i from left to right, and alternating 1) double an edge from each pair of edges uniformly at random and 2) use both edges once, until we reach a_i . We do the same on the other side from a_k, a_{k-1}, \dots , until a_i . This will cause every child except for a_i to be in state 2, and (since k is odd), a_i will be in state 4. The case where every edge entering C is used an even number of times (the bottom picture in the left box of Figure 73) is similar, except we begin the alternation by using both edges. The transition probabilities in this case are $(0, \frac{k-1}{k}, 0, \frac{1}{k})$, where k is the number of children.

On the other hand, with probability $1 - \alpha$, every child transitions to state 4, by either using both edges from each pair if we are in the top case of state 4, or by doubling one edge from each pair uniformly at random if we are in the bottom case of state 4. Overall, the net transition probabilities are then $(0, \alpha \cdot \frac{k-1}{k}, 0, \frac{\alpha}{k} + 1 - \alpha)$. As α ranges from 0 to 1, the transition probability to state 2 ranges from 0 to $\frac{k-1}{k}$. Since this range is most restricted when $k = 3$, we conclude it is always feasible to transition out of state 4 with probabilities $(0, w, 0, 1-w)$ for any $w \in [0, \frac{2}{3}]$.

This finishes the proof of why the P_{odd} Markov chain is achievable for the ranges of probabilities depicted in Figure 65.

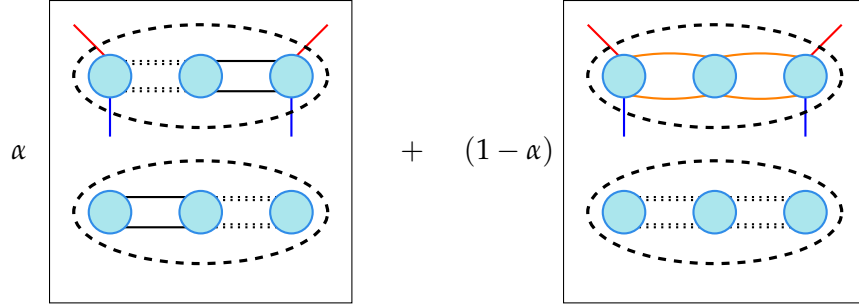


Figure 73: Transitions out of state 4 in the general P_{odd} chain. On the left, we pick one child uniformly at random to transition to state 4 (we visualize this to be the right child in the figure), and the remaining children transition to state 2. On the right, every child transitions to state 4.

□

Note that the rules given in the proof of [Proposition 17.18](#) satisfy the following two invariants: 1) Given that a cycle cut C is in some state, it is equally likely to look like the top picture as the bottom picture of [Figure 63](#), and 2) Each edge in $E^{\rightarrow}(C)$ is used $\frac{1}{2}, \frac{1}{2}, 1, 1$ times in expectation if C is in state 1, 2, 3, 4, respectively.

17.7 Fixed Point and Algorithm

Algorithm 6 A randomized $\frac{4}{3}$ -approximation algorithm for half-integral cycle cut TSP.

Require: A half-integral cycle cut TSP instance $G = (V, E)$ with edge costs c .

Ensure: An Eulerian multi-subgraph T of G with expected cost at most $\frac{4}{3}$ times that of the Subtour LP.

- 1: Compute x , an optimal solution of the Subtour LP.
 - 2: Choose any vertex $r \in V$, and compute the hierarchy \mathcal{H} of critical cuts in $V - \{r\}$.
 - 3: **if** x is not half-integral **or** some cut in \mathcal{H} is not a cycle cut **then**
 - 4: **Fail** ▷ Not a half-integral cycle cut instance
 - 5: **end if**
 - 6: Initialize $T \leftarrow \emptyset$.
 - 7: Sample edges entering $V - \{r\}$ according to the distribution $\mathbf{p} = (\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9})$. Add these edges to T .
 - 8: **for** each cut C in a depth-first search ordering of \mathcal{H} **do**
 - 9: Given the edges in T entering C , sample edges connecting the children of C according to the rules described in [Section 17.6](#) using the specific transition probabilities in [Section 17.7](#). Add these edges to T .
 - 10: **end for**
 - 11: **Return** T .
-

We now give the reader some more intuition by giving a specific example of how to maintain distributions in the feasible region R (as defined in [Definition 17.17](#)), on all the cuts in the hierarchy by choosing appropriate transition probabilities on the Markov chains in [Figure 65](#). This already

gives a $\frac{4}{3}$ -approximation algorithm for half-integral cycle cut TSP, which we describe in Algorithm 6. In Section 17.8, we will extend the ideas here to show that R is the maximal feasible region achievable through our Markov chains.

Specifically, let $\mathbf{p} = (\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9})$ and $\mathbf{q} = (\frac{2}{9}, \frac{4}{9}, \frac{2}{9}, \frac{1}{9})$ (i.e. \mathbf{q} is \mathbf{p} with the first two coordinates swapped). It is easy to check that $\mathbf{p}, \mathbf{q} \in R$. We now show for any half-integral cycle cut instance, it is possible to make the distribution entering any cycle cut to be either \mathbf{p} or \mathbf{q} .

To see this, let C be a cycle cut and suppose C is odd. Set the transition probabilities in P_{odd} to be $x = y = z = w = \frac{2}{3}$. For these probabilities, it is easy to check that $P_{\text{odd}}\mathbf{p} = P_{\text{odd}}\mathbf{q} = \mathbf{p}$.⁵¹ On the other hand, if C is even, setting $z = w = 1$ in P_{even} gives $P_{\text{even}}\mathbf{p} = \mathbf{p}$, and setting $z = \frac{3}{4}, w = 1$ gives $P_{\text{even}}\mathbf{q} = \mathbf{p}$. Thus, as long as the distribution entering C is \mathbf{p} or \mathbf{q} , we can make the distribution on each child of C be either \mathbf{p} (if the child is straight), or \mathbf{q} (if the child is twisted). Since we have freedom in choosing the distribution on the topmost cycle cut $V - \{r\}$, we can simply set it to be \mathbf{p} , and then following the rules given in Section 17.6 with the above transition probabilities will ensure that the distribution on every cut in the hierarchy is either \mathbf{p} or \mathbf{q} .

Proposition 17.19. *Algorithm 6 is a $\frac{4}{3}$ -approximation algorithm for half-integral cycle cut instances of the TSP.*

Proof. By the above reasoning, Algorithm 6 samples from a distribution of Eulerian tours with the property that the distribution of patterns on each cut in the hierarchy is either \mathbf{p} or \mathbf{q} . Under \mathbf{p} and \mathbf{q} , the rules for connecting the children in Section 17.6 guarantee that each edge is used exactly $\frac{4}{3}x_e$ of the time in expectation. The Subtour LP can be solved in polynomial time. The hierarchy of critical cuts can be found efficiently by computing the cactus decomposition of the graph (e.g. [Fle99]). Finally, given the hierarchy, sampling the tour just requires going through the cuts in the hierarchy from the top-down (e.g. using a depth-first search), and for each cut following the rules to sample a multiset of edges inside of it. This takes linear time in the size of the graph. \square

17.8 Characterizing the Feasible Region

We now show that R (as given in Definition 17.17), is the maximal feasible region according to the chains in Figure 65. Recall that by "feasible region", we mean that 1) If the distribution of patterns entering a cycle cut C belongs to R , there is a way to connect up the children of C such that the distribution on each child also belongs to R , and 2) for each $\mathbf{p} \in R$, the corresponding rule for connecting the children of C uses each edge in $E^{\rightarrow}(C)$ at most $\frac{2}{3} = \frac{4}{3}x_e$ of the time in expectation. Informally speaking, R is the set of distributions "that guarantee a $\frac{4}{3}$ -approximation all the way down" the hierarchy of cycle cuts. In particular, the fact that R is nonempty implies the existence of a $\frac{4}{3}$ -approximation algorithm.

Remark 17.20. Note that the distribution $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0)$ lies in R . This is the distribution that, among the four edges entering a cycle cut, uses each pair of edges with equal probability (and possibly doubles zero, one or both of the other edges). We find it nice that such a symmetric distribution is feasible.

Before moving on, note that every distribution in R has a net probability of $\frac{2}{3}$ to be in states 1 or 2, and a probability of $\frac{1}{3}$ to be in states 3 or 4. For any cycle cut C , since all of our rules

⁵¹In fact, it can be checked that for these probabilities, P_{odd} maps *every* distribution (whose first two coordinates sum to $\frac{2}{3}$), to \mathbf{p} .

for connecting the children use each edge in $E \rightarrow (C)$ $\frac{1}{2}$ the time in states 1 and 2, and once in expectation in states 3 and 4, every distribution in R automatically uses each edge exactly $\frac{2}{3} = \frac{4}{3}x_e$ times in expectation. Therefore, checking the feasibility of R boils down to showing that if the distribution of a parent belongs to R , then we can make the distribution of the children also belong to R .

We show this in [Theorem 17.21](#). In other words, R is *sufficient*, in the sense that if the distribution entering a cycle cut belongs to R , then it is possible to get a $\frac{4}{3}$ -approximation all the way down the hierarchy using the Markov chains in [Figure 65](#). We complement this by showing in [Theorem 17.22](#) that R is *necessary*; if the distribution entering a cycle cut does **not** belong to R , then it is impossible to obtain a $\frac{4}{3}$ -approximation using these Markov chains.

Theorem 17.21 (R is sufficient). *If the distribution of patterns entering a cycle cut belongs to R , then there are feasible Markov chains (among the ones shown in [Figure 65](#)) such that the induced distribution entering each child also belongs to R .*

Theorem 17.22 (R is necessary). *Suppose the distribution of patterns entering a cycle cut does **not** belong to R . Then it is not possible to obtain a $\frac{4}{3}$ -approximation using the Markov chains in [Figure 65](#).*

Proof of [Theorem 17.21](#). Let C be any cycle cut in the hierarchy. Suppose the distribution entering C is $(p_1, p_2, p_3, p_4) \in R$. We consider the 2 cases, depending on if C is even or odd. We show that in each case, there is a valid choice of transition probabilities for the corresponding Markov chain (illustrated in [Figure 65](#)), that cause the resulting distribution to also land in R .

Case 1. C is even. Set $w = 1$ and leave z as a variable in P_{even} . Applying the resulting transition matrix to (p_1, p_2, p_3, p_4) yields the distribution

$$\left(\frac{p_1}{2} + zp_2, \frac{p_3}{2} + p_4, \frac{p_1}{2} + (1-z)p_2, \frac{p_3}{2} \right).$$

Let $z = \frac{1}{p_2} \left(\frac{2}{3} - p_4 - \frac{p_1+p_3}{2} \right)$ (this is the value of z that makes the first two components sum to $\frac{2}{3}$). To show that it is valid to set z to this value, we have to show that $z \in [0, 1]$ (since this is the feasible range for z in the P_{even} chain). First, $z \geq 0$ because $\frac{p_1+p_3}{2} + p_4 \leq \frac{2}{3}$.⁵² On the other hand, $z \leq 1$ is equivalent to

$$\frac{p_1}{2} + p_2 + \frac{p_3}{2} + p_4 \geq \frac{2}{3}.$$

Plugging in $p_1 = \frac{2}{3} - p_2$ and $p_3 = \frac{1}{3} - p_4$, this becomes equivalent to $p_2 + p_4 \geq \frac{1}{3}$, which is a constraint defining R . Thus it is valid to set z to this value. Plugging in this value for z and using $p_3 = \frac{1}{3} - p_4$, the resulting distribution becomes

$$(q_1, q_2, q_3, q_4) := \left(\frac{1}{2} - \frac{p_4}{2}, \frac{1}{6} + \frac{p_4}{2}, \frac{1}{6} + \frac{p_4}{2}, \frac{1}{6} - \frac{p_4}{2} \right)$$

It is easy to check that both (q_1, q_2, q_3, q_4) and (q_2, q_1, q_3, q_4) lie in R .

Case 2. C is odd. Set $x = y = z = w = \frac{2}{3}$ in P_{odd} . Applying the resulting transition matrix to (p_1, p_2, p_3, p_4) yields the distribution

$$\left(\frac{2}{3}p_1 + \frac{2}{3}p_2, \frac{2}{3}p_3 + \frac{2}{3}p_4, \frac{1}{3}p_1 + \frac{1}{3}p_2, \frac{1}{3}p_3 + \frac{1}{3}p_4 \right) = \left(\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9} \right).$$

It is easy to check that both $(\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9})$ and $(\frac{2}{9}, \frac{4}{9}, \frac{2}{9}, \frac{1}{9})$ lie in R . □

⁵²Since $p_1 + p_2 = \frac{2}{3}$, $\frac{p_1+p_3}{2} + p_4$ is maximized when $p_1 = \frac{2}{3}$, $p_2 = p_3 = 0$, and $p_4 = \frac{1}{3}$.

Proof of Theorem 17.22. The result follows from the following two statements.

1. Any feasible distribution must have its first two coordinates summing to exactly $\frac{2}{3}$.
2. Given a feasible distribution whose first two coordinates sum to $\frac{2}{3}$, it must in fact be in R .

We will now prove these two statements.

Proof of Statement 1. To show that any feasible distribution must have its first two coordinates summing to exactly $\frac{2}{3}$, consider a general distribution (p_1, p_2, p_3, p_4) . Clearly in order to obtain a $\frac{4}{3}$ -approximation, we must have $p_1 + p_2 \geq \frac{2}{3}$.⁵³ Thus we just need to show $p_1 + p_2$ cannot be strictly larger than $\frac{2}{3}$. To prove this, suppose $p_1 + p_2 > \frac{2}{3}$. We will obtain a contradiction by applying P_{even} twice. Applying P_{even} once with $z = z_1$ and $w = w_1$ (for some $z_1 \in [0, 1], w_1 \in [0, 1]$) to (p_1, p_2, p_3, p_4) , we get the distribution

$$(q_1, q_2, q_3, q_4) := \left(\frac{p_1}{2} + p_2 z_1, \frac{p_3}{2} + p_4 w_1, \frac{p_1}{2} + p_2(1 - z_1), \frac{p_3}{2} + p_4(1 - w_1) \right).$$

In particular, note that $q_1 + q_3 = p_1 + p_2 > \frac{2}{3}$. Applying P_{even} a second time to (q_1, q_2, q_3, q_4) , with $z = z_2$ and $w = w_2$, we get a distribution whose first two coordinates sum to

$$\left(\frac{q_1}{2} + q_2 z_2 \right) + \left(\frac{q_3}{2} + q_4 w_2 \right) \leq \frac{q_1 + q_3}{2} + q_2 + q_4 = 1 - \frac{q_1 + q_3}{2} < \frac{2}{3}.$$

Since the first two coordinates of this distribution sum to strictly less than $\frac{2}{3}$, it cannot give a $\frac{4}{3}$ -approximation.

Proof of Statement 2. Having just shown that any feasible distribution must have its first two coordinates summing to exactly $\frac{2}{3}$, we now show that any such distribution must in fact lie in R . Consider a general distribution whose first two coordinates sum to $\frac{2}{3}$; we can write it as (p_1, p_2, p_3, p_4) where $p_1 + p_2 = \frac{2}{3}$. To show this point lies in R , we just need to show that $p_2 + p_4 \geq \frac{1}{3}$. Applying P_{even} to the input (p_1, p_2, p_3, p_4) , we obtain

$$\left(\frac{1}{2}p_1 + z p_2, \frac{1}{2}p_3 + w p_4, \frac{1}{2}p_1 + (1 - z)p_2, \frac{1}{2}p_3 + (1 - w)p_4 \right).$$

We need the first two components to sum to $\frac{2}{3}$, which means $\frac{1}{2}p_1 + z p_2 + \frac{1}{2}p_3 + w p_4 = \frac{2}{3}$. Plugging in $x, w \leq 1$, we get $\frac{1}{2}p_1 + p_2 + \frac{1}{2}p_3 + p_4 \geq \frac{2}{3}$. Finally, using $p_1 + p_2 + p_3 + p_4 = 1$, we obtain $p_2 + p_4 \geq \frac{1}{3}$. □

Remark 17.23. Algorithm 6 can now be modified to use any initial distribution $\mathbf{p} \in R$, not just $(\frac{4}{9}, \frac{2}{9}, \frac{2}{9}, \frac{1}{9})$. To do this, simply begin by sampling edges entering $V - \{r\}$ according to \mathbf{p} . Then, given the edges entering a parent cycle cut, connect up its children using the rules given in the proof of Proposition 17.18, according to the transition probabilities given in the proof of Theorem 17.21.

Remark 17.24. The "tightness" of the feasible region is respect to our Markov chains in Figure 65. It is possible that there are other patterns / Markov chains that would give rise a larger feasible region.

⁵³Since states 1 and 2 use each edge $\frac{1}{2}$ of the time and states 3 and 4 use each edge once in expectation, $p_1 + p_2 < \frac{2}{3}$ would imply each edge is used strictly more than $1 \cdot \frac{1}{3} + \frac{1}{2} \cdot \frac{2}{3} = \frac{2}{3}$ times in expectation.

17.9 Conclusion and Open Questions

This result leads to several interesting open questions. One such open question is whether our result extends to the case of cycle cuts for non-half-integral solutions. We believe this may be possible through a more refined understanding of the patterns that result from considering non-half-integral solutions.

Clearly a better understanding of what happens in the case of degree cuts is needed to make substantial progress on the overall half-integral case. Recall that in a degree cut, each vertex has degree four, there are no parallel edges, and every non-trivial cut has at least six edges in it. Ideally one would be able to show that any distribution on a parent cut lying in the feasible region of [Fig. 62](#) could be used to induce a distribution on patterns of the children of the degree cut in a subregion of the feasible region with each edge used at most $2/3$ of the time; such a result would lead immediately to a $4/3$ integrality gap for half-integral instances.

18 An Improved Approximation Algorithm for k -ECSM

As mentioned in the introduction, in an instance of the minimum k -edge connected spanning subgraph problem, or k -ECSS, we are given an (undirected) graph $G = (V, E)$ with $n := |V|$ vertices and a cost function $c : E \rightarrow \mathbb{R}_{\geq 0}$, and we want to choose a minimum cost set of edges $F \subseteq E$ such that the subgraph (V, F) is k -edge connected. In its most general form, k -ECSS generalizes several extensively-studied problems in network design such as tree augmentation or cactus augmentation, for which there has been recent exciting progress (e.g. [Fio+18; CTZ21; TZ22; BGA20b]). The k -edge-connected *multi*-subgraph problem, k -ECSM, is a close variant of k -ECSS in which we want to choose a k -edge-connected *multi*-subgraph of G of minimum cost, i.e., we can choose an edge $e \in E$ multiple times. Note that without loss of generality we can assume the cost function c in k -ECSM is a metric, i.e., for any three vertices $x, y, z \in V$, we have $c(x, z) \leq c(x, y) + c(y, z)$.

Around four decades ago, Fredrickson and Jájá [FJ81; FJ82] designed a 2-approximation algorithm for k -ECSS and a $3/2$ -approximation algorithm for k -ECSM. The latter essentially follows by a reduction to the well-known Christofides-Serdyukov approximation algorithm for the traveling salesperson problem (TSP). Over the last four decades, despite a number of papers on the problem [CT00; KR96; Kar99; Gab05; GG08; Gab+09; Pri11; LOS12], the aforementioned approximation factors were only improved in the cases where the underlying graph is unweighted or $k \gg \log n$. Most notably, Gabow, Goemans, Tardos and Williamson [Gab+09] showed that if the graph G is unweighted then k -ECSS and k -ECSM admit $1 + 2/k$ approximation algorithms, i.e., as $k \rightarrow \infty$ the approximation factor approaches 1. The case of k -ECSM where $k = 2$ has received significant attention and (significantly) better than $3/2$ -approximation algorithms were designed for special cases [CR98; BFS16; SV14; Boy+20]. In the general $k = 2$ case, only a $3/2 - \epsilon$ approximation is known where $\epsilon = 10^{-36}$ [KKO22]; we remark this also extends to all even k .

Motivated by [Gab+09], Pritchard posed the following conjecture:

Conjecture 18.1 ([Pri11]). *The k -ECSM problem admits a $1 + O(1)/k$ approximation algorithm.*

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

Theorem 18.2 (Improved Approximation for k -ECSM). *There is a polynomial time randomized algorithm for (weighted) k -ECSM with approximation factor (at most) $1 + \frac{5.06}{\sqrt{k}}$.*

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

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

$$\begin{aligned}
 \min \quad & \sum_{e \in E} x_e c(e) \\
 \text{s.t.} \quad & x(\delta(v)) = k \quad \forall v \in V \\
 & x(\delta(S)) \geq k \quad \forall S \subseteq V, S \neq \emptyset \\
 & x_e \geq 0 \quad \forall e \in E.
 \end{aligned} \tag{79}$$

Note that while in an optimum solution of k -ECSM the degree of each vertex is not necessarily equal to k , since the cost function satisfies the triangle inequality we may assume that in any optimum fractional solution each vertex has (fractional) degree k . This follows from the parsimonious property [GB93b].

We prove [Theorem 18.2](#) by rounding an optimum solution to the above linear program. So, as a corollary we also upper-bound the integrality gap of the above linear program.

Corollary 18.3. *The integrality gap of LP (79) is at most $1 + \frac{5.06}{\sqrt{k}}$.*

18.1 Proof Overview

Before explaining our algorithm, we recall a randomized rounding approach of Karger [Kar99]. Karger showed that if given a solution x to (79) we choose every edge e independently with probability x_e , then the sample is $k - O(\sqrt{k \log n})$ -edge connected with probability close to 1. He then fixes the connectivity of the sample by adding $O(\sqrt{k \log n})$ copies of the minimum spanning tree of G . This gives a randomized $1 + O(\sqrt{\log n/k})$ approximation algorithm for the problem. While this is a very effective procedure for large k , it is not useful when k is a constant or grows slower than $\log n$. We view our result as a refinement of this method using random spanning trees which allows k to be independent of n .

First, we observe that when x is a solution to (79), the vector $2x/k$ is in the spanning tree polytope (after modifying x slightly, see [Corollary 18.7](#) for more details). Following a recent line of works on the traveling salesperson problem [OSS11; KKO21] we write $2x/k$ as a so-called *max-entropy distribution* μ over spanning trees.

Warm-up algorithm and key idea. Our first algorithm, explained in [Section 18.3](#), independently samples $k/2$ spanning trees $T_1, \dots, T_{k/2}$ from μ . Call the (multi-set) union of these trees T^* . Since max entropy distributions are negatively correlated, it is easy to show using Chernoff bounds that any particular cut S has at least $k - O(\sqrt{k \ln k})$ edges with probability at least $1 - O(1/\sqrt{k})$ ⁵⁴. So, in the second step of the algorithm, we add $O(\sqrt{k \ln k})$ additional spanning trees to fix the connectivity of every cut “with high probability.” In other words, after this procedure (which has expected cost $1 + O(\sqrt{\ln k/k})$ times the cost of the LP), *every cut S has at least k edges with probability $1 - O(1/\sqrt{k})$* . One can think of this as a version of Karger’s algorithm which does not fix *every cut* with high probability but instead fixes *each individual cut* with high probability.

A priori, this does not seem like a useful property, because there are exponentially many cuts to bound over. However, we show that (perhaps somewhat surprisingly) there is a way to fix the connectivity of every cut simultaneously by only paying an additional factor of $O(1/k)$ times the cost of the LP in expectation.

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

⁵⁴Of course, one can make this probability much closer to 1 (say $1 - O(1/k^2)$) by only paying a constant factor in the $O(\sqrt{k \ln k})$ term, but it is sufficient to make it $1 - O(1/\sqrt{k})$.

has at least two edges across the cut and therefore it has at least k edges total as desired. This is essentially the key idea of this paper.

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

Main algorithm. The above algorithm is suboptimal in a fairly obvious way. Suppose that a cut S is missing (for example) $\sqrt{k \ln^2 k}$ edges. Then, its connectivity is not fixed by the additional $O(\sqrt{k \ln k})$ spanning trees added by the algorithm. So, the warm-up algorithm simply adds an additional copy of every edge which appeared alone in this cut in its tree. However, it may be that the cut $\delta(S)$ has only one edge in as many as $\Omega(k)$ trees! Therefore, we will add $\Omega(k)$ extra edges to fix the cut instead of just the required $\sqrt{k \ln^2 k}$ edges: a huge overcorrection. [Algorithm 8](#) simply avoids this overcorrection by only adding the number of edges actually missing from the cut, sampling them independently from the set of edges which appeared alone on this cut. It turns out this will let us add only $O(\sqrt{k})$ additional trees instead of $O(\sqrt{k \ln k})$, avoiding the extra $\sqrt{\ln k}$ factor⁵⁵.

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

This bound on the expectation is then enough to complete the argument as follows. Let $n(S)$ be the number of trees T_i for which $|T_i \cap \delta(S)| = 1$. Let e be the unique edge in $\delta(S)$ for some T_i . Then, the probability e needs to double is the expected number of edges missing from $\delta(S)$ divided by $n(S)$. Using the analysis above and that $\mathbb{E}[n(S)] = \Omega(kp_S)$, this can be shown to be $O(\frac{\sqrt{k}e^{-1/p_S}}{kp_S}) = O(1/\sqrt{k})$ in expectation⁵⁶.

Remark 18.4. We note that we expect this algorithm to work for any distribution of spanning trees which is negatively correlated. So, one could for example apply swap rounding [CVZ10] to generate random spanning trees (instead of using the max entropy distribution). However, while the analysis giving $1 + O(\sqrt{\frac{\ln k}{k}})$ approximation in [Section 18.3](#) can easily be modified to give similar bounds for any negatively correlated distribution (since Chernoff bounds can be applied), the proof of [Theorem 18.2](#) in [Section 18.4](#) currently relies on the fact that the distribution of the number of edges in any cut can be written as a sum of independent Bernoullis. So, an extension of [Theorem 18.2](#) to an arbitrary negatively correlated distribution would require a different analysis technique or a generalization of [Lemma 18.18](#).

⁵⁵For a slightly tighter analysis we also include these additional trees in T^* , but this is mostly a superficial difference.

⁵⁶This is not immediate since this is the ratio of the expectations, but we actually need to analyze the expectation of the ratio.

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

18.2 Preliminaries

We will use the same G^0 as in the max entropy algorithm for TSP. We include it here again as a reminder and because the LP solution is slightly different (it is a point in P_{Sub} scaled by $\frac{k}{2}$).

Definition 18.5 (G^0, u_0, v_0). We expand the graph $G = (V, E)$ to a graph G^0 by picking an arbitrary vertex $u \in V$, splitting it into two nodes u_0 and v_0 , and then, for every edge $e = (u, w)$ incident to u , assigning fraction $\frac{x(e)}{2}$ to each of the two edges (u_0, w) and (v_0, w) in G^0 . We set $x((u_0, v_0)) = 0$. Call this expanded graph G^0 , its edge set E^0 , and the resulting fractional solution x^0 , where $x^0(e)$ and $x(e)$ are identical on all other edges. (Note that each of u_0 and v_0 now have fractional degree $k/2$ in x^0 .) In [Corollary 18.7](#) below, we show that $\frac{2}{k} \cdot x^0$ is in the spanning tree polytope for the graph G^0 . For ease of exposition, the algorithm is described as running on G^0 (and spanning trees⁵⁷ of G^0), which has the same edge set as G (when u_0 and v_0 are identified).

We use the notation standard in this thesis, and add the following:

For two sets of edges $F, F' \subseteq E$, we write $F \uplus F'$ to denote the multi-set union of F and F' allowing multiple edges. Note that we always have $|F \uplus F'| = |F| + |F'|$. For a real-valued random variable X , we write $X^+ = \max(0, X)$ to denote the positive part of X .

Definition 18.6 ($S_T(e)$, the ‘‘One-Cut’’ of e in T). For any spanning tree T on the vertex set V_0 , and any edge $e \in T$, let $S_T(e) \subseteq V_0 \setminus \{u_0\}$ be the unique connected component of $T \setminus \{e\}$ which does not contain u_0 . We will call this the **one-cut** of e in T .

Recall that the natural linear programming relaxation for k -ECSM is [\(79\)](#). The solution to this LP can be computed in polynomial time using the ellipsoid method.

18.2.1 Background

The following is an immediate corollary of [Fact 2.3](#), and is also noted in [Corollary 2.4](#).

Corollary 18.7. Let x be the optimal solution of LP [\(79\)](#) and x^0 its extension to G^0 as described in [Definition 18.5](#). Then $\frac{2}{k} \cdot x^0$ is in the spanning tree polytope [\(85\)](#) of G^0 .

This allows us to sample $k/2$ spanning trees independently from $\frac{2}{k} \cdot x$. We will use the max entropy distribution, which is equivalent to a λ -uniform distribution (see [Section 2.3.1](#)).

We will use the fact that any random variable F_T for $F \subseteq E$ is distributed as a $BS(x(F))$. Thus we will study properties of BS random variables. In particular, here we will use Bernoullis as a window into the *variance* of our tree distribution. We start with a fact that comes directly from linearity of expectation and the definition of variance.

⁵⁷A spanning tree in G^0 is a 1-tree in G , that is, a tree plus an edge.

Fact 18.8. If $X = BS(q_1)$ and $Y = BS(q_2)$ are two independent Bernoulli-sum random variables, then $\mathbb{E}[X + Y] = q_1 + q_2$ and $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$.

The following is a consequence of [Theorem 2.17](#).

Corollary 18.9. For any $BS(q)$ with $q \geq 1$, $\mathbb{P}[BS(q) = 0] \leq 1/e$.

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

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

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

Fact 18.10. Given any $0 \leq \epsilon < 1$, let $p_1 \geq p_2 \geq \dots \geq p_m$ be the success probabilities of $m \geq 2$ independent Bernoullis such that $\sum_{i=1}^m p_i = 1 + \epsilon$. Suppose $p_1 \leq \frac{1}{2}(1 + \epsilon)$. Then $\prod_{i=1}^m (1 - p_i) \geq \frac{1}{4}(1 - \epsilon)^2$.

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

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

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

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

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

Theorem 18.11 (Bernstein Inequality for BS Random Variables). Let $X = BS(q)$ be a BS random variable with $\mathbb{E}[X] = q$ and $\text{Var}[X] = \sigma^2$. Then $\forall \lambda > 0$ we have

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

Theorem 18.12 (Multiplicative Chernoff-Hoeffding Bound for BS Random Variables). Let $X = BS(q)$ be a Bernoulli-Sum random variable. Then, for any $0 < \epsilon < 1$ and $q' \leq q$,

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

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

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

18.3 Warm-up: a Simple Algorithm with a $1 + O(\sqrt{\frac{\ln k}{k}})$ -Approximation Ratio.

We first explain a simple algorithm (Algorithm 7) that has a slightly weaker $1 + O(\sqrt{\frac{\ln k}{k}})$ -approximation ratio. We defer our main algorithm (Algorithm 8) and the proof of our main result (Theorem 18.2) to Section 18.4.

Algorithm 7 An Approximation Algorithm for k -ECSM

- 1: Let x^0 be an optimum solution of (79) extended to the graph G^0 as described above.
 - 2: Find weights $\lambda : E^0 \rightarrow \mathbb{R}_{\geq 0}$ such that for any $e \in E^0$, $\mathbb{P}_{\mu_\lambda}[e] \leq \frac{2}{k}x_e^0(1 + 2^{-n})$. ▷ By ??
 - 3: Sample $k/2$ spanning trees $T_1, \dots, T_{k/2} \sim \mu_\lambda$ (in G^0) independently and let $T^* \leftarrow T_1 \uplus \dots \uplus T_{k/2}$.
 - 4: Let B be the disjoint union of an additional $\alpha\sqrt{k/2 - 1}^{58}$ spanning trees sampled from μ_λ . ▷ $\alpha = \Theta(\sqrt{\ln k})$ is a parameter we choose later.
 - 5: **for** $i \in [\frac{k}{2}]$ and $e \in T_i$ **do**
 - 6: **if** $\delta(S_{T_i}(e))_{T^*} < k - \alpha\sqrt{k/2 - 1}$ and $(u_0, v_0) \notin \delta(S_{T_i}(e))$ **then**
 - 7: $F \leftarrow F \uplus \{e\}$.
 - 8: **end if**
 - 9: **end for**
 - 10: **Return** $T^* \uplus B \uplus F$.
-

In the first step of Algorithm 7, we solve (79) on the (slightly) extended graph G^0 . Let x^0 to be the optimal solution. By Corollary 18.7, $(2/k)x^0$ is in the spanning tree polytope. Then in line 2, we find the λ -uniform spanning tree distribution μ_λ where each edge has marginal probability $(2/k)x_e^0$ (ignoring the 2^{-n} relative errors). This step is guaranteed to be done in polynomial time by ??.

In line 3, we independently sample $k/2$ spanning trees⁵⁹ $T_1, \dots, T_{k/2}$ from μ_λ , and let $T^* = T_1 \uplus \dots \uplus T_{k/2}$ to be the (multi-set) union of the samples. It follows that T^* satisfies many desirable properties of the λ -uniform spanning tree distribution:

- i) T^* has the same expectation as the LP solution x^0 , since the marginal probability of each edge is exactly $x^0(e)$;
- ii) For any cut $\delta(S)$ in G , since $\delta(S)_{T^*}$ is distributed as a Bernoulli-sum random variable, Chernoff-type inequalities apply and $\delta(S)_{T^*}$ is highly concentrated around its mean;
- iii) Since T^* is the union of $k/2$ trees, for all cuts we have $\delta(S)_{T^*} \geq k/2$. Moreover, if a cut $\delta(S)$ is not a tree cut of any of the $k/2$ trees, then each of the $k/2$ trees must have at least 2 edges crossing it. Therefore, the number of “bad” cuts of T^* , i.e. those with $\delta(S)_{T^*} < k$, is at most $(n - 1)k/2$ (with probability 1).

To fix the potentially $O(nk)$ bad cuts, we divide them into two types: (i) Cuts S such that $\delta(S)_{T^*} \geq k - \alpha\sqrt{k/2 - 1}$ and (ii) Cuts S where $\delta(S)_{T^*} < k - \alpha\sqrt{k/2 - 1}$, for some $\alpha = \Theta(\sqrt{\ln k})$. We fix all cuts of type (i) by adding $B = \alpha\sqrt{k/2 - 1}$ additional spanning trees as in line 4 of the algorithm (note one could alternatively add $\alpha\sqrt{k/2 - 1}$ copies of the minimum spanning tree as

⁵⁹If k is odd, we sample $\lceil k/2 \rceil$ trees. The bound remains unchanged relative to the analysis we give below as the potential cost of one extra tree is $O(OPT/k)$.

in Karger's algorithm). To fix cuts S of type (ii), we employ the following procedure: for any tree T_i where $\delta(S)_{T_i} = 1$ and S is of type (ii), we add one extra copy of the unique edge of T_i in $\delta(S)$. This procedure is in line 5 to line 9 of the algorithm. Let F be the set of edges added in this step; then the output of our algorithm is $T^* \uplus B \uplus F$ as in line 10.

Now we analyze [Algorithm 7](#).

Theorem 18.13 (Approximation Ratio for [Algorithm 7](#)). *Algorithm 7 outputs a (weighted) k -ECSM with approximation factor (at most) $1 + \sqrt{\frac{8 \ln k}{k}}$.*

We begin by showing that the output of [Algorithm 7](#) is k -edge connected (in G) with probability 1.

Lemma 18.14 (k -Connectivity of the Output). *For any $\alpha \geq 0$, the output of [Algorithm 7](#), $F \uplus B \uplus T^*$ is a k -edge connected subgraph of G .*

Proof. Fix spanning trees $T_1, \dots, T_{k/2}$ in G^0 and a cut S where $(u_0, v_0) \notin \delta(S)$. We show that $\delta(S)_{T^* \uplus F \uplus B} \geq k$. If $\delta(S)_{T^*} \geq k - \alpha \sqrt{k/2 - 1}$, then since B has $\alpha \sqrt{k/2 - 1}$ copies of the minimum spanning tree, $\delta(S)_{T^* \uplus B} \geq k$ and we are done. Otherwise $\delta(S)_{T^*} < k - \alpha \sqrt{k/2 - 1}$. Then, we know that for any tree T_i , if $\delta(S)_{T_i} = 1$, since $(u_0, v_0) \notin \delta(S)_{T_i}$, F has one extra copy of the unique edge of T_i in $\delta(S)$. Therefore, including those cases where an extra copy of the edge e is added, each T_i has at least two edges in $\delta(S)$, so $\delta(S)_{T^* \uplus F} \geq 2 \cdot \frac{k}{2} \geq k$ as desired since there are $\frac{k}{2}$ spanning trees T_i . \square

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

In our algorithm we sample $k/2$ trees $T_1, \dots, T_{k/2}$. The following definition will be useful in this section as well as in [Section 18.4](#). Note it is important to separate the case in which $(u_0, v_0) \in \delta(S)$ for a cut S because in this event, $x^0(\delta(S))$ may be as small as $k/2$, in which case our analysis is not valid. However, since the (u_0, v_0) edge has cost 0, we need not worry about such cuts since they can be trivially satisfied by adding many copies of this edge.

Definition 18.15 (\mathcal{E}_e^i). *For a tree T_i sampled in [Algorithm 7](#) and an edge e , we define \mathcal{E}_e^i to be the event that $e \in T_i \wedge (u_0, v_0) \notin \delta(S_{T_i}(e))$.*

Lemma 18.16. *For any $0 \leq \alpha \leq \sqrt{k}$, $1 \leq i \leq k/2$, and any $e \in E$,*

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

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

Proof. Condition on tree T_i and the event \mathcal{E}_e^i . By [Corollary 2.18](#), for any $1 \leq j \leq k/2$ such that $j \neq i$, $\delta(S_{T_i}(e))_{T_j}$ is a $BS(\mathbb{E}[\delta(S_{T_i}(e))_{T_j}])$ random variable, with $\mathbb{E}[\delta(S_{T_i}(e))_{T_j}] = \frac{2}{k} x(\delta(S_{T_i}(e))) \geq 2$. Also, by definition, $\delta(S_{T_i}(e))_{T_i} = 1$ (with probability 1). Since $T_1, \dots, T_{k/2}$ are independently chosen, by [Corollary 2.18](#) the random variable $\delta(S_{T_i}(e))_{T^*}$ is distributed as $BS(q)$ for $q \geq k - 1$.

Since each T_j has at least one edge in $\delta(S_{T_i}(e))$, $\delta(S_{T_i}(e))_{T^*} \geq k/2$ with probability 1. So, by [Theorem 18.12](#), with $q' = k - 1 - k/2$, when $0 \leq \alpha \leq \sqrt{k/2 - 1}$,

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

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

Proof of [Theorem 18.13](#). Let x be an optimum solution of LP (79). Since the output of the algorithm is always k -edge connected we just need to show $\mathbb{E}[c(F \cup T^* \cup B)] \leq \left(1 + \sqrt{\frac{8 \ln k}{k}}\right) c(x)$. By linearity of expectation,

$$\mathbb{E}[c(T^*)] = \sum_{i \in \lfloor \frac{k}{2} \rfloor} \mathbb{E}[c(T_i)] = \frac{k}{2} \sum_{e \in E} c(e) \mathbb{P}_{\mu_\lambda}[e] = \frac{k}{2} \sum_{e \in E} c(e) \cdot \frac{2}{k} \cdot x_e = c(x),$$

where for simplicity we ignored the $1 + 2^{-n}$ loss in the marginals. On the other hand, since by [Corollary 18.7](#), $\frac{2x}{k}$ is in the spanning tree polytope of G^0 , $c(B) \leq \frac{2c(x)}{k} \cdot \alpha\sqrt{k/2 - 1} \leq \frac{\alpha c(x)}{\sqrt{k/2}}$. It remains to bound the expected cost of F . By [Lemma 18.16](#),

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

Putting these together we get, $\mathbb{E}[c(T^* \cup B \cup F)] \leq (1 + \alpha/\sqrt{k/2} + e^{-\alpha^2/2})c(x)$. Setting $\alpha = \sqrt{\ln \left(\frac{k}{2}\right)}$ finishes the proof. \square

18.4 Improved Algorithm and Proof of Main Theorem

We now introduce our main algorithm that has an approximation ratio of $1 + O\left(\frac{1}{\sqrt{k}}\right)$. Let x^0 be an optimal solution of LP (79) extended to G^0 as above. Our algorithm is given in [Algorithm 8](#). Note for convenience we drop the ceiling in the expression $\frac{k}{2} + \alpha\sqrt{k}$ in all that follows.

Theorem 18.2 (Improved Approximation for k -ECSM). *There is a polynomial time randomized algorithm for (weighted) k -ECSM with approximation factor (at most) $1 + \frac{5.06}{\sqrt{k}}$.*

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

Lemma 18.17 (k -Edge Connectivity of the Output). *The output of [Algorithm 8](#), $F \uplus T^*$ is a k -edge connected subgraph of G .*

Algorithm 8 Algorithm for k -ECSM with Approximation Ratio $1 + O(\frac{1}{\sqrt{k}})$

- 1: Let x^0 be an optimum solution of (79) extended to the graph G^0 as described above.
 - 2: Find weights $\lambda : E^0 \rightarrow \mathbb{R}_{\geq 0}$ such that for any $e \in E^0$, $\mathbb{P}_{\mu_\lambda}[e] \leq \frac{2}{k} \cdot x_e^0 \cdot (1 + 2^{-n})$.
 - 3: Initialize $F \leftarrow \emptyset$.
 - 4: Sample $k/2 + \alpha\sqrt{k}$ spanning trees $T_1, \dots, T_{k/2 + \alpha\sqrt{k}} \sim \mu_\lambda$ (in G^0) independently and let $T^* \leftarrow T_1 \uplus \dots \uplus T_{k/2 + \alpha\sqrt{k}}$.
 - 5: Let $\mathcal{S} \leftarrow \{S_{T_i}(e) : i \in [\frac{k}{2} + \alpha\sqrt{k}], e \in T_i, (u_0, v_0) \notin \delta(S)\}$. $\triangleright \mathcal{S}$ is the set of one-cuts (see Definition 18.6) of $T_i \in T^*$.
 - 6: **for** $S \in \mathcal{S}$ **do**
 - 7: $P(S) := \uplus_{i=1}^{k/2 + \alpha\sqrt{k}} \{e \in T_i : S_{T_i}(e) = S\}$ $\triangleright P(S)$ is the multi-set of $e \in T^*$ with one-cut S .
 - 8: **if** $\delta(S)_{T^*} < k$ **then**
 - 9: **for** $j = 1$ to $k - \delta(S)_{T^*}$ **do**
 - 10: Sample an edge from $P(S)$ uniformly at random and add into F .
 - 11: **end for**
 - 12: **end if**
 - 13: **end for**
 - 14: **Return** $T^* \uplus F$.
-

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

Lemma 18.18 (Variance Upper Bound of Cuts in a Random Spanning Tree). *Let μ_λ be the max-entropy distribution in Algorithm 8. For any $0 \leq p \leq 1$, any $\epsilon \geq 0$ and any $S \subseteq V$ such that $\mathbb{P}_{T \sim \mu_\lambda}[\delta(S)_T = 1] = p$ and $\mathbb{E}_{T \sim \mu_\lambda}[\delta(S)_T] = 2 + \epsilon$, we have $\text{Var}_{T \sim \mu_\lambda}[\delta(S)_T] \leq 4p + 3\epsilon$.*

Proof. By Corollary 2.18, $\delta(S)_T$ is distributed as a BS random variable with $\mathbb{E}_{T \sim \mu_\lambda}[\delta(S)_T] = 2 + \epsilon$ and $\mathbb{P}_{T \sim \mu_\lambda}[\delta(S)_T \geq 1] = 1$. Hence we can write $\delta(S)_T = 1 + X_1 + \dots + X_m$ for some integer $m \geq 2$ ⁶⁰, where X_1, \dots, X_m are independent Bernoulli random variables with success probabilities $b_1 \geq b_2 \geq \dots \geq b_m$. Then from the assumption, $\sum_{i=1}^m b_i = 1 + \epsilon$. By Fact 18.8, we have

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

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

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

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

$$p = \Pr[\forall i, X_i = 0] = (1 - b_1) \prod_{i=2}^m (1 - b_i) \geq (1 - b_1) \left(1 - \sum_{i=2}^m b_i\right) = (1 - b_1) \cdot (b_1 - \epsilon).$$

⁶⁰We remark that the case for $m = 1$ is trivial.

where the fourth step comes from Weierstrass Inequality, and the last step comes from $\sum_{i=1}^m b_i = 1 + \epsilon$. This gives $b_1 \leq \frac{1}{2}(1 + \epsilon - \sqrt{(1 - \epsilon)^2 - 4p})$ or $b_1 \geq \frac{1}{2}(1 + \epsilon + \sqrt{(1 - \epsilon)^2 - 4p})$. Since $4p < 1 - 2\epsilon \leq (1 - \epsilon)^2$, the solutions for b_1 are well-defined. By [Fact 18.10](#) we have $b_1 \geq \frac{1}{2}(1 + \epsilon)$, so $b_1 \geq \frac{1}{2}(1 + \epsilon + \sqrt{(1 - \epsilon)^2 - 4p}) \geq 1 - 2p - \frac{\epsilon}{2}$ (using the square root inequality $\sqrt{1 - x} \geq 1 - x$ for $0 \leq x \leq 1$).

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

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

□

As mentioned in [Section 18.1](#), the following lemma is the key to analyzing [Algorithm 8](#). Roughly speaking, it says that the probability a cut is “bad,” i.e. has fewer than $k - \alpha\sqrt{k}$ edges in T^* , is exponentially small in the probability that $\delta(S)_T = 1$ for $T \sim \mu_\lambda$.

Lemma 18.19 (Expected Augmentation of a Cut). *For any $k \geq 100$ and integer $\alpha \geq 1$ let μ_λ be the max-entropy distribution and T^* be the union of $\frac{k}{2} + \alpha\sqrt{k}$ random spanning trees sampled from μ_λ in [Algorithm 8](#). Then for any $S \subseteq V$,*

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

Proof. We can write the expectation as

$$\begin{aligned} \mathbb{E}_{T^*} \left[(k - \delta(S)_{T^*})^+ \right] &\leq \sum_{i=1}^k \mathbb{P}_{T^*} [\delta(S)_{T^*} \leq k - i] \\ &\leq \sum_{i=0}^{\sqrt{k}} \sum_{j=1}^{\sqrt{k}} \mathbb{P}_{T^*} \left[\delta(S)_{T^*} \leq k - (i \cdot \sqrt{k} + j) \right] \\ &\leq \sum_{i=2\alpha}^{\sqrt{k}} \sqrt{k} \cdot \mathbb{P}_{T^*} \left[\delta(S)_{T^*} \leq k - (i - 2\alpha)\sqrt{k} \right], \end{aligned} \tag{80}$$

where we reindex for convenience in the following argument. Define $\beta \geq 0$ such that $x(\delta(S)) = k + \beta\sqrt{k}$, or equivalently that $\mathbb{E}[\delta(S)_T] = 2(1 + \beta/\sqrt{k})$. By [Lemma 18.18](#), we have

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

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

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

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

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

Applying [Theorem 18.11](#), this is at most

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

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

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

Therefore, we can bound this probability by

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

Therefore,

$$\begin{aligned} \mathbb{E}_{T^*} [(k - \delta(S)_{T^*})^+] &\leq \sqrt{k} \sum_{i=2\alpha}^{\sqrt{k}} \exp\left(\frac{-0.3i}{\max\{k^{-1/2}, p\}}\right) \\ &\stackrel{p, k^{-1/2} \leq 1/e}{\leq} \sqrt{k} \exp\left(\frac{-0.6\alpha}{\max\{k^{-1/2}, p\}}\right) \sum_{i=0}^{\infty} e^{-0.3ei} \\ &\leq 1.8\sqrt{k} \exp\left(\frac{-0.6\alpha}{\max\{k^{-1/2}, p\}}\right) \end{aligned}$$

□

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

Lemma 18.20 (Expected Payment of an Edge for Augmentation). *For any $k \geq 100$ and integer $\alpha \geq 1$, let T^* be the union of $\frac{k}{2} + \alpha\sqrt{k}$ random spanning trees $T_1, \dots, T_{k/2+\alpha\sqrt{k}}$ in [Algorithm 8](#). For any solution x to LP (79),*

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

where F is as defined in [Algorithm 8](#).

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

Let $X_{T_i, e}$ be the number of times that edge e from tree T_i is sampled in [Line 4](#), line 10. We will prove that, letting \mathcal{E}_e^i denote the event $e \in T_i, (u_0, v_0) \notin \delta(S_{T_i}(e))$,

$$\mathbb{E} [X_{T_i, e} \mid \mathcal{E}_e] \leq \frac{7.2e}{\sqrt{k}} e^{-0.6\alpha e} + e^{-\sqrt{k}/2} \quad (81)$$

Then, to prove the lemma,

$$\begin{aligned} \mathbb{E} [c(F)] &= \sum_{e \in E} c(e) \sum_{i=1}^{k/2 + \alpha\sqrt{k}} \mathbb{P} [\mathcal{E}_e^i] \cdot \mathbb{E} [X_{T_i, e} \mid \mathcal{E}_e^i] \\ &\stackrel{(81)}{\leq} \sum_{e \in E} c(e) \sum_{i=1}^{k/2 + \alpha\sqrt{k}} \mathbb{P} [\mathcal{E}_e^i] \left(\frac{7.2e}{\sqrt{k}} e^{-0.6\alpha e} + e^{-\sqrt{k}/2} \right) \\ &\leq \sum_{e \in E} c(e) \left(\frac{k}{2} + \alpha\sqrt{k} \right) \cdot \frac{2}{k} x_e \cdot \left(\frac{7.2e}{\sqrt{k}} e^{-0.6\alpha e} + e^{-\sqrt{k}/2} \right) \\ &= \left(1 + \frac{2\alpha}{\sqrt{k}} \right) \left(\frac{7.2e}{\sqrt{k}} e^{-0.6\alpha e} + e^{-\sqrt{k}/2} \right) c(x) \end{aligned}$$

In the rest of the proof we show (81). First, observe that

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

So, it is enough to upper bound the expected value of the RHS conditioned on $T_i, e \in T_i$. Let $T_{-i}^* = T^* \setminus T_i$ and $P(S)_{-i} := P(S) \setminus T_i$ and note that $|P(S)| = |P(S)_{-i}| + 1$ and $\delta(S)_{T^*} = \delta(S)_{T_{-i}^*} + 1$.

Define $Y_{e, T_i} = \frac{(k - \delta(S)_{T_{-i}^*} + 1)^+}{|P(S)_{-i}| + 1}$. Note that $\mathbb{E} [X_{T_i, e} \mid T_i, e \in T_i] = \mathbb{E} [Y_{e, T_i}]$. We drop the subscript of Y for readability.

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

$$\begin{aligned} \mathbb{E} [Y] &= \mathbb{P} \left[|P(S)_{-i}| \geq \frac{pk}{4} \right] \cdot \mathbb{E} \left[Y \mid |P(S)_{-i}| \geq \frac{pk}{4} \right] \\ &\quad + \mathbb{P} \left[|P(S)_{-i}| < \frac{pk}{4} \right] \cdot \mathbb{E} \left[Y \mid |P(S)_{-i}| < \frac{pk}{4} \right] \end{aligned} \quad (82)$$

We upper bound each term in the RHS separately.

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

Now, if $p > k^{-1/2}$, the RHS is maximized when $p = 1/e$ since $\frac{1}{p} e^{-c/p}$ is an increasing function of p for $0 \leq p \leq 1/e$; note $p \leq 1/e$ (by [Corollary 18.9](#)) and $\alpha \geq 1$. We obtain a bound of $\frac{7.2e}{\sqrt{k}} e^{-0.6\alpha e}$.

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

Next, we bound the second term of (82). First notice $Y \leq 1$ with probability 1; this is because if there are exactly ℓ trees which have (S, \bar{S}) as a one-cut then, $\delta(S)_{T^*} \geq k + 2\alpha\sqrt{k} - \ell$ whereas $|P(S)| = \ell$. Furthermore, $Y \neq 0$ only when $|P(S)_{-i}| \geq 2\alpha\sqrt{k} \geq 2\sqrt{k}$ (for $\alpha \geq 1$). Therefore,

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

To see the last two inequalities notice we must have $p \geq 8k^{-1/2}$ or this event cannot occur. Therefore, since $\mathbb{E} [|P(S)_{-i}|] = p(\frac{k}{2} + \alpha\sqrt{k} - 1)$ the inequality follows by an application of the Chernoff bound (Theorem 18.12).

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

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

Otherwise, suppose $p \leq 2/k$. Then since $\mathbb{E} [|P(S)_{-i}|] \leq 1 + 2\alpha k^{-1/2}$, by Theorem 18.12 (using $\alpha \geq 1$),

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

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

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

which gives (81). Therefore we can bound $\mathbb{E} [Y]$ by $\frac{7.2e}{\sqrt{k}}e^{-0.6\alpha e} + e^{-\sqrt{k}/2}$ for all values of p . \square

Proof of Theorem 18.2. Let x be the optimum solution of (79). From Lemma 18.17, the output of Algorithm 8 is always k -edge connected. Thus it suffices to show that $\mathbb{E} [c(T^* \uplus F)] \leq (1 + \frac{5.06}{\sqrt{k}})c(x)$.

By linearity of expectation,

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

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

Therefore, by Lemma 18.20, $\mathbb{E} [c(T^* \uplus F)]$ is at most

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

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

18.5 Conclusion

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

It still remains open if the integrality gap of the LP is indeed $1 + O(1/k)$ or if there is an approximation algorithm with approximation factor $1 + O(1/k)$. It would also be interesting to find the optimal constant for [Algorithm 8](#). We remark that another open question from Pritchard is whether it can be shown that it is NP-Hard to approximate k -ECSM within a factor of $1 + \Omega(1/k)$.

19 Thin Trees for Laminar Families

19.1 Introduction

Let $G = (V, E)$ be a connected undirected graph. Given any proper $S \subset V$, we use $\delta(S)$ to denote the cut with shores S and $V - S$. A spanning tree T of G is called α -thin if the number of edges of T crossing any given cut of G is at most an α fraction of the total number of edges: $|T \cap \delta(S)| \leq \alpha |\delta(S)|$ for each $S \subseteq V$.

In 2004, Goddyn [God04] made the following conjecture: there exists a function $f : \mathbb{Z}_+ \rightarrow [0, 1]$ with $\lim_{k \rightarrow \infty} f(k)/k = 0$ such that every k -edge-connected graph G has an $f(k)$ -thin spanning tree. This has become known as the *thin tree conjecture*, and it remains open despite substantial efforts.

A natural strengthening of the conjecture, which we will refer to as the *strong thin tree conjecture* makes the same claim, but for $f(k) = C/k$ for some constant C . This conjecture is found explicitly in [Asa+17] and is the best that one could hope for up to constant factors; clearly no k -edge-connected graph has an α -thin tree for any $\alpha < 1/k$. In a different direction, there is also an algorithmic question one can ask: if a thin tree always exists, can we *find* one in polynomial time?

The thin tree conjecture has some nice implications. It implies the *weak 3-flow conjecture* of Jaeger [Jae84]. This has since been resolved, by Thomassen [Tho12], however this would provide an alternate proof. Another application lies in the asymmetric traveling salesman problem (ATSP). As shown by Asadpour, Goemans, Madry, Oveis Gharan and Saberi [Asa+17; OS11], if the constructive form of the strong thin tree conjecture is true, it would yield an $O(1)$ -approximation algorithm to ATSP. This has since been resolved by Svensson, Tarnawski and Végé [STV20] using completely different methods. Nonetheless, a new algorithm stemming from thin trees would be of significant interest. Furthermore, a constant factor approximation algorithm to the *bottleneck* version of the asymmetric traveling salesman problem, where the goal is to minimize the longest edge in the tour rather than the sum, is not known. This would follow from the constructive form of the thin tree conjecture [AKS21].

Although the (strong) thin tree conjecture would no longer imply breakthroughs to these other problems, it remains a natural question in its own right. Turning things around, the positive resolution of these implications can perhaps be viewed as some weak evidence for the conjecture.

For the following discussion, it is useful to observe that the strong thin tree conjecture has the following equivalent formulation. Suppose we are given a graph G as well as a point x in the spanning tree polytope (that is, a convex combination of characteristic vectors of edge sets of spanning trees of G). We say that a spanning tree T is α -thin with respect to x if $|T \cap \delta(S)| \leq \alpha x(\delta(S))$ for every $S \subseteq V$. The conjecture is that there is a universal constant α such that an α -thin tree with respect to x always exists, for any instance and point in the spanning tree polytope. The equivalence follows from the observation that the point x' defined by $x'_e = 2/k$ for all $e \in E$ is in the dominant of the spanning tree polytope for every k -edge-connected graph G , and so there is a point x in the spanning tree polytope with $x_e \leq 2/k$ for all e . An α -thin tree with respect to x is then a $(2\alpha/k)$ -thin tree for the graph.

Progress on the thin tree conjecture. The conjecture is known to hold for some graph classes, most notably planar and bounded genus graphs [OS11]. For general graphs, the best known result is that there always exists an $O(\frac{\text{polyloglog } n}{k})$ -thin tree in any k -connected graph [AO15]. This is non-constructive; constructively, the best known is only $O(\frac{\log n}{\log \log n \cdot k})$ -thinness.

One difficulty with the constructive form of the conjecture is that it's not even clear how to *check* if a given tree T is α -thin, or even $O(\alpha)$ -thin. Nor do we know of a polynomially checkable certificate that can certify thinness. The problem, of course, is that there are an exponential number of cuts to be concerned with. An easier question presents itself: what if we consider an explicitly given family of cuts, and require the thinness condition $|T \cap \delta(S)| \leq \alpha |\delta(S)|$ only for these specific cuts? And one step further: what if we consider a family of cuts with some specific structure?

Explicitly given cut collections. Related questions have been considered from an algorithmic perspective already, independently from the thin tree conjecture. The first class considered was that of *singleton cuts*. Suppose we are given an integer-valued degree bound b_v for each node v of the graph G . The degree bounded spanning tree problem asks for a spanning tree satisfying these bounds, if such a spanning tree exists. This problem is easily seen to be NP-hard, since it captures the question of finding a Hamiltonian path with a specified start and end node. So it is necessary to allow for some relaxation of the degree bounds. Fürer and Raghavachari [FR92] showed that relaxing the degree bounds by 1 additively suffices. That is, they showed how to efficiently find a spanning tree T satisfying $|T \cap \delta(v)| \leq b_v + 1$ for all $v \in V$, if there exists a spanning tree T^* that satisfies the degree bounds exactly.

One can also consider a minimum cost version of the question. Now each edge $e \in E$ has a nonnegative cost $c(e)$, and the goal is to find a *cheapest* spanning tree satisfying the degree bounds (again, assuming one exists). Goemans [Goe06] showed how to efficiently find a spanning tree T which violates the degree bounds by at most an additive 2, and satisfies $c(T) \leq c(T^*)$, where T^* is a minimum cost spanning tree that satisfies all the degree bounds exactly. Singh and Lau [SL15] then showed how to improve the degree violation to just 1, while maintaining the same bound on the cost. They use the method of *iterative relaxation*; we use iterative relaxation as well, so we will discuss this further in the sequel.

That ends the story for degree bounds; what about other families of constraints? So we have a given family \mathcal{F} of subsets, and a “degree bound” b_S for each $S \in \mathcal{F}$. Olver and Zenklus [OZ18] showed how to obtain, constructively, a constant multiplicative violation of all cut constraints if \mathcal{F} is a *chain*; that is, $\mathcal{F} = \{S_1, S_2, \dots, S_t\}$ with $S_1 \subsetneq S_2 \subsetneq \dots \subsetneq S_t$. Linhares and Swamy [LS18] showed that a minimum cost version of this result also holds, if one allows a constant factor approximation in the cost as well as in the cut constraints.

All of these results compare to the natural fractional relaxation. That is, they do not require that there is an actual tree satisfying the degree bounds, merely that there is a point in the spanning tree polytope which does. As such, we can view them in the context of thin trees. They show that weaker versions of the strong thin tree conjecture hold, where the cut bounds are enforced only on singleton cuts, or only on a chain of cuts. We will say that the strong thin tree conjecture holds for a given family \mathcal{F} if given any x in the spanning tree polytope, there is a spanning tree T satisfying $|T \cap \delta(S)| \leq O(1)x(\delta(S))$ for all $S \in \mathcal{F}$.

Our results. Given that the strong thin tree conjecture holds for the family of singletons, and for a chain family, a very natural question presents itself. Suppose that \mathcal{L} is an arbitrary *laminar family* of subsets of V ; that is, for every $S, T \in \mathcal{L}$, $S \cap T$ is either equal to \emptyset , S , or T . Does the strong thin tree conjecture hold for \mathcal{L} ?

We show that this is indeed true. Further, our proof is constructive, and allows for costs. More precisely, given arbitrary nonnegative edge costs, our returned tree has cost within a constant

factor of the cost of the starting fractional solution x .

We briefly sketch our high-level approach, leaving a full overview until [Section 19.6](#). As already mentioned, iterative relaxation has been applied very successfully to the degree-bounded spanning tree problem, so it is a natural candidate approach. However, there is an immediate obstruction. Iterative relaxation for degree-bounded spanning tree is fairly insensitive to the use of the graphic matroid; it works just as well (essentially without changes) if the graphic matroid is replaced by any other matroid.⁶¹ However, the matroid generalization of the laminar-constrained spanning tree problem does *not* have a constant integrality gap, and even obtaining a constant factor multiplicative violation is hard. This was shown by Olver and Zenklusen [[OZ18](#)] already for the chain case. So any successful approach will need to exploit the graphic matroid specifically; it is not clear how to do this directly with iterative relaxation.

We manage to bypass this obstruction and make use of iterative relaxation. We do this by first reducing to a special class of instances that we call \mathcal{L} -aligned, where the fractional solution x has the property that for every set S in the laminar family of constraints \mathcal{L} , the restriction of x to S is a point in the base polytope of the graphic matroid for the graph restricted to S . Our reduction crucially exploits properties of spanning trees, and does not apply to general matroids. We then give an iterative relaxation proof of this \mathcal{L} -aligned case. This part *does* generalize to arbitrary matroids.

Other related work. For laminar families, the most directly comparable work is from 2013 by Bansal, Khandekar, Könemann, Nagarajan, and Peis. They give an *additive* $O(\log n)$ approximation for the laminar constrained spanning tree problem [[Ban+13](#)], improving upon an earlier more general result which given a family of m constraints obtains a violation of $(1 + \epsilon)b + O(\frac{1}{\epsilon} \log m)$ for each bound b [[CVZ10](#)]. As previously mentioned, Olver and Zenklusen [[OZ18](#)] demonstrated a constant factor multiplicative violation for a family of cuts given by a chain. These three results are with respect to the fractional relaxation, and thus also solve the related thin tree problems. Nägele and Zenklusen [[NZ19](#)] demonstrated that in quasi-polynomial time the violation for the chain-constrained spanning tree problem can be improved to a $(1 + \epsilon)$ multiplicative factor, for any $\epsilon > 0$. They further generalize this slightly towards laminar families, by allowing for a family of cuts that form a laminar family of constant *width*, meaning that the maximum number of disjoint sets in the laminar family is bounded by a constant. (Put differently, the number of leaves in the tree representing the laminar family is constant). However, this result is not based on the LP relaxation, and so does not imply anything for the strong thin tree conjecture for chains or constant-width laminar families.

This problem has also been studied for general matroids. Király, Lau and Singh [[KLS12](#)] showed that given a matroid \mathcal{M} and a collection of upper bound constraints, one can achieve an additive violation of $\Delta - 1$ for all constraints, so long as every element of the matroid is in at most Δ constraints. They achieve a similar guarantee if lower bounds (or both lower and upper bounds) are present. Similar results and further generalizations can be found in [[CVZ10](#); [Ban+13](#)].

Pritchard [[Pri11](#)] conjectured that every k -edge-connected graph contains a spanning tree after whose deletion the graph remains $k - f(k)$ connected, where $f(k)$ is any function for which $\lim_{k \rightarrow \infty} f(k)/k = 0$. This can easily be seen as a weakening of the thin tree conjecture. The strong

⁶¹With the notable exception of [[LS18](#)] which solves the bounded degree spanning problem with an additive error of 1 on both lower *and* upper bounds. When translated to the general matroid setting, the additive error is only known to be 2 [[KLS12](#)].

version of this conjecture (which is a consequence of the strong thin tree conjecture) is that $f(k)$ is an absolute constant. Currently the best known bound for this problem (to the best of our knowledge) is $f(k) = \lfloor \frac{k}{2} \rfloor - 1$ by the Nash-Williams theorem [NW61].

There is a natural *spectral* strengthening of thin trees. Let L_H denote the Laplacian of a graph H , and let \preceq denote the Löwner ordering on symmetric matrices⁶². We say T is α -spectrally-thin if $L_T \preceq \alpha L_G$; that is, if $z^T L_T z \leq \alpha z^T L_G z$ for any vector $z \in \mathbb{R}^V$. This is a stronger condition than α -thinness, as can be seen by choosing z to be the characteristic vector of a set $S \subseteq V$. A big advantage of spectral thinness is that it can be efficiently checked. A natural analogue of the strong thin tree conjecture, where connectivity is replaced by the minimum effective conductance, can be derived [HO14] as a consequence of results on the Kadison-Singer problem [MSS13]. This demonstrates that the strong thin tree conjecture holds for edge transitive graphs (or any graph where the minimum edge conductance is within a constant factor of the connectivity). Unfortunately, spectral thinness is too strong a property to directly aid in proving the (strong or weak) thin tree conjecture in general; there are instances where no $o(\sqrt{n}/k)$ -spectrally thin tree exists [HO14; Goe12]. Nonetheless, spectral approaches have been fruitful. The current best result by Anari and Oveis Gharan [AO15] mentioned previously, that $O(\log \log n/k)$ -thin trees exist, makes use of spectral methods in a sophisticated way. Our approach on the other hand is completely combinatorial; we will not make use of any spectral techniques.

19.2 Preliminaries and Results

19.3 Notation

Given a graph $G = (V, E)$ and a subset $S \subseteq V$, let $\delta(S) = \{\{u, v\} : |\{u, v\} \cap S| = 1\}$ denote the set of edges with exactly one endpoint in S . Let $G[S]$ denote the induced graph of G whose vertex set is S , and let $E(S) \subseteq E$ denote the set of edges in $G[S]$. For $\mathcal{P} = \{P_1, \dots, P_k\}$ a partition of a subset of the vertices of G , we let $\delta(\mathcal{P})$ denote the set of edges with endpoints in two different sets P_i . If the choice of G is not clear, we may write, e.g., $\delta_G(S)$ or $\delta_G(\mathcal{P})$.

For any edge weight function $x : E \rightarrow \mathbb{R}$, we write $x(F) := \sum_{e \in F} x(e)$. For $F \subseteq E$, we write $x|_F$ to denote x restricted to F .

19.4 Polyhedral Background

Edmonds [Edm70] gave the following description for the convex hull of the spanning trees of any graph $G = (V, E)$, known as the *spanning tree polytope*.

$$P_{\text{st}}(G) = \{x \in \mathbb{R}_{\geq 0}^E : x(E) = |V| - 1, x(E(S)) \leq |S| - 1 \forall S \subseteq V\}. \quad (83)$$

The following is the natural LP relaxation for the problem given in Definition 19.2.

$$\begin{aligned} \min \quad & \sum_{e \in E} x_e c_e \\ \text{s.t.,} \quad & x(\delta(S)) \leq b_S \quad \forall S \in \mathcal{L}, \\ & x \in P_{\text{st}}(G) \end{aligned} \quad (84)$$

⁶²That is, $A \preceq B$ if $B - A$ is positive semidefinite.

For two $x, x' \in \mathbb{R}^E$, we say x **dominates** x' if $x - x' \geq 0$. Let $P_{\text{st}}^\uparrow(G)$ denote the dominant of the spanning tree polytope of G , that is, the set of points in $x \in \mathbb{R}^E$ which dominate some point in $P_{\text{st}}(G)$. $P_{\text{st}}^\uparrow(G)$ has the following characterization:

$$P_{\text{st}}^\uparrow(G) = \{x \in \mathbb{R}_{\geq 0}^E : x(\delta(\mathcal{P})) \geq |\mathcal{P}| - 1 \quad \forall \text{ partitions } \mathcal{P} \text{ of } V\}. \quad (85)$$

It is well-known that $P_{\text{st}}^\uparrow(G)$ can be separated efficiently.

Theorem 19.1 ([Bar92]). *Given a graph $G = (V, E)$ and a point $x \in \mathbb{R}_{\geq 0}^E$, A partition \mathcal{P} of G minimizing $x(\delta(\mathcal{P})) - (|\mathcal{P}| - 1)$ can be found in polynomial time.*

Suppose $\mathcal{M} = (E, \mathcal{I})$ is a matroid with groundset E and independent sets \mathcal{I} . The *matroid base polytope* of \mathcal{M} , which we will denote $P_{\mathcal{M}}$, is the convex hull of the incidence vectors of all bases of \mathcal{M} . The *rank* of \mathcal{M} , denoted $\text{rank}_{\mathcal{M}}$, is the cardinality of the largest independent set of \mathcal{M} . Given $F \subseteq E$:

1. The *deletion* of F from \mathcal{M} is the matroid on the groundset $\mathcal{M} \setminus F$ with independent sets $\{I \setminus F : I \in \mathcal{I}\}$. If $F = \{e\}$, i.e. it is a singleton, we will use the shorthand $\mathcal{M} - e$.
2. The *restriction* of \mathcal{M} to F , denoted $\mathcal{M}|_F$, is the matroid on groundset F with independent sets $\{I \cap F : I \in \mathcal{I}\}$. This is equivalent to the deletion of $E - F$.
3. The *contraction* of \mathcal{M} by F , denoted \mathcal{M}/F , is the matroid on the groundset $E - F$ with independent sets $\{I \subseteq E - F : I \cup B \in \mathcal{I}\}$, where B is an arbitrary basis of $\mathcal{M}|_F$ (equivalently, an independent set of \mathcal{M} contained in F of largest cardinality). If $F = \{e\}$, i.e. it is a singleton, we will use the shorthand \mathcal{M}/e .

19.5 Our Results

We recall that a family of sets $\mathcal{L} \subseteq 2^V$ is **laminar** if for all $S, T \in \mathcal{L}$, $S \cap T$ is either equal to \emptyset , S , or T .

Definition 19.2 (Laminar constrained spanning tree problem). *Let $G = (V, E)$ be a connected graph, and \mathcal{L} a laminar family on V , with an associated degree bound $b_S \in \mathbb{Z}_{\geq 0}$ for each $S \in \mathcal{L}$. The goal is to find efficiently a spanning tree T for which $|T \cap \delta(S)| \leq \alpha b_S$ for each $S \in \mathcal{L}$, assuming that there does exist a spanning tree T^* satisfying $|T^* \cap \delta(S)| \leq b_S$. In such a case, we say T is an α -approximate solution to the laminar constraints. We assume for convenience that $V \in \mathcal{L}$, though the associated constraint is of course vacuous.*

To solve the above problem, we first determine if LP (84) is feasible, which can be done in polynomial time. If it is not, we may return “no” to the above problem since this would certify that such a tree does not exist. Thus to obtain an α approximation for the problem above, it is enough to obtain an α -thin tree with respect to a solution x of (84).

Definition 19.3 (Laminar α -thin tree for (G, \mathcal{L}, x)). *As input we get a graph $G = (V, E)$, a laminar family \mathcal{L} over V , and a feasible LP solution x to LP (84). Our goal is to find a spanning tree T such that $|T \cap \delta(S)| \leq \alpha x(\delta(S))$ for all $S \in \mathcal{L}$, i.e. a tree that is α thin with respect to x .*

The main result of this paper is the following. We remark it also gives an $O(1)$ approximation in terms of the cost of the tree.

Theorem 19.4. *Given an instance (G, \mathcal{L}, x) , we can in polynomial time find a spanning tree T such that:*

- i) $c(T) \leq (2 + \sqrt{7})c(x) < 5c(x)$, and
- ii) $|T \cap \delta(S)| \leq (2 + \sqrt{7})^2 x(\delta(S)) < 22x(\delta(S))$, i.e., it is a 22-thin tree for (G, \mathcal{L}, x) .

Our theorem can be generalized as follows, which can be used to reduce the cost of the tree arbitrarily close to 2 (at the expense of incurring a larger multiplicative loss).

Theorem 19.5 (Main). *Given an instance (G, \mathcal{L}, x) and any $\eta > 2$, we can in polynomial time find a spanning tree T such that:*

- i) $c(T) \leq \eta c(x)$, and
- ii) $|T \cap \delta(S)| \leq \frac{1}{1-\frac{2}{\eta}}(2\eta + 3)x(\delta(S))$.

We will prove the latter theorem, since the previous theorem follows by setting $\eta = 2 + \sqrt{7}$.

19.6 Proof Overview

A key observation of this paper is the usefulness of the following definition.

Definition 19.6 (\mathcal{L} -aligned). *Given a graph $G = (V, E)$ and a laminar family $\mathcal{L} \subseteq 2^V$ we say a point $x \in P_{\text{st}}(G)$ is \mathcal{L} -aligned if $x|_{E(S)} \in P_{\text{st}}(G[S])$ for all $S \in \mathcal{L}$.*

Note that $G[S]$ should be connected for each set $S \in \mathcal{L}$, otherwise no point can be \mathcal{L} -aligned.

In [Section 19.7](#) and [Section 19.8](#) we show the following two theorems which when combined immediately give [Theorem 19.5](#).

Theorem 19.7 (Laminar thin trees for \mathcal{L} -aligned points). *Given an instance (G, \mathcal{L}, x) for which x is \mathcal{L} -aligned, we can find a tree T of cost at most $c(x)$ in polynomial time for which*

$$|T \cap \delta(S)| \leq 2\lceil x(\delta(S)) \rceil + 1 \leq 2x(\delta(S)) + 3$$

for all $S \in \mathcal{L}$.

Theorem 19.8 (Reduction to \mathcal{L} -aligned points). *For any instance (G, \mathcal{L}, x) and any $\eta > 2$, we can find an instance (G, \mathcal{L}', x') in polynomial time such that:*

- i) x' is \mathcal{L}' -aligned,
- ii) x' is dominated by ηx ,
- iii) If for a spanning tree T there are $\alpha, \beta \geq 0$ such that $|T \cap \delta(S)| \leq \alpha x'(\delta(S)) + \beta$ for all $S \in \mathcal{L}'$, then we have $|T \cap \delta(S)| \leq \frac{1}{1-\frac{2}{\eta}}(\eta\alpha + \beta)x(\delta(S))$ for all $S \in \mathcal{L}$.

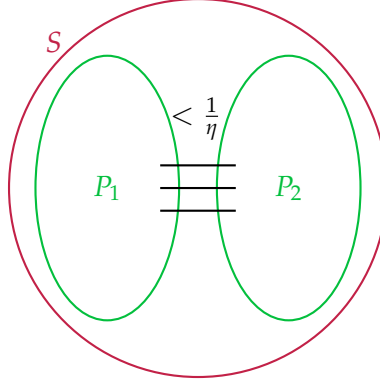


Figure 74: An example of a set which is not η -well-connected (see Definition 19.9). In this case, Algorithm 9 may replace S by P_1 and P_2 in \mathcal{L} .

To obtain our main theorem, given an instance (G, \mathcal{L}) we solve LP (84) to obtain an instance (G, \mathcal{L}, x) . We then apply Theorem 19.8 to obtain a \mathcal{L} -aligned instance (G, \mathcal{L}', x') . Finally, we apply Theorem 19.7 to obtain our tree with the desired properties.

We remark that while Theorem 19.7 can be generalized to hold for any matroid over the edges of a graph and any laminar family (see Section 19.8), Theorem 19.8 cannot be. Olver and Zenklusen [OZ18] showed that there is a matroid and a laminar family of constraints (in fact, their family is a chain, and their matroid simply a partition matroid) with no constant-thin basis, in particular giving a lower bound of $O(\frac{\log n}{\log \log n})$ on the multiplicative violation. Thus it is necessary that one of these two pieces cannot be generalized to all matroids.

Theorem 19.8 is proved via a natural combinatorial procedure which iteratively replaces sets in \mathcal{L} that are far from meeting the criteria $x|_{E(S)} \in P_{\text{st}}(G[S])$ with some partition of them. We first consider the scaling ηx , and show that if $\eta \cdot x|_{E(S)} \in P_{\text{st}}^\uparrow(G[S])$ for all $S \in \mathcal{L}$, then there is a point x' dominated by ηx which is \mathcal{L} -aligned. If not, we iteratively find a minimal cut S for which $\eta \cdot x|_{E(S)} \notin P_{\text{st}}^\uparrow(G[S])$, and then find the partition $\mathcal{P} = \{P_1, \dots, P_k\}$ of S which maximally violates an inequality in $P_{\text{st}}^\uparrow(G[S])$. We then delete S from the laminar family and add P_1, \dots, P_k . We show that $\eta \cdot x|_{E(P_i)} \in P_{\text{st}}^\uparrow(G[P_i])$ for all i . Therefore, by applying this procedure we get closer to obtaining an \mathcal{L} -aligned point. To finish the proof, we show that this process allows us to still effectively maintain (iii) of Theorem 19.8.

Theorem 19.7 (and its generalization to arbitrary matroids) is proved via an iterative relaxation procedure. The criteria that x is \mathcal{L} -aligned is, in some sense, exactly what is needed to make the iterative relaxation procedure work.

19.7 Reduction to \mathcal{L} -aligned Points

The following definition is key to our reduction to \mathcal{L} -aligned points.

Definition 19.9 (Well-connected). *Call a set $S \subseteq V$ η -well-connected if $\eta \cdot x|_{E(S)} \in P_{\text{st}}^\uparrow(G[S])$, i.e., if $\eta x(\delta_{G[S]}(\mathcal{P})) - (|\mathcal{P}| - 1) \geq 0$ for all partitions \mathcal{P} of S .*

We will make use of the following simple fact, that allows us to contract η -well-connected subsets of a given set when evaluating the well-connectedness of a given set S .

Lemma 19.10. Consider a set $S \subseteq V$, and suppose that S_1, \dots, S_r are disjoint subsets of S that are all η -well-connected. Let $G_S = (V_S, E_S)$ be the graph obtained from $G[S]$ after contracting each of S_1, \dots, S_r . Then S is η -well-connected, i.e., $\eta \cdot x_{|E(S)} \in P_{\text{st}}^\uparrow(G[S])$, if and only if $\eta \cdot x_{|E_S} \in P_{\text{st}}^\uparrow(G_S)$.

Proof. Let $y = \eta x_{|E(S)}$.

First, if $y \in P_{\text{st}}^\uparrow(G[S])$, then certainly $y_{|E_S} \in P_{\text{st}}^\uparrow(G_S)$, since given any convex combination of spanning trees of $G[S]$ that dominates y , the same convex combination of the images of these spanning trees upon contracting S_1, \dots, S_r is a convex combination of connected spanning subgraphs of G_S with marginals $y_{|E_S}$.

Conversely, suppose that $y_{|E_S} \in P_{\text{st}}^\uparrow(G_S)$. If $y_{|E_S} = \chi(T)$ for some spanning tree T of G_S , and each $y_{|E(S_i)} = \chi(T_i)$ for some spanning tree T_i of S_i , then the claim is clear; $T \cup T_1 \cup \dots \cup T_r$ is a spanning tree of $G[S]$. But the claim clearly remains true upon taking convex combinations, and moreover taking the dominant of any convex combination. \square

Algorithm 9 Reduction to a new laminar family

- 1: $\mathcal{L}' \leftarrow \emptyset$.
 - 2: **while** \mathcal{L} is nonempty **do**
 - 3: Choose a minimal set $S \in \mathcal{L}$.
 - 4: Let G_S be obtained from $G[S]$ by contracting all the maximal sets in \mathcal{L}' contained in S .
 - 5: Compute a partition \mathcal{P}' of G_S minimizing $\eta x(\delta_{G_S}(\mathcal{P}')) - (|\mathcal{P}'| - 1)$. Let \mathcal{P} be the corresponding partition of S obtained by uncontracting.
 - 6: Delete S from \mathcal{L} and add all parts of \mathcal{P} to \mathcal{L}' .
 - 7: **end while**
 - 8: Return \mathcal{L}' .
-

In this section we prove [Theorem 19.8](#), which heavily relies on [Algorithm 9](#). This algorithm will be used to output the new family \mathcal{L}' in the theorem statement. As such, we first prove some properties of this algorithm.

Lemma 19.11. [Algorithm 9](#) can be implemented in polynomial time.

Proof. In each iteration, $|\mathcal{L}|$ decreases, so there are at most $2|V| - 1$ iterations. Each iteration can be implemented in polynomial time using [Theorem 19.1](#). \square

Lemma 19.12. Consider any graph $G = (V, E)$ and $\eta > 0$. Let \mathcal{P} be a partition of G that minimizes $\eta x(\delta(\mathcal{P})) - (|\mathcal{P}| - 1)$. Then each part of \mathcal{P} is η -well-connected.

Proof. Fix any part $P \in \mathcal{P}$, and consider any partition $\mathcal{Q} = \{Q_1, Q_2, \dots, Q_r\}$ of P . Let \mathcal{P}' be the partition of V' obtained by replacing P with the parts of \mathcal{Q} .

Write $\delta_P(\mathcal{Q})$ for $\delta_{G[P]}(\mathcal{Q})$. Since $|\mathcal{P}'| = |\mathcal{P}| + |\mathcal{Q}| - 1$ and $\eta x(\delta(\mathcal{P}')) = \eta x(\delta(\mathcal{P})) + \eta x(\delta_P(\mathcal{Q}))$, we have

$$\eta x(\delta_P(\mathcal{Q})) - (|\mathcal{Q}| - 1) = \eta x(\delta(\mathcal{P}')) - (|\mathcal{P}'| - 1) - (\eta x(\delta(\mathcal{P})) - (|\mathcal{P}| - 1)).$$

This is nonnegative, by our choice of \mathcal{P} , and so $\eta x \in P_{\text{st}}^\uparrow(G[P])$. \square

Lemma 19.13. The output \mathcal{L}' of [Algorithm 9](#) is a laminar family, and each $S \in \mathcal{L}'$ is η -well-connected.

Proof. We claim that throughout the algorithm, we maintain the invariant that $\mathcal{L} \cup \mathcal{L}'$ is a laminar family, and that each $S \in \mathcal{L}'$ is η -well-connected. Certainly this holds at the start of the algorithm. Consider a partition \mathcal{P}' of G_S generated in step 5. By Lemma 19.12, each part of \mathcal{P}' is η -well-connected. Then since the sets that were contracted in forming G_S are η -well-connected, by Lemma 19.10 all parts of \mathcal{P} are η -well-connected in G . Further, no part of \mathcal{P} crosses a set in \mathcal{L}' , by construction. So the invariant is maintained. \square

The following is the main relevant quality of our reduction.

Lemma 19.14. *Let $S \in \mathcal{L}$ and let \mathcal{L}' be the output of Algorithm 9. Let S_1, \dots, S_ℓ be the unique maximal sets in \mathcal{L}' whose union is S . Then, $\sum_{i=1}^{\ell} x(\delta(S_i)) \leq \frac{1}{1-\frac{2}{\eta}} x(\delta(S)) - \frac{2}{\eta-2}$.*

Proof. Consider the iteration of the algorithm where S is deleted from \mathcal{L} , and a partition \mathcal{P} of S (corresponding to a partition \mathcal{P}' of G_S) is added to \mathcal{L}' . Then $\mathcal{P} = \{S_1, \dots, S_\ell\}$. Note that by the choice of \mathcal{P}' , $\eta x(\delta_{G_S}(\mathcal{P}')) - (|\mathcal{P}'| - 1) \leq 0$ (either \mathcal{P}' is a violated constraint for $P_{\text{st}}^\uparrow(G_S)$, or if G_S is η -well-connected, \mathcal{P}' can be chosen to be the trivial partition of size 1, and equality is attained). Converting this to a statement about \mathcal{P} , we have $\eta x(\delta_{G[S]}(\mathcal{P})) - (|\mathcal{P}| - 1) \leq 0$. Thus

$$\begin{aligned} x(\delta(S)) &= \sum_{i=1}^{\ell} x(\delta(S_i)) - 2x(\delta_{G[S]}(\mathcal{P})) \\ &\geq \sum_{i=1}^{\ell} x(\delta(S_i)) - \frac{2}{\eta} (|\mathcal{P}| - 1) \\ &\geq \left(1 - \frac{2}{\eta}\right) \sum_{i=1}^{\ell} x(\delta(S_i)) + \frac{2}{\eta} \quad (\text{as } x(\delta(S_i)) \geq 1 \text{ for each } S_i). \end{aligned}$$

The claim follows. \square

We now prove Theorem 19.8.

Theorem 19.8 (Reduction to \mathcal{L} -aligned points). *For any instance (G, \mathcal{L}, x) and any $\eta > 2$, we can find an instance (G, \mathcal{L}', x') in polynomial time such that:*

- i) x' is \mathcal{L}' -aligned,
- ii) x' is dominated by ηx ,
- iii) *If for a spanning tree T there are $\alpha, \beta \geq 0$ such that $|T \cap \delta(S)| \leq \alpha x'(\delta(S)) + \beta$ for all $S \in \mathcal{L}'$, then we have $|T \cap \delta(S)| \leq \frac{1}{1-\frac{2}{\eta}} (\eta\alpha + \beta) x(\delta(S))$ for all $S \in \mathcal{L}$.*

Proof. First, apply Algorithm 9 to \mathcal{L} to obtain a new family \mathcal{L}' (which requires only polynomial time by Lemma 19.11). By Lemma 19.13, \mathcal{L}' is a laminar family of η -well-connected sets.

We now show that ηx dominates a point x' which is \mathcal{L}' -aligned, giving i) and ii). Let $G_S = (V_S, E_S)$ denote the graph obtained by restricting to $S \in \mathcal{L}'$ and contracting all children in \mathcal{L}' . By definition of η -well-connected, for any $S \in \mathcal{L}'$, $\eta x|_{E_S} \in P_{\text{st}}^\uparrow(G_S)$. It follows that for every $S \in \mathcal{L}'$ we can find $y_S \in P_{\text{st}}(G_S)$ with $y_S \leq \eta x|_{E_S}$. Combining y_S for each S , we obtain $x' \in P_{\text{st}}(G)$ with $x' \leq \eta x$, and where x' is \mathcal{L} -aligned.

It remains to show (iii). Fix some $S \in \mathcal{L}$. The algorithm replaces S by some partition S_1, S_2, \dots, S_ℓ of S in \mathcal{L}' . Then we have

$$\begin{aligned}
|T \cap \delta(S)| &\leq \sum_{i=1}^{\ell} |T \cap \delta(S_i)| && \text{(since } \bigcup_{i=1}^{\ell} \delta(S_i) \subseteq \delta(S)\text{)} \\
&\leq \sum_{i=1}^{\ell} (\alpha x'(\delta(S_i)) + \beta) && \text{(by assumption)} \\
&\leq \sum_{i=1}^{\ell} (\eta \alpha x(\delta(S_i)) + \beta) && (x' \leq \eta x) \\
&\leq (\eta \alpha + \beta) \sum_{i=1}^{\ell} x(\delta(S_i)) && \text{(since } x(\delta(S_i)) \geq 1 \text{ for all } S_i\text{).}
\end{aligned}$$

By [Lemma 19.14](#), $\sum_{i=1}^{\ell} x(\delta(S_i)) \leq \frac{1}{1-\frac{2}{\eta}} x(\delta(S))$. The claim follows. \square

19.8 Laminar thin trees for \mathcal{L} -aligned points via iterative relaxation

We will now prove [Theorem 19.7](#), or rather a generalization of it where the graphic matroid is replaced by an arbitrary matroid. First, we define the obvious generalization of \mathcal{L} -aligned for a point in the base polytope of a matroid \mathcal{M} .

Definition 19.15. *Given a graph $G = (V, E)$, a matroid \mathcal{M} with groundset E , and a laminar family \mathcal{L} of G , we say that a point $x \in P_{\mathcal{M}}$ is \mathcal{L} -aligned if $x(E(S)) = \text{rank}_{\mathcal{M}}(S)$ for all $S \in \mathcal{L}$.*

(In the case where \mathcal{M} is a graphic matroid, this is just slightly different from the previous definition, if some sets in \mathcal{L} are not connected. The previous definition did not allow for any \mathcal{L} -aligned points in this case, but here it is possible. This relaxation of the definition is irrelevant; there is no real reason to consider disconnected sets in \mathcal{L} , since they could simply be split into their connected components.)

The following is the primary reason it is useful for a point x to be \mathcal{L} -aligned in the iterative relaxation process.

Lemma 19.16. *Let x be \mathcal{L} -aligned. Let $S \in \mathcal{L}$ and let $S_1, \dots, S_k \in \mathcal{L}$ such that $S_i \cap S_j = \emptyset$ for all $1 \leq i, j \leq k, i \neq j$. Let $G_S = (V_S, E_S)$ be the graph arising from contacting S_1, \dots, S_k in the graph $G[S]$. Then, $x(E_S)$ is an integer.*

Proof. Since $x(E(S)) = \text{rank}_{\mathcal{M}}(S)$, it is an integer. Similarly, $x(E(S_i))$ is an integer for all $1 \leq i \leq k$. However $E_S = E(S) \setminus (\cup_{i=1}^k E(S_i))$, from which the claim follows. \square

Next, we define the notion of a *matroid* (rather than a point) being \mathcal{L} -aligned.

Definition 19.17. *Given a graph $G = (V, E)$, a matroid \mathcal{M} with groundset E , and a laminar family \mathcal{L} of G , we say that \mathcal{M} is \mathcal{L} -aligned if for any basis B of \mathcal{M} , and every $S \in \mathcal{L}$, $B \cap E(S)$ is a basis of $\mathcal{M}|_{E(S)}$.*

The relationship between the notion of a matroid being \mathcal{L} -aligned, and a point $x \in P_{\mathcal{M}}$ being \mathcal{L} -aligned, is captured by the following lemma.

Lemma 19.18. *A matroid \mathcal{M} is \mathcal{L} -aligned if and only if for every point $x \in P_{\mathcal{M}}$, x is \mathcal{L} -aligned.*

Proof. First suppose \mathcal{M} is \mathcal{L} -aligned and let $x \in P_{\mathcal{M}}$. Then, we can write x as a convex combination of some bases B_1, \dots, B_k of \mathcal{M} . Since \mathcal{M} is \mathcal{L} -aligned, $B_i \cap E(S)$ is a basis of $\mathcal{M}|_{E(S)}$ for all $S \in \mathcal{L}$. Thus $|B_i \cap E(S)| = \text{rank}(\mathcal{M}|_{E(S)})$ for all i . It follows that $x(E(S)) = \text{rank}(\mathcal{M}|_{E(S)}) = \text{rank}_{\mathcal{M}}(S)$ for all $S \in \mathcal{L}$ as desired, demonstrating that x is \mathcal{L} -aligned.

For the other direction, suppose every point $x \in P_{\mathcal{M}}$ is \mathcal{L} -aligned. Then for any basis B of \mathcal{M} , by taking x to be the characteristic vector of B , we have $|B \cap E(S)| = x(E(S)) = \text{rank}(\mathcal{M}|_{E(S)})$. Thus \mathcal{M} is \mathcal{L} -aligned. \square

In the previous section, we saw how to reduce to the case where x is a point in the base polytope of the graphic matroid that is \mathcal{L} -aligned. It will be more convenient for our purposes to work with a matroid that is \mathcal{L} -aligned; this is a stronger property that will ensure that all fractional points we consider later in the iterative relaxation algorithm are all \mathcal{L} -aligned as well. We can ensure this by *refining* the matroid, in the sense defined in [Lin+20].

Definition 19.19. *Given a matroid \mathcal{M} and a nonempty proper subset R of the groundset, the refinement of \mathcal{M} with respect to R is the matroid \mathcal{M}' obtained as the direct sum of $\mathcal{M}|_R$ and \mathcal{M}/R .*

Note that if \mathcal{M}' is a refinement of \mathcal{M} , then every base of \mathcal{M}' is a base of \mathcal{M} . It is easy to show that for $R \subseteq E$ with $x(R) = \text{rank}_{\mathcal{M}}(R)$, x remains in the base polytope of the matroid obtained by refining \mathcal{M} with respect to R (see [Lin+20] for details). As such, given a point $x \in P_{\mathcal{M}}$ that is \mathcal{L} -aligned, we can repeatedly refine \mathcal{M} by each set of \mathcal{L} in turn, to obtain a new matroid \mathcal{M}' such that $x \in P_{\mathcal{M}'}$ and \mathcal{M}' is \mathcal{L} -aligned. For \mathcal{M} the graphic matroid, this refinement procedure corresponds to taking \mathcal{M}' to be the direct sum of graphic matroids on G_S for each $S \in \mathcal{L}$.

So we consider the generalization of the laminar thin tree problem to matroids, under the restriction that the matroid is aligned with the laminar family. An instance of the problem is defined by a graph $G = (V, E)$, a matroid \mathcal{M} with groundset E , and a laminar family \mathcal{L} with degree bounds b_S for $S \in \mathcal{L}$, such that \mathcal{M} is \mathcal{L} -aligned. Edge costs c_e may also be given. The goal is to find a minimum cost basis of \mathcal{M} satisfying the cut constraints, if a solution exists.

The following LP is the natural relaxation that we will use. Note that since \mathcal{M} is \mathcal{L} -aligned, no explicit additional constraints on x are required; *any* feasible solution must satisfy $x|_{E(S)} \in P_{\mathcal{M}|_{E(S)}}$, and thus must be \mathcal{L} -aligned.

$$\begin{aligned} \min \quad & \sum_{e \in E} x_e c_e \\ \text{s.t.} \quad & x(\delta(S)) \leq b_S \quad \forall S \in \mathcal{L}, \\ & x \in P_{\mathcal{M}}. \end{aligned} \tag{86}$$

Theorem 19.20 (Laminar-constrained matroid basis). *Given an instance $(G, \mathcal{M}, \mathcal{L}, b)$ in which \mathcal{M} is \mathcal{L} -aligned, and where the LP relaxation (86) has a feasible solution x , we can find a basis T of \mathcal{M} in polynomial time for which $c(T) \leq c(x)$ and $|T \cap \delta(S)| \leq 2b_S + 1$ for all $S \in \mathcal{L}$.*

[Theorem 19.7](#) is an immediate consequence, by first refining the graphic matroid as described above.

The algorithm we will use to prove this theorem is shown in [Algorithm 10](#). Our algorithm follows the usual iterative relaxation recipe: it ignores edges set to 0 and 1 and then drops constraints which are close to being satisfied. We have one non-standard step which drops a set in \mathcal{L} if it is *approximately implied* by its immediate parent or child in the family of tight constraints. This non-standard step is what leads to a multiplicative violation instead of an additive one.

Algorithm 10 Procedure LAMCONSTRAINEDBASIS, used to demonstrate [Theorem 19.20](#).

Require: Instance $(G = (V, E, c), \mathcal{M}, \mathcal{L}, b)$ where \mathcal{M} is tight for \mathcal{L} and (86) is feasible.

Ensure: Basis B of \mathcal{M} .

- 1: If $E = \emptyset$, **return** \emptyset .
 - 2: Let x be a basic optimal solution to (86).
 - 3: If there is an edge e with $x_e = 0$, **return** LAMCONSTRAINEDBASIS($G - e, \mathcal{M} - e, \mathcal{L}, b$).
 - 4: If there is an edge e with $x_e = 1$, **return** $\{e\} \cup$ LAMCONSTRAINEDBASIS($G - e, \mathcal{M}/e, \mathcal{L}, b'$), where $b'_S = b_S$ if $e \notin \delta(S)$, and $b'_S = b_S - 1$ if $e \in \delta(S)$.
 - 5: Let $\mathcal{L}_{\text{tight}}$ be the set of cuts $S \in \mathcal{L}$ with $x(\delta(S)) = b_S$.
 - 6: If there is a set $S \in \mathcal{L}_{\text{tight}}$ for which either $\sum_{e \in \delta(S)} (1 - x_e) < 3$, or there is an $S' \neq S \in \mathcal{L}_{\text{tight}}$ with $\delta(S') \subseteq \delta(S)$ and $\sum_{e \in \delta(S) - \delta(S')} (1 - x_e) < 2$, then **return** LAMCONSTRAINEDBASIS($G, \mathcal{M}, \mathcal{L} - \{S\}, b$).
 - 7: **return** "Fail". ▷ Should not reach this line
-

If a recursive call to LAMCONSTRAINEDBASIS returns "Fail", then we consider that the result of the procedure as a whole is also "Fail". We also note that if LAMCONSTRAINEDBASIS is recursively called in any of steps 3, 4 or 6, the required properties of the input to the recursive call are satisfied. In particular, (86) is feasible. For steps 3 and 4, $x_{|E-e}$ is feasible for the smaller instance; for step 6, simply x is. With this in mind, LAMCONSTRAINEDBASIS is well-defined.

We first show that as long as the algorithm does succeed, the returned basis obeys the theorem statement.

Lemma 19.21. *If [Algorithm 10](#) does not return "Fail", the returned set B is a basis and obeys $c(B) \leq c(x)$ and $|B \cap \delta(S)| \leq 2b_S + 1$ for all $S \in \mathcal{L}$.*

Proof. We prove the claim by induction on $|E| + |\mathcal{L}|$. The claim is trivially true if $E = \emptyset$.

So suppose the claim holds for all smaller values of $|E| + |\mathcal{L}|$. If $x_e = 0$ for some e in step 3, then the claim is immediate; as long as the recursive call succeeds, returning a basis B' of $\mathcal{M} - e$ approximately satisfying the constraints, then $B = B'$ is of course a basis of \mathcal{M} still approximately satisfying the constraints. Furthermore, since $c(B') \leq c(x')$ where x' is a basic optimal solution to the problem on $\mathcal{M} - e$, and $x_{|E-\{e\}}$ is feasible for the problem on $\mathcal{M} - e$, $c(B) = c(B') \leq c(x') \leq c(x)$. If $x_e = 1$ for some e in step 4, and the recursive call succeeds and returns a basis B' of \mathcal{M}/e , then $B := B' \cup \{e\}$ is a basis of \mathcal{M} . Further, for any set $S \in \mathcal{L}$ with $e \notin \delta(S)$, we have

$$|B \cap \delta(S)| = |B' \cap \delta(S)| \leq 2b'_S + 1 = 2b_S + 1.$$

On the other hand if $e \in \delta(S)$, we have

$$|B \cap \delta(S)| = |B' \cap \delta(S)| + 1 \leq 2b'_S + 2 < 2b_S + 1.$$

Finally, since $c(B') \leq c(x')$ where x' was a basic optimal solution to the problem on \mathcal{M}/e , and $x_{|E-\{e\}}$ is feasible for the problem on \mathcal{M}/e , $c(B) = c(B') + c(e) \leq c(x') + c(e) \leq c(x)$.

It remains to consider the situation where we drop a constraint in step 6. Suppose a set $S \in \mathcal{L}_{\text{tight}}$ is dropped because $|\delta(S)| - x(\delta(S)) = \sum_{e \in \delta(S)} (1 - x_e) < 3$. Since the constraint is tight, we deduce that $|\delta(S)| - b_S < 3$, and so $|\delta(S)| \leq b_S + 2 \leq 2b_S + 1$ as desired.

Now suppose $S \in \mathcal{L}_{\text{tight}}$ is dropped because there is an $S' \neq S \in \mathcal{L}_{\text{tight}}$ with $\delta(S') \subseteq \delta(S)$ and $\sum_{e \in \delta(S) - \delta(S')} (1 - x_e) < 2$. By tightness, $x(\delta(S) - \delta(S')) = x(\delta(S)) - x(\delta(S')) = b_S - b_{S'}$ is an

integer. Note that either $\delta(S') = \delta(S)$, in which case clearly we can drop the duplicate constraint, or $b_S > b_{S'}$; assume the latter. We have $|\delta(S) - \delta(S')| \leq 1 + b_S - b_{S'}$. Suppose B is any basis satisfying $|B \cap \delta(S')| \leq 2b_{S'} + 1$. Then

$$\begin{aligned} |B \cap \delta(S)| &\leq |B \cap \delta(S')| + |\delta(S) - \delta(S')| \\ &\leq (2b_{S'} + 1) + 1 + b_S - b_{S'} \\ &= b_{S'} + 2 + b_S \\ &\leq (b_S - 1) + 2 + b_S \leq 2b_S + 1. \end{aligned}$$

Of course, dropping a constraint can only decrease the cost of a basic optimal solution to (86), so $c(B) \leq c(x)$ is immediate by induction in this case. \square

Now we are ready to prove the theorem.

Proof of Theorem 19.20. By the above lemma, it is enough to prove that the algorithm succeeds. For this, it suffices to show that whenever the preconditions of LAMCONSTRAINEDBASIS are satisfied, the procedure never reaches step 7.

Suppose for a contradiction that we do reach step 7. By assumption, none of the constraints defining the extreme point x are of the form $x_e = 0$ or $x_e = 1$, so they all come from tight cut constraints and tight matroid constraints. Let $\mathcal{C}_{\text{basis}} = \{C_1, C_2, \dots, C_r\}$, with $C_1 \subsetneq C_2 \subsetneq \dots \subsetneq C_r \subseteq E$ and $\mathcal{L}_{\text{basis}} \subseteq \mathcal{L}_{\text{tight}}$ be such that the constraints $x(\delta(S)) = b_S$ for $S \in \mathcal{L}_{\text{basis}}$ and $x(C) = \text{rank}_{\mathcal{M}}(C)$ for $C \in \mathcal{C}_{\text{basis}}$ are a collection of linearly independent tight constraints defining x . Moreover, choose this basis of tight constraints in such a way that $|\mathcal{C}_{\text{basis}}|$ is as large as possible. The fact that the tight matroid constraints form a chain follows from standard uncrossing arguments (see [Sch03] Chapter 41 or [KLS12]). Since there are precisely $|E|$ defining constraints, we have

$$|E| = |\mathcal{L}_{\text{basis}}| + |\mathcal{C}_{\text{basis}}|.$$

We note that since \mathcal{M} is \mathcal{L} -aligned, the maximality of $\mathcal{C}_{\text{basis}}$ ensures that $E(S) \in \text{span}(\mathcal{C}_{\text{basis}})$ for each $S \in \mathcal{L}$.

Assign 1 splittable token to each $e \in E$; our goal will be to assign these tokens to the constraints of $\mathcal{L}_{\text{basis}}$ and $\mathcal{C}_{\text{basis}}$ so that each tight constraint gets 1 token, and there is something left over. This will be our desired contradiction.

We will assign x_e tokens to C_i for each $e \in C_i - C_{i-1}$. Since $0 < x_e < 1$ for each e , and $x(C_i)$ and $x(C_{i-1})$ are both integers with $x(C_{i-1}) < x(C_i)$, we can deduce that $x(C_i - C_{i-1}) \geq 1$.

Now each edge has $1 - x_e$ tokens remaining. Our token assignment scheme will be as follows. We start with an assignment that is very reminiscent of the scheme for degree bounded spanning trees [SL15]. For each $e = \{u, v\}$, we assign $(1 - x_e)/2$ tokens to the smallest set in $\mathcal{L}_{\text{basis}}$ containing u , and $(1 - x_e)/2$ tokens to the smallest set in $\mathcal{L}_{\text{basis}}$ containing v . After this, we work bottom up on $\mathcal{L}_{\text{basis}}$, and if $S \in \mathcal{L}_{\text{basis}}$ has strictly more than the 1 token needed, we assign the excess to its parent in $\mathcal{L}_{\text{basis}}$.

First, any minimal set $S \in \mathcal{L}_{\text{basis}}$ satisfies $\sum_{e \in \delta(S)} (1 - x_e) \geq 3$, meaning that at least $\frac{3}{2}$ tokens are initially assigned to S . So S receives enough tokens to give a half token as excess to its parent. Inductively, we claim that every set gets 1 token, and moreover, has an excess of at least $\frac{1}{2}$ that it can give to its parent. For any non-minimal $S \in \mathcal{L}_{\text{basis}}$, we have three cases depending on the number of disjoint maximal children of S in $\mathcal{L}_{\text{basis}}$. In each case we will consider the graph $G_S = (V_S, E_S)$ resulting from contracting the maximal children of S in $\mathcal{L}_{\text{basis}}$ in the graph $G[S]$. In Cases 2 and 3 we crucially use that $x(E_S)$ is an integer by Lemma 19.16.

- **Case 1:** S has at least three maximal children in $\mathcal{L}_{\text{basis}}$.

Then inductively, each of these children has an excess of at least $\frac{1}{2}$. This gives us at least $\frac{3}{2}$ tokens for S , as desired.

- **Case 2:** S has exactly two maximal children $A, B \in \mathcal{L}_{\text{basis}}$.

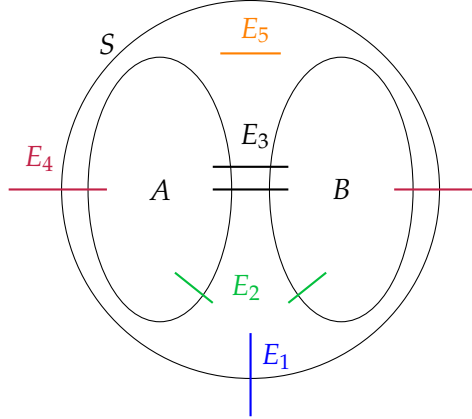


Figure 75: Setting for Case 2. Note some edge sets may be empty.

Inductively, each child has an excess of at least $\frac{1}{2}$, giving us at least one token. Thus we need to collect at least $\frac{1}{2}$ additional tokens.

Consider the edge sets as defined in Fig. 75. In particular,

$$\begin{aligned} E_1 &= \delta(S) \setminus (\delta(A) \cup \delta(B)) \\ E_2 &= (\delta(A) \triangle \delta(B)) \setminus \delta(S) \\ E_3 &= \delta(A) \cap \delta(B) \\ E_4 &= (\delta(A) \cup \delta(B)) \cap \delta(S) \\ E_5 &= E_S \setminus (E_2 \cup E_3) \end{aligned}$$

First we observe that $E_1 \cup E_2 \cup E_5$ is nonempty. For suppose not; then, with χ denoting the incidence vector of a set, we can write

$$\chi(\delta(S)) + 2\chi(E(S)) = \chi(\delta(A)) + \chi(\delta(B)) + 2\chi(E(A)) + 2\chi(E(B)).$$

However, by the maximality of our choice of $\mathcal{C}_{\text{basis}}$, $E(A)$, $E(B)$ and $E(S)$ are all in the span of $\mathcal{C}_{\text{basis}}$, whereas $\delta(S)$, $\delta(A)$ and $\delta(B)$ are all in $\mathcal{L}_{\text{basis}}$. Thus we have a linear dependence among the constraints defined by $\mathcal{C}_{\text{basis}}$ and $\mathcal{L}_{\text{basis}}$, a contradiction.

So $x(E_1) + x(E_2) + x(E_5) > 0$. Therefore, we get

$$\frac{|E_1| + |E_2| - x(E_1) - x(E_2)}{2} + |E_5| - x(E_5) = z - \left(\frac{x(E_1) + x(E_2)}{2} + x(E_5) \right) > 0$$

fractional tokens for some $z \in \mathbb{Z}_{\geq 0}$. We will prove that $\frac{x(E_1) + x(E_2)}{2} + x(E_5)$ is half integral, from which the claim follows. By the integrality of $x(E_S)$ (using \mathcal{L} -alignment) and the

tightness of the constraints on A, B and S , we have that

$$a := x(E_2) + x(E_3) + x(E_5), \quad b := x(E_2) + 2x(E_3) + x(E_4) \quad \text{and} \quad c := x(E_1) + x(E_4)$$

are all integers. Since $a - b/2 + c/2 = \frac{x(E_1) + x(E_2)}{2} + x(E_5)$, the claim follows.

- **Case 3:** S has precisely one maximal child S' in $\mathcal{L}_{\text{basis}}$.

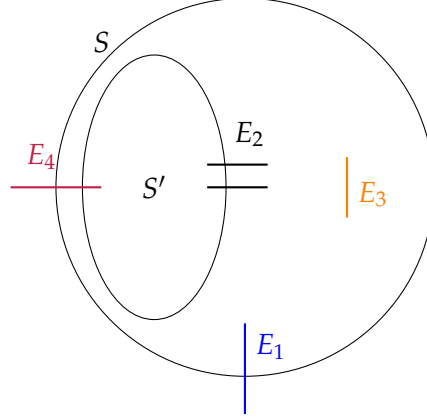


Figure 76: Setting for Case 3.

We need to find 1 token that has been given by edges directly to S , so that the $\frac{1}{2}$ excess token from S' can be carried over as the excess of S .

If $\delta(S') \subseteq \delta(S)$, then because no relaxation step was possible in line 6, $\sum_{e \in \delta(S) - \delta(S')} (1 - x_e) \geq 2$. Since each edge in $\delta(S) - \delta(S')$ contributes $(1 - x_e)/2$ tokens, this gives us our token as necessary. Similarly, if $\delta(S') \supseteq \delta(S)$ we get the desired one token.

So assume that $\delta(S) - \delta(S')$ and $\delta(S') - \delta(S)$ are both nonempty. Let

$$\begin{aligned} E_1 &:= \delta(S) - \delta(S'), \\ E_2 &:= \delta(S') - \delta(S), \\ E_3 &:= E_S - \delta(S'), \quad \text{and} \\ E_4 &:= \delta(S) \cap \delta(S'). \end{aligned}$$

(See [Figure 76](#).)

Let $\delta \in [0, 1)$ be the fractional part of $x(E_4)$. Note that the number of tokens assigned to S is

$$|E_3| - x(E_3) + \frac{1}{2}(|E_1| + |E_2| - x(E_1) - x(E_2)). \quad (87)$$

Also observe that

$$x(E_2) + x(E_3), \quad x(E_1) + x(E_4), \quad \text{and} \quad x(E_2) + x(E_4) \quad (88)$$

are all integer-valued, by tightness of the cut constraints and \mathcal{L} -alignment. We distinguish two subcases.

- $\delta = 0$. Then $x(E_1)$ and $x(E_2)$ are both integers, and moreover since E_1 and E_2 are nonempty, $|E_1| - x(E_1)$ and $|E_2| - x(E_2)$ are both positive integers. This already gives us the desired 1 token by (87).
- $\delta > 0$. Then by (88) the fractional parts of $x(E_1)$ and $x(E_2)$ are both $1 - \delta$, and the fractional part of $x(E_3)$ is then δ . Thus $|E_1| - x(E_1) \geq \delta$ (being positive, with fractional part δ); similarly, $|E_2| - x(E_2) \geq \delta$ and $|E_3| - x(E_3) \geq 1 - \delta$. Substituting into (87), we have at least $1 - \delta + (2\delta)/2 = 1$ tokens assigned to S , as required.

We have demonstrated that all sets in $\mathcal{L}_{\text{basis}}$ receive a full tokens; moreover, any maximal set in $\mathcal{L}_{\text{basis}}$ will have an extra token that is not needed, since it has no parent to give it to. So we have our desired contradiction: $|E| > |\mathcal{C}_{\text{basis}}| + |\mathcal{L}_{\text{basis}}|$. \square

19.9 Conclusion

Besides the (strong) thin tree conjecture, our work leaves open several directions. One fascinating question is whether it is possible to leverage or strengthen our results to give a novel constant factor approximation algorithm for ATSP. While an algorithmic version of the strong thin tree conjecture is sufficient to give a constant factor approximation algorithm for ATSP, it is unclear if it is necessary: indeed, the current constant factor approximation algorithm for ATSP is not known to imply anything about thin trees. We ask if perhaps it is sufficient to focus on thinness for a laminar (or near laminar) family of cuts.

A second open question is whether it is possible to achieve a *minimum cost* tree which violates the degree bounds in a laminar family by any constant factor. One would need to avoid the scaling currently present in our reduction. A natural relaxation of this question is to ask for a $1 + \epsilon$ approximation for arbitrarily small ϵ as has been done for the chain case [LS18].

Finally, we note that our results immediately give a thin tree with respect to the set of minimum cuts of any graph, and we believe it may be possible to extend it to the set of all $(1 + \epsilon)$ near minimum cuts for some small $\epsilon > 0$ using results from [KKO22]. We ask whether it is possible to extend our result to more general families of cuts such as the union of a constant number of laminar families or the set of cuts with at most αk edges in the graph for some constant α significantly larger than 1.

References

- [AKS21] Hyung-Chan An, Robert Kleinberg, and David B. Shmoys. “Approximation Algorithms for the Bottleneck Asymmetric Traveling Salesman Problem”. In: *ACM Transactions on Algorithms* 17.4 (2021). DOI: [10.1145/3478537](https://doi.org/10.1145/3478537) (cit. on p. 232).
- [AO15] Nima Anari and Shayan Oveis Gharan. “Effective-Resistance-Reducing Flows, Spectrally Thin Trees, and Asymmetric TSP”. In: *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*. 2015, pp. 20–39. DOI: [10.1109/FOCS.2015.11](https://doi.org/10.1109/FOCS.2015.11) (cit. on pp. 232, 235).
- [App+07] David L. Applegate, Robert E. Bixby, Vasek Chvatal, and William J. Cook. *The Traveling Salesman Problem: A Computational Study (Princeton Series in Applied Mathematics)*. Princeton, NJ, USA: Princeton University Press, 2007 (cit. on p. 8).

- [Arc+11] Aaron Archer, MohammadHossein Bateni, MohammadTaghi Hajiaghayi, and Howard Karloff. “Improved Approximation Algorithms for Prize-Collecting Steiner Tree and TSP”. In: *SIAM Journal on Computing* 40.2 (2011), pp. 309–332. doi: [10.1137/090771429](https://doi.org/10.1137/090771429) (cit. on p. 10).
- [Aro96] Sanjeev Arora. “Polynomial Time Approximation Schemes for Euclidean TSP and Other Geometric Problems”. In: *FOCS*. 1996, pp. 2–11 (cit. on p. 20).
- [Aro+98] Sanjeev Arora, Michelangelo Grigni, David Karger, Philip Klein, and Andrzej Woloszyn. “A polynomial-time approximation scheme for weighted planar graph TSP”. In: *SODA*. 1998, pp. 33–41 (cit. on p. 20).
- [Asa+10] Arash Asadpour, Michel X. Goemans, Aleksander Madry, Shayan Oveis Gharan, and Amin Saberi. “An $O(\log n / \log \log n)$ -approximation Algorithm for the Asymmetric Traveling Salesman Problem”. In: *SODA*. 2010, pp. 379–389 (cit. on pp. 10, 22, 23).
- [Asa+17] Arash Asadpour, Michel X. Goemans, Aleksander Madry, Shayan Oveis Gharan, and Amin Saberi. “An $O(\log n / \log \log n)$ -Approximation Algorithm for the Asymmetric Traveling Salesman Problem”. In: *Operations Research* 65.4 (2017), pp. 1043–1061 (cit. on pp. 18, 232).
- [Ban+13] Nikhil Bansal, Rohit Khandekar, Jochen Könemann, Viswanath Nagarajan, and Britta Peis. “On generalizations of network design problems with degree bounds”. In: *Mathematical Programming* 141.1 (2013), pp. 479–506. doi: [10.1007/s10107-012-0537-8](https://doi.org/10.1007/s10107-012-0537-8) (cit. on p. 234).
- [Bar92] Francisco Barahona. “Separating from the dominant of the spanning tree polytope”. In: *Operations Research Letters* 12.4 (1992), pp. 201–203. doi: [10.1016/0167-6377\(92\)90045-5](https://doi.org/10.1016/0167-6377(92)90045-5) (cit. on p. 236).
- [BB08] Geneviève Benoit and Sylvia Boyd. “Finding the exact integrality gap for small traveling salesman problems”. In: 33 (2008), pp. 921–931 (cit. on p. 196).
- [BBL09] Julius Borcea, Petter Branden, and Thomas M. Liggett. “Negative dependence and the geometry of polynomials.” In: *Journal of American Mathematical Society* 22 (2009), pp. 521–567 (cit. on pp. 13, 28, 32, 33).
- [BC11a] Sylvia Boyd and Robert Carr. “Finding low cost TSP and 2-matching solutions using certain half-integer subtour vertices”. In: *Discrete Optimization* 8.4 (2011), pp. 525–539 (cit. on p. 9).
- [BC11b] Sylvia Boyd and Robert Carr. “Finding Low Cost TSP and 2-matching Solutions Using Certain Half-Integer Subtour Vertices”. In: *Discrete Optimization* 8 (2011), pp. 525–539 (cit. on p. 196).
- [BEM10] S. Boyd and P. Elliott-Magwood. “Structure of the extreme points of the subtour elimination polytope of the STSP”. In: *Combinatorial Optimization and Discrete Algorithms*. Vol. B23. 2010, pp. 33–47 (cit. on p. 9).
- [Ben95] András A. Benczúr. “A Representation of Cuts within $6/5$ Times the Edge Connectivity with Applications”. In: *FOCS*. 1995, pp. 92–102 (cit. on pp. 13, 61, 134, 136, 139, 151, 174, 175).
- [Ben97] Andras A. Benczúr. “Cut structures and randomized algorithms in edge-connectivity problems”. PhD thesis. MIT, 1997 (cit. on pp. 39, 40, 148).

- [BFS16] Sylvia Boyd, Yao Fu, and Yu Sun. “A $5/4$ -approximation for subcubic 2EC using circulations and obliged edges”. In: *Discrete Applied Mathematics* 209 (2016), pp. 48–58 (cit. on pp. 17, 218).
- [BG08] András A. Benczúr and Michel X. Goemans. “Deformable Polygon Representation and Near-Mincuts”. In: *Building Bridges: Between Mathematics and Computer Science*, M. Groetschel and G.O.H. Katona, Eds., *Bolyai Society Mathematical Studies* 19 (2008), pp. 103–135 (cit. on pp. 13, 40, 61, 134, 136, 138, 139, 148, 151, 174).
- [BGA20a] Jarosław Byrka, Fabrizio Grandoni, and Afrouz Jabal Ameli. “Breaching the 2-Approximation Barrier for Connectivity Augmentation: A Reduction to Steiner Tree”. In: *STOC*. Association for Computing Machinery, 2020, 815–825. ISBN: 9781450369794 (cit. on p. 134).
- [BGA20b] Jaroslaw Byrka, Fabrizio Grandoni, and Afrouz Jabal Ameli. “Breaching the 2 approximation barrier for connectivity augmentation: a reduction to Steiner tree”. In: *STOC*. Ed. by Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy. ACM, 2020, pp. 815–825 (cit. on p. 218).
- [BHH59] Jillian E. Beardwood, John H. Halton, and J. M. Hammersley. “The shortest path through many points”. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 55 (1959), pp. 299–327 (cit. on p. 8).
- [BKN23] Jannis Blauth, Nathan Klein, and Martin Nägele. *A Better-Than-1.6-Approximation for Prize-Collecting TSP*. 2023. arXiv: 2308.06254 [cs.DS] (cit. on p. 10).
- [BN23] Jannis Blauth and Martin Nägele. “An Improved Approximation Guarantee for Prize-Collecting TSP”. In: *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*. STOC 2023. Orlando, FL, USA: Association for Computing Machinery, 2023, 1848–1861. ISBN: 9781450399135. DOI: 10.1145/3564246.3585159 (cit. on p. 10).
- [Boy+20] Sylvia Boyd, Joseph Cheriyan, Robert Cummings, Logan Grout, Sharat Ibrahimpur, Zoltán Szigeti, and Lu Wang. “A $4/3$ -Approximation Algorithm for the Minimum 2-Edge Connected Multisubgraph Problem in the Half-Integral Case”. In: *APPROX/RANDOM*. Ed. by Jarosław Byrka and Raghu Meka. Vol. 176. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2020, 61:1–61:12 (cit. on pp. 17, 218).
- [BP91] S. C. Boyd and William R. Pulleyblank. “Optimizing over the subtour polytope of the travelling salesman problem”. In: *Math. Program.* 49 (1991), pp. 163–187 (cit. on p. 9).
- [BS21] Sylvia Boyd and András Sebő. “The Salesman’s Improved Tours for Fundamental Classes”. In: *Math. Program.* 186.1–2 (2021), 289–307. ISSN: 0025-5610. DOI: 10.1007/s10107-019-01455-3 (cit. on pp. 14, 37, 191, 196, 197).
- [BV22] Traub-V. Blauth J. and J. Vygen. “Improving the approximation ratio for capacitated vehicle routing”. In: *Mathematical Programming* (2022) (cit. on p. 10).
- [Chr76] Nicos Christofides. *Worst Case Analysis of a New Heuristic for the Traveling Salesman Problem*. Report 388. Pittsburgh, PA: Graduate School of Industrial Administration, Carnegie-Mellon University, 1976 (cit. on pp. 8, 9).
- [Coo12] William J. Cook. *In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation*. Princeton, NJ: Princeton University Press, 2012. ISBN: 978-0691152707 (cit. on p. 9).

- [CR98] Robert Carr and R. Ravi. “A New Bound for the 2-Edge Connected Subgraph Problem”. In: *IPCO*. 1998, pp. 112–125 (cit. on pp. 17, 218).
- [CT00] J. Cheriyan and R. Thurimella. “Approximating minimum- size k -connected spanning subgraphs via matching”. In: *SIAM J Comput* 30 (2000), pp. 528–560 (cit. on pp. 17, 218).
- [CTZ21] Federica Cecchetto, Vera Traub, and Rico Zenklusen. “Bridging the Gap between Tree and Connectivity Augmentation: Unified and Stronger Approaches”. In: *Proceedings of the 53rd Annual ACM SIGACT Symposium on Theory of Computing*. New York, NY, USA: Association for Computing Machinery, 2021, 370–383. ISBN: 9781450380539 (cit. on pp. 134, 218).
- [CV00] Robert D. Carr and Santosh Vempala. “Towards a $4/3$ approximation for the asymmetric traveling salesman problem”. In: *SODA*. 2000, pp. 116–125 (cit. on p. 9).
- [CVZ10] Chandra Chekuri, Jan Vondrák, and Rico Zenklusen. “Dependent Randomized Rounding via Exchange Properties of Combinatorial Structures”. In: *FOCS*. 2010, pp. 575–584 (cit. on pp. 220, 234).
- [Dar64] J. N. Darroch. “On the distribution of the number of successes in independent trials”. In: *Ann. Math. Stat.* 36 (1964), pp. 1317–1321 (cit. on p. 32).
- [DFJ54] G Dantzig, R Fulkerson, and S Johnson. “Solution of a large-scale traveling-salesman problem”. In: *Operations Research* 2 (1954), pp. 393–410 (cit. on pp. 8, 10, 19).
- [DHM07] Erik D. Demaine, MohammadTaghi Hajiaghayi, and Bojan Mohar. “Approximation algorithms via contraction decomposition”. In: *SODA*. 2007, pp. 278–287 (cit. on p. 20).
- [DKL76] E.A. Dinits, A.V. Karzanov, and M.V. Lomonosov. “On the structure of a family of minimal weighted cuts in graphs”. In: *Studies in Discrete Mathematics (in Russian)*, ed. A.A. Fridman, 290–306, Nauka (Moskva) (1976) (cit. on pp. 38, 61, 134–136, 139).
- [DT09] Vladimir Deineko and Alexander Tiskin. “Min-weight double-tree shortcutting for Metric TSP: Bounding the approximation ratio”. In: *Electronic Notes in Discrete Mathematics* 32 (2009). DIMAP Workshop on Algorithmic Graph Theory, pp. 19–26. ISSN: 1571-0653. DOI: <https://doi.org/10.1016/j.endm.2009.02.004> (cit. on pp. 259, 260).
- [Edm70] Jack Edmonds. “Submodular functions, matroids and certain polyhedra”. In: *Combinatorial Structures and Their Applications*. New York, NY, USA: Gordon and Breach, 1970, pp. 69–87 (cit. on pp. 20, 21, 235).
- [Edr53] Albert Edrei. “Proof of a Conjecture of Schoenberg on the Generating Function of a Totally Positive Sequence”. In: *Canadian Journal of Mathematics* 5 (1953), 86–94. DOI: [10.4153/CJM-1953-010-3](https://doi.org/10.4153/CJM-1953-010-3) (cit. on p. 27).
- [EJ73] Jack Edmonds and Ellis L. Johnson. “Matching, Euler tours and the Chinese postman”. In: *Mathematical Programming* 5.1 (1973), pp. 88–124 (cit. on p. 21).
- [FF09] Tamás Fleiner and András Frank. *A quick proof for the cactus representation of mincuts*. Tech. rep. QP-2009-03. www.cs.elte.hu/egres. Egerváry Research Group, Budapest, 2009 (cit. on pp. 38, 199).

- [Fio+18] Samuel Fiorini, Martin Gross, Jochen Könnemann, and Laura Sanità. “Approximating Weighted Tree Augmentation via Chvátal-Gomory Cuts”. In: *SODA*. 2018, pp. 817–831 (cit. on p. 218).
- [FJ81] G. N. Fredrickson and Joseph F. JáJá. “Approximation Algorithms for Several Graph Augmentation Problems”. In: *SIAM J. Comput.* 10.2 (1981), pp. 270–283 (cit. on pp. 17, 218).
- [FJ82] G. N. Fredrickson and Joseph F. JáJá. “On the relationship between the biconnectivity augmentation and traveling salesman problem”. In: *Theoretical Computer Science* 19 (1982), pp. 189–201 (cit. on pp. 17, 218).
- [Fle99] Lisa Fleischer. “Building Chain and Cactus Representations of All Minimum Cuts from Hao-Orlin in the Same Asymptotic Run Time”. In: 33 (1999), pp. 51–72 (cit. on p. 214).
- [FM92] Tomás Feder and Milena Mihail. “Balanced matroids”. In: *Proceedings of the twenty-fourth annual ACM symposium on Theory of Computing*. Victoria, British Columbia, Canada: ACM, 1992, pp. 26–38 (cit. on p. 32).
- [FR92] Martin Fürer and Balaji Raghavachari. “Approximating the Minimum Degree Spanning Tree to Within One from the Optimal Degree”. In: *Proceedings of the Third Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 1992, pp. 317–324. DOI: [10.1145/139404.139469](https://doi.org/10.1145/139404.139469) (cit. on p. 233).
- [Fri+22] Zachary Friggstad, Ramin Mousavi, Mirmahdi Rahgoshay, and Mohammad R. Salavatipour. “Improved Approximations for Capacitated Vehicle Routing with Unsplittable Client Demands”. In: *Integer Programming and Combinatorial Optimization: 23rd International Conference, IPCO 2022, Eindhoven, The Netherlands, June 27–29, 2022, Proceedings*. Eindhoven, The Netherlands: Springer-Verlag, 2022, 251–261. ISBN: 978-3-031-06900-0. DOI: [10.1007/978-3-031-06901-7_19](https://doi.org/10.1007/978-3-031-06901-7_19) (cit. on p. 10).
- [Gab05] H. Gabow. “An improved analysis for approximating the smallest k -edge connected spanning subgraph of a multi-graph”. In: *SIAM J Disc Math* 19 (2005), pp. 1–18 (cit. on pp. 17, 218).
- [Gab+09] Harold N. Gabow, Michel X. Goemans, Éva Tardos, and David P. Williamson. “Approximating the smallest k -edge connected spanning subgraph by LP-rounding”. In: *Networks* 53.4 (2009), pp. 345–357 (cit. on pp. 17, 18, 218).
- [GB93a] Michel Goemans and Dimitris Bertsimas. “Survivable Network, Linear Programming Relaxations and the Parsimonious Property”. In: *Math Program* 60 (June 1993). DOI: [10.1007/BF01580607](https://doi.org/10.1007/BF01580607) (cit. on p. 17).
- [GB93b] Michel X. Goemans and Dimitris Bertsimas. “Survivable networks, linear programming relaxations and the parsimonious property”. In: *Math. Program.* 60 (1993), pp. 145–166 (cit. on pp. 10, 219).
- [GG08] H. Gabow and S. Gallagher. “Iterated rounding algorithms for the smallest k -edge connected spanning subgraph”. In: *SIAM J. Comput.* 41 (2008), pp. 61–103 (cit. on pp. 17, 218).

- [GKP95] M. Grigni, E. Koutsoupias, and C. Papadimitriou. “An approximation scheme for planar graph TSP”. In: *FOCS '95: Proceedings of the 36th Annual Symposium on Foundations of Computer Science*. Washington, DC, USA: IEEE Computer Society, 1995, p. 640. ISBN: 0-8186-7183-1 (cit. on p. 20).
- [GL21] Leonid Gurvits and Jonathan Leake. “Capacity lower bounds via productization”. In: *STOC*. 2021, pp. 847–858 (cit. on p. 62).
- [GLS05] David Gamarnik, Moshe Lewenstein, and Maxim Sviridenko. “An improved upper bound for the TSP in cubic 3-edge-connected graphs”. In: *Oper. Res. Lett.* 33.5 (Sept. 2005), pp. 467–474 (cit. on p. 9).
- [God04] Luis A. Goddyn. “Some Open Problems I Like”. available at <http://www.math.sfu.ca/goddyn/Problems/problems.html>. 2004 (cit. on pp. 18, 232).
- [Goe06] Michel X. Goemans. “Minimum Bounded Degree Spanning Trees”. In: *FOCS*. 2006, pp. 273–282 (cit. on pp. 18, 233).
- [Goe09] Michel X. Goemans. *Combining Approximation Algorithms for the Prize-Collecting TSP*. 2009. arXiv: 0910.0553 [cs.DS] (cit. on p. 10).
- [Goe12] Michel Goemans. *Thin spanning trees*. Lecture at “Graph Theory @ Georgia Tech: Conference Honoring the 50th Birthday of Robin Thomas”. 2012 (cit. on p. 235).
- [Goe95a] Michel X. Goemans. “Worst-Case Comparison of Valid Inequalities for the TSP”. In: *MATH. PROG* 69 (1995), pp. 335–349 (cit. on p. 9).
- [Goe95b] Michel X. Goemans. “Worst-case comparison of valid inequalities for the TSP”. In: 69 (1995), pp. 335–349 (cit. on p. 196).
- [Gup+22] Anupam Gupta, Euiwoong Lee, Jason Li, Marcin Mucha, Heather Newman, and Sherry Sarkar. “Matroid-Based TSP Rounding for Half-Integral Solutions”. In: *Integer Programming and Combinatorial Optimization*. Ed. by Karen Aardal and Laura Sanità. Lecture Notes in Computer Science 13265. See also <https://arxiv.org/pdf/2111.09290.pdf>. 2022, pp. 305–318 (cit. on pp. 42, 196, 197).
- [Gur06] Leonid Gurvits. “Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications”. In: *STOC*. Ed. by Jon M. Kleinberg. ACM, 2006, pp. 417–426 (cit. on p. 62).
- [Gur08] Leonid Gurvits. “Van der Waerden/Schrijver-Valiant like Conjectures and Stable (aka Hyperbolic) Homogeneous Polynomials: One Theorem for all”. In: *Electr. J. Comb.* 15.1 (2008) (cit. on p. 62).
- [GW17] Kyle Genova and David P. Williamson. “An Experimental Evaluation of the Best-of-Many Christofides’ Algorithm for the Traveling Salesman Problem”. In: *Algorithmica* 78.4 (2017), pp. 1109–1130 (cit. on p. 15).
- [HK70] M. Held and R.M. Karp. “The traveling salesman problem and minimum spanning trees”. In: *Operations Research* 18 (1970), pp. 1138–1162 (cit. on pp. 10, 20).
- [HLP52] G. H. Hardy, J. E. Littlewood, and G. Polya. *Inequalities*. Cambridge Univ. Press, 1952 (cit. on p. 32).

- [HN19] Arash Haddadan and Alantha Newman. “Towards Improving Christofides Algorithm for Half-Integer TSP”. In: *ESA*. Ed. by Michael A. Bender, Ola Svensson, and Grzegorz Herman. Vol. 144. LIPIcs. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019, 56:1–56:12 (cit. on pp. 9, 42, 196).
- [HNR17] Arash Haddadan, Alantha Newman, and R. Ravi. “Cover and Conquer: Augmenting Decompositions for Connectivity Problems”. abs/1707.05387. 2017 (cit. on p. 9).
- [HO14] Nicholas J. A. Harvey and Neil Olver. “Pipage Rounding, Pessimistic Estimators and Matrix Concentration”. In: *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2014, pp. 926–945 (cit. on p. 235).
- [Hoe56] W. Hoeffding. “On the distribution of the number of successes in independent trials”. In: *Ann. Math. Statist.* 27 (1956), pp. 713–721 (cit. on pp. 29, 96).
- [Jae84] F. Jaeger. “On circular flows in graphs, Finite and Infinite Sets”. In: *Colloquia Mathematica Societatis Janos Bolyai* 37 (1984), pp. 391–402 (cit. on pp. 18, 232).
- [Jai01] Kamal Jain. “A Factor 2 Approximation Algorithm for the Generalized Steiner Network Problem”. In: *Combinatorica* 21 (2001), pp. 39–60 (cit. on p. 134).
- [Kar00] David R. Karger. “Minimum cuts in near-linear time”. In: *J. ACM* 47.1 (Jan. 2000), pp. 46–76 (cit. on p. 134).
- [Kar+21] Anna R. Karlin, Nathan Klein, Shayan Oveis Gharan, and Xinzhi Zhang. *An Improved Approximation Algorithm for the Minimum k -Edge Connected Multi-Subgraph Problem*. 2021. DOI: [10.48550/ARXIV.2101.05921](https://arxiv.org/abs/10.48550/ARXIV.2101.05921) (cit. on p. 134).
- [Kar72] Richard M. Karp. “Reducibility Among Combinatorial Problems”. In: *Complexity of Computer Computations*. 1972, pp. 85–103 (cit. on pp. 8, 20, 257).
- [Kar99] D. Karger. “Random sampling in cut, flow, and network design problems”. In: *Math OR* 24 (1999), pp. 383–413 (cit. on pp. 17, 218, 219).
- [KGV83] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. “Optimization by Simulated Annealing”. In: *Science* 220.4598 (1983), pp. 671–680. DOI: [10.1126/science.220.4598.671](https://doi.org/10.1126/science.220.4598.671). eprint: <https://www.science.org/doi/pdf/10.1126/science.220.4598.671> (cit. on p. 8).
- [KKO20] Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. “An improved approximation algorithm for TSP in the half integral case”. In: *STOC*. Ed. by Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy. ACM, 2020, pp. 28–39 (cit. on pp. 9, 42, 134, 135, 191, 192, 196).
- [KKO21] Anna R. Karlin, Nathan Klein, and Shayan Oveis Gharan. “A (Slightly) Improved Approximation Algorithm for Metric TSP”. In: *STOC*. ACM, 2021 (cit. on pp. 61, 137, 144, 162–164, 167, 169, 172, 173, 183, 185, 187, 192, 219, 277).
- [KKO22] Anna Karlin, Nathan Klein, and Shayan Oveis Gharan. “A (Slightly) Improved Bound on the Integrality Gap of the Subtour LP for TSP”. In: *FOCS*. IEEE Computer Society, 2022, pp. 844–855 (cit. on pp. 61, 137, 167, 169, 172, 173, 176, 177, 179, 189, 218, 247).
- [Kle05] Philip N. Klein. “A linear-time approximation scheme for planar weighted TSP”. In: *FOCS*. 2005, pp. 647–657 (cit. on p. 20).

- [KLS12] Tamás Király, Lap Chi Lau, and Mohit Singh. “Degree Bounded Matroids and Submodular Flows”. In: *Combinatorica* 32 (2012), pp. 703–720. DOI: [10.1007/s00493-012-2760-6](https://doi.org/10.1007/s00493-012-2760-6) (cit. on pp. 234, 244).
- [KLS15] Marek Karpinski, Michael Lampis, and Richard Schmieid. “New inapproximability bounds for TSP”. In: *Journal of Computer and System Sciences* 81.8 (2015), pp. 1665–1677. ISSN: 0022-0000 (cit. on p. 9).
- [KP09] *A Near-Linear Time Algorithm for Constructing a Cactus Representation of Minimum Cuts*. 2009, pp. 246–255 (cit. on p. 134).
- [KR96] S. Khuller and B. Raghavachari. “Improved approximation algorithms for uniform connectivity problems”. In: *J Algorithms* 21 (1996), pp. 434–450 (cit. on pp. 17, 218).
- [Lin+20] André Linhares, Neil Olver, Chaitanya Swamy, and Rico Zenklusen. “Approximate multi-matroid intersection via iterative refinement”. In: *Mathematical Programming* 183.1 (2020), pp. 397–418. DOI: [10.1007/s10107-020-01524-y](https://doi.org/10.1007/s10107-020-01524-y) (cit. on p. 242).
- [LOS12] Bundit Laekhanukit, Shayan Oveis Gharan, and Mohit Singh. “A Rounding by Sampling Approach to the Minimum Size k -Arc Connected Subgraph Problem”. In: *ICALP (1)*. 2012, pp. 606–616 (cit. on pp. 17, 218).
- [LRS11] Lap-Chi Lau, R. Ravi, and Mohit Singh. *Iterative Methods in Combinatorial Optimization*. 1st. New York, NY, USA: Cambridge University Press, 2011 (cit. on p. 134).
- [LS18] André Linhares and Chaitanya Swamy. “Approximating Min-Cost Chain-Constrained Spanning Trees: A Reduction from Weighted to Unweighted Problems”. In: *Mathematical Programming* 172 (2018), pp. 17–34. DOI: [10.1007/s10107-017-1150-7](https://doi.org/10.1007/s10107-017-1150-7) (cit. on pp. 233, 234, 247).
- [Mit99] Joseph SB Mitchell. “Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric TSP, k -MST, and related problems”. In: *SIAM Journal on Computing* 28.4 (1999), pp. 1298–1309 (cit. on p. 20).
- [MS16] Tobias Mömke and Ola Svensson. “Removing and Adding Edges for the Traveling Salesman Problem”. In: 63 (2016). Article 2 (cit. on pp. 20, 196).
- [MSS13] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. “Interlacing Families II: Mixed Characteristic Polynomials and the Kadison-Singer Problem”. In: *FOCS*. 2013 (cit. on p. 235).
- [NNI94] Hiroshi Nagamochi, Kazuhiro Nishimura, and Toshihide Ibaraki. “Computing All Small Cuts in an Undirected Network”. In: *SIAM Journal on Discrete Mathematics* 10 (1994), pp. 469–481 (cit. on p. 175).
- [NW61] C. St. J. A. Nash-Williams. “Edge disjoint spanning trees of finite graphs”. In: *J. London Math. Soc.* 36 (1961), pp. 445–45 (cit. on p. 235).
- [NZ19] Martin Nägele and Rico Zenklusen. “A New Dynamic Programming Approach for Spanning Trees with Chain Constraints and Beyond”. In: *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2019, pp. 1550–1569 (cit. on p. 234).
- [OS11] Shayan Oveis Gharan and Amin Saberi. “The Asymmetric Traveling Salesman Problem on Graphs with Bounded Genus”. In: *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*. 2011, pp. 967–975 (cit. on pp. 18, 232).

- [OSS11] Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. “A Randomized Rounding Approach to the Traveling Salesman Problem”. In: *FOCS*. IEEE Computer Society, 2011, pp. 550–559. ISBN: 978-0-7695-4571-4 (cit. on pp. 19, 20, 24, 33, 38, 53, 61, 63, 144, 167, 219).
- [OZ18] Neil Olver and Rico Zenklusen. “Chain-Constrained Spanning Trees”. In: *Mathematical Programming* 167 (2018), pp. 293–314. DOI: [10.1007/s10107-017-1126-7](https://doi.org/10.1007/s10107-017-1126-7) (cit. on pp. 18, 233, 234, 238).
- [Pri11] David Pritchard. “k-Edge-Connectivity: Approximation and LP Relaxation”. In: *Proceedings of the Ninth International Workshop on Approximation and Online Algorithms (WAOA)*. Ed. by Klaus Jansen and Roberto Solis-Oba. Springer Berlin Heidelberg, 2011, pp. 225–236. DOI: [10.1007/978-3-642-18318-8_20](https://doi.org/10.1007/978-3-642-18318-8_20) (cit. on pp. 17, 18, 218, 234).
- [PS23] Aditya Pillai and Mohit Singh. *Linear Programming based Reductions for Multiple Visit TSP and Vehicle Routing Problems*. 2023. arXiv: [2308.11742](https://arxiv.org/abs/2308.11742) [cs.DS] (cit. on p. 10).
- [PV84] Christos H Papadimitriou and Umesh V Vazirani. “On two geometric problems related to the travelling salesman problem”. In: *Journal of Algorithms* 5.2 (1984), pp. 231–246. ISSN: 0196-6774. DOI: [https://doi.org/10.1016/0196-6774\(84\)90029-4](https://doi.org/10.1016/0196-6774(84)90029-4) (cit. on p. 258).
- [RS05] Qazi Ibadur Rahman and Gerhard Schmeisser. *Analytic Theory of Polynomials*. Oxford: Oxford University Press, 2005. ISBN: 9780198534938 0198534930 (cit. on p. 28).
- [Sch03] Alexander Schrijver. *Combinatorial Optimization*. English. Vol. 2. Algorithms and Combinatorics. Springer, 2003. ISBN: 978-3-540-44389-6 (cit. on pp. 11, 20, 244).
- [Ser78] A. I. Serdyukov. “O nekotorykh ekstremal’nykh obkhodakh v grafakh”. In: *Upravlyayemye sistemy* 17 (1978), pp. 76–79 (cit. on p. 9).
- [SL15] Mohit Singh and Lap Chi Lau. “Approximating minimum bounded degree spanning trees to within one of optimal”. In: *Journal of the ACM* 62.1 (2015). DOI: [10.1145/2629366](https://doi.org/10.1145/2629366) (cit. on pp. 18, 233, 244).
- [STV20] Ola Svensson, Jakub Tarnawski, and László A. Végh. “A Constant-Factor Approximation Algorithm for the Asymmetric Traveling Salesman Problem”. In: *Journal of the ACM* 67.6 (2020). DOI: [10.1145/3424306](https://doi.org/10.1145/3424306) (cit. on pp. 18, 20, 232).
- [SV12] András Sebö and Jens Vygen. “Shorter Tours by Nicer Ears:” CoRR abs/1201.1870. 2012 (cit. on p. 20).
- [SV14] András Sebö and Jens Vygen. “Shorter tours by nicer ears: 7/5-Approximation for the graph-TSP, 3/2 for the path version, and 4/3 for two-edge-connected subgraphs”. In: *Combinatorica* 34.5 (2014), pp. 597–629 (cit. on pp. 17, 218).
- [SV19] Damian Straszak and Nisheeth K. Vishnoi. “Maximum Entropy Distributions: Bit Complexity and Stability”. In: *COLT*. Ed. by Alina Beygelzimer and Daniel Hsu. Vol. 99. Proceedings of Machine Learning Research. PMLR, 2019, pp. 2861–2891 (cit. on pp. 23, 65, 166).
- [SW90] D. B. Shmoys and D. P. Williamson. “Analyzing the Held-Karp TSP bound: a monotonicity property with application”. In: *Inf. Process. Lett.* 35.6 (Sept. 1990), pp. 281–285 (cit. on pp. 9, 11, 19, 22).

- [SWZ12] Frans Schalekamp, David P. Williamson, and Anke van Zuylen. “A proof of the Boyd-Carr conjecture”. In: *SODA*. 2012, pp. 1477–1486 (cit. on pp. 9, 71).
- [SWZ13] Frans Schalekamp, David P. Williamson, and Anke van Zuylen. “2-Matchings, the Traveling Salesman Problem, and the Subtour LP: A Proof of the Boyd-Carr Conjecture”. In: *Mathematics of Operations Research* 39.2 (2013), pp. 403–417 (cit. on pp. 71, 196).
- [Tho12] Carsten Thomassen. “The weak 3-flow conjecture and the weak circular flow conjecture”. In: *Journal of Combinatorial Theory, Series B* 102.2 (2012), pp. 521–529. doi: [10.1016/j.jctb.2011.09.003](https://doi.org/10.1016/j.jctb.2011.09.003) (cit. on pp. 18, 232).
- [TV22] Vera Traub and Jens Vygen. “An Improved Approximation Algorithm for The Asymmetric Traveling Salesman Problem”. In: *SIAM Journal on Computing* 51.1 (2022), pp. 139–173. doi: [10.1137/20M1339313](https://doi.org/10.1137/20M1339313). eprint: <https://doi.org/10.1137/20M1339313> (cit. on p. 20).
- [TVZ20] Vera Traub, Jens Vygen, and Rico Zenklusen. “Reducing path TSP to TSP”. In: *STOC*. Ed. by Konstantin Makarychev, Yury Makarychev, Madhur Tulsiani, Gautam Kamath, and Julia Chuzhoy. ACM, 2020, pp. 14–27 (cit. on pp. 10, 16, 167).
- [TZ22] Vera Traub and Rico Zenklusen. “A Better-Than-2 Approximation for Weighted Tree Augmentation”. In: *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*. 2022, pp. 1–12. doi: [10.1109/FOCS52979.2021.00010](https://doi.org/10.1109/FOCS52979.2021.00010) (cit. on p. 218).
- [Wol80] Laurence A. Wolsey. “Heuristic analysis, linear programming and branch and bound”. In: *Combinatorial Optimization II*. Vol. 13. Mathematical Programming Studies. Springer Berlin Heidelberg, 1980, pp. 121–134 (cit. on pp. 9, 11, 19, 22, 63, 137).
- [WS11] David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011. doi: [10.1017/CB09780511921735](https://doi.org/10.1017/CB09780511921735) (cit. on p. 9).

A Preliminary Background

A.1 Graphs, Walks, and Circuits

Let $G = (V, E)$ denote a graph with vertex set V and edge set E . We will always use n to denote $|V|$. We will sometimes use $\{u, v\}$ to denote the edge between u and v .

A **walk** in a graph is a sequence of edges that joins a sequence of vertices. In other words, a sequence of edges (e_1, \dots, e_k) is a walk if there exists a sequence of vertices (v_1, \dots, v_{k+1}) (not necessarily distinct) such that $e_i = \{v_i, v_{i+1}\}$. A walk is **closed** if $v_1 = v_{k+1}$.

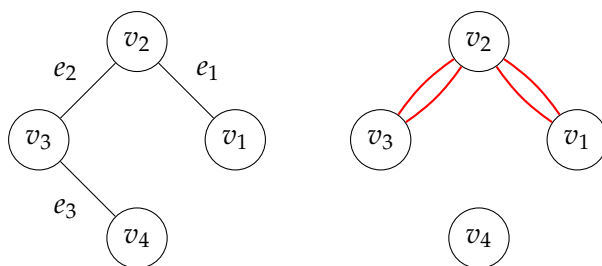


Figure 77: On the right is the closed walk $\{e_1, e_2, e_2, e_1\}$ on the graph to the left. This walk is associated with the vertex sequence $\{v_1, v_2, v_3, v_2, v_1\}$.

A **cycle** is a closed walk in which only the first and last vertices in its vertex sequence appear twice. A **Hamiltonian cycle** is a cycle of length n , i.e. one that visits all vertices. It is NP-Hard to recognize if a graph contains a Hamiltonian cycle [Kar72].

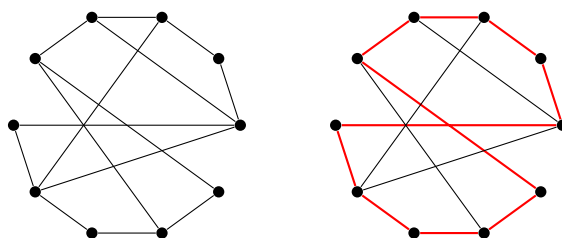


Figure 78: On the left, a Hamiltonian graph. On the right is a Hamiltonian cycle of this graph highlighted in red.

A **circuit** is a closed walk which does not repeat edges. A circuit is **Eulerian** if it uses every edge of the graph. Thus, every Hamiltonian cycle is an Eulerian circuit, however the converse is not true. We say a graph is Eulerian if it has an Eulerian circuit. The following simple fact is due to Euler:

Fact A.1. *A connected multigraph is Eulerian if and only if every vertex has even degree.*

This demonstrates that while it is NP-Hard to determine if a graph has a closed walk that visits every *vertex* exactly once (except the starting vertex), it is trivial to determine if a graph has a walk which visits every *edge* exactly once.

A.2 Metric Completion and Equivalent Formulations of Metric TSP

Let $G = (V, E)$ be a connected graph with weight function $c \in \mathbb{R}^E$ such that c_e denotes the cost of edge e . For a set of edges $F \subseteq E$, we will let $c(F)$ denote $\sum_{e \in F} c_e$.

We define the **metric completion** of G as a complete graph such that $c_{\{u,v\}}$ is equal to the weight of the shortest path between u and v in G . (Note that this may decrease the value of c_e for some edges e , as in the below example.)

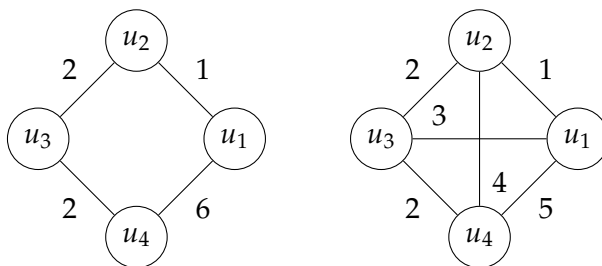


Figure 79: On the right is the metric completion of the graph on the left.

Fact A.2. *The following problems are equivalent:*

1. *Given a complete graph $G = (V, E)$ and a symmetric cost function $c : V \times V \rightarrow \mathbb{R}_{\geq 0}$ with forms a metric, find a minimum cost Hamiltonian cycle. This is classically known as the **metric traveling salesperson problem**, or metric TSP.*
2. *Given a connected graph $G = (V, E)$ with a weight function $c \in \mathbb{R}^E$, find a minimum cost multi-set of edges F such that the multigraph $G' = (V, F)$ is connected and Eulerian, i.e. a minimum cost TSP tour.*
3. *Given a complete graph $G = (V, E)$ and a symmetric cost function $c : V \times V \rightarrow \mathbb{R}_{\geq 0}$ which forms a metric, find a minimum cost multi-set of edges F such that the multigraph $G' = (V, F)$ is connected and Eulerian, i.e. a minimum cost TSP tour.*

The equivalence of these problems follows from the fact that (1) the metric completion of a graph forms a metric, and (2) any optimal TSP tour of a graph G can be shortcut to a Hamiltonian cycle of the same cost in the metric completion of G . In general, any TSP tour can be shortcut to a Hamiltonian cycle of *no greater cost* in the metric completion (by a simple application of the triangle inequality).

Remark A.3. It is important to note that there can be many ways to shortcut a TSP tour to a Hamiltonian cycle, and finding the one which produces the Hamiltonian cycle of minimum cost is NP-Hard, even in the Euclidean case [PV84]. Thus there is a subtlety to be aware of when going from (2) or (3) to (1): if your solution to the former is approximate, in translating to (1) you may sometimes improve your solution. However, determining how much you can improve is NP-Hard. **Thus in this work we ignore all potential improvements gleaned from shortcutting.** We make no claim that this is the correct approach, and indeed learning how to exploit the gain from shortcutting is an interesting open direction. So, we will always work with (3), and since the graph is given by a metric we can assume that the optimal solution is a Hamiltonian cycle. By [Fact A.1](#), our goal is therefore to select the cheapest multi-set of edges which connects the graph

and has even degree at every vertex. In this section we will continue to address shortcutting, but in future sections we will ignore it.

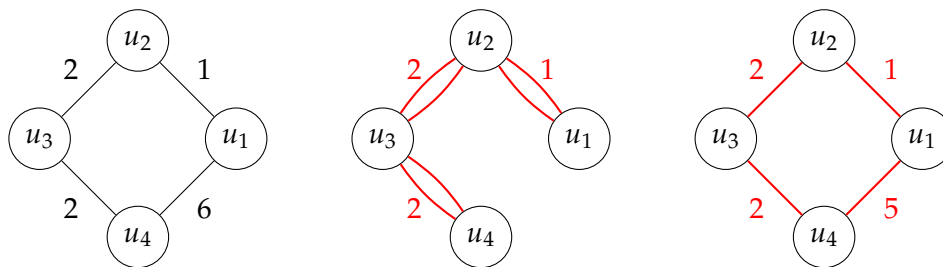


Figure 80: On the left is a metric TSP instance given by (3). In the middle is the optimal solution of weight 10 as in (3). On the right, the optimal solution in the metric completion as in (2).

A.3 The Double Tree Algorithm

The double tree algorithm is a fast and simple 2-approximation for TSP. As in all future discussions in this paper, our input will be symmetric metric $c : V \times V \rightarrow \mathbb{R}_{\geq 0}$ and our goal will be to output a minimum cost Eulerian multi-subgraph.

This algorithm simply chooses a minimum spanning tree T and returns $T \uplus T$, where for two sets A, B we let $A \uplus B$ denote their multi-set union.

Algorithm 11 Double Tree Algorithm for TSP

- 1: Find a minimum cost spanning tree T of G
 - 2: Return $T \uplus T$, i.e. return the Eulerian multi-subgraph containing two copies of each edge in T .
-

Since T is connected, the output is connected. Furthermore, since every edge in T appears twice, every vertex has even degree, so $T \uplus T$ is Eulerian. Therefore this algorithm always produces a valid solution. Finally we notice that $c(T) \leq c(OPT)$, since OPT connects the graph and T is the cheapest set of edges that connects the graph. Thus the cost of the solution returned by the algorithm is $2c(T) \leq 2c(OPT)$, as desired.

Fig. 80 is an example solution produced by the double tree algorithm. In this case it is optimal. However, the analysis is tight, as Fig. 81 illustrates.

The graph in Fig. 81 is Hamiltonian, so there is a solution of cost n , however the double tree algorithm returns a solution of cost $2n - 2$, which tends to a 2-approximation as $n \rightarrow \infty$. (The same is true for any Hamiltonian graphic instance of TSP.) If shortcutting is allowed (see Remark A.3), this instance is not a lower bound example, however a lower bound of 2 is given in [DT09] which works against even the optimal shortcutting.

A.4 The Christofides-Serdyukov Algorithm

For a set of vertices $O \subseteq V$ with even parity, an O -Join $M \subseteq E$ is a set of edges such that every vertex in O is adjacent to an odd number of edges in F and every vertex not in O is adjacent to an even number of edges in F . A minimum cost O -Join can be computed in polynomial time by finding a minimum weight perfect matching on the metric completion of G .

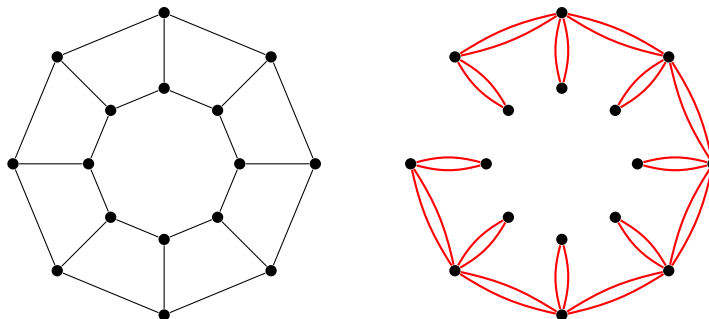


Figure 81: A tight example for the double tree algorithm (without shortcutting). Let $c_e = 1$ for all edges.

Algorithm 12 Christofides-Serdyukov Algorithm

- 1: Find a minimum cost spanning tree T of G
 - 2: Let O be the set of vertices with odd degree in T . Compute the minimum cost O -Join M .
 - 3: Return $T \uplus M$.
-

Every vertex with odd degree in T has odd degree in M , so it has even degree in $T \uplus M$. Similarly, every vertex with even degree in T has even degree in M , so it again has even degree in $T \uplus M$. Therefore $T \uplus M$ is Eulerian and the algorithm is correct.

We now show that this is a $3/2$ approximation. As above, we use that $c(T) \leq c(OPT)$. It remains to show that $c(M) \leq \frac{1}{2}c(OPT)$. Using [Fact A.2](#), we may assume OPT is a Hamiltonian cycle. Now, OPT itself can be decomposed into the union of two O -Joins, as visualized here:

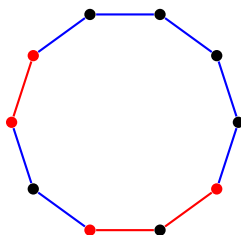


Figure 82: The set O of odd vertices of the minimum spanning tree is marked in red along the optimal Hamiltonian cycle. Notice OPT can be decomposed into two disjoint O -Joins, marked in blue and red. Therefore the minimum cost O -Join must cost at most $OPT/2$.

It follows that the minimum cost O -Join costs at most $OPT/2$, completing the proof. This analysis is tight, even allowing shortcutting, as seen in the following [Fig. 83](#).⁶³

A.5 Integrality Gap

Here we define the integrality gap of a polyhedron P in the positive orthant as the smallest number $\alpha \geq 1$ such that for every $x \in P$ and every positive weight function c , there exists an

⁶³In this example we do not use an optimal shortcutting. For a simple lower bound of $\frac{3}{2}$ even under optimal shortcutting, see for example [\[DT09\]](#).

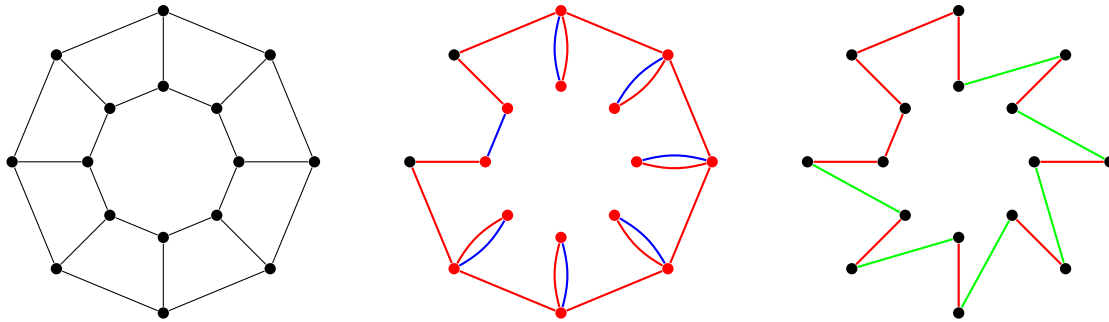


Figure 83: A tight example for Christofides. Let $c_e = 1$ for all edges. In red is the minimum spanning tree, and marked in red are the odd vertices. In blue is a minimum cost O -Join. The cost will be $\frac{3}{2}n - 2$, but the graph is Hamiltonian, so this is a lower bound of $3/2$ as $n \rightarrow \infty$. Note we can force Christofides to pick the red edges by making their costs just slightly cheaper than the remaining edges. On the right is a possible shortcutting of the Eulerian graph which also costs $\frac{3}{2}n - 2$, where the red edges have cost 1 and the green edges have cost 2. Note that this can be easily generalized: *any* Hamiltonian graphic instance with a spanning tree with all but a $o(1)$ fraction of vertices of even degree is a lower bound of $3/2$ for Christofides.

$x' \in P \cap \mathbb{Z}$ such that $c(x') \leq ac(x)$.⁶⁴

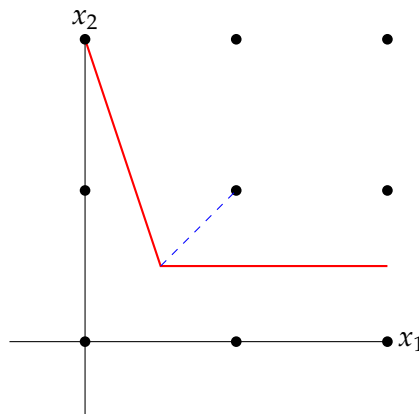


Figure 84: The integrality gap of this polyhedron is 2. It is achieved by the weight function $(1, 1)$ and the point $(0.5, 0.5)$. This has value 1, but the best integer coordinate has value 2.

For a convex polyhedron P , this is equivalent to studying how much the values of the following two programs can differ:

$$\begin{aligned} \min \quad & c^T z \\ \text{s.t.} \quad & Az \geq b \\ & z \in \mathbb{Z}_{\geq 0} \end{aligned} \quad (89)$$

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq b \\ & x \in \mathbb{R}_{\geq 0} \end{aligned} \quad (90)$$

Where A is the polyhedron P of interest in inequality form (possibly given by a separation oracle) and c is an arbitrary non-negative cost function. In other words, the integrality gap is the

⁶⁴Of course, an analogous definition can be given for maximization problems.

largest possible value of $\frac{c(z)}{c(x)}$ over all cost functions c where z, x are optimal solutions to the above programs (1) and (2) respectively.

This quantity is of significant interest when designing approximation algorithms, since often (2) can be solved efficiently but not (1). In this situation, proving that the integrality gap is α implies solving (2) estimate the cost of (1) up to a factor α .

B Proofs from Section 6

In all of the following lemmas, we assume that $\epsilon_\eta \leq \epsilon_{1/2}^2$ and $12\epsilon_{1/1} \leq \epsilon_{1/2}$.

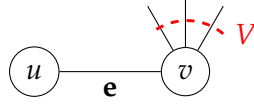


Figure 85: Setting of Lemma 6.28

Lemma 6.28. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \leq 1/2 - \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$ then, \mathbf{e} is 2-1-1 happy with probability at least $0.005\epsilon_{1/2}^2$.*

Proof. Let A, B, C be the degree partitioning of $\delta(u)$. Let $V := \delta(v)_{-\mathbf{e}}$ (see Fig. 85). Condition u, v be trees, \mathbf{e} and C to 0, let ν be the resulting measure. This happens with probability at least 0.5 and increases marginals in $A_{-\mathbf{e}}, B_{-\mathbf{e}}, V$ by at most $x_{\mathbf{e}} + 2\epsilon_{1/1} + \epsilon_\eta \leq x_{\mathbf{e}} + 2.1\epsilon_{1/1}$ and by tree conditioning decreases marginals by at most $2\epsilon_\eta$. After conditioning, we have

$$\begin{aligned} \mathbb{E}_\nu [A_T] &\in x(A) - x_{\mathbf{e}(A)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \subset [0.5, 1.5], \text{ similarly } \mathbb{E}_\nu [B_T] \subset [0.5, 1.5] \\ \mathbb{E}_\nu [V_T] &\in x(\delta(v)) - x_{\mathbf{e}} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \subset [1.5, 2.01] \\ \mathbb{E}_\nu [B_T + V_T] &\in x(B) + x(\delta(v)) - x_{\mathbf{e}} - x_{\mathbf{e}(B)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \subset [2 + 1.8\epsilon_{1/2}, 3.01], \\ \mathbb{E}_\nu [A_T + B_T] &\in x(A) + x(B) - x_{\mathbf{e}(A)} - x_{\mathbf{e}(B)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \subset [1.5, 2.01], \\ \mathbb{E}_\nu [A_T + B_T + V_T] &\in x(A) + x(B) + x(\delta(v)) - x_{\mathbf{e}} - x_{\mathbf{e}(A)} - x_{\mathbf{e}(B)} + [-2\epsilon_\eta, x_{\mathbf{e}} + 2.1\epsilon_{1/1}] \\ &\subset [3 + 1.75\epsilon_{1/2}, 4.01]. \end{aligned}$$

where we used $\epsilon_{1/2} \leq 0.001$ and $12\epsilon_{1/1} < \epsilon_{1/2}$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)}, x_{\mathbf{e}(A)} + x_{\mathbf{e}(B)} \leq x_{\mathbf{e}} \leq 1/2 - \epsilon_{1/2}$. It immediately follows from Proposition 6.8 that $\mathbb{P}_\nu [A_T = B_T = 1, V_T = 2]$ is at least a constant. In the rest of the proof, we do a more refined analysis. Using $A_T + B_T \geq 1, V_T \geq 1$,

$$\begin{aligned} \mathbb{P}_\nu [A_T + B_T + V_T = 4] &\geq (1.75\epsilon_{1/2})e^{-1.75\epsilon_{1/2}} \geq 1.7\epsilon_{1/2}, && \text{(Lemma 6.4)} \\ \mathbb{P}_\nu [A_T + B_T \geq 2], \mathbb{P}_\nu [V_T \geq 2] &\geq 0.39, && \text{(Lemma 6.5)} \\ \mathbb{P}_\nu [A_T + B_T \leq 2], \mathbb{P}_\nu [V_T \leq 2] &\geq 0.5, && \text{(Markov, } A_T + B_T \geq 1, V_T \geq 1 \text{ under } \nu) \\ \mathbb{P}_\nu [A_T \leq 1] &\geq 0.25, \mathbb{P}_\nu [B_T + V_T \leq 3] \geq 0.33. && \text{(Markov's Inequality and } V_T \geq 1 \text{ under } \nu) \\ \mathbb{P}_\nu [A_T \geq 1] &\geq 0.39, \mathbb{P}_\nu [B_T + V_T \geq 3] \geq 1.75\epsilon_{1/2}, && \text{(Lemma 6.5)} \end{aligned}$$

It follows by Corollary 6.12 (with $\epsilon = 0.195, p_m \geq 1 - 2\epsilon \geq 0.6$) that

$$\mathbb{P}_\nu [V_T = 2 | A_T + B_T + V_T = 4] \geq 0.13.$$

Note that since $V_T \geq 1, A_T + B_T \geq 1$ with probability 1, we apply Corollary 6.12 to $V_T - 1, A_T + B_T - 1$.

Furthermore, by Lemma 6.11, $\mathbb{P}_\nu [A_T \geq 1 | A_T + B_T + V_T = 4] \geq 0.128, \mathbb{P}_\nu [A_T \leq 1 | A_T + B_T + V_T = 4] \geq 0.43\epsilon_{1/2}$. The same holds for B_T . Therefore, by Corollary 6.12 (with $\epsilon = 0.055\epsilon_{1/2}$), using that $\epsilon_{1/2} < 0.001$,

$$\mathbb{P}_\nu [A_T = 1 | A_T + B_T = 2, V_T = 2] \geq 0.05\epsilon_{1/2}.$$

Putting these together we have

$$\begin{aligned}
\mathbb{P}[e \text{ 2-1-1 happy}] &\geq 0.5\mathbb{P}_\nu[A_T = B_T = 1, V_T = 2] \\
&= 0.5\mathbb{P}_\nu[A_T + B_T + V_T = 4] \mathbb{P}_\nu[V_T = 2|A_T + B_T + V_T = 4] \\
&\quad \cdot \mathbb{P}_\nu[A_T = 1|V_T = 2, A_T + B_T = 2] \\
&\geq 0.5(1.7\epsilon_{1/2})(0.13)(0.05\epsilon_{1/2}) \geq 0.005\epsilon_{1/2}^2
\end{aligned}$$

as desired. \square

Lemma 6.29. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a top edge bundle such that $x_{\mathbf{e}} \geq 1/2 + \epsilon_{1/2}$. If $\epsilon_{1/2} \leq 0.001$, then, \mathbf{e} is 2-1-1 happy with respect to u with probability at least $0.006\epsilon_{1/2}^2$.*

Proof. Let A, B, C be the **degree partitioning** of the edges in $\delta(u)$, $V = \delta_{-\mathbf{e}}(v)$. Condition u, v be trees, $C_T = 0$ and $u \cup v$ to be a tree (in order). This happens with probability at least $\frac{1}{2} + \epsilon_{1/2} - 3\epsilon_\eta - 2\epsilon_{1/1} \geq 0.5$. Let ν be the resulting measure restricted to edges in A, B, V . Note that ν on edges in A, B, V is SR. This is because ν is a product of two strongly Rayleigh distribution on the following two disjoint set of edges (i) the edges between u, v and (ii) the edges in $A_{-\mathbf{e}}, B_{-\mathbf{e}}, V$.

Furthermore, observe that under ν , every set of edges in $A_{-\mathbf{e}}, B_{-\mathbf{e}}, V$ increases by at most $2\epsilon_{1/1} + \epsilon_\eta < 0.2\epsilon_{1/2}$ (using $12\epsilon_{1/1} \leq \epsilon_{1/2}$), and decreases by at most $1 - x_{\mathbf{e}} + 2\epsilon_\eta$. Therefore,

$$\begin{aligned}
\mathbb{E}_\nu[A_T] &\in x(A) + [-(1 - x_{\mathbf{e}}) - 2\epsilon_\eta, 1 - x_{\mathbf{e}} + 0.2\epsilon_{1/2}] \subset [0.5, 1.5], \text{ similarly, } \mathbb{E}_\nu[B_T] \in [0.5, 1.5] \\
\mathbb{E}_\nu[V_T] &\in x(\delta(v)) - x_{\mathbf{e}} + [-(1 - x_{\mathbf{e}}) - 2\epsilon_\eta, 0.2\epsilon_{1/2}] \subset [0.995, 1.5]. \\
\mathbb{E}_\nu[A_T + B_T] &\in x(A) + x(B) + 1 - x_{e(A)} - x_{e(B)} + [-(1 - x_{\mathbf{e}}) - 2\epsilon_\eta, 0.2\epsilon_{1/2}] \subset [1.995, 2.5], \\
\mathbb{E}_\nu[B_T + V_T] &\in x(B) + x(\delta(v)) - x_{\mathbf{e}} + [-(1 - x_{\mathbf{e}}) - 2\epsilon_\eta, 1 - x_{\mathbf{e}} + 0.2\epsilon_{1/2}] \subset [1.99, 3 - 1.75\epsilon_{1/2}]. \\
\mathbb{E}_\nu[A_T + B_T + V_T] &\in x(A) + x(B) + x(\delta(v)) + 1 - x_{\mathbf{e}} - x_{e(A)} - x_{e(B)} + [-(1 - x_{\mathbf{e}}) - 2\epsilon_\eta, 0.2\epsilon_{1/2}] \\
&\subset [2.99, 4 - 1.75\epsilon_{1/2}].
\end{aligned}$$

where in the upper bound on $\mathbb{E}_\nu[A_T], \mathbb{E}_\nu[B_T], \mathbb{E}_\nu[B_T + V_T]$ we used that the marginals of edges in the bundle \mathbf{e} can only increase by $1 - x_{\mathbf{e}}$ (in total) when conditioning $u \cup v$ to be a tree. So,

$$\begin{aligned}
\mathbb{P}_\nu[A_T + B_T + V_T = 3] &\geq \epsilon_{1/2}, && \text{(By Theorem 2.17)} \\
\mathbb{P}_\nu[A_T + B_T \geq 2] &\geq 0.63, \mathbb{P}_\nu[V_T \geq 1] \geq 0.63 && \text{(Lemma 6.5, } A_T + B_T \geq 1) \\
\mathbb{P}_\nu[A_T + B_T \leq 2] &\geq 0.25, \mathbb{P}_\nu[V_T \leq 1] \geq 0.25, && \text{(Markov Inequality, } A_T + B_T \geq 1) \\
\mathbb{P}_\nu[A_T \geq 1] &\geq 0.39, \mathbb{P}_\nu[B_T + V_T \geq 2] \geq 0.59 && \text{(Lemma 6.5)} \\
\mathbb{P}_\nu[A_T \leq 1] &\geq 0.25, \mathbb{P}_\nu[B_T + V_T \leq 2] \geq 1.75\epsilon_{1/2}, && \text{(Markov, In worst case } \mathbb{P}[B_T + V_T < 2] = 0)
\end{aligned}$$

It follows by [Corollary 6.12](#) (with $\epsilon = 0.157, p_m = 0.68$) that

$$\mathbb{P}_\nu[A_T + B_T = 2|A_T + B_T + V_T = 3] \geq 0.12.$$

Note that since $A_T + B_T \geq 1$ with probability 1, we apply [Corollary 6.12](#) to $A_T + B_T - 1, V_T$.

Furthermore, by [Lemma 6.11](#), $\mathbb{P}_\nu[A_T \geq 1|A_T + B_T + V_T = 3] \geq 0.68\epsilon_{1/2}$ and $\mathbb{P}_\nu[A_T \leq 1|A_T + B_T + V_T = 3] \geq 0.147$. By symmetry, the same holds for B_T . Therefore, by [Corollary 6.12](#),

$$\mathbb{P}_\nu[A_T = 1|A_T + B_T = 2, V_T = 1] \geq 0.09\epsilon_{1/2}.$$

where we used $\epsilon_{1/2} < 0.001$.

Finally,

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy}] \geq (0.09\epsilon_{1/2})0.12(\epsilon_{1/2})0.5 \geq 0.005\epsilon_{1/2}^2,$$

as desired. \square

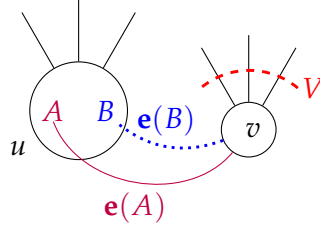


Figure 86: Setting of Lemma B.1

Lemma B.1. For a good half top edge bundle $\mathbf{e} = (\mathbf{u}, \mathbf{v})$, let A, B, C be the degree partitioning of $\delta(u)$, and let $V = \delta(v)_{-\mathbf{e}}$ (see Fig. 86). If $\epsilon_{1/2} \leq 0.001$, $x_{\mathbf{e}(B)} \leq \epsilon_{1/2}$, and $\mathbb{P}[(A_{-\mathbf{e}})_T + V_T \leq 1] \geq 5\epsilon_{1/2}$ then \mathbf{e} is 2-1-1 good,

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t. } u] \geq 0.005\epsilon_{1/2}^2$$

Proof. The proof is similar to Lemma 6.29. We condition u, v to be trees, $C_T = 0$, $u \cup v$ to be a tree. Let ν be the resulting SR measure on edges in A, B, V . The main difference is since $x_{\mathbf{e}} \not\geq 1/2 + \epsilon_{1/2}$ we use the lemma's assumptions to lower bound $\mathbb{P}_{\nu}[A_T + B_T + V_T = 3]$, $\mathbb{P}_{\nu}[A_T + V_T \leq 2]$, $\mathbb{P}_{\nu}[B_T + V_T \leq 2]$.

First, since \mathbf{e} is 2-2 good, by Lemma 6.22 and negative association,

$$\mathbb{P}_{\nu}[(\delta(u)_{-\mathbf{e}})_T + V_T \leq 2] \geq \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + V_T \leq 2] - \mathbb{P}[C_T = 0] \geq 0.4\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_{\eta} \geq 0.22\epsilon_{1/2},$$

where we used $\epsilon_{1/1} \leq \epsilon_{1/2}/12$. Letting $p_i = \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + V_T = i]$, we therefore have $p_{\leq 2} \geq 0.22\epsilon_{1/2}$. In addition, by Lemma 6.4, $p_3 \geq 1/4$. If $p_2 < 0.2\epsilon_{1/2}$, then from $p_2/p_3 \leq 0.8\epsilon_{1/2}$, we could use log-concavity to derive a contradiction to $p_{\leq 2} \geq 0.22\epsilon_{1/2}$ (analogously to what's done in the proof of Lemma 6.1). Therefore, we must have

$$\mathbb{P}_{\nu}[A_T + B_T + V_T = 3] = \mathbb{P}_{\nu}[(\delta(u)_{-\mathbf{e}})_T + V_T = 2] \geq 0.2\epsilon_{1/2}.$$

Next, notice since $\mathbb{P}[u, v, u \cup v \text{ trees}, C_T = 0] \geq 0.49$, by the lemma's assumption, $\mathbb{P}_{\nu}[\mathbf{e}(B)] \leq 2.01\epsilon_{1/2}$. Therefore,

$$\mathbb{E}_{\nu}[B_T + V_T] \leq x(V) + x(B) + 1.01\epsilon_{1/2} + 2\epsilon_{1/1} + \epsilon_{\eta} \leq 2.51.$$

So, by Markov, $\mathbb{P}_{\nu}[B_T + V_T \leq 2] \geq 0.15$. Finally, by negative association,

$$\mathbb{P}_{\nu}[A_T + V_T \leq 2] \geq \mathbb{P}_{\nu}[(A_{-\mathbf{e}})_T + V_T \leq 1] \geq \mathbb{P}[(A_{-\mathbf{e}})_T + V_T \leq 1] - \mathbb{P}[C_T = 0] \geq 4.8\epsilon_{1/2}$$

where we used the lemma's assumption.

Now, following the same line of arguments as in Lemma 6.29, we have $\mathbb{P}_{\nu}[A_T + B_T = 2 | A_T + B_T + V_T = 3] \geq 0.12$. Also, $\mathbb{P}_{\nu}[A_T \geq 1 | A_T + B_T + V_T = 3] \geq 3.02$, which implies $\mathbb{P}_{\nu}[A_T = 1 | A_T + B_T = 2, V_T = 1] \geq 0.42\epsilon$. This implies

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy}] \geq (0.42\epsilon_{1/2})0.12(0.2\epsilon_{1/2})0.498 \geq 0.005\epsilon_{1/2}^2$$

as desired. \square

Lemma 6.30. Let $\mathbf{e} = (\mathbf{v}, \mathbf{u})$ and $\mathbf{f} = (\mathbf{v}, \mathbf{w})$ be good half top edge bundles and let A, B, C be the *degree partitioning* of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(B)} \leq \epsilon_{1/2}$. Then, one of \mathbf{e}, \mathbf{f} is 2-1-1 happy with probability at least $0.005\epsilon_{1/2}^2$.

Proof. Let $U = \delta(u)_{-\mathbf{e}}$. By [Lemma 6.7](#), we can assume, without loss of generality, that

$$\mathbb{E}[U_T | \mathbf{f} \notin T, u, v, w \text{ tree}] \leq x(U_T) + 0.405 + 3\epsilon_\eta. \quad (91)$$

On the other hand,

$$\begin{aligned} \mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T] &\geq \mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T | \mathbf{f} \notin T, u, v, w \text{ tree}] \mathbb{P}[\mathbf{f} \notin T, u, v, w \text{ tree}] \\ &\geq \mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T | \mathbf{f} \notin T, u, v, w \text{ tree}] 0.49 \end{aligned}$$

So,

$$\mathbb{E}[(A_{-\mathbf{e}-\mathbf{f}})_T | \mathbf{f} \notin T, u, v, w \text{ tree}] \leq \frac{1}{0.49} x(A_{-\mathbf{e}-\mathbf{f}}) \leq \frac{1}{0.49} (4\epsilon_{1/2} + \epsilon_\eta) \leq 8.2\epsilon_{1/2}. \quad (92)$$

Combining (91) and (92), we get $\mathbb{E}[U_T + (A_{-\mathbf{e}})_T | \mathbf{f} \notin T, u, v, w \text{ tree}] \leq 1.91$ where we used $\epsilon_{1/2} \leq 0.001$. Therefore, using [Lemma 6.4](#), we get

$$\mathbb{P}[U_T + (A_{-\mathbf{e}})_T \leq 1] \geq 0.49 \mathbb{P}[U_T + (A_{-\mathbf{e}})_T \leq 1 | \mathbf{f} \notin T, u, v, w \text{ tree}] \geq 0.01,$$

Since $\epsilon_{1/2} \leq 0.001$, by [Lemma B.1](#), \mathbf{e} is 2-1-1 good. \square

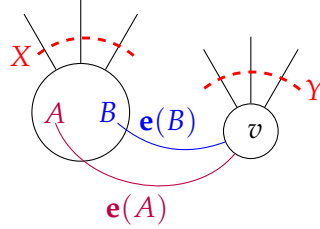


Figure 87: Setting of [Lemma 6.31](#).

Lemma 6.31. Let $\mathbf{e} = (\mathbf{u}, \mathbf{v})$ be a good half edge bundle and let A, B, C be the *degree partitioning* of $\delta(u)$ (see [Fig. 87](#)). If $\epsilon_{1/2} \leq 0.001$ and $x_{\mathbf{e}(A)}, x_{\mathbf{e}(B)} \geq \epsilon_{1/2}$, then

$$\mathbb{P}[\mathbf{e} \text{ 2-1-1 happy w.r.t } u] \geq 0.02\epsilon_{1/2}^2.$$

Proof. Condition C_T to be zero, u, v and $u \cup v$ be trees. This happens with probability at least 0.49. Let ν be the resulting measure. Let $X = A_{-\mathbf{e}} \cup B_{-\mathbf{e}}, Y = \delta(v)_{-\mathbf{e}}$. Since \mathbf{e} is 2-2 good by [Lemma 6.22](#) and stochastic dominance,

$$\mathbb{P}_\nu[X_T + Y_T \leq 2] \geq \mathbb{P}[(\delta(u)_{-\mathbf{e}})_T + Y_T \leq 2] - \mathbb{P}[C_T = 0] \geq 0.4\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta \geq 0.22\epsilon_{1/2},$$

where we used $\epsilon_{1/1} < 12\epsilon_{1/2}$. It follows by log-concavity of $X_T + Y_T$ that $\mathbb{P}_\nu[X_T + Y_T = 2] \geq 0.2\epsilon_{1/2}$. Now,

$$\mathbb{E}_\nu[X_T], \mathbb{E}_\nu[Y_T] \in [1 - 3\epsilon_{1/1}, 1.5 + \epsilon_{1/2} + 2\epsilon_{1/1} + 3\epsilon_\eta] \subset [0.995, 1.51]$$

So,

$$\mathbb{P}_\nu [X_T \geq 1], \mathbb{P}_\nu [Y_T \geq 1] \geq 0.63, \quad (\text{Lemma 6.5})$$

$$\mathbb{P}_\nu [X_T \leq 1], \mathbb{P}_\nu [Y_T \leq 1] \geq 0.245. \quad (\text{Markov})$$

Therefore, by [Corollary 6.12](#) $\mathbb{P}_\nu [X_T = 1 | X_T + Y_T = 2] \geq 0.119$.

$$\mathbb{P}_\nu [X_T = Y_T = 1] \geq (0.2\epsilon_{1/2})0.119 \geq 0.023\epsilon_{1/2},$$

Let \mathcal{E} be the event $\{X_T = Y_T = 1 | \nu\}$. Note that in ν we always choose exactly 1 edge from the \mathbf{e} bundle and that is independent of edges in X, Y , in particular the above event. Therefore, we can correct the parity of A, B by choosing from e_A or e_B . It follows that

$$\mathbb{P} [\mathbf{e} \text{ 2-1-1 happy w.r.t } u] \geq \mathbb{P}_\nu [\mathcal{E}] (1.99\epsilon_{1/2})0.49 \geq 0.02\epsilon_{1/2}^2,$$

where we used that $\mathbb{E}_\nu [\mathbf{e}(A)_T] \geq 1.99\epsilon_{1/2}$, and the same fact for $\mathbf{e}(B)_T$. To see why this latter fact is true, observe that conditioned on u, v trees, we always sample at most one edge between u, v . Therefore, since under ν we choose exactly one edge between u, v , the probability of choosing from $e(A)$ (and similarly choosing from $\mathbf{e}(B)$) is at least

$$\frac{\mathbb{E} [\mathbf{e}(A)_T | u, v \text{ trees}, C_T = 0]}{\mathbb{P} [\mathbf{e} | u, v \text{ trees}, C_T = 0]} \geq \frac{x_{\mathbf{e}(A)} - 2\epsilon_\eta}{x_{\mathbf{e}} + 3\epsilon_{1/1}} \geq \frac{\epsilon_{1/2} - 2\epsilon_\eta}{1/2 + 1.3\epsilon_{1/2}} \geq 1.99\epsilon_{1/2}$$

as desired. □

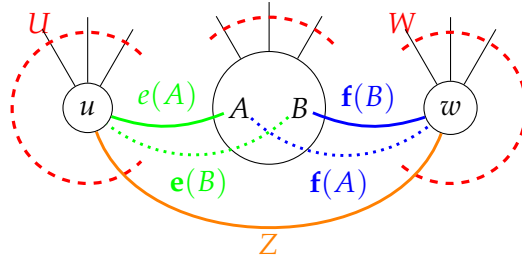


Figure 88: Setting of [Lemma 6.34](#). We assume that the dotted green/blue edges are at most $\epsilon_{1/2}$. Note that edges of C are not shown.

Lemma 6.34. *Let $\mathbf{e} = (\mathbf{u}, \mathbf{v}), \mathbf{f} = (\mathbf{v}, \mathbf{w})$ be two good top half edge bundles and let A, B, C be [degree partitioning](#) of $\delta(v)$ such that $x_{\mathbf{e}(B)}, x_{\mathbf{f}(A)} \leq \epsilon_{1/2}$. If \mathbf{e}, \mathbf{f} are not 2-1-1 good with respect to v , and $\epsilon_{1/2} \leq 0.0002$, then \mathbf{e}, \mathbf{f} are 2-2-2 happy with probability at least 0.01.*

Proof. First, observe that by [Lemma B.1](#) if $\mathbb{P} [U_T + (A_{-\mathbf{e}})_T \leq 1] \geq 0.25\epsilon$, where $\epsilon \geq 20\epsilon_{1/2}$ is a constant that we fix later, then \mathbf{e} is 2-1-1 good, which is a contradiction. So, assume, $\mathbb{P} [U_T + (A_{-\mathbf{e}})_T \geq 2] \geq 1 - 0.25\epsilon$. Furthermore, let $q = \mathbb{P} [U_T + (A_{-\mathbf{e}})_T \geq 3]$. Since $x(U) + x(A_{-\mathbf{e}}) \leq 2 + 3\epsilon_{1/2} + 2\epsilon_{1/1} + 3\epsilon_\eta \leq 2 + 3.2\epsilon_{1/2}$ (where we used $x_{\mathbf{e}(A)} \geq x_{\mathbf{e}} - x_{\mathbf{e}(B)} - x_C \geq 1/2 - 2\epsilon_{1/2} - 2\epsilon_{1/1} - \epsilon_\eta$ and where we used $12\epsilon_{1/1} \leq \epsilon_{1/2}$),

$$2(1 - q - 0.25\epsilon) + 3q \leq 2 + 3.2\epsilon_{1/2}.$$

This implies that $q \leq 0.5\epsilon + 3.2\epsilon_{1/2} \leq 0.75\epsilon$ (for $\epsilon \geq 13\epsilon_{1/2}$). Therefore,

$$\mathbb{P}[U_T + (A_{-\mathbf{e}})_T = 2], \mathbb{P}[W_T + (B_{-\mathbf{f}})_T = 2] \geq 1 - \epsilon \quad (93)$$

where the second inequality follows by a similar argument.

Claim B.2. *Let $Z = \delta(u) \cap \delta(w)$. If $\epsilon < 1/15$, then either $\mathbb{E}[Z|u, v, w \text{ tree}] \leq 3\epsilon$ or $\mathbb{E}[Z|u, v, w \text{ tree}] \geq (1 - 3\epsilon)$.*

Proof. For the whole proof we work with μ conditioned on u, v, w are trees. Let $z = \mathbb{E}[Z]$. Let $D = U \cup W \cup A_{-\mathbf{e}} \cup B_{-\mathbf{f}} \setminus Z$. Note that $D_T + 2Z_T = U_T \cup W_T \cup (A_{-\mathbf{e}})_T \cup (B_{-\mathbf{f}})_T$. By Eq. (93) and a union bound $\mathbb{P}[D_T + 2Z_T = 4] \geq 1 - 2\epsilon - 3\epsilon_\eta$. Therefore,

$$2.1\epsilon \geq 2\epsilon + 3\epsilon_\eta \geq \mathbb{P}[D_T + 2Z_T \neq 4] \geq \mathbb{P}[D_T = 3] \geq \sqrt{\mathbb{P}[D_T = 2] \mathbb{P}[D_T = 4]}$$

where the last inequality follows by log-concavity. On the other hand,

$$\begin{aligned} z &= \mathbb{P}[Z = 1] \leq \mathbb{P}[D_T = 2, Z = 1] + \mathbb{P}[D_T + 2Z_T \neq 4] \leq \mathbb{P}[D_T = 2] + 2.1\epsilon, \\ 1 - z &= \mathbb{P}[Z = 0] \leq \mathbb{P}[D_T = 4, Z = 0] + \mathbb{P}[D_T + 2Z_T \neq 4] \leq \mathbb{P}[D_T = 4] + 2.1\epsilon \end{aligned}$$

Putting everything together,

$$(2.1\epsilon)^2 \geq (z - 2.1\epsilon)(1 - z - 2.1\epsilon) = z(1 - z) - 2.1\epsilon + 2.1\epsilon^2.$$

Therefore, using $\epsilon \leq 1/15$, we get that either $z \leq 3\epsilon$ or $z \geq 1 - 3\epsilon$. \square

So, for the rest of proof we assume $\mathbb{E}[Z_T|u, v, w \text{ trees}] < 3\epsilon$. A similar proof shows \mathbf{e}, \mathbf{f} are 2-2-2 good when $\mathbb{E}[Z_T|u, v, w \text{ trees}] > 1 - 3\epsilon$. We run the following conditionings in order: u, v, w trees, $Z_T = 0, C_T = 0, \mathbf{e}(B), \mathbf{f} \notin T, \mathbf{e}(A) \in T$. Note that $\mathbf{e}(A) \in T$ is equivalent to $u \cup v$ be a tree. Call this event \mathcal{E} (i.e., the event that all things we conditioned on happen). First, notice

$$\mathbb{P}[\mathcal{E}] \geq (1 - 3\epsilon_\eta)(1 - 3\epsilon - 2\epsilon_{1/1} - \epsilon_\eta - \epsilon_{1/2} - (1/2 + \epsilon_{1/2}))(1/2 - 3\epsilon_{1/2}) \geq 0.22 \geq 1/5 \quad (94)$$

Moreover, since all of these conditionings correspond to upward/downward events, $\mu|\mathcal{E}$ is strongly Rayleigh. The main statement we will show is that

$$\mathbb{P}[\mathbf{e}, \mathbf{f} \text{ 2-2-2 happy}|\mathcal{E}] \geq \mathbb{P}[U_T = (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0, W_T = 2|\mathcal{E}] = \Omega(1).$$

The main insight of the proof is that Eq. (93) holds (up to a larger constant of ϵ), even after conditioning $\mathcal{E}, B_{-\mathbf{f}} = 0, A_{-\mathbf{e}} = 1$; so, we can bound the preceding event by just a union bound. The main non-trivial statement is to argue that the expectations of $B_{-\mathbf{f}}$ and $A_{-\mathbf{e}}$ do not change so much under \mathcal{E} .

Combining (93) and (94),

$$\mathbb{P}[U_T + (A_{-\mathbf{e}})_T = 2|\mathcal{E}], \mathbb{P}[W_T + (B_{-\mathbf{f}})_T = 2|\mathcal{E}] \geq 1 - 5\epsilon. \quad (95)$$

We claim that

$$\mathbb{E}[B_T|\mathcal{E}] = \mathbb{E}[(B_{-\mathbf{f}})_T|\mathcal{E}] \leq x(B_{-\mathbf{f}}) + 3\epsilon_\eta + 3\epsilon_{1/1} + \epsilon_{1/2} + 35\epsilon \leq 0.66 \quad (96)$$

using $\epsilon_{1/2} < 0.0002$ and $\epsilon = 20\epsilon_{1/2}$. To see this, observe that after each conditioning in \mathcal{E} either all marginals increase or all decrease. Furthermore, the events $C_T = 0, Z_T = 0, \mathbf{e}(B)_T = 0$ can increase marginals by at most $3\epsilon_\eta + 3\epsilon_{1/1} + \epsilon_{1/2}$; the only other event that can increase $B_{-\mathbf{f}}$ is $\mathbf{f} \notin T$. Now we know $\mathbb{P}[(B_{-\mathbf{f}})_T + W_T = 2|\mathcal{E}] \geq 1 - 5\epsilon$ before and after conditioning $\mathbf{f} \notin T$. Therefore, by Corollary 2.19, $2 - 10\epsilon \leq \mathbb{E}[(B_{-\mathbf{f}})_T + W_T] \leq 2 + 25\epsilon$. But if $\mathbb{E}[(B_{-\mathbf{f}})_T]$ increased by more than 35ϵ , then either before conditioning $\mathbf{f} \notin T$, $\mathbb{E}[(B_{-\mathbf{f}})_T + W_T] < 2 - 10\epsilon$ or afterwards it is more than $2 + 25\epsilon$, which is a contradiction, and completes the proof of (96). A similar argument shows that $\mathbb{E}[(A_{-\mathbf{e}})_T|\mathcal{E}] \leq 0.66$.

We also claim that

$$\mathbb{E}[(A_{-\mathbf{e}})_T|\mathcal{E}] \geq x(A_{-\mathbf{e}}) - 3\epsilon_\eta - 35\epsilon \geq 0.33.$$

As above, everything conditioned on in \mathcal{E} increases $\mathbb{E}[(A_{-\mathbf{e}})_T]$ except for possibly $\mathbf{e}(A) \in T$. As above, we know that $\mathbb{P}[U_T + (A_{-\mathbf{e}})_T = 2|\mathcal{E}] \geq 1 - 5\epsilon$ before and after $\mathbf{e}(A) \notin T$. So again applying Corollary 2.19, we see that it can't decrease by more than 35ϵ .

It follows that

$$0.33 \leq \mathbb{E}[(A_{-\mathbf{e}})_T|\mathcal{E}] \leq \mathbb{E}[(A_{-\mathbf{e}})_T|\mathcal{E}, (B_{-\mathbf{f}})_T = 0] \leq 0.66 + 0.66 \leq 1.32.$$

So, by Lemma 6.4 and Theorem 2.17, $\mathbb{P}[(A_{-\mathbf{e}})_T = 1|\mathcal{E}, (B_{-\mathbf{f}})_T = 0] \geq 0.33e^{-.33} \geq 0.237$.

Therefore, by Lemma 6.4

$$\mathbb{P}[\mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0] \geq (0.22)(0.39)(0.23) \geq 0.019.$$

Therefore, by (95)

$$\mathbb{P}[U_T = 1|\mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0], \mathbb{P}[W_T = 2|\mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0] \geq 1 - 5\epsilon/0.019$$

Finally, by union bound

$$\mathbb{P}[U_T = 1, W_T = 2|\mathcal{E}, (A_{-\mathbf{e}})_T = 1, (B_{-\mathbf{f}})_T = 0] \geq 1 - \epsilon/0.009$$

Using $\epsilon = 20\epsilon_{1/2}$ and $\epsilon_{1/2} \leq 0.0002$ this means both of the above events happens, so \mathbf{e}, \mathbf{f} are 2-2-2-happy with probability $0.019(1 - \epsilon/0.009) > 0.01$ as desired. \square

C Cuts crossed on one side

This is the first appendix for the integrality gap result, [Theorem 1.2](#).

C.1 Cuts crossed on one side

In [Theorem 11.2](#), we found a vector which satisfies all cuts crossed on both sides. Consequently, we can study the structure of cuts which remain after deleting all cuts crossed on both sides, i.e. connected components of cuts crossed on one side. The arguments in this section closely follow [Section 5.4](#). Even though in that section, these families were handled using OPT edges, the extension to charging to LP edges is very natural in this setting and requires little modification.

Lemma C.1. *Suppose \mathcal{C} is a connected component of cuts in $\mathcal{N}_{\eta,1}$ with $|\mathcal{C}| \geq 2$. If $\eta \leq \frac{2}{5}$, the corresponding polygon P of \mathcal{C} has no inside atoms.*

Proof. By [Lemma 10.17](#), to show that the polygon of \mathcal{C} has no inside atoms it is sufficient to show that there are no k -cycles for any integer k . Since $\eta \leq 2/5$, by [Lemma 10.16](#) there are no 3 or 4-cycles. By way of contradiction suppose there was a k -cycle $C_1, \dots, C_k \in \mathcal{C}$ with $k \geq 5$. Then, perhaps after renaming, we can assume that C_1, C_2, C_3 do not contain the root, C_1 and C_3 each cross C_2 , and $C_1 \cap C_3 = \emptyset$. Then by [Lemma 10.12](#), $C_2 \in \mathcal{N}_{\eta,2}$, which is a contradiction since we assumed $\mathcal{C} \subseteq \mathcal{N}_1$. Therefore, P has no k -cycles for $k \geq 5$. So, P has no inside atoms. \square

Also note that by [Lemma 10.12](#), if \mathcal{C} is a connected component of cuts in $\mathcal{N}_{\eta,1}$, then every cut $C \in \mathcal{C}$ is crossed on one side in the polygon of \mathcal{C} . (In other words, deleting the cuts in $\mathcal{N}_{\eta,2}$ does not allow a cut previously crossed on one side to be crossed on both sides in its new polygon.)

C.2 Notation and results that we re-use

As before suppose \mathcal{C} is a connected component of cuts crossed on one side with corresponding polygon P . Now assume P has outside atoms a_0, \dots, a_{m-1} , and WLOG assume a_0 is the root (recall there are no inside atoms).

Definition C.2 (Leftmost and Rightmost cuts). *We call any cut $C \in \mathcal{C}$ with leftmost atom a_1 a leftmost cut of P , and any cut $C \in \mathcal{C}$ with rightmost atom a_{m-1} a rightmost cut of P . We also call a_1 the leftmost atom of P (resp. a_{m-1} the rightmost atom).*

Recall that in [Theorem 5.9](#), we showed that polygons of cuts crossed on one side have a simple structure. In particular, they look like a near-integral cycle. We repeat the theorem here:

Theorem C.3 (Structure of Polygons of $\mathcal{N}_{\eta,1}$). *For $\epsilon_\eta \geq 7\eta$ and any polygon of cuts crossed on one side with atoms $a_0 \dots a_{m-1}$ (where a_0 is the root) the following holds:*

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

We will also re-use the following definitions:

Definition C.4 (*A, B, C-Polygon Partition*). The *A, B, C-polygon partition* of a polygon P is a partition of edges of $\delta(a_0)$ into sets $A = E(a_1, a_0)$ and $B = E(a_{m-1}, a_0)$, $C = \delta(a_0) \setminus A \setminus B$.

Definition C.5 (*Happy Polygons*). For a spanning tree T , we say that a polygon P of cuts crossed on one side is happy if

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

We say that P is left-happy (respectively right-happy) if

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

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

Definition C.6 (*Happy Cut*). We say a leftmost cut $L \in \mathcal{C}$ is happy if

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

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

Note that, by definition, if leftmost cut L is happy and P is left happy then L is even, i.e., $\delta(L)_T = 2$. Similarly, a_1 is even if it is happy and P is left-happy.

Definition C.7 (*Relevant Cuts*). Define the family of relevant cuts of a polygon P representing a connected component $\mathcal{C} \subseteq \mathcal{N}_{\eta,1}$ as follows:

$$\mathcal{C}_+ = \mathcal{C} \cup \{a_i : 1 \leq i \leq m-1 \wedge x(\delta(a_i)) \leq 2 + \eta\}.$$

Lemma C.8 (Restatement of [Lemma 5.28](#)). There is a mapping of cuts in \mathcal{C}_+ to the collections of edges $E(a_1, a_2), \dots, E(a_{m-2}, a_{m-1})$ such that each set $E(a_i, a_{i+1})$ has at most 4 cuts mapped to it, every cut $C \in \mathcal{C}_+$ containing atoms a_i through a_j is mapped to either $E(a_{i-1}, a_i)$ or $E(a_j, a_{j+1})$ (or both), and every atom of the polygon in \mathcal{C}' gets mapped to two (not necessarily distinct) groups of edges $E(a_i, a_{i+1}), E(a_j, a_{j+1})$.

Note in the following three statements, we gain a factor of two compared to their previous incarnations as we look at η -near min cuts instead of 2η -near min cuts.

Lemma C.9 (Restatement of [Lemma 5.26](#)). For every cut $A \in \mathcal{C}$ that is not a leftmost or a rightmost cut, $\mathbb{P}[\delta(A)_T = 2] \geq 1 - 11\eta$.

Lemma C.10 (Restatement of [Lemma 5.27](#)). For any atom $a_i \neq a_0$ that is not the leftmost or the rightmost atom we have

$$\mathbb{P}[\delta(a_i)_T = 2] \geq 1 - 21\eta.$$

Lemma C.11 (Restatement of [Lemma 5.30](#)). For every leftmost or rightmost cut A in P that is an η -near min cut, $\mathbb{P}[A \text{ happy}] \geq 1 - 5\eta$, and for the leftmost atom a_1 (resp. rightmost atom a_{m-1}), if it is an η -near min cut then $\mathbb{P}[a_1 \text{ happy}] \geq 1 - 12\eta$ (resp. $\mathbb{P}[a_{m-1} \text{ happy}] \geq 1 - 12\eta$).

C.3 Main theorem for cuts crossed on one side

The following is the extension of [Theorem 5.24](#) to the case when we do not use OPT. This is the key theorem used to deal with components cuts crossed on one side and the atoms in \mathcal{N}_η which compose them.

Theorem C.12 (Happy Polygons (Similar to [Theorem 5.24](#))). *Let $G = (V, E, x)$ for an LP solution x . Let μ be an arbitrary distribution of spanning trees with marginals x . For any $\alpha > 0$, $\eta \leq 1/10$, and $\epsilon_\eta = 7\eta$, there is a random vector $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (as a function of $T \sim \mu$) such that*

- For a connected component \mathcal{C} of cuts crossed on one side with corresponding polygon P and atoms a_0, a_1, \dots, a_{m-1} and cycle partition A, B, C the following holds:
 - For any cut $S \in \mathcal{C}_+$ which is not a leftmost/rightmost cut/atom if $\delta(S)_T$ is odd then we have $s^*(\delta(S)) \geq \alpha(1 - \epsilon_\eta)$,
 - If P is left happy, then for any $S \in \mathcal{C}_+$ that is a leftmost cut or the leftmost atom, if $\delta(S)_T$ is odd, then we have $s^*(\delta(S)) \geq \alpha(1 - \epsilon_\eta)$.
 - Similarly, if P is right happy then for any cut $S \in \mathcal{C}_+$ that is a rightmost cut or the rightmost atom, if $\delta(S)_T$ is odd, then $s^*(\delta(S)) \geq \alpha(1 - \epsilon_\eta)$.
- $\mathbb{E}[s_e^*] \leq 44\alpha\eta x_e$ for all $e \in E$.

Before proving the theorem, we study a special case.

Lemma C.13 ([Theorem C.12](#) Holds for Triangles). *Let $S = X \cup Y$ where X, Y, S are ϵ_η -near min cuts which do not cross. Then, letting X be a_1 and Y be a_2 (and $a_0 = \overline{X \cup Y}$) [Theorem C.12](#) holds.*

Proof. In this case this system has cycle partition $A = E(a_1, a_0), B = E(a_2, a_0), C = \emptyset$. For the edges $E(a_1, a_2)$ we define an increase event \mathcal{I} when at least one of $T \cap E(X), T \cap E(Y), T \cap E(S)$ is not a tree. Whenever this happens we define $s_e^* = \alpha x_e$ for all $e \in E(a_1, a_2)$. If S is left-happy we need to show when $\delta(X)_T$ is odd, then $s^*(\delta(X)) \geq \alpha(1 - \epsilon_\eta)$. This is because when S is left-happy we have $A_T = 1$ (and $C_T = 0$), so either the increase event \mathcal{I} does not happen and we get $\delta(X)_T = 2$ or it happens in which case $s^*(\delta(X)) = \alpha \cdot x(E(a_1, a_2)) \geq \alpha(1 - \epsilon_\eta)$ by ???. Finally, observe that by [Corollary 2.29](#), $\mathbb{P}[\mathcal{I}] \leq 3\epsilon_\eta/2$, so $\mathbb{E}[s_e^*] = 1.5\epsilon_\eta\alpha x_e < 44\alpha\eta x_e$ for all $e \in E(a_1, a_2)$. \square

Proof of [Theorem C.12](#). Fix a connected component \mathcal{C} of \mathcal{N}_1 with corresponding polygon P . Fix $1 < i < m$. By [Theorem C.3](#) $x(E(a_{i-1}, a_i)) \geq 1 - \epsilon_\eta$.

For the at most four cuts mapped to $E(a_{i-1}, a_i)$ in [Lemma C.8](#), we define the following three events:

- i) A leftmost cut mapped to $E(a_{i-1}, a_i)$ is not happy.
- ii) A rightmost cut mapped to $E(a_{i-1}, a_i)$ is not happy.
- iii) A cut which is not leftmost or rightmost mapped to $E(a_{i-1}, a_i)$ is odd.

Observe that the cuts in (i) and (ii) are assigned to $E(a_{i-1}, a_i)$ in [Lemma C.8](#). We say an atom a is singly-mapped to $E(a_{i-1}, a_i)$ if in the matching a is only mapped to $E(a_{i-1}, a_i)$ once, otherwise we say it is doubly-mapped to $E(a_{i-1}, a_i)$.

We say an event $\mathcal{I}(E(a_{i-1}, a_i))$ occurs if either (i), (ii), or (iii) occurs. If $\mathcal{I}(E(a_{i-1}, a_i))$ occurs then for all $e \in E(a_{i-1}, a_i)$, we set:

$$s_e^* = \begin{cases} \alpha x_e & \text{If (i),(ii), or (iii) occurred for at least one non-atom cut in } \mathcal{C}', \text{ or for an atom} \\ & \text{which is doubly-mapped to } E(a_{i-1}, a_i) \\ \alpha x_e/2 & \text{Otherwise.} \end{cases}$$

If $\mathcal{I}(E(a_{i-1}, a_i))$ does not occur we set $s_e^* = 0$ for all $e \in E(a_{i-1}, a_i)$.

First, observe that for any non-atom cut $S \in \mathcal{C}_+$ (i.e. any relevant cut) that is not a leftmost or a rightmost cut/atom, if $\delta(S)_T$ is odd, then if $E(a_{i-1}, a_i)$ is the set of edges that S is mapped to, it satisfies $s^*(\delta(S)) \geq \alpha \cdot x(E(a_{i-1}, a_i)) \geq \alpha(1 - \epsilon_\eta)$. So, these cuts satisfy the conditions of the theorem.

The same inequality holds for non-leftmost/rightmost atom cuts $a \in \mathcal{C}'$ which are doubly-mapped to $E(a_{i-1}, a_i)$. For non-leftmost/rightmost atom cuts $a \in \mathcal{C}'$ which are singly-mapped to $E(a_{i-1}, a_i)$, a is mapped (possibly even twice) to another edge $E(a_{j-1}, a_j)$ (note $j = i - 1$ or $i + 1$), and in this case $s^*(\delta(S)) \geq \alpha/2 \cdot 2(1 - \epsilon_\eta) = \alpha(1 - \epsilon_\eta)$, and again the above inequality holds.

Now, suppose $S \in \mathcal{C}$ is a leftmost cut of P and $\delta(S)_T$ is odd, and the rightmost atom of S is a_{i-1} (i.e. it is mapped to $E(a_{i-1}, a_i)$). If P is not left-happy then there is nothing to prove. If P is left-happy, we may assume S is not happy. Then $\mathcal{I}(E(a_{i-1}, a_i))$ happens, so as in the above inequality $s^*(\delta(S)) \geq \alpha(1 - \epsilon_\eta)$. We obtain the same condition for rightmost cuts and leftmost/rightmost atoms that are assigned to P (note leftmost/rightmost atoms are always doubly-mapped: a_1 to $E(a_1, a_2)$ and a_{m-1} to $E(a_{m-2}, a_{m-1})$).

It remains to upper bound $\mathbb{E}[s_e^*]$ for any edge $e \in E(a_{i-1}, a_i)$. By [Lemma C.8](#), at most four cuts are mapped to $E(a_{i-1}, a_i)$.

First suppose exactly one atom is doubly-mapped to $E(a_{i-1}, a_i)$. Then there are at most three cuts mapped to $E(a_{i-1}, a_i)$, including that atom. The probability of an event of type (i) or (ii) occurring for the leftmost or rightmost atom is at most $1 - 12\eta$ by [Lemma C.11](#). Atoms which are not leftmost or rightmost are even with probability at least $1 - 21\eta$ by [Lemma C.10](#). Therefore, in the worst case, the doubly-mapped atom is not leftmost or rightmost. For the remaining two cuts, leftmost and rightmost cuts are happy with probability at least $1 - 5\eta$ by [Lemma C.11](#), and (non-atom) non leftmost/rightmost cuts are even with probability at least $1 - 11\eta$ by [Lemma C.9](#). Therefore in the worst case the remaining two (non-atom) cuts mapped to $E(a_{i-1}, a_i)$ are not leftmost/rightmost. Therefore, if an atom is doubly-mapped to $E(a_{i-1}, a_i)$, for any $e \in E(a_{i-1}, a_i)$ we have

$$\mathbb{E}[s^*(e)] \leq 21\eta\alpha x_e + 2 \cdot 11\eta\alpha x_e < 44\eta\alpha x_e$$

Note if two atoms are doubly-mapped to $E(a_{i-1}, a_i)$, there are no other mapped cuts and in the worst case the atoms are not leftmost/rightmost, so for any $e \in E(a_{i-1}, a_i)$,

$$\mathbb{E}[s^*(e)] \leq 2 \cdot 21\eta\alpha x_e < 44\eta\alpha x_e$$

Otherwise, any atoms mapped to $E(a_{i-1}, a_i)$ are singly-mapped. In this case, if only an atom cut is odd/unhappy, we set $s^*(e) = x_e\alpha/2$. The probability of an event of type (i) or (ii) occurring for the leftmost or rightmost atom is at most $1 - 12\eta$ by [Lemma C.11](#), so we can bound the contribution of this event to $\mathbb{E}[s^*(e)]$ by $12\eta\alpha x_e/2$. Atoms which are not leftmost or rightmost are even with probability at least $1 - 21\eta$ by [Lemma C.10](#), and so we can bound their contribution by

$21\eta\alpha x_e/2$. Therefore, in the worst case four non-leftmost/rightmost *non*-atom cuts are mapped to $E(a_{i-1}, a_i)$, in which case, for any $e \in E(a_{i-1}, a_i)$,

$$\mathbb{E}[s^*(e)] \leq 4 \cdot 11\eta\alpha x_e = 44\eta\alpha x_e$$

as desired. □

D Proof of Theorem 12.1

In this section, we use the previous section and [Theorem 11.2](#) to prove [Theorem 12.1](#), the main technical ingredient required to prove the integrality gap result.

Definition D.1 (Hierarchy). *For an LP solution x^0 with support $E_0 = E \cup \{e_0\}$ where x is x^0 restricted to E , a hierarchy $\mathcal{H} \subseteq \mathcal{N}_{\epsilon_\eta}$ is a laminar family with root $V \setminus \{u_0, v_0\}$, where every cut $S \in \mathcal{H}$ is called either a “near-cycle” cut or a degree cut. In the special case that S has exactly two children we call it a triangle cut. Furthermore, every cut S is the union of its children. For any (non-root) cut $S \in \mathcal{H}$, define the parent of S , $\mathfrak{p}(S)$, to be the smallest cut $S' \in \mathcal{H}$ such that $S \subsetneq S'$.*

For a cut $S \in \mathcal{H}$, let $\mathcal{A}(S) := \{a \in \mathcal{H} : \mathfrak{p}(a) = S\}$. If S is called a “near-cycle” cut, then we can order cuts in $\mathcal{A}(S)$, a_1, \dots, a_{m-1} such that

- $A = E(\bar{S}, a_1), B = E(a_{m-1}, \bar{S})$ satisfy $x(A), x(B) \geq 1 - \epsilon_\eta$.
- For any $1 \leq i < m - 1$, $x(E(a_i, a_{i+1})) \geq 1 - \epsilon_\eta$.
- $C = \cup_{i=2}^{m-2} E(a_i, \bar{S})$ satisfies $x(C) \leq \epsilon_\eta$.

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

*We abuse notation and for an edge $e = (u, v)$ that is not a neighbor of u_0, v_0 , we write $\mathfrak{p}(e)$ to denote the smallest⁶⁵ cut $S' \in \mathcal{H}$ such that $u, v \in S'$. We say edge e is a **bottom edge** if $\mathfrak{p}(e)$ is a polygon cut and we say it is a **top edge** if $\mathfrak{p}(e)$ is a degree cut.*

The terminology of the above differs slightly from [Definition 5.31](#), where we replace “polygon” cut with “near-cycle” cut and “polygon” partition with “near-cycle” partition.

By [Theorem C.3](#), an example of a near-cycle cut is the union of non-root atoms of a connected component of cuts crossed on one side (i.e. its outer polygon cut). Another example is the non-root atoms of a connected component of minimum cuts (i.e. a cycle of a cactus of length at least four).

In the following, we will define a hierarchy \mathcal{H} satisfying the above definition such that every cut $S \in \mathcal{N}_{\eta, \leq 1}$ is either in \mathcal{H} or there is a near-cycle cut $P \in \mathcal{H}$ representing a connected component \mathcal{C} such that $S \in \mathcal{C}$.

Recall the “main payment theorem.”

Theorem D.2. *For an LP solution x^0 where x is x^0 restricted to E and a hierarchy \mathcal{H} for some $\epsilon_\eta \leq 10^{-10}$ and any $\beta > 0$, the maximum entropy distribution μ with marginals x satisfies the following:*

- i) *There is a set of good edges $E_g \subseteq E \setminus \delta(\{u_0, v_0\})$ such that any bottom edge e is in E_g and for any (non-root) $S \in \mathcal{H}$ such that $\mathfrak{p}(S)$ is not a near-cycle cut, we have $x(E_g \cap \delta(S)) \geq 3/4$.*

⁶⁵in the sense of the number of vertices that it contains

- ii) There is a random vector $s : E_g \rightarrow \mathbb{R}$ (as a function of $T \sim \mu$) such that for all e , $s_e \geq -x_e\beta$ (with probability 1), and
- iii) If a near-cycle cut S with cycle partition A, B, C is not left happy, then for any set $F \subseteq E$ with $\mathfrak{p}(e) = S$ for all $e \in F$ and $x(F) \geq 1 - \epsilon_\eta/2$, we have

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

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

- iv) For every cut $S \in \mathcal{H}$ such that $\mathfrak{p}(S)$ is not an near-cycle cut, if $\delta(S)_T$ is odd, then $s(\delta(S)) \geq 0$.
- v) For a good edge $e \in E_g$, $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ (where $\epsilon_P \geq 3.12 \cdot 10^{-16}$).

We will shortly show how the main payment theorem along with [Theorem C.12](#) and [Theorem 11.2](#) implies the following:

Theorem D.3. Let x^0 be a feasible solution of LP (2) with support $E_0 = E \cup \{e_0\}$ with x the restriction of x^0 to E . Let μ be the max entropy distribution with marginals x . For $\eta \leq 10^{-12}$, $\beta > 0$, there is a set $E_g \subset E \setminus \delta(\{u_0, v_0\})$ of good edges and two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$) such that

- (i) For each edge $e \in E_g$, $s_e \geq -x_e\beta$ and for any $e \in E \setminus E_g$, $s_e = 0$.
- (ii) For each η -near min cut S , including those for which $\{u_0, v_0\} \in \delta(S)$, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.
- (iii) We have $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ for all edges $e \in E_g$ and $\mathbb{E}[s_e^*] \leq 125\eta\beta x_e$ for all edges $e \in E$, where ϵ_P is defined in [Theorem 5.33](#).
- (iv) For every cut S crossed on at most one side such that $S \neq \overline{\{u_0, v_0\}}$, $x(\delta(S) \cap E_g) \geq 3/4$.

We will first use it to prove the appendix theorem, which we already showed implies [Theorem 1.2](#):

Theorem 12.1 (Combination of [Theorem 9.8](#) and [Theorem 11.2](#)). Let x^0 be a solution of LP (2) with support $E_0 = E \cup \{e_0\}$, and x be x^0 restricted to E . Let $\eta \leq 10^{-12}$, $\beta > 0$ and let μ be the max-entropy distribution with marginals x . Then there are two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$), such that

- i) For each edge $e \in E$, $s_e \geq -x_e\beta$ (with probability 1).
- ii) For each $S \in \mathcal{N}_\eta$, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.
- iii) For every edge e , $\mathbb{E}[s_e^*] \leq 125\eta\beta x_e$ and $\mathbb{E}[s_e] \leq -\frac{1}{3}x_e\epsilon_P\beta$, where ϵ_P is defined in [Theorem D.2](#).

Proof of Theorem 12.1. Let E_g be the good edges defined in [Theorem D.3](#) and let $E_b := E \setminus E_g$ be the set of bad edges; in particular, note all edges in $\delta(\{u_0, v_0\})$ are bad edges. We define a new vector $\tilde{s} : E \cup \{e_0\} \rightarrow \mathbb{R}$ as follows:

$$\tilde{s}(e) \leftarrow \begin{cases} \infty & \text{if } e = e_0 \\ -x_e(4\beta/5)(1 - 2\eta) & \text{if } e \in E_b, \\ x_e(4\beta/3) & \text{otherwise.} \end{cases} \quad (97)$$

Let \tilde{s}^* be the vector s^* from [Theorem 11.2](#) called with $\alpha = 2\beta$. We claim that for any η -near minimum cut S such that $\delta(S)_T$ is odd, we have

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

To check this note by (iv) of [Theorem D.3](#) for every set $S \in \mathcal{N}_{\eta, \leq 1}$ such that $S \neq V \setminus \{u_0, v_0\}$, we have $x(E_g \cap \delta(S)) \geq \frac{3}{4}$, so we have

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \geq \tilde{s}(\delta(S)) = \frac{4\beta}{3}x(E_g \cap \delta(S)) - \frac{4\beta}{5}(1 - 2\eta)x(E_b \cap \delta(S)) \geq 0. \quad (98)$$

For $S = V \setminus \{u_0, v_0\}$, we have $\delta(S)_T = \delta(u_0)_T + \delta(v_0)_T = 2$ with probability 1, so condition ii) is satisfied for these cuts as well. Finally, consider cuts $S \in \mathcal{N}_{\eta, 2}$. By [Theorem 11.2](#), if $\delta(S)_T$ is odd, then $\tilde{s}^*(\delta(S)) \geq \alpha(1 - \eta) = 2\beta(1 - \eta)$. Therefore, in such a case we have:

$$\tilde{s}(\delta(S)) + \tilde{s}^*(\delta(S)) \geq 2\beta(1 - \eta) - \frac{4\beta}{5}(1 - 2\eta)x(\delta(S)) \geq 0 \quad (99)$$

where we use that $x(\delta(S)) \leq 2 + \eta$.

Now, we are ready to define s, s^* . Let \hat{s}, \hat{s}^* be the s, s^* of [Theorem D.3](#) respectively. Define $s = \gamma\tilde{s} + (1 - \gamma)\hat{s}$ and similarly define $s^* = \gamma\tilde{s}^* + (1 - \gamma)\hat{s}^*$ for some γ that we choose later. We prove all three conclusions of [Theorem 12.1](#) for s, s^* . (i) follows by (i) of [Theorem D.3](#) and [Eq. \(97\)](#). (ii) follows by (ii) of [Theorem D.3](#) and [Eqs. \(98\)](#) and [\(99\)](#) above. It remains to verify (iii). For edge $e \in E$, $\mathbb{E}[s_e^*] \leq 125\eta\beta x_e$ by (iii) of [Theorem D.3](#) and the construction of s^* . On the other hand, by (iii) of [Theorem D.3](#) and [Eq. \(97\)](#),

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

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

So it remains to prove the following:

Theorem D.3. *Let x^0 be a feasible solution of LP (2) with support $E_0 = E \cup \{e_0\}$ with x the restriction of x^0 to E . Let μ be the max entropy distribution with marginals x . For $\eta \leq 10^{-12}$, $\beta > 0$, there is a set $E_g \subset E \setminus \delta(\{u_0, v_0\})$ of good edges and two functions $s : E_0 \rightarrow \mathbb{R}$ and $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ (as functions of $T \sim \mu$) such that*

- (i) *For each edge $e \in E_g$, $s_e \geq -x_e\beta$ and for any $e \in E \setminus E_g$, $s_e = 0$.*
- (ii) *For each η -near min cut S , including those for which $\{u_0, v_0\} \in \delta(S)$, if $\delta(S)_T$ is odd, then $s(\delta(S)) + s^*(\delta(S)) \geq 0$.*
- (iii) *We have $\mathbb{E}[s_e] \leq -\epsilon_P\beta x_e$ for all edges $e \in E_g$ and $\mathbb{E}[s_e^*] \leq 125\eta\beta x_e$ for all edges $e \in E$, where ϵ_P is defined in [Theorem 5.33](#).*
- (iv) *For every cut S crossed on at most one side such that $S \neq \overline{\{u_0, v_0\}}$, $x(\delta(S) \cap E_g) \geq 3/4$.*

Proof. We start by explaining how to construct \mathcal{H} . Run the following procedure on $\mathcal{N}_{\eta, \leq 1}$ (of x): For every connected component \mathcal{C} of $\mathcal{N}_{\eta, \leq 1}$, if $|\mathcal{C}| = 1$ then add the unique cut in \mathcal{C} to the hierarchy. Otherwise, \mathcal{C} corresponds to a polygon P of cuts crossed on one side with atoms a_0, \dots, a_{m-1} (for some $m > 3$). By [Lemma C.1](#) all these atoms are outside atoms. Add a_1, \dots, a_{m-1} to \mathcal{H} ⁶⁶ and $\cup_{i=1}^{m-1} a_i$ to \mathcal{H} . Note that since $x(\delta(\{u_0, v_0\})) = 2$, the root of the hierarchy is always $V \setminus \{u_0, v_0\}$.

Now, we name every cut in the hierarchy. For a cut S , if there is a connected component of at least two cuts with union equal to S , then call S a near-cycle cut with A, B, C partition as defined in [Definition D.1](#). If S is a cut with exactly two children X, Y in the hierarchy (i.e. a triangle), then let $A = E(X, \bar{X} \setminus Y)$, $B = E(Y, \bar{Y} \setminus X)$ and $C = \emptyset$. Otherwise, call S a degree cut.

Fact D.4 ([\[KKO21, Fact 4.34\]](#)). *The above procedure produces a valid hierarchy.*

The following observation simply follows from the fact that the new cuts that we introduce in the above hierarchy, i.e., atoms and union of non-root atoms of a polygon, are not crossed and are never part of a non-singleton connected component.

Fact D.5. *The set of non-singleton connected components $\mathcal{C}_1, \mathcal{C}_2, \dots$ that the above procedure produces are in one-to-one correspondence to the set of non-singleton connected components of $\mathcal{N}_{\eta, \leq 1}$.*

Let E_g and s be defined as in [Theorem 5.33](#) for the hierarchy defined above, and let $s_{e_0} = \infty$. Also, let s^* be the sum of the $s^* : E \rightarrow \mathbb{R}_{\geq 0}$ vectors from [Theorem 11.2](#) and [Theorem C.12](#) called with $\alpha = \frac{2+\eta}{1-\epsilon_\eta} \beta$. (i) follows from (ii) of [Theorem 5.33](#). Then, $\mathbb{E}[s_e^*] \leq (18 + 44)\eta \left(\frac{2+\eta}{1-\epsilon_\eta} \beta\right) \leq 125\eta\beta$ follows from [Theorem 11.2](#) and [Theorem C.12](#) and using that $\eta \leq 10^{-12}$ and $\epsilon_\eta = 7\eta$. Also, $\mathbb{E}[s_e] \leq -\epsilon_p \beta x_e$ for edges $e \in E_g$ follows from (v) of [Theorem 5.33](#).

Now, we verify (iv): For any (non-root) cut $S \in \mathcal{H}$ such that $p(S)$ is not a near-cycle cut $x(\delta(S) \cap E_g) \geq 3/4$ by (i) of [Theorem 5.33](#). The only remaining case is η -near minimum cuts which are either atoms or near minimum cuts in a polygon. Fix such a set S in a polygon P . Let S' be the union of the non-root atoms of P . Then by [Lemma 2.37](#), $x(\delta(S) \cap \delta(S')) \leq 1 + \epsilon_\eta$. All edges in $\delta(S) \setminus \delta(S')$ are bottom edges, so by (i) of [Theorem 5.33](#) are in E_g . Therefore, $x(\delta(S) \cap E_g) \geq 1 - \epsilon_\eta \geq 3/4$.

It remains to verify (ii): We consider 5 groups of cuts:

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

Type 2: Cuts $S \in \mathcal{N}_{\eta, 2}$. By [Theorem 11.2](#) and the fact that $\alpha = \frac{2+\eta}{1-\epsilon_\eta} \beta$, if $\delta(S)_T$ is odd then

$$s^*(\delta(S)) \geq \frac{2+\eta}{1-\epsilon_\eta} \beta (1-\eta) \geq (2+\eta)\beta \geq -s(\delta(S))$$

where we use that $s_e \geq -\beta x_e$ for all edges e and $x(\delta(S)) \leq 2 + \eta$.

Type 3: Cuts $S \in \mathcal{H} \cap \mathcal{N}_\eta$ where $p(S)$ is not a near-cycle cut. By (iv) of [Theorem 5.33](#) and that $s^* \geq 0$ the inequality follows.

Type 4: Cuts S such that either $S \in \mathcal{N}_{\eta, \leq 1} \setminus \mathcal{H}$ or $S \in \mathcal{H} \cap \mathcal{N}_\eta$ and $p(S)$ is a (non-triangle) near-cycle cut. In this case either S is an atom or an η -near minimum cut of a non-singleton connected component \mathcal{C} of $\mathcal{N}_{\eta, \leq 1}$ with corresponding polygon P of cuts crossed on one side

⁶⁶Notice that an atom may already correspond to a connected component, in such a case we do not need to add it in this step.

and the cycle partition A, B, C . If S is not a leftmost cut/atom or a rightmost cut/atom, then by [Theorem C.12](#), whenever $\delta(S)_T$ is odd, we have (similar to Type 2):

$$s^*(\delta(S)) \geq \frac{2+\eta}{1-\epsilon_\eta} \beta(1-\epsilon_\eta) = (2+\eta)\beta \geq -s(\delta(S)) \quad (100)$$

Otherwise, suppose S is a leftmost cut. If P is left-happy then by [Theorem C.12](#), similar to above, $s^*(\delta(S)) + s(\delta(S)) \geq 0$ if $\delta(S)_T$ is odd. Otherwise, let S' be the union of the non-root atoms of P and $F = \delta(S) \setminus \delta(S')$. By [Lemma 2.37](#), we have $x(F) \geq 1 - \epsilon_\eta/2$. Therefore, by (iii) of [Theorem 5.33](#) we have

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

as desired. Note that since S is a leftmost cut, we always have $A \subseteq \delta(S)$. But C may have an unpredictable intersection with $\delta(S)$; in particular, in the worst case only edges of C with negative slack belong to $\delta(S)$. This is why we need to use $s^-(C)$ instead of $s(C)$. A similar argument holds when S is the leftmost atom or a rightmost cut/atom.

Type 5: Cuts $S \in \mathcal{H} \cap \mathcal{N}_\eta$ where $p(S)$ is a triangle P . This is similar to the previous case except we use [Lemma C.13](#) to argue that the inequality is satisfied when P is left/right happy. \square