# Lecture 20

# $\left(\frac{5}{3} - \epsilon\right)$-approximation Algorithm for TSPP - Part 2*

This is a continuation of the previous lecture describing the An-Kleinberg-Shmoys algorithm for the Traveling Salesman Path Problem [AKS11].

## 20.1 Review

We begin by reviewing notes and definitions from the previous lecture.

First, we give an LP relaxation for TSPP:

$$
\begin{aligned}
& \min c(x) := \sum_{e \in E} c_e x_e \\
\text{s.t. } & x(\partial S) \geq 1 \quad \text{for separating cuts, i.e. } |S \cap \{s,t\}| = 1 \text{, with } |S| > 1 \\
& x(\partial S) \geq 2 \quad \text{for non-separating cuts with } |S| > 1 \\
& x(\partial S) = 2 \quad \text{for cuts with } |S| = 1 \\
& x_e \geq 0 \quad \forall e \in E
\end{aligned}
$$

This polytope is a subset of the spanning tree polytope. As a result, optimal $x^* = \sum_{i \leq \binom{n}{2}} \lambda_i \mathbf{1}_{A_i}$, where $A_i$ are the actual spanning trees with $s, t$ as leaves. Based on this, the [AKS11] algorithm goes as follows:

1. Solve the TSPP LP relaxation to get a solution $x^*$.
2. Write $x^*$ as a convex combination of spanning trees $A_i$ that have $s$ and $t$ as leaves to get $x^* = \sum_{i \leq \binom{n}{2}} \lambda_i \mathbf{1}_{A_i}$.
3. Pick a spanning tree $A$ at random from this distribution (choose $A_i$ with probability $\lambda_i$).

*Lecturer: Ryan O'Donnell. Scribe: Alex Beutel.

4. Let $T = T_A$ be the set of vertices in $A$ whose degree parity needs to be fixed. As described above, $|T|$ is even.

5. Take $M$ to be the minimum cost matching on $T$.

6. Then $A \cup M$ has an Eulerian path from $s$ to $t$. Shortcut to avoid taking the same edge twice and return the resulting path.

We will make heavy use of two other properties in this lecture:

1. *Fractional T-join dominator*: $y$ such that $y_e \geq 0$ and $y(\partial S) \geq 1$ if $S$ is an odd-T-cut ($|S \cap T|$ is odd).

2. Property $\circledast$: If $S$ is a separating odd-T-cut, then $\mathbf{1}_A(\partial S) \geq 2$.

## 20.2 A $1.658\bar{3}$-approximation

**Goal** We would like to find a mapping of $(x^*, A) \to y$, some fractional $T_A$-join dominator such that the expected cost

$$\mathop{\mathbf{E}}_A[c(y)] \leq 0.658\bar{3}c(x^*).$$

We note that $\mathbf{E}_A[c(M)] \leq \mathbf{E}_A[c(y)]$ and $\mathbf{E}[c(A)] = c(x^*)$. Therefore, $\mathbf{E}[\text{cost}] \leq 1.658\bar{3}c(x^*)$.

**Idea 1** We first try taking $y = \frac{1}{3}x^* + \frac{1}{3}\mathbf{1}_A$. We see that the expected cost

$$E[c(y)] = \frac{1}{3}c(x^*) + \frac{1}{3}\,\mathbf{E}[c(\mathbf{1}_A)] = \frac{2}{3}c(x^*).$$

This therefore gives a $\frac{5}{3}$-approximation. However, we must prove that this is valid (that $y$ is a fractional $T$-join dominator):

C1. $S$ is non-separating: We know that $x^*(\partial S) \geq 2$ and $\mathbf{1}_A(\partial S) \geq 1$. Therefore, $y(\partial S) = \frac{1}{3}x^*(\partial S) + \frac{1}{3}\mathbf{1}_A(\partial S) \geq \frac{2}{3} + \frac{1}{3} = 1$.

C2. $S$ is separating: We know that $x^*(\partial S) \geq 1$ and $\mathbf{1}_A(\partial S) \geq 2$, based on $\circledast$. Therefore, $y(\partial S) \geq \frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 2 = 1$.

We see that we get the $\frac{5}{3}$-approximation. It is going to be hard to say anything smarter about non-separating cuts, but we can improve upon the analysis of the separating ones.

**Idea 2** We now consider $y = 0.35 \cdot x^* + 0.3 \cdot \mathbf{1}_A + \ldots$, where we expect we will have to add another term at the end. For a non-separating $S$ we see that $y(\partial S) \geq .35 \cdot 2 + .3 \cdot 1 + \ldots \geq 1$. However, for a separating $S$ we find that we only know that $y(\partial S) \geq .35 \cdot 1 + .3 \cdot 2 + \ldots = .95 + \ldots$. Thus, we describe $y = .35 \cdot x^* + .3 \cdot \mathbf{1}_A + .05 \cdot f$ with the goal that $\mathbf{E}_A[c(f)] \leq \frac{1}{6} c(x^*)$ so that $\mathbf{E}[c(y)] \leq .65c(x^*) + .05 \cdot \frac{1}{6} c(x^*) = .658\bar{3}c(x^*)$. Therefore in the case of a separating $S$, we find that $y(\partial S) \geq .95 + .05 \cdot f(\partial S)$.

We now focus on different $f$ terms that will provide an acceptable answer when $S$ is a separating cut (since all positive $f$ will satisfy the case for a non-separating $S$).

We know that if $x^*(\partial S)$ is sufficiently bigger than 1 we are covered. For example, if $x^*(\partial S) \geq \frac{8}{7}$ then $y(\partial S) \geq .35 \cdot \frac{8}{7} + .6 = 1$ and the $f$ term is unnecessary.

**Definition 20.1.** A separating cut is $\tau$-*narrow* if $x^*(\partial S) < 1 + \tau$. While, we define this in the general sense, $\tau$ previously (and in all future cases) is in fact equal to $\frac{1}{7}$.

Interestingly, we are fine if $x^*(\partial S) = 1$. For every choice of $A$, we know that $S$ will not be an odd-$T_A$-cut. Thus, in checking the fractional $T$-join dominator conditions for $y$, we do not need to worry about $S$. $x^*(\partial S) = \mathbf{E}_A[\mathbf{1}_A(\partial S)]$ and $\mathbf{1}_A(\partial S) \geq 1$ always. Therefore we know that $\mathbf{E}_A[\mathbf{1}_A(\partial S) = 1$ and thus $\mathbf{1}_A(\partial S) = 1$ always. Therefore, $S$ is not an odd-$T_A$-cut by $\circledast$. Even if $x^*(\partial S) = 1.01$, we know that $\mathbf{Pr}[\mathbf{1}_A(\partial S) \geq 2]$ must be very small.

**Proposition 20.2.** *If $S$ is $\tau$-narrow, then $\mathbf{Pr}_A[S$ is an odd-$T_A$-cut$] < \tau$*

*Proof.* We know that $\mathbf{E}[\mathbf{1}_A(\partial S)] = x^*(\partial S) < 1 + \tau$. Therefore, $\mathbf{Pr}[\mathbf{1}_A(\partial S) \geq 2] < \tau$ and $\mathbf{Pr}[S$ is an odd-$T_A$-cut$] \leq \mathbf{Pr}[\mathbf{1}_A(\partial S) \geq 2]$ by $\circledast$. $\square$

**Theorem 20.3.** *Given $x^*$, the $\tau$-narrow separating cuts $S_1, \ldots S_k$ have $\partial S_i$ "almost" disjoint.*

For proving this we will begin by pretending that they are in fact truly disjoint.

**Definition 20.4.** $f = \sum_{i \leq k} \mathbf{1}[S_i \text{ is an odd-}T_A\text{-cut}] \cdot x^*\big|_{\partial S_i}$, where $x^*\big|_{\partial S_i}$ puts 0 on edges not on the boundary.

If $S$ is a separating odd-$T_A$-cut, then $y(\partial S) \geq .35 \cdot x^*(\partial S) + .3 \cdot 2 + .05 \cdot x^*(\partial S) \geq 1$. Therefore, this definition of $f$ produces a fractional $T$-join dominator.

As stated earlier, our goal is for $\mathbf{E}_A[c(f)] \leq \frac{1}{6} c(x^*)$. We see here that $\mathbf{E}_A[c(f)] = \sum_i \mathbf{Pr}_A[S_i \text{ is an odd-}T_A\text{-cut}] \cdot c(x^*\big|_{\partial S_i}) \leq \tau \sum_i c(x^*\big|_{\partial S_i}) = \tau c(x^*)$. There is no double-counting in the summation because of the assumed "disjointness" of the $\partial S_i$'s. Therefore, this gives us $\mathbf{E}_A[c(f)] \leq \frac{1}{7} c(x^*)$. Because the $\partial S_i$'s are only *almost* disjoint, this is relaxed from $\frac{1}{7}$ to $\frac{1}{6}$.

**Claim 20.5.** *If $S$ and $S'$ are separating $\tau$-narrow cuts, then they don't cross.*

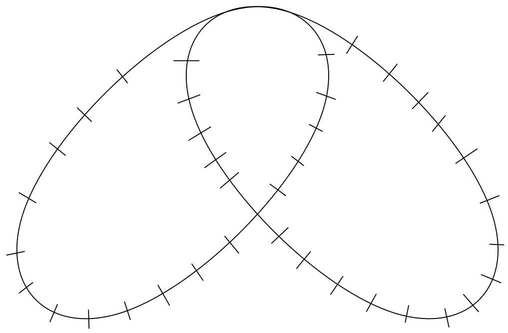


Figure 20.1: Two crossing separating cuts, $S$ and $S'$, with edges out of $S \Delta S'$.

*Proof.* Suppose $S$ and $S'$ cross. $S \Delta S'$ has two non-empty pieces, where each piece is a non-separating cut (since $s$ is of course in the intersection of the two). Therefore, by the LP definition, there must be at least 2 edge weights out of each piece. Therefore, $x^*(\partial S) + x^*(\partial S') \geq 4$. However, both $x^*(\partial S) \leq 1 + \tau$ and $x^*(\partial S') \leq 1 + \tau$. This is a contradiction. Definitely include the picture, but this proof doesn't make sense to me. □

Because the separating $\tau$-narrow cuts do not cross, we can consider the cuts as being sequential subsets either around $s$ or $t$, as shown in Figure 20.2. We can label vertices based on the largest cut they fall within. Therefore, the set $C_i$ is the set of vertices as shown in the picture above, the set $L_i$ is the set $C_{i-1}$ and the set $R_i = C_{i+1}$.

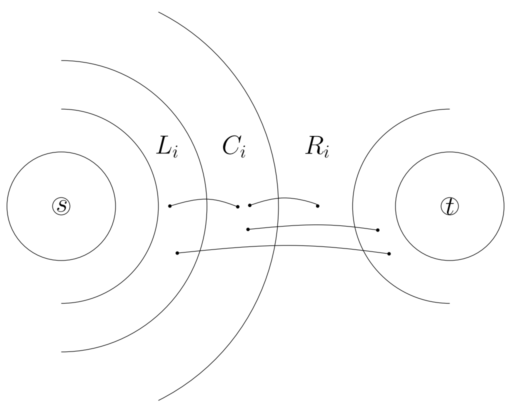



Figure 20.2: Concurrent $\tau$-narrow cuts with edges between different regions, each labeled by $C_i$ as described in the text.

**Idea** Instead of putting $x^*|_{\partial S_i}$ into $f$, just put $x^*|_{E(C_i,R_i)}$ because we know that $E(C_i, R_i)$ are disjoint for all $i$.

**Lemma 20.6.** *For all $i$, we know that $x^*(E(C_i, R_i)) \geq 1 - \tau = \frac{6}{7}$. Therefore, we can put $\frac{7}{6} \cdot x^*|_{E(C_i,R_i)}$ into $f$.*

This makes $y$ valid and the expected cost becomes $\frac{7}{6} \cdot \frac{1}{7} \cdot c(x^*) = \frac{1}{6}c(x^*)$.

*Proof.* Because the cuts are $\tau$-narrow we know that $x^*(L, C) + x^*(L, R) = x^*(L, C \cup R) \leq 1 + \tau$, and from the LP since $S = C$, we know that $x^*(L, C) + x^*(C, R) = x^*(C, L \cup R) \geq 2$. If we subtract the first statement from the second we see that $x^*(C, R) - x^*(L, R) \geq 1 - \tau$ and thus $x^*(C, R) \geq 1 - \tau$. □

## 20.3 The Hirsch Conjecture

We now switch topics entirely to other new, interesting work. We begin with some basic definitions.

**Definition 20.7.** $P$ is a convex polytope in $\mathbb{R}^d$ with $n$ facets (and thus $(d-1)$-dimensional faces).

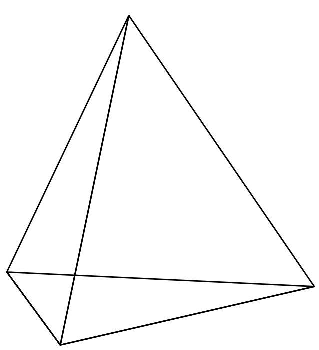
**Definition 20.8.** $G(P)$ is the graph of the vertices of $P$.

**Definition 20.9.** $\delta(P)$ is the diameter of $G(P)$ (the maximum distance between any two points).

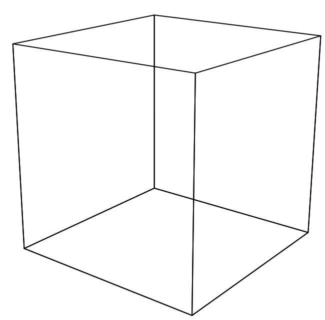
**Definition 20.10.** $\Delta(d, n)$ is the largest possible $\delta(P)$ for a $P$ in $\mathbb{R}^d$ with $n$ facets. This bounds the running time of the simplex algorithm using "clairvoyant" pivoting.

**1957 Hirsch Conjecture** $\Delta(d, n) \leq n - d$

The original conjecture allowed an unbounded $P$ and was disproved in 1967 by Klee and Walkup. We will focus on the case of a bounded $P$. This has been verified in many cases. We look for example at the simplex case of a convex hull of $d+1$ points in general positions. For example, for a basic tetrahedron we find that $n = d + 1$ and thus $n - d = 1$. Looking at Figure 20.3(a) we see that $\delta(P) = 1$. Similarly for a cube we find that $n = 2d$, $n - d = d$ and $\delta(P) = d$.



(a)



(b)

Figure 20.3: Basic tetrahedron (a) and cube (b) as discussed in the text.

Much progress has been made since:

[KW67] stated that the Hirsch conjecture was equivalent to the statement that $\forall u, v$ it is possible to go from $u$ to $v$ such that each step enters a new facet.

[Larman 1970]: $\Delta(d,n) \le \frac{1}{8}2^d \cdot n$

[Barnette 1974]: $\Delta(d,n) \le \frac{1}{12}2^d \cdot n$

[Kalai-Kleitman 1992]: $\Delta(d,n) \le n^{\log d+1}$

Finally, in 2010 [Santos] proved the conjecture false showing that $\Delta(43,86) \ge 44$. [MSW11] found a completely explicit case for which $d = 20$, $n = 40$, 36442 vertices, and $\delta(P) = 21$.

**Hähnle Conjecture** $\Delta(d,n) \le d(n-d)+1$.

Additionally, the poly-Hirsch conjecture is still open, stating $\Delta(d,n) \le n^{O(1)}$.

## 20.4 A Rough Explanation of Santos's Theorem (2010)

We will roughly outline the method used by Santos to disprove the Hirsch Conjecture.

**Definition 20.11.** Polytope $P$ with $n \ge 2d$ is a *spindle* if it has two vertices $u, v$ which don't share a facet but every facet touches $u$ or $v$.

A spindle can be thought of as the intersection of two "cones" with apices $u, v$ such that the $\bar{uv}$ is inside the intersection.

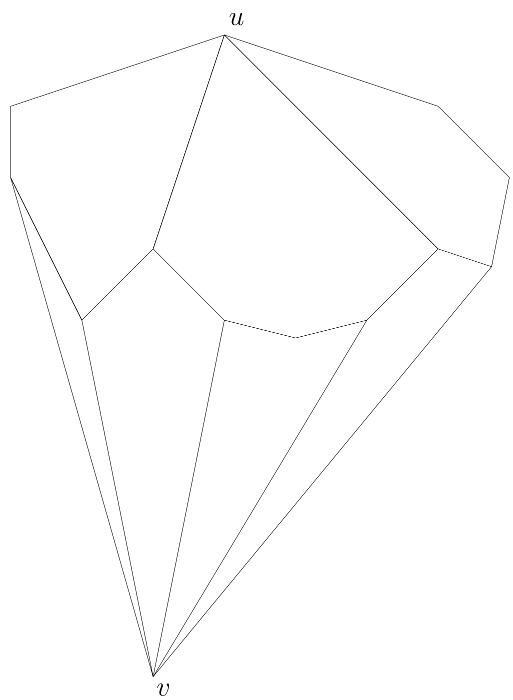



Figure 20.4: A spindle as described previously.

**Definition 20.12.** Length $= \operatorname{dist}(u,v)$.

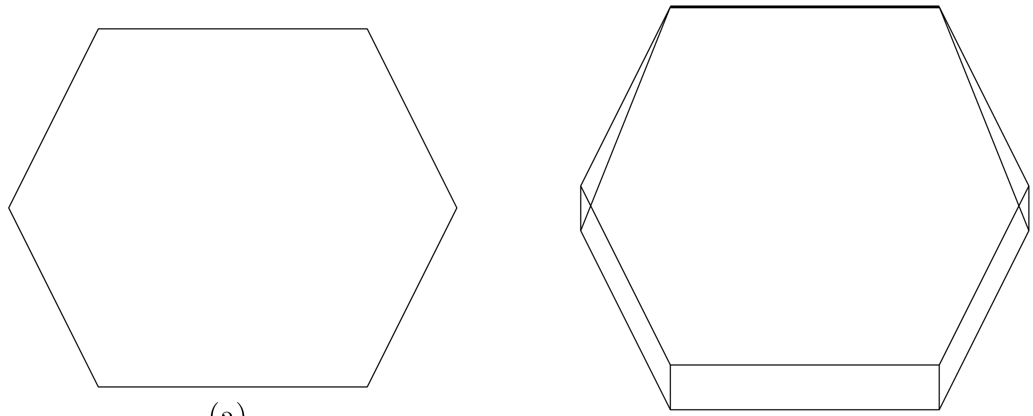



Figure 20.5: An example of the pivot transformation from hexagon (a) to 3-dimensional polytope (b).

**Theorem 20.13.** *If there exists a spindle $P$ with $n > 2d$ and length $l$, then there exists a spindle $P'$ in dimension $d+1$ with $n+1$ facets and length $\geq l+1$.*

The process of creating $P'$ from $P$ is known as the "Santos Wedge Operation." We can of course see the effect of doing this wedge operation repeatedly: $(d,n,l) \rightarrow (d+1,n+1,l+1) \rightarrow (d+2,n+2,l+2) \rightarrow \dots$. If we do this $n-2d$ times then we get $(n-d, 2n-2d, l+n-2d)$. If $l > d$ then diam $\geq l+n-2d > n-d$. Therefore, to disprove the Hirsch conjecture, we can construct a $d$-dimensional spindle of length $> d$.

*Proof.* The [KW] wedge picks one pair of connected vertices to keep constant and doubles all vertices. Such that for vertex $u$ and new vertex $u'$, there is an edge between $u$ and $u'$ as well as added edges between $u'$ and the copied versions of the neighbors of $u$. This transformation is shown in Figure 20.5.

If $n > 2d$ then $u$ or $v$ is degenerate (we will assume $u$). Take any facet that contains $u$ and wedge on it. This is almost the new spindle. We must add a small amount of perturbation. □

**Theorem 20.14.** *There exists a spindle in 5 dimensions with length 6.*

Santos gives two proofs. First he explicitly gives an example that can be verified on the computer. Second, there is a nice conceptual analysis of combinatorial maps on the 3-sphere.

It can be shown that there does not exist a 3-dimensional spindle with length 4. Additionally, [Thomas] showed that there does not exist a 4-dimensional spindle with length 5. As such, a 5-dimensional spindle is the smallest one for which we can have length $>$ dimensions.

This results in a disproof of the Hirsch Conjecture but an explicit polytope is not given. The example for which $\Delta(20,4) \geq 21$ is explicit.

# Bibliography

[AKS11] Hyung-Chan An, Robert Kleinberg, and David B. Shmoys. Improving christofides' algorithm for the s-t path tsp. *CoRR*, abs/1110.4604, 2011. 20, 20.1