Recent Advances in Approximation Algorithms **Spring 2015**

Lecture 12: Randomized Rounding Algorithms for Symmetric TSP

Lecturer: Shayan Oveis Gharan *May 6th*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications.*

We defined the symmetric TSP in lecture 6. The best known approximation algorithm for STSP is the 3/2 approximation algorithm of Christofides. The algorithm is very easy to describe, first choose a minimum cost spanning tree and then add that minimum cost Eulerian augmentation to the tree, that is the minimum cost set of edges to make the tree Eulerian. An undirected graph is Eulerian when every vertex has even degree. Therefore, the minimum cost Eulerian augmentation of a given subgraph is the minimum cost matching on the odd degree vertices of that subgraph.

In this lecture we use the strong negative dependence property of random spanning trees to design a $3/2-\epsilon_0$ approximation algorithm for STSP on graph metrics where $\epsilon_0 > 0$ is a constant that is independent of the size of the problem. A metric $c(.,.)$ is a graph metric if there is an unweighted graph where $c(.,.)$ is the shortest path metric of that graph. The material of this lecture are based on the work of Oveis Gharan, Saberi and Singh [OSS11]. The basic idea of the algorithm of [OSS11] is quite similar to Christofides, the main difference is to choose a random spanning tree (as opposed to a minimum spanning tree) based on the solution of the LP relaxation of STSP. Then we add the minimum cost matching on the odd degree vertices of the tree. The main crux of the analysis is to use the properties of random spanning tree distributions to bound the expected cost of the matching.

After [OSS11] several groups managed to design improved approximation algorithms for STSP on graph metrics [MS11; Muc12; SV12] with approximation factors $1.46, 1.44, 1.4$ respectively. All of these improved results mainly use graph theoretic techniques. We will not talk about the details of these three results in this course. It remains an open problem to improve the Christofides' approximation algorithm for general metrics.

12.1 Randomized Rounding Algorithm

The basic idea of the algorithm is very similar to the algorithm that we described in Lecture 6-8 for ATSP. Let x be an optimal solution of the Held-Karp relaxation for STSP,

$$\begin{aligned} \text{minimize} \quad & \sum_{u,v} c(u,v)x(\{u,v\}) \\ \text{subject to} \quad & \sum_{u\in S, v\in \overline{S}} x(\{u,v\}) \geq 2 \quad \forall\, S \subsetneq V \\ & \sum_{v\in V} x(\{u,v\}) = 2 \qquad \forall\, u \in V \\ & x \geq 0. \end{aligned} \tag{12.1}$$

$z = (1-1/n)x$ and $G = (V, E, z)$ be the underlying graph. The reader should not confuse z with the variables that we used in the last lecture to write a real stable polynomial. Here, we used the letter z for a fractional spanning tree to be consistent with the notation of Lectures 6-8. We write z as a maximum entropy distribution of spanning trees of G, so there is a weighted uniform distribution μ of spanning trees with marginals z. Let T be a sample of μ and O be the odd degree vertices of T.

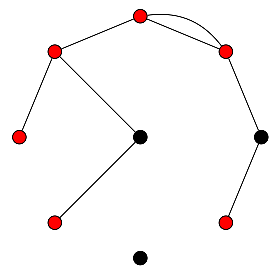

Figure 12.1: An O-join of the red vertices. Observe that an O-join is a union of paths connecting the red vertices

As alluded to in the introduction, we want to add the minimum cost Eulerian augmentation of T in G. A set $F \subseteq E$ is an *O-join* of T if every vertex $v \in O$ has an odd degree in F and every other vertex has an even degree. See Figure 12.1 for an example of an O-join. Observe that an O-join is a union of paths where each path is connecting two distinct vertices of O. It follows that the minimum cost Eulerian augmentation of T is the minimum cost O-join, F of T. Note that $c(F)$ is the same as the cost of the minimum matching on odd degree vertices of T in the metric completion of graph G. In other words, by the triangle inequality, instead of choosing a path between odd degree vertices u, v of T we may simply choose the edge $\{u, v\}$, and by triangle inequality the cost of this edge is at most the cost of the path in the O-join.

Edmonds and Johnson [EJ73] showed that for any graph $G = (V, E)$ and cost function $c : E \to \mathbb{R}_+$, and $O \subseteq V$ with even number of vertices, the minimum cost of an O-join equals the optimum value of the following linear program, so the optimum is always integral.

$$\begin{array}{lll} \text{minimize} & c(y) & \\ \text{subject to} & y(\delta(S)) \geq 1 & \forall S \subseteq V, |S \cap O| \text{ odd} \\ & y \geq 0 & \forall e \in E \end{array} \tag{12.2}$$

Now we are ready to describe a randomized rounding algorithm for STSP.

Algorithm 12.1 A Randomized Rounding Algorithm for Symmetric TSP.

Let x be an optimal solution of (12.1), $z = (1 - 1/n)x$ and $G = (V, E, z)$ be the support graph of z.
Calculate the maximum entropy distribution μ with marginals z and choose a sample T.
Let O be the odd degree vertices of T. Use (12.2) to find the minimum cost O-join of T, F.
Return $T \cup F$.

It was shown in [OSS11] that a variant of the above algorithm has an approximation factor of $3/2 - \epsilon_0$ for STSP on graph metrics. In this course we will not talk about [OSS11] in its full generality, instead our focus is on part of the proof that utilizes strongly Rayleigh measures and random spanning tree distribution. More specifically, we prove the following theorem.

Theorem 12.1. *Suppose for any set $S \subseteq V$ of size at least 2 ($|S| \geq 2$), $x(\delta(S)) \geq 2 + \epsilon$, then for any cost function $c : E \to \mathbb{R}_+$ satisfying the triangle inequality, the approximation factor of Algorithm 12.1 is at most $3/2 - \Omega(\epsilon)$.*

12.2 Good Edges, Random Spanning Trees and T-join Polytope

Since $T \sim \mu$ and μ preserves z as the marginals,

$$\mathbb{E}[c(T)] = \sum_{e \in E} c(e) z(e) = c(z) = (1 - 1/n) c(x).$$

So, to prove Theorem 12.1 it is enough to show that $\mathbb{E}[c(F)] \leq (1/2 - \Omega(\epsilon)) c(x)$.

Theorem 12.2. *Suppose for any set $S \subseteq V$ of size at least 2 ($|S| \geq 2$), $z(\delta(S)) \geq 2 + \epsilon$. Then for any cost function $c : E \to \mathbb{R}_+$,*

$$\mathbb{E}[c(F)] \leq (1/2 - \Omega(\epsilon)) c(x).$$

First, let us see that for any spanning tree T, $c(F) \leq c(x)/2$. Similar to the proof idea of rounding by sampling algorithm for ATSP all we need to do is to construct a feasible (fractional) solution of the O-join polytope of cost at most $c(x)/2$. To do that, it is enough to let $y(e) = x(e)/2$ for all $e \in E$. Since $x(\delta(S)) \geq 2$ for all $S \subseteq V$, we get $y(\delta(S)) \geq 1$ for all S. Therefore,

$$c(F) \leq c(y) = c(x)/2.$$

Note that in the above construction $y(\delta(S)) \geq 1$ for all $S \subset V$. But, in order to have a feasible solution of (12.2) we only need that $y(\delta(S)) \geq 1$ whenever $|S \cap O|$ is odd. First, observe that $|S \cap O|$ is odd if and only if $|T(S, \overline{S})|$ is odd. The idea is to use the randomness of T to assign a slightly smaller value $y(e) = x(e)(1/2 - \epsilon')$ to some of the edges (with a constant probability) while preserving the feasibility of y.

Before, getting into the details of the construction of y, let us give a simple example showing that to obtain a bound $\mathbb{E}[c(F)] \leq (1/2 - O(\epsilon)) c(x)$ we have to use that T is chosen from a maximum entropy distribution.

Example 12.3. *Let G be a complete graph, $c(e) = 1$ for all e and $x(e) = \frac{1}{n-1}$ for all $e \in E$. Now, suppose T is an arbitrary minimum spanning tree, in particular, let T be a star. Then, all vertices of T (except possibly one) have an odd degree. Therefore, $\mathbb{E}[c(F)] \approx n/2$. Even if μ preserves the marginals of z (but is not a maximum entropy distribution), it may be a convex combination of n stars each rooted at a vertex of G.*

On the other hand, if μ is a maximum entropy distribution, then using strong concentration bounds we can show that any vertex v has degree 2 in $T \sim \mu$ with a constant probability,

$$\mathbb{P}[|T(\delta(v))| = 2] \geq \Omega(1),$$

see Theorem 12.7 for the details. It follows that in expectation a constant fraction, γ, of vertices of T have even degree. But then, the expected cost of F is at most

$$\mathbb{E}[c(F)] = \frac{1}{2}\mathbb{E}[|\{v : |T(\delta(v))| \text{ is odd}\}|] \leq \frac{(1-\gamma)n}{2} = (1-\gamma)c(x)/2.$$

We say an edge e is *even* if both of its endpoints have degree 2 in T[1]. In this case we say $\mathcal{E}_e$ occurs. We say e is *good* if it is even with probability at least δ,

$$\mathbb{P}[\mathcal{E}_e] \geq \gamma,$$

for some constant δ that we fix later in the proof.

Lemma 12.4. *Suppose that every edge $e \in E$ is good. Then,*

$$\mathbb{E}[c(F)] \leq (1/2 - \Omega(\epsilon \cdot \gamma)) c(x).$$

[1]Tor the argument in this section we may also define even edges to be those where both endpoints have even degree.

Proof. All we need to do is to construct a vector y that is a feasible solution of (12.2) such that $\mathbb{E}[c(y)] \leq (1/2 - \Omega(\gamma \cdot \epsilon))c(x)$. We construct y as follows, for any edge e,

$$y(e) = \begin{cases} \frac{x(e)}{2+\epsilon} & \text{if } \mathcal{E} \text{ occurs,} \\ \frac{x(e)}{2} & \text{otherwise.} \end{cases}$$

Now, for any set S, if $|S| \geq 2$, then

$$y(\delta(S)) \geq \frac{x(\delta(S))}{2+\epsilon} \geq 1,$$

by the assumption of Theorem 12.1. On the other hand, if $|S| = \{v\}$, then if v has an even degree, there is nothing to prove. If v has an odd degree, then $y(\delta(v)) = x(\delta(v)) = 1$.

It remains to bound the cost of y. Since every edge is good,

$$\begin{aligned} \mathbb{E}[c(y)] &\leq \sum_e \frac{x(e)}{2}(1 - \mathbb{P}[\mathcal{E}_e]\,\epsilon/4) \\ &\leq c(x)(1/2 - \Omega(\epsilon \cdot \gamma)). \end{aligned}$$

□

Note that in the proof of the above lemma we crucially used that all nonsingleton cuts of x have size at least $2 + \epsilon$. In the general case, to show that Algorithm 12.1 improves the Christofides' algorithm for graph metrics one needs to exploit the structure of the near minimum cuts of x, i.e., those of size at most $2 + \epsilon$. This structure is very well studied in the last decade by Benczúr and Goemans [Ben95; BG08], also see [OSS11] for the extensions needed to analyze Algorithm 12.1.

Now all we need to show is that every edge of G is good. Unfortunately, this is not necessarily true if $x(e) \approx 1/2$, but it is true in all other cases.

Lemma 12.5. *There is an absolute constant $\gamma > 0$ (independent of n) such that for any edge $e \in E$, if $x(e) < .49$ or $x(e) > .51$ then*

$$\mathbb{P}[\mathcal{E}_e] \geq \gamma.$$

We will not talk about the special case where $x(e) \approx 1/2$ in this lecture. It is easy to see that the above two lemmas (together with the spacial case where $x(e) \approx 1/2$) imply the main theorem 12.1.

In the rest of this lecture we use the properties of strongly Rayleigh distributions to prove a weaker version of the above lemma.

Lemma 12.6. *For any pair of vertices $u, v \in V$, if there is no edge between u, v in G, then both of them have degree exactly 2 in $T \sim \mu$ with probability at least $1/100$.*

The proof of the above lemma contains almost all of the ideas necessary to prove Lemma 12.5 but it needs a significantly lighter calculation.

Upshot. In Lectures 6-8 we used the negative correlation property to prove that a random spanning tree with marginals z is $O(\log n/\log\log n)$-thin with high probability. In that context, our main use of randomness was to prove concentration of measures, that is all of the cuts have small number of edges in T relative to their size in z. Here, on the other hand, we are concerned with the parity of the number edges of the sampled tree in the (degree) cuts. Trivially, we can not expect a tree to have an even number of edges in every cut, but using randomness we can expect to have even number of edges in at least 1/3 of the cuts. Although this can be proven, it is not enough to prove the above lemma as it says nothing about the correlation of parity in different cuts. To prove the above lemma, we would like to say the degree of u is independent of the degree of v in a random spanning tree. This is not necessarily true in general, but as we will see in the next section we can use the properties of strongly Rayleigh measures to prove the above lemma.

12.3 Existence of Good Edges

From now on we fix two vertices u, v and we prove Lemma 12.6 for these vertices. Let $X = |T(\delta(u))|$ be the degree of u and $Y = |T(\delta(v))|$ be the degree of v. For the simplicity we assume that $\mathbb{E}[X] = \mathbb{E}[Y] = 2$ (note that the right value is $\mathbb{E}[X] = \mathbb{E}[Y] = 2(1 - 1/n)$).

First, we show that $\mathbb{P}[X = 2] \geq 1/4$. To see this we use the fact that the degree distribution of any vertex is the same as the sum of independent Bernoulli random variables (see Lecture 11 for the proof). So say $B_1, \ldots, B_m$ are independent Bernoulli's such that the law of $B_1 + \cdots + B_m$ is the same as the law of X. The question is what is the minimum possible value of $\mathbb{P}[B_1 + \cdots + B_m = 2]$? This question is answered by Hoeffding several decades ago.

Theorem 12.7 (Hoeffding [Hoe56, Corollary 2.1])**.** *For any set of independent Bernoulli's $B_1, \ldots, B_n$ with priors $p_1, \ldots, p_m$ and any function $g : \mathbb{R} \to \mathbb{R}$, $\mathbb{E}[g(B_1 + \cdots + B_m)]$ is minimized (or maximized) when each p_i takes one of the values $0/1/p$ for a fixed p, i.e., $\{p_1, \ldots, p_m\} \subseteq \{0, p, 1\}$.*

Since $\mathbb{E}[B_1 + \cdots + B_m] = \mathbb{E}[X] = 2$, by the above theorem, in the worst case each B_i occurs with probability $2/m$, so

$$\mathbb{P}[X = 2] \geq 2(1 - 2/m)^m \geq 1/4.$$

Now, we get to the interesting part of the argument, we would like to say $\mathbb{P}[X = 2, Y = 2] \geq \Omega(1)$. Unfortunately, we cannot use union bound because $\mathbb{P}[X = 2] < 1/2$. The basic idea is to use the Bayes rule,

$$\mathbb{P}[X = 2, Y = 2] = \mathbb{P}[Y = 2|X = 2]\,\mathbb{P}[X = 2] \geq \mathbb{P}[Y = 2|X = 2]/4.$$

The main question is how to lower bound $\mathbb{P}[Y = 2|X = 2]$. First, we need to understand the measure $\{.|X = 2\}$. It turns out that by the truncation and projection closure properties this is also a strongly Rayleigh measure[2]. So the law of Y is the same as a sum of independent Bernoulli's. Therefore, if $1.01 < \mathbb{E}[Y|X = 2] < 2.99$ we can again use Theorem 12.7 and we are done. But, if $\mathbb{E}[Y|X = 2] \geq 3$, then it could be that $Y = 3$ with probability 1 and we are doomed. The crux of the proof is to see how does the expectation change under conditioning. We will use negative association and stochastic dominance properties to control the changes of expectation under conditioning.

We do a slightly different conditioning.

$$\mathbb{P}[X = 2, Y = 2] = \mathbb{P}[X = 2|X + Y = 4]\,\mathbb{P}[X + Y = 4] \geq \mathbb{P}[X = 2|X + Y = 4]/5,$$

where we can again invoke Theorem 12.7 to show that $\mathbb{P}[X + Y = 4] \geq 1/5$. Now, let $\mu' = \mu|_{\delta(u)\cup\delta(v)}$ be the projection of μ onto $\delta(u) \cup \delta(v)$, and μ'_4 be μ' truncated to samples of size 4. Since μ is SR, so is μ', μ'_4.

Since μ'_4 is SR, X is log concave, so

$$\mathbb{P}_{\mu'_4}[X = 2] \geq \sqrt{\mathbb{P}_{\mu'_4}[X = 1] \cdot \mathbb{P}_{\mu'_4}[X = 3]}. \tag{12.3}$$

[2] Recall that in the truncation operation we can only condition on the size of the sample of a strongly Rayleigh distribution. To study Y conditioned on $X = 2$ we first need to project on $E \setminus \delta(u)$ then truncate this to $n - 3$. This is because $X = 2$ if and only $|T \cap (E \setminus \delta(u))| = n - 3$.

Note that under μ'_4, X can only take values $1, 2, 3$, it cannot be zero because the degree of u is at least 1 and it cannot be 4 because the degree of v is at least 1. We use negative association and stochastic dominance to show that

$$\begin{aligned} \mathbb{P}_{\mu'_4}[X \le 2] = \mathbb{P}_{\mu'_4}[Y \ge 2] &\quad \ge 1/8, \\ \mathbb{P}_{\mu'_4}[X \ge 2] = \mathbb{P}_{\mu'_4}[Y \le 2] &\quad \ge 1/8. \end{aligned} \tag{12.4}$$

Observe that by (12.3) this implies that $\mathbb{P}_{\mu'_4}[X = 2] \ge 1/8$ and we are done.

First, by stochastic dominance property,

$$\mathbb{P}_{\mu'}[X \le 2 | X + Y = 4] \ge \mathbb{P}_{\mu'}[X \le 2 | X + Y = 5] \ge \dots,$$

so

$$\mathbb{P}_{\mu'}[X \le 2 | X + Y = 4] \ge \mathbb{P}_{\mu'}[X \le 2 | X + Y \ge 4].$$

Similarly,

$$\mathbb{P}_{\mu'}[Y \ge 2 | X + Y = 4] \ge \mathbb{P}_{\mu'}[Y \ge 2 | X + Y \le 3].$$

Therefore,

$$\begin{aligned} \mathbb{P}_{\mu'}[X \le 2 | X + Y = 4] &+ \mathbb{P}_{\mu'}[Y \ge 2 | X + Y = 4] \\ &\ge \mathbb{P}_{\mu'}[X \le 2 | X + Y \ge 4] + \mathbb{P}_{\mu'}[Y \ge 2 | X + Y \le 3] \\ &= \mathbb{P}_{\mu'}[X \le 2, Y \ge 2 | X + Y \ge 4] + \mathbb{P}_{\mu'}[Y \ge 2, X \le 2 | X + Y \le 3] \\ &\ge \mathbb{P}_{\mu'}[X \le 2, Y \ge 2] \end{aligned} \tag{12.5}$$

Since $f = \mathbb{I}[X \le 2]$ is a nonincreasing function and $g = \mathbb{I}[Y \ge 2]$ is a nondecreasing function, and since the two functions are disjointly supported as $\delta(u) \cap \delta(v) = \emptyset$ by negative association property of μ', f, g are positively correlated,

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

Therefore, by (12.5)

$$\begin{aligned} \mathbb{P}_{\mu'_4}[X \le 2] + \mathbb{P}_{\mu'_4}[Y \ge 2] &\ge \mathbb{E}[f \cdot g] \\ &\ge \mathbb{E}[f] \cdot \mathbb{E}[g] \\ &= \mathbb{P}_{\mu}[X \le 2] \cdot \mathbb{P}_{\mu}[Y \ge 2] \ge 1/4. \end{aligned}$$

The last inequality follows by another application of Theorem 12.7 and we leave it as an exercise. This proves Equation 12.4 which implies that $\mathbb{P}[X = 2 | X + Y = 4] \ge 1/8$. This completes the proof of Lemma 12.6.

12.4 Conclusion

It is a fascinating open problem to extend Theorem 12.1 to all feasible solutions of the Held-Karp relaxation. Perhaps one of the easiest hard cases is when the vector x is half integral, i.e., when $x(e)$ is either $1/2$ or 1 for every edge of G. It is conjectured that these are integrality gap examples of the Held-Karp LP relaxation for Symmetric TSP [SWZ12]. Note that in this case any cut of value larger than 2 has size at least 2.5. So, we only need to study structure of cuts of value exactly 2. This structure is very well studied and it is known as the cactus representation [DKL76; FF09].

It is also interesting to design tight examples for Algorithm 12.1 for general metrics. The worst example that we are aware is the complete graph that we illustrated in Example 12.3, that gives an approximation factor of $5/4$ in expectation.

References

[Ben95] A. A. Benczúr. "A Representation of Cuts within 6/5 Times the Edge Connectivity with Applications". In: *FOCS*. 1995, pp. 92–102 (cit. on p. 12-4).

[BG08] A. A. Benczúr and M. 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 p. 12-4).

[DKL76] E. Dinits, A. Karzanov, and M. 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 p. 12-6).

[EJ73] J. Edmonds and E. L. Johnson. "Matching, Euler tours and the Chinese postman". In: *Mathematical Programming* 5.1 (1973), pp. 88–124 (cit. on p. 12-2).

[FF09] T. Fleiner and A. 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 p. 12-6).

[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 p. 12-5).

[MS11] T. Moemke and O. Svensson. "Approximating Graphic TSP by Matchings". In: *FOCS*. 2011, pp. 560–569 (cit. on p. 12-1).

[Muc12] M. Mucha. "$\frac{13}{9}$-approximation for graphic TSP." In: *STACS*. 2012, pp. 30–41 (cit. on p. 12-1).

[OSS11] S. Oveis Gharan, A. Saberi, and M. Singh. "A Randomized Rounding Approach to the Traveling Salesman Problem". In: *FOCS*. 2011, pp. 550–559 (cit. on pp. 12-1, 12-2, 12-4).

[SV12] A. Sebö and J. Vygen. "Shorter Tours by Nicer Ears:" CoRR abs/1201.1870. 2012 (cit. on p. 12-1).

[SWZ12] F. Schalekamp, D. P. Williamson, and A. van Zuylen. "A proof of the Boyd-Carr conjecture". In: *SODA*. 2012, pp. 1477–1486 (cit. on p. 12-6).