CS 798 - Algorithmic Spectral Graph Theory, Fall 2015, Waterloo

Lecture 22: Strongly Rayleigh measures

Recently, there is significant progress in better approximations for traveling salesman problems.

We study some underlying mathematics behind these developments, including the concept of strongly Rayleigh probability measures and an interesting generalization of Weaver's conjecture using it.

Thin tree

Given an undirected graph $G=(V,E)$, for $0<\alpha<1$, we say a spanning tree T is α-thin if for all $S \subseteq V$, we have $|\delta_T(S)| \leq \alpha \cdot |\delta_G(S)|$.

In words, a spanning tree is α-thin if it uses at most α fraction of edges in every cut.

There is a very interesting conjecture about the existence of a thin tree.

Goddyn's conjecture Every k-edge-connected graph has a $O(\frac{1}{k})$-thin spanning tree.

This is a strong conjecture as the condition is clearly necessary.

Besides being an interesting conjecture on its own, it has significant implication to the asymmetric traveling salesman problem (ATSP).

ATSP

Given a directed graph, with a cost $c(e)$ on each edge e, the asymmetric traveling salesman problem is to find a minimum cost tour that visits every vertex at least once.

This is the directed version of the traveling salesman problem.

It is relatively easy to design a $O(\log n)$-approximation algorithm for ATSP.

Not long ago, the best known approximation algorithm is a $O(\log n / \log\log n)$-approximation algorithm, using some properties of random spanning trees that can be understood using strongly Rayleigh measures.

It is observed that a positive resolution of the thin tree conjecture would lead to a constant factor approximation algorithm for ATSP.

Theorem. If there is an algorithm to find a $\frac{c}{k}$-thin tree in a k-edge-connected graph, then there is a c-approximation algorithm for ATSP.

We won't prove this theorem in this course as the techniques used are quite different (LP-based).
The basic idea is a Christofides-type algorithm: find a minimum spanning tree and then augment it into a tour. In the undirected case, we add a min-cost perfect matching between odd degree vertices, and this gives a 1.5-approximation algorithm for TSP.
In the directed case, we use a flow to augment the tree, and it can be shown that if the tree is thin, then a small multiple of the LP solution can be used to find a feasible flow.
See the course notes of Oveis Gharan if you are interested.

Spectrally thin tree

It can be shown that a random spanning tree is a $O(\frac{\log n}{\log\log n} \cdot \frac{1}{k})$-thin tree.
The argument is similar to cut sparsification, using Chernoff bound and careful union bound.
As mentioned before, we can apply Chernoff bound because the edges in a random spanning tree are negatively correlated.

As in graph sparsification, we consider a spectral generalization of thinness of a tree.
We say a tree T in α-spectrally-thin if $L_T \precsim \alpha L_G$.
Note that it is a stronger notion than (combinatorically) thinness, as $L_T \precsim \alpha L_G$ implies that $|\delta_T(S)| = \chi_S^T L_T \chi_S \leq \alpha \chi_S^T L_G \chi_S = \alpha |\delta_G(S)|$, just as in spectral sparsification.
One advantage of this stronger notion is that it is easier to reason, e.g. given a tree, it is easy to check whether it is α-spectrally-thin, while there is no known method to check whether it is (combinatorically) α-thin.
Moreover, the result by Marcus-Spielman-Srivastava implies non-trivial sufficient condition for the existence of a spectrally thin tree.

Corollary: If the maximum effective resistance of an edge in G is α, then there exists a $O(\alpha)$-spectrally-thin tree.

We also won't prove this result; see Harvey-Olver or the blogpost by Srivastava for details.
The idea is that if maximum effective resistance is α, then we can apply MSS theorem to

partition the edge set of G into two subsets, such that each is a good spectral approximation of G.

We recursively apply MSS in the subgraphs (with slightly weaker bounds on maximum effective resistance) until we cannot apply again. by that time there will be $O(\frac{1}{\alpha})$ subgraphs, each is connected and $O(\alpha)$-spectrally thin.

This gives hope that the techniques developed by MSS can be used to prove Goddyn's conjecture.

The corollary of the MSS theorem gives a spectrally thin tree, which is combinatorically thin, but it requires a stronger assumption that the max effective resistance is small, which is not necessarily satisfied in a k-edge-connected graph.

The recent breakthrough made by Anari and Oveis Gharan, in a very high level, can be seen as a way to reduce the combinatorial problem to the spectral problem and use a generalization of MSS theorem to solve the problem.

The reduction, however, is quite complicated and technically challenging.

We will discuss their generalization of MSS theorems, and briefly mention how they apply it.

Let's start with the underlying mathematics.

Strongly Rayleigh measure

Let $\mu: 2^{[m]} \to \mathbb{R}$ be a probability distribution of a set of m elements, i.e. $\mu(S) \geq 0$ for every subset $S \subseteq [m]$ and $\sum_{S \subseteq [m]} \mu(S) = 1$.

An important example to keep in mind is the probability distribution of random spanning trees in which the ground set $[m]$ is the set of edges and $\mu(T) = \frac{1}{N}$ if the edges in T form a spanning tree where N is the number of spanning trees in the graph, and zero otherwise. More generally, we can consider distributions where each spanning tree has probability proportional to the product of the edge weights.

Given a probability distribution $\mu: 2^{[m]} \to \mathbb{R}$, its generating polynomial $g(z_1, \ldots, z_m)$ is defined as

$$g(z_1, \ldots, z_m) := \sum_{S \subseteq [m]} \mu(S) \prod_{i \in S} z_i$$, i.e. the coefficient of the monomial $\prod_{i \in S} z_i$ is $\mu(S)$.

A probability distribution is called <u>strongly Rayleigh</u> if its generating polynomial is real stable.

Before we discuss properties and examples of strongly Rayleigh measures, we first see some more operations preserving real stability.

Closure operations

Given a real stable polynomial $f(z_1,\dots,z_m)$, we have seen in L19 that substitution by real numbers (i.e. $f(a,z_2,\dots,z_m)$ for $a\in\mathbb{R}$) and applying $(1+t\frac{\partial}{\partial z_i})$ are real-stability preserving.

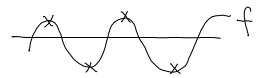
We see two more simple real-stability preserving operations that we will use today.

① Reweighting: If $f(z_1,\dots,z_m)$ is real stable, then $f(w_1z_1,\dots,w_mz_m)$ is real stable if $w_i\geq 0\ \forall i$.

(The proof is straighforward.)

② Differentiation: If $f(z_1,\dots,z_m)$ is real stable, then $\frac{\partial}{\partial z_i}f(z_1,\dots,z_m)$ is real stable.

First, consider the one-dimensional case. If $f(z)$ is real-rooted, then it is easy to see that $f'(z)$ is real-rooted, as between two roots of f there is a root of f', and so f' interlaces f.



More generally, the Gauss-Lucas theorem says that the (complex) roots of $f'(z)$ is in the convex hull of the roots of $f(z)$.

Then, the general case follows from the one-dimensional case, as we can fix all the other variables and consider the univariate restriction as we did in L19.

Properties of strongly Rayleigh measures

Some closure properties of strongly Rayleigh measures follow from the closure properties of real-stability.

① Conditioning: For a variable x_i, the conditional probability distribution $\mu_{|x_i=0}$ and $\mu_{|x_i=1}$ are also strongly Rayleigh measures.

To see this, let $g(x_1,\dots,x_m)$ be the generating polynomial of μ.

The generating polynomial of $\mu_{|x_i=0}$ is simply $g(x_1,\dots,x_{i-1},0,x_{i+1},\dots,x_m)$, and thus is real stable because of real substitution, and hence $\mu_{|x_i=0}$ is strongly Rayleigh.

The generating polynomial of $\mu_{|x_i=1}$ is $x_i\frac{\partial}{\partial x_i}g(x_1,\dots,x_m)$, and is real stable by differentiation.

Okay, we need to rescale the polynomial so that the sum of coefficients remains one.

Therefore, fixing the value of a subset of variables, the conditional probability measure remains strongly Rayleigh.

② Negative Correlation : This is probably the most useful property of strongly Rayleigh measures, as it allows us to apply Chernoff bounds to prove concentration.

The simplest form of negative correlation is $Pr(x_i=1 \mid x_j=1) \le Pr(x_i=1)$ for any two elements i, j.

Note that the probability $Pr(x_i=1)$ can be read from the generating polynomial $g(x_1,\dots,x_m)$ of μ,

as $Pr(x_i=1) = \frac{\partial}{\partial x_i} g \big|_{x_1=\dots=x_{i-1}=x_{i+1}=x_m=1} = \sum_{\substack{S\subseteq[m] \\ i\in S}} \mu(S)$, the sum of coefficients containing i.

Therefore, the negative correlation inequality can be rewritten as $Pr(x_i=1, x_j=1) \le Pr(x_i=1)Pr(x_j=1)$,

and this can be expressed as $\left(\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} g(\vec{1})\right)\cdot g(\vec{1}) \le \left(\frac{\partial}{\partial x_i} g(\vec{1})\right)\left(\frac{\partial}{\partial x_j} g(\vec{1})\right)$.

Strongly Rayleigh measures satisfy this inequality for any values (other than $\vec{1}$).

<u>Proposition</u> $\left(\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} g(z)\right)\cdot g(z) \le \left(\frac{\partial}{\partial x_i} g(z)\right)\left(\frac{\partial}{\partial x_j} g(z)\right)$.

<u>proof</u> For any $z\in\mathbb{R}^m$, consider the bivariate restriction $f(s,t) = g(z_1,\dots,z_{i-1}, z_i+s, z_{i+1},\dots, z_j+t,\dots z_m)$.

Note that $f(s,t) = g(z) + \left(\frac{\partial}{\partial x_i} g(z)\right)s + \left(\frac{\partial}{\partial x_j} g(z)\right)t + \left(\frac{\partial}{\partial x_i}\frac{\partial}{\partial x_j} g(z)\right)st$, and f is real stable.

The proposition would follow if we show that for any bivariate real-stable polynomial

$f(s,t) = a + bs + ct + dst$, we have $bc \ge ad$.

Consider when $f(s,t)=0$.

This implies that $s = -\frac{a+ct}{b+dt}$, and thus $s+\frac{c}{d} = \frac{bc-ad}{bd+d^2t}$.

Suppose, by contradiction, that $bc-ad<0$.

If $d\ne0$, then whenever $Im(t)>0$, we have $Im(s)>0$, contradicting real stability.

If $d=0$, then $s = -\frac{a}{b}-\frac{ct}{b}$, then our assumption $bc<0$ would imply the same contradiction. □

Using the simple form of negative correlation as bases, one can inductively prove the following strong form of negative association:

For any two non-decreasing functions f and g depending on disjoint subsets of variables,

we have $E[f(x_1,\dots,x_n)]\,E[g(x_1,\dots,x_n)] \ge E[f(x_1,\dots,x_n)\,g(x_1,\dots,x_n)]$.

This is proved by Feder and Mihail.

Determinantal measure

We now study an important class of strongly Rayleigh measures.

<u>Definition</u> A probability distribution $\mu: 2^{[m]}\to\mathbb{R}$ is determinantal if there exists a matrix $A\in\mathbb{R}^{m\times m}$ such

that $\sum_{T: S\subseteq T} \mu(T) = \det(A_{S,S})$, where $A_{S,S}$ is the $|S|\times|S|$ submatrix of A restricting to the rows and columns corresponding to S.

<u>proposition</u> If μ is determinantal with a matrix $0 \preceq A \preceq I$, then μ is strongly Rayleigh.

<u>proof</u> Consider $g(z) = \det(I - A + AZ)$, where Z is the diagonal matrix with $Z_{i,i} = z_i$.

We claim that g is real stable and the coefficient of the monomial $\prod_{i\in S} z_i$ is $\mu(S)$.

First, to see real stability, note that $g(z) = \det(I - A + AZ) = \det(A)\det(A^{-1} - I + Z)$.

So, $g(z)$ can be written in the form $\det(B + \sum_{i=1}^{m} z_i A_i)$ where $B \succeq 0$ and $A_i \succeq 0\ \forall i$, and this is real stable as shown in L19 (not quite, but very close, using continuity argument).

To check the coefficients, let $S = \{1, \dots, k\}$ and define $x_S \in \{0,1\}^n$ as $x_S(i) = \begin{cases} 1 & \text{if } i \in S \\ 0 & \text{otherwise} \end{cases}$, then $g(x_S)$ is equal to $\sum_{T\subseteq S} \mu(T)$, and we should check that $\det\left(I - A + A\begin{pmatrix} \overbrace{1 \ddots 1}^{k} & \\ & 0 \ddots 0 \end{pmatrix}\right)$ gives the same result.

Note that if these results agree for every subset S, then we can conclude that the coefficient on the monomial $\prod_{i\in S} z_i$ is $\mu(S)$.

To see it, $\det\left(I - A + A\begin{pmatrix} \overbrace{1 \ddots 1}^{k} & \\ & 0 \ddots 0 \end{pmatrix}\right) = \det\begin{pmatrix} I_k & -A_{S,\bar{S}} \\ 0 & I_{n-k} - A_{\bar{S},\bar{S}} \end{pmatrix} = \det(I_{n-k} - A_{\bar{S},\bar{S}})$.

Observe that
$$\begin{aligned}
\det(I - A_{\bar{S},\bar{S}}) &= 1 + \sum_{i=1}^{n-k} (-1)^i \sum_{T\subseteq \bar{S} : |T| = i} \det(A_{T,T}) \\
&= 1 + \sum_{i=1}^{n-k} (-1)^i \sum_{T\subseteq \bar{S} : |T| = i} \Pr(\text{output contains } T) && \text{by definition of determinental measure} \\
&= 1 - \Pr(\text{output contains some element in } \bar{S}) && \leftarrow \text{inclusion-exclusion principle} \\
&= \Pr(\text{output contains no element in } \bar{S}) \\
&= \sum_{T\subseteq S} \mu(T). \ \square
\end{aligned}$$

One interesting example of determinantal measure is the uniform random spanning tree measure.

Let G be an undirected graph with m edges.

Let Y be the $m\times m$ matrix where each row and column corresponds to an edge, and
$Y_{e,f} = \langle L_G^{\dagger/2} b_e, L_G^{\dagger/2} b_f \rangle$ where $b_e = x_i - x_j$ for an edge ij.

Burton and Pemantle proved that the uniform random spanning tree measure is determinantal using this matrix. In the following, we use the notation $T\sim\mu$ to denote a random spanning tree.

<u>Theorem</u> Let F be a subset of edges. Then $\Pr_{T\sim\mu}(F\subseteq T) = \det(Y_{F,F})$.

We won't prove this theorem. The proof is about one page long, using inclusion-exclusion and induction similar to the above result. See the course notes (L03) by Oveis Gharan for a proof.

To get some feeling. $\Pr_{T\sim\mu}(e\in T) = \det(Y_{e,e}) = \langle L_G^{+/2} b_e, L_G^{+/2} b_e\rangle = b_e L_G^{+} b_e = \mathrm{Reff}(e)$, agreeing with what we know.

More generally, it is also true in the setting when each edge has a weight, and a spanning appears with probability proportional to the product of the edge weights.

This theorem shows that random spanning trees enjoy the strong negative correlation properties for strongly Rayleigh measures, even if we do conditioning on any subset of variables.

These properties were used successfully in studying the (symmetric) traveling salesman problem.

Again. See the course notes of Oveis Gharan for more information about this.

Generalization of MSS using strongly Rayleigh measures

Recall that Marcus-Spielman-Srivastava proved the following theorem.

Theorem Let $v_1, \ldots, v_m$ be independent random vectors in $\mathbb{R}^n$ with finite support such that

$$\sum_{i=1}^{m} E[v_i v_i^T] = I \quad \text{and} \quad E\|v_i\|^2 \le \varepsilon \quad \text{for all } 1 \le i \le m,$$

$$\text{then} \quad \Pr\left[\left\| \sum_{i=1}^{m} v_i v_i^T \right\| \le (1+\sqrt{\varepsilon})^2 \right] > 0.$$

This result crucially relies on the assumption that the vectors $v_1, \ldots, v_m$ are independent r.v.

Anari and Oveis Gharan proved the following beautiful generalization.

Theorem Let μ be a homogenous strongly Rayleigh measure on $[m]$ with the maximum probability of an element is ε_1 (i.e. $\Pr_{S\sim\mu}(i\in S) \le \varepsilon_1, \ \forall i$). Given $v_1, \ldots, v_m \in \mathbb{R}^n$ with $\sum_{i=1}^{m} v_i v_i^T = I$ and $\|v_i\|^2 \le \varepsilon_2$ for $1 \le i \le m$, we have

$$\Pr_{S\sim\mu}\left[\left\| \sum_{i\in S} v_i v_i^T \right\| \le 4(\varepsilon_1+\varepsilon_2) + 2(\varepsilon_1+\varepsilon_2)^2 \right] > 0.$$

Let's see that it can be used to prove directly the corollary about spectrally thin tree, without using recursion.

As before, we set $v_i = L^{+/2} b_{e_i}$ so that $\sum_{i=1}^{m} v_i v_i^T = I$ and $\|v_i\|^2 = \mathrm{Reff}(e_i) \le \alpha$.

We can set the probability measure μ to be the uniform random spanning tree measure, so that it is strongly Rayleigh and homogenous (i.e. each monomial with non-zero coefficient is of the same degree), and furthermore $\Pr_{T\sim\mu}(e\in T) = R_{eff}(e) \le \alpha$.

So, we can apply the new theorem with $\varepsilon_1 = \varepsilon_2 = \alpha$, and directly get a spanning tree T with $\|\sum_{i\in T} v_i v_i^T\| \le O(\alpha)$.

The key advantage of the theorem is that we can guarantee that the output is a spanning tree, so that we get connectivity for free (instead of worrying about minimum eigenvalue), and this is really important in attacking the thin tree conjecture as we will discuss.

Proof ideas

The proof is also based on the method of interlacing families, with the same two key steps:

① Proving the family $\det(xI - \sum_{i\in S} v_i v_i^T)_{S\in supp(\mu)}$ forms an interlacing family.

② Bound the maxroot of $\mathbb{E}_{S\sim\mu} \det(xI - \sum_{i\in S} v_i v_i^T) = \sum_S \mu(S) \det(xI - \sum_{i\in S} v_i v_i^T)$.

Recall that both steps depend crucially on a key formula writing the expected characteristic polynomial as applying some differential operations on a multivariate polynomial.

Anari and Oveis Gharan proved a generalization incorporating the probability measure.

<u>Theorem</u> Let $v_i \in \mathbb{R}^n$ and the degree of the homogenous polynomial be d.

Then $$x^{d-n} \sum_S \mu(S) \det(x^2 I - \sum_{i\in S} 2 v_i v_i^T) = \prod_{i=1}^m \left(1 - \frac{\partial^2}{\partial z_i}\right) g(x\vec{1}+z) \cdot \det(xI + \sum_{i=1}^m z_i v_i v_i^T)\Big|_{z_1=\dots=z_m=0}$$

<u>proof</u> The main idea is to write two polynomials, one with $\mu(S)$ as the coefficient of $\prod_{i\in S} z_i$ and another with $\det(x^2 I - \sum_{i\in S} 2 v_i v_i^T)$ as the coefficient of $\prod_{i\in S} z_i$.

If this can be done, then we can multiply this two polynomial, and take out $\mu(S)\det(x^2 I - \sum_{i\in S} 2 v_i v_i^T)$ as the coefficient of $\prod_{i\in S} z_i^2$ by differenting with $\prod_{i\in S} \frac{\partial^2}{\partial z_i}$ on the product of that two polynomials and substituting zero.

This really is the main idea but it requires some care to work it out.

Let's start with the left hand side. Let $A_i = v_i v_i^T$ (so that there is no confusion with tree T)

$$\sum_T \mu(T) \det(xI + t_i \sum_{i\in T} A_i) = \sum_T \mu(T) \prod_{i\in T}\left(1 + \frac{t_i \partial}{\partial z_i}\right) \det(xI + \sum_{i\in T} z_i A_i)\Big|_{z_1=\dots=z_m=0} \quad \text{(multilinear formula)}$$

$$= \sum_T \mu(T) \sum_{S\subseteq T} \left(\prod_{i\in S} \frac{t_i \partial}{\partial z_i}\right) \det(xI + \sum_{i=1}^m z_i A_i)\Big|_{z_1=\dots=z_m=0} \quad \text{(expanding and including all variables)}$$

$$= \sum \left(\sum \mu(T) \right) \cdot \prod \frac{t_i \partial}{\partial z_i} \det(xI + \sum^m z_i A_i) \Big|$$

$S \subseteq T \quad i \in S \; \partial z_i) \quad \ldots \quad i=1 \quad \ldots \quad |_{z_1=\ldots=z_m=0}$ including all variables)

$$= \sum_S \Big(\sum_{T: S \subseteq T} \mu(T) \Big) \cdot \prod_{i \in S} \frac{t_i \partial}{\partial z_i} \det(xI + \sum_{i=1}^m z_i A_i) \Big|_{z_1 = \ldots = z_m = 0} .$$

Let's come up one polynomial with coefficient $\sum_{T: S \subseteq T} \mu(T)$ on the monomial $\prod_{i \in S} z_i$, and another polynomial with coefficient $\prod_{i \in S} \frac{\partial}{\partial z_i} \det(xI + \sum_{i=1}^m z_i A_i) |_{z_1 = \ldots = z_m = 0}$ on $\prod_{i \in S} z_i$.

Consider $g(x\vec{1} + z) = \sum_T \mu(T) \prod_{i \in T} (x + z_i)$. Each T with $S \subseteq T$ will contribute $\mu(T) x^{|T| - |S|}$ to $\prod_{i \in S} z_i$.

Therefore, the coefficient of $\prod_{i \in S} z_i$ is $\sum_{T: S \subseteq T} \mu(T) x^{|T|-|S|} = \sum_{T: S \subseteq T} \mu(T) x^{d - |S|}$ since μ is homogenous.

Consider $f(z) = \det(xI + \sum_{i=1}^m z_i A_i)$.

Then, the coefficient of $\prod_{i \in S} z_i$ is $\prod_{i \in S} \frac{\partial}{\partial z_i} \det(xI + \sum_{i=1}^m z_i A_i) |_{z_1 = \ldots = z_m = 0}$.

So, since both f and g are multilinear, the coefficient of $\prod_{i \in S} z_i^2$ is the product of the above coefficients, so $\prod_{i \in S} \frac{\partial}{\partial z_i^2} f \cdot g \big|_{z_1 = \ldots = z_m = 0} = 2^{|S|} x^{d - |S|} \Big(\sum_{T: S \subseteq T} \mu(T) \Big) \cdot \prod_{i \in S} \frac{\partial}{\partial z_i} \det(xI + \sum_{i=1}^m z_i A_i) \big|_{z_1 = \ldots = z_m = 0}$

Therefore, $\prod_{i=1}^m \big(1 - \frac{\partial}{\partial z_i^2}\big) f \cdot g \big|_{z_1 = \ldots = z_m = 0} = \sum_{S \subseteq [m]} (-1)^{|S|} 2^{|S|} x^{d - |S|} \Big(\sum_{T: S \subseteq T} \mu(T) \Big) \prod_{i \in S} \frac{\partial}{\partial z_i} \det(xI + \sum_{i=1}^m z_i A_i) \big|_{z_1 = \ldots = z_m = 0}$.

The LHS of this equality is precisely the RHS of the theorem.

The RHS of this equality is equal to x^d times the expected characteristic polynomial with $t_i = \frac{-2}{x}$.

Hence, $\prod_{i=1}^m \big(1 - \frac{\partial}{\partial z_i^2}\big) f \cdot g \big|_{z_1 = \ldots = z_m = 0} = x^d \sum_T \mu(T) \det(xI - \frac{2}{x} \sum_{i=1}^m A_i)$

$$= x^{d-n} \sum_T \mu(T) \det(x^2 I - 2 \sum_{i=1}^m A_i). \quad \square$$

Once this formula is established, the remaining proof is quite similar to that in MSS.

First, the formula shows that the expected characteristic polynomial is real-rooted, because g is real stable as μ is strongly Rayleigh, and so $f \cdot g$ is because both f, g are, and then the operation $1 - \frac{\partial^2}{\partial z_i}$ is real-stability preserving because it is equal to $(1 - \frac{\partial}{\partial z_i})(1 + \frac{\partial}{\partial z_i})$ and both are real-stability preserving by L19.

Using this, one can prove that the family forms an interlacing family.

In the second step, they upper bound the maxroot of the expected characteristic polynomial using the multivariate barrier argument.

The potential function is the same $\frac{\partial}{\partial z_i} \log p(z_1, \ldots, z_m)$.

They need to also compute the second derivative because of $1-\frac{\partial^{-}}{\partial z_i}$.

One main difference is that now the target upper bound is $O(\varepsilon_1+\varepsilon_2)$, much smaller than $(1+\sqrt{\varepsilon})^2$ as in MSS, but it turns out that the $1-\frac{\partial^2}{\partial z_i}$ operation allows much smaller shift.

There are also some additional technical details but all the main ideas are mentioned.

Applications to ATSP

The fundamental building block of the new ATSP result is the following application.

Theorem Given a graph $G=(V,E)$ and a set $F\subseteq E$ such that (V,F) is k-edge-connected, if $R_{eff}(e)\le\varepsilon$ for all $e\in F$, then G has a $O(\frac{1}{k}+\varepsilon)$ spectrally thin tree in F.

proof idea Since F is k-edge-connected, there are at least $k/2$ edge disjoint spanning trees in F.

This implies that there is a point in the spanning tree polytope with max value $= O(\frac{1}{k})$.

It turns out that this point implies that there is a reweighting of the edges, so that the weighted random spanning tree distribution μ has maximum edge probability $O(\frac{1}{k})$, i.e. $\varepsilon_1 = O(\frac{1}{k})$. This proof is based on the use of maximum entropy distribution.

The assumption about effective resistance implies that $\varepsilon_2=\varepsilon$, and then the main theorem implies this theorem. □

With the theorem, their strategy is to add "short-cut" edges into the graph, so that it doesn't change the cut structures much, while creating many edges with small effective resistance.

They do it iteratively so that the edges with small effective resistance form a k-edge-connected subgraph.

The most difficult step is to show the existence of good short-cut edges, which they prove by a delicate analysis of a semidefinite program.

There are still 80 pages after the theorem above!

In the end, they prove that any k-edge-connected graph has a $O(\frac{\log\log n}{k})$-thin tree.

However, it doesn't imply a polynomial time $O(\log\log n)$-approximation algorithm for ATSP yet, as the theorem that we presented today is non-constructive.

It is a very interesting open problem to make this approach constructive.

Also, it is very interesting to find new applications of these new techniques.

References

- The Kadison-Singer problem for strongly Rayleigh measures and applications to asymmetric TSP, by Anari and Oveis Gharan, 2015.
- Effective-resistance-reducing-flows, spectrally thin trees, and asymmetric TSP, by Anari and Oveis Gharan, 2015.
- Negative dependence and the geometry of polynomials, by Borcea, Branden, Liggett, 2008.
- Course notes of "Recent developments in approximation algorithms", by Oveis Gharan, 2015.
- Course notes of "The geometry of polynomials in algorithms, combinatorics and probability", by Srivastava, 2015.
- Pipage rounding, pessimistic estimators and matrix concentration, by Harvey and Olver, 2014.