13. Short codes and sum of squares.

Lecturer: Massimo Lauria

Disclaimer: this lecture note has not yet been reviewed by the main lecturer. It is released as it is for the convenience of the students.

http://www.csc.kth.se/~lauria/sos14/

This lecture is based on the paper [1]. For the purpose of this lecture we work with regular graphs, that is graphs where all vertices have the same number of edges incident to them. The central notion for this lecture is the following notion of expansion of a set of vertices in a graph G.

[1] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. In *Proceedings of the 53rd Annual Symposium on Foundations of Computer Science (FOCS '12)*, pages 370–379, 2012

Definition 1. *For a regular graph $G = (V, E)$ and a set of vertices $S \subseteq V$, the expansion $\Phi(S)$ of S is defined as*

$$\Phi(S) = \Pr_{\{x,y\}\in E}(y \notin S \mid x \in S).$$

The volume $\mu(S)$ of a set $S \subseteq V$ is defined as $\mu(S) = \Pr_{\{x,y\}\in E}(x \in S)$.

As we are working with regular graphs, the volume is also equal to $\mu(S) = \Pr_{x\in V}(x \in S)$. If we take the whole set V the expansion we get is $\Phi(V) = 0$, because all vertices are in V while none are outside of V. Also, if we look at a single vertex $v \in V$, then we define the expansion $\Phi(\{v\}) = 1$ as every edge connects two distinct vertices. Furthermore we have $\Phi(\phi) = 1$, in which ϕ denotes the empty set.

We also identify the d-regular graph G with its adjacency matrix, where rows and columns correspond to vertices in the graph and $G_{x,y} = 1/d$ if $\{x, y\}$ is an edge in G and 0 otherwise. For two real-valued functions $f, h\colon V \to \mathbb{R}$, we define the scalar product $\langle f, h\rangle$ to be the expectation of $f(x)h(x)$ taken over vertices $x \in V$, that is $\langle f, h\rangle = \mathbb{E}_{x\in V} f(x)h(x)$. Note that this definition is a linear factor off from the standard definition we would get from viewing f and h as two vectors indexed by $x \in V$. In order to preserve known relations we also define the p-norm $\|f\|_p = \left(\mathbb{E}_{x\in V} f(x)^p\right)^{1/p}$. We drop the subscript when $p = 2$, that is $\|f\| = \|f\|_2$.

For two function $f, h\colon V \to \mathbb{R}$, we are interested in the following bilinear form $\langle f, Gh\rangle$. Note that we can write the form as follows:

$$\langle f, Gh\rangle = \mathop{\mathbb{E}}_{\{x,y\}\in E} f(x)h(y), \tag{1}$$

because $[Gh]_x = \mathbb{E}_{\{x,y\}\in E} h(y)$ as by definition of the matrix G we sum up all $h(y)$ that correspond to neighbors y of vertex x and then divide by d, which is the number of neighbors of x. Combining this observation with the definition of the scalar product we get the equation in (1).

Using bilinear forms we can write the expansion of a vertex set S as

$$\Phi(S) = \frac{\langle \mathbb{1}_S, G(\mathbb{1} - \mathbb{1}_S)\rangle}{\|\mathbb{1}_S\|^2}, \tag{2}$$

where $\mathbb{1}$ is a vector consisting of only 1's, while $\mathbb{1}_S$ has 1 at positions $v \in S$ and 0 elsewhere. This follows from noting that $\Pr_{\{x,y\}\in E}(x \in S \wedge y \in$

$S') = \mathbb{E}_{\{x,y\}\in E} \mathbb{1}_S(x)\mathbb{1}_{S'}(y)$ for any two sets of vertices $S, S' \subseteq V$, and the expectation is just the scalar product $\langle \mathbb{1}_S, G\mathbb{1}_{S'} \rangle$. Also $\mu(S) = \|\mathbb{1}_S\|^2$. The previous equation can also be written as

$$\Phi(S) = 1 - \frac{\langle \mathbb{1}_S, G\mathbb{1}_S \rangle}{\|\mathbb{1}_S\|^2}. \tag{3}$$

Definition 2. *Consider a regular graph $G = (V, E)$. We say that G is a δ-small set expander if for every set of vertices $S \subseteq V$ such that $\mu(S) \leq \delta$ it holds that $\Phi(S) \geq 0.9$.*

A family of graphs $\{G_N\}$ is a small-set-expanders if there exists a $\delta_0 > 0$ such that every graph in the family is a δ_0-small set expander.

Note that this definition of expansion differs from the "standard" one in that the "standard" definition sets δ to $1/2$. As we can take a much smaller δ we can require larger expansion.

The motivation for this work is the Small Set Expansion Hypothesis, which states that it is NP-hard to distinguish between the case when a given graph is a small set expander and the case when the graph has a small set with very bad expansion (close to 0). The importance of this hypothesis is that it implies the Unique Games Conjecture, which is an important conjecture in the hardness of approximation.

In particular, when we consider Sum-Of-Squares proof system, the Small Set Expansion Hypothesis predicts that there exists a family of small set expanders that requires a degree $n^{\Omega(1)}$ to prove that the graph is not a small set expander. As a small number of large eigenvalues implies the existence of small degree Sum-Of-Squares proof, it means that we want to find a family of small set expanders that has $n^{\Omega(1)}$ eigenvalues close to 1.

In this lecture we discuss the result that shows that there exists a family of expanders that almost has the wanted property of having a many eigenvalues close to 1. However, the result fails at providing the degree lower bound for the Sum-Of-Squares proof system as the constructed expanders still have small degree proofs.

We define $R_\varepsilon(G)$ to be the number of eigenvalues of the graph G that are at least $1 - \varepsilon$. For example, if we take the graph G to be a disjoint union of $1/\delta_0$ "standard" expanders on $\delta_0 N$ vertices, then G is a small set expander as every subgraph is an expander. Also, we have that $R_\varepsilon(G) = 1/\delta_0$, because eigenvalues that contribute come only from the partitions S that select each disjoint unit. But this does not work for us as $R_\varepsilon = O(1)$ and does not grow with the size of the graph. In the rest of the lecture we discuss how to get such graphs.

We need the following result, which can be viewed as sort of inverse Cauchy-Schwarz inequality.

Theorem 3 (Hypercontractivity)**.** *For a function f on the hypercube $\{0,1\}^R$, of degree k (when f is written as a multilinear polynomial over reals), it holds that*

$$\mathop{\mathbb{E}}_{\{0,1\}^R} f^4 \leq 9^k \left(\mathop{\mathbb{E}}_{\{0,1\}^R} f^2 \right)^2.$$

The family of graphs which will have the desired properties is the family of so-called Cayley graphs whose vertices are elements of some subspace of $\mathbb{F}_2^R$,

and edges are defined by the difference relation between vertices. Formally, we have the following definition.

Definition 4 (Cayley graph)**.** *For some subspace $\mathcal{D}$ of $\mathbb{F}_2^n$ and a subset $\mathcal{T} \subseteq \mathbb{F}_2^R$, the Cayley graph $\mathrm{Cay}(\mathcal{D}, \mathcal{T})$ is a graph such that $\mathcal{D}$ is the set of vertices and there is an edge between x and y in $\mathcal{D}$ if $(x - y) \in \mathcal{T}$.*

First we show a simpler result that a particular type of Cayley graphs are good expanders with $\log n$ large eigenvalues, and then we discuss how to strengthen this result so that the number of large eigenvalues is close to $n^{\Omega(1)}$. We show the following theorem.

Theorem 5. *Let $\mathcal{T}$ be some subset of $\mathbb{F}_2^R$. If*

1. *$\mathbb{E}_{t\in\mathcal{T}} |t| \leq \varepsilon R$ and*
2. *for every $S \subseteq [R]$ such that $|S| > k$, it holds $\lambda_S \leq 0.01$,*

then the Cayley graph $\mathrm{Cay}(\mathbb{F}_2^R, \mathcal{T})$ has $R_{4\varepsilon}(\mathrm{Cay}(\mathbb{F}_2^R, \mathcal{T})) \geq R/2$ and $\Phi(S) \geq 0.99 - 2^k \mu(S)^{1/4}$ for every subset of vertices S.

The useful thing to note here is that the character functions $\chi_S(x) = (-1)^{\sum_{i\in S} x_i}$, for every subset $S \subseteq [R]$, are eigenvectors of the Cayley graph $G = \mathrm{Cay}(\mathbb{F}_2^R, \mathcal{T})$. An equivalent representation of $\chi_S(x)$ is as a degree $|S|$ polynomial $\chi_S(x) = \prod_{i\in S}(1 - 2x_i)$. To see that χ_S are eigenvectors of G we take the previous observation that $[Gh]_x = \mathbb{E}_{\{x,y\}\in E} h(y)$ and apply it to the character noting that neighbors of x in the Cayley graph are $(x + t)$ for $t \in \mathcal{T}$:

$$(G\chi_S)_x = \mathop{\mathbb{E}}_{t\in\mathcal{T}} \chi_S(x + t) \tag{4}$$

$$= \mathop{\mathbb{E}}_{t\in\mathcal{T}} \chi_S(x)\chi_S(t) \tag{5}$$

$$= \chi_S(x) \mathop{\mathbb{E}}_{t\in\mathcal{T}} (-1)^{\sum_{i\in S} t_i}. \tag{6}$$

The second equality follows from the property of the character function, and we can see that $\mathbb{E}_{t\in\mathcal{T}}(-1)^{\sum_{i\in S} t_i}$ does not depend on x, but only on S. Hence, we denote it by λ_S and we have $G\chi_S = \lambda_S\chi_S$.

Lemma 6. *If $\mathbb{E}_{t\in\mathcal{T}} |t| \leq \varepsilon R$, then $R_{4\varepsilon} \geq R/2$.*

Proof. By the previous observation $\lambda_{\{i\}}$ is equal to

$$\lambda_{\{i\}} = \mathop{\mathbb{E}}_{t\in\mathcal{T}} (-1)^{t_i} = 1 - 2 \Pr_{t\in\mathcal{T}}(t_i = 1), \tag{7}$$

where the last equality follows from the observation that $(-1)^{t_i} = 1 - 2\Pr(t_i = 1)$ and summing up over all $\mathcal{T}$ and dividing by $|\mathcal{T}|$ gives the probability $\Pr_{t\in\mathcal{T}}(t_i = 1)$. As we assume that $\mathbb{E}_{t\in\mathcal{T}} |t| \leq \varepsilon R$, we have that for at least $R/2$ coordinates i the probability $\Pr_{t\in\mathcal{T}}(t_i = 1)$ is at most 2ε. Hence, at least $R/2$ eigenvalues $\lambda_{\{i\}}$ have their value lower bounded by $1 - 4\varepsilon$ and we proved the lemma. □

With the previous lemma we have shown the part of the Theorem 5 that states that we have many eigenvalues close to 1. Now we will show the second part that such graphs have the requirement on the $\Phi(S)$.

Lemma 7. *If for every $S \subseteq [R]$ such that $|S| > k$, it holds $\lambda_S \leq 0.01$, then $\Phi(S) \geq 0.99 - 2^k \mu(S)^{1/4}$ for all S.*

Proof. We take the projection operator P that projects a vector into subspace U defined as $U = \text{Span}\{\chi_S \mid \lambda_S \geq 0.01\}$. By the contrapositive of the assumption of the lemma, we have that for every χ_S that generates U it holds that $|S| \leq k$. Hence, we have that U is a subspace of degree k polynomials.

Write the projector P as the matrix and then it holds that

$$0.01I + P \succeq G, \tag{8}$$

where I is the identity matrix. This is because χ_S are eigenvectors of G. Hence, if we look at P and G as written in basis defined by χ_S, we have that G would be the diagonal matrix with λ_S on its diagonal and P is a diagonal matrix with 1 at positions corresponding to $\lambda_S \geq 0.01$ and 0 elsewhere. Hence, $0.01I + P$ is a diagonal matrix with all values on diagonal greater than the values on the diagonal of matrix G, because when $\lambda_S \geq 0.01$ the diagonal of P is set to 1 and otherwise the identity matrix sets it to 0.01.

From the previous equation and properties of positive semidefinite matrices we get that $\langle f, Gf \rangle \leq \langle f, (0.01 + P)f \rangle$. Hence, for any vector $\mathbb{1}_S$ we have

$$\langle \mathbb{1}_S, G\mathbb{1}_S \rangle \leq 0.01 \|\mathbb{1}_S\|^2 + \langle \mathbb{1}_S, P\mathbb{1}_S \rangle. \tag{9}$$

Note that $\|\mathbb{1}_S\|^2 = \mu(S)$, so we proceed by bounding $\langle \mathbb{1}_S, P\mathbb{1}_S \rangle$ in terms of $\|\mathbb{1}_S\|^2$. We use Hölder's inequality to bound this scalar product:

$$\langle f, g \rangle \leq \|f\|_p \|q\|_q, \tag{10}$$

where $1/p + 1/q = 1$, $p, q \geq 1$. Applying this inequality to $\langle \mathbb{1}_S, P\mathbb{1}_S \rangle$ and setting $p = 4/3$ and $q = 4$, we get

$$\langle \mathbb{1}_S, P\mathbb{1}_S \rangle \leq \|\mathbb{1}_S\|_{4/3} \|P\mathbb{1}_S\|_4. \tag{11}$$

As $\mathbb{1}_S$ is 0-1 valued, we have that $\|\mathbb{1}_S\|_{4/3} = \mu(S)^{3/4}$. By hypercontractivity we have that for $\|P\mathbb{1}_S\|_4$ there is a constant k such that $\|P\mathbb{1}_S\|_4 \leq 2^k \|P\mathbb{1}_S\|_2$. For 2-norm it holds that when projecting we have the following inequality $\|Px\| \leq \|x\|$ and, hence, it holds that $\|P\mathbb{1}_S\|_4 \leq 2^k \mu(S)^{1/2}$. Thus, we have

$$\langle \mathbb{1}_S, P\mathbb{1}_S \rangle \leq \mu(S)^{3/4} 2^k \mu(S)^{1/2} \leq 2^k \mu(S)^{1/2} \mu(S). \tag{12}$$

Putting (12) into (9), we get that

$$\langle \mathbb{1}_S, G\mathbb{1}_S \rangle \leq \left(0.01 + 2^k \mu(S)^{1/4}\right) \mu(S), \tag{13}$$

and as $\Phi(S) = 1 - \frac{\langle \mathbb{1}_S, G\mathbb{1}_S \rangle}{\|\mathbb{1}_S\|^2}$, we have that $\Phi(S) \geq 0.99 - 2^k \mu(S)^{1/4}$ for any set S. □

Taking the Cayley graph with points defined by all vectors in $\mathbb{F}_2^R$ does not give us the result, as we have shown that there are only R large eigenvalues, while the whole space has 2^R points. So, we cannot beat the bound of $\log n$ large eigenvalues. But, if we require just that $R_\varepsilon = \Omega(\log n)$, then we can take $\mathcal{T}$ to be all vectors that have weight ε, which turns out to satisfy the conditions of the theorem.

To actually get close to polynomial bounds we will use a subspace $\mathcal{D}$ of $\mathbb{F}_2^R$ that is much smaller than the whole space $\mathbb{F}_2^R$. We can show the following theorem.

Theorem 8. *Let $\mathcal{T}$ be some subset of $\mathbb{F}_2^R$ and $\mathcal{D}$ be the linear space spanned by $\mathcal{T}$. Let $\mathcal{C} = \mathcal{D}^\perp$ be the dual of $\mathcal{D}$ (set of vectors orthogonal to $\mathcal{D}$). If*

1. $\mathbb{E}_{t\in\mathcal{T}} |t| \leq \varepsilon R$,
2. *for every $\alpha \in \mathbb{F}_2^R$ such that the distance $\mathrm{dist}(\alpha, \mathcal{C})$ between α and $\mathcal{C}$ is greater than k, it holds $\mathbb{E}_{t\in\mathcal{T}}(-1)^{\langle\alpha,t\rangle} \leq 0.01$, and*
3. $k \leq (\mathrm{dist}(\mathcal{C}) - 1)/4$,

then the Cayley graph $\mathrm{Cay}(\mathcal{D}, \mathcal{T})$ has $R_{4\varepsilon}(\mathrm{Cay}(\mathcal{D}, \mathcal{T})) \geq R/2$ and $\Phi(S) \geq 0.99 - 2^k \mu(S)^{1/4}$ for every subset of vertices S.

The first two conditions are analogous to the two conditions in Theorem 5. In the proof there are not much more changes and the goal is just to do the same calculations as in Theorem 5. The two things that change are that some characters might become the same in this subspace, that is $\chi_S(x) = \chi_{S'}(x)$ for all $x \in \mathcal{D}$ and where $S \neq S'$. But this is not a big problem and we still get the same number $R/2$ of large eigenvalues. The second problem is that we need the hypercontractivity to hold on this restricted space $\mathcal{D}$. Here the third condition restricting the smallest distance between elements of $\mathcal{C}$ helps, and we can show that vectors are sufficiently far apart making hypercontractivity true on $\mathcal{D}$.

Choosing $\mathcal{D}$ to be the appropriate Reed-Muller code and $\mathcal{T}$ to be the set of its minimal codewords, we get that the Cayley graphs $\mathrm{Cay}(\mathcal{D}, \mathcal{T})$ are small set expanders with almost $n^{\Omega(1)}$ eigenvalues close to 1. Formally, we have the following corollary.

Corollary 9. *There exists a family of graphs $\{G_n\}_n$ such that every G_n is a small set expander and $R_\varepsilon = 2^{(\log n)^{1/\log(1/\varepsilon)}}$, where n is the number of vertices in the graph.*

Hence, we have shown that there is a family of small set expanders with a lot of eigenvalues close to 1. Unfortunately, these expanders do not give the degree lower bound for the Sum-Of-Squares proof system, as they can be refuted in constant degree. In order to make this formal, the system we are refuting is

$$\sum_{v\in V} x_v = \delta|V|$$
$$\sum_{\{u,v\}\in V} x_u x_v \geq 0.1\delta|E|$$
$$x_v^2 - x_v = 0$$

We interpret the variables to denote that $v \in S$ if $x_v = 1$, and $v \notin S$ if $x_v = 0$. The first condition then states that we are looking at set S of size $\delta|V|$, that is $\mu(S) = \delta$. The second condition states that such a set S has at least $0.1\mu(S)$ fraction of the edges in it and, hence, at most $0.9\mu(S)|E|$ edges are crossing its border giving us $\Phi(S) \leq 0.9$. The third condition is just the standard one that encodes that either we have vertex v in S or not. Thus, the constraints encode that there exists a set with measure $\mu(S) = \delta$ and expansion $\Phi(S) \leq 0.9$.

In refuting this claim on our small set expanders, the main point is that we can run our proof that the constructed Cayley graphs are good expanders in small degree Sum-Of-Squares proof, that is we can prove Lemma 7.

The main problem in using the Sum-Of-Squares to prove that the graph is a small set expander is in proving the hypercontractivity inequality using the Sum-Of-Squares. That is, the problem is to prove that

$$\mathop{\mathbb{E}}_{\{0,1\}^R} (P\mathbb{1}_S)^4 \leq 9^k \left(\mathop{\mathbb{E}}_{\{0,1\}^R} \mathbb{1}_S^2 \right)^2 .$$

But, in our case both the left-hand and the right-hand side will be degree 4 polynomials, so proving that their difference is greater than or equal to 0 in the Sum-Of-Squares proof system involves only degree 4 reasoning. Plugging this into the rest of the proof we can get a constant degree Sum-Of-Squares refutation of the fact that our graph is not an expander.

References

[BGH+12] Boaz Barak, Parikshit Gopalan, Johan Håstad, Raghu Meka, Prasad Raghavendra, and David Steurer. Making the long code shorter. In *Proceedings of the 53rd Annual Symposium on Foundations of Computer Science (FOCS '12)*, pages 370–379, 2012.