Lecture 8: Huang's proof

Rajat Mittal

IIT Kanpur

We have seen many complexity measures and relations between them. Figure 1 summarizes them.

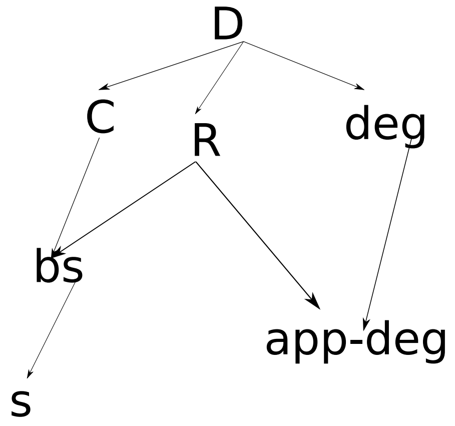

Fig. 1. Relations between complexity measures. Arrow from A to B implies $A = \Omega(B)$.

You saw previously that all these complexity measures in Figure 1 are polynomially related (with the exception of sensitivity).

In a breakthrough, Huang [3] recently showed that even sensitivity is polynomially related to all these measures (called sensitivity conjecture). For a perspective on this result, the conjecture was open for around 30 years, and finally Huang settled it by giving a beautiful proof which can arguably fit in one page.

1 Huang's proof of sensitivity conjecture

The result was shown by introducing a new quantity, spectral sensitivity, denoted $\lambda(f)$ (introduced in [3], formalized in [1]). It was a lower bound on sensitivity (follows easily from the definition), recently it has been shown to be a lower bound on approximate degree [1] (*not* needed for the proof of sensitivity conjecture). This modifies the relationship diagram to Figure 2.

Huang provided a polynomial upper bound on the degree of a function f using $\lambda(f)$.

Exercise 1. Why is that sufficient?

First, we define spectral sensitivity for a Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$. To define spectral sensitivity, we need the concept of *sensitivity graph* of the function f, a subgraph of Boolean hypercube.

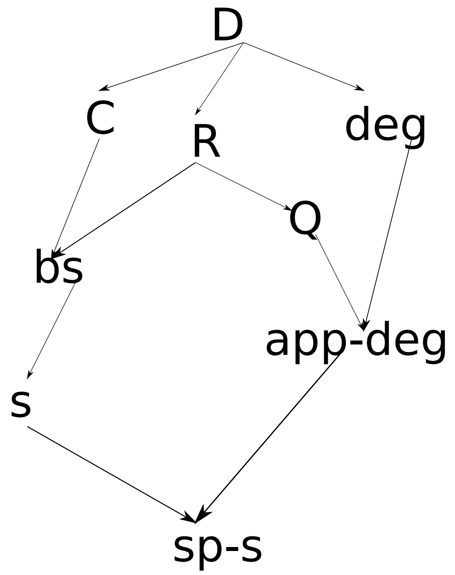

Fig. 2. Relations diagram with spectral sensitivity. Arrow from A to B implies $A = \Omega(B)$.

Exercise 2. What is a Boolean hypercube (as a graph)?

The sensitivity graph of f, say G_f, is a subgraph of Boolean hypercube, i.e., there are 2^n vertices (for each input). An edge x, y is present in G_f iff $f(x) \neq f(y)$ and x, y is an edge in Boolean hypercube (they have Hamming distance 1).

Exercise 3. Find a function f whose sensitivity graph is the Boolean hypercube itself.

Exercise 4. How many edges are there in the sensitivity graph of OR_n.

Exercise 5. Show a subgraph of Boolean hypercube which is not a sensitivity graph for any function f.

We are interested in the eigenvalues of the adjacency matrix, say A_f, of the graph G_f. We first notice that the graph G_f is bipartite.

Exercise 6. Show that Boolean hypercube is bipartite.

That means, if u is an eigenvalue of G_f, then so is $-u$ (assignment). That means we can talk about the maximum eigenvalue (without clarifying if absolute value needs to be taken before taking maximum).

The spectral sensitivity of f, called $\lambda(f)$, is the maximum eigenvalue (also called spectral norm) of the adjacency matrix of G_f.

Since the eigenvalue of a matrix is bounded by the maximum row sum (why), $\lambda(f) \leq \mathrm{s}(f)$. For $\lambda(f) \leq \widetilde{\deg}(f)$, refer to [1]. This completes the relationships given in Figure 2.

The main result of this section is the following upper bound on $\deg(f)$ in terms of λ settling sensitivity conjecture.

Theorem 1 ([3]).

For any Boolean function $f : \{0,1\}^n \to \{0,1\}$,

$$\deg(f) \leq \lambda(f)^2.$$

The first simplification is that we can assume $\deg(f) = n$. If not, pick the monomial in the polynomial representation of f with highest degree, and set all other variables to some values. For the restricted function, $\deg(f)$ is same but $\lambda(f)$ can only be smaller (assignment).

That means we can assume $\deg(f) = n$ (any counterexample to Theorem 1 can be converted into a counterexample with full degree). In other words, we just need to prove that $\lambda(f) \geq \sqrt{n}$ when $\deg(f) = n$.

What can we say about sensitivity graph of f when $\deg(f) = n$? Define $V_0 = \{x : f(x) = \text{PARITY}(x)\}$ and $V_1 = \{x : f(x) \neq \text{PARITY}(x)\}$.

Exercise 7. Show that $\deg(f) = n$ is equivalent to saying that $|V_0| \neq |V_1|$.

The problems statement changes to, given that $|V_0| > 2^{n-1}$ (if $|V_0| < |V_1|$ then consider $1 - f$), show that $\lambda(f) \geq \sqrt{n}$.

Exercise 8. Show that there is no edge between V_0 and V_1. Inside V_0 (and V_1), the edges are exactly the edges of Boolean hypercube.

This means that the eigenvalues of G_f are union of eigenvalues of the subgraph on V_0 and V_1. We know that inside V_0 and V_1, the edges are exactly like the Boolean hypercube. In other words we are interested in the eigenvalues of the induced subgraph on V_0 and V_1. For any V with more than half the vertices, we need to show that the induced subgraph from Boolean hypercube (say G_V) has eigenvalue more than $\sqrt{n}$. This will finish the proof.

An interesting lemma relates the eigenvalues of the induced subgraph with the eigenvalues of the original graph. It is called *Cauchy's interlacing theorem* [3], we will only use the following special case of it.

Lemma 1. *Let G be a graph on k vertices and its eigenvalues be $\lambda_1 \leq \lambda_2 \cdots \leq \lambda_k$. If G_V is the induced subgraph on V with l vertices, then*

$$\|G_V\| \geq \lambda_l,$$

where $\|G_V\|$ denotes the maximum eigenvalue of G_V.

Proof. The adjacency matrix of G is an $k \times k$ matrix. The eigenvectors corresponding to bigger eigenvalues, $\{\lambda_k, \lambda_{k-1}, \cdots, \lambda_l\}$, span a vector space of dimension $k - l + 1$, say $\mathcal{S}_1$. The vector space corresponding to l standard basis vectors e_v where $v \in V$, say $\mathcal{S}_2$, spans a subspace of dimension l.

Exercise 9. Since the sum of dimensions of $\mathcal{S}_1$ and $\mathcal{S}_2$ is more than k, show that their intersection is non-empty.

For the common vector v, $Av = A_V v$ (where A, A_V are the adjacency matrices of G, G_V respectively), and the length of Av is more than λ_l times the length of v. So, we get

$$\|G_V\| := \|A_V\| \geq \lambda_l.$$

□

The adjacency matrix of Boolean hypercube (say H_n) has dimensions $2^n \times 2^n$. Arrange the eigenvalues of H in increasing order, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{2^n}$. From Lemma 1, the maximum eigenvalue of G_V is more than $\lambda_{2^{n-1}+1}$.

What is $\lambda_{2^{n-1}+1}$? You will show in the assignment that the eigenvalues of Boolean hypercube has very simple structure. It has eigenvalue $-n + 2k$ with multiplicity $\binom{n}{k}$.

Exercise 10. What bound will this give on $\|G_V\|$ when $|V| > 2^{n-1}$?

Unfortunately the interlacing theorem applied on H_n doesn't seem to be of much help. It turns out, a small modification of the adjacency matrix of H_n will do the trick. The idea is to introduce a signing of the Boolean hypercube (assign -1 to some of the 1 entries of the matrix), that *compresses* the eigenvalues. In particular, we will try to make half the eigenvalues $\sqrt{n}$ and other half to be $-\sqrt{n}$. Applying interlacing theorem on that signed matrix will give the result.

Proof of Theorem 1. The main idea of the proof is to construct a *signing* of the adjacency matrix of the Boolean hypercube. A signing of a $\{0,1\}$ matrix is assigning negative sign to some non-zero entries of the matrix. Let A_s be a signing of a $\{0,1\}$ matrix, then you will show in the assignment

$$\|A\| \geq \|A_s\|.$$

We will construct a signing s of Boolean hypercube such that half of its eigenvalues (2^{n-1} of them) will be $\sqrt{n}$ and the other half will be $-\sqrt{n}$. If A is the adjacency matrix of G_V,

$$\|A_V\| \geq \|(A_s)_V\|.$$

Here A is the adjacency matrix of H_n and A_V denote the induced matrix on the subset V.

By Lemma 1, $\|(A_s)_V\|$ should be greater than the $2^{n-1}+1$ highest eigenvalue of A_s, which is $\sqrt{n}$.

The only task is to construct the signing with required properties. Notice that we want to have half the eigenvalues $\sqrt{n}$ and other half to be $-\sqrt{n}$. This implies that we want,

$$A_S^2 = nI.$$

(The trace being 0 ensures that there are equal number of $\sqrt{n}$ and $-\sqrt{n}$ eigenvalues).

Looking at every non-diagonal entry of A_S^2, it basically arises from two sums from a 4-cycle in the Boolean hypercube. If (x,y) are at Hamming distance 2 differing at i,j,

$$A_S^2(x,y) = A_S^2(x,x^{\oplus i})A_S^2(y,y^{\oplus j}) + A_S^2(x,x^{\oplus j})A_S^2(y,y^{\oplus i}).$$

This can only be 0, if all 4-cycles have odd number of -1's.

Exercise 11. Convince yourself of the previous statement. Also check that diagonal entries are fine.

In other words, a signing with odd number of -1's will finish the proof. Such a signing is trivial for $n=2$.

Exercise 12. Can you construct such a signing for $n=3$?

Formally, the signing can be defined inductively by,

$$(A_1)_s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, (A_n)_s = \begin{pmatrix} A_{n-1} & I \\ I & -A_{n-1} \end{pmatrix}$$

You can easily show the following properties of this signing by induction.

- $(A_n)_s$ is a signing of H_n (it follows the structure of Boolean hypercube).
- Trace of $(A_n)_s$ is 0.
- $(A_n)_s^2 = nI$.

From the third property, each eigenvalue is either $\sqrt{n}$ or $-\sqrt{n}$. From the trace property, the multiplicity of each eigenvalue is 2^{n-1}. Thus, we have the signing with required property, showing that if V is a subset of vertices of H_n such that $|V| > 2^{n-1}$, then $\|G_V\| \geq \sqrt{n}$.

By the discussion before the proof, this implies that $\lambda(f) \geq \sqrt{n}$ for any f with degree n. □

Huang's result, using the already known relationship between block sensitivity and degree [4], implies that $\mathrm{bs}(f) = O(\mathrm{s}(f)^4)$. We only know a function for which $\mathrm{bs}(f) = \Omega(\mathrm{s}(f)^2)$ [5]. It is an open problem to bridge this gap.

There have been interesting developments after this discovery, as mentioned before, it was proven that $\lambda(f) = O(\widetilde{\deg}(f))$ in [1]. They were able to use this to show that for any Boolean function f, $\deg(f) = O(\widetilde{\deg}(f)^2)$. This is known to be optimal by OR function.

2 Certificate games

Certificate games provide a nice algorithmic way to look at many of the combinatorial measures. They can be used to model sensitivity, spectral sensitivity, fractional block sensitivity (not studied in this course). For details, look at [2].

A certificate game for a Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$ is played by two cooperating players, say Alice and Bob, against a verifier. The players *can't communicate*, but can decide on a strategy beforehand. Once the game starts (players can't contact each other), the verifier sends them two inputs, $x \in f^{-1}(0)$ to Alice and $y \in f^{-1}(1)$ to Bob. Alice replies with an index $i \in [n]$ and Bob with an index $j \in [n]$. They win when $i = j$ and $x_i \neq y_i$. The inverse of the worst case winning probability (over all x, y pairs) is called the CG value of the function (for the best protocol).

Exercise 13. Suppose Alice and Bob don't use any randomness. Show that they can't win on every pair of input for Parity.

The game can be studied in multiple settings. The players can use randomness at their own end, though these random bits are not known to other party (private randomness). Another model is when they can have shared randomness (public randomness model). In another setting, we can assume that they can use quantum entanglement.

We will restrict ourselves to one of the simplest setting, private randomness and also assume that the question pair x, y is at Hamming distance 1. We will show that under such restrictions, for any Boolean function f, $CG(f) = \lambda(f)^2$. For other such results, please see [2].

Before we look at the relationship of spectral sensitivity with certificate games (the particular setting we have in mind). Let us first explore these quantities independently.

The private randomness certificate game strategy for a function f is characterized by the probability distribution $\{p_{x,i}\}_i$ for every input x. Notice that every input, depending upon whether it is a 1 or a 0-input, can only be sent to one of the party, either Alice or Bob. So, a strategy is a *collection of probability distributions on indices for each input.*

Exercise 14. What is the probability of winning on input pair x, y?

Notice that since we are in the Hamming distance 1 case, the input pair x, y differ at a single bit $i_{x,y}$. So the winning probability is $p_{x,i_{x,y}} p_{y,i_{x,y}}$, because they answer with these probabilities independently. The winning probability of the strategy is the worst case success probability, $\min_{(x,y)} p_{x,i_{x,y}} p_{y,i_{x,y}}$, where (x, y) have Hamming distance 1. The CG value of the game is the inverse of the maximum winning probability over all strategies.

Exercise 15. Show that the CG value of the game is at least the sensitvity of the function.

On the other hand, $\lambda(f)$ is the maximum eigenvalue of the sensitivity graph (precisely, the adjacency matrix of this graph). Notice that every row of this matrix has at most n ones, where these ones can only be in row x if the column y is at Hamming distance 1 from x. Remember that the y which is at Hamming distance one from x and differs on index i is represented as x^i.

We are ready to prove the main result of this section.

Theorem 2. *For any Boolean function $f : \{0,1\}^n \rightarrow \{0,1\}$, $CG(f) = \lambda(f)^2$.*

Note 1. There is no order notation involved here, the equality is exact. Also, the theorem follows from other equalities [2,1], we give a direct proof here.

Proof. The proof has two parts. It is easier to show that $CG(f) \leq \lambda(f)^2$. We need to come up with a strategy which wins with probability $\lambda(f)^2$. Let v be the principal eigenvector of the sensitivity graph of f. Since all entries of adjacency matrix are non-negative, by Frobenius-Perron theorem, all entries of v are non-negative.

Exercise 16. Suppose M is a matrix with all entries being non-negative. Show that there exists a principal eigenvector which is non-negative entrywise.

Given a non-negative v, it is easy to design the strategy,

$$p_{x,i} = \frac{v_{x^i}}{\sum_j v_{x^j}}.$$

Exercise 17. Show that the winning probability of this strategy is $\frac{1}{\lambda(f)^2}$.

Coming to the other part, we need to show that $CG(f) \geq \lambda(f)^2$. Let G_f denote the adjacency matrix of the sensitivity graph of f.

Exercise 18. Show that $CG(f) \geq \lambda(f)^2$ iff the matrix $I - \frac{1}{\sqrt{CG(f)}} G_f$ is positive semidefinite. Hint: use spectral decomposition of symmetric matrices.

Note 2. An $n \times n$ symmetric matrix A is positive semidefinite if $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$. Another equivalent way is to show that all eigenvalues are non-negative.

Let $P := I - \frac{1}{\sqrt{CG(f)}} G_f$. We will show that P is the sum of n positive semidefinite matrices, $P = P_1 + P_2 + \cdots + P_n$. Why is that enough?

The matrix P_i corresponds to the i-th index (as you might have guessed). Look at all pairs x, y such that they differ at index i, are at Hamming distance 1, and $f(x) \neq f(y)$ (notice that these pairs are disjoint). Let $\{p_{x,i}\}_i$ be the strategy which wins with probability $\frac{1}{CG(f)}$. Set,

$$P_i(x,x) = p_{x,i}, P_i(y,y) = p_{y,i}, P_i(x,y) = P_i(y,x) = \frac{1}{\sqrt{CG(f)}}.$$

Exercise 19. Show that this 2×2 matrix is positive semidefinite.

Since P_i can be viewed as a block diagonal matrix with these 2×2 submatrices as diagonals (after rearranging columns and rows), it is positive semidefinite.

Exercise 20. What properties of positive semidefinite matrices do you need to show the previous statement?

Looking at $\sum_i P_i$, diagonal entries sum up to 1 because $p_{x,i}$'s form a probability distribution. Also, a non-diagonal entry x, y is non-zero if and only if x, y are at Hamming distance 1 and $f(x) \neq f(y)$. Actually, the entry will be exactly $\frac{1}{\sqrt{CG(f)}}$ contributed by the i where x, y differ.

So, we showed that P is the sum of positive semidefinite matrices and hence it is itself positive semidefinite. This proves the remaining part, $CG(f) \geq \lambda(f)^2$. □

3 Assignment

Exercise 21. Suppose A is the adjacency matrix of a bipartite graph. Show that if u is an eigenvalue of A, then so is $-u$.

Exercise 22. Why is the λ of restricted function smaller than the λ of the original function?

Exercise 23. Let H_n be the Boolean hypercube on n elements. Show that H_n has eigenvalue $-n + 2k$ with multiplicity $\binom{n}{k}$ for $0 \leq k \leq n$.

Hint: Use induction and structure of the adjacency matrix of Boolean hypercube.

Exercise 24. Just by looking at the eigenvalues of Boolean hypercube and Cauchy's interlacing theorem, you can come up with a statement like: if the degree n coefficient of f is *big enough* then $\lambda(f) \geq \sqrt{n}$. Make this statement precise and prove it.

Exercise 25. Show that if A_s is a signing of A, then

$$\|A\| \geq \|A_s\|.$$

Exercise 26. What is the spectral sensitivity of PARITY?

Exercise 27. Show that CG value of parity is $\Theta(n^2)$.

Exercise 28. Show that $CG(f) \leq \mathrm{s}(f)^2$.

References

1. S. Aaronson, S. Ben-David, R. Kothari, S. Rao, and A. Tal. Degree vs. approximate degree and quantum implications of huang's sensitivity theorem. *STOC 2021*, 2021.
2. Sourav Chakraborty, Anna Gal, Sophie Laplante, Rajat Mittal, and Anupa Sunny. Certificate games. *ITCS 2023*, 2023.
3. H. Huang. Induced subgraphs of hypercubes and a proof of the sensitivity conjecture. *Annals of Mathematics, Volume 190, Pages 949-955*, 2019.
4. N. Nisan and M. Szegedy. On the degree of boolean functions as real polynomials. *Computational Complexity, volume 4, pages 301–313*, 1994.
5. D. Rubinstein. Sensitivity vs. block sensitivity of boolean functions. *Combinatorica, Volume 15, Pages 297–299*, 1995.