Notes for Lecture 16

Today we begin the proof of the PCP theorem.

Theorem 1 $\mathbf{NP} \subseteq \mathbf{PCP}_{c=1,s=\frac{1}{2}}(O(\log(n)), O(1))$

To do this we will construct instances of constraint satisfaction problems (CSPs) for which it is hard to distinguish the case in which the CSP is satisfiable from the case in which every assignment contradicts a constant fraction of constraints.

We will work with the type of CSPs where each constraint has two variables, but where each variable can take on a non-boolean (but constant-size) range of values.

Definition 1 Max-2-CSP-Σ
Input: variables $x_1, \ldots, x_n$ that range over Σ, a collection of binary constraints.
Goal: find an assignment that maximizes that number of satisfied constaints.

Definition 2 *If $\mathcal{C}$ is a CSP, we call opt($\mathcal{C}$) the fraction of constraints which are satisfied by the optimal assignment.*

The following is the version of the PCP Theorem that we will prove.

Theorem 2 *There exists a Σ_0, a polynomial time reduction R, and a $\delta_0 > 0$ such that*

- *R is a reduction from 3-coloring to Max-2-CSP-Σ_0.*
- *If G is 3-colorable, then $R(G)$ is satisfiable.*
- *If G is not 3-colorable, then opt($\mathcal{C}$) $\leq 1 - \delta_0$.*

This theorem implies the PCP theorem because given a graph G, we can define a valid proof to be a binary encoding of a solution to the constraint satisfaction problem $R(G)$. Given an alleged proof, the verifier randomly picks $O(\frac{1}{\delta_0})$ constraints to check, reads an assignment for the variables in such constraints from the proof, and accepts if and only if all constraints are satisfied.

The verifier uses $O(\log(n))$ random bits and reads $O(\frac{1}{\delta_0} \log |\Sigma_0|)$ bits of the proof. (We assume that the assignment to the n variables is encoded as a string of $n \log |\Sigma_0|$ bits.) If R works as in the theorem statement, then if G is three colorable, the CSP is satisfiable and there exists a valid proof that is accepted with probability 1. Furthermore, if G is not three colorable, then, for every alleged proof, a δ_0 fraction of the constraints in $R(G)$ will not be satisfied. Therefore, with probability at least $\frac{1}{2}$ the verifier will choose a constraints that is not satisfied, and thus reject.

Observe that 2-CSP-$\{a, b, c\}$ is at least as hard is 3-coloring because 3-coloring can be set up as a 2-CSP over a three-element range. We see from the theorem statement that

- $\text{opt}(G) = 1 \Rightarrow opt(R(G)) = 1$.
- $\text{opt}(G) \leq 1 - \frac{1}{|E|} \Rightarrow opt(R(G)) \leq 1 - \delta_0$.

The idea will be to create R by amplifying the fraction of unsatisfied constraints by a constant factor while only increasing the number of constraints by a linear amount and applying this amplification a logarithmic number of times. We can restate the theorem as follows:

Theorem 3 (restated) *There is δ_0, Σ_0, $|\Sigma_0| \geq 3$, and polynomial time R mapping inputs of Max-2-CSP-Σ_0 to Max-2-CSP-Σ_0 such that*

1. *# of constraints of $R(\mathcal{C}) = O(\#$of constraints of $R(\mathcal{C}))$.*
2. *$opt(\mathcal{C}) = 1 \Rightarrow opt(R(G)) = 1$.*
3. *$opt(\mathcal{C}) \leq 1 - \delta \Rightarrow opt(R(\mathcal{C})) \leq 1 - 2\delta$ if $\delta < \delta_0$.*

We prove this theorem using two lemmas. The first lemma will amplify the number of unsatisfiable constraints, but will also increase the range size. The second lemma will reduce the range size, but will decrease the number of unsatisfiable constraints.

Lemma 4 (Amplification) *$\forall \Sigma_0$, $\forall c$, there exists Σ and a poly-time R_1, mapping Max-2-CSP-Σ_0 to Max-2-CSP-Σ such that R satisfies 1) and 2) in Theorem 3 and $opt(\mathcal{C}) \leq 1 - \delta \Rightarrow opt(R_1(\mathcal{C})) \leq 1 - c\delta$ provided that $c \leq \delta_0$.*

Lemma 5 (Range Reduction) *$\exists \Sigma_0$, $\exists c_0$, such that for all Σ, there exists a poly-time R_2, mapping Max-2-CSP-Σ to Max-2-CSP-Σ_0 such that R satisfies 1) and 2) in Theorem 3 and $opt(\mathcal{C}) \leq 1 - \delta \Rightarrow opt(R_2(\mathcal{C})) \leq 1 - \delta/c_0$.*

To get the theorem from these two lemmas, let $c = 2c_0$ in Lemma 4, then the composition $R_2(R_1(\cdot))$ solves the theorem because:

$$opt(\mathcal{C}) \leq 1 - \delta \Rightarrow opt(R_1(\mathcal{C})) \leq 1 - c\delta = 1 - 2c_0\delta \Rightarrow opt(R_2(R_1(\mathcal{C}))) \leq 1 - 2\delta$$

We conclude today's lecture with a preliminary result that will be helpful in the proof of Lemma 4, by showing that without loss of generality we can work with instances of Max-2-CSP-Σ whose constraint graph is a bounded-degree expander.

If C is an instance of Max-2-CSP-Σ, its *constraint graph* is a multi-graph $G_C = (V, E)$ that has one vertex for every variable of C, and that has one edge for every constraint of C, joining the vertices corresponding to the two variables that appear in the constraint.

First we convert the graph to a bounded degree graph. We do this in much the same way as we reduced 3SAT to 3SAT where each variable occurs at most some constant number of times.

Let $\mathcal{C}$ be a set of constraints for a Max-2-CSP-Σ over variables $x_1, \ldots, x_n$ where x_i occurs m_i times. For every i introduce variables $y_i^1, \ldots, y_i^{m_i}$ and construct a k-regular graph G_i with m_i vertices of edge expansion at least 1 (note that k is a constant). Now construct a new Max-2-CSP-Σ $\mathcal{C}'$ over the y_i^j variables as follows:

- For each constraint $f(x_i, x_j)$ in $\mathcal{C}$ where f is the ath occurrence of x_i and the bth occurrence of x_j create a new constraint $f(y_i^a, y_j^b)$.
- For every i, for every edge $(a, b) \in G_i$ create a constraint $y_i^a = y_i^b$.

The # of constraints in $\mathcal{C} = \frac{1}{2}\sum_i m_i$

The # of constraints in $\mathcal{C}' = \frac{1}{2}\sum_i m_i + \sum_i \frac{km_i}{2} = \frac{k+1}{2}\sum_i m_i = O($# of constraints in $\mathcal{C})$.

As we saw with the 3SAT reduction, the minimum number of constraints violated in $\mathcal{C}'$ is the same as the minimum number of constraints violated in $\mathcal{C}$.

Thus we obtain a regular degree-d graph (d = k + 1).

Now we would like to, by adding vacuous constraints, make this graph into an expander.

Claim 6 *For $i \in \{1, 2\}$, let $G_i = (V, E_i)$ be a degree d-regular graph with adjacency matrix M_i. Let G be the $2d$-regular graph obtained by taking the disjoint union of the edges of G_1 and the edges of G_2, so that the transition matrix M of G satisfies $M = \frac{1}{2}M_1 + \frac{1}{2}M_2$.*

Then $\bar{\lambda}_2(G) \leq \frac{1}{2} + \frac{1}{2}\bar{\lambda}_2(G_2)$.

Proof:

$$\bar{\lambda}_2(G) = \max_{x \perp (1,\ldots,1)} \frac{|xMx^T|}{xx^T} \leq \max_{x \perp (1,\ldots,1)} \frac{\frac{1}{2}|xM_1x^T| + \frac{1}{2}|xM_2x^T|}{xx^T} \leq \frac{1}{2} + \frac{1}{2}\bar{\lambda}_2(G_2) \ .$$

□

Let $\mathcal{C}$ be a CSP with a d-regular constraint graph. Let $\mathcal{C}_{EXP}$ be a CSP with d-regular λ-expanding constraint graph and constraints that are always trivially satisfied. Then $\mathcal{C} + \mathcal{C}_{EXP}$ has a constraint graph which is $2d$-regular and is $\left(\frac{1}{2} + \frac{1}{2}\lambda\right)$-expanding.

Furthermore, if $\mathcal{C}$ is satisfiable, then $\mathcal{C} + \mathcal{C}_{EXP}$ is also satisfiable. If $\mathrm{opt}(\mathcal{C}) \leq 1 - \delta$ then $\mathrm{opt}(\mathcal{C} + \mathcal{C}_{EXP}) \leq 1 - \delta/2$.