U.C. Berkeley
CS294: PCP and Hardness of Approximation
Professor Luca Trevisan

Handout N13
March 8, 2006
Scribe: Grant Schoenebeck

Notes for Lecture 13

Today we begin our proof of the PCP theorem in earnest.

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 CSPs for which it is hard to distinguish if the CSP is satisfiable or not close to satisfiable. We will look at CSPs where each constraint has two variables, but where each variable can take on more than 2 values.

Definition 2 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 3 *If $\mathcal{C}$ is a CSP, we call opt($\mathcal{C}$) the fraction of constraints which are satisfied by the optimal assignment.*

Theorem 4 *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-$\Sigma 0$.*
- *G is colorable implies $R(G)$ is satisfiable.*
- *G is not colorable implies that 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 an solution to the Max-2-CSP-Σ $R(G)$. Then we can construct a verifier that randomly picks $O(\frac{1}{\delta_0})$ constraints to check, and if they are all satisfied, the verifier accepts, otherwise, the verifier rejects.

It is easy to see that the verifier uses $O(log(n))$ random bits and reads $O(\frac{1}{\delta_0}\log|\Sigma_0|)$ bits of the proof. If R works as in the theorem statement, then if G is three colorable, the CSP is satisfiable and there exists a valid proof. Furthermore, if G is not three colorable, then a δ_0 fraction of the constraints will not be satisfied. Therefore, with probability at least $\frac{1}{2}$ the verifier will choose a constraints that is not satisfied.

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 character alphabet. 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 5 (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 $\delta < \delta_0$.*

Looking through the three requirements, it is fairly straightforward to verify that each of them is required. In particular, because we are going to apply this theorem a logarithmic number of times to obtain the previous theorem, we cannot increase the number of constraints by more than a linear amount; otherwise, we will end up with a super-polynomial number of constraints.

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

Lemma 6 (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* **??** *and $opt(\mathcal{C}) \leq 1-\delta \Rightarrow opt(R_1(\mathcal{C})) \leq 1 - c\delta$ provided that $c \leq \delta_0$.*

Lemma 7 (Alphabet 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* **??** *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 **??**, 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 can use a constraint graph to visualize the the variables involved in a Max-2-CSP-Σ. Each vertex of the graph is a variable, and each edge is a constraint which is incident to the two vertices which represent the variables in the constraint. If $\mathcal{C}$ is a collection of m constraints over the variables: $x_1, \ldots, x_n$, then we get a graph with n vertices and m edges. We will want this graph to be an expander in order to prove Lemma 1. The rest of the lecture will illustrate how to massage this graph into an expander.

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).

Example 1

$$\begin{aligned} \Sigma &= \{0,1,2,3,4\} \\ x_1 &\neq x_2 \\ x_2 &\neq x_3 \\ x_3 &\neq x_4 \\ x_2 - x_1 \equiv 2 &\mod 5 \\ x_3 + x_4 \equiv 1 &\mod 5 \\ x_1 + x_4 \equiv 4 &\mod 5 \\ x_1 - x_4 \equiv 3 &\mod 5 \end{aligned}$$

You can see that x_i occurs $4, 3, 3, 4$ times for $i = 1, 2, 3, 4$ respectively. So, for example, we create new variables for x_1: $y_1^1, y_1^2, y_1^3, y_1^4$. For our constant degree expander graphs, we can just use the 3-cycle for x_2 and x_3, and we can use the 4-cycle for x_1 and x_4.

Using the reduction we get the following CSP:

$$
\begin{aligned}
\Sigma &= \{0,1,2,3,4\} \\
y_1^1 &\neq y_2^1 \\
y_2^2 &\neq y_3^1 \\
y_3^2 &\neq y_4^1 \\
y_2^3 - y_1^2 &\equiv 2 \mod 5 \\
y_3^3 + y_4^2 &\equiv 1 \mod 5 \\
y_1^3 + y_4^3 &\equiv 4 \mod 5 \\
y_1^4 - y_4^4 &\equiv 3 \mod 5 \\
y_1^1 &= y_1^2 \\
y_1^2 &= y_1^3 \\
y_1^3 &= y_1^4 \\
y_1^4 &= y_1^1 \\
&\vdots \\
y_4^1 &= y_4^2 \\
y_4^2 &= y_4^3 \\
y_4^3 &= y_4^4 \\
y_4^4 &= y_4^1
\end{aligned}
$$

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

Claim 8 *For $i \in \{1,2\}$, let $G_i = (V, E_i)$ be a degree d-regular graph with adjacency matrix M_i. Let $d = |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_n|$ be the eigenvalues of M_2. Let $M' = M_1 + M_2$ be the adjacency matrix for the corresponding $2d$-regular graph, and let $2d = |\lambda'_1| \geq |\lambda'_2| \geq \cdots \geq |\lambda'_n|$ be the eigenvalues of M'. Then, for any λ, if $|\lambda_2| \leq \lambda$, $|\lambda'_2| \leq d + |\lambda_2|$.*

PROOF:

$$
|\lambda'_2| = \max_{x \perp (1,\ldots,1)} \frac{|xM'x^\top|}{xx^\top} = \max_{x \perp (1,\ldots,1)} \frac{|xM_1x^\top + xM_2x^\top|}{xx^\top} = \max_{x \perp (1,\ldots,1)} \frac{|xM_1x^\top|}{xx^\top} \max_{x \perp (1,\ldots,1)} \frac{|xM_2x^\top|}{xx^\top} \leq d + |\lambda_2|
$$

□

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 $d + \lambda$-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 $\mathrm{opt}(\mathcal{C} + \mathcal{C}_{EXP}) \leq 1 - \delta/2$.