

Lecture 5

Lecturer: Madhu Sudan

Scribe: Vincenzo Bonifaci

Today we will prove that systems of affine equations admit PCPs if we allow the verifier to use a polynomial number of random bits. We will take a path that will help us later when we will have to show PCPs for systems of quadratic equations.

We introduce some useful notation. If a_1, \dots, a_n are n bits, we denote by $E(a_1, \dots, a_n)$ the 2^n -bit string giving, for every linear function of n variables l , the value of l on (a_1, \dots, a_n) :

$$E(a_1, \dots, a_n)[l] = l(a_1, \dots, a_n).$$

We saw yesterday that $\delta(E(a_1, \dots, a_n), E(a'_1, \dots, a'_n)) \geq \frac{1}{2}$ if $(a_1, \dots, a_n) \neq (a'_1, \dots, a'_n)$.

Suppose we are given a polynomial P_1 that is linear and homogenous. We want to verify if π is a proof that P_1 has a zero (obviously, a linear homogenous function always has a zero but this is not true for affine or quadratic equations). The verifier will operate in two steps:

1. Verify that π is close to $E(a_1, \dots, a_n)$ for some a_1, \dots, a_n .
2. Verify that $P_1(a_1, \dots, a_n) = 0$.

Now let us consider the different cases that can occur and what the verifier should do in order to be a $\text{PCP}_{1,s}[\text{poly}(n), O(1)]$ verifier. Let p_0 be any positive constant less than $4/9$ (this fact will be needed later).

- Case 1. For some a_1, \dots, a_n , $\pi = E(a_1, \dots, a_n)$ and $P_1(a_1, \dots, a_n) = 0$. In this case the verifier should accept with probability 1, since the proof is in the expected format and correct.
- Case 2. For some a_1, \dots, a_n , $\delta(\pi, E(a_1, \dots, a_n)) \leq p_0$, but $P_1(a_1, \dots, a_n) = 1$. The proof is near to the expected format but wrong. The verifier should reject in step 2 with probability at least some constant, say p_2 .
- Case 3. For every a_1, \dots, a_n , $\delta(\pi, E(a_1, \dots, a_n)) > p_0$. The proof is not in the expected format. The verifier should reject in step 1 with probability at least some constant, say p_1 .
- Case 4. For some a_1, \dots, a_n , $\delta(\pi, E(a_1, \dots, a_n)) \leq p_0$ and $P_1(a_1, \dots, a_n) = 0$, but $\pi \neq E(a_1, \dots, a_n)$. In this case the verifier can do anything: it could accept because the polynomial has a zero, but it could also reject because the proof is not in the expected format. This is what allows PCP to work.

With these parameters, we will have a soundness s of $1 - \min\{p_1, p_2\}$. Notice that it can be improved to an arbitrary constant by running the verifier a constant number of times; the completeness remains perfect.

Now consider step 2 of the verifier V . It has to check if $P_1(a_1, \dots, a_n) = 0$ without knowing a_1, \dots, a_n . In the lucky case that $\pi = E(a_1, \dots, a_n)$, there is a simple way to do this: query $\pi[P_1]$ and accept iff it is zero. However, this does not work if π is even slightly corrupted. But we can exploit the linearity of the functions represented in π in the following way.

Step 2: Pick linear function l_1 uniformly at random. Accept iff $\pi[P_1 + l_1] - \pi[l_1] = 0$.

Now if we suppose not to be in case 3 (which will be handled in step 1), $\pi[l_1] \neq l_1(a_1, \dots, a_n)$ with probability at most p_0 . Also $\Pr_{l_1}[\pi[l_1 + P_1] \neq (l_1 + P_1)(a_1, \dots, a_n)]$ is at most p_0 . Notice that these two events are not independent; however the probability that either of them occurs is at most the union bound:

$$\Pr_{l_1}[\pi[l_1] \neq l_1(a_1, \dots, a_n) \vee \pi[l_1 + P_1] \neq (l_1 + P_1)(a_1, \dots, a_n)] \leq 2p_0.$$

Thus with probability at least $1 - 2p_0$

$$\pi[l_1 + P_1] - \pi[l_1] = (l_1 + P_1)(a_1, \dots, a_n) - l_1(a_1, \dots, a_n) = P_1(a_1, \dots, a_n)$$

which means that in case 2 we will reject with probability at least $p_2 := 1 - 2p_0$. Notice how we were able to simulate one query to $E(a_1, \dots, a_n)$ with two queries to π .

Let's consider step 1 now. We can use a similar idea to check that π is near to the expected format.

Step 1: Pick linear functions l_1, l_2 uniformly at random. Reject if $\pi[l_1] + \pi[l_2] \neq \pi[l_1 + l_2]$.

In case 1, when the proof is correct, everything is working. For case 3 we have to bound

$$\Pr_{l_1, l_2}[\pi[l_1] + \pi[l_2] \neq \pi[l_1 + l_2]].$$

A result of Blum, Luby and Rubinfeld [?] is going to help us.

Theorem 1 (BLR) *If $\epsilon < 2/9$ and*

$$\Pr_{l_1, l_2}[\pi[l_1] + \pi[l_2] \neq \pi[l_1 + l_2]] \leq \epsilon,$$

then there are a_1, \dots, a_n s.t.

$$\delta(\pi, E(a_1, \dots, a_n)) \leq 2\epsilon.$$

The same theorem can be generalized to groups.

Theorem 2 *For every pair of finite groups G and H , if $f : G \rightarrow H$ is such that*

$$\Pr_{x, y \in G}[f(x) \cdot f(y) \neq f(x \cdot y)] \leq \epsilon,$$

then there is a homomorphism¹ $\phi : G \rightarrow H$ such that

$$\Pr_{x \in G}[f(x) \neq \phi(x)] \leq 2\epsilon$$

provided that $\epsilon < 2/9$.

This result is also tight as shown by Ben-Or et al. [?].

Theorem 3 *Let $G = \mathbb{Z}_{3^m}^+$, $H = \mathbb{Z}_{3^{m-1}}^+$ and*

$$f(x) = \begin{cases} 1 & \text{if } x \equiv 1 \pmod{3} \\ 0 & \text{if } x \equiv 0 \pmod{3} \\ -1 & \text{if } x \equiv -1 \pmod{3}. \end{cases}$$

Then $\Pr_{x, y}[f(x) \cdot f(y) \neq f(x \cdot y)] = 2/9$ and

$$\delta(f, \phi) > 1 - \frac{1}{3^{m-1}}.$$

We can now end the analysis of case 3. Suppose by contradiction the verifier is rejecting with probability less than $p_1 := p_0/2$ and let $\epsilon = p_0/2$. By BLR, we have $\delta(\pi, E(a_1, \dots, a_n)) \leq p_0$, but this is inconsistent with case 3.

Thus we ended our analysis for one linear equation. Remember that eventually we want to handle a system of m quadratic equations. In the rest of this lecture we show how to move from one linear equation to a system of m affine equations.

¹ ϕ is a homomorphism if $\forall x, y \in G, \phi(x) \cdot \phi(y) = \phi(x \cdot y)$.

An affine equation has the form $\sum_i c_i x_i + c_0 = 0$. In order to handle a single affine equation we can simply change step 2 so that it checks if $\pi[l_1 + P_1] - \pi[l_1] = c_0$.

To verify a system of m equations we cannot simply repeat m times the second step, because querying $O(m)$ bits of the proof is too expensive, so we use a different idea. Define v_j to be 0 if the j -th equation is satisfied and 1 otherwise. We want to check if the polynomial $T(z) = \sum_{j=1}^m v_j z_j$ is identically zero.

Without loss of generality let the j -th equation be $P_j(x_1, \dots, x_n) = c_0^{(j)}$ where P_j is linear and homogenous. Then $\sum_j v_j z_j = 0$ iff

$$\sum_j z_j P_j = \sum_j z_j c_0^{(j)}.$$

The verifier now performs step 2 as follows: it picks a linear function l_1 uniformly at random, m values $\alpha_1, \dots, \alpha_m$ uniformly at random from $\{0, 1\}$, and accepts iff

$$\pi\left[\sum_j \alpha_j P_j + l_1\right] - \pi[l_1] = \sum_j \alpha_j c_0^{(j)}.$$

In cases 1 and 3 everything is still working fine. For case 2, some equation is not satisfied; that is, $\sum_j v_j z_j$ is not identically zero. Then

$$\Pr_{\alpha_1, \dots, \alpha_m \in \{0, 1\}} \left[\sum_j \alpha_j P_j(a_1, \dots, a_n) = \sum_j \alpha_j c_0^{(j)} \right] \leq \frac{1}{2}$$

since if a polynomial $Q(x_1, \dots, x_n)$ is not identically zero, $\Pr_{x_1, \dots, x_n \in S} [Q(x_1, \dots, x_n) = 0]$ is at most $\frac{\deg Q}{|S|}$. Thus the probability that V does not reject is at most $\frac{1}{2} + 2p_0$, and V will reject with probability at least $p_2 := \frac{1}{2} - 2p_0$.

So let us recap what the verifier does.

Step 1: Pick linear functions l_1, l_2 uniformly at random. Reject if

$$\pi[l_1] + \pi[l_2] \neq \pi[l_1 + l_2].$$

Step 2: Pick $\alpha_1, \dots, \alpha_m \in \{0, 1\}$ independently and uniformly at random. Define $\bar{P} = \sum_j \alpha_j P_j$ and $\bar{c} = \sum_j \alpha_j c_0^{(j)}$. Pick linear function l_1 uniformly at random. Accept iff

$$\pi[\bar{P} + l_1] - \pi[l_1] = \bar{c}.$$