

# PCP - Huji Course, Ex 2.

David Ponnarovsky

July 27, 2023

## 1 Ex 1. Sumchecking with coefficients.

We would like to verify that a given polynomial box  $P$  satisfies that  $\sum_{x \in [d]^m} \varphi(x) f_P(x) = 0$  by accessing to at most  $O(md)$  variables. For any function  $\varphi : [d]^m \rightarrow \mathbb{F}_q$ . Denote by  $\varphi' : \mathbb{F}_q^m \rightarrow \mathbb{F}_q$  the extension of  $\varphi$  into a polynomial over  $\mathbb{F}_q^m$ . We saw in that lectures (and also in the previous assignment) that there is such a unique extension.

We are going to split the section into three, first we are going to show how to verify that  $\sum_{x \in [d]^m} f_P(x) = 0$ . When the polynomial is a function into  $\mathbb{F}_q$ . (I think, but not sure, that in the lecture we saw only the case when  $q = 2$ ). Then in the second part we will show how can one reduce the coefficients case into the non-coefficients case. Finally, in the last part, we combine all together to show that the construction achieves the requirements.

### 1.1 Over non binary field.

Let's define a series of polynomial boxes  $f_i$  such that:

$$f_0 = f$$
$$f_{i+1}(x_1, \dots, x_{m-i}) = \sum_{y \in [d]} f_i(x_1, \dots, x_{m-i}, x_{m-i+1} = y)$$

Our verifier will ask for a proof which is a list of  $f_0, f_1, f_2, \dots, f_m$ . Now, notice that if  $f$  is an honest assignment then  $f_m$  is just the summation of  $f$  over the cube  $[d]^m$ . So it is sufficient to show the existence of a verifier that rejects with high probability any string far from being encoded by the previous structure.

- 1 Sample uniformly random  $i \sim [m]$  and check that  $f_i$  is a codeword of the polynomial code in  $m$  variables at degree at most  $m \cdot d$ .
- 2  $r_1, r_2, \dots, r_m \leftarrow$  sample uniformly  $m$  points of  $[d]$
- 3 **for**  $i \in [1, m]$  **do**
- 4     Check if  $f_{i+1}(r_1, \dots, r_{m-i-1}, x_{m-i}) - \sum_{y \in [d]} f_i(r_1, r_2, \dots, r_{m-i-1}, x_{m-i}, y)$  is the zero polynomial by a random test that uses at most single query. (Here  $x_{m-i} \in [d]$  is the only variable)
- 5
- 6     If not then reject.
- 7 **end**
- 8 Accept if  $f_0 = 0$

*Proof.* For convenience let's denote by  $g_i(x_{m-i})$  the difference that has been queried in line number 3.

1. Correctness. Easy. If the assignment is honest then by definition  $g_i = 0$  for any  $i \in [m]$  and therefore for any  $x_{m-i}$  we will have that  $g_i(x_{m-i}) = 0$ . So, in that case iteration will pass. And whole proof will be accepted with probability 1.
2. Soundness. Assume that there is any  $\deg f_{i+1} \leq \deg f_i$ , Thus asking

□

## 1.2 Coefficients $\mapsto$ non-coefficients.

Now as we proved in the classes  $\deg f \cdot g \leq \deg f + \deg g$ . Therefore we can redact the problem of verifying whether the weight summation is zero by considering the summation of the polynomial  $\varphi' \cdot f$  over the cube  $[d]^m$ .

- 1 Sample uniformly random  $x \sim [d]^m$  and check that  $\varphi'(x) = \varphi(x)$
- 2 Check that  $\varphi$  is a polynomial at degree at most  $d \cdot m$ .
- 3 If both of the checks passed, accept.

*Proof.*

□

## 1.3 Combine all.

- 1 Use the first tester to check the validity of the pair  $(\varphi', \varphi)$ .
- 2 Check that the degree of  $\xi$  is at most  $2md$
- 3 Check that the polynomial  $f \cdot \varphi' - \xi$  is the zero polynomial.
- 4 Using the first verifier, accept if the summation of  $\xi$  over the cube  $[d]^m$  is zero.

*Proof.*

□

## 2 Ex 2.

The question concerns with the following test: [\[COMMENT\] rewrite again.](#)

- 1 Choose  $x, y \in \{\pm 1\}^k$  independently.
- 2 Choose  $\mu \in \{\pm 1\}$ .
- 3 Choose a random noise  $z \in \{\pm 1\}^k$  such that  $z_i$  gets +1 with probability  $1 - \varepsilon$ .
- 4 Accept if  $\mu f(\mu x) \cdot g(y) = f(z \cdot xc^{-1}(y))$

### 2.1 2.a.

Let  $f = \chi_{\{i\}}, g = \chi_{\{j\}}$  and  $j = c(i)$ . In that case it holds that:

$$\begin{aligned} \mu f(\mu x) \cdot g(y) &= \mu \chi_{\{i\}}(\mu x) \chi_{\{j\}}(y) = \mu^2 x_i y_j = x_i y_j \\ f(z \cdot xc^{-1}(y)) &= \chi_{\{i\}}(zxc^{-1}(y)) = z_i x_i y_j \end{aligned}$$

Thus, the test pass only if  $z_i = 1$  and it given that this event happens with probability  $1 - \varepsilon$ .

### 2.2 2.b.

Denote by  $\alpha_I \in \mathbb{R}$  and  $\beta_I \in \mathbb{R}$  the coefficients of  $f, g$  over the character  $\chi_{\{I\}}$ .

$$\begin{aligned} &\mathbf{E} [\mu f(\mu x) \cdot g(y) f(z \cdot xc^{-1}(y))] \\ &= \sum_{I, J, K} \alpha_I \alpha_K \beta_J \mathbf{E} [\mu \chi_{\{I\}}(\mu x) \chi_{\{J\}}(y) \chi_{\{K\}}(zxc^{-1}(y))] \\ &= \sum_{I, J, K} \alpha_I \alpha_K \beta_J \mathbf{E} [\mathbf{E} [\mu \chi_{\{I\}}(\mu x) \chi_{\{J\}}(y) \chi_{\{K\}}(zxc^{-1}(y)) | \mu]] \\ &= \sum_{I, J, K} \alpha_I \alpha_K \beta_J \frac{1}{2} \left( (-1)^{|I|+1} + 1 \right) \mathbf{E} [\chi_{\{I\}}(x) \chi_{\{J\}}(y) \chi_{\{K\}}(zxc^{-1}(y))] \end{aligned}$$

Thus, all the elements in which  $|I|$  is even contribute zero for the exception. Now, let's apply the conditional expectation formula again conditioning over  $I, J, K, x, y$ :

$$\begin{aligned}
&= \sum_{I, J, K, |I| \text{ is odd}} \alpha_I \alpha_K \beta_J \mathbf{E} \left[ \mathbf{E} \left[ \chi_{\{I\}}(x) \chi_{\{J\}}(y) \chi_{\{K\}}(zc^{-1}(y)) \mid I, J, K \right] \right] \\
&= \sum_{I, J, K, |I| \text{ is odd}} \alpha_I \alpha_K \beta_J \mathbf{E} \left[ \sum_{\xi=0}^{|K|} \binom{|K|}{\xi} (-\varepsilon)^\xi (1-\varepsilon)^{|K|-\xi} \chi_{\{I\}}(x) \chi_{\{J\}}(y) \chi_{\{K\}}(xc^{-1}(y)) \right] \\
&= \sum_{I, J, K, |I| \text{ is odd}} \alpha_I \alpha_K \beta_J \mathbf{E} \left[ (1-2\varepsilon)^{|K|} \chi_{\{I\}}(x) \chi_{\{J\}}(y) \chi_{\{K\}}(xc^{-1}(y)) \right]
\end{aligned}$$

Let us denote by  $C^{-1}(K)$  the indices  $C^{-1}(K) = \{j : \exists i \in K, c(i) = j\}$ . Then we get that:

$$\chi_{\{K\}}(xc^{-1}(y)) = \prod_{i \in K} x_i y_{c_i} = \chi_{\{K\}}(K) \chi_{\{C^{-1}(K)\}}(y)$$

Recall that for any  $I, J \subset [n]$  it holds that:

$$\mathbf{E} [\chi_{\{I\}}(x) \chi_{\{J\}}(x)] = \mathbf{E} [\chi_{\{I \Delta J\}}(x)] = \mathbf{1}_{I=J}$$

And therefore the above can be simplified into:

$$\sum_{|I| \text{ is odd}} \alpha_I^2 \beta_{C^{-1}(I)} (1-2\varepsilon)^{|I|}$$

### 2.3 2.c

First let's bound from below the expectation by the given that  $f$  and  $g$  pass the test with probability at least  $\frac{1}{2} + \delta$ :

$$\begin{aligned}
&\mathbf{E} [\mu f(\mu x) \cdot g(y) f(z \cdot xc^{-1}(y))] \\
&= \mathbf{Pr} [\mu f(\mu x) \cdot g(y) = f(z \cdot xc^{-1}(y))] - \mathbf{Pr} [\mu f(\mu x) \cdot g(y) \neq f(z \cdot xc^{-1}(y))] \\
&\geq \frac{1}{2} + \delta - \left( \frac{1}{2} - \delta \right) = 2\delta
\end{aligned}$$

Thus in total the inequality of the above section becomes:

$$\sum_{|I| \text{ is odd}} \alpha_I^2 \beta_{C^{-1}(I)} (1-2\varepsilon)^{|I|} \geq 2\delta$$

Using Cauchy-Schwartz to bound from above, we obtain:

$$\begin{aligned}
4\delta^2 &\leq \left( \sum_{|I| \text{ is odd}} \alpha_I^2 \beta_{C^{-1}(I)} (1-2\varepsilon)^{|I|} \right)^2 \leq \sum_{|I| \text{ is odd}} \alpha_I^2 \cdot \sum_{|I| \text{ is odd}} \alpha_I^2 \beta_{C^{-1}(I)}^2 (1-2\varepsilon)^{2|I|} \\
&\leq \sum_{|I| \text{ is odd}} \alpha_I^2 \beta_{C^{-1}(I)}^2 (1-2\varepsilon)^{2|I|}
\end{aligned}$$

Now let's denote by  $\eta \in (0, 1)$  a threshold parameter and separate the above summation into two part, when the first part sums up the elements in which  $|I| \leq \eta n$  and the second sums elements in which  $|I| \geq \eta n$ :

$$\begin{aligned}
4\delta^2 &\leq \sum_{|I| \text{ is odd}, |I| \leq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2 (1-2\varepsilon)^{2|I|} + \sum_{|I| \text{ is odd}, |I| \geq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2 (1-2\varepsilon)^{2|I|} \\
&\leq \sum_{|I| \text{ is odd}, |I| \leq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2 (1-2\varepsilon)^{2|I|} + (1-2\varepsilon)^{2\eta n} \sum_{|I| \text{ is odd}, |I| \geq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2 \\
&\leq \sum_{|I| \text{ is odd}, |I| \leq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2 + (1-2\varepsilon)^{2\eta n} \sum_{|I| \text{ is odd}, |I| \geq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2
\end{aligned}$$

When in the last transition we use the fact that  $1-2\varepsilon < 1$ . By picking  $\eta$  such that  $(1-2\varepsilon)^{2\eta n} = \Theta(\delta^3)$  we have that for a family of tests:

$$3\delta^2 \leq \sum_{|I| \text{ is odd}, |I| \leq \eta n} \alpha_I^2 \beta_{C^{-1}(I)}^2 \quad (1)$$

As the summation is over  $I$  at odd size, the empty set is not counted in the summation, namely there must be a non empty  $I$  such that  $|I| \leq \frac{1}{2} \log_{1-2\varepsilon}(\delta^3)$  and  $\alpha_I \beta_{C^{-1}(I)}$  have non zero weight. Thus we can define:

$$L_f = \left\{ I : |I| \leq \frac{1}{2} \log_{1-2\varepsilon}(\delta^3) \text{ and } |I| \text{ is odd} \right\}$$

$$M_g = \left\{ C^{-1}(I) : |I| \leq \frac{1}{2} \log_{1-2\varepsilon}(\delta^3) \text{ and } |I| \text{ is odd} \right\}$$

## 2.4 Ex 3. The label cover problem.

**The reduction.** Let  $\langle G = (V, E), \{c_e\} \rangle$  be a given instance of the Label cover problem. For each edge  $e = \{v, u\} \in E$  define the test  $T_\varepsilon(c_e)$  as defined above, Thus in total we define a  $|E|$  tests, denote them by  $T$ . Consider the language  $L$  such that a test collection  $T$  is in  $L$  if there exists function  $f \times V$  such that the probability:

$$\Pr [T_\varepsilon(c_{\{v,u\}}) \text{ accepts on } f_v, f_u] \geq \frac{1}{2} + \delta$$

For every  $\{v, u\} \in E$ . A probabilistic verifier takes a candidate  $f \times V : \pm \times V \rightarrow \pm$ , picks a random edge  $e \in E$  and then check  $T_\varepsilon(c_e)$  over the functions  $f_v, f_u$ .

**Completeness.** Suppose that  $\langle G = (V, E), \{c_e\} \rangle \in (\mu, 1)$ -Label Cover then either there exists a labeling  $A$  such that  $c_{vu}(A(v)) = A_u$  for any  $\{v, u\} \in E$  or that any labeling satisfies at most  $\mu$  constraints. For completeness, let's assume the first case, and denote by  $A$  the satisfying labeling. Consider the function  $f \times V : \pm \times V \rightarrow \pm$  defined as follow:  $f_v = \chi_{\{A(v)\}}$ , So by the first section of part 2 we have that any of the test accepts with probability  $1 - \varepsilon$ . That it, as we pick a test uniformly random, the existences of satisfying labeling for the label cover problem give a function that pass the test with probability  $1 - \varepsilon$ .

**Soundness.** Now, assume the second case, namely that any labeling satisfies at most  $\mu$  constraints. Also assume through contradiction that there exists an assignment that satisfies more than  $\frac{1}{2} + \delta$  equations, so by the same arguments we use in section 2.b we have that the expectation of the product  $\mathbf{E} [\mu f_v(\mu x) f_u(y) f_v(zxc^{-1}(y))] \geq 2\delta$  when here, in addition for taking the expectation over the  $x, y, z, \mu$  we also summing on the edges  $\{v, u\} \in E$ .

Now we are about to show that for at least  $\delta$  of tests the product  $\mathbf{E} [\mu f_v(\mu x) f_u(y) f_v(zxc^{-1}(y)) | u, v]$  conditioned on the test is greater than  $\delta$ . For convenient let's use the notation  $\mathbf{E}[\cdot] \geq \delta$  for referring to tests that the average in on their product is greater than  $\delta$ , and by the same manner let's use the notation  $\mathbf{E}[\cdot] \leq \delta$ . So:

$$\begin{aligned} 2\delta &\leq \mathbf{E} [\mu f_v(\mu x) f_u(y) f_v(zxc^{-1}(y))] = \\ &\Pr [u, v \text{ s.t } \mathbf{E}[\cdot] \geq \delta] \mathbf{E} [\mu f_v(\mu x) f_u(y) f_v(zxc^{-1}(y)) | u, v \text{ s.t } \mathbf{E}[\cdot] \geq \delta] + \\ &\Pr [u, v \text{ s.t } \mathbf{E}[\cdot] \leq \delta] \mathbf{E} [\mu f_v(\mu x) f_u(y) f_v(zxc^{-1}(y)) | u, v \text{ s.t } \mathbf{E}[\cdot] \leq \delta] \\ &\leq \Pr [u, v \text{ s.t } \mathbf{E}[\cdot] \geq \delta] \cdot 1 + \Pr [u, v \text{ s.t } \mathbf{E}[\cdot] \leq \delta] \cdot \delta \\ &\leq \Pr [u, v \text{ s.t } \mathbf{E}[\cdot] \geq \delta] + \delta \end{aligned}$$

Thus for at least  $\delta$  fraction of the tests equation 1 holds. Now consider the follow probabilistic assignment, for any vertex  $v$  we choose a set  $I \subset [n]$  at probability that equals to the projection of  $f_v$  on  $\chi_{\{I\}}$ , then picking uniformly from the support of  $I$  a label for  $v$ .

### 3 Part 3.

**Label cover when the aleph-bet depends on the vertex.** Instead of showing reduction into the general label cover we will show a reduction to a similar problem in which vertices can have an additional restriction on the valid characters that one can set on. In formal, we will say that  $\langle G, \{\Sigma_v : v \in V\}, \{c_e : e \in E\} \rangle$  instance of Generalized-Label-Cover if there is an labeling  $A : V \rightarrow \Sigma$  such that for any  $\{v, u\} \in E$  it holds that  $c_e A(v) = A(u)$  and in addition for any  $v \in V$  we have that  $A(v) \in \Sigma_v \subset \Sigma$ .

**The reduction.** Define the Bipartite graph  $G = (L, R, E)$ . Associate the left vertices with the variables and the right with the closures. Define  $\{u, v\}$  to be an edge if the literal which associate with the vertex  $u$  is in the closure associate with vertex  $v$ . For the alphabet take  $\Sigma = \mathbb{Z}_2^3$ . For any right vertex  $v \in R$  define  $\Sigma_v$  be all the assignments for which the  $v$ -closure is satisfied and for any left vertex  $u$  define  $\Sigma_u = \{(1, 0, 0), (0, 0, 0)\}$ . Finally define  $c_e$  for  $e = \{v \in R, u \in L\}$  to be the projection of  $\sigma \in \Sigma$ , setted on  $v$ , to the coordinate corresponding with  $u$ . For example, assume that  $v$  associate with  $x \vee y \vee z$  and let  $u$  be the vertex associate with  $x$ , And assume that  $A(v) = (1, 0, 1)$ , then  $c_e A_v = (1, 0, 0)$ .

**Completeness.** Suppose that  $\varphi \in \text{E3-CNF-SAT}$  and let  $x \in \mathbb{F}_2^*$  be the assignment that satisfies  $\varphi$ . That is,  $\varphi(x) = \text{True}$ . Let  $A$  be the labeling that sets for any vertex on the left the bit matched to that literal by  $x$  followed by zeros padding. And for any right vertex the triple of the bits corresponding to literals involving in the associated closure. By the fact that  $x$  satisfies  $\varphi$  any closure in  $\varphi$  is satisfied by  $x$  and therefore each of the right vertices (closures) see on his local view a character of  $\Sigma_v$ . In addition by the definition of the construction any pair of connected vertices satisfies the edge restriction.

**Soundness.** Suppose that  $\varphi \in \text{E3-CNF-SAT}$  but not satisfiable and  $\langle G, \{\Sigma_v : v \in V\}, \{c_e : e \in E\} \rangle$  is an instance obtained by the reduction above. Assume towards contradiction that there exists labeling  $A$  such that more than  $\mu' = 6\mu$  of the restriction  $\{c_e\}$  are satisfied.

Define by  $\alpha_i$  to be the number of right vertices which satisfy exactly  $i$  edges, that is,

$$\alpha_i = |\{v \in R : |\{c_e A(v) = A(u) : u \in L\}| = i\}|$$

**Claim 1.** For any labeling  $A$  such that  $\alpha_3 \geq \mu$  there exists an assignment  $x \in \mathbb{F}_2^*$  satisfies at least  $\mu$  portion of the restrictions.

*Proof.* The proof is trivial. □

**Claim 2.** For any labeling  $A$  that satisfy  $\xi$  constraints, there exists labeling  $A'$  such that any constraint that satisfied by  $A$  also satisfied by  $A'$  and in addition  $\alpha_0 = \alpha_1 = 0$ . Put it differently, we can assume that  $\alpha_0 = \alpha_1 = 0$ .

*Proof.* Let  $v \in R$  be a vertex that satisfies less than two edges. Recall that  $\Sigma_v$  contains all the triple that satisfy the closure associated with  $v$ . By the fact that for any 3-CNF closure there is exactly one assignment which does not satisfy it, It follows that  $|\Sigma_v| = 2^3 - 1 = 7 \geq 2^2$ . Therefore, we can replace  $A(v)$  by a triple that agree with the first two vertices connected to it. □

Using the above claim we can infer that  $\alpha_2 + \alpha_3 = |R|$  and in addition  $2 \cdot \alpha_2 + 3 \cdot \alpha_3 \geq \mu' \cdot 3|R|$ . Thus,  $\alpha_3 \geq (3\mu' - 2)|R|$ . Particularly if  $\mu' \geq \frac{\mu+2}{3}$  then  $\alpha_3 \geq \mu|R|$ , Combining the claim above we get a contradiction to the fact that  $\varphi \in (\mu, 1)$  gap-3E-CNF-SAT and not satisfiable.