Hardness Of Approximation

Lecture 8: Direct Product testing

Instructor: Irit Dinur and Amey Bhangale *Scribe: Amey Bhangale*

1 Direct Product Testing

Last week we finished the proof of the parallel repetition theorem using the direct product theorem. In this lecture, we will complete the full proof by proving the direct product theorem.

Recall the setting in the direct product testing. We have a collection of strings $\{f_S\}$, for every subset S of $[n]$ of size at most k. Here $f_S \in \{0,1\}^S$. These are supposedly restrictions of a fixed n bit string. We need to check if this is the case.

Definition 1.1. *$\{f_S\}$ is called* perfect *if there exists $h : [n] \to \{0,1\}$ such that $f_S = h|_S$ for every $S \subset [n]$, $|S| \leq k$.*

Our goal is to test if a given collection is close to perfect.

Agreement Test with $\rho \in (0,1)$:

1. Choose $(v_i, v_2, \ldots, v_k) \in V^k$ uniformly at random and let $S = \cup_i v_i$.
2. For each $i \in [k]$, with probability ρ set $v_i' = v_i$ and with probability $(1-\rho)$ set v_i' to be a uniformly random element from V. Let $S' = \cup_i v_i'$.
3. Accept iff $f_S|_r = f_{S'}|_r$, where $r = S \cap S'$.

The following theorem is our main goal today (this appeared as Theorem 1.5 in notes for lecture 7).

Theorem 1.2. *Let $\rho = 0.1$ and choose $\alpha = 10^{-5}$. There exists $0 < \gamma < 1$ (that depends on ρ and α) such that if $\{f_S\}$ passes the agreement test with probability at least $\epsilon := (1-\gamma)^k$, then there exists some $r \subset [n]$, $|r| \approx \rho k$ and a function $h_r : [n] \to \{0,1\}$ such that*

$$Pr_{S \supseteq r}\left[f_S \overset{\geq(1-\alpha)k}{=} h|_S\right] \geq \epsilon^{O(1)},$$

where the notation $\overset{\geq(1-\alpha)k}{=}$ means that the two strings disagree on at most αk locations.

Note that the amount of sets S that contain r is a very small, sub-constant, fraction of all sets, roughly $n^{-|r|}$.

A few remarks about the theorem.

- Naively one would expect a stronger theorem, that guarantees a global function $h : [n] \to \{0,1\}$ such that $Pr[h|_S = f_S]$ is noticeable when the acceptance probability is above $exp(-k)$. However, this is too strong a hope when the acceptance probability is $exp(-k)$ as we demonstrate in Section 1.2.

  It turns out that if the acceptance probability is $k^{-O(1)}$ then one can in fact conclude that there is a global function agreeing on $k^{-O(1)}$ fraction of the f_S [DG08].

- As we saw in the previous lecture, even this less-global conclusion, as stated in the theorem, is enough to prove the parallel repetition theorem with exponential decay.

How can one find the global function h_r? Clearly, the natural strategy of defining $h_r(x)$ by the value that is most popular among all $\{f_S(x) \mid x \in S\}$ is not going to work. Here is simple counter example to this strategy. For each S assign a random string from $\{0^k, 1^k\}$ to f_S. Clearly, the agreement test passes with probability at least $1/2$, but the plurality strategy gives a random function.

To overcome such examples, the overall idea is to *zoom-in* to a small subset of $\{S\}$ such that we enjoy much stronger agreement among the sets form the subset.

1.1 Graphs associated with the test

The test distribution gives rise to different weighted graphs. Let's take a look at a few natural graphs related to the test.

1. DP_ρ: In this graph, the vertex set is V^k (ordered tuples) and the edge set is given by the agreement test. Starting from $(v_1, v_2, \ldots, v_k)$, we move to a neighbor by the following process - for all $i \in [k]$ independently, we keep v_i as it is with probability ρ and with the remaining probability, we select a random vertex from V. The good thing about this graph is that it is a product graph and hence it is easier to analyze analytically. For example we can get the eigenvalues of this graph by just knowing the eigenvalues of the base graph.

2. Folded DP_ρ graph: Here we glue together all the ordered tuples. Thus, in this graph the vertex set is all subsets of V of size at most k and the edge weight is the total weight of moving from subset S to subset T in the graph DP_ρ. The expansion behavior on this folded graph dictates the expansion behavior on the graph DP_ρ.

3. Johnson graph $J(n,k)$. Here n is the size of the universe and the vertex set is all k-sized subsets of n. (S, S') is an edge in this graph iff $S \cap S' = k-1$. This graph is studied widely in the literature. The reason this graph is related to the above two graphs is that a short walk (say of length $(1-\rho)k$ when $k \ll n$) on $J(n,k)$ is similar to the previous graphs.

1.2 Expansion of small sets in DP_ρ

It is very instructive to think of the cases when the direct product test passes with non-negligible probability but there is no global function agreeing with $\{f_S\}$. To this end, consider the folded graph DP_ρ. In this graph, there are many small sets which *do not* expand. For example for a fixed subset r of size $\ll k$ (even size 1) and consider the family of sets $\{S \mid S \supset r\}$. In a typical step in this graph we are keeping each element in r roughly with probability ρ. Thus, with probability roughly $\rho^{|r|}$ we stay in the same set $\{S \mid S \supset r\}$. This gives a way to create a collection $\{f_S\}$ which will pass the agreement test with non-negligible probability. For every r of size ρk, take a random function g_r and set $\{f_S \mid S \supset r\}$ with respect to g_r (if not assigned previously). In this case, there is no global function correlated with $\{f_S\}$ but the agreement test passes with probability at least $\rho^{|r|}$. This is precisely because in the agreement test we end up selecting (S, S') from $\{S \mid S \supset r\}$ for some r with probability $\rho^{|r|}$. Thus, in this respect Theorem 1.2 is tight!

Thus, studying small set expansion property in these graphs is instrumental in analyzing such tests. This is different from the global expansion of the graph which is dictated, through Cheeger's inequality, by the second eigenvalue of the associated adjacency matrix.

We will rely on the following theorem regarding small set expansion in the graph DP_ρ.

Theorem 1.3 ([MOR⁺06])**.** *Suppose $A, B \subseteq [n]^k$ of size at least ϵ then*

$$Pr_{(x,y)\in DP_\rho}[x \in A \ \&\ y \in B] \geq \epsilon^{\frac{2-\sqrt{\rho}}{1-\sqrt{\rho}}}$$

A few simple observations regarding the above theorem.

- If $\rho = 0$ then x and y are totally uncorrelated and hence we get that the probability of the event $x \in A$ and $y \in B$ is ϵ^2, as expected.
- If $\rho = 1$ then x and y are perfectly correlated and if A and B are disjoint then we do get the probability to be 0. 3) when ρ is somewhere in between, say $1/2$, then the lemma non-trivially says that no matter which sets A and B we take, we have a considerable chance that $x \in A$ and $y \in$B.

1.3 Proof of Theorem 1.2

Definition 1.4. *A restriction r is "good" if there exists $g : r \to \{0,1\}$ such that the set*

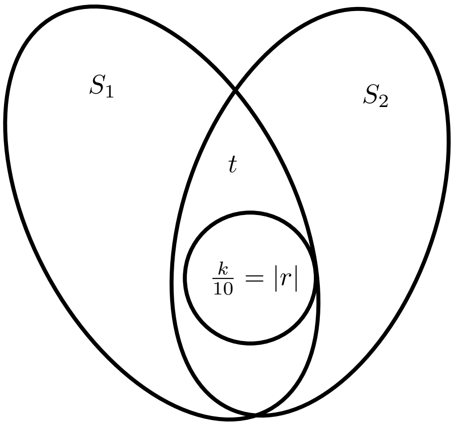
$$Z_r^g = \{S \supset r \mid f_S|_r = g\}$$

is of size at least $\Omega(\epsilon)$.

Claim 1.5. *There are at least $\Omega(\epsilon)$ fraction of good r, where r is distributed according to the test distribution.*

Proof. This follows from a simple averaging argument. □

Consider the following distribution D_1: (r, t, S, S') - Select $r \sim B(k, 1/10)$, $|t - r| \sim B(k, 4/10)$, $v_1, \ldots, v_t$. Then choose $S \setminus t$ and $S' \setminus t$.



Definition 1.6. *r_0 is β-excellent if*

$$Pr_{(r,t,S,S')\sim D_1}[f_S|_r = f_{S'}|_r \text{ but } f_S|_t \overset{\geq \beta k}{\neq} f_{S'}|_t \mid r = r_0] < exp(-k).$$

The notation $\overset{\geq \beta k}{\neq}$ means that the two string disagree on at least βk locations.

In other words, r_0 is excellent if for a typical pair of sets $S, S' \in Z_{r_0}^g$ agreeing on r_0 and whose intersection is more than r_0 also agree on (most of) the remaining intersection. This property is crucial in arguing that the plurality vote from the set $\{f_S \mid S \in Z_{r_0}^g\}$ is going to be consistent with many $\{f_S \mid S \in Z_{r_0}^g\}$.

Claim 1.7. *There are at least* $(1 - exp(-k))$ *fraction of* r *which are* β*-excellent.*

Proof. Consider r_0 and consider the family of sets $\{S \supset r_0\}$. Now, based on $f_S|_{r_0}$, we can partition the sets $\{S \supset r_0\}$ into at most $exp(r_0)$ parts. Consider a subgraph of the graph $DP_{\rho'}$ on $\{S \supset r_0\}$ where we only consider edges whose both end points are inside the same part in the partition. For an edge (S, S') let $t = (S \cap S') \setminus r_0$. We will call an edge (S, S') good if f_S and f'_S agree t on at least $(1 - \beta)$ fraction of points. Otherwise we call the edge bad. The excellence property precisely means that the fraction of bad edges is $exp(-k)$.

An alternate way of choosing D_1 is to first select t from the appropriate binomial distribution $B(k, 5/10)$ and then select r as a subset of t. According to this distribution given that the event $f_S|_t \overset{\geq \beta k}{\neq} f_{S'}|_t$ occurs, the probability that $f_S|_r = f_{S'}|_r$ is $2^{\Omega(-\beta k)}$. This is because while choosing r we will have to miss each one of the βk elements from t on which $f_S, f_{S'}$ disagree. Thus,

$$Pr_{(r,t,S,S') \sim D_1}\left[f_S|_r = f_{S'}|_r \text{ but } f_S|_t \overset{\geq \beta k}{\neq} f_{S'}|_t\right] \leq 2^{\Omega(-\beta k)}.$$

An averaging argument shows that there are at most η fraction for r_0 such that

$$Pr_{(r,t,S,S') \sim D_1}\left[f_S|_r = f_{S'}|_r \text{ but } f_S|_t \overset{\geq \beta k}{\neq} f_{S'}|_t \mid r = r_0\right] \geq \frac{2^{\Omega(-\beta k))}}{\eta}.$$

The rest of the r_0 are excellent, setting $\eta = exp(-k)$ proves the claim. □

We have a following simple corollary.

Corollary 1.8. *There are at least* $poly(\epsilon)$ *fraction of* r *which are both good and excellent.*

The following claim finishes the proof of the direct product theorem.

Claim 1.9. *If* r *is good (large* Z_r^g*) and excellent then there exists* $h_r : V \to \{0, 1\}$ *such that*

$$Pr_{S \sim Z_r^g}[f_S \overset{\geq \alpha k}{\neq} h_r|_S] \leq \epsilon^{O(1)}$$

We define the function h_r on $x \in [n] \setminus r$ by taking plurality of $\{f_S(x) \mid S \in Z_r^g, x \in S \setminus r\}$.

We will give a proof sketch here. For a more rigorous proof see [IKW12]. Before proceeding, let us see why it should work. The reason why plurality works is because we are in the *high acceptance regime* inside Z_r^g. In other words, inside the set Z_r^g, if we look at a pair of sets whose intersection is more than r then with high probability (w.p close to 1) they agree on most of the intersection. This is precisely the excellence property! Thus, once we zoom-in to Z_r^g, we have a direct product test (a slight variation as we are only considering whether they *mostly* agree or not inside the intersection instead of a *total* agreement) which accepts with probability close to 1.

In order to use the excellence property, it is desirable to consider the graph $DP_{\rho'}$ where $\rho' = 5 \cdot \rho$. In this graph, we can label edges (S, S') as 'good' if $f_S|_{S \cap S'} \overset{\geq (1-\beta)\rho' k}{=} f_{S'}|_{S \cap S'}$, and 'bad' otherwise. Since r is excellent, there are many 'good' edges. These good edges will contribute towards showing $f_S \overset{\geq (1-\alpha)k}{=} h_r|_S$, provided there are many 'good' edges inside Z_r^g. This is because f's opinion on only Z_r^g is considered while defining h_r. A priori, it is not clear why it should be the case that many 'good' edges are inside Z_r^g. This is where we use Lemma 1.3. Thus, using this lemma, there are many good edges inside Z_r^g and the *plurality decoding* works.

Proof. (Sketch) Suppose the claim in not true, this means for a random $S \in Z_r^g$, f_S and $h_r(S)$ disagree on at least αk locations with probability $\epsilon^{O(1)}$. Select a random set $e \subseteq S \setminus r$ of size $0.4k$. Then, by simple application of Chernoff bound, we get that $h_r(e)$ and $f_S(e)$ disagree on at least $\alpha/2$ fraction of the locations with high probability. However, since we define h_r by taking the plurality vote, for a random e, $h_r(e)$ should agree with at least $\Omega(\epsilon)$ fraction of $f_S|_e$ on at least $\Omega(1)$ fraction of locations. These two contradict the excellence property. The starting assumption claims that for a random e and S containing e, $h_r(e)$ and $f_S|_e$ disagree on many locations, whereas the plurality condition would imply that $h_r(e)$ and $f_S|_e$ should agree on $\Omega(1)$ fraction of points. Both these properties imply that for a random e and S, S' containing e in Z_r^g, $f_S|_e$ and $f_{S'}|_e$ disagree on many locations, contradicting the excellence property of r. □

References

[DG08] Irit Dinur and Elazar Goldenberg. Locally testing direct product in the low error range. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 613–622. IEEE, 2008.

[IKW12] Russell Impagliazzo, Valentine Kabanets, and Avi Wigderson. New direct-product testers and 2-query pcps. *SIAM Journal on Computing*, 41(6):1722–1768, 2012.

[MOR$^+$06] Elchanan Mossel, Ryan O'Donnell, Oded Regev, Jeffrey E Steif, and Benny Sudakov. Non-interactive correlation distillation, inhomogeneous markov chains, and the reverse bonami-beckner inequality. *Israel Journal of Mathematics*, 154(1):299–336, 2006.