PCPs and HDX - Lecture 11

January 23, 2017

1 Direct product and large gap amplification

In the previous lecture, we have seen that tight hardness of approximation results come by first attaining a large gap (by amplification through taking direct product) and next by using the long code gadget. The transformation through which amplification is achieved, was a simple direct product transformation. Roughly speaking, it contains a new vertex for every t original vertices, and also a new vertex for every t original constraints.

The intuition driving this construction is that a natural way to satisfy a direct product instance is by fixing an assignment for the vertices, and then, assign each t-tuple according to this assignment.

Is this really the best possible assignment to the direct product instance? As hinted in the previous lecture, the answer is no. There are better-than-direct-product strategies, that rely on correlations. There has been extensive study trying to understand the relationship between the value of the original instance and the value of the direct product instance, and we will not get into these works at this point.

Nevertheless, in a qualitative sense, the behavior of this construction is more or less as if direct product assignments are the best possible. Indeed, this is the way to view the parallel repetition theorem of Raz.

2 From Parallel Repetition to Agreement testing

In the previous lecture, we have seen a transformation from a 3SAT instance Φ to the clause versus variable label cover instance G and from there, to the t-cause versus t-variable label cover instance $G^{\otimes t}$.

- if $\mathrm{val}(\Phi) = 1$ then $val(G^{\otimes t}) = 1$
- if $\mathrm{val}(\Phi) < 1 - \delta$ then $val(G^{\otimes t}) < \exp(-t)$

whereas the first item is easy to see, this is the so-called parallel repetition theorem.

Inside the parallel repetition theorem, lies an agreement testing question.

Lemma 2.1. *Let a, b be an assignment to $G^{\otimes t}$. If $val(G^{\otimes t}, a, b) > \delta$ then $agree_D(a) > \delta^2$ where D is the distribution given by: for each $i = 1..t$, choose a random clause c_i and then u_i, v_i two independent variables inside the clause.*

To understand what this theorem means, we need to define what is the *agreement parameter* of an assignment.

Let X be a collection of subsets of some universe V. A local assignment over X is a set of local functions, one per $S \in X$, $a_S : S \to \Sigma$. Given a distribution over pairs of subsets S_1, S_2 the agreement of $a = \{a_S\}$ with respect to distribution is

$$agree_D(a) = \Pr_{S_1,S_2 \sim D}[a_{S_1}(v) = a_{S_2}(v) \text{ for all } v \in S_1 \cap S_2]$$

Proof. For each clause c_i let p_i be the conditional probability that edges incident to c_i are satisfied by a, b. Then $\mathbb{E}_i[p_i] = \delta$. For any $c = (c_1, \ldots, c_t)$ and two neigbors $(v_1, \ldots, v_t)$ and $(u_1, \ldots, u_t)$, if both constrants are satisfied then for all i in which $u_i = v_i$ t must hold that $a_u(i) = a_v(i)$. Therefore $agree(a)$ is at least the probability of choosing c and two neigbors u, v and requiring both to be satisfied. But this is exactly

$$\mathbb{E}_i[p_i^2] \geq (\mathbb{E}_i[p_i])^2 = \delta^2$$

□

Let us summarize what we know about the connection between soundness of G^t and the structure of a:

- if $agr(a) = 1$, Under mild conditions of connectivity of Φ, a is derived from a global function.
- if a is derived from global function, then $val(G^{\otimes t}, a) = val(G, a)^t$
- if a is arbitrary and $val(G^{\otimes t}, a) > \delta$ then $agree(a) > \delta^2$. So a behaves locally like a global function. Does it mean that a is close to a global function?

The last question has he property testing flavor, but there is an important difference: it deals with the low acceptance regime. An easier question might have been: if $agree(a) > 1 - \delta$ prove that $a \approx g^t$. This is the type of question we have faced in the proof of the basic PCP theorem using gap amplification. In that section the acceptance probability was always close to 1.

Before we continue, I would like to describe the cleaner setting for studying the agreement question.

3 Agreement tests

In the basic setting there is a local assignment, i.e. a collection $a = \{a_S\}$ of local functions.

A global assignment is derived from a global assignment if $a = g^{\otimes t}$.

A local test for being global is the agreement test: choose two subsets, s_1, s_2 accept if and only if their local functions a_{s_1} and a_{s_2} agree on every element in their intersection. Namely, for every $v \in s_1 \cap s_2$, $a_{s_1}(v) = a_{s_2}(v)$.

Clearly, every global assignment satisfies this test with probability one. It does not matter what the distribution on pairs of subsets is, in fact. This is equivalent to saying that all agreement tests have perfect completeness.

What about soundness? Soundness of an agreement test depends on the exact distribution on the pair of subsets. [By the way, it is quite natural to also study agreement tests that query more than just two subsets. However, two is the smallest number of queries such a test can make, so it is naturally more interesting. Moreover it has the most interesting applications towards PCPs].