PCP and Limits of Approximation Algorithms - I 4 Mar, 2015

Problem Set 1

- Due Date: **17 Mar (Tue), 2015**
- Collaboration is encouraged, but all writeups must be done individually.
- Indicate names of all collaborators.
- The length of the problem statement is NOT reflective of the problem difficulty.
- Refering sources other than the lecture notes is discouraged, since for some of the problems a Google search will reveal the solution. But if you do use an outside source (text books, lecture notes, any material available online), do mention the same in your writeup.

---

1. **Logarithmic randomness is necessary**

   Show that if $SAT \in PCP_{1,\frac{1}{2}}[r(n), O(1)]$ for $r(n) = o(\log n)$, then $P = NP$.

2. **[linearity test of 3 functions]**

   Consider the following modification of the BLR-linearity test towards testing linearity of 3 functions $f, g, h : \{0,1\}^n \to \{1,-1\}$ simultaneously.

   $$\begin{array}{ll}\textsf{BLR-3-Test}^{f,g,h} : \text{“} & 1.\ \text{Choose } y, z \in_R \{0,1\}^n \text{ independently} \\ & 2.\ \text{Query } f(y), g(z), \text{ and } h(y+z) \\ & 3.\ \text{Accept if } f(y)g(z)h(y+z) = 1. \qquad\qquad\qquad \text{”}\end{array}$$

   Clearly, if the three functions f, g, h are the same linear function, then the above test accepts with probability 1. Suppose one of the three functions f, g, h (say f) and its negation (i.e., $-f$) is δ-far from linear (this means $\max_\alpha |\hat{f}_\alpha| \leq 1 - 2\delta$), show that

   $$\Pr_{y,z}[\textsf{BLR-3-Test}^{f,g,h} \text{ rejects }] \geq \delta.$$

   [Hint: The Cauchy-Schwarz inequality $\left(\sum a_i b_i\right)^2 \leq \left(\sum a_i^2\right) \cdot \left(\sum a_i^2\right)$ may come useful.]

3. **[recycling queries in linearity test]**

   In lecture, we analyzed the soundness of the BLR-Test to show that if f is $(1/2 - \varepsilon)$-far from linear, then the test accepts with probability at most $1/2 + \varepsilon$. If we repeat this test k times, we obtain a linearity test which makes $3k$ queries and has the following property: if f is $(1/2 - \varepsilon)$-far from linear, then the test accepts with probability at most $(1/2 + \varepsilon)^k = 1/2^k + \delta$. Thus every additional 3 queries improves the soundness by a factor of 1/2. In this problem, we show that this can be considerably improved.

Assume that both f and $-f$ are $(1-\varepsilon)/2$-far from linear (i.e., $\max_\alpha |\hat{f}_\alpha| \leq \varepsilon$). Consider the following linearity test (parameterized by k).

Test_k^f : " 1. Choose $z_1, z_2, \ldots, z_k \in_R \{0,1\}^n$

2. For each distinct pair $(i,j) \in \{1,\ldots,k\}$

Check if $f(z_i)f(z_j)f(z_i+z_j) = 1$.

3. Accept if all the tests pass. "

Observe that this test makes at most $k + \binom{k}{2}$ queries. We will show below that the soundness of the test is roughly $2^{-\binom{k}{2}}$, thus showing that every additional query improves the soundness by a factor of $1/2$ (almost).

Assume that both f and $-f$ are $(1-\varepsilon)/2$-far from linear.

(a) Show that the acceptance probability of the above test is given by

$$\begin{aligned}
\Pr[\mathsf{acc}] &= \mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{i,j}\left(\frac{1+f(z_i)f(z_j)f(z_i+z_j)}{2}\right)\right] \\
&= \frac{1}{2^{\binom{k}{2}}} \cdot \sum_{S \subseteq \binom{[k]}{2}} \mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]
\end{aligned}$$

(b) Consider any term in the above summation corresponding to a non-empty S (i.e., $\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]$). Suppose $(1,2) \in S$. Show that

$$\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right]$$

is upper bounded by $\mathbb{E}_{z_1,z_2}[f(z_1+z_2)g(z_1)h(z_2)]$ for some functions $g,h : \{0,1\}^n \to \{0,1\}$.

[Hint: Fix all the variables other than z_1 and z_2 such that that the expectation is maximized.]

(c) Use the result of Problem 2 to conclude that the expression in the above (for non-empty sums) is at most ε (i.e., $\mathbb{E}_{z_1,\ldots,z_k}\left[\prod_{(i,j)\in S} f(z_i)f(z_j)f(z_i+z_j)\right] \leq \varepsilon$ for non-empty S).

(d) Conclude that $\Pr[\mathsf{acc}]$ is at most $2^{-\binom{k}{2}} + \varepsilon$.

4. **[polynomial decoding: short list of polynomials]**

Let $A : \mathbb{F}^m \to \mathbb{F}$ be any function (not necessarily a low degree polynomial). Let $p_1, p_2, \ldots, p_t : \mathbb{F}^m \to \mathbb{F}$ be the list of *all* degree d polynomials such that $\Pr_x[A(x) = p_i(x)] \geq \delta$. In other words, $p_1, \ldots, p_t$ is the list of *all* polynomials that have each agreement at least δ with the function A. Assume $\delta \geq 2\sqrt{d/q}$. Prove that $t \leq 2/\delta$. Hence, there are not too many low-degree polynomials that have considerable agreement with two polynomials.

[Hint: Use the fact that two low degree polynomial agree on at most d/q fraction of points (Schwartz-Zippel Lemma)]

5. **[low degree testing to list of polynomials]**

   In lecture, we showed that if there is a list of low-degree polynomials that agrees with the space oracle then low-degree test theorem is true. In this problem, we will show the converse of this statement.

   Suppose there exists a function $f : (0,1) \to (0,1)$ such that the following is true.

   "[Low Degree Test Theorem] For every function $A : \mathbb{F}^m \to \mathbb{F}$ and $A : \mathcal{S}_k^m \to P_{m,d}$ that satisfies

$$\Pr_{s,x}\left[A(s)(x) = A(x)\right] \geq \gamma,$$

   we have

$$\Pr_{x}\left[A(x) = Q(x)\right] \geq f(\gamma)$$

   for some polynomial Q of degree at most d (end of Low Degree Test Theorem)"

   (recall that we proved the above in lecture for the function $f(\gamma) = \gamma^2 - \varepsilon$)

   Let $\varepsilon_0 = \sqrt{d/q}$ and $\delta \in (\varepsilon_0, 1)$. Set $\delta' = f(\delta - \varepsilon_0) - \varepsilon_0 \geq 2\varepsilon_0$. Prove that for any function $B : \mathbb{F}^m \to \mathbb{F}$, there exists a list of at most $t \leq 2/\delta'$ polynomials $Q_1, \ldots, Q_t : \mathbb{F}^m \to \mathbb{F}$ of degree at most d such that

$$\Pr_{s \in \mathcal{S}_k^m, x \in s}\left[B(s)(x) \neq B(x) \wedge (\exists i, Q_i|_s \equiv B(s))\right] \geq 1 - \delta.$$

   You may assume the result of Problem 4. We will prove the above statement as follows. Suppose for contradiction that the statement if false.

   Let $Q_1, Q_2, \ldots, Q_t$ be the list of polynomials that have at least δ' agreement with B. By Problem 4, $t \leq 2/\delta'$. Suppose the statement was false. Consider the following 3 events for a random $s \in \mathcal{S}_k^m$ and $x \in s$.

   - $C : B(s)(x) = B(x)$
   - $P : \exists i \in [t], B(x) = Q_i(x)$
   - $S : \exists i \in [t], B(s) \equiv Q_i|_s$

   (a) Show that $\Pr[C \wedge \bar{S}] > \delta$. $\bar{S}$ denotes the event "not S"

   (b) Argue using Schwartz-Zippel Lemma, $\Pr[C \wedge \bar{P} | \bar{S}] \leq \varepsilon_0$.

   (c) Conclude from the previous two parts that $\Pr[C \wedge \bar{P}] > \delta - \varepsilon_0$.

   (d) Construct a new oracle $B' : \mathbb{F}^m \to \mathbb{F}$ as follows: let Q' be an arbitrary polynomial of degree exactly $d+1$. Set $B'(x)$ to be $Q'(x)$ on all points x that satisfy P and $B(x)$ elsewhere. Let the space oracle of B' be the same as that of B. Show from the previous part that

$$\Pr\left[B(s)(x) = B'(x)\right] > \delta - \varepsilon_0.$$

   (e) Conclude from the low-degree test theorem that there exists a polynomial Q of degree at most d such that $\Pr[Q'(x) = Q(x)] \geq f(\delta - \varepsilon_0)$. Argue that Q and Q' are distinct polynomials and hence,

$$\Pr[B'(x) = Q(x) \wedge B'(x) \neq B(x)] \leq \Pr[Q'(x) = Q(x)] \leq \frac{d+1}{q} \leq \varepsilon_0.$$

(f) Argue that $\Pr[B(x) = Q(x) = B'(x)] \geq f(\delta - \varepsilon_0) - \varepsilon_0 = \delta'$.

(g) Conclude from above that there exists a $i \in [t]$ such that $Q \equiv Q_i$ (i.e., Q and Q_i are identical polynomials)

(h) Conclude that $\delta' \leq \Pr[B(x) = Q_i(x) = B'(x)] \leq \Pr[Q'(x) = Q(x)] \leq \varepsilon_0$, which is a contradiction.