Probabilistically Checkable Proofs Spring 2008

Exercise no. 4

Due: May 21st, 2008

In this exercise we rely on the following fact: In the proof that we saw in class for $NP \subseteq PCP_{1,\frac{1}{2}}[O(\log n), O(1)]_{\{0,1\}^{\text{poly}\log n}}$ we in fact showed the following stronger theorem:

Theorem 1 (PCP verifier) *Let $f : \{0,1\}^n \to \{0,1\}$ be an efficiently computable Boolean function. Then, there is an efficient verifier that given oracle access to a string y and to a proof π over alphabet $\{0,1\}^{\text{poly}\log n}$, uses $O(\log n)$ randomness, makes $O(1)$ queries to y and π, and achieves the following:*

- ***Completeness:*** *If y is the low degree extension of some $x \in \{0,1\}^n$ (for some specific choice of parameters for the low degree extension), and $f(x) = 1$, then there exists a proof π that the verifier always accepts.*
- ***Soundness:*** *If either (i) y is 0.2-far from being the low degree extension of some $x \in \{0,1\}^n$; or (ii) y is 0.8-close to the low degree extension of some $x \in \{0,1\}^n$, but $f(x) = 0$; then for any proof π, the verifier rejects with probability at least $\frac{1}{2}$.*

Solve the following problems:

1. **PCP of Proximity.** Based on Theorem 1, show that SAT has a $(O(\log n), O(1))_{\{0,1\}^{\text{poly}\log n}}$-restricted verifier of proximity with completeness 1 and constant soundness and proximity parameters.

2. **The Arora-Safra composition.** For natural numbers q and n, let T be an efficiently decidable q-ary predicate on n bit strings. An *(r, q')-restricted inner verifier for T* is an efficient verifier that given oracle access to q strings $y_1, \ldots, y_q$ and to a proof π, uses randomness r, makes q' queries to $y_1, \ldots, y_q$ and π, and achieves the following:

   - **Completeness:** If $y_1, \ldots, y_q$ are, respectively, the low degree extensions of $x_1, \ldots, x_q \in \{0,1\}^n$, and $T(x_1, \ldots, x_q) = \mathit{true}$, then there exists a proof π that the verifier always accepts.
   - **Soundness:** If either (i) there is $i \in [q]$ such that y_i is 0.2-far from being the low degree extension of some $x_i \in \{0,1\}^n$; or (ii) for each $i \in [q]$, y_i is 0.8-close to the low degree extension of some $x_i \in \{0,1\}^n$, *but* $T(x_1, \ldots, x_q) = \mathit{false}$; then for any proof π, the verifier rejects with some constant probability.

   (a) **Inner verifier.** Based on Theorem 1, show that any efficiently decidable predicate T on q n-bit strings has a $(O(q \log n), O(q))_{\{0,1\}^{\text{poly}\log qn}}$-restricted inner verifier.

   (b) **Composition Theorem.** Prove the following composition theorem:

   - Let V_1 be a $(r, q)_{\Sigma_1}$-restricted verifier for SAT. Assume that the test that V_1 performs on its q queries is given by the efficiently decidable q-ary predicate T, where the answers to the q queries are given as binary strings of length $n' = \lceil \log |\Sigma_1| \rceil$.
   - Let V_2 be a $(r', q')_{\Sigma_2}$-restricted inner verifier for T.

Then, there is a $(r + r', q')_{\Sigma_2}$-restricted verifier V for SAT. If V_1 and V_2 have perfect completeness, then so does V. If V_1 has soundness s_1 and V_2 has soundness s_2, then V has soundness $s_1 + s_2$.

3. (a) **The Expander Mixing Lemma.** Let $G = (V, E)$ be an (n, d, λ)-expander. Prove that for any $S, T \subseteq V$,

$$\left| E(S,T) - |S|\,|T|\,\frac{d}{n} \right| \leq \lambda \sqrt{|S|\,|T|}$$

   (b) **Randomness-efficient amplification.** Show that for any r and $\varepsilon > 0$,

$$PCP_{\frac{2}{3},\frac{1}{3}}[r, O(1)] \subseteq PCP_{1-\varepsilon,\varepsilon}[r, \mathrm{poly}(\frac{1}{\varepsilon})]$$

4. **Robustization.** Show that any language $L \in PCP_{c,s}[r,q]_{\Sigma}$ has a PCP verifier with the same completeness c, soundness s and randomness r that makes $q' = O(q \log |\Sigma|)$ queries to a binary proof and has robust soundness $\delta = \Omega(\frac{1}{q})$.

   Hint: Use the existence of explicit codes with constant rate and linear distance, i.e., functions $E_k : \{0,1\}^k \to \{0,1\}^{O(k)}$ such that for every two $x \neq y \in \{0,1\}^k$ it holds that $\Delta(E_k(x), E_k(y)) \geq \frac{1}{3}$.

5. **Quadratic testing.** Prove the following claims:

   (a) **Tensor product.** For any four vectors $a, b, x, y \in \{0,1\}^n$, it holds that $\langle a \otimes b, x \otimes y \rangle = \langle a, x \rangle \cdot \langle b, y \rangle$.

   (b) **Lemma.** For any $n \times n$ non-zero matrix M over the field $\{0,1\}$, when picking two vectors $x, y \in \{0,1\}^n$ uniformly at random, with probability at most $\frac{3}{4}$, $x^T M y = 0$.

   (c) **Quadratic testing.** The quadratic functions encoding is a mapping $QF : \{0,1\}^{n-1} \to \{0,1\}^{2^{n^2}}$ that maps $(a_2, \ldots, a_n) \in \{0,1\}^{n-1}$ to $H(a' \otimes a')$, where $a' = (1, a_2, \ldots, a_n)$ and H is the Hadamard encoding. Show that QF is locally testable, i.e., there exists a randomized procedure $\mathcal{A}$ that given a word $w \in \{0,1\}^{2^{n^2}}$ makes $O(1)$ queries to w to decide whether w is close to a QF codeword:

   - **Completeness:** If $w = QF(a)$ for some $a \in \{0,1\}^{n-1}$, then $\mathcal{A}$ always accepts.
   - **Soundness:** If for every $a \in \{0,1\}^{n-1}$, $dist(w, QF(a)) \geq \delta$, then $\mathcal{A}$ rejects with probability at least $\Omega(\delta)$.

◇ **The PCP Theorem.** *Bonus question (20 points):* Prove that $NP \subseteq PCP_{1,\frac{1}{2}}[O(\log n), O(1)]$. You may use the composition theorem we saw in class or the Arora-Safra composition from Question 2.