6.896 Sublinear Time Algorithms	November 16, 2004
Homework 3	
Lecturer: Ronitt Rubinfeld and Eli Ben-Sasson	*Due Date: November 16, 2004*

1 Definitions

In this problem set we work over the binary field with elements $\{0, 1\}$ and addition and multiplication modulo two.

For $x, y \in \{0, 1\}^n$ let $\delta(x, y) = \frac{|\{i \in [n] : x_i \neq y_i\}|}{n}$ be the normalized Hamming distance. For $S \subset \{0, 1\}^n$ let $\delta(x, S) = \min_{y \in S} \delta(x, y)$. A function $\ell : \{0, 1\}^n \to \{0, 1\}$ is called linear iff $\ell(x + y) = \ell(x) + \ell(y)$ for all $x, y \in \{0, 1\}^n$. A code $\mathcal{C} : \{0, 1\}^n \to \{0, 1\}^m$ is called (q, δ, ϵ)-linear locally decodable if for any linear function ℓ there exists a local decoder D_ℓ making q random queries to (possibly corrupted) word $w \in \{0, 1\}^m$, such that the following holds:

$$\text{If } \exists x \in \{0,1\}^n, \delta(w, \mathcal{C}(x)) \leq \delta, \text{Then } \Pr[D_\ell^w = \ell(x)] \geq 1 - \epsilon]$$

The Hadamard code is the mapping $H^n : \{0, 1\}^n \to \{0, 1\}^{2^n}$. Each entry of a Hadamard codeword is labeled by a linear function $\ell(x)$ and $H^n_\ell(x) = \ell(x)$. (There are 2^n linear functions, each of the form $\ell(x) = \sum_{i=1}^n a_i x_i$ for $a_1, \ldots, a_n \in \{0, 1\}$).

For $x \in \{0, 1\}^n$ let $x \cdot x^T \in \{0, 1\}^{n^2}$ be the outer product of x with itself. I.e. $x \cdot x^T$ is a $n \times n$ matrix, whose (i, j) entry is $x_i \cdot x_j$.

Let $P^{n^2} : \{0, 1\}^n \to \{0, 1\}^{2^{n^2}}$ be the encoding mapping $x \in \{0, 1\}^n$ to $H^{n^2}(x \cdot x^T)$. Notice the codewords of P^{n^2} are a subset of the codewords of H^{n^2} (i.e. P^{n^2} is a sub-code of H^{n^2}).

A *constraint* is a polynomial. The *degree* of the constraint is the degree of the polynomial. The degree of a set of constraints is the maximal degree of a constraint in the set.

A Boolean circuit with n inputs and size $m \geq n$ is described by a sequence of gates $\{g_i\}_{i=1}^m$, where $g_i = x_i$ for $i \leq n$ (these are the input gates) and for $n < i \leq m$ we have either $g_i = \neg g_j$ (negation) for some $j < i$, or $g_i = g_j \wedge g_k$ (conjunction) for some $j, k < i$. The output of the circuit on input $x_1, \ldots x_n \in \{0, 1\}$ is the value of g_m on this input. A circuit is satisfiable if there exists an input for which the output is 1. Let SAT be the language of all satisfiable circuits.

A language L is said to belong to $PCP_{c,s}[q, L(\cdot)]$ if there exists a probabilistic polynomial time verifier V that on input ϕ of size n, V reads ϕ, tosses some coins, makes q queries to a proof of length $L(n)$ (over the binary alphabet), and the following hold:

- Completeness: If $\phi \in L$ then there exists a proof such that V accepts π with probability $\geq c$.
- Soundness: If $\phi \notin L$ then for every "bad proof" $\pi \in \{0, 1\}^{L(n)}$, V rejects π with probability at least s.

2 Homework Questions

Prove the following statements:

1. The Hadamard code is $(2, 1/8, 1/4)$-linear locally decodable.
2. The code P^{n^2} is locally testable with constant query complexity. Hint: notice the diagonal of $x \cdot x^T$ equals x. Use the previous question and rely on the local testability of the Hadamard code (discussed in class).
3. For $\mathcal{C} = \{C_i\}_{i=1}^k$ a set of degree two constraints over (Boolean) variables $y_1, \ldots, y_m$, the following property is locally testable with constant query complexity, perfect completeness and constant, non-zero soundness:

   $$SAT(\mathcal{C}) = \{z \in \{0,1\}^{m^2} : z = P^{m^2}(y) \text{ and } y \text{ satisfies } \mathcal{C}\}$$

   Hint: Take a sum of a *random* subset of constraints. Rely on the local testability and linear local decodability of P^{m^2} (previous questions)
4. There exist $s > 0$ and integer q such that SAT belongs to $PCP_{1,s}[q, 2^{m^2}]$. In words, SAT has PCP proofs with perfect completeness, constant non-zero soundness, constant query complexity, and proof length $O(2^{m^2})$ (with respect to a Boolean circuit of size m). Hints:
   - Let $y \in \{0,1\}^m$ be the value of all gates in the circuit on a certain input $x \in \{0,1\}^n$.
   - Define a set of degree two constraints $\mathcal{C}$ such that y satisfies $\mathcal{C}$ iff y is a "good" description of the circuit's gates on a satisfying assignment.
   - Argue that all properties needed to verify satisfiability of the circuit can be locally verified by a constant number of queries into $P^{m^2}(y)$.