HW3

Guy Levy 206865362

1: MIP vs PCP.

1.1 From MIP to PCP.

Will show: L has 2 message MIP $(l_v, l_p) \Rightarrow L$ has a $(k \cdot l_p)$-query PCP of length $k \cdot 2^{l_v} \cdot l_p$

Let $(V, (P_i)_{i=1}^k)$ be a 2 message MIP protocol for L.

The PCP proof text:
The truth tables for all functions that are defined by the MIP interaction in the following way:
$\forall i \in [k] \quad f_i(m_{vi}) = m_{p_i}$ such that m_{vi} is V's query to P_i and m_{p_i} is P_i's answer.
Since different P_i's cannot interact with each other, and since we can assume WLOG that P_i 's are deterministic,
each message from V to P_i defines his answer and so the f_i's are well defined.
The proof's length is $k \cdot 2^{l_v} \cdot l_p$ (the size of k truth tables for functions $f_i : \{m : |m| = l_v\} \to \{a : |a| = l_p\}$)

The PCP verifier:
V' simulates V to get his messages $(m_{vi})_{i=1}^k$,
queries the PCP proof for $(f_i(m_{vi}))_{i=1}^k$ (k queries for l_p bits each, so $k \cdot l_p$ 1-bit-queries),
then simulates V to accept or reject according to $(f_i(m_{vi}))_{i=1}^k$.

Completeness and soundness are immediatly implied from the MIP's completeness and soundness.

1.2 From PCP to MIP.

First we construct an MIP with q provers but very large soundness error: $1 - \frac{1}{\text{poly}(q)}$.

given input $x \in \{0,1\}^*$ describe $(V, P_1, \ldots, P_q)(x)$.
If $x \in L$ each of the q honest provers runs the same algorithm to construct some specific PCP proof text for $x \in L$.
Denote $\tilde{V}$ as the verifier of the PCP protocol.

V flips coin $b \leftarrow \{0,1\}$.

If $b = 0$: V simulates $\tilde{V}$ to get his queries.
For each $i \in [q]$, V asks P_i for the bit specified in the PCP by the i th query.

V, which now has all answers, simulates $\tilde{V}$ to accept or reject according to those answers.

If $b = 1$: V chooses random coordinate for a bit in the proof $i \in [m]$. accepts iff all provers P_i answered the same.

Completeness:
Let $x \in L \Rightarrow$,
If $b = 0$, all honest $P_1, \ldots, P_q$ return all correct answers to queries from the original PCP.
So V accepts with probability 1.
If $b = 1$, all honest $P_1, \ldots, P_q$ run the same (deterministic) algoritm, and given query i, will return the i th bit of the PCP proof text.
So V accepts by its definition.

Soundness:
Let $x \notin L$ and let $P_1^*, \ldots, P_q^*$ be some series of provers.

$$Pr[(V, P_1^*, \ldots, P_q^*)(x) = 1] =$$

$$\frac{1}{2}Pr[(V, P_1^*, \ldots, P_q^*)(x) = 1 | b = 0] + \frac{1}{2}Pr[(V, P_1^*, \ldots, P_q^*)(x) = 1 | b = 1]$$

$$= \frac{1}{2}Pr[\tilde{V} \text{ accepts answers from provers}] + \frac{1}{2}Pr_i[\text{all provers argee on input } i]$$

Case A:
Assume $\exists j \neq k \in [q] \ : \ Pr_i[P_j^*(i) \neq P_k^*(i)] \geq \frac{1}{100q^2}$:
$SE \text{ (soundness error)} \leq \frac{1}{2} + \frac{1}{2}Pr_i[\text{all } P_j^*(i) \text{ agree}] \leq Pr_i[P_j^*(i) = P_k^*(i)] \leq 1 - \frac{1}{100q^2}$

Case B:
Assume $\forall j, k \in [q] \ : \ Pr_i[P_j^*(i) \neq P_k^*(i)] \leq \frac{1}{100q^2}$:

$$SE =$$

$$\frac{1}{2}Pr_V[\tilde{V} \text{ accepts answers from provers}] + \frac{1}{2}Pr_i[\text{all provers agree on input } i]$$

$$\leq \frac{1}{2}Pr_V[\tilde{V} \text{ accepts answers from provers}] + \frac{1}{2}$$

Assume now, for the sake of contradiction that:
$Pr_V[\tilde{V} \text{ accepts answers from provers}] > \frac{3}{4}$.
Remember $\tilde{V}$ has soundness $\frac{1}{2}$ for proofs of x.
Denote for each $j \in [q]$: π^j is the PCP proof text dictated by P_j^*'s answers. i.e $\pi_i^j = P_j^*(i)$
Claim $\pi \triangleq \pi^1$ is a PCP proof for x.
$Pr[\tilde{V} \text{ accepts } \pi] \geq$
$Pr[\tilde{V} \text{ accepts } \pi \wedge \forall j \in [q] \ : \pi_i^j = \pi_i] =$
$Pr[\tilde{V} \text{ accepts } \pi \mid \forall j \in [q] \ : \pi_i^j = \pi_i] \cdot Pr[\forall j \in [q] \ : \pi_i^j = \pi_i] =$
$Pr[\tilde{V} \text{ accepts answers from provers}] \cdot Pr[\text{all prover's outputs argee given input } i] >$

$\frac{3}{4} \cdot Pr[\text{all prover's outputs argee given input } i] \geq^{\text{lemma } (*)}$
$\frac{3}{4} \cdot (1 - \frac{1}{100q}) \geq$
$\frac{3}{4} \cdot \frac{99}{100}$ which is a contradiction to the assumption of $\tilde{V}$ soundness of $\frac{1}{2}$.
Conclude from the contradiction: $SE \leq \frac{7}{8}$ for case B and so $SE \leq 1 - \frac{1}{100q^2}$ soundness error overall.

Proof of lemma $(*)$:
$Pr[\text{all prover's outputs argee given input } i] = 1 - Pr[\text{exists pair of provers which disagr}$
$= 1 - Pr[\text{exists a prover } P_j^* \text{ which disagrees with } P_1^*] \geq$
$1 - \sum_{j=2}^{q} Pr_i[P_1^*(i) \neq P_j^*(i)] \geq 1 - q \cdot \frac{1}{100q^2} = 1 - \frac{1}{100q}$

So now we have an MIP with soundness error $1 - \frac{1}{100q^2}$ and q provers.
We now describe an MIP with constant soundness error.

Given input x, we repeat the previous protocol $100q^2$ times and so we use $100q^3$ provers overall.
We accept iff all the repetitions accepted and reject otherwise.

Completeness:
If $x \in L$ all repetitions of the previous protocol will accept with probability 1, so we accept with probability 1.

Soundness:
If $x \notin L$:

$$Pr[\text{acc}] = Pr[\text{all repetitions accept}] =$$

$$\Pi_{i=1}^{100q^2} Pr[\text{rep } i \text{ accepts}] = (1 - \frac{1}{100q^2})^{100q^2} < \frac{1}{2}$$

1.3 MIPs for NP

Let $L \in NP$
$\Rightarrow$(PCP theorem) L has PCP which uses $O(1)$ queries and $O(log(n))$ randomness ($\text{poly}(n)$ length proof, denote as m)
$\Rightarrow$(1.2) L has $\text{poly}(O(1)) = O(1)$ provers, 2 message MIP in which the verifier messages have length $O(log(m)) = O(log(\text{poly}(n))) = O(log(n))$
and prover messages have length 1.

2: The Reed Solomon Code is not Locally Testable

With the definitions from the guildlines,
let $z \in \mathbb{F}^{|S|}$ be some string.

(guildline 1)
Sampling some f uniformly from all functions from $\mathbb{F}$ to $\mathbb{F}$,
is the same as sampling its truth table uniformly from $\mathbb{F}^{|\mathbb{F}|}$,
which is the same as sampling random $r \in \mathbb{F}$ for every possible input $x \in \mathbb{F}$.
Thus the following holds:

$$Pr_{f \leftarrow D_1}[f(S) = z] = \Pi_{i=1}^{|S|} Pr[f(i) = z_i] = (\frac{1}{|\mathbb{F}|})^{|S|}$$

for the analysis in the D_0 case, will use counting:

$$Pr_{f \leftarrow D_0}[f(S) = z] =$$

$$\frac{|\{p \mid deg(p) \leq d \wedge p(S) = z\}|}{|\{p \mid deg(p) \leq d\}|} =^{(*)} \frac{|\mathbb{F}|^{d+1-|S|}}{|\mathbb{F}|^{d+1}} = (\frac{1}{|\mathbb{F}|})^{|S|}$$

Explanation for $(*)$: from the Fundamental Theorem of Algebra,
to define, uniquely, a polynomial from $\mathbb{F}$ to $\mathbb{F}$ of degree d,
we need $d+1$ points.
So given $|S| < d+2$ points, we need some extra $d+1-|S|$ points to uniquely define some specific p
which satisfies all points.

So we see $Pr_{f \leftarrow D_0}[f(S) = z] = Pr_{f \leftarrow D_1}[f(S) = z]$

(guildline 2)
Let A be some non-adaptive algorithm, which reads less than $d+2$ field elements and outputs a single bit.
Then, for some S - some series of less than $d+2$ field elements:

$E_{f \leftarrow D_0}[A^f] =$
$Pr_{f \leftarrow D_0}[A^f(S) = 1] =$
$\sum_z Pr_A[A^f(S) = 1 \mid f(S) = z] \cdot Pr_{f \leftarrow D_0}[f(S) = z] =$
$\sum_z Pr_A[A^f(S) = 1 \mid f(S) = z] \cdot Pr_{f \leftarrow D_1}[f(S) = z] =$
$Pr_{f \leftarrow D_1}[A^f(S) = 1] =$
$E_{f \leftarrow D_1}[A^f]$

(guildline 3)
$E_{f \leftarrow D_1}[A^f] =$
$Pr_{f \leftarrow D_1}[A^f(S) = 1] =$
$\sum_{h \in D_1} Pr[A^f(S) = 1 \mid f = h] \cdot Pr_{f \leftarrow D_1}[f = h] =$

$\sum_{h \in D_1'} Pr[A^f(S) = 1 \mid f = h] \cdot Pr_{f \leftarrow D_1}[f = h] +$
$\sum_{h \in D_1 \setminus D_1'} Pr[A^f(S) = 1 \mid f = h] \cdot Pr_{f \leftarrow D_1}[f = h] \leq$

$\sum_{h \in D_1'} Pr[A^f(S) = 1 \mid f = h] \cdot Pr_{f \leftarrow D_1'}[f = h] +$
$\sum_{h \in D_1 \setminus D_1'} Pr[A^f(S) = 1 \mid f = h] \cdot Pr_{f \leftarrow D_1}[f = h] =$

$Pr_{f\leftarrow D_1'}[A^f(S)=1]+\sum_{h\in D_1\setminus D_1'}Pr[A^f(S)=1 \mid f=h]\cdot Pr_{f\leftarrow D_1}[f=h]=$

$E_{f\leftarrow D_1'}[A^f]+\sum_{h\in D_1\setminus D_1'}Pr[A^f(S)=1 \mid f=h]\cdot Pr_{f\leftarrow D_1}[f=h]$

Left to show (to prove claim in guildline 3) that:

$$\sum_{h\in D_1\setminus D_1'}Pr[A^f(S)=1 \mid f=h]\cdot Pr_{f\leftarrow D_1}[f=h]=2^{-\Omega(n)}$$

$\sum_{h\in D_1\setminus D_1'}Pr[A^f(S)=1 \mid f=h]\cdot Pr_{f\leftarrow D_1}[f=h]\leq$
$\sum_{h\in D_1\setminus D_1'}Pr_{f\leftarrow D_1}[f=h]\leq\frac{|D_1\setminus D_1'|}{|D_1|}\leq^{(\diamond)}\frac{|\mathbb{F}|^{d+1+0.2|\mathbb{F}|}}{|\mathbb{F}|^{|\mathbb{F}|}}=2^{-\Omega(n)}$

Explanation for $(\diamond)$:
$|D_1\setminus D_1'|=|\{f \mid \exists \text{ polynomial } p \text{ of degree } d \text{ such that } \Delta(p,f)<0.1\}|\leq$ (f is 0.1-close to some p if p can be changed on 10% of its input to become f)
$|\{p \mid p \text{ polynomial of degree } d\}|\cdot\binom{|\mathbb{F}|}{0.1|\mathbb{F}|}\cdot|\mathbb{F}|^{0.1|\mathbb{F}|}\leq$
$|\mathbb{F}|^{d+1}\cdot|\mathbb{F}|^{0.1|\mathbb{F}|}\cdot|\mathbb{F}|^{0.1|\mathbb{F}|}=|\mathbb{F}|^{d+1+0.2|\mathbb{F}|}$

(guildline 4)
So we have proven:
(1) $E_{f\leftarrow D_0}[A^f]=E_{f\leftarrow D_1}[A^f]$
(2) $E_{f\leftarrow D_1}[A^f]\leq E_{f\leftarrow D_1'}[A^f]+2^{-\Omega(n)}$

It follows that:
$E_{f\leftarrow D_0}[A^f]-E_{f\leftarrow D_1'}[A^f]\leq 2^{-\Omega(n)}$

Or, put slightly differently, denoting A's random string explicitly as r:
$E_{f\leftarrow D_0}[Pr_r[A^f(r)=1]]-E_{f\leftarrow D_1'}[Pr_r[A^f(r)=1]]\leq 2^{-\Omega(n)}$

From E's linearity:
$\Rightarrow E_{f\leftarrow D_0,f'\leftarrow D_1'}[Pr_r[A^f(r)=1]-Pr_r[A^f(r)=1]]\leq 2^{-\Omega(n)}$
$\Rightarrow \exists f\in D_0, f'\in D_1' \text{ such that } Pr_r[A^f(r)=1]-Pr_r[A^f(r)=1]\leq 2^{-\Omega(n)}$

So we found a YES instance f and a NO instance f' of the RS code, which cannot be distinguished by a PPT machine with $d+2$ queries to the function - but with only negligible probability.

Therefore the gap between completeness and soundness must be exponentially small.

$\Rightarrow$ RS is not $(d+2,0.1)$-testable.