HW2

Guy Levy 206865362

1.1 .

$P : \mathbb{F}^n \to \mathbb{F}$
n variate, d total degree.

$Pr_{z\in\mathbb{F}^n}[P(z) = 0] \leq \frac{d}{|\mathbb{F}|}$

1 . induction on n:

base case: $n = 1$ $P : \mathbb{F} \to \mathbb{F}$, $P \not\equiv 0$ univariate polynomial over field $\mathbb{F}$ having degree d. (which is also its total degree). Thus from the Fundamental Theorem of Algebra $P(z) = 0$ has no more than d solutions, i.e.
$Pr_{z\in\mathbb{F}^n}[P(z) = 0] \leq \frac{d}{|\mathbb{F}|}$

step $(n + 1)$: assume for any $P' : \mathbb{F}^n \to \mathbb{F}$ of total degree d:
$Pr_{z\in\mathbb{F}^n}[P'(z) = 0] \leq \frac{d}{|\mathbb{F}|}$
let $P : \mathbb{F}^{n+1} \to \mathbb{F}$ be polynomial of total degree d.
$Pr_{z\in\mathbb{F}^{n+1}}[P'(z_1, \ldots, z_{n+1}) = 0]$
$\leq \sum_{a\in\mathbb{F}} Pr_{z\in\mathbb{F}^{n+1}}[P'(z_1, \ldots, z_{n+1}) = 0 \mid z_1 = a] \cdot Pr_{z_1\in\mathbb{F}}[z_1 = a]$
$= \sum_{a\in\mathbb{F}} Pr_{z\in\mathbb{F}^n}[P'(a, z_1, \ldots, z_n) = 0] \cdot Pr_{z\in\mathbb{F}}[z = a]$
$\leq \sum_{a\in\mathbb{F}} \frac{d}{|\mathbb{F}|} \cdot \frac{1}{|\mathbb{F}|} = \frac{|\mathbb{F}|}{|\mathbb{F}|} \cdot \frac{d}{|\mathbb{F}|} = \frac{d}{|\mathbb{F}|}$

2 . let $n \in \mathbb{N}$, $d \leq |\mathbb{F}| - 1$

claim:
$p(z_1, \ldots, z_n) = \Pi_{k=1}^{d}(z_1 - k)$
satisfies $Pr_{z\in\mathbb{F}^n}[p(z) = 0] = \frac{d}{|\mathbb{F}|}$

Thats because $p(z) = 0$ iff $1 \leq z_1 \leq d$.
and randomly chosen z has $\frac{d}{|\mathbb{F}|}$ probability of its first coordinate to satisfy $1 \leq z_1 \leq d$.

1.2 .

lemma 1.2:
guildlines:

guildline 1.
define $I : \mathbb{F}^{2n} \to \mathbb{F}^n$ so that I is a multilinear individual degree 1 polynomial s.t:
$\forall x, x' \in \{0,1\}^n$
if $x = x'$ then $I(x, x') = 1$ else $I(x, x') = 0$

claim $I(z) = \Pi_i^n(1 - z_i - z_{i+n} + 2z_i z_{i+n})$ satisfies the above.
I is indeed multilinear.

if $\exists i \in [n]$: $z_i \neq z_{i+n}$ then $1 - z_i - z_{i+n} + 2z_i z_{i+n} = 1 - 1 = 0$ so $I(z) = 0$

else $\forall i \in [n]$: either $z_i = z_{i+n} = 1$ and then $1 - z_i - z_{i+n} + 2z_i z_{i+n} = 1 - 2 + 2 = 1$ so $I(z) = 1$
or $z_i = z_{i+n} = 0$ and then $1 - z_i - z_{i+n} + 2z_i z_{i+n} = 1 - 0 = 1$ so $I(z) = 1$

guildline 2.
observe that if $\forall x \in \{0,1\}^n : Q(x) = 0$ then $\forall z \in \mathbb{F}^n$ it holds that
$\sum_{x \in \{0,1\}^n} Q_z(x) = 0$ where $Q_z(x) = Q(x) \cdot I(x, z)$

guildline 3.
$Pr_z[\sum_{x \in \{0,1\}^n} Q_z(x) = 0 \mid \exists x \in \{0,1\}^n : Q(x) \neq 0]$
$= Pr_z[\sum_{x \in \{0,1\}^n} Q(x) \cdot I(x, z) = 0 \mid \exists x \in \{0,1\}^n : Q(x) \neq 0]$
$\leq \frac{nd+1}{|\mathbb{F}|}$ from polynomial identity lemma and $(*)$,$(**)$ of which proof is later

$(*)$ Q individual degree d, I individual degree 1.
$t_z(x) = \sum_{x \in \{0,1\}^n} Q(x) \cdot I(x, z)$ is total degree of at most $nd + 1$

$(**)$ $\exists x' \in \{0,1\}^n : Q(x) \neq 0$ so for $z = x'$: $\sum_{x \in \{0,1\}^n} Q(x) \cdot I(x, x') = Q(x') \cdot 1 \neq 0$
so not the 0 polynomial

conclude:
$Pr_z[\sum_{x \in \{0,1\}^n} Q_z(x) \neq 0 \mid \exists x \in \{0,1\}^n : Q(x) \neq 0] \geq 1 - \frac{nd+1}{|\mathbb{F}|}$

guildline 4.
given $Q : \mathbb{F}^n \to \mathbb{F}$, V chooses $z \leftarrow \{0,1\}^n$ randomly and sends ot to P.

initiate sumcheck V_{sc}, P_{sc} on $Q_z(x) = Q(x) \cdot I(x, z)$
(given orcale access to Q, V supplies V_{sc} with orcale access to Q_z by accesssing $Q(x)$ and multiplying with $I(x, z)$)
V accepts or rejects accordingly.

completeness:

assume $Q(x) = 0 \quad \forall x \in \{0,1\}^n$

$Q_z(x) = Q(x) \cdot I(x,z) = 0 \quad \forall x \in \{0,1\}^n$

$Pr[V \text{ accepts}] = Pr[V_{sc} \text{ accepts}] = 1$

soundness:
assume $\exists x \in \{0,1\}^n : Q(x) \neq 0$
$Pr[V \text{ accepts}] = Pr[V_{sc} \text{ accepts}]$
$= Pr[V_{sc} \text{ accepts} \mid \sum Q_z(x) = 0] \cdot Pr[\sum Q_z(x) = 0] + Pr[V_{sc} \text{ accepts} \mid \sum Q_z(x) \neq 0] \cdot$
$\leq 1 \cdot \frac{nd+1}{|\mathbb{F}|} + \frac{nd+1}{|\mathbb{F}|} \cdot 1 = O(\frac{nd}{|\mathbb{F}|})$

complexity:
V can compute I in $poly(n, d, log|\mathbb{F}|)$ so that will be the cost of every orcale query by V_{sc} (considering also V queries its orcale in $O(1)$).
V_{sc}'s communication is $poly(n)$, so overall the protocol runs in $poly(n, d, log|\mathbb{F}|)$.

number of queries:
V queries Q exactly once for every time V_{sc} queries Q_z which is n times (number of coordinates of input to Q_z).
So V queries Q n times.

2 . (did not manage solve fully)

Show $\text{GapClique}_{\frac{1}{2}}(k)$ is NP-Hard.

guildline 1:
Let $L \in NP$. goal: show $L \leq_p \text{GapClique}_{\frac{1}{2}}(k)$.

guildline 2:
Let $x \in L$. Consider the PCP verifier V_x for L that makes q queries and uses r random bits (and has x hardcoded).

guildline 3:
Construct $G_x = (V, E), \quad V \subseteq \{0,1\}^q \times \{0,1\}^r, \quad E \subseteq V \times V$
that includes only vertices $\omega \circ \rho$ such that V_x accepts the answers $\omega \in \{0,1\}^q$ when using the random string $\rho \in \{0,1\}^r$
$V = \{\omega \circ \rho \mid V_x^\rho \text{ accpets and } \omega \text{ is the answers to its queries } \}$
(using the superscript notation to specify the random string V_x runs with)

guildline 4:

Put an edge between two vertices iff they are consistant, namely the answers of neighboring vertices should agree on each common query

$$E = \{(\omega \circ \rho_1, \omega \circ \rho_2) \mid \omega \in \{0,1\}^q, \rho_1, \rho_2 \in \{0,1\}\}$$

guildline 5:
Choosing $k = \frac{2^r}{2^q}$ show that if $x \in L$ then the graph has a clique of size k whereas if $x \notin L$ then it can have a clique of at most $\frac{k}{2}$

$x \in L$
$\Rightarrow$ $Pr[V_x() = 1] = 1$
$\Rightarrow$ all 2^r runs accept.
$\Rightarrow$ G_x has 2^r vertices $\omega \circ \rho$
notice there are only 2^q possibilities for ω so from the pigeon hole principle there is some ω^* which $k = \frac{2^r}{2^q}$ vertices agree on. i.e. G_x has clique of size k

$x \notin L$
$\Rightarrow$ $Pr[V_x() = 1] \leq \frac{1}{2}$
$\Rightarrow$ no more than 2^{r-1} of rhos accept.
assume for the sake of contradiction G_x has a clique of size $\frac{k}{2} + 1 = \frac{2^r}{2^q} + 1$
$\Rightarrow$ could not solve :((hopefully some contradiction)

So the karp reduction $f(x) = G_x$ satisfies:
$x \in L \Rightarrow G_x \in \mathrm{GapClique}_{\frac{1}{2}}(k)[YES]$
$x \notin L \Rightarrow G_x \in \mathrm{GapClique}_{\frac{1}{2}}(k)[NO]$

i.e. of we can solve $\mathrm{GapClique}_{\frac{1}{2}}(k)$ in polynomial time, we can decide L in polynomial time.
conclude $\forall L \in NP : L \leq_p \mathrm{GapClique}_{\frac{1}{2}}(k)$ i.e. $\mathrm{GapClique}_{\frac{1}{2}}(k)$ is NP hard.

3 .

$A \in_R \{0,1\}^{n \times k}$
$C : \{0,1\}^k \to \{0,1\}^n$
$C(x) = Ax$

Let $\varepsilon > 0$, assume $n = \Omega(\frac{k}{\varepsilon^2})$
show C has relative distance of at least $\frac{1}{2} - \varepsilon$ with probability 0.99.

C is linear so enough to show for $x = 0$: $Pr_A[\forall x \neq 0 : \bar{\omega}(C(x)) \geq \frac{1}{2} - \varepsilon] \geq 0.99$
equivalently: $Pr_A[\exists x \neq 0 : \bar{\omega}(C(x)) < \frac{1}{2} - \varepsilon] < 0.01$

analyze for for some fixed $x \neq 0$:
$Pr_A[\bar{\omega}(C(x)) < \frac{1}{2} - \varepsilon]$
$= Pr_A[\bar{\omega}(Ax) < \frac{1}{2} - \varepsilon]$
$= Pr_A[\sum_i^n (Ax)_i < n(\frac{1}{2} - \varepsilon)]$

interlude:
notation E_i is an indicator for the event $(Ax)_i = 1$.
claim $Pr_A[E_i] = \frac{1}{2}$
proof: $x \neq 0 \;\; (Ax)_i = \langle row_i A, x \rangle$
saw in previous HW that $Pr[\langle row_i A, x \rangle = 0] = \frac{1}{|\mathbb{F}|} = \frac{1}{2}$
$\Rightarrow 1 - Pr[\langle row_i A, x \rangle = 1] = \frac{1}{2}$
$\Rightarrow Pr[\langle row_i A, x \rangle = 1] = \frac{1}{2}$

continue analysis:

$= Pr[\frac{\sum E_i}{n} < \frac{1}{2} - \varepsilon]$
$= Pr[\frac{1}{2} - \frac{\sum E_i}{n} > \varepsilon]$
$< e^{-\frac{\varepsilon^2}{4 \cdot \frac{1}{4}} n}$ (chernoff)
$= e^{-\varepsilon^2 n}$
$\leq e^{-k} \quad (n = \Omega(\frac{k}{\varepsilon^2}))$

So, using the union bound:
$Pr_A[\exists x \neq 0 \;:\; \bar{\omega}(C(x)) < \frac{1}{2} - \varepsilon]$
$< \sum_{x \neq 0} Pr_A[\bar{\omega}(C(x)) < \frac{1}{2} - \varepsilon]$
$\leq 2^k \cdot e^{-k} = (\frac{e}{2})^{-k} \leq 0.01$ (goal)

That holds for $k \geq -log_{e/2}(0.01) \approx 15.007$ i.e. $k \geq 16$