CSS.330.1 : PCPs
Limits of Approximation Algorithms

Lecture 02 (2023-2-3)

Instructor: Prahladh Harsha

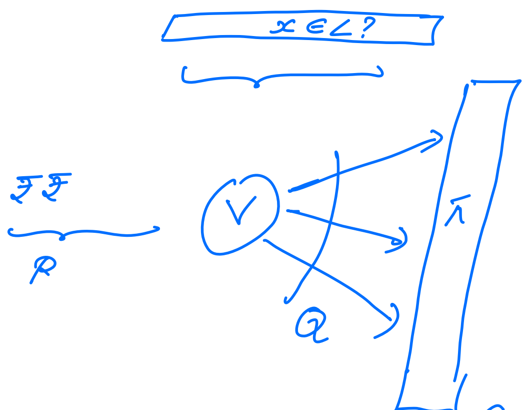
Today

- Linearity Testing (BLR, Fourier)
- Constant-query exponential-sized PCPs

Recall the PCP Thm (the proof-checking viewpoint).

PCP Theorem: $\exists\ Q,\ \underset{\alpha \in (0,1)}{\exists}\ \forall\ L \in NP,\ \exists\ c$

$$L \in PCP_{1,1-\alpha}[c \log n, Q]$$



Key Difference:

V makes only Q queries into the proof.

(cost: randomized verifier).

① Q: can be made as small as 3.

② Randomness needed: O(log n)

Qn: Are there properties $\mathcal{P}$ (sets of strings) such that one can check if $f \in \mathcal{P}$ by probing f at very few locations?

Tester distinguish string $\in P$

string far from P

(far - usually Hamming metric).

Linearity Testing:

$f: \{0,1\}^n \to \{0,1\}$ $\quad (\{0,1\} - GF(2))$

f is linear if

$$f(x+y) = f(x) + f(y), \quad \forall x, y \in \{0,1\}^n$$

f is linear iff $\exists\, \alpha \in \{0,1\}^n$

$$\underbrace{f(x) = \sum \alpha_i x_i}_{f = \ell_\alpha} \quad \forall x \in \{0,1\}^n$$

Qn: Given a fn: $f: \{0,1\}^n \to \{0,1\}$ (as an oracle)

check if f is linear or far from linear?

"far from linear:"

$$\delta(f) = \min_\alpha \delta(f, \ell_\alpha) = \min_\alpha \Pr_x [f(x) \neq \ell_\alpha(x)]$$

Blum-Luby-Rubinfeld '91

BLR^f: 1. Pick $x, y \in_R \{0,1\}^n$
2. Query f at $x, y, x+y$
3. Accept if $f(x+y) = f(x) + f(y)$.

$\varepsilon(f)$ = Rejection probability of test
$= \Pr_{x,y}[f(x+y) \neq f(x) + f(y)]$.

Understand: $\varepsilon(f)$ vs $\delta(f)$.

Completeness: If $\delta(f) = 0$ (ie f is linear)
$\Downarrow$
$\varepsilon(f) = 0$

Soundness: $\varepsilon(f)$ is small $\overset{??}{\Rightarrow}$ $\delta(f)$ is small

Lemma [BLR, Coppersmith]
If $\varepsilon(f) < 2/9 \Rightarrow \delta(f) \leq 2\varepsilon(f)$.

Proof:

φ – self corrected function.

$\varphi : \{0,1\}^n \to \{0,1\}$
$\varphi(x) \triangleq \underset{y \in \{0,1\}^n}{\text{Plurality}}\{f(x+y) - f(y)\}$

SubClaim 1: $\delta(f, \varphi) < 2\varepsilon(f)$

Pf: $BAD = \{x \in \{0,1\}^n \mid \Pr_y[f(x) \neq f(x+y) - f(y)] \geq 1/2\}$.

$x \notin BAD \Rightarrow \varphi(x) = f(x)$

$\delta(f, \varphi) \leq \Pr_x[x \in BAD]$

$\varepsilon(f) = \Pr_{x,y}[f(x) + f(y) \neq f(x+y)]$

$\geq \Pr_x[x \in BAD] \cdot \Pr_{x,y}[f(x) \neq f(x+y) - f(y) \mid x \in BAD]$

$\geq \Pr_x[x \in BAD] \cdot 1/2 \geq \delta(f, \varphi)/2$. $\square$

SubClaim 2: $\forall x$, $\Pr_y[\varphi(x) = f(x+y) - f(y)] > 2/3$.

(i.e., self-corrected fn is not just the most popular value, but actually the overwhelming majority).

Pf: Fix x

$\Pr_{y_1, y_2}[f(x+y_1) - f(y_1) = f(x+y_2) - f(y_2)]$

$= \Pr_{y_1, y_2}[f(x+y_1) + f(y_2) = f(x+y_2) + f(y_1)]$

$\geq \Pr_{y_1, y_2}[f(x+y_1) + f(y_2) = f(x+y_1+y_2) = f(x+y_2) + f(y_1)]$

$\geq 1 - 2\varepsilon(f) > 5/9$ . . . $(*)$

For each b $\quad p_b = \Pr_y [f(x+y) - f(y) = b]$

$\therefore \sum_b p_b = 1$ $\quad$ (Recall $\varphi(x) = \arg\max_b p_b$)

(*) $\therefore \sum_b p_b^2 > 5/9.$

$\Rightarrow (\max p_b) \sum_b p_b > 5/9$

$\Rightarrow \max p_b > 5/9.$

But,

$(\max p_b)^2 + (1 - \max p_b)^2 \geq \sum p_b^2 > 5/9$

$2(\max p_b)^2 - 2(\max p_b) + 4/9 > 0.$

$\left(\alpha^2 - \alpha + 2/9 > 0 \qquad (\alpha - 2/3)(\alpha - 1/3) > 0 \right)$

Hence, $\max p_b > 2/3$ $\quad$ (since $\max p_b > 5/9 > 1/3$).

$\boxtimes$

SubClaim 3: φ is linear

Proof: Fix $x, y \in \{0,1\}^n$

$\varphi(x) = f(x+z) - f(z)$ but for $< 1/3$ of z's

$\varphi(y) = f(z) - f(z-y)$ but for $< 1/3$ of z's

$\varphi(x+y) = f(x+z) - f(z-y)$ but for $< 1/3$ of z's.

Hence $\exists$ one z, for which all the above 3 are true.

But then

$$\varphi(x) + \varphi(y) = f(x+z) - f(z-y)$$
$$= \varphi(x+y).$$ $\boxtimes$

Hence concludes Lemma $\boxtimes$.

Observations :

① Proof doesn't use Domain $= \{0,1\}^n$ or Range $= \{0,1\}$

Works for the following more general case.

G, H - arbitrary Abelian Groups

$f: G \to H$

Homomorphism : $\forall\, x, y \in G$, $f(x) +_H f(y) = f(x +_G y)$

② Coppersmith's counterexample that $2/9$ is tight.

$G = H = \mathbb{Z}/3n\mathbb{Z}$ for n - large integer

$f: G \to G$

$$f(x) = \begin{cases} 0 & \text{if } x \equiv 0 \pmod 3 \\ 1 & \text{if } x \equiv 1 \pmod 3 \\ 3n-1 & \text{if } x \equiv -1 \pmod 3 \end{cases}$$

Obs: $\delta(f, \mathrm{Hom}(G,H)) = 2/3$.

$x \backslash y$	0	1	-1
0	✓	✓	✓
1	✓	×	✓
-1	✓	✓	×

$\varepsilon(f) = 2/9$

③. $x, y \in_R G$, were independent elts of G.

Can one use lesser randomness?

Yes: Shpilka - Wigderson.

x – uniform elt of G

y – ε-biased set of G.

(problem set 1).

Alternative Analysis via Fourier

$G = \{0,1\}^n$; $\quad H = \{0,1\}$

$= (GF(2))^n \quad\quad = GF(2)$.

Bellare - Coppersmith - Håstad - Kiwi - Sudan '96.

$\ell_\alpha(x) = \sum \alpha_i x_i$. $\quad \alpha \in \{0,1\}^n$.

$\{0,1\} \longrightarrow \{\pm 1\}$

$b \to (-1)^b$

$\chi_\alpha(x) = (-1)^{\ell_\alpha(x)} = (-1)^{\sum \alpha_i x_i}$

$\mathcal{F} = \{ f : \{0,1\}^n \to \mathbb{R} \}$, 2^n-dimensional $\mathbb{R}$-space.

Equip w/ inner product

$$\langle f, g \rangle = \mathbb{E}[f(x) g(x)]$$

<u>Observation:</u> $\{\chi_\alpha\}$ – form an orthonormal basis.

<u>Pf:</u> $\forall \alpha$, $\mathbb{E}_x [\chi_\alpha(x)] = 0$ if $\alpha \neq 0^n$

$\forall \alpha \neq \beta$ $\mathbb{E}_x [\chi_\alpha(x) \chi_\beta(x)] = \mathbb{E}_x [\chi_{\alpha \oplus \beta}(x)]$

$= 0$

$$\langle \chi_\alpha, \chi_\beta \rangle = \begin{cases} 0 & \text{if } \alpha \neq \beta \\ 1 & \text{if } \alpha = \beta \end{cases}$$

Any f can be written as

$$f = \sum_\alpha \hat{f}_\alpha \chi_\alpha$$

where $\underbrace{\hat{f}_\alpha = \langle f, \chi_\alpha \rangle}_{\text{Fourier coefficients}}$

$$\langle f, g\rangle = \sum_\alpha \hat{f}_\alpha \cdot \hat{g}_\alpha \quad \text{(Plancherel's identity)}$$

$$\langle f, f\rangle = \sum_\alpha \hat{f}_\alpha^2 \quad \text{(Parseval's identity)}.$$

For Boolean fns (ie, $f(x) \in \{\pm 1\}$
$\langle f, f\rangle = 1$, $\sum \hat{f}_\alpha^2 = 1$)

BLR^f: 1. Pick $x, y \in_R \{0,1\}^n$
2. Query f at $x, y, x+y$
3. Accept if $f(x+y) = f(x) \cdot f(y)$.

Analyse using Fourier.

$$\begin{aligned}
\delta(f) &= \min_\alpha \delta(f, \chi_\alpha) \\
&= \min_\alpha \Pr_x [f(x) \neq \chi_\alpha(x)] \\
&= \min_\alpha \left(1 - \Pr_x [f(x) = \chi_\alpha(x)]\right) \\
&= \min_\alpha \left(1 - \mathbb{E}_x [\mathbb{1}(f(x) \cdot \chi_\alpha(x) = 1)]\right) \\
&= \min_\alpha \left\{1 - \mathbb{E}_x \left[\left(\frac{1 + f(x)\chi_\alpha(x)}{2}\right)\right]\right\} \quad \text{Arithmetization} \\
&= \min_\alpha \left\{\frac{1 - \mathbb{E}_x [f(x)\chi_\alpha(x)]}{2}\right\} \\
&= \min_\alpha \left\{\frac{1 - \hat{f}_\alpha}{2}\right\} = \frac{1 - \max_\alpha \hat{f}_\alpha}{2}
\end{aligned}$$

Claim: $\delta(f) = \dfrac{1 - \max_\alpha \hat{f}_\alpha}{2}$.

$$\varepsilon(f) = \Pr_{x,y}\left[f(x)f(y)f(x+y) \neq 1\right]$$

Arithmetization

$$= \mathbb{E}_{x,y}\left[\frac{1 - f(x)f(y)f(x+y)}{2}\right]$$

$$= \frac{1}{2} - \frac{1}{2}\mathbb{E}_{x,y}\left[f(x)f(y)f(x+y)\right]$$

$$\mathbb{E}_{x,y}\left[f(x)f(y)f(x+y)\right] = \mathbb{E}_{x,y}\left[\sum_\alpha \hat{f}_\alpha \chi_\alpha(x) \sum_\beta \hat{f}_\beta \chi_\beta(y) \sum_\gamma \hat{f}_\gamma \chi_\gamma(x+y)\right]$$

$$= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \,\mathbb{E}_{x,y}\left[\chi_\alpha(x)\chi_\beta(y)\chi_\gamma(x+y)\right]$$

$$= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \,\mathbb{E}_{x,y}\left[\chi_\alpha(x)\chi_\beta(y)\chi_\gamma(x)\chi_\gamma(y)\right]$$

$$= \sum_{\alpha,\beta,\gamma} \hat{f}_\alpha \hat{f}_\beta \hat{f}_\gamma \,\mathbb{E}_{x}\left[\chi_\alpha(x)\chi_\gamma(x)\right] \cdot \mathbb{E}_{y}\left[\chi_\beta(y)\chi_\gamma(y)\right]$$

$$= \sum_\alpha \hat{f}_\alpha^3$$

$$\varepsilon(f) = \frac{1}{2}\left(1 - \sum_\alpha \hat{f}_\alpha^3\right)$$

$$\geq \frac{1}{2}\left(1 - \left(\max_\alpha \hat{f}_\alpha\right)\sum_\alpha \hat{f}_\alpha^2\right)$$

$$= \frac{1}{2}\left(1 - \max_\alpha \hat{f}_\alpha\right) \quad \text{(using Booleanity of } f\text{)}$$

$$= \delta(f)$$

Lemma [BCHKS]: $\delta(f) \leq \varepsilon(f)$.

Part II: Constant-query exponential-sized PCP for CIRCUIT-SAT.

[Arora-Lund-Motwani-Sudan-Szegedy 92].

Circuit SAT:

Input: Circuit C.

Goal: To check if $\exists z$, $C(z) = 1$

Assumptions: C - 1. AND, NOT gates of arity 2.

2. # gates (input, output, internal) $= n$.

3. $\exists z \in \{0,1\}^n$ (setting to all gates) that respects the functionality of all gates.

(constraint) $P_i : \{0,1\}^n \to \{0,1\}$

$\mathbb{F}_2^n \to \mathbb{F}_2$

$$P_i(z) = \begin{cases} z_i - 1 & \text{if } i \text{ - output gate} \\ z_i - (1-z_j) & \text{if } i \text{ - NOT gate whose input is the gate } j \\ \end{cases}$$

$$\left\{\begin{array}{lll} z_i - z_j z_k & & \text{if } i\text{- AND gate w/ inputs from gate } j \text{ \& } k \\ 0 & & \text{if } i\text{- input gate} \end{array}\right.$$

Check: $\exists\, z \in \mathbb{F}_2^n$, $\forall\, i$ gates $P_i(z) = 0$.

Goal: Write "z" in a PCP-format that allows for local checking

PCP = Encoding of z using an ECC which is locally testable.

Code - Hadamard / Walsh-Hadamard (WH) code.

$$WH: \{0,1\}^k \to \{0,1\}^{2^k}$$

$$z \longmapsto (\ell_\alpha(z))_{\alpha \in \mathbb{F}_2^k}$$

Assume: P_i - has only linear constraints.

π - "$WH(z)$" . $f: \{0,1\}^n \to \{0,1\}$

<u>Verifier Test:</u> $f = \ell_z$.

1. Linearity Test:

(a) Pick $x, y \in_R \mathbb{F}_2^n$

(b) Query f at $x, y, x+y$

(c) Reject if $f(x+y) \neq f(x) + f(y)$.

2. Circuit Consistency Test.

(a). Pick $a_1 \ldots a_n \in_R \{0,1\}$

$$P(z) = \sum a_i \cdot P_i(z)$$

$$= \ell_\alpha(z) + b$$

(b). Pick $x \in \{0,1\}^n$

Query f at x, $x+\alpha$

Reject if $f(x+\alpha) - f(x) + b \neq 0$.

<u>Soundness Claim:</u> Assuming P has no quadratic constraints.

If $\Pr[\text{Verifier}^f \text{ acc}] \geq 1-\delta$

$$\Downarrow$$

$\exists\ z \in \{0,1\}^n$ s.t (i) f is δ-close to $WH(z)$
ℓ_z

(ii) Z is a satisfying assignment to C.

Handle Quadratic Constraints:

$$z \longrightarrow \ell_z = WH(z)$$

Quadratic Encoding

$$\{0,1\}^k \to \{0,1\}^{2^{k^2}}$$

$$z \mapsto \text{quad}_z$$

where $\text{quad}_z : \{0,1\}^{k^2} \to \{0,1\}$)

$$M \mapsto \sum_{i,j} M_{ij} z_i z_j)$$

Obs:

(1) quad_z is linear (ie, $\text{quad}_z(M) + \text{quad}_z(N) = \text{quad}_z(M+N)$).

(2) $f : \{0,1\}^{k^2} \to \{0,1\}$ is linear if there exists a matrix B st

$$f(M) = \langle M, B \rangle$$

But if $f = \text{quad}_z$ $B = zz^T$

Qn: Given $f : \{0,1\}^n \to \{0,1\}$ $f = \ell_z$

$F : \{0,1\}^{n^2} \to \{0,1\}$. $F = \ell_B$

need to check $B = zz^T$

Suggestion: Pick $x, y \in_R \{0,1\}^n$

$$x^T B y = x^T z z^T y$$

$$\langle xy^T, B\rangle = \langle x, z\rangle \langle y, z\rangle$$
$$F(xy^T) = f(x)\cdot f(y)$$

Quadratic Correlation Test^{f,F} $\left(f: \{0,1\}^n \to \{0,1\},\ F: \{0,1\}^{n^2} \to \{0,1\}\right)$

1. Pick $x, y \in_R \{0,1\}^n$
2. Pick $N \in_R \{0,1\}^{n^2}$
3. Query F at $xy^T + N$ & N
   f at x, y.
4. Accept if $F(xy^T+N) - F(N) = f(x)\cdot f(y)$.

Completeness: $f = \ell_z$ & $F = \ell_{zz^T}$

$\Downarrow$

$\Pr[\text{Quad Cor Test}^{f,F} \text{ acc}] = 1.$

Soundness: f is δ-close to ℓ_z
$(\delta \in (0, 1/4))$. & F is δ-close to ℓ_B and $B \neq zz^T$

$\Downarrow$

$\Pr[\text{Quad Corr Test}^{f,F} \text{ acc}] \leq \frac{3}{4} + 4\delta.$

__Pf:__ BAD events

(1) $f(x) \neq \ell_z(x)$ — $\leq \delta.$

(2) $f(y) \neq \ell_z(y)$ $\leq \delta$

(3) $F(N) \neq \ell_B(N)$ $\leq \delta$

(4) $F(xy^T + N) \neq \ell_B(xy^T + N)$ $\leq \delta$.

(5) $\langle xy^T, B \rangle = \langle x, z \rangle \cdot \langle y, z \rangle$

ie, $\Pr_{x,y}\left[x^T (B - zz^T) y = 0 \right]$

Let $C = B - zz^T$

Suppose $C_{ij} \neq 0$

$\Pr_{x,y}\left[x^T C y = 0 \right]$

W/ prob $1/2$ over choice of y

Cy is a non-zero vector.

Conditioned on that

w/ prob $1/2$ over x, $x^T C y \neq 0$.

PCP verifier:

Input: Circuit C.

Goal: $\exists z$, st $C(z) = 1$

Proof: $f : \{0,1\}^n \to \{0,1\}$ $(f = \ell_z)$

$F: \{0,1\}^{n^2} \to \{0,1\}$ $\quad (F = quad_2)$

Verifier$^{f,F}(C)$:

1. Linearity Test^{f,F}
   (a) Pick $x, y \in_R \{0,1\}^n$, check if $f(x)+f(y) = f(x+y)$
   (b) Pick $M, N \in_R \{0,1\}^{n^2}$, check if
   $$F(M) + F(N) = F(M+N)$$

2. Quad-Correlation Test^{f,F}
   (a) Pick $x, y \in_R \{0,1\}^n$, $N \in \{0,1\}^{n^2}$
   (b) Accept if $F(xy^T + N) - F(N) = f(x) f(y)$.

3. Circuit-Consistency test^f
   (a) Pick $a_1 \dots a_n \in_R \{0,1\}$
   $$P(z) = \sum a_i \cdot P_i(z)$$
   $$= \langle B, zz^T \rangle + \langle \alpha, z \rangle + b.$$
   (b) Query f at α, $\alpha + r$ $\quad (r \in_R \{0,1\}^n$
   $\quad F$ at B, $B+N$ $\quad N \in_R \{0,1\}^{n^2})$

   accept if $F(B+N) - F(N) + f(\alpha + r) - f(r) + b = 0$.

Completeness: If $C(Z)=1$, then $f = \ell_z$, $F = quad_z$
$$\Pr[\mathrm{Ver}^{f,F}(C) = acc] = 1.$$

Soundness: $\exists\, \delta_0 \in (0,1)$, $\forall\, \delta \leq \delta_0$

$$\Pr[\mathrm{Ver}^{f,F} \text{ acc}] \geq 1-\delta$$

$$\Downarrow$$

$\exists\ C(z)=1$, s.t. f is δ-close to ℓ_z

& F is δ-close to quad_z.

Proof.

BAD events.

(1) f is δ-far from being linear $\leq 1-\delta$

$\exists\, z,\ \delta(f, \ell_z) \leq \delta.$

(2) F is δ-far from being linear $\leq 1-\delta$

$\exists\, B,\ \delta(F, \ell_B) \leq \delta$

(3) $B \neq zz^T$ $\leq \frac{3}{4} + 4\delta$

Assume $B = zz^T$ $\leq 1-\delta$

$(\delta \leq 1/20)$

(4) $C(z) \neq 1$ $\leq \frac{1}{2} + 4\delta$

$\leq 1-\delta.$

$(\delta \leq 1/20)$

$\boxtimes$..

PCP: Verifier

#queries = 9

$\#\text{randomness} = O(n^2)$

$$\text{CIRCUIT-SAT} \in \text{PCP}_{1,\, 1-\frac{1}{20}}[O(n^2),\, 9]$$