Polynomial identity testing (PIT)

- PIT is the following algorithmic problem:

  Given an arithmetic circuit $C(\bar{x})$, over a ring R, test whether C is identically zero.

  (We want an algorithm that runs in time polynomial in size(C).)

- We will focus on the case of R being a field $\mathbb{F} = \mathbb{Q}$ or $\mathbb{F}_q$.

Theorem [Schwartz, Zippel et al] PIT $\in$ CoRP.

Proof:

- Let $C(\bar{x})$ be the given circuit of size s, over $\mathbb{F} = \mathbb{F}_q$.
- $\Rightarrow \deg C < s^s$.
- We could assume $|\mathbb{F}| > 2 \cdot s^s$, otherwise we can use an appropriate field extension.

(Fast constructions are known due to [Adleman, Lenstra '86])

- The algorithm is simply a random evaluation:

    0) Pick an $S \subseteq \mathbb{F}$ of size $2 \cdot s^s$.

    1) Pick a <u>random</u> $(a_1, \ldots, a_n) \in S^n$.

    2) If $C(\bar{a}) = 0$ then OUTPUT <u>Zero</u>
    else " <u>nonZero</u>.

- It has been proved before (in an Assignment) that: if C is a nonzero polynomial then

$$\Pr_{\bar{a} \in S^n}[C(\bar{a}) \neq 0] > 1 - \frac{\deg C}{2s^s} > \frac{1}{2}.$$

- Clearly, $C(\bar{a})$ can be computed in time $\mathrm{poly}(s, \lg|\mathbb{F}|)$.

- In the case when $\mathbb{F} = \mathbb{Q}$, $C(\bar{a})$ may be doubly-exp. large!

In that case, we pick a random prime p & evaluate $C(\bar{a}) \bmod p$.

(Exercise: compute the error probability)

no mistake on identities (coRP)

- Thus, in all cases PIT has a randomized poly-time algorithm. $\square$

- Note that in the above algorithm the specifics of the circuit C were not used. (Only the size bound was needed.)

- Such an algorithm is called a blackbox identity test.

over a ring extension & mod primes

(One can only evaluate a blackbox.)

Definition: For a family $\mathcal{C}$ of n-variate circuits of size s, a hitting-set $\mathcal{H} \subseteq \mathbb{F}^n$ is a poly(s)-sized set of points such that: If $C \in \mathcal{C}$ is nonzero then $\exists \bar{\alpha} \in \mathcal{H}$, $C(\bar{\alpha}) \neq 0$.

Or, $\mathcal{H}$ hits $\mathcal{C}$.

existence of small hitting-sets

Lemma: Let $S \subseteq \mathbb{F}$ be of size s^{3s} & $\mathcal{C}$ be the family of size-s circuits, n-variate, over $\mathbb{F}_q$.
Then, a random $\bar{a} \in S^n$ hits $\mathcal{C}$.

Proof:

- $\Pr_{\bar{a} \in S^n} [\exists 0 \neq C \in \mathcal{C},\ C(\bar{a}) = 0]$

$$\leq |\mathcal{C}| \cdot \frac{s^s}{|S|} < q^s \cdot \frac{s^s}{s^{3s}} \leq s^{-s}.$$ (Assume $q < s$)

$$\Rightarrow \Pr_{\bar{a}} [\forall 0 \neq C \in \mathcal{C},\ C(\bar{a}) \neq 0] > 1 - s^{-s}.$$

□

OPEN (Derandomization): Can a hitting-set be computed in det. poly-time ?

- Given $\mathcal{H}$, by interpolation, we can find polynomials $(p_1(y), \ldots, p_n(y)) =: \bar{p}(y)$ such that their first few values, on fixing y, give us the points in $\mathcal{H}$.
Also, $\deg p_i \leq |\mathcal{H}|$.

- This motivates us to define arithmetic analogs of prgs (pseudorandom generators).

Defn: $\{(\underbrace{p_1^n(y), \ldots, p_n^n(y)}_{=: f(n)}) \mid n \in \mathbb{N}\} =: f$ is called an $s(n)$-hsg against $\mathcal{C}$, if

hitting-set generator

- each $p_j^n(y)$ has deg $\leq s(n)$ & is computable in time poly$(s(n))$,
- for any nonzero $C \in \mathcal{C}$ on n-variables,

$$C(p_1^n(y), \ldots, p_n^n(y)) \neq 0 .$$

depending on $\mathcal{C}$ one might want to go $\rightarrow$ mod $g(y)$.

Derandomization Qn: Do efficient hsg exist?

- Apart from being a fundamental qn., this is also related to proving lower bounds (close to VP $\neq$ VNP).

- A PIT algorithm would imply some lower bound:

Thm [Kabanets, Impagliazzo '03]: PIT $\in$ P $\Rightarrow$ NEXP $\not\subseteq$ P/poly or VNP $\neq$ VP.

- We will skip this proof & instead focus on the implications of an efficient hsg (& a converse!).

Thm [Agrawal '05]: Let f be an $s(n)$-hsg against $\mathcal{C}$. Then, there is a multi-linear polynomial computable in poly($s(n)$)-time that is not in $\mathcal{C}$.

Assume $s(n) \leq 2^{n/2}$.

Proof:

- Consider $f(n) = (p_1(y), \ldots, p_n(y))$ for a large enough n.
- Define $\ell(n) := \lg s(n)$ & $m := 2\ell \leq n$.
- The idea is to consider an annihilating polynomial $q(X_1, \ldots, X_m)$ for $(p_1(y), \ldots, p_m(y))$.

- In particular, $q(\overline{x}) = \sum_{S \subseteq [m]} c_S \cdot X_S$

s.t. $c_S \in \mathbb{F}$ & $q(p_1(y), \ldots, p_m(y)) = 0$.

- This sets up a linear system in the unknowns c_S:

$$\#\text{unknowns} = 2^m,$$
$$\#\text{equations} \leq m \cdot s$$

$\Rightarrow$ $\exists$ a nontrivial solution $(\because 2^m > ms)$.

- Moreover, the solution can be computed in time $\text{poly}(2^m) = \text{poly}(s(n))$.

- Since q vanishes on $f(n)$, we deduce that $q \notin \mathcal{C}$. (Also, q is m-var. & computable in $2^{O(m)}$-time) $\square$

If $\mathcal{C} = VP$, $\Rightarrow$ then q_m is $2^{\Omega(m)}$-hard

– Is there a converse to this?

Does "$VNP \neq VP$ $\Rightarrow$ efficient hsg for VP"?

- We can prove a weaker claim (both strengthening the premise & weakening the conclusion!).

<u>Thm</u> [KI'03, Agrawal-Vinay'08]: Let $\{q_m\}_{m\geq 1}$ be a multilinear polynomial family, computable in E, that is not computable by subexponential sized arithmetic circuits.

Then, there is an efficient variable reduction for VP circuits, from n to $O(\lg n)$ variables, that preserves nonzeroness.

Recall: deg & s are poly(n).

(This implies an $n^{O(\lg n)}$-hsg for VP circuits.)

<u>Proof</u>:

- Let C be a nonzero circuit of size $s = s(n)$ computing a polynomial of deg $\leq s$ (wlog).
- We wish to reduce its variables,

preserving the nonzeroness.

We will utilize the q_m's, for "small" m, to feed into C.

- For this we need a set-family called Nisan-Wigderson designs.

Defn: Let $\ell > n > d$. A collection $\mathcal{I} = \{I_1, \ldots, I_m\}$ of n-size subsets of $[\ell]$ is an (ℓ, n, d)-design if: $|I_j \cap I_k| \leq d$, $\forall\, j \neq k \in [m]$.

Lemma [NW'94]: There is an algorithm that on input (ℓ, n, d), $(\ell > 10n^2/d)$, outputs an (ℓ, n, d)-design $\mathcal{I}$ having $m \geq 2^{d/10}$ subsets, in time $2^{O(\ell)}$.

Pf:

- A greedy approach works.
- Details skipped. $\square$

- Say, C has n variables $z_1, \dots, z_n$.
- Let $\mathcal{I} = \{S_1, \dots, S_n\}$ be a $(c\lg n, d\lg n, 10\lg n)$-design, for suitable constants $c > d > 10$.

Note that by the previous Lemma, $\mathcal{I}$ can be constructed in $\mathrm{poly}(n)$-time.

- Now we map $\{z_1, \dots, z_n\}$ to $\bar{x} := \{x_1, \dots, x_{c\lg n}\}$ as follows:

$$z_i \mapsto p_i := q_{d\lg n}(\bar{x}_{S_i})$$

where $\bar{x}_{S_i}$ is the $(d\lg n)$-tuple given by the indices in S_i.

<u>Claim</u>: $C(p_1, \dots, p_n) \neq 0$.

<u>Pf</u>:
- Suppose not.
- As $C(\bar{z}) \neq 0$ but $C(\bar{p}) = 0$, there is a $j \in [n]$ s.t. $C(p_1, \dots, p_j, z_{j+1}, \dots, z_n) = 0$ but $C(p_1, \dots, p_{j-1}, z_j, \dots, z_n) \neq 0$.
- $\Rightarrow (z_j - p_j) \mid C(p_1, \dots, p_{j-1}, z_j, \dots, z_n)$.

• Now we can fix $z_{j+1}, \ldots, z_n$ & the x_i's that do not occur in p_j to random values from the field.

• This reduces us to the case:

$$(z_j - p_j) \mid C'(p_1'(\bar{x}_{S_1 \cap S_j}), \ldots, p_{j-1}'(\bar{x}_{S_{j-1} \cap S_j}), z_j) \neq 0.$$

• Note that $|S_k \cap S_j| \leq 10 \lg n$, for $k \neq j$.

$\Rightarrow$ The above circuit $C'(p_1', \ldots, z_j)$ has size $< s + n^{11}$.

(As p_1' etc. can be written as a sum of $2^{10 \lg n}$ monomials.)

• We could now invoke Kaltofen ('89) VP circuit factorization algorithm (in the blackbox setting!).

$\Rightarrow$ p_j has a VP circuit of size s^e, where e is a constant independent of d & c.

- Since $p_j = q_{d\lg n}(\bar{x}_{S_j})$ was assumed

complexity $2^{\Omega(d\lg n)}$ → to be a "hard" polynomial, we can deduce a contradiction by taking d suitably larger than e.

$\Rightarrow C(p_1, \dots, p_n) \neq 0.$

$\square$

- Note that $C(\bar{p})$ is $(c\lg n)$-variate & $\deg = O(s\lg n)$.

- Thus, $C(\bar{p})$ has sparsity at most $(s\lg n)^{O(\lg n)} = n^{O(\lg n)}$.

- Finally, one needs to design an efficient hsg for sparse polynomials.

(When the arity m is small then one can simply take $[0 \dots \deg]^m$ as a hitting-set.)

PIT for shallow circuits.

- Suppose we solve PIT for the depth-4 or depth-3 models.
  What will that imply for PIT for VP ?

- Let us consider a very special depth-4, called diagonal depth-4 circuits:

$$C(x_1,..,x_n) = \sum_{i\in[k]} f_i^{\,d}$$

where f_i is a sparsity ω polynomial in $\mathbb{F}[x_1,..,x_n]$ of deg $\leq \delta$.

Thm [Agrawal-Vinay '08]: If there is an efficient hsg against diagonal depth-4 model (even assuming $n, \delta = O(\lg s)$ & $d = \omega(1)$), then there is an efficient variable reduction for VP circuits, from n to $O(\lg n)$, that preserves nonzeroness.

ch $\mathbb{F}=0$ is assumed

Proof:

- Let f be a poly(s)-hsg against the said diagonal depth-4 of size s.

consider an annihilator of $f(1), \ldots, f(O(\lg s))$

- By the "hsg ⇒ hard poly." theorem, we get a family of multilinear polynomials $\{q_m\}_{m\geq 1}$ that is computable in $2^{O(m)}$ time but requires diagonal depth-4 of size $2^{\Omega(m)}$.

Claim: $\{q_m\}_{m\geq 1}$ requires VP circuits of size $2^{\Omega(m)}$.

Pf:
- Let there be a circuit computing $q_m(x_1, \ldots, x_m)$ in size $s = s_m$ & degree $d = d_m$. (with $s_m = 2^{o(m)}$)
- By the depth-reduction we have a circuit C in $\Sigma\Pi^{5^t}\Sigma\Pi^{m/2^t}$ of size $\binom{s+5^t}{5^t} + s \cdot \binom{m+d/2^t}{d/2^t}$. for any $t \in [\lg d_m]$. (Note: $d_m = m$.)

- This can be seen by first bringing the circuit to $O(\lg m)$-depth & product-fanin 5. Moreover, each child of a product gate has degree at most half that of the product.

- Now we divide the circuit in two parts – <u>top part</u> having t product layers & the bottom part.
- We convert each of these parts to a depth-2 circuit ($\Sigma\Pi$).
- The top part gives an s-variate, $\deg \leq 5^t$ polynomial.
- The bottom part gives several $\Sigma\Pi$ circuits, each m-variate & $\deg \leq m/2^t$.

- Combining the two parts we get a $\Sigma\Pi^{5^t}\Sigma\Pi^{m/2^t}$ circuit of size $\binom{s+5^t}{5^t} +$ $s \cdot \binom{m+m/2^t}{m/2^t}$.

- Pick $t = \log_5 \sqrt{m/\lg s} = \omega(1)$. $\left[\Rightarrow 2^t = \left(\frac{m}{\lg s}\right)^{1/2\lg 5}.\right]$

$\Rightarrow$ size $= s^{O(5^t)} + s \cdot (2^t)^{O(m/2^t)} = 2^{O(\sqrt{m \lg s})} +$
$s \cdot 2^{O(mt/2^t)} = 2^{o(m)}$.

$\Rightarrow q_m$ has a $\Sigma\Pi^{\omega(1)}\Sigma\Pi^{m/\omega(1)}$ circuit of size $2^{o(m)}$.

- By Fischer's trick this can be immediately written as a $\Sigma\wedge^{\omega(1)}\Sigma\Pi^{m/\omega(1)}$ circuit of size $2^{o(m)}$.
- This contradicts the hardness of q_m.

$\Rightarrow \{q_m\}_{m\geq 1}$ has VP complexity $2^{\Omega(m)}$ as well. $\square$

- Now by "hard poly. $\Rightarrow$ hsg" theorem, we can use q_m to design an efficient $n \mapsto O(\lg n)$ variable reduction that preserves the nonzeroness of VP circuits.

$\square$

Corollary: An efficient hsg for $\Sigma^s \wedge^{\omega(1)} \Sigma^s \Pi^{O(\lg s)}$ circuits in $O(\lg s)$ variables over $\mathbb{F}$ (of char. $=0$)
$\Rightarrow n^{O(\lg n)}$-hsg for VP (over $\mathbb{F}$).

[AGS'18] Variables can be bootstrapped in PIT.

Some PIT algorithms

- PIT results are known only for very special cases.

- The motivating cases for PIT techniques have been –
  $\Sigma\Pi$ (or sparse), $\Sigma\wedge\Sigma$ (diagonal depth-3), set-multilinear $\Sigma\Pi\Sigma$ (& ROABP), $\Sigma^k\Pi\Sigma$ (bounded top fanin depth-3), occur-k depth-4.

Prg for $\Sigma\Pi$ (sparse PIT)

- Let C be a $\Sigma\Pi$ circuit in $\mathbb{F}[x_1, \ldots, x_n]$.

- size(C) constitutes n, degree $<d$ & the number of monomials s in the polynomial C.

- PIT is trivial if C is given explicitly.

- However, when C is a blackbox, the PIT becomes more interesting.

- <u>Idea</u>: Kronecker map $x_i \mapsto t^{d^i}$, followed by polynomial division.

▷ For $\varphi : x_i \mapsto t^{d^i}$, $i \in [n]$, & a polynomial $f(\bar{x})$ of $\deg < d$, we have: $f \neq 0 \Rightarrow \varphi(f) \neq 0$.

<u>Proof</u>:

- φ sends a monomial $\bar{x}^{\bar{e}}$ to $t^{\bar{e}\cdot\bar{d}}$, where $\bar{d} := (d, d^2, \ldots, d^n)$.
- Since $\bar{e} \in [0 \ldots d-1]^n$, $\bar{e}\cdot\bar{d}$ can be seen as a d-ary number with digits $\bar{e}$.

$\Rightarrow$ such $\bar{e}$ are mapped to distinct values. $\square$

- We can reduce the degree by going modulo $t^r - 1$, for "small" prime r's.

▷ $t^{\bar{e}\cdot\bar{d}} \equiv t^{\bar{e}'\cdot\bar{d}} \bmod \langle t^r - 1\rangle$ iff
$\bar{e}\cdot\bar{d} \equiv \bar{e}'\cdot\bar{d} \pmod r$ iff
$(\bar{e}-\bar{e}')\cdot\bar{d} \equiv 0 \pmod r$.

- Note that $|(\bar{e}-\bar{e}')\cdot\bar{d}| < 2d^{n+1}$.
- Thus, if $\bar{e}\neq\bar{e}'$ then $(\bar{e}-\bar{e}')\cdot\bar{d}$ has at most $\lg 2d^{n+1}$ prime factors.

- By the prime number theorem, there are $> \lg 2d^{n+1}$ many primes smaller than $\tilde{O}(n\lg d)$.

▷ Thus, if $\bar{x}^{\bar{e}_1}, \ldots, \bar{x}^{\bar{e}_s}$ are distinct monomials then $\bar{e}_1\cdot\bar{d} \not\equiv \bar{e}_i\cdot\bar{d} \pmod r$, for $i\in[2\ldots s]$, for some prime $r = \tilde{O}(s\cdot n\lg d)$.

<u>Thm</u>: C has a blackbox PIT algo. that takes poly$(sn\lg d)$-time.

Hsg for tiny depth-3 suffices

- It was shown, in the last lecture, that efficient hsg for a "tiny" case of $\Sigma\wedge\Sigma\Pi$ will imply quasi-poly for VP.

This, of course, can also be brought down to depth-3.

→ brute-force is $s^{O(\lg s)}$.

Theorem: An efficient hsg for $\Sigma\Pi\Sigma^{o(\lg s)}$ size-s circuits in $O(\lg s)$ variables over $\mathbb{F}$ ($\mathrm{ch}\,\mathbb{F}=0$) $\Rightarrow$ $n^{O(\lg n)}$-hsg for VP.

Proof:

- As we have seen in the previous proof: an efficient hsg gives us a multilinear polynomial family $\{q_m\}_{m\geq 1}$ that requires $\Sigma\Pi\Sigma^{o(m)}$ circuits of size $2^{\Omega(m)}$.
- As before, if $\{q_m\}_{m\geq 1}$ has a VP circuit C of size $s = 2^{o(m)}$, then it can be

reduced to a $\Sigma\Lambda^{\omega(1)}\Sigma\Pi^{m/\omega(1)}$ circuit of size $2^{o(m)}$,

(Fischer's trick)

- which can be further reduced to a $\Sigma\Lambda^{\omega(1)}\Sigma\Lambda^{m/\omega(1)}\Sigma$ circuit of size $2^{o(m)}$,
- which, by the duality trick on the top Λ-gate & by factorization, converts to $\Sigma\Pi\Sigma^{o(m)}$ of size $2^{o(m)}$.

$\Rightarrow$ contradiction to $\{q_m\}_m$'s hardness.

$\Rightarrow \{q_m\}_m$ requires VP circuits of size $2^{\Omega(m)}$.

$\Rightarrow$ $n^{\omega(\lg n)}$-hsg for VP. $\square$

- Thus, all we need for PIT is to understand "tiny" depth-3 or tiny diagonal depth-4.

- How about diagonal depth-3?

Some results are known, but not completely understood.