CS711 Course Project
Randomness Efficient Identity Testing

Vijaykrishna, Satyabrata, Mohan

October 31, 2020

Abstract

We are given an n-variate polynomial with m monomials, total degree δ. This paper by Spielman and Klivans performs randomized poly-time identity testing with $O(\log(mn\delta))$ random bits as opposed to previous work that needed $\Omega(n)$ random bits. The techniques developed also help in designing deterministic poly-time algorithms for sparse identity testing and sparse interpolation.

1 Problem definition and setting

We are given an n-variate polynomial P with total degree δ and number of monomials m. For two vectors d and a, let $\langle d, a\rangle$ denote the inner product. Also, x^d denotes the monomial $x_1^{d_1} \cdots x_n^{d_n}$. The algorithms in the paper will work in the *blackbox* setting. Below is the main theorem of the paper:

Theorem 1. *Let $|\mathbb{F}| \geq (n\delta/\varepsilon)^6$ and $P \in \mathbb{F}[x_1, \cdots, x_n]$ with m monomials, total degree δ given as a blackbox. There is a randomized algorithm that tests if the polynomial is zero with success probability $1 - \varepsilon$ using $O(\log(mn\delta/\varepsilon))$ random bits and runs in time polynomial in n, $\log(\delta)$, $\log(\varepsilon^{-1})$. Moreover, the algorithm queries the polynomial at points with bit length $O(\log(n\delta/\varepsilon))$.*

In order to prove this we will have to prove a few other important lemmas along the way, which we do in the coming sections.

2 Reducing to a univariate polynomial

The first step is to reduce the problem to a univariate identity testing problem, which would then be more tractable. Let P be denoted as

$$\sum_{j=1}^{m} c_j x^{d^{(j)}}$$

To reduce it to a univariate, one could try to substitute $x_i = y^i$, however this can make the result zero (e.g $x_1 x_2 - x_3$). In general, let a be a vector such that the substitution done is $x_i = y^{a_i}$. Upon this substitution, a monomial in P changes as follows:

$$x^d \to y^{\langle d,a \rangle}$$

We will define a collection of t vectors $a^{(1)}, \cdots, a^{(t)}$ such that a good fraction $(1-\varepsilon)$ of them yield non-zero polynomials when used for the substitution. For this purpose, we define

$$a_i^{(k)} = k^{i-1} \mod p$$

where p is a prime slightly larger than t.

Lemma 2. *Let p be a prime larger than t and δ. Then, for all j*

$$\Pr_{1 \le k \le t}[\forall j' \neq j, \langle d^{(j)}, a^{(k)} \rangle \neq \langle d^{(j')}, a^{(k)} \rangle] \geq 1 - mn/t$$

Proof. Recall that the entries of $d^{(j)}$ are in $\{0, \cdots, \delta\}$. For a particular j'

$$\begin{aligned}
\Pr[\langle d^{(j)}, a^{(k)} \rangle = \langle d^{(j')}, a^{(k)} \rangle] &= \Pr[\langle d^{(j)} - d^{(j')}, a^{(k)} \rangle = 0] \\
&= \Pr\left[\sum_{i=1}^{n} (d_i^{(j)} - d_i^{(j')}) a_i^{(k)} = 0\right] \\
&\leq \Pr\left[\left(\sum_{i=1}^{n} (d_i^{(j)} - d_i^{(j')}) a_i^{(k)}\right) \mod p = 0\right] \\
&= \Pr_{1 \le k \le t}\left[\left(\sum_{i=1}^{n} e_i (k^{i-1} \mod p)\right) \mod p = 0\right] \\
&\leq n/t
\end{aligned}$$

where $e_i = (d_i^{(j)} - d_i^{(j')}) \bmod p$. The last line is basically computing a univariate degree n polynomial over $\mathbb{F}_p$ and asking for the fraction of zeroes (recall that p is greater than t and δ). So we can bound it by n/t. Union bound over all j' gives the statement of the lemma as required. $\square$

3 Reducing the degree

Notice that from lemma 2, we can set $t = mn/\varepsilon$ and p to a prime less than $2 \cdot \max(\delta, mn/\varepsilon)$. On the substitution $x_i = y^{a_i^{(k)}}$, we will get a non-zero univariate $P'(y)$ with probability $1-\varepsilon$. The degree of P' is at most $p\delta$, since each entry of $a^{(k)}$ is in $\{0, \cdots, p\}$ and total degree of P is δ. The issue is that m can be very large (as much as $O(n^\delta)$), so $\deg(P')$ can be very large which is not desirable for the further processing of P'. This section deals with modifying the procedure to give a smaller degree univariate. First, an easy modification of lemma 2 gives:

Lemma 3. *Let $d^{(1)}, \cdots, d^{(m)}$ be distinct vectors with entries in $\{0, 1, \cdots, \delta\}$ and p be a prime greater than t and δ. Then*

$$\Pr_{1 \leq k \leq t}[\langle d^{(j)}, a^{(k)} \rangle \text{ are distinct for all } j] \geq 1 - m^2 n/t$$

Next we will prove theorem 5. For this, we will use the following isolation lemma without proof.

Lemma 4. *Let C be any collection of distinct linear forms in variables $z_1, \cdots, z_l$ with coefficients in $\{0, \cdots, K\}$. Let $z_1, \cdots, z_l$ be chosen uniformly at random from $\{0, \cdots, Kl/\varepsilon\}$. Then with probability at least $1-\varepsilon$, there is a unique form of minimal value.*

Theorem 5. *There is a randomized polynomial time algorithm which maps a non zero polynomial $P \in \mathbb{F}[x_1, \cdots, x_n]$ of total degree at most δ with m monomials to a non-zero univariate polynomial of degree $O(n^6\delta^6/\varepsilon^5)$. The algorithm succeeds with probability $1 - \varepsilon$ and uses* $\mathrm{polylog}(mn/\varepsilon)$ *random bits.*

Proof. Let us first get a vector a satisfying the conditions of lemma 3 (with $\varepsilon/2$). We can represent each element of a using a q-bit number where $q = \log(4m^2n/\varepsilon)$. Let us split these bits into l buckets each of size q/l. Then we can define vectors $\{b_1, \cdots, b_l\}$ with elements in $\{0, \cdots, 2^{q/l}\}$ where b_i is obtained from a by taking the i^{th} block from each element of a.

Let d denote the characteristic vector of any particular monomial. Notice that by construction

$$\langle d, a \rangle = \left\langle d, \sum_{r=0}^{l-1} 2^{rq/l} b_r \right\rangle$$

Also, by the condition of lemma 3, for two different monomials

$$\langle d, a \rangle \neq \langle d', a \rangle$$

The above two observations make it natural to define a linear form in variables $y_1, \cdots, y_r$ associated with each monomial x^d as follows:

$$L_d := \left\langle d, \sum_{r=0}^{l-1} y_r b_r \right\rangle = \sum_{r=0}^{l-1} \langle d, b_r \rangle y_r$$

Note that the coefficients $\langle d, b_r \rangle \in \{0, \cdots, \delta\, 2^{q/l}\}$. These linear forms are distinct for $d \neq d'$ because they differ at the point

$$(y_1, \cdots, y_l) = (1, 2^{q/l}, 2^{2q/l}, \cdots, 2^{(l-1)q/l})$$

Now we can apply lemma 4 on $C = \{L_1, \cdots, L_m\}$ in variables $y_1, \cdots, y_l$ and $K = \delta\, 2^{q/l}$. So if we pick y_i's uniformly at random from $\{0, \cdots, \delta l \cdot 2^{q/l}/\varepsilon\}$, then we get a unique linear form of minimal value with probability $1 - \varepsilon$.

Given all this, it suffices to put the following substitution in the original polynomial: $x_i = h^{z_i}$ where $z = \sum_{r=0}^{l-1} y_r b_r$. This is because any particular monomial x^d will become

$$x^d \to h^{\langle d, z \rangle} = h^{L_d}$$

Since there is a unique minimal valued L_d, $P'(h)$ will have at least this monomial and so it won't become zero. Substituting $l = q/\log(\delta n/\varepsilon)$ gives the bounds stated in the theorem. □

4 Performing the test

From theorem 5, given a non-zero $P \in \mathbb{F}[x_1, \cdots, x_n]$, we have obtained a way to get a univariate $P'(y)$ which is probably non zero. The degree of the univariate as we have seen is $O(n^6 \delta^6/\varepsilon^5)$. So, if we check non-zeroness of $P'(y)$ at y picked randomly from a set of size $O((n\delta/\varepsilon)^6)$, this will work with probability $1 - \varepsilon$.

Of course, we can't feed in y directly, since all we have is a blackbox for $P(x_1, \cdots, x_n)$. However, notice that by construction

$$P'(y) = P(y^{z_1}, \cdots, y^{z_n})$$

So we just feed $x_i = y^{z_i}$ into the blackbox. Since $z_i \leq \deg(P')$, an upper bound for the bit length of x_i is $\mathrm{poly}(n, \delta, \varepsilon^{-1})$. This is not prohibitively large and for purposes of brevity we omit a part that improves this, since the main focus is on the small number of random bits. This gives us the following slightly weaker form of theorem 1.

Theorem 6. *Let* $|\mathbb{F}| \geq \Omega((n\delta/\varepsilon)^6)$ *and* $P \in \mathbb{F}[x_1, \cdots, x_n]$ *with at most* m *monomials, total degree* δ*. There is a randomized algorithm that tests if the polynomial is zero with success probability* $1 - \varepsilon$ *using* $\mathrm{polylog}(mn\delta/\varepsilon)$ *random bits and runs in time polynomial in* $n, \delta, \varepsilon^{-1}$*. Moreover, the algorithm queries the polynomial at points with bit length polynomial in* $n, \delta, \varepsilon^{-1}$*.*

5 Deterministic algorithms

Using the methods presented, we can actually get deterministic algorithms for sparse polynomials i.e when $m, \delta = \mathrm{poly}(n)$.

5.1 Sparse Identity Testing

When $m, \delta = \mathrm{poly}(n)$, notice that the number of possible z's that can be sampled is $\mathrm{poly}(n)$. Therefore, in such cases we don't have to rely on randomness as we can just go over all these z's one by one. This immediately gives a deterministic poly-time identity testing algorithm for sparse polynomials. Also this is independent of the characteristic, which is something previous work depended on.

5.2 Sparse interpolation

We will work with the reals for simplicity. Let us set $t = m^2n + 1$, then there is some $a \in \{a^{(1)}, \cdots, a^{(t)}\}$ satisfying the conditions of lemma 3. On using this a, we get a univariate $P'(y)$ of the form

$$\sum_j c_j y^{\langle d_j, a \rangle}$$

where the $\langle d_j, a \rangle$ are distinct for distinct j. The degree of P' is at most $2m^2n\delta$. So by querying at $2m^2n\delta + 1$ points we can find the c_j's using interpolation. The issue is that from $\langle d_j, a \rangle$ we cannot determine d_j. For this, we do the substitution $x_i = p_i y^{a_i}$, where $p = (p_1, \cdots, p_n)$ is a list of distinct primes. Then the form of $P'(y)$ will be

$$\sum_j c_j p^{d_j} y^{\langle d_j, a \rangle}$$

This is very similar to the previous univariate in terms of the monomials of y. However, the interpolation will now output the coefficients as $c'_j = c_j p^{d_j}$. We already know c_j, so d_j can be obtained by looking at the unique prime factorization of c'_j/c_j.