Quantum Weak Parity Problem

Mohammad Bavarian 6.854 project

December 17, 2012

Abstract

In this work we shall investigate quantum query complexity of weak PARITY problem. Weak Parity is the following natural query problem: What is the minimum number of queries necessary to compute PARITY$(x_1, x_2, x_3, \ldots, x_n)$ on at least $\frac{1}{2} + \epsilon$ fraction of the inputs. For randomized classical or exact quantum machines this relaxed PARITY problem remains as hard as original PARITY problem. Although Bounded error quantum machines needed almost as many queries as classical machines for original PARITY problem, they outperform classical algorithms in this relaxed problem. We will show upper $O(\frac{n}{\sqrt{\log(1/\epsilon)}})$ and lower bound of $\Omega(\max\{\frac{n}{\log(1/\epsilon)}, \sqrt{\log n}\})$. After this, we shall investigate the connection of this problem in the setting of $\epsilon = 2^{-n}$ with important extremal problems over boolean hypercube and we shall give some improvements in that case.

1 Introduction

One of earliest examples in study of black-box quantum problems is the PARITY problem. For computing the PARITY of $X = (x_1, x_2, \ldots, x_n)$ a randomized machine must query all the inputs because even a single unqueried input can change the value of output by itself. A quantum machine, on the other hand, can get away with only $\frac{n}{2}$ queries using Deutsch-Josza algorithm which is optimal by direct application of polynomial method.

Now consider the following scenario: We are given access to the query box to inputs $X = \{x_i\}_{i=1}^n$ and we still hope to compute the PAIRTY of the $\{x_i\}_{i=1}^n$'s. But unfortunately, for one reason or another, we are told that we can only query the box at most k number of times.[1] If $k < \frac{n}{2}$, we cannot hope to compute $PARITY_n$ correctly on all inputs. So the one natural thing we could try to do with our limited number of queries is to try to maximize the number of inputs that we will answer correctly.

[1] Perhaps the box becomes faulty and unstable after k queries.

Problem 1 (Weak Parity Problem- First Formulation) *Given access to a black box $X = (x_1, x_2, \ldots, x_n)$ and a parameter k, what is the maximum size of the set $A \subseteq \{0,1\}^n$ where there exist a quantum algorithm U with at most k queries to X that satisfies $U(x) = PARITY(x)$ with probability $2/3$ for all $x \in A$.*

Remark Note that U does not have to satisfy any guarantee on the set A^c. So we can interpret Weak Parity Problem as an attempt to identify low complexity *partial functions* agreeing with PARITY on their domain $A \subseteq Q_n$.

The first observation is that a classical machine limited to query the box $k < n$ number of times, might as well not query the box at all, and produce say 0 immediately because no matter what the algorithm does, it would be successful at most on $\frac{1}{2}$ of the inputs.
To turn above problem into the familiar form of query minimization, we could use the following equivalent formulation:

Problem 2 (Weak Parity Problem- Second Formulation) *What is the minimum number of queries for an algorithm that computes $PARITY_n$ with bounded error on a set of fractional size at least $(\frac{1}{2} + \epsilon)2^n$ of boolean hypercube?*

1.1 Outline

In this work, we shall present a general upper bound of $O(\frac{n}{\sqrt{\log(1/\epsilon)}})$ for Weak Parity problem. Then we apply polynomial method to derive a lower bounds of $\Omega(\frac{n}{\log(1/\epsilon)})$. Using adversary method then we show a different absolute lower bounds of $\Omega(\sqrt{\log n})$ using only $\epsilon > 0$. This final lower bound is interesting as our general lower bound does not give us anything more than $\Omega(1)$ for computing PARITY on the regime where $\epsilon = O(2^{-n})$. This lower bound shows that a superconstant number of queries is always necessary for any nontrivial success for Weak Parity problem.
To prove the later lower bound we need a more combinatorial approach. We will use an extremal result over the hypercube due to Chung et al [CFGS88] to prove this:

Theorem 1 *Any quantum algorithm computing weakly PARITY on a subset of size as small as $2^{n-1} + 1$ of $Q_n = \{0,1\}^n$ must make at least $\Omega(\sqrt{\alpha(n)})$ queries where $\alpha(n) = \frac{1}{2}\log n - \frac{1}{2}\log\log(n) + \frac{1}{2}$.*

On how many queries is really necessary we have the following conjecture:

Conjecture 1 *No algorithm with query complexity $k \ll \sqrt{n}$ can compute $PARITY_n$ on more than half of the inputs.*

The motivation from this conjecture comes partly from the connections to sensitivity conjecture.[KK04] It turns out that one of the most natural approaches to proving lower bounds for Weak Parity in this regime go through very similar scenarios as in the case of sensitivity conjecture. Hence we believe that proving a lower bound of $\Omega(n^{\delta})$ for weak Parity in this regime would be a major breakthrough.

We make a first step in improving above lower bound on logarithmic regime by showing that

Theorem 2 *Any quantum algorithm computing weakly $PARITY_n$ on a subset of size as small as $2^{n-1}+1$ of $Q_n = \{0,1\}^n$ must make at least $\sqrt{\alpha_1(n)}n$ queries where*

$$\alpha_1(n) = (\frac{2}{3} - o(1)) \log n$$

We hope that a straightforward but tedious generalization of our method for theorem 2 can go up to $(1-\delta)\log n$ for any $\delta > 0$.The above result has the following rather interesting consequence :For any boolean function $f : \{0,1\}^n \to \{0,1\}$ we have $\alpha^{s(f)} > \deg(f)$ for any $\alpha > 2$.

To best of our knowledge, this is the best upper bound on $\deg(f)$ in terms of $s(f)$. In the interest of space, we shall not present the proof of 2 nor its generalization for $(1-\delta)\log n$ in this work. However, in the appendix, we shall present the main corollary, and leave the details of the of generalization to follow-up work. We shall prove,

Theorem 3 *For any boolean function f, we have $s(f) \geq \frac{2}{3}(1-o(1))\log(\deg(f))$ or more precisely*

$$\deg(f) \leq \sqrt{s(f)}(2\sqrt{2})^{s(f)}$$

In the manuscript, we shall not prove the theorem 2. Instead, we directly argue for theorem 3. We hope that the proof of this theorem captures most of the important ideas for the general improvement and can be used as a guide for the future work.

2 Acknowledgement

This work wouldn't have happened without the help of Scott Aaronson. I thank him for his technical help and his encouragement he provided. All the result here are either due to him or were obtained with his collaboration.

3 Algorithms for Weak Parity

Theorem 4 *There is an algorithm computing $PARITY_n$ on $\frac{1}{2}+\epsilon$ fraction of Q_n using $O(\frac{n}{\sqrt{log(1/\epsilon)}})$ queries.*

Proof The idea of the proof is to first consider the case $\epsilon = \frac{1}{2^n}$. In that case $OR(x_1, x_2, x_3, \ldots, x_n)$ agrees with $PARITY_n$ on all odd inputs, i.e. $x \in \{0,1\}^n$ where $x_1 + x_2 + \ldots + x_n = 1$ and also on the input $x = (0, 0, 0, \ldots, 0)$. Since OR_n can be computed by $O(\sqrt{n})$ queries using Grover's algorithm. We get our result for $\epsilon = \frac{1}{2^n}$.

Now for general ϵ, We will use the above idea as follows: Let $m = \lfloor \log(1/\epsilon) \rfloor$ and $s = \lceil \frac{n}{m} \rceil$. Then we take $y_1 = x_1 + x_2 + \ldots + x_s$ and $y_2 = x_{s+1} + \ldots + x_{2s}$ and so on. Notice that $PARITY(x) = PARITY(y_1, y_2, \ldots, y_m)$. To weakly compute Parity on $\frac{1}{2}+\epsilon$ fraction of inputs, we compute the function $g(x) = OR(y_1, y_2, y_3, \ldots, y_m)$. The above function would agree with PARITY on $\frac{1}{2} + 2^{-m}$ and can be computed with $O(\sqrt{m}.s) = O(\frac{n}{\sqrt{\log(1/\epsilon)}})$. ■

Remark An important observation is that above algorithms' speedup was all due to the case $\epsilon = \frac{1}{2^n}$. Moreover, if we managed to improve the above upper bound for any $\epsilon > 0$ we will get an improvement for all $\epsilon' > \epsilon$. So in some sense, the bottleneck in improving the above algorithm is improving the case $\epsilon = \frac{1}{2^n}$. However as the conjecture 1 states we believe that this is impossible.

4 Lower Bounds for Weak Parity

4.0.1 First Attempt: Adversary method

Let $Q_n = \{0,1\}^n$ be the boolean hypercube with the canonical graph structure connecting two points when their Hamming distance is one. We first need the following two easy lemmas:

Lemma 1 *For any set A of size $(\frac{1}{2}+\epsilon)2^n$ subset of boolean hypercube $Q_n = \{0,1\}^n$, the induced subgraph on A has induced average degree at least ϵn.*

Proof This is simply by counting edges. The vertices of A has total degree $(\frac{1}{2}+\epsilon)n2^n$ degree out of which at most $(\frac{1}{2}-\epsilon)n2^n$ belong to the cut $E(A, A^c)$. Hence, the average degree is at least $\frac{2\epsilon n}{2(\frac{1}{2}+\epsilon)} > \epsilon n$. ■

Lemma 2 *Any graph $G = (V, E)$ with average degree $d^* = d_{ave}$ has a induced subgraph where all the degrees are $\geq \frac{d^*}{2}$*

Proof Greedily delete any vertex with degree $< \frac{d^*}{2}$. Continue this until either you delete every vertex or you are left with an induced subgraph satisfying the above min-degree condition. Now notice that the first event cannot occur, because if the graph gets empty, this means that deleting $< \frac{d^*}{2}|V| = |E|$ edges has emptied the graph. But this is impossible. ■

Theorem 5 *The quantum query complexity of Weak Parity Problem over $\frac{1}{2} + \epsilon$ fraction of hypercube is at least $\Omega(\epsilon n)$.*

Proof Consider the set A where the quantum algorithm computes the Parity correctly(with bounded error). Since $|A| = (1/2 + \epsilon)2^n$ by above lemmas there exist a $B \subset A$ such that the induced degree over B are at least $\epsilon n/2$. Let $X = B \cap \{x \in Q_n || x| = \text{odd}\}$ and $Y = B \cap \{x \in Q_n || x| = \text{even}\}$. Consider the relation between $R(X, Y)$ that you put every x in relation between all of its neighbors in y. By a direct application of adversary method distinguishing X from Y requires $\Omega(\epsilon n)$ queries. ■

This lower bound will be superseded by the next lower bound using polynomial method. But since it's quite general and has potential application to other problems in Weak query model it was presented.

4.1 Polynomial Method Lower Bound

Here we shall present the main query lower bound for Weak Parity problem of $\Omega(\frac{n}{\log(1/\epsilon)})$. This crucially uses the self-reducibility of Parity.

Theorem 6 *Any quantum algorithm which on a set A subseteqQ_n of size $(\frac{1}{2} + \epsilon)2^n$ correctly outputs PARITY with probability $\geq 2/3$ must make at least $\Omega(\frac{n}{\log(1/\epsilon)})$ queries.*

Proof Let T the query complexity of the algorithm C for weak Parity. We will use the weak algorithm to produce an algorithm C' that decides PARITY on *all inputs* with probability strictly greater than $1/2$ and query complexity $O(T \log(1/\epsilon))$. The algorithm is as follows: First it chooses $Y = (y_1, y_2, \ldots, y_n) \in \{0, 1\}^n$ uniformly at random. Let $b = y_0 + y_1 + \ldots + y_n$ be the parity of Y. Then C' runs algorithm C on the input $Z = X + Y$ for $l = M \log(1/\epsilon)$ times computing the majority of the outputs. Let r be the the majority of the outputs of l different tuns of C on Z. Then C' will outputs $b + \alpha$ as $PARITY(X)$. Now let's see why this algorithm works:

First of all, given access to the query box for actual input X, one can easily simulate the queries $Z = X + Y = (x_0 + y_0, x_1 + y_1, \ldots, x_n + y_n)$. More formally, to simulate a single query to Z, you use one query to the blackbox for querying X and then you run a single query to Y which one can easily implemented with a simple (classical) circuit.

Now the important thing is that Z is uniformly at random from boolean hypercube. So with probability $\frac{1}{2} + \epsilon$ this point is going to be in the set where weak algorithm correctly computes PAIRITY with probability $\geq 2/3$. Now the idea is that if we run the weak algorithm for $l = M \log(1/\epsilon)$ times on Z and take the majority of the answer, if $Z \in A$ then then the majority agrees with $PARITY(Z)$ with probability $> 1 - \frac{\epsilon}{2}$ for $M = 200 \log(1/\epsilon)$.

Now what's the probability of success of this algorithm.

$$\begin{aligned}\mathbb{P}[C'(X) = PARITY(X)] &\geq \mathbb{P}[Z \in A]\mathbb{P}[r = PARITY(Z)] \\ &\geq (\frac{1}{2} + \epsilon)(1 - \frac{\epsilon}{2}) > \frac{1}{2}\end{aligned}$$

Since any algorithm computing PARITY on all inputs with $> 1/2$ probability of success can be seen to require $\frac{n}{2}$ queries by polynomial method the result follows. [BBC$^+$01]

■

4.1.1 Understanding the Gap Better

The main upper and lower bounds we have presented are formulated in terms of the second formulation of weak PARITY problem. To understand the gap between this upper and lower bounds, it's more instructive to go back to the original formulation 1 of this problem.

Corollary 1 (theorem 4 and 6 recast) *A quantum machine limited to make only k queries to $X = (x_1, x_2, \ldots, x_n)$ can output PARITY correctly with bounded error on a set of fractional size at most $\frac{1}{2} + 2^{-\Omega(\frac{n}{k})}$*

On the other hand, there exist an algorithm that makes k queries and outputs PARITY correctly on a set of fractional size $\frac{1}{2} + 2^{-O(\frac{n^2}{k^2})}$.

In this language, It's clear that if $k = n^c$ for some $0 < c < 1$ the set A where one could correctly outputs PARITY obeys

$$\frac{1}{2} + 2^{-O(n^{2(1-c)})} \leq \frac{|A|}{2^n} \leq \frac{1}{2} + 2^{-\Omega(n^{1-c})}$$

So we see that if $k = n^c$ for $c < 1/2$ the gap is very large. The lower bound on the size is just a trivial $\frac{1}{2}$ while the upper bound remains the non-trivial $\frac{1}{2} + 2^{-\Omega(n^{1-c})}$. As stated in conjecture 1 we believe that the answer that the lower bound here is closer to the truth. In the next section, we shall present some evidence for this conjecture and presents the combinatorial questions related to this question.

5 Weak Parity with $k \ll \sqrt{n}$ queries?

Our lower bound in previous section has not ruled out an algorithm that queries $X = (x_1, x_2, \ldots, x_n)$ only a constant number of times and compute PARITY on a set larger than half of the boolean hypercube. Indeed, in this section we shall *rule out this possibility* using arguments about sensitivity. As indicated in the conjecture 1, we believe that the true query complexity of any algorithm for Weak-Parity on a set larger than half is $\Omega(\sqrt{n})$. However, we believe it might be unlikely for us to prove a lower bound of $\Omega(n^\delta)$ for any $\delta > 0$ given the current techniques.

5.1 Some Extremal Problems over Hypercube

The following combinatorial problem investigated originally by [GL92, CFGS88] is the key to our results.

Theorem 7 (Chung et al) *Let $A \subseteq Q_n$ have size at least $2^{n-1}+1$. The induced subgraph on A has a vertex of degree $\alpha(n)$. where*

$$\alpha(n) = \frac{1}{2}\log(n) - \frac{1}{2}\log\log n + \frac{1}{2}$$

Corollary 2 *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a boolean function. Let $s(f)$ denote the sensitivity of f and $\deg(f)$ denote the degree of f as a polynomial over $\mathbb{R}$. (The same as the largest size of a Fourier coefficient) Then we have*

$$s(f)4^{s(f)} \geq \deg(f)$$

Proof Let f be our boolean function with $m = \deg(f)$. Wlog assume $x_1x_2\ldots x_m$ is a maxonomial of f. Consider the function $h(x_1, x_2, \ldots, x_m) = f(x_1, x_2, \ldots, x_m, 1, 1, 1, , \ldots, 1)$. So h is a degree m function over Q_m which is restriction of f. Hence, $s(h) \leq s(f)$. So we just need to prove the result for h.

Let $g(x) = (\prod_{i=1}^m x_i)\, h(x)$. We have

$$g(x) = \left(\prod_{i=1}^n x_i\right) \sum_{S\subseteq[m]} \widetilde{h}(S) \prod_{i\in S} x_i = \sum_{S\subseteq[m]} \widetilde{h}(S) \prod_{i\notin S} x_i$$

Since h was full degree $\widetilde{h}([n]) \neq 0$ hence $\widetilde{g}(\emptyset) = \mathbb{E}[g] \neq 0$. So it means taking $A = \{x \in Q_n | g(x) = 1\}$ either A or A^c have size larger than $2^{n-1} + 1$. Wlog $|A| \geq 2^{n-1} + 1$. Apply the above theorem to A to conclude there exist $x \in A$ such that x as $\alpha(m)$ neighbors in A.[2] Now notice that $h(x) =\neq h(y)$ for all those $\alpha(m)$ neighbors y. The result follows. ■

Remark In the appendix we shall present an improvement to this corollary.

To prove the theorem 7 We need following lemma:

Lemma 3 *Let $V \subseteq Q_n$ and let d be the average degree of vertices in the induced subgraph on V. Then*

$$|V| \geq 2^d$$

Proof The statement is equivalent to proving $|V| \log(|V|) \geq 2|E|$ where $|E|$is the number of induces edges over V. Wlog assume $V_1 = V \cap \{x_1 = 0\}$ and $V_2 = V \cap \{x_1 = 1\}$ such that V_i's are both non-empty. Assume $|V_1| \leq |V_2|$ and notice that the number of edges in direction 1 in E is at most $|V_1|$. Now by induction on size of $|V|$ we have

$$2|E| \leq 2|E_1|+2|E_2|+|V_1| \leq |V_1| \log(|V_1|)+|V_2| \log(|V_2|)+|V_1| \leq (|V_1|+|V_2|) \log(|V_1|+|V_2|)$$

The last inequality follows from the inequality $H(p) \geq 2p$ for $0 < p \leq 1/2$ for binary entropy $H(p)$. ■

Now we shall prove the theorem 7 bound using the lemma. We shall somewhat modify the proof of Graham et al which makes the proof easier to generalize.

Proof of **theorem 7**: Wlog assume set A has size exactly $2^{n-1} + 1$. Consider some direction i. Consider the two to one mapping $X = (x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n) \to (x_1, x_2, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n)$. Since the map injects to a set of size $2^{n-1} < |A|$ there are two elements mapped to the same point, i.e. we can find a pair $p, p^i \in S$ where p^i denote the point p with i-th bit flipped.

Consider the set H of the points on the $x_i = 0$ subcube such that x, x^i either both belong to A or both do not belong to A,

$$H = H_1 \cup H_2 \qquad H_1 = \{x \in Q_n \,|\, x_i = 0 \ (x, x^i) \in A\} \quad H_2 = \{y \in Q_n \,|\, y_i = 0 \ (y, y^i) \notin A\}$$

Since $p \in H_1$ this set is nonempty. We shall prove that H_1 is large using the fact that H is large. Let Δ be the max-degree of A. It is easy to see that each element $x \in H$ has $n - 2\Delta + 1$ neighbors in H_2.(because at most $\Delta - 1$ neighbors of x and at most the same number of neighbors of x^i will be in A)

[2] Recall that $\alpha(m) = \frac{1}{2} \log(m) - \frac{1}{2} \log \log m + \frac{1}{2}$

For every pair (x, x^i) such that $x \notin H$ exactly one of them belong to A. So $|A| - |A^c| = 2 = \frac{|H_1|-|H_2|}{2}$. Hence, $|H_2| = |H_1| - 1$. Since every element of H_1 has $n - 2\Delta + 1$ neighbors in H_2 we see that elements of H_2 have $> n - 2\Delta + 1$ neighbors in H_1 *on average*. Now it follows that the average induced degree in H is at least $n - 2\Delta + 1$. From the lemma it follows that $|H| \geq 2^{n-2\Delta+1}$ Now it follows that $|H_1| \geq 2^{n-2\Delta}$. This is exactly the number of edges in A in direction i.

Now we know that the total number of edges in A are bounded by $\Delta(2^{n-1}+1)$. Hence we see that there is some direction such that the number of edges in that direction is bounded by $\frac{\Delta}{n}(2^{n-1}+1)$. Combining the bounds we get $\frac{\Delta}{n}(2^{n-1}+1) \geq 2^{n-2\Delta}$ that will give us the desired result. ■

The consequences of this for weak PARITY problem is clear. Applying grover lower bound to a point with degree $\alpha(n)$ we get that,

Corollary 3 *Any quantum algorithm computing weakly PARITY on a subset of size as small as* $2^{n-1}+1$ *of* $Q_n = \{0,1\}^n$ *must make at least* $\sqrt{\alpha(n)}$ *queries where* $\alpha(n) = \frac{1}{2}\log n - \frac{1}{2}\log\log(n) + \frac{1}{2}$.

Remark In the appendix, we shall generalize the techniques of this proof slightly to get an improvement to the corollary 2.In improving the relationship between $s(f)$ and $\deg(f)$ instead of improving the theorem 7, we shall directly apply our techniques of proof of theorem 7 to the corollary. The reason we decide to present the improvement to the corollary 2 but not to 7 is mostly for the sake of clarity.

References

[BBC+01] Robert Beals, Harry Buhrman, Richard Cleve, Michele Mosca, and Ronald de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, July 2001.

[CFGS88] F. R. K. Chung, Zoltán Füredi, R. L. Graham, and P. Seymour. On induced subgraphs of the cube. *J. Comb. Theory Ser. A*, 49(1):180–187, September 1988.

[GL92] C. Gotsman and N. Linial. The equivalence of two problems on the cube. *J. Comb. Theory Ser. A*, 61(1):142–146, September 1992.

[KK04] Claire Kenyon and Samuel Kutin. Sensitivity, block sensitivity, and l-block sensitivity of boolean functions. *Inf. Comput.*, 189(1):43–53, February 2004.

A Improving The Relationship Between $\deg(f)$ and $s(f)$

Now we shall sketch how to improve corollary 2 to $deg(f) \leq \sqrt{s(f)}(2\sqrt{2})^{s(f)}$. I believe that with some care the result of this section can be generalized to get $\deg(f) \leq c^{s(f)}$ for any $c > 2$ and also get the improvement $\Delta(A) \geq (1-\delta)\log n$ analogous to theorem 7. But we shall leave this further generalization that require *some care* to future work and just stick the simplest improvement that captures the idea. However, we mention that to best of our knowledge even this simple generalization might be the best known upper bound on $\deg(f)$ based on sensitivity.

The proof is a simple generalization of the proof for theorem 7. The idea is to instead of counting sensitive pairs in one direction to consider and count higher order structures of this kind instead.

Theorem 8 *For any boolean function f, we have $s(f) \geq \frac{2}{3}(1-o(1))\deg(f)$. More precisely,*

$$deg(f) \leq \sqrt{s(f)}(2\sqrt{2})^{s(f)}$$

Proof As shown in the proof of weaker version of this theorem, we can wlog assume that f is full degree n. Given two directions (i,j) let $D_{i,j}(f)$ denote the D_iD_jf the derivative of f with respect of i-th and then j-th direction. We have

$$D_{i,j}(f)(x) = \sum_{\{i,j\}\in S\subset[n]} \widetilde{f}(S)x_{S\setminus\{i,j\}}$$

Since f is full degree $D_{i,j}(f)$ is also a non-zero polynomial. As a result since $D_{i,j}(f)$ has non-zero Fourier coefficients it must be non-zero at some point x. Now we shall interpret the fact that $D_{i,j}(f)(x) \neq 0$ in terms of sensitivity at points (x, x^i, x^j, x^{ij}).

Consider the small subcube formed by the four points (x, x^i, x^j, x^{ij}). I claim we must have that one of these 4 points has sensitivity 2 inside this square, i.e. wlog $f(x) = 0$ and $f(x^i) = f(x^j) = 1$. If there is no sensitivity inside this cube then $f(x) = f(x^i) = f(x^j) = f(x^{ij})$ which is impossible because $D_{ij}(f)(x) = \frac{f(x)-f(x^i)-f(x^j)+f(x^{ij})}{4} \neq 0$. So assume $f(x) = 0$ and $f(x^i) = 1$. If both x and x^i has only sensitivity one in the square we must have $f(x^j) = 0$ and $f(x^{ij}) = 1$ which is again in contrast with $D_{ij}(f)(x) \neq 0$.

So assume wlog that $f(x) = 0$ and $f(x^i) = f(x^j) = 1$ and $x_i = x_j = 0$. Consider the set $H = \{x \in Q_n \,|\, x_i = x_j = 0\,,\, f(x^i) = f(x^j) = 1\,,\, f(x) = 0\}$. It is immediate that this set H has $n - 3s(f) + 2$ minimum degree for its vertices. Hence, we get that size of this set is at least $2^{n-3s(f)+2}$. For each $x \in H$ we have a triple (x, x^i, x^j) where $f(x)$ is sensitive with respect to both i and j direction. On the other hand number of such triples is at most $\binom{s(f)}{2}2^n$. So there must be a pair of directions with less than $\frac{\binom{s(f)}{2}}{\binom{n}{2}}2^n$ such triples.

Hence it follows that $2^{n-3s(f)+2} \leq \frac{\binom{s(f)}{2}}{\binom{n}{2}} 2^n$. From this we get

$$\deg(f) \leq \sqrt{s(f)}(2\sqrt{2})^{s(f)}$$

which is a significant improvement to previous bound of $\deg(f) \leq s(f)4^{s(f)}$. Using above ideas with higher and higher derivatives we hope that we get more improvements. This will require a more elaborate analysis of the condition $D_{i_1,i_2,\ldots,i_k}(f) \neq 0$ in terms of the sensitivity induced in that k-cube. ■

B Why Is Section 4.1 Called Polynomial Method?

We shall give a completely pure-mathematical version of the proof of theorem 6 here. Let $p(x_1, x_2, \ldots, x_n)$ be the polynomial of degree $2T$ that represent the acceptance probability of our Weak-Parity algorithm. Consider the following mapping $f(\alpha) = 3\alpha^2 - 2\alpha^3$. One can easily check that the dynamic under f always remain between $[0, 1]$ and f satisfies the symmetry condition $f(1-\alpha) = 1 - f(\alpha)$. and has two attractive fixed points at 0 and 1 and an repulsive fixed point at $1/2$. Moreover for $\alpha \leq 1/3$ we have $f(\alpha) \leq \alpha^{1.2}$.

Using this, consider the following iteration. Weak $p_0(x) = p(x)$ and define $p_i(x) = 3p_{i-1}(x)^2 - 2p_i(x)^3$. After $m = O(\log\log(1/\epsilon))$ iteration $p_m(x)$ would be a polynomial of degree $O(T\log(1/\epsilon))$. Furthermore for every $x \in A$ we have $|p_m(x) - PARITY(x)| \leq \epsilon/2$. This is because the error after m iteration is at most $\left(\frac{1}{3}\right)^{1.2^m} \leq \epsilon/2$ for some $m = O(\log\log(1/\epsilon))$. Now if we consider the polynomial

$$q(y) = \frac{1}{2}\mathbb{E}_{|z|=\text{even}} p(x+z) + \frac{1}{2}(1 - \mathbb{E}_{|z|=\text{odd}} p(x+z))$$

This is a polynomial of degree $O(T\log(1/\epsilon))$ that satisfies $|q(y) - PARITY(y)| < \frac{1}{2}$. Using the symmetrization of Minsky and Pappert[BBC+01] and direct application of polynomial method to number of roots of symmetrized polynomial at $1/2$ the lower bound of $T = \Omega(\frac{n}{\log(1/\epsilon)})$ follows.

C Relations to Sensitivity Conjecture

For the remainder of the section we shall denote boolean hypercube with its natural graph structure as $Q_n = \{0, 1\}^n$.

In order to appreciate the connections between this problem and sensitivity conjecture. we need the following definitions over the boolean hypercube.

Definition 1 *For any $A \subseteq Q_n$ define $\Lambda(A)$ as the the maximum degree of the induced subgraph on A. Similarly, define $\Gamma(A) = \max\{\Lambda(A), \Lambda(A^c)\}$.*

It turns out that sensitivity conjecture is *equivalent* to lower bounds on $\Gamma(A)$ for $|A| > 2^{n-1}$ while lower bounds on Weak PARITY problem *follows* from lower bounds on $\Lambda(A)$. Of course the most natural conjecture that proves *both* conjectures is the following

Conjecture 2 (Subsumes Sensitivity and WEAK PARITY Conjectures 1) *There exist $\delta > 0$ such that for any set $A \subseteq Q_n$ with $|A| > 2^{n-1}$ we have $\Lambda(A) = \Omega(n^\delta)$.*

Remark From the examples so far constructed in the literature the upper bound shown on the above δ is $1/2$ that follows from the following example [CFGS88]: Assume n is a square and consider the function

$$f(x) = AND(OR(x_1, x_2, \ldots, x_{\sqrt{n}}), OR(x_{\sqrt{n}+1}, \ldots, x_{2\sqrt{n}}), \ldots, OR(x_{n-\sqrt{n}+1}, x_{n-\sqrt{n}+2}, \ldots, x_n))$$

Now take the set $B = \{x \in Q_n \mid PARITY(x) = f(x)\}$. Then we have the following:

- $|B| = 2^{n-1} \pm 1$ and $|B^c| = 2^{n-1} \mp 1$.
- $\Lambda(B) = \Lambda(B^c) = \sqrt{n}$ hence $\Gamma(B) = \sqrt{n}$.

It's author's opinion that the conjecture 2 is one of the most approachable versions of the statements that imply sensitivity conjecture. Indeed arguing about large subsets of boolean hypercube and applying randomness vs structure, i.e. very careful inductions, type ideas might be the way to go forward for proving such statements.

The relation between the conjecture 2 and weak PARITY conjecture is straightforward. Because of lower bound for Grover problem, a $\Omega(n^\delta)$ lower bound on on $\Lambda(A)$ implies a $\Omega(n^{\delta/2})$ lower bound for Weak PARITY query complexity on A. As to the relation to sensitivity we have,

Theorem 9 *If there exist $\delta > 0$ such that $\Gamma(A) = \Omega(n^\delta)$ for any $|A| > 2^n$ then sensitivity conjecture holds. More precisely $s(f) = \Omega(\deg(f)^\delta)$.*

Remark Since $\Gamma(A) \geq \Lambda(A)$ we see that sensitivity conjecture is implied by the conjecture 2

Proof First let's see how sensitivity conjecture implies a lower bound on the $\Gamma(A)$. Wlog, take $Q_n = \{-1, 1\}^n$ so Fourier analysis becomes more natural.

Define the function $g : \{-1, 1\}^n \to \{-1, 1\}$ as $g(x) = 1$ iff $x \in A$. Since $|A| > 2^n - 1$ we have $\mathbb{E}[g] > 0$. Consider $h(x) = g(x) \prod_{i=1}^n x_i$. Since $\mathbb{E}[g] > 0$ we see that h satisfies

$\deg(h) = n$. Assuming sensitivity conjecture this means $s(h) = \Omega(n^\delta)$. So take a $x \in Q_n$ such that $s(h, x) = \Omega(n^\delta)$. If $x \in A$ this means all those $\Omega(n^\delta)$ neighbors of x are also in A, hence $\Lambda(A) = \Omega(n^\delta)$. Similarly, if $x \in A^c$ it would mean that $\Lambda(A^c) = \Omega(n^\delta)$. In any case, $\Gamma(A) = \Omega(n^\delta)$ easily follows.

The trick to prove the other direction is also similar. Consider a function $g(x)$ we want to show $s(g) = \Omega(\deg(g)^\delta)$. Wlog let $x_1 x_2 \dots x_m$ be a maxonomial in Fourier expansion of g so $m = \deg(g)$. Take the restriction of g to those m coordinates. The lower bound on the sensitivity follows by applying $\Gamma(A) = \Omega(m^\delta)$ where $A \subseteq Q_m$ is defined by

$$A = \{x \in Q_m | \, g(x_1, x_2, \dots, x_m, 1, 1, \dots, 1) \prod_{i=1}^{m} x_i = 1\}$$

■

In summary, we see that the connection between sensitivity and weak PARITY is that a natural (stronger) conjecture than sensitivity implies both. This is the combinatorial conjecture that $\Lambda(A) = \Omega(n^\delta)$ for any $|A| > 2^{n-1}$. On the other hand, if one replaces $\Lambda(A)$ with $\Gamma(A) = \max\{\Lambda(A), \Lambda(A^c)\}$ this is indeed *equivalent* with sensitivity conjecture.