Property Testing Lower Bounds via Communication Complexity

Lecturer: Venkatesh Medabalimi

1 Outline

In this lecture we discuss a beautiful application of Communication Complexity to come up with lower bounds in the area of property testing. We discuss the work - `Property Testing lower bounds via Communication Complexity` ([3], journal version) . This work provides testing lower bounds for many properties by introducing a new technique that seems to be able to say a lot. We begin by a brief introduction to Property Testing and then see how it relates to Communication Complexity at a high level. We then describe the general reduction technique that is employed throughout this lecture. We recollect some of the communication complexity problems and then give the reductions based on the general technique to get the testing lower bounds for a variety of properties.

2 Property Testing

Property testing studies "what can be determined about a large object with limited access to it ?" A typical property testing setup has

- A large object
- A property of interest
- A tester with query access to the large object

The goal of the tester is to

- Accept the object if it has property $\mathcal{P}$
- Reject if the object is "far" from having $\mathcal{P}$

But how far is "far" ? An object is ϵ far from having a property $\mathcal{P}$, if one has to modify the object on ϵ fraction of the places to match the closest object with the property $\mathcal{P}$. Making the above more specific, in this lecture we will be using the following setting

- The large object: A boolean function f on n bits.
- A property of interest : X (some subset of functions)
- The Testers goal given an ϵ is to :
  - Accept with probability at-least $\frac{2}{3}$ if f has the property $\mathcal{P}$.
  - Reject with probability at-least $\frac{2}{3}$ if f has to be modified on at-least ϵ fraction of the inputs to be in $\mathcal{P}$.

Query Complexity($\mathcal{P}, \epsilon, \mathcal{A}$) is the number of times f has to be queried by the testing algorithm $\mathcal{A}$ with the above goals. Some examples of properties that one might be interested in testing are:- whether f is

- a linear function ?
- isomorphic to a given function ?
- a k-junta ?
- a monotone function ?
- a dictator ?
- a half-space ?
- an s-sparse polynomial ?
- computable by a size s decision tree ?

Query complexity upperbounds for testing properties are usually known. (Apparently) Lower bounds are hard to come by and no significant technique beyond Yao's minimax lemma is known. We need lower bounds !! Based on a new technique that uses communication complexity we shall discuss testing lower bounds for the following properties

- k-linearity
- k-junta
- functions with low Fourier degree
- Class of linear functions from $GF(3)^n$ to $GF(3)$ that have only 0-1 coefficients.
- Monotonicity and Submodularity
- Computability by width-4 OBDD
- Concise Representations

Communication complexity has been a good tool for coming up with lower bounds for many problems. Before we know more on the connection between communication complexity and property testing, we shall recollect some problems in communication complexity using which we get our lower bounds. we know several problems with large Communication Complexity: Set-disjointness, inner-product, graph hamming distance etc. On the basis of these results Communication complexity has evolved as a means for proving lower bounds in other areas. We have discussed some of these applications in lectures 5 and 6. The areas include streaming algorithms, circuit complexity and data structures.

3 Property Testing & Communication Complexity : Any connections ?

There are two similarities that one can observe between communication complexity and property testing,

- both have parties with unbounded computational power: tester and communicating players.
- both involve algorithms by parties with restricted access to input.

Using these similarities we reduce communication problems to testing problems. In particular, we wish to relate number of bits communicated to compute some function to the number of queries required to test a property. The overview of the reduction procedure we use is as follows. We perform the reduction in two steps. For these steps we need to define an intermediate communication problem $C_\phi^{\mathcal{P}}$. In $C_\phi^{\mathcal{P}}$, Alice and Bob get functions f, g as their arguments and their goal is to solve the property testing problem for a combination function of f and g, $h(x) = \phi(f(x), g(x), x)$ for the property of interest $\mathcal{P}$ (the property for which we wish to get a lower bound on query complexity). With the help of $C_\phi^{\mathcal{P}}$ we perform the reduction in the following two steps.

- Reduce a communication complexity problem C to $C_\phi^{\mathcal{P}}$. (giving us $R(C_\phi^{\mathcal{P}}) \geq R(C)$)
- Reduce $C_\phi^{\mathcal{P}}$ to the problem of property testing $\mathcal{P}$. (giving us a lower bound for property testing.)

We now describe the general technique in greater detail, first we need some definitions.Let S be some finite set.

Definition (Combining Operator) A combining operator takes $f, g : \{0,1\}^n \to S$ and gives $h : \{0,1\}^n \to R$.

Definition (Simple Combining Operator) A combining operator ϕ is simple if $\forall\ x,\ h(x) = \phi(f(x), g(x), x)$, i.e can be computed from $x, f(x)$ and $g(x)$.

Let $\mathcal{P}$ be a property (that has an algorithm) with query complexity Q. $C_\phi^{\mathcal{P}}$ is the following communication problem. Alice receives a function $f : \{0,1\}^n \to S$, Bob receives $g : \{0,1\}^n \to S$. The have to compute

$$C_\phi^{\mathcal{P}}(f,g) = \begin{cases} 1 & \text{if } h \text{ has property } \mathcal{P} \\ 0 & \text{if } h \text{ is } \epsilon\text{-far from having } \mathcal{P} \end{cases}$$

The following lemma relates the randomized communication complexity of $C_\phi^{\mathcal{P}}(f,g)$, $R(C_\phi^{\mathcal{P}})$ to the randomized query complexity of $\mathcal{P}$, $Q(\mathcal{P})$ when ϕ is a simple combination operator.

Lemma 1 *(Main Reduction Lemma) Consider functions from $\{0,1\}^n \to S$, for any property $\mathcal{P}$, a simple combining operator ϕ and communication game $C_\phi^{\mathcal{P}}$ we have $R(C_\phi^{\mathcal{P}}) \leq 2Q(\mathcal{P})\lceil \log |S| \rceil$*

Proof The proof is easy to see. Alice and Bob want to test if $h(x) = \phi(f(x), g(x), x)$ has the property $\mathcal{P}$. They can run an identical randomized property tester for $\mathcal{P}$ using common randomness. When they both wish to see what value does h take at some x, since ϕ is a smile combination operator they can compute it using just $f(x), g(x)$ and x. To do this Alice can communicate to Bob the value $f(x)$ in $\lceil \log |S| \rceil$ bits and similarly Bob can communicate $g(x)$ to Alice costing in all $2\lceil \log |S| \rceil$ bits per query simulated. This gives us $R(C_\phi^{\mathcal{P}}) \leq 2Q(\mathcal{P})\lceil \log |S| \rceil$.

The general scheme we employ for showing a lower bound for a given property is to

- identify a suitable communication complexity problem $C(x, y)$.
- reduce C to $C_\phi^{\mathcal{P}}$ using an appropriate simple combination operator. This reduction gives us

$$R(C_\phi^{\mathcal{P}}) \;\geq\; R(C)$$

- From the main reduction lemma 1 we have

$$2Q(\mathcal{P})\lceil \log |S| \rceil \geq R(C_\phi^{\mathcal{P}})$$

- So using both the main reduction lemma and the reduction from C to $C_\phi^{\mathcal{P}}$ we have

$$2Q(\mathcal{P})\lceil \log |S| \rceil \quad \geq \quad R(C_\phi^{\mathcal{P}}) \;\geq\; R(C)$$

4 Communication Complexity problems we use

The Communication Complexity problems we will be interested in and their lower bounds are recollected below.

Set-Disjointness

Given n bit strings x, y representing sets $A, B \subseteq [n]$ respectively, $DISJ_n(x, y) = 1$ if $|A \cap B| > 0$ and $DISJ_n(x, y) = 0$ if A, B are disjoint. Given strings x, y Alice and Bob have to compute $DISJ_n(x, y)$. We know that $R(DISJ_n) = \Omega(n)$ [6],[7]. The complexity is known to be the same even if there is a promise that $|A \cap B| \leq 1$. [1]

k-BAL-DISJ

This is a balanced version of $DISJ_n$ where Alice receives a set $A \subseteq [n]$ of size $\lfloor k/2 \rfloor + 1$ and Bob receives a set B of size $\lceil k/2 \rceil + 1$ and there is a promise that $|A \cap B| \leq 1$.

Lemma 2 *For all $0 \leq k \leq n - 2$ we have $R(k\text{-BAL-DISJ}) = \Omega(\min k, n - k)$*

Gap-Equality

Alice and Bob, given $x, y \in \{0, 1\}^n$ compute

$$GEQ_{n,t}(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } \Delta(x, y) = t \\ * & \text{otherwise} \end{cases}$$

Note that $R^0(GEQ) = O(1)$ from the randomized complexity of Equality problem. For protocols that err only on the other side, the communication complexity is much higher. It is known that $R^1(GEQ)$ is $\Omega(n)$.

Gap-Hamming-Distance

Given x, y Alice and Bob compute

$$GHD_{n,t}(x,y) = \begin{cases} 1 & \text{if } \Delta(x,y) \geq n/2 + t \\ 0 & \text{if } \Delta(x,y) \leq n/2 - t \\ * & \text{otherwise} \end{cases}$$

For $t = \theta(\sqrt{n})$ a lower bound of $R(GHD) = \Omega(n)$ is known.

Extended-Gap-Hamming Distance

In the extended version x, y come with the promise that $|x| = |y| = k/2$ and one wishes to distinguish $\Delta(x,y) \geq k/2+t$ from $\Delta(x,y) \leq k/2-t$. It is known that $\forall t$ and $\forall k \leq n$, $R(GHD_{n,k,t}) = \Omega(\min\{(k/t)^2, k\} - \log k)$

5 Testing $k-$linearity, k-junta, Functions with low Fourier degree

Definition K-Linear: A function $f : \{0,1\}^n \to \{0,1\}$ is $k-$linear if it is of the form $f(x) = \sum_{i \in S} x_i \pmod 2$ where $|S| \leq k$.

Definition K-Junta: A function $f : \{0,1\}^n \to \{0,1\}$ is a $k - junta$ if $\exists$ a set $J \subset S$ s.t $|J| \leq k$ and $\forall x, y \in \{0,1\}^n$ s.t $x_i = y_i$, $\forall i \in J$ we have $f(x) = f(y)$.

Definition Fourier degree of a function: Every function $f : \{0,1\}^n \to \{-1,1\}$ has a unique Fourier expansion $f = \sum_{S \subseteq [n]} \hat{f}(S)\chi_S(x)$ where $\chi_S(x) = (-1)^{\sum_{i \in S} x_i}$. We say a function $f : \{0,1\}^n \to \{-1,1\}$ has Fourier degree k if the size of the largest set S s.t $\hat{f}(S) \neq 0$ is at-most k.

Linear functions have some nice separation properties amongst themselves as well as with juntas and functions with low fourier degree. In particular, one can see that all linear functions are $1/2$ far from eachother. The lemma below helps us disguise the set disjointness problem as a property testing problem in the space of linear functions.

Lemma 3 *(distance to $(k+2)$-linear functions) If f is $k+2$-linear then its $1/2$ far from*

- *k- linear functions.*
- *k- juntas.*
- *functions with Fourier degree at-most k.*

Proof Its easy to see the first two. We will prove the last one, which actually implies first two. Consider a function g with Fourier degree k, the distance between f and g is given by the expectation $E_x[f(x)g(x)]$ under a uniform distribution on x. By Parseval's theorem we have

$$E_x[f(x)g(x)] = \sum_{S \subseteq [n]} \hat{f}(S)\hat{g}(S)$$

f is $k+2$-linear so its Fourier expansion has $T \subseteq [n]$ s.t $\hat{f}(T) > 0, |T| = k+2$ and $\hat{f}(S) = 0\ \forall S \neq T$. Also since g has Fourier degree k, $g(T) = 0$. So $E_x[f(x)g(x)] = 0 \implies$ the distance is $\frac{1}{2}$. So a $k+2$ linear function is 1/2 far from functions with Fourier degree at-most k.

Theorem 4 *Fix $1 < k < n-1$, the query complexity of k-linearity, k-junta and functions of Fourier degree atmost k is atleast $\Omega(\min\{k, n-k\})$ (known upper bounds are $O(k \log k), O(k \log k)$ [2] and $2^{O(d)}$ respectively)*

Proof We perform the reduction from k-BAL-DISJ$_n$. Alice get a set $A \subseteq [n]$ of size $\lfloor k/2 \rfloor + 1$ and Bob receives a set $B \subseteq [n]$ of size $\lceil k/2 \rceil + 1$ and there is a promise that $|A \cap B| \leq 1$. Alice now uses the parity function corresponding to set A and Bob uses the parity function on set B. Consider the communication game $C_\phi^{\mathcal{P}}$ (We shall leave undisclosed what $\mathcal{P}$ is for the time being) defined by the combination operator is $\phi(f, g) = f \oplus g$. If $|A \cap B| = 0$, $h(x)$ is $k+2$ linear. If $|A \cap B| = 1$, $h(x)$ is k-linear. Note that a k linear function is also a $k-$junta and is a function with Fourier degree at-most k. Also by lemma 3 we have that when A, B are disjoint the resultant $k+2$-linear function is 1/2 far from all the *k-linear* functions, *k-juntas* and functions with Fourier degree at-most k. So our $\mathcal{P}$ in the reduction can effectively be any of these three properties. So we have $R(K$-BAL-DISJ$) \leq \min\{R(C_\oplus^{k\text{-}JUNTA}), R(C_\oplus^{DEGREE-k}), R(C_\oplus^{K-SPARSE})\}$ and correspondingly we get a lower bound for the query complexity of all these properties through the main reduction lemma1, since $\oplus$ is a simple combination operator.

More recently tight bounds for testing $k-$linearity for $k \approx n/2$ were obtained in [4] using connections with geometry of the boolean hypercube.

6 Testing the class of linear functions from $GF(3)^n$ to $GF(3)$ that have only 0-1 coefficients

Definition Linear function from $GF(3)^n$ to $GF(3)$: Functions $f : \{0,1,2\}^n \to \{0,1,2\}$ of the form $f(x) = \sum_{i \in S \subseteq [n]} a_i x_i$ where $a_i \in \{0,1,2\}, \forall i$.

Definition Linear function from $GF(3)^n$ to $GF(3)$ with only 0-1 coefficients: Functions f : $\{0,1,2\}^n \to \{0,1,2\}$ of the form $f(x) = \sum_{i \in S \subseteq [n]} x_i$

Theorem 5 *Testing the class of linear functions from $GF(3)^n$ to $GF(3)$ that have only 0-1 coefficients requires $\Omega(n)$ queries.*

Proof (due to Oded Goldreich) Lets denote the function class of property of interest by $\{0,1\}$-LIN. We use the combination operator $\phi(f, g) = f + g$. We give a reduction from set disjointness to $C_\phi^{\{0,1\}-LIN}$. Since ϕ is a simple combination operator the result will then follow from the main reduction lemma 1. Consider the communication complexity problem of set disjointness. Let $A, B \subseteq [n]$ be the inputs of the two parties Alice and Bob. Based on these sets they can build linear functions $f, g \in \{0,1\}$-LIN such that the coefficient of an element is 0 if it does not belong to the set and its 1 if it does belong. Now consider the combining operator $\phi(f, g) = f + g$, if $|A \cap B| \geq 1$, $\phi(f, g) = h$ is a linear function in $GF(3)$ that has 2 as a coefficient of one of its variables. If $|A \cap B| = 0$, then $h \in \{0,1\}$-LIN. Let Q be the query complexity of $\mathcal{P}$. We can show that when the sets don't intersect the resulting function h is $\frac{2}{3}$ from having the property $\{0,1\}$-LIN. Now consider the difference $h-l$ where $l \in \{0,1\}$-LIN. We can show that the distance to any function in $\{0,1\}$-LIN is at-least $\frac{2}{3}$. By Schwartz zippel lemma we have that function $h-l$ for any linear function l will have the value 0 on at-most $\frac{1}{3}$ of the inputs. Which means h will have to be modified on at-least $\frac{2}{3}$ of the domain to fall in $\{0,1\}$-LIN and so h is $\frac{2}{3}$ far from the function class $\{0,1\}$-LIN.

7 Testing computability by width-4 OBDD

Definition OBDD: Given two finite sets X, Y, consider functions $f : X^n \to Y$, OBDD are directed acyclic graphs of say $n+1$ levels with nodes at each node having $|X|$ edges branching out from them. The nodes at the last level are the sink nodes and map to elements in Y. Each level is used to pick the edge with value that the variable corresponding to that level takes. A given input evaluates to the sink node that the traversal ends up at. Width of an OBDD is the maximum over number of nodes present at any level.

Theorem 6 *Testing the class of functions computable by width 4-OBDD requires $\Omega(n)$ queries.*

Proof We first construct functions that are hard in the sense that they can't be computed by width 4-OBDDs. Consider the following functions from $\{0,1\}^4 \to 0,1$

$$\begin{aligned}
\psi_0(x_1,x_2,x_3,x_4) &= 0 \\
\psi_1(x_1,x_2,x_3,x_4) &= x_1x_3 \\
\psi_2(x_1,x_2,x_3,x_4) &= x_2x_4 \\
\psi_3(x_1,x_2,x_3,x_4) &= x_1x_3 \oplus x_2x_4
\end{aligned}$$

Let $\tilde{n} = \lceil \frac{n-1}{4} \rceil$, Consider the class of functions obtained by the following definition

$$h(z) = x_1 \oplus \sum_{j=1}^{\tilde{n}} \psi_{z_j}(x_{4j-2}, x_{4j-1}, x_{4j}, x_{4j+1}) \text{ where } z \in \{0,1,2,3\}^{\tilde{n}}$$

Due to a lemma by Goldreich, we know that when $z \in \{0,1,2\}^{\tilde{n}}$, h is computable by a 4-OBDD and $\frac{1}{16}$ far to 4-OBDD computable functions when z has atleast single j such that $z_j = 3$. We now make use of this observation and do a reduction from DISJ$_{\tilde{n}}$. Alice uses $a \in \{0,1\}^n$ with 1 denoting presence of an element and Bob uses $b \in \{0,2\}^n$ with 2 denoting presence of an element to construct functions from the above class. The combining operator just adds up the indices, $\phi(f_a, g_b) = h_{a+b}$. Due to the way we defined the functions ψ_i above this evaluates to $h_z(x) = h_a(x) \oplus h_b(x) \oplus x_1$ which is a simple combination operator. We have one or more 3s when there is an intersection and when they are disjoint the entires consist of only 0,1 or 2. The reduction from $DISJ_{\tilde{n}}$ to $C_\phi^{4\text{-}OBDD\ computable}$ now falls in place due to the above mentioned lemma by Goldreich. From the main reduction lemma 1 it now follows that

$$2Q(4\text{-}OBDD \text{ computable}) \geq R(C^{4\text{-}OBDD \text{ computable}}) \geq R(DISJ_{\tilde{n}}) = \Omega(\tilde{n}) = \Omega(n)$$

8 Testing Monotonicity and Submodularity

Let $R \subseteq \mathbb{Z}$.

Definition Monotonicity: A function $f : \{0,1\}^n \to R$ is monotone if $\forall x \leq y$ i.e x, y that satisfy $x_i \leq y_i, \forall i$ we have $f(x) \leq f(y)$

Theorem 7 *Testing $f : \{0,1\}^n \to R$ for monotonicity needs $\Omega(\min\{n, |R|^2\})$ queries. (known upper bound is $O(n \log |R|)$ from.. [5])*

Proof Here we discuss the cases, $R = \mathbb{Z}$, $|R| = \sqrt{n}$ and $R = o(\sqrt{n})$. So we only leave out the case when $|R|$ and $\sqrt{n}$ are within multiplicative constants.

- Let R = $\mathbb{Z}$. We perform a reduction from DISJ$_n$ to montonicity testing. First Consider the combination operator $\phi(f(x), g(x), x) = 2|x| + f(x) + g(x)$. Consider the communication game $C_\phi^{Monotone}$. Now we perform the reduction from DISJ$_n$ to $C_\phi^{Monotone}$. Let Alice and Bob recieve set $A, B \subseteq [n]$ respectively. They then use the character functions $\chi_A(x)$ and $\chi_B(x)$ in the communication game. We claim that if A, B are disjoint $h(x) = \phi(f(x), g(x), x)$ is monotone and if $|A \cap B| \geq 1$ then the resulting h is 1/8 far from the class of $Monotone$ functions. We now prove this claim. Fix $i \in [n]$, let $x_0, x_1 \in \{0,1\}^n$ be the vectors with i^{th} bit set to 0 and 1 respectively.
  - When A, B are disjoint, if $i \notin A, i \notin B$ then $\chi_A(x_1) = \chi_A(x_0)$ and $\chi_B(x_1)\chi_A(x_0)$
  $$h(x_1) - h(x_0) = 2 > 0$$
  if $i \in A, i \notin B$ then $\chi_A(x_1) = -\chi_A(x_0)$ and $\chi_B(x_1)\chi_A(x_0)$
  $$h(x_1) - h(x_0) = 2 - 2\chi_A(x_0) \geq 0$$
  similarly when if $i \in B, i \notin A$ we have $h(x_1) - h(x_0) \geq 0$.
  So $h(x)$ is monotone when A, B are disjoint.
  - When $|A \cap B| \geq 1$, if $i \in A \cap B$ we have $\chi_A(x_1) = -\chi_A(x_0)$ and $\chi_B(x_1) = -\chi_B(x_0)$. So
  $$h(x_1) - h(x_0) = 2 - 2\chi_A(x_0) - 2\chi_B(x_0) > 0$$
  This means that $h(x_1) - h(x_0)$ could take the value -2 if $\chi_A(x_0) = \chi_B(x_0) = 1$. Consider the 2^{n-1} pairs of points of the form (x'_0, x'_1) obtainied by taking them along direction i. Amongst these pairs exactly 1/4 of them are such that the cardinality of $|a \cap x_0|$ and cardinality of $|b \cap x_0|$ are both even.(here a, b are bit vector representations of sets A, B.) So $h(x)$ has to be modified on atleast 1/4 of these 2^{n-1} pairs. We just used one such $i \in A \cap B$. This gives us that when $|A \cap B| \geq 1$, $h(x)$ is atleast 1/8 far from the class of monotone functions.

This completes the reduction from DISJ$_n$ to $C_\phi^{Monotone}$ and since ϕ is a simple combination operator the result follows from the main reduction lemma 1.

- When $|R| = \sqrt{n}$, we construct h' an appropriately scaled and truncated version of h for this restricted range R used in the above argument.

$$h'(x) = \begin{cases} -\sqrt{n}/2 & \text{when } |x| - n/2 < -\sqrt{n}/2 + 1 \\ \sqrt{n}/2 & \text{when } |x| - n/2 > \sqrt{n}/2 - 1 \\ |x| - n/2 + \frac{\chi_A(x)+\chi_B(x)}{2} & \text{when } ||x| - n/2| \leq \sqrt{n}/2 - 1 \end{cases}$$

By central limit theorem, h' falls to the extremes of the range R over only a constant fraction of x's. After this observation the proof is similar to above case, since h' is just $h/2$ around $n/2$ above. When A, B are disjoint h' is monotone and constant fraction from monotone when they intersect.

- For the last case, if $R = o(\sqrt{n})$, we show a reduction from the above case where $R = \sqrt{n}$. Let $g : \{0.1\}^m \to R$ be a function such that $m = |R|^2$. From $g(x)$ we can create the function $h : \{0,1\}^n \to R$ such that $h(x,y) = g(x)$ where $x \in \{0,1\}^m$ and $y \in \{0,1\}^{n-m}$. Now note that if $g(x)$ is monotone so is $h(x,y)$. Also if $g(x)$ is ϵ far from monotone, $h(x,y)$ is ϵ far from monotone because, for a given padding y_0 the fraction of changes that need to be made remains the same for h as in g and one has to make these many changes for all possible paddings y_0 to the input. This completes the reduction from the problem of property testing such a $g : \{0,1\}^m \to R$ with $|R| = \sqrt{m}$ to testing $f : \{0,1\}^n \to R$ with $|R| = o(\sqrt{n})$, giving us the required lower bound for this case, $\Omega(m) = \Omega(|R|^2)$.

This concludes the proof that testing monotonicity requires $\Omega(\min\{n, |R|^2\})$ queries. Note that, these upper and lower bounds do not say much for the case of Boolean functions $f : \{0,1\}^n \to \{0,1\}$.

Definition Submodularity: The real valued function $f : \{0,1\}^n \to \mathbb{R}$ is submodular if for every $x, y \in \{0,1\}^n$, $f(x \vee y) + f(x \wedge y) \leq f(x) + f(y)$.

The following lemma due to Seshadri and Vondrák [8] shows that submodularity testing is at-least as hard as monotonicity testing.

Lemma 8 *Given the function $f : \{0,1\}^n \to \mathbb{R}$, there exists a function, $g : \{0,1\}^{n+1} \to \mathbb{R}$ with the following properties:*

- *If f is monotone, then g is submodular.*
- *If f is ϵ-far from monotone, then g is $\epsilon/2$-far from sub-modular.*
- *For each $x \in \{0,1\}^{n+1}$, the value of $g(x)$ can be determined with 2 queries to f.*

Given a problem instance f for monotonicity testing one can run the sub-modularity tester for g, the oracle access to g can be obtained through oracle for f using 2 queries per each required query to g and the first two parts of the lemma guarantee correctness of sub-modularity testing imply correctness of monotonicity testing for f. From the lower bound for monotonicity testing we obtain that sub modularity testing requires at-least $\Omega(n)$ queries. Known upper bound for sub-modularity testing is $2^{O(\sqrt{n}\log n)}$ [8].

References

[1] Z. Bar-Yossef, TS Jayram, R. Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. In *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, pages 209–218. IEEE, 2002.

[2] E. Blais. Testing juntas nearly optimally. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 151–158. ACM, 2009.

[3] E. Blais, J. Brody, and K. Matulef. Property testing lower bounds via communication complexity. In *Computational Complexity (CCC), 2011 IEEE 26th Annual Conference on*, pages 210–220. IEEE, 2011.

[4] E. Blais and D. Kane. Tight bounds for testing k-linearity. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 435–446, 2012.

[5] Y. Dodis, O. Goldreich, E. Lehman, S. Raskhodnikova, D. Ron, and A. Samorodnitsky. Improved testing algorithms for monotonicity. *Randomization, Approximation, and Combinatorial Optimization. Algorithms and Techniques*, pages 97–108, 1999.

[6] B. Kalyanasundaram and G. Schintger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4):545–557, 1992.

[7] A.A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[8] C. Seshadhri and J. Vondrák. Is submodularity testable? *arXiv preprint arXiv:1008.0831*, 2010.