Outline

1、Deterministic lifting theorem for block sensitivity

2、$\Omega(n)$ lower bound for Set Disjointness

3、Intuition of $\Omega(n)$ lower bound for Tseitin Formula

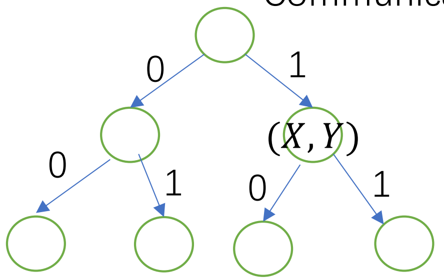
Proof overview

There exists a constant $c \geq 0$ *such that the following holds. Let* $f: \{0,1\}^n \to \{0,1\}$ *be any boolean function, and let* $g: [q] \times [q] \to \{0,1\}$ *be low discrepancy gadget function such that* $q \geq c$*. Then*

$$P^{CC}(f \circ g^n) = \Omega(bs(f) \cdot \log q)$$

Raz-McKenzie's simulation [RM99,GPW15]

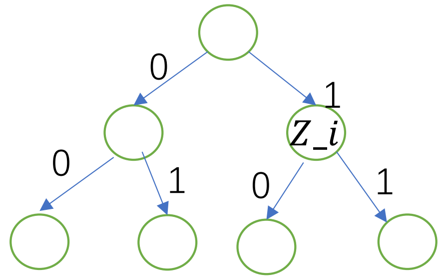
Communication by Alice or Bob



Convert communication protocol of $f \circ \mathrm{g}^n$

to decision tree for $f:\{0,1\}^n \rightarrow \{0,1\}$

Query



To show the simulation is correct,

1) **Thickness Lemma:** Average (degree) to worst (degree) reduction (or sunflower lemma [LMMPZ20]) to keep the disperser property:

$$\mathrm{g}^n(X_I, Y_I) = \{0,1\}^I$$ (g is gadget function and I is unfixed coordinates)

To show the simulation is efficient,

2) **Projection Lemma:** Potential function argument and probabilistic method to ensure

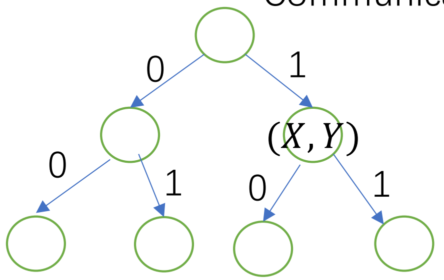
$\mathrm{g}(X_i, Y_i) = Z_i$ and potential function decreases by at least Ω(log q) in each "query iteration"

Raz-McKenzie' s simulation [RM99,GPW15]

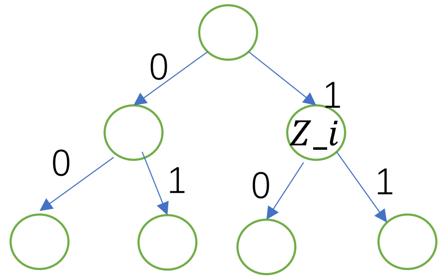
$f \circ g^n$

$f : \{0,1\}^n \to \{0,1\}$

Communication by Alice or Bob



Query



However, **Thickness Lemma** has gadget size barrier:

Average (degree) to worst (degree) reduction (or sunflower lemma [LMMPZ20]) need q= **Ω(n)**

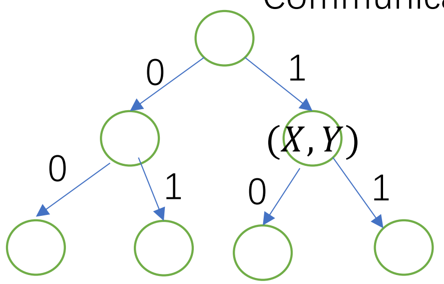
Gadget size is a fundamental parameter in lifting theorems and their applications [GP16, GJW16, GR18].

Our simulation

$f \circ g^n$

$f: \{0,1\}^n \to \{0,1\}$

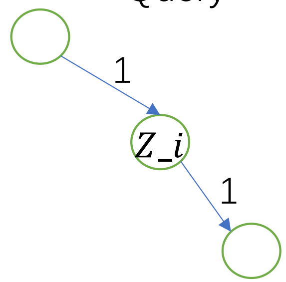
Communication by Alice or Bob



Find a communication path of $f \circ g^n$

by a decision tree path of $f: \{0,1\}^n \to \{0,1\}$

Query



To show the simulation is correct, we follow the decision tree path in **each "query iteration"**

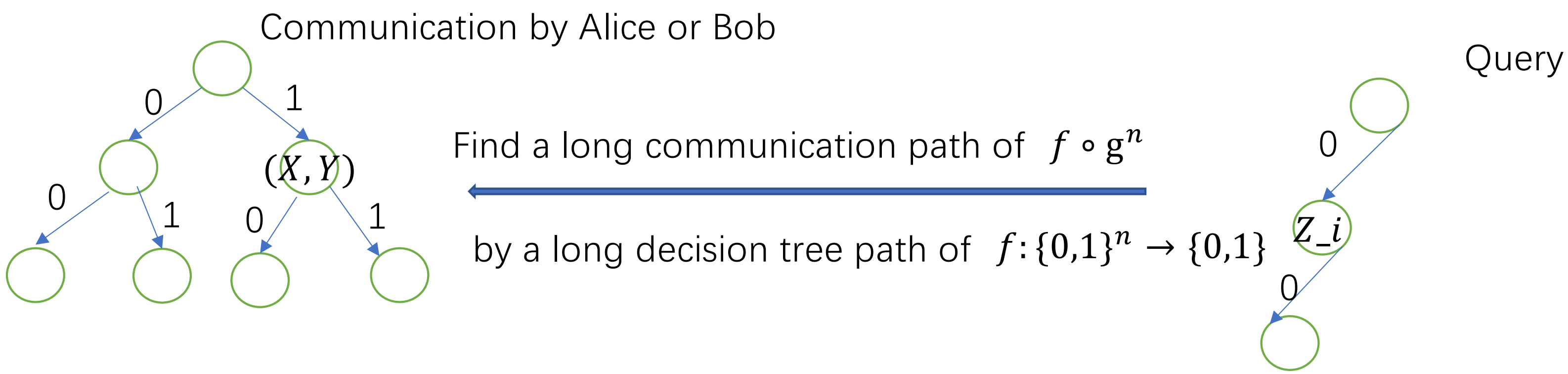
To show the simulation is efficient,

2) **Projection Lemma:** Potential function argument and low-discrepancy property [CFKMP19] to ensure

potential function decreases by at least $\Omega(\log q)$ in each "query iteration"

Our simulation

$f \circ g^n$



How to keep a long path ? Block sensitivity of $f: \{0,1\}^n \to \{0,1\}$ (a lower bound of query complexity)

WLOG Assume $bs(f, 0^n) = n,$ we could find a query path with $\Omega(n)$ length by always answer 0.

We need a new **Projection Lemma:**

1) Projection condition: $e_i \notin g^n(X, Y)$

2) Potential function: density of 0^n

3) Keep $g(X_i, Y_i) = 0$

Key Observation: Our projection can be delayed

Projection Lemma

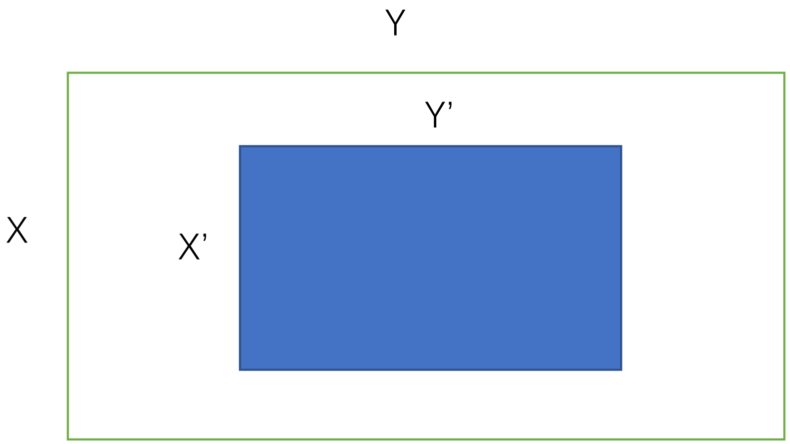
WLOG Assume $bs(f, 0^n) = n$,

1) Potential function: density of 0^n We define $D_0^I \coloneqq \{(x, y) \in [q]^I \times [q]^I : g^I(x, y) = 0^I\}$.

For each $I \subseteq [n]$ and $R = X \times Y \subseteq [q]^I \times [q]^I$. $we\ define\ it's\ potential\ function\ as$

$$E(R) := \log \frac{|R \cap D_0^I|}{|D_0^I|}$$

2) Projection condition: if $e_i \notin \mathrm{g}^n(X, Y)$, we do projection on i and keep $\mathrm{g}(X_i, Y_i) = 0$



Do projection on Alice's side: Find a $u \in [q]$

$$X' = \{x \in X : x_i = u\} \quad and \quad Y' = \{y \in Y : g(u, y_i) = 0\}$$

Do projection on Bob's side: Find a $\mathrm{v} \in [q]$

$$Y' = \{y \in Y : y_i = v\} \quad and \quad X' = \{x \in X : g(x\_i, v) = 0\}$$

Potential function argument

Assume $bs(f, 0^n) = n$, $\quad E(R) := \log \frac{|R \cap D_0^I|}{|D_0^I|}$

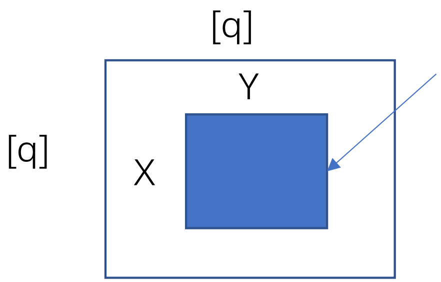
By average argument,

potential function decrease by at least O(1) in each "communication iteration"

if $e_i \notin g^n(X, Y)$, we do projection on i and keep $g(X_i, Y_i) = 0$

potential function increase by at least Ω(log q) in each "query iteration"

Low-discrepancy gadget: Fix $(x_{I\setminus i}, y_{I\setminus i})$ with $g^{I\setminus i}(x_{I\setminus i}, y_{I\setminus i}) = 0^{I\setminus i}$, If $e_i \notin g^I(X_I, Y_I)$ then $|X_i||Y_i| \le q$



$|X||Y| \le q$ When $1 \notin g(X, Y)$

By average argument, either we could do projection on Alice' side or Bob' s side to increase potential function.

$\Omega(n)$ lower bound for Set Disjointness

Previous proofs:

Entropy argument [Raz92]

Information complexity paradigm [BYJKS03] :

1、Direct sum argument

2、information complexity of AND is $\Omega(1)$:
Average encoding theorem + “Cut-and-Paste Lemma”

Our proofs:

Potential function argument:

1、Using entropy as potential function

2、Potential function increase by at least $\Omega(1)$ in each “query iteration” :

Projection Lemma + + “Cut-and-Paste Lemma”

Our goal

Let P be the hard input distribution $D_0^{[n]} = \{(x,y):\ \wedge^n(x,y) = 0^n\}$ $D_i^{[n]} = \{(x,y):\ \wedge^n(x,y) = e_i\}$

Main Lemma: $\sum \mathrm{P}\left(\mathrm{R} \cap D_i^{[n]}\right) + 2^{-\Omega(n)} \geq \Omega(\mathrm{P}\left(\mathrm{R} \cap D_0^{[n]}\right))$ (Corruption bound [Raz92])

For any rectangle R

Either

Projection Lemma + "Cut-and-Paste Lemma"

$$\sum \mathrm{P}\left(\mathrm{R} \cap D_i^{[n]}\right) \leq \epsilon \cdot \mathrm{P}\left(\mathrm{R} \cap D_0^{[n]}\right) \Longrightarrow \mathrm{P}\left(R \cap D_0^{[n]}\right) \leq 2^{-\Omega(n)}$$

Or

$$\sum \mathrm{P}\left(\mathrm{R} \cap D_i^{[n]}\right) \geq \epsilon \cdot \mathrm{P}\left(\mathrm{R} \cap D_0^{[n]}\right)$$

Hard input distribution P

We define the a hard distribution distribution P as follows,

1、Randomly sample a bit $b \in \{0,1\}$ and $i \in [n]$

2、 If $b = 0$, randomly sample (x, y) in $D_0^{[n]}$

If $b = 1$, randomly sample (x, y) in $D_\mathrm{i}^{[n]}$

Potential function: entropy of Q

Q be the distribution of P condition on $D_0^I \cap R$

$$E(Q) = H(Q) - |I| \log 3$$

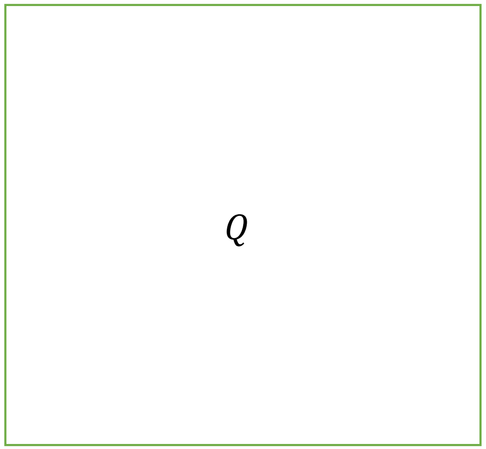
For any family of distributions $Q_1, \ldots Q_l$, and let Q be a linear combination of them, i.e., $\mathrm{Q} = \sum_i p_i \cdot Q_i$, the potential function of Q is defined as $E(Q) = \sum_j p_j \cdot E(Q_j)$.

Deterministic: Density of D_0^I

Randomized: Entropy of the distribution of D_0^I

How to do Projection?

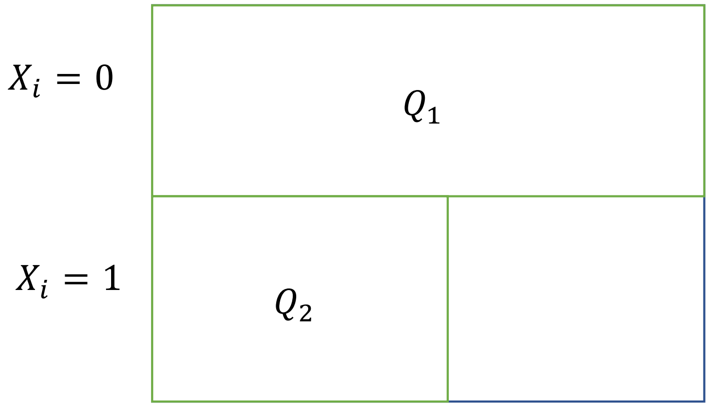
$E(Q) = H(Q) - |I| \log 3$



Do projection on Alice's side

and coordinate i

$E(Q) = p_1 \cdot E(Q_1) + p_2 \cdot E(Q_2).$



$p_1 = Q(X_i = 0)\ and\ \ p_2 = Q(X_i = 1)$

Similarly, we can do projection on Bob's side

Analysis of Projection Lemma

Our Goal: $E\left(Q_{I\setminus i}\right) = E(Q_I) + \Omega(1)$

Chain Rule of entropy: Let W be the random variable that $\Pr[W = 0] = p_1$ *and* $\Pr[W = 1] = p_2$

$$H(Q_I) = H(W) + H\left(Q_{I\setminus i} \middle| W\right) + H(Q_i \,| Q_{I\setminus i}, W)$$

$$E\left(Q_{I\setminus i}\right) = E(Q_I) + \log 3 - H(W) + \Pr[W = 0] \cdot H(Q_i \,| Q_{I\setminus i}, W = 0)$$

Since $H(Q_i \,| Q_{I\setminus i}, W = 1) = 0$, we need bound $\log 3 - H(W) + \Pr[W = 0] \cdot H(Q_i \,| Q_{I\setminus i}, W = 0) = \Omega(1)$

Let $t < 1$ be a constant, Assume $H(Q_i \,| Q_{I\setminus i}, W = 0) \leq t$

$$\log 3 - H(W) + \Pr[W = 0] \cdot \mathrm{t} = \log\left(\frac{3}{1 + 2^t}\right) = \Omega(1)$$

"Cut-and-Paste Lemma" (Connections between bias and potential function)

How to bound $H(Q_i \mid Q_{I\setminus i}, W = 0) \leq t$?

	Y = 0	Y = 1
X = 0	0	0
X = 1	0	1

Lemma:

If $\Pr_{(x,y)\sim P}[x \wedge y = 1] \leq \epsilon \cdot \Pr_{(x,y)\sim P}[x \wedge y = 0]$ where ϵ is a tiny constant. then either

$$\Pr_{(x,y)\sim Q}[x = 1 \mid y = 0] \leq \epsilon \cdot \Pr_{(x,y)\sim Q}[x = 0 | y = 0]$$

or

$$\Pr_{(x,y)\sim Q}[y = 1 | x = 0] \leq \epsilon \cdot \Pr_{(x,y)\sim Q}[y = 0 | x = 0]$$

Either entropy loss of X $(H(Q_i \mid Q_{I\setminus i}, X\_i = 0) \leq t)$ is Ω(1) or entropy loss of Y $(H(Q_i \mid Q_{I\setminus i}, Y\_i = 0) \leq t)$ is Ω(1)

So, we can do projection either on Alice' s side or Bob' s side.

Connections between information complexity

Projection Lemma + Chain Rule | Direct sum argument

K-UDISJ	Potential function in our simulation	Information complexity paradigm
$\Omega(n/k^4)$	Entropy	Entropy argument [AMS99] [Raz92]
$\Omega(n/k^2)$	KL divergence	Hellinger distance [BYJKS03]
$\Omega(n/k)$	Shifting + "KL divergence" (not finished yet)	Hellinger distance [Jay09] " KL divergence" [Gro09]

Intuition of $\Omega(n)$ lower bound for Tseitin Formula

Raz-McKenzie simulation : Maintain full set : $\mathrm{g}^n(X_I, Y_I) = \{0,1\}^I$

Our simulation in block sensitivity : density of **one string** with high block sensitivity

Question: How to maintain a subset set which is hard enough to prove lower bounds ?

We try it by use Tseitin Formula as a example

Tseitin Formula

Let $G = (V, E, l)$ be a connected labelled expander graph of maximum degree d where the labelling $l: V \rightarrow \{0,1\}$ has odd Hamming weight.

The **Tseitin formula** $\mathbf{Tse_G}$ associated with G is the d- CSP that has the edges $e \in E$ as variables and for each node $v \in V$ there is a constraint C_v defined by $\mathrm{C_v}(\alpha) = 1$ if and only if $\bigoplus_{e \in N(v)} \alpha(e) = l(v)$

It follows from a simple parity argument that is unsatisfiable.

Communication version: $\mathbf{Tse_G} \circ \boldsymbol{g^m(x, y)}$ $\quad \boldsymbol{where\ g\ is\ the\ constant\ gadget\ function}$

Previous results:

$\Omega(\mathrm{m}/\log m)$ lower bound in [GP14] via critical block sensitivity and $\Omega(m)$ lower bound in [PR17] via “degree to rank lifting”

$\Omega(n)$ query lower bound for Tseitin Formula

potential function argument：

1、Define a potential function which is large at the beginning and small in the end

2、In each query round, we set the value of edge e to maximize the potential function.

3、 Proving potential function is still large after $\Omega(n)$ queries.

$\Omega(n)$ query lower bound for Tseitin Formula

For assignment $\alpha \in \{0,1\}^m$, viol(α) be the set of unsatisfiable nodes

potential function ：density of 1-violation

Let $\mathrm{U} \subseteq \{0,1\}^m$ be the set of possible assignment. $R(U)$ = {v : there is a $\alpha \in U$, viol(α) = v }

Claim 1:

Let $\mathrm{U} \subseteq \{0,1\}^m$ be the set of all possible assignment, then for any query of $\mathrm{U} = \mathrm{U}_0 \cup \mathrm{U}_1$, exist $b \in \{0,1\}$, $|R(U_b)| \geq R(U)/2$

Claim 2:

Let $\mathrm{U} \subseteq \{0,1\}^m$ be the set of all possible assignment, If there is a $v \in [n]$ is violated by any assignment $\alpha \in U$, then $|R(U)| \leq 1$.

By Claim 1 and Claim 2, we only get a $\Omega(\log n)$ lower bound

$\Omega(n)$ query lower bound for Tseitin Formula

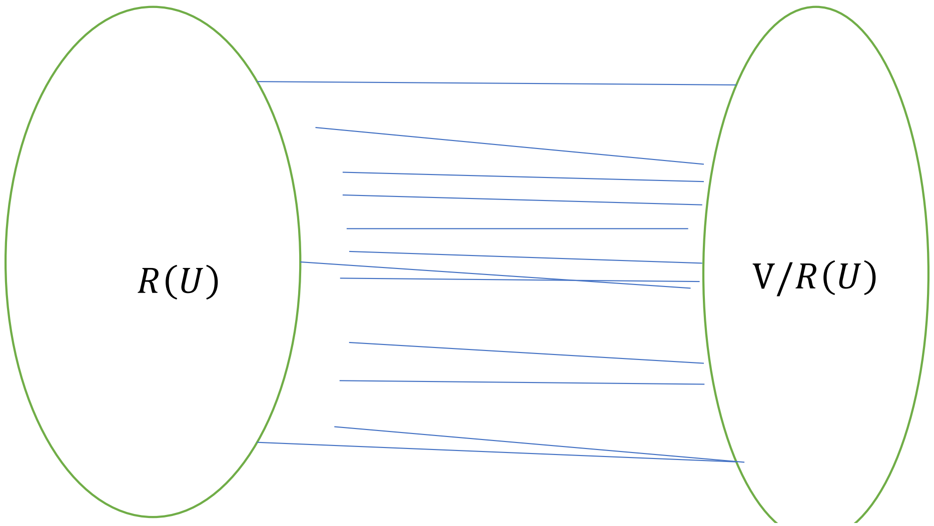
$R(U)$ = {v : there is a $\alpha \in U$, viol(α) = v }

Gap amplification for $R(U)$: (Using edge expansion of expander graph to maintain the subset $R(U)$)

Lemma 1.5 (edge expansion)**.** *Suppose G is a λ eigenvalue expander. Then for every $S \subseteq V$ with $|S| \leq n/2$ we have*

$$E(S, V - S) \geq \frac{d - \lambda}{2}|S|$$

When $\lambda = d - \Omega(1)$, then there is a constant c, $E(S, V - S) \geq c|S|$.



Observation 1:

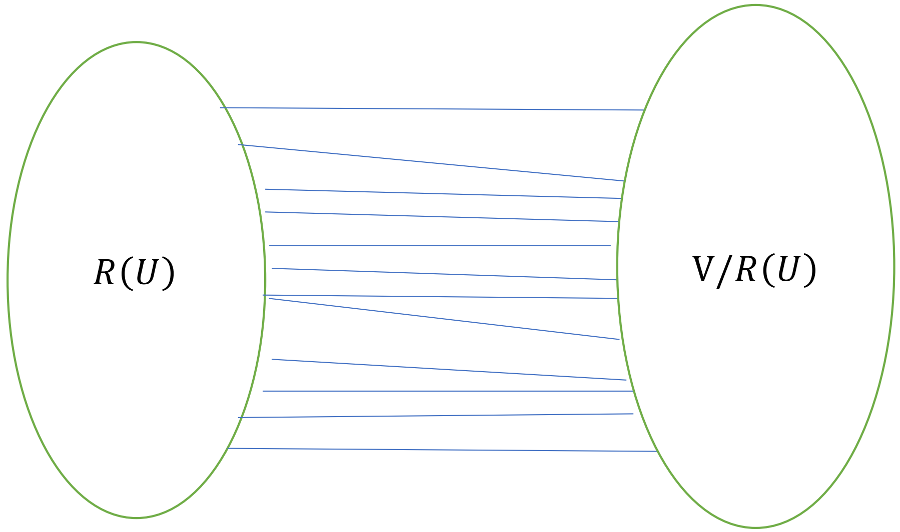
$$if \quad \frac{n}{2} \geq R(U) \geq \frac{n}{4},$$

$$then\ E\big(R(U), V - R(U)\big) = \Omega(n)$$

$\Omega(n)$ query lower bound for Tseitin Formula

$R(U)$ = {v : there is a $\alpha \in U$, viol(α) = v }

Gap amplification for $R(U)$: (Using edge expansion of expander graph to maintain the subset $R(U)$)



Observation 2:

Claim 1.6. *All edges in $E(R(U), V - R(U))$ must be queried.*

Proof. If $e = (u, v) \in E(R(U), V - R(U))$ is not queried, since there is assignment α with $\mathsf{Viol}(\alpha) = \{v\}$. then there is a assignment α' with $\mathsf{Viol}(\alpha) = \{u\}$. α' is obtained from α by flipping the values of edge e. $\square$

$\Omega(n)$ query lower bound for Tseitin Formula

For assignment $\alpha \in \{0,1\}^m$, viol(α) be the set of unsatisfiable nodes

potential function：density of 1-violation

Let $\mathrm{U} \subseteq \{0,1\}^m$ be the set of possible assignment. $R(U)$ = {v : there is a $\alpha \in U$, viol(α) = v }
and $R(U) = n$ at the beginning.

Claim 1:

Let $\mathrm{U} \subseteq \{0,1\}^m$ be the set of all possible assignment, then for
any query of $\mathrm{U} = \mathrm{U}_0 \cup \mathrm{U}_1$, exist $b \in \{0,1\}$, $|R(U_b)| \geq R(U)/2$

1、$if\ \frac{n}{2} \geq R(U) \geq \frac{n}{4}, then\ E\big(R(U), V - R(U)\big) = \Omega(n)$

2、All edges in $E\big(R(U), V - R(U)\big)$ must be queried.

Gap amplification !

Claim 2:

Let $\mathrm{U} \subseteq \{0,1\}^m$ be the set of all possible assignment, If there is a
$v \in [n]$ is violated by any assignment $\alpha \in U$, then $|R(U)| \leq 1$.

Communication lower bound for Tseitin Formula

How to extend the **Gap amplification lemma** to communication version?

Reference

[Raz92] On the distributional complexity of set disjointness

[RM99] Separation of the monotone nc hierarchy.

[AMS99] The space complexity of approximating the frequency moments

[BYJKS03] An information statistics approach to data stream and communication complexity

[Jay09] Hellinger Strikes Back: A Note on the Multi-party Information Complexity of AND

[Gro09] Asymptotically optimal lower bounds on the NIH-multi-party information complexity of the AND-function and disjointness

[GP14] Communication Lower Bounds via Critical Block Sensitivity

[GPW15] Deterministic Communication vs. Partition Number

[PR17] Strongly exponential lower bounds for monotone computation

[CFKMP19] Query-to-communication lifting using low-discrepancy gadgets