Multiparty Communication Complexity: Strong Fooling Sets and Applications to Streaming Statistics

Jiahui Liu, Ying Sheng, Anand Sundaram

Columbia University

May 5, 2017

Introduction

- 2-party Communication Complexity
- Multiparty Communication Complexity

Why multiparty?

We have seen a number of two-party communication complexity problems such as MAJ, XOR, EQ and their applications to obtaining lower bounds in circuit complexity, boolean functions and algorithms. Generally, the two-party communication complexity is a well-studied model [KN97][And79], while multi-party is relatively more mysterious.

Two models for Multiparty CC

- **Number-in-hand(NIH)**: The t players each holds an n-bit input $x_i \in \{0,1\}^n$. They wish to compute a joint function of their input $f(x_1, ..., x_t)$. They can communicate until one of them figures out the value of $f(x_1, ..., x_t)$ and returns it.
- **Number-on-the-forehead(NOF)**: Each player can see the inputs of all other players but cannot see his own input. Other setting is the same as in number-in-hand model.

Equivalent in two party communication case, entirely different for multi-party communication!
The NIH is useful in proving space lower bounds for streaming algorithms. The NOF model has applications to circuit complexity but since every player has a relatively large amount of information, the lower bound is hard to obtain. We do not cover NOF here.

Settings of Communication

- **blackboard model**: any message sent by a player is written on a board and all players can see.
- **message-passing model(private message)**: a player p_i sends messages to another specific player p_j.
- **coordinator model**: we have a $(t+1)$-th player as a coordinator who does not get any input. Players send messages only to the coordinator but not to each other.

Observation: the communication complexity in the coordinator model is just that from the message-passing model with a $\log t$ multiplicative factor. Each player p_i can send message (m, j) to the coordinator, where m is the original message he wants to send to p_j and j tells the coordinator to forward to player p_j. (if communication is not one-way)

Settings of Communication

- Sometimes the message can only be one-way
- Simultaneous Message: all players send messages simultaneously to coordinator

Deterministic and Randomized Protocols

Similar to 2-party CC, we can have deterministic and randomized Protocols,

- Deterministic: Bound analysis methods used in deterministic problems are often purely combinatorial
- Randomized: the players have access to public or private coins, which are equivalent to an infinite string of independent random bits. The protocol is allowed to return a wrong answer with probability of a small constant ϵ.
  Various methods: symmetrization, information complexity, direct sum arguments, etc.

Applications of NIH multiparty CC

A data stream is a sequence of data that is too large to be entirely stored in memory. Streaming algorithms are algorithms for computing, and often estimating some statistics with respect to an input data stream. The input may be examined in a few passes, but typically just one.
Problems people are usually interested

- frequency moments
- distinct elements
- heavy hitters
- streaming on graphs: edge connectivity, maximum bipartite matching, etc.

NIH multiparty communication complexity can be used to lower bound the space complexity of deterministic and randomized streaming algorithms.

What we will cover

- Strong Fooling Sets technique for lower bounds of deterministic protocols
- Space lower bounds for frequency moments
- Graph streams and graph communication problems
- Our own ideas inspired by strong fooling sets technique
- A glance of randomized protocols(if time permits)

Strong Fooling Sets: Concepts

Definition

Let $\mathcal{X}_1, ..., \mathcal{X}_t$ be finite sets. Put $\mathcal{X} := \mathcal{X}_1 \times ... \times \mathcal{X}_t$. $f : \mathcal{X} \rightsquigarrow \mathcal{Z}$ is a partial function. Or say, $f : \mathcal{X} \rightsquigarrow \mathcal{Z} \cup \{\star\}$, where $\star$ is a special "do-not-care" value. Denote t players as $PLR_1, ..., PLR_t$. Suppose the input for each PLR_i is $x_i \in \mathcal{X}_i$. The players should then communicate according to a deterministic protocol Π to output $\Pi^0(x) \in \mathcal{Z}$. Π computes f if

$$\forall x \in \mathcal{X} : f(x) \neq \star \Rightarrow \Pi^0(x) = f(x)$$

- Discreet(Private) Protocol
- Blackboard Protocol

Discrete Protocol & Blackboard Protocol

- Discreet(Private) Protocol:

$$cost(\Pi) := \max_{i,x} |\Pi_i(x)| = \min\{B : \Pi \text{ is B-bounded}\}$$

$$DD(f) := \min\{cost(\Pi) : \text{discreet protocol } \Pi \text{ computes } f\}$$

- Blackboard Protocol:

$$cost^{tot}(\Pi) := \max_{x}(|Pi_1^w(x)| + \cdots + |\Pi_t^w(x)|)$$

$$BB^{tot}(f) := \min\{cost^{tot}(\Pi) : \Pi \text{ computes } f\}$$

Relation:

$$cost^{tot}(\Pi) \leq \frac{t}{2} cost(\Pi) \tag{1}$$

Weak Fooling Set bound

- Combinatorial Rectangle: $\mathcal{Y}_1 \times \cdots \times \mathcal{Y}_t \subseteq \mathcal{X}$
- Each equivalence class of a blackboard protocol Π is a rectangle in $\mathcal{X}$. In particular, if $\mathcal{F} \subseteq \mathcal{X}$ lies within an equivalence class, then so does $span(\mathcal{F})$. Consequently, if Π computes a partial function f, then f is constant on $span(\mathcal{F})$.
- suppose the communication game is $f : \mathcal{X} \rightsquigarrow \mathcal{Z}$. A set $\mathcal{F}$ is a K-weak-fooling-set for f if
  $\forall \mathcal{F}' \subseteq \mathcal{F}$ with $|\mathcal{F}'| > K, span(\mathcal{F}')$ is not constant on f

Weak Fooling Set bound

Lemma (Weak-fooling-set bound)

Suppose that $f : \mathcal{X} \rightsquigarrow \mathcal{Z}$ specifies a t-player communication game and that f has a K-weak-fooling-set $\mathcal{F}$. Then

$$BB^{tot}(f) \geq \log \frac{|\mathcal{F}|}{K} \Rightarrow DD(f) \geq \frac{2}{t} \log \frac{|\mathcal{F}|}{K}$$

Neighborhood

Definition

Let $\mathcal{G} \subseteq \mathcal{X}$ be nonempty. A *neighborhood* within $\mathcal{G}$ is a t-tuple $\mathcal{N} = (\mathcal{H}_1, ..., \mathcal{H}_t)$ where each $\mathcal{H}_i \subseteq \mathcal{G}$ and $core(\mathcal{N}) := \mathcal{H}_1 \cap ... \cap \mathcal{H}_t$ is nonempty. And $wid(\mathcal{N}) := \min\{|H_1|, ..., |H_t|\}$.

Lemma (Generalized Rectangle Property)

Let Π be a discrete protocol on input space $\mathcal{X}$. Let $\mathcal{N}$ be a neighborhood within $\mathcal{X}$ such that Π is smooth on $\mathcal{N}$. Then $span(\mathcal{N})$ lies within an equivalence class of Π. Consequently, if Π computes a partial function f, then f is constant on $span(\mathcal{N})$.

Proof.

$(x_1, ...x_t) \to (y_1, x_2, ...x_t) \to \cdots \to (y_1, ..., y_t)$ □

Strong Fooling Set bound

Lemma

Let Π be a B-bounded t-player discreet protocol on input space $\mathcal{X}$. Let $\mathcal{G} \subseteq \mathcal{X}$. Then there exists a neighborhood $\mathcal{N}$ within $\mathcal{G}$ such that Π is smooth on $\mathcal{N}$ and $wid(\mathcal{N}) \geq |\mathcal{G}|/(t2^B)$.

Lemma

Suppose the finite set S is partitioned into L blocks and $s \in_R S$ is picked uniformly at random. For every real $A > 0$,
$\Pr[|\textit{block contains } s| < |S|/(AL)] < 1/A$.

Probabilistic method:

$$\Pr\left[wid(\mathcal{N}_x) \geq \frac{\mathcal{G}}{t2^B}\right] = 1 - \Pr\left[\exists i : |H_i| < \frac{\mathcal{G}}{t2^B}\right]$$

$$\geq 1 - \sum_{i=1}^{t} \Pr\left[|H_i| < \frac{\mathcal{G}}{t2^B}\right]$$

$$> 1 - \sum_{i=1}^{t} \frac{1}{t} = 0$$

Strong Fooling Set bound

Definition

Let $f : \mathcal{X} \rightsquigarrow \mathcal{Z}$ specify a communication game and let $\mathcal{F} \subseteq \mathcal{X}$. $\mathcal{F}$ is a K-fooling-set for f, if for every neighborhood $\mathcal{N}$ within $\mathcal{F}$,

$$wid(\mathcal{N}) > K \Longrightarrow f \text{ is nonconstant on } span(\mathcal{N})$$

- For a B-bounded discreet protocol Π for f. If $\mathcal{F}$ is a K-fooling-set for f.
- There exists a neighborhood $\mathcal{N}$ within $\mathcal{F} \subseteq \mathcal{X}$ such that Π is smooth on $\mathcal{N}$ and $wid(\mathcal{N}) \geq |\mathcal{F}|/(t2^B)$.
- f is constant on $span(\mathcal{N})$.
- Then there is $wid(\mathcal{N}) \geq |\mathcal{F}|/(t2^B) \leq K$.
- Thus, $DD(f) \geq \log \frac{|\mathcal{F}|}{tK}$ (**Strong Fooling Set bound**)

m,mm

Lower bounds for EQ-DIST

$$\text{EQ-DIST}_{n,t}(x_1, ..., x_t) = \begin{cases} 1, & \text{if } x_1 = \cdots = x_t, \\ 0, & \text{if } x_i \neq x_j, \forall i, j, s.t. 1 \leq i < j \leq t, \\ \star, & \text{otherwise.} \end{cases}$$

$\mathcal{F} := \{x^{\otimes t} : x \in \{0,1\}^n\}$ is a $(t-1)$-fooling set of f. If $wid(\mathcal{N}) \geqslant t$

- 1-instance: $\emptyset \neq core(\mathcal{N}) = H_1 \cap \cdots \cap H_t \subseteq span(\mathcal{N})$
- 0-instance: since $wid(\mathcal{N}) \geqslant t$, let each player i select an element in H_i that is different with all prior selections.

$$DD(EQ - DIST_{n,t}) \geq \log \frac{2^n}{t(t-1)} = n - \log(t^2 - 1) \geq n - 2\log t$$

EQ-SPRD Problem

Input: $|x_1| = \cdots = |x_t| = \lceil \beta n \rceil$

$$\text{EQ-SPRD}_{n,t}^{\beta,\gamma}(x_1, ..., x_t) = \begin{cases} 1, & \text{if } x_1 = \cdots = x_t, \\ 0, & \text{if } |x_1 \cup \cdots \cup x_t| \geq \gamma n, \\ \star, & \text{otherwise.} \end{cases}$$

Lower bounds for EQ-SPRD – large t

Theorem

For all values $0 < \beta < \gamma \leq 1$ *and sufficiently large* n*, if* $t \geq \gamma n$*, then* $DD(EQ\text{-}SPRD^{\beta,\gamma}_{n,t}) \geq (\beta \log(1/\gamma))n - \log t$

$\mathcal{F} = \{x^{\otimes t} : x \in \{0,1\}^n\}$ is a $\binom{\lfloor \gamma n \rfloor}{\beta n}$-fooling set, when $t \geq \gamma n$. If $wid(\mathcal{N}) > \binom{\lfloor \gamma n \rfloor}{\beta n}$:

- 1-instance: $\emptyset \neq core(\mathcal{N}) = H_1 \cap \cdots \cap H_t \subseteq span(\mathcal{N})$
- 0-instance: let each player i select a subset y_i in H_i that enlarge the current union $U_{i-1} = y_1 \cup \cdots \cup y_{i-1}$ by 1.

$$DD(\text{EQ-SPRD}^{\beta,\gamma}_{n,t}) \geq \log \frac{\binom{n}{\beta n}}{t\binom{\lfloor \gamma n \rfloor}{\beta n}} \geq \log \frac{\left(\frac{n}{\gamma n}\right)^{\beta n}}{t} \geq \beta n \log \frac{1}{\gamma} - \log t$$

Lower bounds for EQUALITY – small t

Theorem

For all values $t \geq 2, \beta > 0, \gamma \leq \beta t(1 - e\beta t) > \beta$, *and sufficiently large integral* n, *we have*

$DD(EQ\text{-}SPRD^{\beta,\gamma}_{n,t}) \geq 2e\beta^2 n - 2\log t - \Theta(1)$

Proof: Using Error Correcting Code to reduce it to EQ-DIST problem.

Lower bounds for EQUALITY – small t

- Suppose $\mathcal{C}$ is a set of subsets of $[n]$, i.e. $\mathcal{C} \subseteq 2^{[n]}$, it is an (r, s, n)-packing if:
  - $\forall A \in \mathcal{C}$, $|A| = s$
  - $\forall A, B \in \mathcal{C}, A \neq B$, $|A \cap B| \leq r$
- For all values $0 \leq r \leq s \leq n$, there exists an (r, s, n)-packing $\mathcal{C}$ such that

$$|\mathcal{C}| \geq \binom{n}{r} \Big/ \binom{s}{r}^2 .$$

- Set $s = \lceil \beta n \rceil$, $r = 2\lceil e\beta s \rceil$

$$|\mathcal{C}| \geq \frac{\binom{n}{r}}{\binom{s}{r}^2} \geq \left(\frac{n}{s}\right)^r \left(\frac{r}{es}\right)^r = \left(\frac{nr}{es^2}\right)^r$$

Lower bounds for EQUALITY – small t

To find an injection:

$$r \log \frac{nr}{es^2} \geq r \log \frac{2e\beta sn}{es^2} = r \log \frac{2\beta n}{\lceil \beta n \rceil} = 2e\beta^2 n - \Theta(1) \geq N$$

Reduction:

- Obviously, the 1-input for EQ-DIST is also 1-input for EQ-SPRD.
- For 0-input $(x_1, ..., x_t)$ maps to $(y_1, ..., y_t)$ in EQ-SPRD. Since all y_i's are different, from the packing property, there is:

$$|x_1 \cup \cdots \cup x_t| \geq ts - \binom{t}{2} r = ts - t(t-1)\lceil e\beta s \rceil \geq \lceil \beta n \rceil t(1 - e\beta t) \geq \gamma n$$

so, it is also an 0-input for EQ-SPRD.

$$DD(EQ - SPRD_{n,t}^{\beta,\gamma}) \geq DD(EQ - DIST_{N,t}) \geq 2e\beta^2 n - 2\log t - \Theta(1)$$

Failed attempt to improve lower bound

Recall:

Theorem

For all values $0 < \beta < \gamma \leq 1$ *and sufficiently large* n*, if* $t \geq \gamma n$*, then* $DD(\text{EQ-SPRD}^{\beta,\gamma}_{n,t}) \geq (\beta \log(1/\gamma))n - \log t$

$\mathcal{F} = \{x^{\otimes t} : x \in \{0,1\}^n\}$ is a $\binom{\lfloor \gamma n \rfloor}{\beta n}$-fooling set, when $t \geq \gamma n$.

We guess that the constraint $t \geq \gamma n$ could be relaxed.

- Intuitively, every time the next player may could enlarge more than 1 element. So we tried to consider it in expectation, then use the probabilistic method. It is hard to bound, and actually union bound is far from enough.
- If there is not any constraints, there is a contra-example: Just consider every subsets contains a common element. We are not sure, if it is possible to improve if we add some constraints.

Discussions

- For EQ-DIST: trivial upper bound $n+1$, so the lower bound$(n-2\log t)$ is tight up to lower order terms.
- For EQ-SPRD: trivial upper bound $2\log\binom{n}{\lceil\beta n\rceil} \leq 2H(\beta)n$ when $\beta < 1/2$. So, the lower bound is tight on the order of n.
- The upper bound for EQ-DIST under blackboard protocol could reach $O(n/t)$. So the lower bound $n-2\log t$ confirmed the separation between blackboard protocol and discreet(private) protocol.

Streaming Lower Bounds: Frequency Moments

Let $x = (x_1, ..., x_t), x_i \in \{0,1\}^n$ be an input for EQ-SPRD$^{\beta,\gamma}_{n,t}$. The corresponding frequency vector $f = x_1 + ... + x_t$.

Small moments: k < 1

For $k < 1$, we have (from [CK16]):

$$\text{EQ-SPRD}^{\beta,\gamma}_{n,t}(x) = 1 \Rightarrow F_k(f) = \beta t^k n$$

$$\text{EQ-SPRD}^{\beta,\gamma}_{n,t}(x) = 0 \Rightarrow F_k(f) \geq F_0(f) \geq \gamma n$$

As a result:

Theorem

For each $k \in [0,1)$, every deterministic s-space p-pass α-estimator for F_k satisfies $ps = \Omega(max\{n^{1-k}/\alpha, n/\alpha^{2/(1-k)}\})$.

Streaming Lower Bounds: Frequency Moments

Let $x = (x_1, ..., x_t), x_i \in \{0,1\}^n$ be an input for EQ-SPRD$^{\beta,\gamma}_{n,t}$. The corresponding frequency vector $f = x_1 + ... + x_t$.

Intermediate case: k = 1

F_1 is just stream length, so trivially:

$$ps = \Omega(\log m)$$

Large moments: k > 1

Here we have, by arguments from convexity ([CK16]):

$$ps = \Omega(n/\alpha^{2k/(k-1)})$$

By concentration bounds on power sums of random variables and existential arguments about mapping matrices, this can be improved to:

$$\Omega(n/\alpha^{k/(k-1)}(\log\log\alpha)^2)$$

Streaming Lower Bounds: Frequency Moments

Upper bounds (also from [CK16])

Theorem

For integers $p > 1$, and reals $k > 0$ and $\alpha > 1$, there is a family of deterministic p-pass α-estimators for F_k with the following guarantees on their space usage, s:

- *When $k = 0$, we have $ps = \lceil n/\alpha \rceil + O(\log n)$.*
- *When $0 < k < 1$, we have $ps = O(n \log m/\alpha^{1/(1-k)})$*
- *When $k = 1$, at $p = 1$ we have $s \leq \lceil \log m \rceil$, trivially.*
- *When $k > 1$, we have $ps = O(n \log m/\alpha^{1/(k-1)})$*

Streaming Lower Bounds: Frequency Moments

Longer Streams

Use more players to make a longer stream? Worsens the best lower bounds, at least for EQ-DIST.

1. $DD(\text{EQ-DIST}_{n,t}) \geq n - 2 \cdot \log t$ ([CK16])
2. $DD(\text{EQ-DIST}_{n,t}) \leq n + 1$ ([CK16])
3. $DD(\text{EQ-DIST}_{n,t}) \leq 2(n - \log t + 2)$
   (tighter upper bound for many players with streaming protocol)

Streaming Lower Bounds: Frequency Moments

Longer Streams

What about lengthening the stream by just having each player put multiple copies of their input in the stream? But longer streams don't produce a larger ratio between the F_k of YES and NO instances or allow larger fooling sets.

For a fixed stream length m:

$$n\lfloor\frac{m}{n}\rfloor^k \leq F_k \leq m^k$$

For any two streams s_1 and s_2 of length m:

$$\frac{s_1}{s_2} \leq n^{k-1}$$

So we can only argue that at most $\log_\alpha n^{k-1}$ distinct streams of length m must be pairwise distinguished by an α-estimator for F_k.

Streaming Lower Bounds: Frequency Moments

Longer streams

Can show some relationship between space complexity and length.

Define an infinite sequence $M = m_1, m_2 \ldots$ where:

$$m_i = n^{2i}$$

$$M = n^2, n^4, n^6, \ldots$$

Can show for all $i < j$, for any stream s_1 of length m_i and any stream s_2 of length m_j:

$$\frac{F_k(s_2)}{F_k(s_1)} \geq n^k$$

So any α-estimator for non-trivial $\alpha < n^{k-1}$ must be able to estimate $\log_{n^2}(m)$, which requires $\Omega(\log \log m)$ bits as $m \gg n$.

Streaming Lower Bounds: Frequency Moments

Turnstile Model

Some vector $x \in \{-m, -m+1, \ldots, m-1, m\}^n$ as stream of element-wise $\pm$ increments

Negative updates can cancel out positive updates in the turnstile model to produce a small F_k for a large stream

Have lower bounds ([WW15]):

Theorem

For any $\epsilon \in (0,1)$ and any $\delta \geq 2$, any algorithm obtaining a $(1+\epsilon)$ approximation with probability at least $1-\delta$ to F_k requires $\Omega\left(n^{\frac{1-2}{k}} \epsilon^{-2} \log m \log \frac{1}{\delta}\right)$. This also implies that algorithms with high success probability (where $\delta = O(\frac{1}{n})$), including all deterministic algorithms, require $\Omega\left(n^{\frac{1-2}{k}} \epsilon^{-2} \log m \log n\right)$ space.

Streaming Lower Bounds: Frequency Moments

Augmentation by problem copies

What if there are g groups of t players, and for each group i, all its members have i copies of every element from subset of size βn in an instance of EQ-SPRD, but together the players must compute the output of each EQ-SPRD problem?

The players can just pass along g sketches in parallel and only use streams of length βn. The number of copies held doesn't need to influence the stream reduction, only which sketch is updated.

Reduction to varying length streams

Seems promising, but how?

Graph Streams: Maximum Matching Size Estimation(MMSE)

- **Input:** A bipartite graph $G_n = (V_1, V_2, E)$, with one set of vertices $|V_2| = n$ and the other set of vertices $|V_1| = O(n)$. The stream gives each vertex $u \in V_1$ and all its neighbors in V_2 every time("vertex-arrival" model).
- **Output**: An estimate of edges in a maximum cardinality matching.

Graph Streams: Maximum Matching Size Estimation(MMSE)

Previous Results on MMSE

- A 2-estimator algorithm: maintain a maximal matching using $n\lceil \log n \rceil$ space.
- An $O(\sqrt{n})$-estimator algorithm: maintain a randomized sketch using $O(\mathrm{poly}\log n)$ space.
- [KKS14]: an $O(\mathrm{poly}\log n)$-estimator using $O(\mathrm{poly}\log n)$ space over randomly ordered streams.
- [EHL$^+$15]: $\Omega(\sqrt{n})$ space lower bound for randomized one-pass $(3/2-\epsilon)$-estimators and $\Omega(n)$ space lower bound for deterministic one-pass $(3/2-\epsilon)$-estimators. This paper used reduction from a communication problem called Boolean Hidden Matching.

Most of the results are more algorithmic.

MMSE: Results from Lower Bounds of EQ-DR

There are $t = \lfloor \gamma n \rfloor$ players. Each player receives a $\lceil \beta n \rceil$ subset of n. The problem is to distinguish the the case when all subsets are equal from the case when each player picks a representative element from his/her subset so that these representatives are distinct.

$$EQ - DR^{\beta}_{n,t} = \begin{cases} 1, & \text{if } |x_1| = |x_2| = ... = |x_t| = \lceil \beta n \rceil \\ & \text{and } x_1 = ... = x_t \\ 0, & \text{if } |x_1| = |x_2| = ... = |x_t| = \lceil \beta n \rceil \\ & \text{and } \exists g : [t] \rightarrow x_1 \cup ... \cup x_t \\ & \text{such that } g \text{ is injective and } g(i) \in x_i, \forall i \in [t] \\ *, & \text{otherwise} \end{cases}$$

MMSE: Results from Lower Bounds of EQ-DR

A weaker version of EQ-SPRD!
The proof for lower bound of Equal-vs-Spread applies to analyzing lower bound proof of EQ-DR by letting $\gamma = t/n$ and using the same construction for $\mathbf{y} = \text{span}(\mathcal{N})$. Then we can easily get the lower bound: for all values $0 < \beta < 1$, $\epsilon > 0$ and sufficiently large n, if $(\beta + \epsilon)n \leq t < n$, then we have

$$\text{DD}(\text{EQ-DR}^{\beta}_{n,t}) \geq (\beta \log(n/t))n - \log t$$

MMSE: Results from Lower Bounds of EQ-DR

Then we can reduce EQ-DR$^{\beta}_{n,t}$ to MMSE:

- set the two sets of vertices to be $V_1 = \{u_1, ..., u_t\}$ and $V_2 = [n]$.
- For $i \in [t]$, player i adds edges $\{\{u_i, j\}, j \in x_i\}$ to construct a graph G.
- By definition of EQ-DR$^{\beta}_{n,t}$, there is an injective mapping from $[t]$ to $[n]$.
- If EQ-DR$^{\beta}_{n,t}(x_1, ..., x_t) = 1$, then an MCM in G has size βn
- if EQ-DR$^{\beta}_{n,t}(x_1, ..., x_t) = 0$, then an MCM in G has size t

To optimize the original EQ-DR$^{\beta}_{n,t}$ lower bound, we set $t = n/e$ and $\beta = 1/(\alpha e) - 1/n$ for some $\alpha < t/\beta n$. Then we can obtain a lower bound of $(\frac{n}{e\alpha} \log e - \log n)/2$ for an $\alpha-$estimator.

Maximum Matching in Simultaneous Message Model

Consider a related problem from [DNO14]:

n players together have a bipartite graph $G = (V_1, V_2, E)$, with $|V_1| = |V_2| = n$. Each player gets as input the set of neighbors of a vertex in V_1. They send a (possibly) randomized message to a coordinator **simultaneously** who has to output a perfect matching M, but M may contain edges not in E. The goal is to maximize $|M \cap E|$.

This problem can be modified that the coordinator has to just estimate the maximum matching size. And the lower bound reduced from EQ-DR in [CK16] can also be generalized to the SM model.

Edge Connectivity

Divert to 2-party CC here....

Dynamic graph connectivity problem XCONN:

Alice and Bob get inputs E_A and E_B which are edges on the vertex set $[n]$. They should determine whether the graph $E_A \oplus E_B := (E_A \cup E_B) - (E_A \cap E_B)$ is connected.

Reduction EQ to XCONN

Alice and Bob each has a complete graph with size $n/2$. And they hold the vectors indicate the edges between these two components.

- Alice adds a complete graph on $[n/2]$ and Bob adds a complete graph on $[n] - [n/2]$.
- The inputs for $EQ_{n^2/4}$ are encoded on the edges in $[n/2] \times ([n] - [n/2])$: if Alice and Bob hold a same edge, this corresponds to one equal bit
- When EQ $= 1$, XCONN $= 0$; when EQ $= 0$, XCONN $= 1$.

Thus the communication complexity of XCONN is at least $\frac{n^2}{4}$.

Edge Connectivity with strong promise

Still hard!
Even with the promise that $(E_A \oplus E_B)$ is disconnected or $(n/2-1)$-connected(i.e. at least need to remove $(n/2-1)$ edges to disconnect it), the communication complexity of determine connectivity is still $\Omega(n^2)$.
Do the same reduction as the above from EQ_{N^2} where $N = \Omega(n)$ and use a binary Error Correcting Code of size 2^{N^2} , block length $n^2/4$, and distance $n/2-1$.

- if EQ $= 1$, $E_A \oplus E_B$ is disconnected
- if EQ $=0$, then the corresponding ECC has a distance of $(n/2-1)$, so there are at least $n/2-1$ edges from Alice's set $[n/2]$ to Bob's $[n]-[n/2]$ and each of them holds a complete graph, so $E_A \oplus E_B$ is $(n/2-1)$-connected.

Separation between Deterministic and Randomized protocols

Randomized protocol is more powerful.

Example

- Separation of randomized and deterministic protocols for XCONN: $O(n \log^3 n)$ randomized protocol upper bound from [AGM12].
  Alice can send Bob the sketch for connectivity. Bob can solve XCONN using this sketch.
- The results in [CK16] show that deterministic approximation of frequency moments within any constant factor requires linear space(except $k = 1$), hence demonstrating a separation from randomized streaming which require only $\tilde{O}(1)$ space for $k \leq 2$ and $o(n)$ for $k > 2$ to $(1 \pm \epsilon)$-approximate F_k.

Randomized Protocols

While randomized protocol is more powerful, in analyzing lower bounds for them, sometimes need to take use of deterministic protocols.
One of the most important principles that relate deterministic to randomized protocols is Yao's Minimax Principle, which gives an equivalence between two kinds of randomness in algorithms: randomness inside the algorithm itself, and randomness on the inputs. A lot of randomized lower bounds can be proved using its idea.

Randomized Protocols

Fix some model of computation for computing a Boolean function F. Let $R_\epsilon(F)$ be the minimal complexity among all randomized algorithms that compute $F(x)$ with success probability at least $1-\epsilon$, for all inputs x. Let $D^\mu_\epsilon(F)$ be the minimal complexity among all deterministic algorithms that compute F correctly on a fraction of at least $(1-\epsilon)$ of all inputs, weighed according to a distribution μ on the inputs. Yao's principle tells us that these two complexities are equal if we look at the "hardest" input distribution μ:

$$R_\epsilon(F) = \max_\mu D^\mu_\epsilon(F)$$

Therefore in order to prove a lower bound for randomized protocol it suffices to find a hard distribution and prove a distributional lower bound for it.

Randomized Protocols

See more randomized techniques in our report....!

- Symmetrization
- Information Complexity
- Direct Sum argument

And various applications to randomized streaming algorithms.

Other Applications of NIH?

Applications of NIH deterministic multiparty communication complexity beyond streaming algorithms lower bound exist.
The parity decision tree complexity $D_{\oplus}(f)$ and the t-party XOR functions $F(x_1, ...x_t) = f(x_1 \oplus ... \oplus x_t)$ have the relationship $CC^{(k)}(F) \leq k \cdot D_{\oplus}(f)$
[Yao15] shows that for 4-party XOR:

$$D_{\oplus}(f) \leq O(CC^{(4)}(F)^5)$$

Open Questions

- The lower bounds discussed in this paper are all on the insertion-only model; could these techniques be used to prove tighter lower bounds for deterministic estimators in the turnstile model?
- The l.b. and u.b. of F_k estimation when $k \leq 1$ are not very close. Could we improve it?
- Although the authors said that the lower bound of F_k estimation when $0 \leq k < 1$ is tight, there is still a $\log m$ gap, where the m is the length of the stream. Could we still improve it? The lower bound even do not contains m. Actually, the estimator for F_k and entropy given in this paper is very simple and rough. Though it is lying in the algorithm side, but still worth to think could we improve the upper bound? The estimator given for entropy using 2-pass, is that possible just using 1-pass?

- The lower bounds shown in graph streams are all consider bipartite graph. Could we do something on non-bipartite graph? Or could we do something on MCM rather than MMSE?
- Let's formulate the problems in the strong fooling sets paper into a randomized context. Can we apply the techniques for randomized protocols presented in this report(symmetrization, information complexity) to them?
- Following the above question, what kind of hard input distribution can we construct for EQ-SPRD and EQ-DIST to apply the above techniques?

Conclusions

- The Strong Fooling Sets method is a powerful tool for deterministic NIH multiparty communication complexity. Some bounds obtained in [CK16] seem to be not very tight, but it's also non-trivial to tighten these bounds using only tools from the paper. We may need to utilize other tools or develop new ones.

Conclusions

- Throughout this project, we have seen many examples demonstrating sharp contrast, i.e. separation between different models. A slight change in the communication model can result in lower/upper bounds with significant difference and also a change in analysis methods. For example: private-message versus blackboard model; simultaneous message passing versus general one-way protocol, not to mention with-promise versus no-promise and deterministic versus randomized.
- Promise on the input of a problem sometimes helps with analysis of lower bounds and reduction to other problems, but does not necessarily lower the difficulty of the problem. For example, the edge connectivity problem with strong promise still requires $\Omega(n^2)$ space.

Conclusions in terms of good research...

- We surveyed on a few techniques used for analyzing multiparty communication complexity for this project. However, we have learned throughout the progress that it's hard to start a research project from a specific technique unless one has profound insight into its applications. It's more reasonable to actually start from a specific problem and go on to fix and detail our model step by step, meanwhile looking for possible techniques to solve it.
- It's worthwhile to continue this project by attempting a few problems we identified even though we got stuck for now. For some of the open problems we identified, it would help to make them more specific and eliminate some of our invalid conjectures with future research.


Kook Jin Ahn, Sudipto Guha, and Andrew McGregor.
Analyzing graph structure via linear measurements.
In *Proceedings of the Twenty-third Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '12, pages 459–467, Philadelphia, PA, USA, 2012. Society for Industrial and Applied Mathematics.


C-C Yao Andrew.
Some complexity questions related to distributed computing.
In *Proc. 11th STOC*, pages 209–213, 1979.


Amit Chakrabarti and Sagar Kale.
Strong fooling sets for multi-player communication with applications to deterministic estimation of stream statistics.
In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 41–50. IEEE, 2016.


Shahar Dobzinski, Noam Nisan, and Sigal Oren.
Economic efficiency requires interaction.

In *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing*, STOC '14, pages 233–242, New York, NY, USA, 2014. ACM.

Hossein Esfandiari, Mohammad T. Hajiaghayi, Vahid Liaghat, Morteza Monemizadeh, and Krzysztof Onak.
Streaming algorithms for estimating the matching size in planar graphs and beyond.
In *Proceedings of the Twenty-sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '15, pages 1217–1233, Philadelphia, PA, USA, 2015. Society for Industrial and Applied Mathematics.

Michael Kapralov, Sanjeev Khanna, and Madhu Sudan.
Approximating matching size from random streams.
In *Proceedings of the Twenty-fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '14, pages 734–751, Philadelphia, PA, USA, 2014. Society for Industrial and Applied Mathematics.

Eyal Kushilevitz and Noam Nisan.

Communication Complexity.
Cambridge University Press, New York, NY, USA, 1997.

Omri Weinstein and David P Woodruff.
The simultaneous communication of disjointness with applications to data streams.
In *International Colloquium on Automata, Languages, and Programming*, pages 1082–1093. Springer, 2015.

Penghui Yao.
Parity decision tree complexity and 4-party communication complexity of xor-functions are polynomially equivalent.
CoRR, abs/1506.02936, 2015.