Gadgetless Lifting Beats Round Elimination: Improved Lower Bounds for Pointer Chasing

Xinyu Mao **Guangxu Yang** Jiapeng Zhang

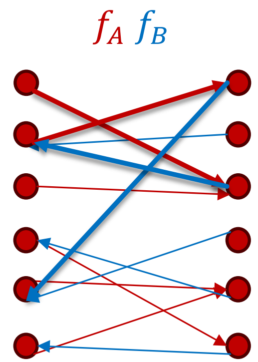
USC University of Southern California

The pointer chasing problem

The pointer chasing problem

Pointer chasing is a natural problem where many rounds of communication is useful.

- Alice holds $f_A \in [n]^n$, Bob hold $f_B \in [n]^n$.
- The k step pointer chasing function $PC_k: [n]^n \times [n]^n \to \{0,1\}$
  - $pt_0 := 1$
  - for odd r's, $pt_r := f_A(pt_{r-1})$
  - for even r's, $pt_r := f_B(pt_{r-1})$
  - $PC_k(f_A, f_B) := pt_k \bmod 2$.



There is a k-round protocol for PC_k with $O(k \log n)$ communication bits.

The pointer chasing problem

Theorem (Yehudayoff 2016). Any randomized $(k-1)$-round protocol for PC_k that is correct with probability 0.9 requires $\Omega\left(\frac{n}{k} - k\log n\right)$ bits of communication. Round Elimination

The (k log$n$) loss is associated with many round elimination- based analysis [NW91, Kla00, KNTSZ07, FKM+09, Yeh20].

In this paper, we avoid the k log n loss via the gadgetless lifting.

This work $\Omega\left(\frac{n}{k} + k\right)$ lower bound via a completely different proof

Pointer chasing has wide applications in circuit lower bounds [NW91, KPPY84] , distributed computation [NDSP11]

streaming algorithms [FKM+09, GO16, ACK19] property testing [CG18] differential privacy [JMR20] learning [CPP22]

transformer architecture [PNP24].

Gadgetless Lifting

Query to communication lifting theorems

- Let $g: \{0,1\}^q \times \{0,1\}^q \to \{0,1\}$ be a **gadget** function with low discrepancy
- Consider functions of the form $f \circ g^n$ for some outer function $f: \{0,1\}^n \to \{0,1\}$,

$$(f \circ g^n)\big((x_1, y_1), \dots, (x_n y_n)\big) \coloneqq f\big(g(x_1, y_1), \dots, g(x_n, y_n)\big).$$

$CC(f \circ g^n) = \Omega(Q(f) \cdot q)$, where $Q(f)$ denotes the query complexity of f.

Lifting theorems show that the query type protocols are optimal for $f \circ g^n$!

- Not all functions can be written as $f \circ g^n$. Pointer chasing is not a composed function
- Often need $q = \Omega(n)$: Proving lift theorems for constant gadget size q is very hard and has many implications [RM97, WYY17, CKLM19, GPW18, CFK+19, LMM+20]

A simple class of $k-1$ rounds protocols for pointer chasing

- Alice choose a subset $I \subseteq [n]$ of size $S := 10\frac{n}{k}$ uniformly at random, and then send $f_A(I)$ to Bob.
- Alice and Bob run the naïve (k rounds) protocol, but they can skip one round if the pointer falls into I.
- If the skip round never happens, Alice and Bob simply abort at the last round.
- The skip round event happen with high probability.

This is a $(k-1)$ –rounds randomized protocol for PC_k with $\widetilde{\Theta}\left(\frac{n}{k}+k\right)$ communication bits.

This protocol is not a query type protocol but Alice and Bob only send values of some coordinate to each other ! How to prove this protocol is optimal ?

Gadgetless lifting

- Identify a simple class of protocols $\mathcal{K}$.
- Prove lower bound for these simple protocols.
- Prove that every protocol can be "simulated" by a combination of simple protocols.

$$CC(f) := \min_{\Pi:\Pi \text{ computes } f} CC(\Pi) \approx \min_{\Pi \in \mathcal{K}} CC(\Pi) =: CC_{\mathcal{K}}(\Pi).$$

- For pointer chasing, $\mathcal{K}$ is the set of protocols where Alice and Bob only send values of some coordinate to each other.

Decomposition and Sampling Process

Density restoring partition

Def. For a random variable $\boldsymbol{X}$, its min-entropy is defined as $\mathbf{H}_\infty(\boldsymbol{X}) := \log \frac{1}{\max_x \Pr[\boldsymbol{X}=x]}$.

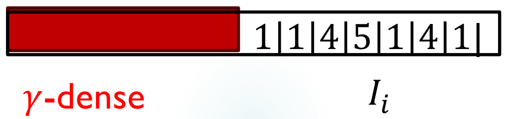
Def. We say a random variable $\boldsymbol{X}$ over $[n]^J$ is γ-dense if $\mathbf{H}_\infty\big(\boldsymbol{X}(I)\big) \geq \gamma \log n\ |I|$ for all $I \subseteq J$.

For a set X, $\boldsymbol{X}$:= uniform distribution over X.

Lemma 1 ([GPW17]). For any $X \subseteq [n]^J$, there is a partition $X = X^1 \cup \cdots \cup X^r$ and each X^i is associated with a set I_i with the following properties.

- X^i is fixed on I_i: there exists some $\alpha_i \in [n]^{I_i}$ such that $x(I_i) = \alpha_i$ for all $x \in X^i$.
- $\boldsymbol{X}^i(J \setminus I_i)$ is γ-dense.
- $\mathbf{D}_\infty\left(\boldsymbol{X}^i(J \setminus I_i)\right) \leq \mathbf{D}_\infty(X) - (1-\gamma)\log n\,|I_i| + \delta_i$ where $\delta_i = \log \frac{|X|}{|\cup_{j \geq i} X^j|}$.

▶ $\mathbf{D}_\infty(\boldsymbol{X}) := |J| \log n - \mathbf{H}_\infty(\boldsymbol{X})$ if $\boldsymbol{X}$ is supported on $[n]^J$.



Decomposition and sampling process $DS(\Pi)$

Input: A protocol tree Π under uniform input distribution
Output: A rectangle $R = X \times Y \subseteq [n]^n \times [n]^n$, $J_A, J_B \subseteq [n]$.
Initialization: $X \coloneqq Y \coloneqq [n]^n, J_A \coloneqq J_B \coloneqq [n]$, skip $\coloneqq$ $false$, $r \coloneqq 0, v \coloneqq root$.

1. Partition X into $X = X^0 \cup X^1$ according to node v in the protocol tree
2. Sample $\boldsymbol{b} \in \{0,1\}$ such that $\Pr[\boldsymbol{b} = b] = \frac{|X^b|}{|X|}$.
3. Update $X \coloneqq X^{\boldsymbol{b}}, v \coloneqq u_{\boldsymbol{b}}$.
4. If u_b is owned by Bob:
   - Further partition X into $X = X^0 \cup X^1$ where $X^b \coloneqq \{f_A \in X : f_A(z_{r-1}) \bmod 2 = b\}$.
   - Sample $\boldsymbol{b} \in \{0,1\}$ such that $\Pr[\boldsymbol{b} = b] = \frac{|X^b|}{|X|}$.
   - Update $X \coloneqq X^{\boldsymbol{b}}, r \coloneqq r + 1$.
5. Let $X = X^1 \cup \cdots \cup X^m$ be density restoring partition of X with associated $I_1, \ldots, I_m$.
6. Sample a random element $\boldsymbol{j} \in [m]$ such that $\Pr[\boldsymbol{j} = j] = \frac{|X^j|}{|X|}$ for $j \in [m]$.
7. Update $X \coloneqq X^j, J_A \coloneqq J_A \setminus I_j$.
8. If u_b is owned by Bob $z_{r-1} \notin J_B$, skip := true.

Suppose Alice owns node v and u_0, u_1 be the children of v in the protocol tree.

As a new round begins, we do an extra partition to fix the parity of pt_r.

1|1|4|5|1|4|1|

γ-dense

X_{I_j} is fixed;
X_{J_A} is dense.

Loop invariant

Input: A protocol Π
Output: A rectangle $R = X \times Y \subseteq [n]^n \times [n]^n$, $J_A, J_B \subseteq [n]$.
Initialization: $X := Y := [n]^n, J_A := J_B := [n]$, skip $:= false, r := 0, v := root$.

Lemma 2. *Set $\gamma := 1 - \frac{0.1}{\log n}$. Then in the running of $DS(\Pi)$, we have the following loop invariants:*

After each iteration,

- $X \times Y \subseteq \Pi_v$.
- $\boldsymbol{X}(J_A), \boldsymbol{Y}(J_B)$ *are γ-dense.*
- *There exists some $\alpha_A \in [n]^{\overline{J_A}}, \alpha_B \in [n]^{\overline{J_B}}$ such that $x(\overline{J_A}) = \alpha_A, y(\overline{J_B}) = \alpha_B$ for all $x \in X, y \in Y$.*
- *There exists some $z_r \in [n]$ such that $pt_r(f_A, f_B) = z_r$ for all $f_A \in X$, $f_B \in Y$.*

In the round elimination method, it should fix pt_r in each round,
but we only fix the parity of pt_r and use the density restoring partition helps to fix pt_r.
This is why we save the $k \log n$ factor

Relating accuracy and *average fixed size*

Lemma. If $DS(\Pi)$ outputs $(R = X \times Y, J_A, J_B)$ then

$$\Pr_{(f_A, f_B)}[\Pi(f_A, f_B) = PC_k(f_A, f_B)] \le \frac{2^{0.1}}{2} + \frac{2^{0.1}}{n} \cdot k \cdot \mathbf{E}[|\overline{J_A}| + |\overline{J_B}|]$$

$$\mathbf{E}[|\overline{J_A}| + |\overline{J_B}|] \ge \Omega(n/k)$$

Based on the loop invariant , the proof is similar to the analysis of simple protocols. We omitted the proof in this talk.

Average fixed size is bounded by communication: A density increment argument

$$\mathbf{D}_\infty(\boldsymbol{X}) \coloneqq |J| \log n - \mathbf{H}_\infty(\boldsymbol{X})$$

- In the running of $DS(\Pi)$, our density function is :

$$D_\infty(R) \coloneqq D_\infty\big(X(J_A)\big) + D_\infty\big(Y(J_B)\big).$$

- In the beginning, $D_\infty([n]^n \times [n]^n) = 0.$
- In expectation (over the choice of $\boldsymbol{b}$), each communication bit/new round (assume Alice speaks) increase $D_\infty(R)$ by at most 1:

$$\frac{|X^0|}{|X|}\log\frac{|X^0|}{|X|} + \frac{|X^1|}{|X|}\log\frac{|X^1|}{|X|} \leq 1.$$

- In expectation (over the choice of $\boldsymbol{j}$), $D_\infty(R)$ decreases by at least $0.1\ \mathbf{E}_j\big[|I_j|\big] - 1.$
  - $\mathbf{D}_\infty\left(\boldsymbol{X}^i(J \setminus I_i)\right) \leq \mathbf{D}_\infty(X) - 0.1\ |I_i| + \delta_i$ where $\delta_i = \log\frac{|X|}{|\cup_{j\geq i} X^j|}.$
  - $\mathbf{E}_j[\delta_j] = \sum_j p_j\ \delta_j = \sum_j p_j \log\frac{1}{\sum_{t\geq j} p_t} \leq \int_0^1 \frac{1}{1-x}\,dx \leq 1.$ where $p_j \coloneqq \frac{|X^j|}{|X|}$
- $D_\infty(R) \geq 0 \rightarrow \mathbf{E}[|\overline{J_A}| + |\overline{J_B}|] = \mathbf{E}[|I_1| + |I_2| + \cdots +] \leq O(C).$

Proof outline

- **The decomposition and sampling process**:

  Use **density restoring partition** to decompose the behavior of Π into the combination of simple protocols (i.e., fixing some coordinates).

- Relating accuracy and **average fixed size** . $\mathbf{E}[|I|] \geq \Omega(n/k)$.

  By the analysis of simple protocols.

- **Average fixed size** is bounded by communication . $\mathrm{CC} \geq \mathbf{E}[|I|]$.

  By a density increment argument.

Thank you for listening ☺