

# A Decision Tree Based Analysis of communication complexity for FORK Game

Guangxu Yang

February 5, 2021

## 1 FORK game

**Definition 1.1** (FORK game). Define a relation  $\mathbf{FORK}_{w,l} \subseteq [w]^l \times [w]^l \times [l] \cup \{0\}$ , we implicitly pad  $x \in [w]^l$  and  $y \in [w]^l$  with additional 0 and  $l+1$  positions such that  $x_0 = y_0 = 1$ ,  $x_{l+1} \neq y_{l+1}$ . This ensures that  $\forall x, y \in [w]^l, \exists i \in \{0, \dots, l\}$  such that  $(x, y, i) \in \mathbf{FORK}_{w,l}$  where  $(x, y, i) \in \mathbf{FORK}_{w,l} \iff (x_i = y_i \text{ and } x_{i+1} \neq y_{i+1})$ .

**Definition 1.2** (Switch search problem). Define the “Switch” relation:

$$\mathbf{Switch}_l = \{(z, i) \in \{0, 1\}^l \times \{0, 1, \dots, l\} \mid z_i = 1, z_{i+1} = 0\}$$

where we use  $z_0 = 1$  and  $z_{l+1} = 0$ , i.e., we are given  $l$  bits and wish to find a “switching point”, a position  $i$  where a 1-bit flips into a 0-bit. If  $z = 0^l$  we must output  $i = 0$  and if  $z = 1^l$  we must output  $i = l$ .

We note that the FORK game  $\mathbf{FORK}_{w,l} = \mathbf{Switch}_l \circ \mathbf{Eq}_w$  where  $z_i = \mathbf{Eq}(x_i, y_i)$ .

## 2 Communication complexity of FORK game

### 2.1 Query complexity of switch

Let  $z \in \{0, 1, *\}^l$  be string. we call  $i$  active coordinate if  $z_i = *$ .

Let  $I \in [l]$  be the set of active coordinates. and  $I_0$  be the first half coordinates of  $I$  and  $I_1$  be the last half coordinates of  $I$ .

**Claim 2.1.**  $\mathbf{P}^{\text{dt}}(\mathbf{Switch}) = \log l$ .

*Proof.* Let  $T$  be a decision tree for  $\mathbf{Switch}$ , we show there is a path of length  $\Omega(\log l)$  in  $T$ . Starting from the root of  $T$ , we find a path step by step. In each step  $t$ , we reach a node  $v_t$  in  $T$  of depth  $t$ . At the beginning, we set  $z = *^l$  and  $|I| = l$

**Find long path in decision tree:**

1. If the decision query a coordinate  $i \in I_0$ , we fix  $z_i = 1$  for all coordinates in  $I_0$  and  $I = I_1$ .
2. If the decision query a coordinate  $i \in I_1$ , we fix  $z_i = 0$  for all coordinates in  $I_1$  and  $I = I_0$ .
3. Let  $v_{t+1} = v_{t, z_i}$ .

In each query round, we fix at most  $|I|/2$  coordinates. At the end,  $|I| \leq 1$ , the query rounds at least  $\log l$ .  $\square$

Let  $S$  be the set of long query paths we could find in above proof. In next section, we will show that we could find a long communication path which is consistent with a long query path in  $S$ .

## 2.2 Communication complexity of FORK game

Let  $z \in \{0, 1, *\}^n$  be string. we call  $i$  unfixed coordinate if  $z_i = *$ . Let  $I$  be the set of unfixed coordinates and  $I_0$  be the first half coordinates of  $I$  and  $I_1$  be the last half coordinates of  $I$ .

We get  $z^1$  by set  $z_i^1 = z_i$  if  $z_i \in \{0, 1\}$  and  $z_i^1 = 1$  if  $z_i = *$ .

We get  $z^0$  by set  $z_i^0 = z_i$  if  $z_i \in \{0, 1\}$ ,  $z_i^0 = 1$  if  $i \in I_0$  and  $z_i = *$  and  $z_i^0 = 0$  if  $i \in I_1$  and  $z_i = *$ .

**Definition 2.2** (Potential function). *Let  $X \times Y \subseteq [w]^l \times [w]^l$  be a rectangle and  $z$  be a string. The potential function is defined by*

$$\alpha(X, Y, z) = \frac{|\{(x_I, y_I) : (x, y) \in X \times Y, \mathbf{Eq}_w^I(x, y) = z^1\}|}{w^I}$$

**Claim 2.3.** *Let  $X \times Y$  be a rectangle and  $I \subseteq [n]$  be the unfixed coordinates, then for any partition of  $Y = Y_0 \cup Y_1$ , there is a  $Y_b$  such that  $\alpha(X, Y_b) \geq \alpha(X, Y)/2$ .*

**Claim 2.4.** *Let  $X \times Y$  be a rectangle and  $I \subseteq [n]$  be the unfixed coordinates, then for any partition of  $X = X_0 \cup X_1$ , there is a  $X_b$  such that  $\alpha(X_b, Y) \geq \alpha(X, Y)/2$ .*

**Lemma 2.5** (Disperser property). *Let  $X \times Y$  be a rectangle and  $I$  be the set of unfixed coordinates, If  $\alpha(X, Y, z) \geq \frac{16}{w}$ , then*

$$\{\mathbf{Eq}_w^I(x, y) : x \in X, y \in Y\} = \{0, 1\}^I$$

*Proof.* We do an induction on  $|I| = l$ . The base case of  $l = 1$  follows from equality function. For the induction step, assume the lemma holds for all integers at most  $l - 1$ . We need prove that any string  $z \in \{\mathbf{Eq}_w^l(x, y) : x \in X, y \in Y\}$  when  $\alpha(X, Y, z) \geq \frac{16}{w}$ .

let  $I_0 \subseteq I$  be the set of coordinates that  $z_i = 0$  and  $I_1 \subseteq I$  be the set of coordinates that  $z_i = 1$ . If  $|I_0| \leq l - 1$ , by average argument, there is a sub-rectangle  $X' \times Y'$  and  $q^{I_1} \in [w]^{|I_1|}$  with  $X_{I_1} = Y_{I_1} = q^{I_1}$  such that  $I_0$  be the set of unfixed coordinates,  $\alpha(X', Y', z) \geq \frac{16}{w}$ . Since the lemma holds for all integers at most  $l - 1$ ,  $z \in \{\mathbf{Eq}_w^l(x, y) : x \in X, y \in Y\}$ .

If  $|I_0| = l$ , we prove  $0^l \in \{\mathbf{Eq}_w^l(x, y) : x \in X, y \in Y\}$  by probabilistic method. Let  $A = A_1 \times \dots \times A_l$  where each  $A_i$  is subset of  $[w]$  with size  $w/2$ . if we ensure that exist a set  $A$  such that there is a  $x \in X \cap A$  and  $y \in Y \cap A^c$  for  $\mathbf{Eq}_w^l(x, y) = 0^l$  then the lemma immediately holds.

So it remains to show that there exist such set  $A$ . We choose  $A_i$ s as follows: first choose at random  $w/2$  strings  $v^1, \dots, v^{w/2}$  each of length  $l$ . Then we define  $A_i$  to include the  $i$ -th letter in each of these  $w/2$  strings and extend it into a set of size  $w/2$  randomly (Note that the resulting sets  $A_1, \dots, A_l$  are indeed random and independent.). For any  $x \in A$ ,  $\Pr[x \in X] \geq \alpha$  and any  $y \in A^c$ ,  $\Pr[y \in Y] \geq \alpha$ . We call  $A$   $X$ -good if  $X \cap A \neq \emptyset$  and  $Y$ -good if  $Y \cap A^c \neq \emptyset$ . We note that if there is a  $A$  both  $X$ -good and  $Y$ -good the lemma hold. We could prove that:

$$\Pr[A \text{ is not } X\text{-good}] \leq (1 - \alpha)^{w/2} \leq e^{-\alpha w/2}$$

and

$$\Pr[A \text{ is not } Y\text{-good}] \leq (1 - \alpha)^{w/2} \leq e^{-\alpha w/2}$$

Therefore, the probability that either  $A$  is not  $X$ -good or  $A$  is not  $Y$ -good is at most  $2e^{-\alpha w/2} < 1$ . There is a  $A$  both  $X$ -good and  $Y$ -good, the lemma holds.  $\square$

**Lemma 2.6** (Projection lemma). *[GS95] Let  $X \times Y$  be a rectangle and  $I$  be the set of unfixed coordinates, If  $\alpha(X, Y, z) \geq 4w^{-2/3}$  then there is a  $b \in \{0, 1\}$ , we could fix  $z_i = b$  for all coordinates in  $I_{1-b}$  to get a sub-rectangle  $X' \times Y'$  and  $I = I_b$  such that  $\alpha(X', Y', z) \geq w^{-1/3}$ .*

**Theorem 2.7.**  $\mathbf{P}^{\text{cc}}(\mathbf{FORK}_{w,l}) = \Omega(\log l \log w)$ .

*Proof.* Let  $T$  be a protocol tree for  $(S \circ \text{Eq}_w^l)$ , we show there is a path of length  $\Omega(\log l \log w)$  in  $T$ . Starting from the root of  $T$ , we find a path step by step. In each step  $t$ , we reach a node  $v_t$  in  $T$  of depth  $t$ . At the beginning, We set  $z = *^n$ . Let  $(X_t, Y_t)$  be a rectangle which consistent with current node  $v_t$  and  $z$ . At the beginning,  $\alpha(X, Y, z) = 1$ .

### Find long path in protocol tree:

1. In each communication round, If it is Bob's turn in the protocol, then Bob sends a bit  $b \in \{0, 1\}$  that maximizes the potential function  $\alpha(X_t, Y_t)$ . Set  $\mathcal{X}_t = \mathcal{X}_{t-1}, \mathcal{Y}_t = \mathcal{Y}_{t-1, b}$ . If it is Alice's turn in the protocol, then Bob sends a bit  $b \in \{0, 1\}$  that maximizes the potential function  $\alpha(X_t, Y_t)$ . Set  $\mathcal{X}_t = \mathcal{X}_{t-1, b}, \mathcal{Y}_t = \mathcal{Y}_{t-1}$ .
2. If  $8w^{-2/3} \geq \alpha(X, Y, z) \geq 4w^{-2/3}$ , we do projection. By projection lemma, there is a  $b$  such that we could fix  $z_i = b$  for all coordinates in  $I_{1-b}$  to get a sub-rectangle  $(X_{t+1}, Y_{t+1})$  and  $\alpha(X_{t+1}, Y_{t+1}, z) \geq w^{-1/3}$ . Set  $I = I_b$ .

By disperser property, we know that we would do projection at least  $\log l$  times and by Claim 2.3 and Claim 2.4, the communication rounds between two projection is at least  $\log(\frac{1}{4}w^{1/3})$ . As a result, the communication complexity of FORK game is  $\Omega(\log l \log w)$ .  $\square$

## 3 Open problems

**Problem 3.1** (Direct sum of FORK game by Or Meir). *The  $m$ -fold direct sum of FORK relation on  $l$  bits, denoted  $U_l^{\otimes m}$  is the communication problem in which Alice and Bob get matrices  $X, Y \in [w]^{m \times l}$  that each row is a instance of FORK relation. They are required to output a tuple  $(j_1, \dots, j_m) \in [l]^m$  such that for every row  $i \in [m]$  holds that  $\mathbf{Eq}(x_{i, j_i}, y_{i, j_i}) \neq \mathbf{Eq}_w(x_{i, j_i+1}, y_{i, j_i+1})$ .*

**Remark:** We note that when  $w = 2$ ,  $\mathbf{P}^{\text{cc}}(U_l^{\otimes m}) = \Omega(m \log l)$ . [Mei18] prove the Direct Sum of Universal Relations by Raz-McKenzie simulation and Rank argument.

**Definition 3.2.** *Given a deterministic decision-tree  $T$  over  $\{0, 1\}^l$ , the 0-depth of  $T$  is the maximum number of queries which are answered 0, in any root-to-leaf path of  $T$ . The 0-query complexity of  $f$ , denoted  $\mathbf{P}_0^{\text{dt}}(f)$ , to be the smallest 0-depth of  $T$ , taken over deterministic decision-trees  $T$  which solve the search problem associated with  $f$ .*

**Theorem 3.3** (by [LM19]). *For any Boolean relation  $f : \{0, 1\}^l \times \mathcal{C}$ , whenever  $n \geq \cdot \log p$*

$$\mathbf{P}^{\text{cc}}(f \circ \text{Eq}_n^q) = \Omega(n \cdot \mathbf{P}_0^{\text{dt}}(f))$$

**Problem 3.4.** *Improving the size of the gadget for the simulation theorems of [LM19]*

**Lemma 3.5** (Disperser property of index gadget). *[LMZ20] Let  $X \times Y$  be a rectangle and  $q \geq \log n$ . Assume that  $X$  is  $\log n$ -spread and  $Y \geq \frac{2^{qn}}{2^q}$  then*

$$\{\mathbf{Ind}(x, y) : x \in X, y \in Y\} = \{0, 1\}^n$$

**Problem 3.6.** *If we could change the condition of projection in [LMZ20] to prove a  $P$  lifting theorem with  $\log n$  gadget size ?*

**Problem 3.7.** *In [dRMN<sup>+</sup>20], They prove communication lower bound of  $\mathbf{Search}(\text{Peb}_G) \circ \text{Eq}^n$ . Could we prove the same lower bound by disperser property of equality gadget and path simulation ?*

## References

- [dRMN<sup>+</sup>20] Susanna F. de Rezende, Or Meir, Jakob Nordström, Toniann Pitassi, Robert Robere, and Marc Vinyals. Lifting with simple gadgets and applications to circuit and proof complexity, 2020.
- [GS95] Michelangelo Grigni and Michael Sipser. Monotone separation of logarithmic space from logarithmic depth. *Journal of Computer and System Sciences*, 50(3):433–437, 1995.
- [LM19] Bruno Loff and Sagnik Mukhopadhyay. Lifting theorems for equality. In *36th International Symposium on Theoretical Aspects of Computer Science (STACS 2019)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2019.
- [LMZ20] Shachar Lovett, Raghu Meka, and Jiapeng Zhang. Improved lifting theorems via robust sunflowers. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 27, page 48, 2020.
- [Mei18] Or Meir. The direct sum of universal relations. *Information Processing Letters*, 136:105–111, 2018.