# A survey on communication complexity and its application to lower bound

Binghui Peng

January 25, 2018

## 1 Introduction

Communication complexity theory, initiated by Andrew Yao[3] thirty-five years ago, is a central branch of theoretical computer science. The theory studies the minimum amount of communication, measured in bits, required in order to compute functions whose arguments are distributed among several parties. Besides its own interests, it also develop useful tool and technique to prove lower bound. Furthermore, communication complexity provide fundamental complexity result which can in turn be used to prove other lower bound via reduction. In this survey, we mainly focus on the application of communication complexity to prove lower bound in other area.

This survey organize as follow: in section 2, we introduce basic definition of communication complexity and summarize some fundamental results of communication complexity. In section 3, we introduce a recent lower bound progress with help of communication complexity, we details the definition of extension complexity for polytope and Yannakakiss Lemma, a celebrated results build connection between extented formulation and communication complexity. In section 4, we fulfill the proof of [?paper], proving a $\Omega(2^{\sqrt{n}})$ lower bound for extension complexity for TSP polytope. In section 5, we summary the survey and provide more application of communication complexity and some related work..

## 2 Communication Complexity

### 2.1 Basic Definition of Communication Complexity

#### 2.1.1 Deterministic Model

There are various model of communication, the original one steps from Yao's paper is two-party deterministic model. Consider a function $f : X \times Y \to \{0, 1\}$, where $X$ and $Y$ are finite sets. The model features two cooperating parties, traditionally called Alice and Bob. Alice receives an input $x \in X$, Bob receives an input $y \in Y$, and their objective is to compute $f(x, y)$. To this end, Alice and Bob communicate back and forth according to an agreed-upon protocol. The cost of a given communication protocol is the maximum number of bits exchanged on any input pair $(x, y)$. The deterministic communication complexity of f, denoted $D(f)$ is the least cost of a communication protocol for $f$.

In order to derive lower bound for deterministic communication model, couples of method has been develop.

**Fooling set method** A straightforward technique for proving communication lower bounds is the fooling set method, which works by identifying a large set of inputs no two of which can occupy the same f-monochromatic rectangle. Formally, a fooling set for $f : X \times Y \to \{0, 1\}$ is any subset $S \subseteq X \times Y$ with the following two properties: ($i$) f is constant on $S$ and ($ii$) if $(x, y)$ and $(x^\star, y^\star)$ are two distinct elements of $S$, then f is not constant on $\{(x, y), (x, y^\star), (x^\star, y), (x^\star, y^\star)\}$. A direct observation is an f-monochromatic rectangle can contain at most one element of $S$. Therefore, any partition (or even cover) of $X \times Y$ by f-monochromatic rectangles must feature a rectangle for each point in the fooling set $S$. Formally, we have

**Theorem** (Fooling set method). Let $f : X \times Y \to \{0, 1\}$ be a given communication problem. If $S$ is a fooling set for $f$, then

$$D(f) \geq log_2(|S|)$$

### 2.1.2 Nondeterministic Model

In order to proceeds to funcy application of communication complexity, we introduce nondeterministic model here. In nondeterministic protocoal, Alice and Bob are given a proof or advice string by a prover, which can depend on both of their inputs; the communication cost is the worst-case length of the proof; and a protocol is said to compute an output $z \in \{0, 1\}$ of a function$f$ if $f(x, y) = z$ if and only if there exists proof such that both Alice and Bob accept.

A remark here is that some techniques like fooling set method of rectangle size method(we will introduce later) also applies in nondeterministic protocol.

### 2.1.3 Remark

We do not introduce either randomize protocol multiparty protocol or useful technique like Yao's minmax Lemma, due to page limits, but they are still important topic and we refer [2] to more informative introduction.

## 2.2 A Central Problem — DISJOINTNESS

### 2.2.1 Summarize of previous result

In this subsection, we introduce a fundamental problem in communication model, the DISJOINTNESS[2]. It plays the central role in communication complexity, just like SAT in NPC problem. Tim Roughgarden have said "If you only remember one problem that is hard for communication protocols, it should be the disjointness problem.". The formal definition is

$$DISJ_n(A, B) = \begin{cases} 1 & A \cap B = \emptyset \\ 0 & otherwise \end{cases}$$

The following tabular summarize important results on DISJOINTNESS.

| | |
|---|---|
| Deteministic | $\Omega(n)$ |
| Nondeterministic | $\Omega(n)$ |
| Randomize | $\Omega(n)$[5] |
| Multiparty(deterministic/randomize) | $\Omega(n/k)$[6] |

### 2.2.2 Uniqueness Disjointness

The Unique Disjointness is a promise version of disjointness with
(1) inputs (X, Y) of Disjointness with

$$\|X \cap Y\| = 0$$

(0) inputs (X, Y) of Disjointness with

$$|X \cap Y| = 1$$

We have the following theorem

**Theorem** (Fiorini et al. 2015; Kaibel and Weltge 2015)[1] The nondeterministic communication complexity of Unique-Disjointness is $\Omega(n)$

**Proof:** We briefly sketch the idea of proof here. We proof with rectangle size method, the 1 input is at least $3^n$. We need the following technical lemma.

**Lemma** Every 1-rectangle of Unique Disjointness contains at most $2^n$ 1-inputs.

The lemma can be proved by induction: for 1-rectangle $R = A \times B$ and $x \in A, y \in B$ with 0 in last $n-k$ coordinate, the number of 1-input is at most $2^k$. We partition all 1 input into $(x1, y0), (x0, y1), (x0, y0)$ and let $S_1, S_2$ contains first two type respectively. We partition the third type to $S_1$ or $S_2$ with the requirement that $(x, y)$ do not appear twice in same set(this can be done easily). As a consequence, in $S_1$ and $S_2$, we can ignore the last coodinate and by induction each sub-rectangle has 1 input at most $2^{k-1}$ □

# 3 Extended Formulations

In this section, we introduce one application of communication complexity. instead of deeping into detail, we introduce basic definition and crucial lemma this section. Basically, extended formulation(extended complexity) capture the number of inequality to describes a polytope(linear projection are allowed here) The rigious definition is given below:
An extended formulation (EF) of a polytope $P \subseteq R^d$ is a linear system

$$E^{=}x + F^{=}y = g^{=}, E^{\leq}x + F^{\leq}y \leq g^{\leq}$$

in variables $(x, y) \in R^{d+r}$, where $E^{=}, F^{=}, E^{\leq}, F^{\leq}$ are real matrices with $d, k, d, k$ columns respectively, and $g^{=}, g^{\leq}$ are column vectors, such that $x \in P$ if and only if there exists $y$ such that holds.

**Example** Let $x_\pi = (\pi(1), \cdots, \pi(n))$ with $\pi \in S$ is a permutation. The permutahedron is the convex hull of all $n!$ such vectors. The permutahedron is known to have $2^{n/2} - 2$ facets (see e.g. Goemans (2014)), so a polynomial-sized linear description would seem out of reach. However, by adding $n^2$ auxiliary variables $y_{ij}$ where $y_{ij}$ intends to indicates whether $\pi(i) = j$ and satiesfies

$$\sum_{j=1}^{n} y_i j \leq 1$$

$$\sum_{i=1}^{n} y_i j \leq 1$$

$$y_{ij} \geq 0$$

$$x_i = \sum_{j=1}^{n} j y_{ij}$$

If we take the newly formed polytope $R^{n+n^2}$ and project onto the x-coordinates, we get exactly the permutahedron. This is an example shows that extended formulation cound be much smaller than we 'directly' describe the polytope.

In order to apply communication complexity to extended formulation, we design the following contrived communication problem.

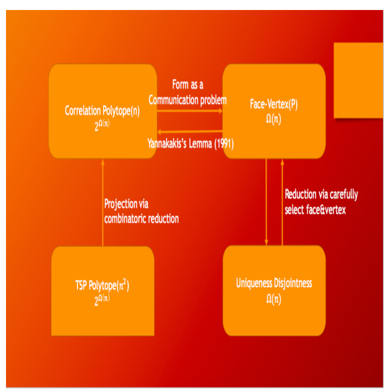
**Face-Vertex(P)**For a polytope $P$, in the corresponding Face-Vertex(P) problem, Alice gets a face $f$ of $P$ (in the form of a supporting hyperplane a, b) and Bob gets a vertex $v$ of P . The function $FV(f, v)$ is defined as 1 if $v$ does not belong to $f$,and 0 if $v \in f$. Equivalently,$FV(f, v) = 1$if and only if $a^T v < b$, where $a, b$ is a supporting hyperplane that induces f. A key result is the following:

**Lemma** (Yannakakiss Lemma (1991)) If the polytope $P$ admits an extended formulation $Q$ with $r$ inequalities, then the nondeterministic communication complexity of Face-Vertex(P) is at most $log_2(r)$.

Yannakakiss Lemma is quite intuitive. Due to page limit, we do not present rigous proof here, but the key idea is that a prover can name a facet $f$ of Q such that:
(i) there exists $y_v$ such that $(v, y_v) \notin f$ and
(ii) for every $(x, y) \in Q$ with $x \in f$,$(x, y) \in f$

# 4 $2^{\Omega(\sqrt{n})}$ Lower Bound for TSP polytope

With the neccesary material presented so far, we can prove the $2^{\Omega(\sqrt{n})}$ Lower Bound for TSP polytope. The following picture is a view of the whole proof.




## 4.1 Correlation Polytope

Given a 0-1 $n$ bit vector $x$, we consider the corresponding (symmetric and rank-1) outer product $xx^T$. For a positive integer $n$, we define COR as the convex hull of all $2^n$ such vectors $xx^T$ (ranging over $x \in \{0, 1\}^n$). This is a polytope in $R^{n^2}$ , and its vertices are precisely the points $xx^T$ with $x \in \{0, 1\}^n$. We have the following theorem

**Theorem** The nondeterministic communication complexity of Face-Vertex(COR) is $\Omega(n)$.
Proof: The proof was done via reduction to uniqueness disjointness problem. First, we have the following technical lemma.

**Lemma** (Fiorini et al. 2015) For every subset $f_S$ $\{1, 2 \cdots n\}$, there is a face $f_S$ of COR such that for every $R \subseteq \{1, 2, \cdots, n\}$ with characteristic vector $x_R$ and corresponding vector $v_R = x_R x_R^T$ of COR,

$$v_R \in f \text{ if and only if } |S \cap U| = 1$$

With this lemma, we can clearly reduce VERTEX-FACE(COR) to uniqueness disjointness, which has lower bound of $\Omega(n)$, and by Yannakakis's lemma, we can prove the $\Omega 2^n$ extension complexity for correlation polytope.The above lemma itself is not hard to proof and the idea is to find out the supporting plane. □

## 4.2 TSP Polytope

The last step is to prove $EF(TSP(n^2)) \geq EF(COR(n))$. We prove it by showing tha polytope $COR(n)$ is the linear projection of a face of $TSP(O(n^2))$. Formally,

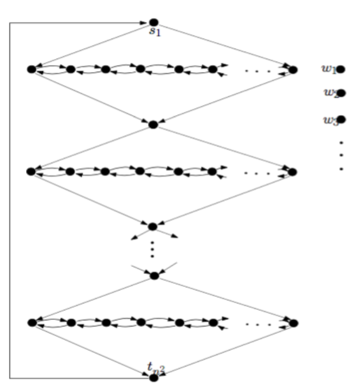
**Lemma** For each $n$, there exists a positive integer $q = O(n^2)$ such that $TSP(q)$ contains a face that is an extension of $COR(n)$.

The proof preceeds by the following two step(notice the reduction is purely combinatoric and analog to reduction done when we prove that hamitonian cycle is NPC)

(i) Define a 3SAT formula $\phi(n)$ with $n^2$ variables such that the satisfying assignments of $_n$ bijectively correspond to the matrices $bb^T$, where $b \in \{0, 1\}^n$
(ii) Construct a directed graph $D_n$ with $O(n^2)$ vertices such that each directed tour of $D_n$ defines a satisfying assignment of $\phi_n$, and conversely each satisfying assignment of $\phi_n$ has at least one corresponding directed tour in $D_n$. Modify this to undirect case and maintain an bijective relation.
Step (i): Define

$$\phi_n := \wedge_{i,j \in [n], i \neq j} (C_{ii} C_{jj} \vee \overline{C_{ij}}) \wedge (C_{ii} \vee \overline{C_{jj}} \vee \overline{C_{ij}}) \wedge (\overline{C_{ii}} \vee C_{jj} \vee \overline{C_{ij}}) \wedge (\overline{C_{ii}} \vee \overline{C_{jj}} \vee \overline{C_{ij}})$$

Just notice that each gadget model the computation $C_{ij} = C_{ii} \wedge Cjj$
Step (ii): Step 2 is completely analog to traditional reduction from 3SAT to Hamitonian Path.(See the following picture to get reminded) Get thing together, we have




**Theorem** The extension complexity of the TSP polytope TSP(n) is $2^{\Omega(n^{1/2})}$.[1]

# 5 Summary

Communication Complexity is really a powerful tool to prove lower bound. It has wide application in proving lower bound for data structure, streamming algorithm, property test and algorithmic game theory etc. For more interesting application, I refer to [4].

# Reference


[1] Fiorini, Samuel, et al. “Exponential lower bounds for polytopes in combinatorial optimization”
[2] Sherstov, Alexander A. “Communication complexity theory: Thirty-five years of set disjointness.”
[3] Yao, Andrew Chi-Chih. “Some complexity questions related to distributive computing”
[4] Tim Roughgardens lecture note: `http://theory.stanford.edu/~tim/w15/w15.html`
[5] Razborov, A.A. “On the distributional complexity of disjointness
[6] Nisan, N. “ The communication complexity of approximate set packing”