Multiparty Communication Complexity for Set Disjointness: A Lower Bound

Sammy Luo, Brian Shimanuki

May 6, 2016

Abstract

The standard model of communication complexity involves two parties each starting with partial information on the inputs to a problem and exchanging a limited amount of information to solve the problem jointly. In this project we investigate an extension of this model to more than two parties. In particular, we focus on an extension of the set disjointness problem to this k-party setting, where the parties wish to compute whether the intersection of the k input sets is nonempty. The main result is the derivation of a lower bound of $\Omega\left(\left(\frac{n}{4^k}\right)^{1/4}\right)$ on the randomized communication complexity for $\text{DISJ}_{n,k}$. This is taken from Sherstov (STOC '12).

We will begin by introducing the number-on-the-forehead model for multiparty communication, where each party has access to all inputs but their own. We discuss its importance and relevance to communication complexity theory in general. We introduce the k-party version of the set disjointness problem and extend a simple deterministic protocol from Grolmusz (1994) which yields an upper bound of $O(\log^2 n + k^2 n/2^k)$ on the problem's complexity. We then proceed to describe a series of attempts at pushing the lower bound closer to the upper bound, illustrating the nuances of some techniques used for working with this communication model, culminating in the stated main result.

Finally we discuss some direct product results for the problem, and compare the difficulty of bounding the complexity of set disjointness and its cousin, the generalized inner product.

1 Introduction

1.1 Multiparty Communication

In general, communication complexity theory addresses the problem of computing a function when no party has access to the entire input. In the typical model, two parties, Alice and Bob, have access to inputs $x \in X$ and $y \in Y$ respectively, and wish to evaluate a function $f : X \times Y \to \{0, 1\}$ on input (x, y) using as few bits of communication as possible. The smallest number of bits that needs to be communicated in a protocol evaluating f is called the *communication complexity* of f. In the case of *randomized communication complexity*, we allow the parties to utilize randomness and only require them to succeed with some probability $1 - \epsilon$. Results in communication complexity theory have many connections with other areas of complexity theory, such as circuit complexity and proof complexity.

It is natural to consider generalizations of the 2-party model to a setup with $k \geq 3$ parties. One useful model is the so-called *number-on-the-forehead* model, introduced by Chandra, Furst, and Lipton [4]. In this model, k parties want to compute $f(x_1, \dots, x_k)$ for some function $f : X_1 \times X_2 \times \dots \times X_k \to -1, 1$, where the ith party has access to all inputs except x_i. (This can be thought of as the ith party has x_i written on their forehead, and can see the inputs on everyone else's foreheads.)

Many questions that have been thoroughly analyzed for the 2-party case remain open in the general k-party setting, where lower bounds on communication complexity are much more difficult to prove. The difficulty in proving lower bounds arises from the overlap in the inputs known to different parties. Nevertheless, many of the methods for studying problems in the 2-party setting have natural generalizations for tackling the corresponding multiparty versions of the problems.

1.2 Set Disjointness

One question that remains open in the k-party setting is the communication complexity of the set disjointness problem. The problem is as follows: Given input sets $S_1, \dots, S_k \subseteq \{1, 2, \dots, n\}$, determine whether they have empty intersection, i.e. whether $S_1 \cap S_2 \cap \dots \cap S_k = \emptyset$. Represented as a formula, we have

$$\mathrm{DISJ}_{n,k}(x_1, x_2, \dots, x_k) = \bigwedge_{j=1}^{n} \bigvee_{i=1}^{k} \overline{x_{ij}} = \neg \bigvee_{j=1}^{n} \bigwedge_{i=1}^{k} x_{ij},$$

where x_i is the vector whose jth component is 1 if $j \in S_i$ and 0 otherwise.

Another formulation of DISJ comes from a composition of functions. Namely, we can define $G^g_{n,k} : \{0,1\}^{n \times k} \to \{-1, +1\}$ by $G^g_{n,k}(x_1, \dots, x_k) = g(\bigwedge_i x_{i1}, \dots, \bigwedge_i x_{in})$. Note that $\mathrm{DISJ}_{n,k}$ is just $G^{\mathrm{NOR}}_{n,k}$. Another function of interest is the generalized inner product $\mathrm{GIP}_{n,k}(x_1, \dots, x_k) = \bigoplus_{j=1}^n \bigwedge_{i=1}^k x_{ij}$, which is $G^{\mathrm{PARITY}}_{n,k}$. We will briefly compare these in Section 8.

The unique set disjointness problem, $\mathrm{UDISJ}_{n,k}$, is the promise version of $\mathrm{DISJ}_{n,k}$: we are guaranteed that the intersection of the input sets has size either 0 or 1. The promise version of a problem can be no harder to solve than the original problem, so giving lower bounds on the communication complexity of UDISJ will give corresponding lower bounds on that of DISJ. Note that the function $\mathrm{UDISJ}_{n,k}$ is a *partial function*, a function f whose domain $\operatorname{dom} f$ is a subset of $X_1 \times \dots \times X_k$. Where the distinction is relevant, a function whose domain is the whole set will be called a *total* function.

The 2-party case of the set disjointness problem is well studied. A tight lower bound of $n + 1$ is known for the deterministic communication complexity, and a tight lower bound of $\Omega(n)$ for the randomized communication complexity was shown by Kalyanasundaram and Schnitger [8].

Progress on the general k-party case has been significantly more difficult. In 1994, Grolmusz [7] proved an upper bound of $O(\log^2 n + k^2 n/2^k)$ for the deterministic complexity that remains the best

Bound	**Reference**	**Year**
$O\left(\log^2 n + \frac{k^2 n}{2^k}\right)$	Grolmusz [7]	1994
$\Omega\left(\frac{\log n}{k}\right)$	Tesson [16]	2003
	Beame et al. [3]	2006
$\Omega\left(\left(\frac{n}{2^{2^k} k}\right)^{1/(k+1)}\right)$	Lee and Shraibman [10]	2007
	Chattopadhyay and Ada [5]	2008
$\Omega\left(\left(\frac{2^{\Omega(\sqrt{k \log n})}}{2^{k^2}}\right)^{1/(k+1)}\right)$	Beame and Huynh-Ngoc [2]	2009
$\Omega\left(\left(\frac{n}{4^k}\right)^{1/4}\right)$	Sherstov [13]	2012
$\Omega\left(\frac{\sqrt{n}}{2^k k}\right)$	Sherstov [15]	2014

Table 1: Randomized communication complexity for k-party set disjointness [15]

known. Progress on improving the corresponding lower bounds has been made incrementally and will be described later in this paper. The state of the art is Sherstov's lower bound of $\Omega(\sqrt{n}/2^k k)$ on the randomized complexity [15], proven in 2014. This paper's focus is a treatment of the derivation of an earlier, slightly weaker result by Sherstov, whose proof is more instructive for showcasing the kinds of techniques used in tackling this class of problems. The result, proven by Sherstov in 2012 [13], is stated below:

Theorem 1.1. *Set disjointness has randomized communication complexity*

$$R_{1/3}(\mathrm{DISJ}_{n,k}) = \Omega\left(\left(\frac{n}{4^k}\right)^{1/4}\right).$$

2 An Upper Bound

Theorem 2.1. $D(\mathrm{DISJ}_{n,k}) = O\left(\frac{k^2 n}{2^k}\right)$.

We describe a deterministic protocol for set disjointess using at most $(k^2+k-1)\left\lceil \frac{n}{2^{k-1}-1} \right\rceil$ bits of communication based off of work by Grolmusz [7], proving Theorem 2.1. In fact, we show a protocol for a more general problem, the problem of determining the size of the intersection. This trivially solves the disjointness problem as the sets are disjoint iff the size of the intersection is 0.

Consider the input $\{0,1\}^{n\times k}$ as a binary matrix where the ith player see all but the ith column. The goal is to determine the number of rows of all 1s.

2.1 The Protocol

The protocol is divided into two parts. The first part is for each block of at most $2^{k-1}-1$ rows, to determine a row which does not exist in the block. The second part is to use the known nonexistent row to determine the number of rows of all 1s within the block. The players can then sum the number of all 1 rows over all blocks.

We give a protocol for this problem on $\{0,1\}^{n\times k}$ for $n \leq 2^{k-1}-1$.

The first player knows $k-1$ columns. Then by the Pigeonhole Principle there is sequence $\alpha' \in \{0,1\}^{k-1}$ which does not occur in any portion of the input visible to the first player. Thus $\alpha = (0, \alpha_1, \ldots, \alpha_{k-1})$ does not occur. The first player can announce α using $k-1$ bits. All players then know that α does not occur.

Given an α which all players know does not occur as a row, we show how to compute the number of all 1 rows. Without loss of generality, α is of the form $(0,\dots,0,1,\dots,1)$ since the players can renumber themselves according to α. Define y_i as the number of rows in the input of the form $(0,\dots,0,1,\dots,1)$ starting with i 0s. Now supposing α contains l 0s, $y_l = 0$ is the number of times α occurs. For each $i < l$, using k bits, the ith player announces z_i, the number of rows of the form $(0,\dots,0,*,1,\dots,1)$ where the $*$ occurs at the ith column and is unknown to the ith player. Note that $z_i = y_i + y_{i+1}$. Each player can then privately compute $\sum_{i \text{ even}} z_i - \sum_{i \text{ odd}} z_i = y_0 + y_l = y_0$. This is the number of all 1 rows.

2.2 Cost Analysis

For each of the $\left\lceil \frac{n}{2^{k-1}-1} \right\rceil$ blocks, the protocol uses $k-1$ bits to determine α, and uses k bits to determine z_i for $0 \le i < l$. Since $l \le k$, we have this is a total of $\left\lceil \frac{n}{2^{k-1}-1} \right\rceil ((k-1) + k \cdot k) = (k^2+k-1)\left\lceil \frac{n}{2^{k-1}-1} \right\rceil$ bits as claimed.

3 Preliminaries

For simplicity, we define Boolean functions to use the range $\{-1,+1\}$ with -1 corresponding to true. We let $[m]$ denote the set $\{1,\dots,m\}$. Finally, let $f \circ g$ denote the composition of functions f and g, that is, for functions $f : \{-1,+1\}^n \to \{-1,+1\}, g : X \to \{-1,+1\}$ we define the composition $f \circ g : X^n \to \{-1,+1\}$ by $(f \circ g)(x_1,\dots,x_n) = f(g(x_1),\dots,g(x_n))$.

3.1 Cylinder Intersections

In 2-party communication complexity, we can model protocols as a decision tree with nodes marked by rectangles. In multiparty communication, we use a generalized analogue, which are called cylinder intersections, as introduced by Babai, Nisan, and Szegedy [1].

Definition 3.1 (Cylinder Intersections)**.** A k-dimensional cylinder intersection is a function $\chi : X_1 \times X_2 \times \cdots \times X_k \to \{0,1\}$ of the form $\chi(x_1,\dots .x_k) = \prod_{i=1}^{k} \chi_i(x_1,\dots,x_{i-1},x_{i+1},\dots,x_k)$ with $\chi_i : X_1 \times \cdots \times X_{i-1} \times X_{i+1} \times \cdots \times X_k \to \{0,1\}$.

Thus a k-dimensional cylinder is the product of k functions which do not depend on one of the k coordinates. In the forehead model, this corresponds to a player not knowing the number on his own forehead. At any point in the protocol, what each player knows is independent of his own number.

Note that when $k = 2$, we get that $\chi(x_1,x_2) = \chi_1(x_2)\chi_2(x_1) = \mathbb{1}[x_2 \in S_1] \cdot \mathbb{1}[x_1 \in S_2] = \mathbb{1}[(x_1,x_2) \in S_2 \times S_1]$ for some $S_1, S_2 \subseteq [n]$. This is exactly the indicator for a rectangle.

Recall that in the case of a 2-party protocol, we can divide the input space into rectangles. More explicitly, if $\Pi : X \times Y \to \{-1,+1\}$ is a deterministic protocol with cost c, then $X \times Y$ can be divided into 2^c (possibly empty) disjoint rectangles which lead to the 2^c possible transcripts. The same reasoning on cylinder intersections generalizes to the following proposition in the k-party model.

Proposition 3.2. *Let $\Pi : X_1 \times \cdots \times X_k$ be a deterministic k-party communication protocol with cost c. Then*

$$\Pi = \sum_{i=1}^{2^c} a_i \chi_i$$

for some cylinder intersections $\chi_1,\dots,\chi_{2_c}$ with pairwise disjoint support and some coefficients $a_1,\dots,a_{2^c} \in \{-1,+1\}$.

Now a randomized protocol with cost c is a probability distribution on deterministic protocols of cost exactly c. So the following corollaries are directly implied by Proposition 3.2.

Corollary 3.3. *Let F be a (possibly partial) Boolean function on $X_1 \times \cdots \times X_k$. If $R_\epsilon(F) = c$, then*

$$\begin{aligned} &|F(x_1, \ldots, x_k) - \Pi(x_1, \ldots, x_k)| \leq \frac{\epsilon}{1-\epsilon}, && (x_1, \ldots, x_k) \in \operatorname{dom} F, \\ &|\Pi(x_1, \ldots, x_k)| \leq \frac{1}{1-\epsilon}, && (x_1, \ldots, x_k) \in X_1 \times \cdots \times X_k, \end{aligned}$$

where $\Pi = \sum_\chi a_\chi \chi$ is a linear combination of cylinder intersections with $\sum_\chi |a_\chi| \leq \frac{2^c}{1-\epsilon}$.

Corollary 3.4. *Let |pi be a randomized k-party protocol with domain $X_1 \times X_2 \times \cdots \times X_k$ and cost c. Then*

$$P[\Pi(x_1, \ldots, x_k) = -1] \equiv \sum_\chi a_\chi \chi(x_1, \ldots, x_k)$$

on $X_1 \times \cdots \times X_k$, where each χ is a cylinder intersection and $\sum_\chi |a_\chi| \leq 2^c$.

3.2 Fourier Transform

We define the multidimensional Fourier transform on the Boolean hypercube $\{-1, +1\}^n$.

Definition 3.5 (Fourier Transform)**.** Define $\chi_S(x) = \prod_{i \in S} x_i$ for $S \in [n]$. Note that these form an orthonormal basis over the vector space of real functions over the Boolean hypercube. Every function $\phi : \{-1, +1\}^n \to \mathbb{R}$ has a unique representation $\phi(x) = \sum_{S \in [n]} \hat{\phi}(S) \chi_S(x)$, where $\hat{\phi}(S) = 2^{-n} \sum_{x \in \{-1,+1\}^n} \phi(x) \chi_S(x)$ are called the *Fourier coefficients* of ϕ.

3.3 Polynomial Approximation

Another useful technique, closely related to the consideration of the Fourier transform of a function ϕ, is that of approximating ϕ by a low-degree polynomial.

Definition 3.6. For a partial function $\phi : X \to \mathbb{R}$ and a nonnegative real ϵ, the *ϵ-approximate degree of ϕ*, denoted $\deg_\epsilon(\phi)$, is defined as the least degree of a real polynomial p with

$$\begin{cases} |\phi(x) - p(x)| \leq \epsilon & \text{if } x \in \operatorname{dom} \phi, \\ |p(x)| \leq 1 + \epsilon & \text{otherwise.} \end{cases}$$

Define $E(\phi, d)$ to be the smallest ϵ such that $\deg_\epsilon(\phi) \leq d$.

Notice that if ϕ is a partial Boolean function, then a polynomial p that ϵ-approximates any extension of ϕ to a total Boolean function f also gives an ϵ-approximation to ϕ, since $|p(x)| \leq |f(x) - p(x)| + |f(x)| \leq \epsilon + 1$ when $x \notin \operatorname{dom} \phi$.

We will make use of a few results on approximate degrees. The first is a result due to Sherstov [12], relating approximate degree to an inequality involving a function with no low-order Fourier coefficients:

Theorem 3.7. *Given a partial real-valued function ϕ on $\{-1, +1\}^n$, we have $\deg_\epsilon(\phi) > d$ if and only if there exists $\psi : \{-1, +1\}^n \to \mathbb{R}$ such that*

$$\sum_{x \in \operatorname{dom} \phi} \phi(x) \psi(x) - \sum_{x \notin \operatorname{dom} \phi} |\psi(x)| - \epsilon \|\psi\|_1 > 0,$$

and $\hat{\psi}(S) = 0$ for $|S| \leq d$.

Let $\widetilde{\mathrm{AND}}_n$ be the promise version of the function AND_n, so its domain consists of sets of inputs at most one of which is false. We have the following bound proven by Nisan and Szegedy [11]:

Theorem 3.8.

$$\deg_{1/3}(\widetilde{\mathrm{AND}}_n) = \Theta(\sqrt{n}),$$
$$\deg_{1/3}(\mathrm{AND}_n) = \Theta(\sqrt{n}).$$

Corollary 3.9.

$$\deg_{1/3}(\mathrm{NOR}_n) = \deg_{1/3}(\mathrm{AND}_n) = \Theta(\sqrt{n}).$$

The proof of Theorem 3.7 is an application of LP duality; see [12]. Theorem 3.8 follows from a simple analytic result that gives a $\Theta(\sqrt{n})$ lower bound on the degree of a polynomial given a large derivative at one point and some bounds on its values at n consecutive integers; see [11]. Corollary 3.9 follows from the bijection $\mathrm{NOR}_n(x_1, \ldots, x_n) = \mathrm{AND}_n(1-x_1, \ldots, 1-x_n)$, which doesn't change the degree.

The importance of Theorem 3.8 to the disjointness problem comes from the observation that for any m, r, k we have $\mathrm{DISJ}_{mr,k} = \mathrm{AND}_m \circ \mathrm{DISJ}_{r,k}$ and $\mathrm{UDISJ}_{mr,k} = \widetilde{\mathrm{AND}}_m \circ \mathrm{UDISJ}_{r,k}$. Similarly, we take notice of Corollary 3.9 because $\mathrm{DISJ}_{n,k}$ is equivalent to $G^{\mathrm{NOR}}_{n,k}$.

4 Discrepancy

In the case of randomized communication complexity, our analysis needs some refinement beyond the direct use of cylinder intersections. The concept of *discrepancy* is an important tool that serves to generalize the use of monochromatic rectangles in the 2-party case, or of cylinder intersections in the k-party case, to the analysis of a randomized protocol.

In the 2-party setting, the discrepancy of a (total) function $f : X \times Y \to \{-1, +1\}$ with respect to a probability distribution P is defined as

$$\mathrm{disc}_P(f) = \max_{R=S\times T} |P(R \cap f^{-1}(1)) - P(R \cap f^{-1}(-1))|$$
$$= \max_R \left| \sum_{(x,y)\in X\times Y} f(x,y)P(x,y)\chi_R(x,y) \right|,$$

where R ranges over all rectangles and χ_R is the indicator function for the rectangle R (i.e. $\chi_R(x,y) = 1$ if $(x,y) \in R$ and $\chi_R(x,y) = 0$ otherwise).

The last expression for discrepancy above extends well to the k-party setting by simply replacing rectangles by general cylinder intersections. We have, for a total function F,

$$\mathrm{disc}_P(F) = \max_\chi \left| \sum_{x \in X_1\times X_2\times\cdots\times X_k} F(x)P(x)\chi(x) \right|,$$

where χ ranges over all cylinder intersections.

We can extend this definition to a partial function F by setting

$$\mathrm{disc}_P(F) = \sum_{x\notin \mathrm{dom}\, F} P(x) + \max_\chi \left| \sum_{x\in \mathrm{dom}\, F} F(x)P(x)\chi(x) \right|.$$

We define the overall discrepancy to be $\mathrm{disc}(F) = \min_P \mathrm{disc}_P(F)$, the least discrepancy over all possible probability distributions.

Intuitively, discrepancy is a measure of how strongly a function correlates with a particular cylinder intersection, and computing it plays a similar role to finding a Fourier coefficient of maximal magnitude.

It has been long established [6] for the 2-party model that low discrepancy leads to lower bounds on communication complexity. Namely, for any function $F : \{0,1\}^n \to \{-1,+1\}$, we have $2^{R_\epsilon(F)} \geq \frac{1-2\epsilon}{\mathrm{disc}(F)}$. This also holds in the k-party model as well as in the partial functions extension. Sherstov [13] proves the following theorem.

Theorem 4.1 (Discrepancy Method)**.** *Let F be a (possibly partial) Boolean function on $X_1 \times \cdots \times X_k$. Then*

$$2^{R_\epsilon(F)} \geq \frac{1-2\epsilon}{\mathrm{disc}(F)}.$$

A stronger technique is to find a function Ψ which is highly correlated with F but has low correlation with all cylinder intersections. This method is referred to as the generalized discrepancy method, which is proved in [13].

Theorem 4.2 (Generalized Discrepancy Method)**.** *Let $F : X_1 \times \cdots \times X_k \to \{-1,+1\}$ be a (possibly partial) Boolean function. Then for every nonzero $\Psi : X_1 \times \cdots \times X_k \to \mathbb{R}$,*

$$2^{R_\epsilon(F)} \geq \frac{1-\epsilon}{\max_\chi |\langle \chi, \Psi \rangle|} \left(\sum_{x \in \mathrm{dom}\, F} F(x)\Psi(x) - \sum_{x \notin \mathrm{dom}\, F} |\Psi(x)| - \frac{\epsilon}{1-\epsilon} \|\Psi\|_1 \right),$$

where $\max_\chi$ *is over cylinder intersections χ.*

We have the following useful upper bound on discrepancy given in [13], which has a very simple proof:

Proposition 4.3. *For a (total) function $F : X \times Y \to -1, 1$ and a probability distribution P on $X \times Y$, define Φ to be the $|X| \times |Y|$ matrix $\Phi = [F(x,y)P(x,y)]_{x \in X, y \in Y}$. Then*

$$\mathrm{disc}_P(F) \leq \|\Phi\| \sqrt{|X||Y|}.$$

Here $\|\Phi\|$ denotes the spectral norm of Φ as a linear map, i.e. $\|\Phi\| = \max_{x \neq 0} \frac{\|\Phi x\|}{\|x\|}$, where x ranges over nonzero vectors in $\mathbb{R}^X$.

Proof. We have $\mathrm{disc}_P(F) = \max_R \left| \sum_{(x,y) \in X \times Y} F(x,y)P(x,y)\chi_R(x,y) \right|$. Pick χ_R attaining this maximum. Say $R = S \times T$, and let $1_S, 1_T$ be the indicator functions for S, T respectively, and let v_S, v_T be the corresponding indicator vectors in $\mathbb{R}^X, \mathbb{R}^Y$ respectively. Then

$$\mathrm{disc}_P(F) = \left| \sum_{(x,y) \in X \times Y} F(x,y)P(x,y)1_S(x)1_T(y) \right| = |v_T \cdot (\Phi v_S)| \leq \|v_T\| \|\Phi v_S\|$$

$$\leq \|v_T\| \|\Phi\| \|v_S\| \leq \|\Phi\| \sqrt{|X||Y|}.$$

Here we have used Cauchy-Schwarz as well as the definition of the spectral norm. □

Sherstov uses this to prove a general discrepancy-type result relevant to the analysis of an XOR of m copies of $\mathrm{UDISJ}_{n,k}$. Before stating this result, we make a few definitions to simplify our notation. Specifically, when $X_1, \ldots, X_r$ are any 0,1-matrices or vectors with n rows, let $D(X_1, \ldots, X_r) = -1$ if for $1 \leq j \leq n$, one of the X_i has a 0 in row j, and let $D(X_1, \ldots, X_r) = 1$ otherwise. So when $x_1, \ldots, x_k$ are vectors, we have $D(x_1, \ldots, x_k) = \mathrm{DISJ}_{n,k}(x_1, \ldots, x_k)$.

Let $\mathcal{U}_n$ be the uniform distribution on the set $\{0,1\}^n$ of 0, 1-vectors in $\mathbb{R}^n$, and let $\mu_{n,k}$ be the uniform distribution on $n \times k$ 0, 1-matrices such that exactly one row consists of all ones, thus admitting at most one coordinate at which disjointness can fail. Hence the support of $\mathcal{U}_n \times \mu_{n,k-1}$, viewed as a distribution on sets of k vectors, is contained in the domain of $\text{UDISJ}_{n,k}$.

Finally, recall that for a distribution P on a set X, the notation $x \sim P$ means that $x \in X$ is a random variable distributed according to P. For any function F, the expectation with respect to P is $\mathbb{E}_{x\sim P} F(x) = \sum_{x\in X} P(x)F(x)$.

Now, define

$$\Gamma_k(n_1, n_2, \dots, n_m) = \max_{\chi} \left| \mathop{\mathbb{E}}_{(x^1,W^1),\dots,(x^m,W^m)} \left[\chi \cdot \prod_{i=1}^{m} D(x^i, W^i) \right] \right|,$$

for any positive integers $n_1, \dots, n_m$. Here $(x^i, W^i) \sim \mathcal{U}_{n_i} \times \mu_{n_i,k}$, and χ ranges over cylinder intersections of dimension $k+1$ on $X_1 \times \cdots X_{k+1} = (\{0,1\}^{n_1+\cdots+n_m})^{k+1}$, where X_j consists of the jth columns of each (x^i, W^i). That is, we can view the overall distribution $\mu^{(m)}$ of all the (x^i, W^i) as a distribution over $(k+1)$-tuples of collections of m vectors $(v_1, \dots, v_m) \in \{0,1\}^{n_1} \times \cdots \times \{0,1\}^{n_m}$.
.

The connection with the discrepancy of unique disjointness becomes clear when we notice that $\Gamma_k(n_1, n_2, \dots, n_m) = \text{disc}_{\mu^{(m)}}(F)$, where $F = \prod_{i=1}^{m} D(x^i, W^i)$, viewed as a function on $X_1 \times \cdots X_{k+1} = (\{0,1\}^{n_1+\cdots+n_m})^{k+1}$. So $F = -1$ if and only if an odd number of the (x^i, W^i) satisfy disjointness. Each parameter (x^i, W^i) corresponds to an instance of UDISJ, so F computes the XOR of m instances of unique disjointness. That is,

$$\Gamma_k(n_1, n_2, \dots, n_m) = \text{disc}_{\mu^{(m)}}(\text{UDISJ}_{n_1,k+1} \oplus \cdots \oplus \text{UDISJ}_{n_m,k+1}).$$

Sherstov [13] gives the following bound:

Theorem 4.4. *For all positive integers $n_1, n_2, \dots, n_m$ and k,*

$$\Gamma_k(n_1, n_2, \dots, n_m) \le \frac{(2^k - 1)^m}{\sqrt{n_1 n_2 \cdots n_m}}.$$

The proof given in [13] uses induction on k, reducing to a smaller-dimensional cylinder intersection through use of conditional independence and bounding of the expectations that arise. It is instructive to look at the proof of the base case, expanding on the proof sketch that Sherstov gives.

Proposition 4.5. *For all positive integers $n_1, n_2, \dots, n_m$,*

$$\Gamma_1(n_1, n_2, \dots, n_m) \le \frac{1}{\sqrt{n_1 n_2 \cdots n_m}}.$$

Proof. Note that for $k = 1$, each W^i in the definition of Γ_k reduces to a vector with exactly one nonzero component. Thus, W^i is effectively a uniformly random integer $j_i \in \{1, 2, \dots, n_i\}$. Then $D(x^i, W^i) = -1$ if and only if $x^i_{j_i}$, the j_ith entry of x^i, is 0. So

$$\Gamma_1(n_1, n_2, \dots, n_m) = \text{disc}_{\mu^{(m)}} \left(\prod_{i=1}^{m} (-1)^{1+x^i_{j_i}} \right).$$

As in Proposition 4.3, then, we construct a matrix of the form $\Phi = [F(x,y)P(x,y)]_{x\in X, y\in Y}$, where $x = (x^1, \dots, x^m) \in X = \{0,1\}^{n_1} \times \cdots \times \{0,1\}^{n_m}$ and $y = (j_1, \dots, j_m) \in Y = [1, n_1] \times \cdots \times [1, n_m]$. So $P(x,y) = \frac{1}{\prod_{i=1}^{m} n_i 2^{n_i}}$ for all x, y, and $F(x,y) = \prod_{i=1}^{m} (-1)^{1+x^i_{j_i}}$. Since $F \cdot P$ is a product of m

functions $F_i = \frac{(-1)^{1+x^i_{j_i}}}{n_i 2^{n_i}}$ such that each F_i only depends on (x_i, j_i), this matrix can be decomposed into a tensor product $\bigotimes_{i=1}^m \Phi_i$ of m matrices $\Phi_i = \left[\frac{(-1)^{1+x^i_{j_i}}}{n_i 2^{n_i}}\right]_{x^i \in \{0,1\}^{n_i}, j_i \in [1,n_i]}$.

The spectral norm is multiplicative with respect to tensor products, so

$$\|\Phi\| = \prod_{i=1}^m \|\Phi_i\| = \prod_{i=1}^m \frac{\left\| \left[-(-1)^{x^i_{j_i}}\right]_{x^i \in \{0,1\}^{n_i}, j_i \in [1,n_i]} \right\|}{n_i 2^{n_i}}.$$

Let $M_i = \left[(-1)^{x^i_{j_i}}\right]_{x^i \in \{0,1\}^{n_i}, j_i \in [1,n_i]}$. The rows of M_i are all possible $1, -1$-vectors, so for any v, when $\|M_i v\|^2$ is expanded, all the cross terms $v_i v_j$ cancel out and we are left with $2^{n_i} \sum v_i^2 = 2^{n_i} \|v\|^2$. So $\|M_i\| = \sqrt{2^{n_i}}$, and thus

$$\|\Phi\| = \prod_{i=1}^m \frac{\sqrt{2^{n_i}}}{n_i 2^{n_i}}.$$

Then Proposition 4.3 gives

$$\Gamma_1(n_1, \ldots, n_m) = \operatorname{disc}_{\mu^{(m)}}(F) \leq \|\Phi\| \sqrt{2^{n_1 + \cdots + n_m} n_1 \cdots n_m} = \frac{1}{\sqrt{n_1 n_2 \cdots n_m}},$$

which is the desired result. □

5 The Pattern Matrix Method and an Early Lower Bound

The goal of this section is to see an earlier lower bound result while gaining familiarity with the generalized discrepancy method.

In an earlier work, Sherstov [12] developed the pattern matrix method for 2-party quantum communication lower bounds, which built upon the generalized discrepancy method by generating an appropriate probability distribution. Chattopadhyay and Ada [5] extended the idea to higher dimensions for multiple players.

The goal of these methods is to transform the problem into one computing $F_k^f(x, y_1, y_2, \ldots, y_{k-1}) = f(x|_{\sigma(y_1, \ldots, y_{k-1})})$ for a particular function $f : \{0,1\}^m \to \{-1, +1\}$, where $x|_{\sigma(y_1, \ldots, y_{k-1})}$ is an m-element subset of the bits of x defined by the *selector* $\sigma(y_1, \ldots, y_{k-1})$. A function g is then chosen along with a probability distribution μ such that f and g (and hence F_k^f and F_k^g) are highly correlated over μ but F_k^g is almost uncorrelated with all cylinder intersections, thus allowing for lower bounds via discrepancy.

The selector used in [5] is to let $y_i \in [\ell]^m$ for some ℓ with $x \in \{0,1\}^n$ with $n = \ell^{k-1} m$. x can be thought of as an m-tuple of k-dimensional arrays with size ℓ in each dimension. The selector σ is defined by letting the jth element of each of the y_i be an index in one of the dimensions of the jth array of x. That is,

$$\left(x|_{\sigma(y_1, \ldots, y_{k-1})}\right)_j = x_{j, y_{1,j}, y_{2,j}, \ldots, y_{k-1,j}}.$$

This formulation allows us to perform the following reduction.

Lemma 5.1. *For $n = \ell^{k-1} m$, if $f : \{0,1\}^m \to \{-1, +1\}$ and $f' : \{0,1\}^n \to \{-1, +1\}$ have the property that $f(z) = f'(z')$ when z and z' contain the same number of 1s, then*

$$R_\epsilon(F_k^f) \leq R_\epsilon(G_{n,k}^{f'}).$$

Proof. There are functions $\Gamma_i : [\ell]^m \to \{0,1\}^n$, namely the bitmask for the coordinate in the ith dimension of each array, such that $F_k^f(x, y_1, \ldots, y_{k-1}) = G_{n,k}^{f'}(x, \Gamma_1(y_1), \ldots, \Gamma_{k-1}(y_{k-1}))$ for all $x, y_1, \ldots, y_{k-1}$. Thus players for F_k^f can privately convert their inputs and use the protocol for $G_{n,k}^{f'}$. □

Remark. The functions $f = \mathrm{NOR}_m$ and $f' = \mathrm{NOR}_n$ satisfy the conditions of this lemma, and will lead to bounds on $\mathrm{DISJ}_{n,k}$.

We bound the complexity of F_k^f in terms of discrepancy.

Lemma 5.2. *Let μ be a probability distribution on $\{-1,+1\}^m$ and let λ be the probability distribution on $\{-1,+1\}^n \times ([\ell]^m)^{k-1}$ defined from μ as $\lambda(x, y_1, \ldots, y_{k-1}) = \frac{\mu(x|_{\sigma(y_1,\ldots,y_{k-1})})}{\ell^{m(k-1)}2^{n-m}}$. If $f : \{-1,+1\}^m \to \{-1,+1\}$ satisfies $\mathbb{E}_{x\sim\mu} f(x)\chi_S(x) = 0$ for all $|S| < d$ and some positive integer d, then*

$$\left(\mathrm{disc}_\lambda(F_k^f)\right)^{2^{k-1}} \leq \sum_{j=d}^{(k-1)m} \binom{(k-1)m}{j} \left(\frac{2^{2^{k-1}-1}}{\ell - 1}\right)^j.$$

Remark. For $\ell - 1 \geq \frac{2^{2^k}(k-1)em}{d}$ and $d > 2$, we have $\mathrm{disc}_\lambda(F_k^f) \leq \frac{1}{2^{d/2^{k-1}}}$.

Proof sketch. The full proof can be found in [5]. The main idea is to compute the discrepancy with respect to an arbitrary cylinder intersection and bound it using algebraic and Fourier properties.

We have the discrepancy with respect to cylinder intersection χ is

$$\begin{aligned}\mathrm{disc}_\lambda(F_k^f) &= \left|\sum_{x,y_1,\ldots,y_{k-1}} F_k^f(x, y_1, \ldots, y_{k-1})\chi(x, y_1, \ldots, y_{k-1})\lambda(x, y_1, \ldots, y_{k-1})\right| \\ &= \left|\mathop{\mathbb{E}}_{x,y_1,\ldots,y_{k-1}} F_k^f(x, y_1, \ldots, y_{k-1})\chi(x, y_1, \ldots, y_{k-1})\mu(x|_{\sigma(y_1,\ldots,y_{k-1})})\right|.\end{aligned}$$

With repeated applications of the triangle inequality and the Cauchy-Schwartz inequality, this can be transformed to

$$\left(\mathrm{disc}_\lambda(F_k^f)\right)^{2^{k-1}} \leq \mathop{\mathbb{E}}_{y_{1,0},y_{1,1},\ldots,y_{k-1,0},y_{k-1,1}} \left|\mathop{\mathbb{E}}_x \prod_{u\in\{0,1\}^{k-1}} F_k^f(x, y_{1,u_1}, \ldots, y_{k-1,u_{k-1}})\mu(x|_{\sigma(y_1,\ldots,y_{k-1})})\right|.$$

Let $r = \sum_i |y_{i,0} \cap y_{i,1}|$. We make the following claims, of which the proofs can be found in [5].

Claim 5.3.

$$\left|\mathop{\mathbb{E}}_x \prod_{u\in\{0,1\}^{k-1}} F_k^f(x, y_{1,u_1}, \ldots, y_{k-1,u_{k-1}})\mu(x|_{\sigma(y_1,\ldots,y_{k-1})})\right| \leq 2^{(2^{k-1}-1)r}.$$

This follows from μ being a probability distribution.

Claim 5.4. *If $r < d$,*

$$\left|\mathop{\mathbb{E}}_x \prod_{u\in\{0,1\}^{k-1}} F_k^f(x, y_{1,u_1}, \ldots, y_{k-1,u_{k-1}})\mu(x|_{\sigma(y_1,\ldots,y_{k-1})})\right| = 0.$$

This claim can be proved using the Fourier theoretic properties of $\hat{f}(S) = 0$ for $|S| < d$.

Together with some probability and combinatorial identities, the bounds made in these claims can be used to obtain the theorem. □

We are now in a position to prove a bound on $G_k^{f'}$.

Theorem 5.5. *Let $f : \{0,1\}^m \to \{-1,+1\}$ have δ-approximate degree d. Let $n \geq \left(\frac{2^{2^k}(k-1)e}{d}\right)^{k-1} m^k$, and $f' : \{0,1\}^n \to \{-1,+1\}$ satisfy $f(z) = f(z')$ when z and z' contain the same number of 1s. Then*

$$R_\epsilon(G_k^{f'}) \geq \frac{d}{2^{k-1}} + \log(\delta - 2\epsilon).$$

Proof. By Theorem 3.7, there is a function $\psi : \{0,1\}^m \to \mathbb{R}$ such that $\langle f, \psi\rangle > \delta$ and $\langle \psi, \chi_S\rangle = 0$ for $|S| < d$. Then there is a function $g = \operatorname{sgn} \circ \psi$ and probability distribution $\mu = \frac{\psi}{\|\psi\|_1}$ such that $\mathbb{E}_{x\sim\mu} f(x)g(x) > \delta\, \mathbb{E}_{x\sim\mu} |g(x)| = \delta$ and $\mathbb{E}_{x\sim\mu} g(x)\chi_S(x) = 0$ for $|S| < d$. These g and μ satisfy Lemma 5.2, and so we have

$$\operatorname{disc}_\lambda(F_k^g) \leq \frac{1}{2^{d/2^{k-1}}},$$

for λ as defined in Theorem 3.7 and $\ell \geq 2^{2^k}(k-1)em/d$. Since $n = \ell^{k-1}m$, this holds for $n \geq (\frac{2^{2^k}(k-1)e}{d})^{k-1}m^k$. Now we have $\mathbb{E}_{x\sim\lambda} F_k^f(x)F_k^g(x) = \mathbb{E}_{x\sim\mu} f(x)g(x) > \delta$. A restatement of the generalized discrepancy method from Theorem 4.2 for total functions is the following.

Corollary 5.6 (Generalized Discrepancy Method)**.** *For total functions $F : X_1 \times \cdots \times X_k \to \{-1,+1\}$ and $\Psi : X_1 \times \cdots \times X_k \to \mathbb{R}$,*

$$2^{R_\epsilon(F)} \geq \frac{(1-\epsilon)\langle F, \Psi\rangle - \epsilon\|\Psi\|_1}{\max_\chi |\langle \chi, \Psi\rangle|} \geq \frac{\langle F, \Psi\rangle - 2\epsilon\|\Psi\|_1}{\max_\chi |\langle \chi, \Psi\rangle|}$$

Let $\Psi(x) = F_k^g(x)\lambda(x)$. Note that $\|\Psi\|_1 = 1$. By Corollary 5.6, we have

$$\begin{aligned} 2^{R_\epsilon(F_k^f)} &\geq \frac{\mathbb{E}_{x\sim\lambda} F_k^f(x)F_k^g(x) - 2\epsilon}{\max_\chi |\mathbb{E}_{x\sim\lambda} \chi(x)F_k^g(x)|} \\ &> \frac{\delta - 2\epsilon}{\operatorname{disc}_\lambda(F_k^g)}. \end{aligned}$$

Applying the value for $\operatorname{disc}_\lambda(F_k^g)$ and rearranging yields the desired result. □

Corollary 5.7.

$$R_\epsilon(\mathrm{DISJ}_k) = \Omega\left(\frac{n^{\frac{1}{k+1}}}{2^{2^k}(k-1)2^{k-1}}\right)$$

for constant $\epsilon > 0$.

Proof. Let $f = \mathrm{NOR}_m$ and $f' = \mathrm{NOR}_n$. By Theorem 3.9, we have $\deg_{1/3}(\mathrm{NOR}_m) = \Theta(\sqrt{m})$. Setting $n = (\frac{2^{2^k}(k-1)e}{\deg_{1/3}(\mathrm{NOR}_m)})^{k-1}m^k$, and rearranging yields the corollary for $\epsilon < 1/3$. This can be reduced to arbitrary ϵ by an amplification. □

6 The Main Theorem

The key to Sherstov's proof of Theorem 1.1 is the following bound relating approximate degree to randomized communication complexity:

Theorem 6.1. *Let f be a partial Boolean function on $\{-1,+1\}^n$, and let $F = f \circ \mathrm{UDISJ}_{r,k}$. For any $\epsilon, \delta \geq 0$, let $c = R_\epsilon(F)$ be the randomized communication complexity with error ϵ of F as a k-party communication problem, and let $d = \deg_\delta(f)$. Then we have*

$$2^c \geq (\delta - \epsilon(1+\delta)) \left(\frac{d\sqrt{r}}{2^k e n} \right)^d .$$

Given this theorem, the proof of Theorem 1.1 is straightforward:

Proof of Theorem 1.1. By Theorem 3.8, for some $c > 0$ we have $\deg_{1/3}(\widetilde{\mathrm{AND}}_m) > c\sqrt{m}$ for all m. Since $\mathrm{UDISJ}_{mr,k} = \widetilde{\mathrm{AND}}_m \circ \mathrm{UDISJ}_{r,k}$, applying Theorem 6.1 with $\epsilon = 1/5, \delta = 1/3, n = m, r = 4^{k+2}m/c^2, f = \widetilde{\mathrm{AND}}_m$ gives $d \geq c\sqrt{m}$ and

$$R_{1/5}(\mathrm{UDISJ}_{mr,k}) \geq \log\left(1/15 + d\log\left(\frac{d2^{k+2}\sqrt{m}}{c2^k em}\right)\right) \geq \Omega\left(\sqrt{m}\log\left(\frac{4}{e}\right)\right) = \Omega(\sqrt{m}).$$

Replacing $R_{1/5}$ by $R_{1/3}$ should only change the bound by a constant factor. Now, setting $m \approx \sqrt{c^2 n/4^{k+2}}$ gives $mr \approx n$, so

$$R_{1/3}(\mathrm{DISJ}_{n,k}) \geq R_{1/3}(\mathrm{UDISJ}_{n,k}) \geq \Omega\left(\frac{n}{4^k}\right)^{1/4},$$

as claimed. □

We present two proofs of Theorem 6.1 given by Sherstov in [13], rearranged and clarified here to highlight the motivation behind each step.

6.1 Primal Proof

The first proof is described as "primal" because, unlike previous work on results of this nature, it does not rely on switching to the dual view of the problem. The method used in this proof resurfaces in the proofs of direct product results later in the paper.

Primal proof of Theorem 6.1. The idea of this proof is to construct a low-degree polynomial approximating f given a protocol attaining communication cost $c = R_\epsilon(F)$.

As in Section 4, we think of $F(X_1, \ldots, X_n)$ in the communication protocol setting as a function on k inputs $Y_1, \ldots, Y_k$, where $Y_j \in \{0,1\}^{r \times n}$ consists of the jth columns of each X_i. Then by 3.3, F is approximated by a linear combination $\Pi = \sum_\chi a_\chi \chi$ of k-dimensional cylinder intersections $\chi(Y_1, \ldots, Y_k) = \chi(X_1, \ldots, X_n)$, with $\sum_\chi |a_\chi| \leq 2^c/(1-\epsilon)$.

We want to convert this approximation into an approximation of f by a low-degree polynomial. To translate the approximation from F to f, we will introduce an operator M sending each function $G : (\{0,1\}^{r\times k})^n \to \{-1,+1\}$ to a function $MG : \{0,1\}^n \to \{-1,+1\}$. Since $F(X_1, \ldots, X_n) = f(\mathrm{UDISJ}_{r,k}(X_1), \ldots, \mathrm{UDISJ}_{r,k}(X_n))$, M will be an averaging operator that retains information about the values of $\mathrm{UDISJ}_{r,k}$ on the inputs.

Let $\mu = \mathcal{U}_r \times \mu_{r,k-1}$, and let μ_{+1}, μ_{-1} be the probability distributions that μ induces on $\mathrm{UDISJ}_{r,k}^{-1}(-1), \mathrm{UDISJ}_{r,k}^{-1}(1)$ respectively. So, for example, $\mu_{+1}(X_0)$ is the probability of choosing $X = X_0$ when sampling X according to μ conditioned on $\mathrm{UDISJ}_{r,k}(X) = 1$. Let $X = (x, W) \sim \mu$.

For any fixed W, $\mathrm{UDISJ}_{r,k}(x,W)$ only depends on one bit of x, so since x is picked uniformly, $(x,W) \in \mathrm{UDISJ}_{r,k}^{-1}(-1)$ with probability $\frac{1}{2}$. Thus for $z = \pm 1$ we have

$$\mu_z(x,W) = \begin{cases} 2\mu(x,W) & \text{if } D(x,W) = z, \\ 0 & \text{otherwise.} \end{cases}$$

Thus $\mu = \frac{\mu_{+1}+\mu_{-1}}{2}$.

Now we can define the operator M as follows: given a function $G : (\{0,1\}^{r\times k})^n$, for any $z = (z_1, \cdots, z_n) \in \{-1,+1\}^n$ let

$$(MG)(z) = \mathop{\mathbb{E}}_{X_1\sim\mu_{z_1},\ldots,X_n\sim\mu_{z_n}}[G(X_1,\ldots,X_n)].$$

So, for $z \in \operatorname{dom} f$,

$$\begin{aligned} MF(z) &= \mathop{\mathbb{E}}_{X_1\sim\mu_{z_1},\ldots,X_n\sim\mu_{z_n}}[f(\mathrm{UDISJ}_{r,k}(X_1),\ldots,\mathrm{UDISJ}_{r,k}(X_n))] \\ &= \mathop{\mathbb{E}}_{X_1\sim\mu_{z_1},\ldots,X_n\sim\mu_{z_n}}[f(z_1,\ldots,z_n)] = f(z), \end{aligned}$$

as wanted. Note that $\|MG\|_\infty \le \|G\|_\infty$ for any G by the triangle inequality. Also, M is linear by linearity of expectation, so $M\Pi = \sum_\chi a_\chi M\chi$.

By definition we have $M\chi(x) = \sum_{S\subseteq\{1,\ldots,n\}} \widehat{M\chi}(S)\prod_{i\in S} x_i$, so keeping only the terms with $|S| < d$ gives a degree $d-1$ approximation of $M\chi$, extending to an approximation of $M\Pi$ and thus of $MF = f$.

Let $p(x) = \sum_\chi a_\chi \sum_{|S|\le d-1} \widehat{M\chi}(S)\prod_{i\in S} x_i$. Then p is a degree $d-1$ polynomial with

$$\|M\Pi - p\|_\infty = \left\| \sum_\chi a_\chi \sum_{|S|\ge d} \widehat{M\chi}(S) \prod_{i\in S} x_i \right\|_\infty \le \sum_\chi |a_\chi| \sum_{|S|\ge d} |\widehat{M\chi}(S)|.$$

Now we bound the Fourier coefficients $\widehat{M\chi}(S) = \mathbb{E}_{z\in\{-1,+1\}^n}(M\chi)(z)\prod_{i\in S} z_i$. Since $\mu = \frac{\mu_{+1}+\mu_{-1}}{2}$, the expectation over z smooths the expectation over $\mu_{z_1}\times\cdots\times\mu_{z_n}$ into an expectation over $\mu\times\cdots\times\mu$. That is, since $z_i = D(X_i)$ by construction,

$$|\widehat{M\chi}(S)| = \left| \mathop{\mathbb{E}}_{X_1,\ldots,X_n\sim\mu}\left[\chi(X_1,\ldots,X_n)\prod_{i\in S} D(X_i)\right]\right|.$$

Fix $S = \{i_1,\ldots,i_m\}$ and let $\{1,\ldots,n\}\setminus S = \{j_1,\ldots,j_{n-m}\}$. For any choice $A_1,\ldots,A_{n-m}$ of $X_{j_1},\ldots,X_{j_{n-m}}$, let $\chi_{A_1,\ldots,A_{n-m}}(X_{i_1},\ldots,X_{i,m}) = \chi(X_1,\ldots,X_n)$ evaluated with $X_{j_i} = A_i$ for all i. Then each $\chi_{A_1,\ldots,A_{n-m}}$ is still a k-dimensional cylinder intersection, since a factor of χ independent of an input Y_j will still be independent of Y_j when some of the columns of each input are fixed. So

$$\begin{aligned} &\left| \mathop{\mathbb{E}}_{X_1,\ldots,X_n\sim\mu}\left[\chi(X_1,\ldots,X_n)\prod_{i\in S} D(X_i)\right]\right| \\ \le& \mathop{\mathbb{E}}_{A_{j_1},\ldots,A_{j_{n-m}}\sim\mu}\left| \mathop{\mathbb{E}}_{X_{i_1},\ldots,X_{i_m}\sim\mu}\left[\chi_{A_{j_1},\ldots,A_{j_{n-m}}}(X_{i_1},\ldots,X_{i_m})\prod_{i\in S} D(X_i)\right]\right| \\ \le& \max_\chi \left| \mathop{\mathbb{E}}_{X_{i_1},\ldots,X_{i_m}\sim\mu}\left[\chi\cdot\prod_{l=1}^m D(X_{i_l})\right]\right| = \Gamma_{k-1}(r,r,\ldots,r), \end{aligned}$$

where Γ_{k-1} has $m = |S|$ inputs of r. By Theorem 4.4 we have $\Gamma_{k-1}(r, r, \ldots, r) \leq \frac{(2^{k-1}-1)^m}{r^{m/2}} \leq \left(\frac{2^{k-1}}{\sqrt{r}}\right)^m$. So,

$$\|M\Pi - p\|_\infty \leq \sum_\chi |a_\chi| \sum_{|S|\geq d} |\widehat{M\chi}(S)| \leq \sum_\chi |a_\chi| \sum_{|S|\geq d} \left(\frac{2^{k-1}}{\sqrt{r}}\right)^{|S|}$$
$$\leq \frac{2^c}{1-\epsilon} \sum_{m=d}^{n} \binom{n}{m} \left(\frac{2^{k-1}}{\sqrt{r}}\right)^m.$$

When $\frac{2^k en}{d\sqrt{r}} \leq 1$, applying binomial sum bounds then gives $\|M\Pi - p\|_\infty \leq \frac{2^c}{1-\epsilon}\left(\frac{2^k en}{d\sqrt{r}}\right)^d$. Otherwise, replace p by 0. Then $\|M\Pi\|_\infty \leq \|\Pi\|_\infty \leq \frac{1}{1-\epsilon}$ by the specifications of the approximation of F by Π, so we still have $\|M\Pi - p\|_\infty \leq \frac{2^c}{1-\epsilon}\left(\frac{2^k en}{d\sqrt{r}}\right)^d$. Hence, $E(M\Pi, d-1) \leq \frac{2^c}{1-\epsilon}\left(\frac{2^k en}{d\sqrt{r}}\right)^d$.

So, for some polynomial p of degree at most $d-1$, we have, for $x \in \operatorname{dom} f$,

$$|f(x) - p(x)| \leq |f(x) - M\Pi(x)| + \|M\Pi - p\|_\infty \leq \frac{\epsilon}{1-\epsilon} + E(M\Pi, d-1).$$

For $x \notin \operatorname{dom} f$, we have

$$|p(x)| \leq \|M\Pi\|_\infty + \|M\Pi - p\|_\infty \leq \frac{1}{1-\epsilon} + E(M\Pi, d-1) = \frac{\epsilon}{1-\epsilon} + E(M\Pi, d-1) + 1.$$

So

$$E(f, d-1) \leq \frac{\epsilon}{1-\epsilon} + E(M\Pi, d-1) \leq \frac{\epsilon}{1-\epsilon} + \frac{2^c}{1-\epsilon}\left(\frac{2^k en}{d\sqrt{r}}\right)^d.$$

Since $\deg_\delta(f) = d$, we must have $\delta < E(f, d-1) \leq \frac{\epsilon}{1-\epsilon} + \frac{2^c}{1-\epsilon}\left(\frac{2^k en}{d\sqrt{r}}\right)^d$, which rearranges to the bound desired.

□

6.2 Dual Proof

Unlike previous analyses using the generalized discrepancy method and pattern matrix techniques, Sherstov gives an explicit probability distribution, for which a selector can be applied which exploits conditional properties of the distribution. Namely, the probability distribution acts on a subspace of $(\{0,1\}^{r\times k})^n$, which allows a selector to work which would not have worked on the entire $(\{0,1\}^{r\times k})^n$ space.

With this selector, we construct our highly correlated function Ψ and then proceed to bound its inner product with all k-dimensional cylinder intersections. These allow us to use the generalized discrepancy method to place a bound on F.

Dual proof of Theorem 6.1. We begin by dividing the input $(\{0,1\}^{r\times k})^n$ into n blocks of $\{0,1\}^{r\times k}$. Our selector will select one bit from each block.

Consider the same probability distribution $\mu = \mathcal{U}_R \times \mu_{r,k-1}$ on the domain of $\text{UDISJ}_{r,k}$. Let $d = \deg_\delta(f)$. By Theorem 3.7, there is a function $\psi : \{-1,+1\}^n \to \mathbb{R}$ which satisfies

$$\sum_{x\in\operatorname{dom} f} f(x)\psi(x) - \sum_{x\notin\operatorname{dom} f} |\psi(x)| > \delta,$$
$$\|\psi\|_1 = 1,$$
$$\hat{\psi}(S) = 0, \qquad |S| < d.$$

Recall that each element A from our probability distribution μ has at most one row which can contain all 1s, namely the unique row of all 1s in the last $k-1$ columns as generated by $\mu_{r,k-1}$. Thus if we view the columns of A as $(x, y_1, \ldots, y_{k-1})$, the columns $y_1, \ldots, y_{k-1}$ can act as a selector for the first column in that $x|_{S(y_1,\ldots,y_{k-1})} = x_j$ where j is the row of all 1s in $y_1, \ldots, y_{k-1}$.

Define $\Psi : (\{0,1\}^{r\times k})^n \to \mathbb{R}$ by

$$\Psi(X_1, \ldots, X_n) = 2^n \psi(\mathrm{DISJ}_{r,k}(X_1), \ldots, \mathrm{DISJ}_{r,k}(X_n)) \prod_{i=1}^{n} \mu(X_i)$$

for blocks X_i.

Since μ distributes equally onto $\mathrm{UDISJ}_{r,k}^{-1}(-1)$ and $\mathrm{UDISJ}_{r,k}^{-1}(+1)$,

$$\|\Psi\|_1 = 2^n \mathop{\mathbb{E}}_{x\in\{-1,+1\}^n} |\psi(x)| = 1.$$

Similarly,

$$\begin{aligned}\sum_{\operatorname{dom} F} F(X_1,\ldots,X_n)\Psi(X_1,\ldots,X_n) - \sum_{\overline{\operatorname{dom} F}} |\Psi(X_1,\ldots,X_n)| &= \sum_{x\in\operatorname{dom} f} f(x)\psi(x) - \sum_{x\notin\operatorname{dom} f} |\psi(x)| \\ &> \delta.\end{aligned}$$

We now bound $\langle \Psi, \chi\rangle$ for all cylinder intersections χ. Since $\hat{\psi}(S) = 0$ for $|S| < d$,

$$\begin{aligned}|\langle \Psi, \chi\rangle| &\le 2^n \sum_{|S|\ge d} \left|\hat{\psi}(S)\right| \left| \mathop{\mathbb{E}}_{X_1,\ldots,X_n\sim\mu} \chi(X_1,\ldots,X_n) \prod_{i\in S} \mathrm{DISJ}_{r,k}(X_i)\right| \\ &\le 2^n \sum_{|S|\ge d} \left|\hat{\psi}(S)\right| \Gamma_{k-1}(\underbrace{r,\ldots,r}_{|S|}).\end{aligned}$$

Since $\max_{S\subseteq[n]} \left|\hat{\psi}(S)\right| \le 2^{-n}\|\psi\|_1$,

$$|\langle \Psi, \chi\rangle| \le \sum_{|S|\ge d} \Gamma_{k-1}(\underbrace{r,\ldots,r}_{|S|})$$

By Theorem 4.4,

$$|\langle \Psi, \chi\rangle| \le \sum_{|S|\ge d} \left(\frac{2^{k-1}}{\sqrt{r}}\right)^{|S|}.$$

Since $|\langle \Psi, \chi\rangle| \le \|\Psi\|_1 \|\chi\|_\infty = 1$,

$$\begin{aligned}|\langle \Psi, \chi\rangle| &\le \min\left\{1, \sum_{i=d}^{n} \binom{n}{i} \left(\frac{2^{k-1}}{\sqrt{r}}\right)^{i}\right\} \\ &\le \left(\frac{2^k en}{d\sqrt{r}}\right)^d.\end{aligned}$$

Applying the generalized discrepancy method as given by Theorem 4.2 to the bound for $|\langle \Psi, \chi\rangle|$ and the value for $\|\Psi\|_1$,

$$\begin{aligned}2^{R_\epsilon(F)} &\ge \frac{1-\epsilon}{\left(\frac{2^k en}{d\sqrt{r}}\right)^d} \left(\delta - \frac{\epsilon}{1-\epsilon}\right) \\ &= (\delta - \epsilon(1+\delta)) \left(\frac{d\sqrt{r}}{2^k en}\right)^d.\end{aligned}$$

□

7 Direct Product Theorems

Given a computational problem, it is natural to ask how well the cost of solving l instances of the problem scales with l. In certain situations, we can expect that the cost of solving l instances is approximately that of solving each of them individually, i.e. l times the cost of solving one instance of the problem. A result of this form is called a *direct product theorem*.

More formally, in the setting of a k-party communication problem F, for given ϵ, l, m we are interested in finding the randomized communication complexity $R_{\epsilon,m}(F, F, \ldots, F) = R_{\epsilon,m}(F^{(l)})$ of a protocol that, given l instances of the problem, correctly solves at least $l - m$ instances with probability at least $1 - \epsilon$. A result guaranteeing that close to $lR_\epsilon(F)$ complexity is needed for $\epsilon = 1 - 2^{-O(l)}$ is known as a *threshold direct product theorem* (TDPF). If we require $m = 0$, such a result is a *strong direct product theorem* (SDPF). SPDFs tend to be difficult to prove, and often simply not true.

The situation for the 2-party disjointness problem has been fully analyzed. Klauck proved the following SDPT in 2010 [9]:

Theorem 7.1. *For some absolute constant $\alpha > 0$ and every l,*

$$R_{1-2^{-\alpha l},0}(\mathrm{DISJ}_{n,2}^{(l)}) = l\Omega(n).$$

Klauck's proof reduces the problem $\overline{\mathrm{DISJ}}_{n,2}^{(k)}$ to the problem $\mathrm{SEARCH}_{\binom{N}{k}}(x, y)$ of finding k indices in which strings x, y of length N match and then uses linear programming duality to show a lower bound for this problem.

For the general k-party case, Sherstov proved a TDPF of unique disjointness in [13]:

Theorem 7.2. *For some absolute constant $\alpha > 0$ and every l,*

$$R_{1-2^{-\alpha l},\alpha l}(\mathrm{UDISJ}_{n,k}^{(l)}) = l\Omega\left(\frac{n}{4^k}\right)^{1/4}.$$

The proof of Sherstov's result echoes the primal proof of Theorem 1.1, converting an efficient communication protocol into a low-degree polynomial approximation. Specifically, we will give polynomials approximating the probability of a given output string $z = (z_1, \ldots, z_l) \in \{-1, +1\}^l$ being produced from each input. To this end we define the notion of approximants:

Definition 7.3. A (σ, m, l)-*approximant* for a partial Boolean function f is a system of real-valued functions $\{\phi_z\}$ on X^l, indexed by $z \in \{-1, +1\}^l$, such that we have

$$\sum_{z\in\{-1,+1\}^l} |\phi_z(x^1, \ldots, x^l)| \le 1 \; \forall x^1, \ldots, x^l \in X,$$

$$\sum_{\substack{z\in\{-1,+1\}^l \\ |\{i:z_i=-1\}|\le m}} |\phi_{(z_1 f(x^1),\ldots,z_l f(x^l))}(x^1, \ldots, x^l)| \ge \sigma \; \forall x^1, \ldots, x^l \in \operatorname{dom} f.$$

In this definition, ϕ_z is intuitively seen to represent the probability that f outputs the string z given an input $(x^1, \ldots, x^l)$. Interpreting a (σ, m, l)-approximant this way means that any given input has a probability at least σ of getting at least $l - m$ of the correct outputs.

Sherstov states the relevant polynomial approximation theorem:

Theorem 7.4. *For any sufficiently small $\beta > 0$, every $(2^{-\beta l}, \beta l, l)$-approximant $\{\phi_z\}$ of a partial Boolean function f on $\{-1, +1\}^n$ satisfies*

$$\max_{z\in\{-1,+1\}^l} \{\deg \phi_z\} \ge \beta l \deg_{1/3}(f).$$

Sherstov proves this theorem in [14]. The proof applies Theorem 3.7 to f and then proceeds with bounding in a manner resembling the generalized discrepancy method.

Using this and the proof method of the primal proof of Theorem 6.1, we can prove the following direct product theorem about compositions with unique disjointness:

Theorem 7.5. *Fix a sufficiently small $\alpha > 0$. Let f be a partial Boolean function on $\{-1,+1\}^n$, let $d = \deg_\delta(f)$, let $r = 4^k\lceil \frac{n}{\alpha d}\rceil^2$, and let $F = f \circ \mathrm{UDISJ}_{r,k}$. Then*

$$R_{1-2^{-\alpha l},\alpha l}(F^{(l)}) \geq \alpha l d.$$

Proof. We proceed as in the primal proof of Theorem 6.1. Define μ, μ_{+1}, μ_{-1} as there. This time we define the averaging operator M as follows: given a function $\phi : (\{0,1\}^{r\times k})^{ln}$, for any $x = (x_{1,1}, \cdots, x_{l,n}) \in \{-1,+1\}^{ln}$ let

$$(M\phi)(x) = \mathop{\mathbb{E}}_{X_{1,1}\sim\mu_{x_{1,1}},\ldots,X_{l,n}\sim\mu_{x_{l,n}}}[\phi(X_{1,1},\ldots,X_{l,n})].$$

If $\{\phi_z\}$ is a (σ, m, l)-approximant for F, we then have, letting $x^i = (x_{i,1},\ldots,x_{i,n})$ for $1 \leq i \leq l$,

$$\sum_{z\in\{-1,+1\}^l} |(M\phi_z)(x^1,\ldots,x^l)| = \sum_{z\in\{-1,+1\}^l} \left| \mathop{\mathbb{E}}_{X_{1,1}\sim\mu_{x_{1,1}},\ldots,X_{l,n}\sim\mu_{x_{l,n}}}[\phi_z(X_{1,1},\ldots,X_{l,n})]\right|$$

$$\leq \mathop{\mathbb{E}}_{X_{1,1}\sim\mu_{x_{1,1}},\ldots,X_{l,n}\sim\mu_{x_{l,n}}} \sum_{z\in\{-1,+1\}^l} |\phi_z(X_{1,1},\ldots,X_{l,n})|$$

$$\leq \mathop{\mathbb{E}}_{X_{1,1}\sim\mu_{x_{1,1}},\ldots,X_{l,n}\sim\mu_{x_{l,n}}} 1 = 1.$$

Similarly, letting $X^i = (X_{i,1},\ldots,X_{i,n})$ for $1 \leq i \leq l$, for $(x^1,\ldots,x^l) \in \operatorname{dom} f$ we have

$$\sum_{\substack{z\in\{-1,+1\}^l\\ |\{i:z_i=-1\}|\leq m}} (M\phi_{(z_1 f(x^1),\ldots,z_l f(x^l))})(x^1,\ldots,x^l)$$

$$= \mathop{\mathbb{E}}_{X_{1,1}\sim\mu_{x_{1,1}},\ldots,X_{l,n}\sim\mu_{x_{l,n}}} \sum_{\substack{z\in\{-1,+1\}^l\\ |\{i:z_i=-1\}|\leq m}} \phi_{(z_1 f(x^1),\ldots,z_l f(x^l))}(X^1,\ldots,X^l)$$

$$= \mathop{\mathbb{E}}_{X_{1,1}\sim\mu_{x_{1,1}},\ldots,X_{l,n}\sim\mu_{x_{l,n}}} \sum_{\substack{z\in\{-1,+1\}^l\\ |\{i:z_i=-1\}|\leq m}} \phi_{(z_1 F(X^1),\ldots,z_l F(X^l))}(X^1,\ldots,X^l) \geq \sigma,$$

where we have used the fact that $F(X^i) = f(D(X_{i,1}),\ldots,D(X_{i,n})) = f(x)$ when $(X_{i,1},\ldots,X_{i,n}) \sim \mu_{x_{i,1}} \times \cdots \times \mu_{x_{i,n}}$. So, $\{M\phi_z\}$ is a (σ, m, l)-approximant for f.

Suppose we have a randomized protocol Π that attains $c = R_{1-2^{-\alpha l},\alpha l}(F^{(l)})$, where α is a constant we will pick later. For any $z \in \{-1,+1\}^l$, we can define a protocol Π_z that runs Π and accepts if and only if Π outputs z. Letting $\phi_z(X^1,\ldots,X^l)$ give the probability that Π_z accepts on input $(X^1,\ldots,X^l)$, we see that ϕ_z is a $(2^{-\alpha l}, \alpha l, l)$-approximant for F. Since Π_z has cost c, by Corollary 3.4 we have $\phi_z = \sum a_\chi \chi$ for some k-dimensional cylinder intersections χ with $\sum |a_\chi| \leq 2^c$.

So, applying the bounding argument in the primal proof of Theorem 6.1 to $M\phi_z$, we get for any D that

$$E(M\phi_z, D-1) \leq 2^c \left(\frac{2^k eln}{D\sqrt{r}}\right)^D.$$

Assume for the sake of contradiction that $c < \alpha ld$. Setting $D = \beta ld$ then gives

$$E(M\phi_z, D-1) \leq 2^{\alpha ld} \left(\frac{2^k eln}{\beta ld 2^k \lceil \frac{n}{\alpha d} \rceil} \right)^{\beta ld} \leq 2^{\alpha ld} (\alpha e/\beta)^{\beta ld}.$$

So there is a system of polynomials $\{p_z\}$ with degree less than βld satisfying $\|M\phi_z - p_z\|_\infty \leq 2^{\alpha ld} (\alpha e/\beta)^{\beta ld}$ for each z. Taking α sufficiently small compared to β makes $\{p_z - 2^{\alpha ld}(\alpha e/\beta)^{\beta ld}\}$ satisfy the necessary properties for a $(2^{-\beta ld}, \beta l, l)$-approximant for f. But $\deg(p_z) < \beta ld$ for all z, contradicting Theorem 7.4. So in fact we have $c \geq \alpha ld$, as desired.

□

Theorem 7.2 follows as a corollary:

Proof of Theorem 7.2. This follows from Theorem 7.5 in the same way that Theorem 1.1 follows from Theorem 6.1. Specifically, since $\deg_{1/3}(\widetilde{\text{AND}}_n) \geq c\sqrt{n}$ for some c, letting $f = \widetilde{\text{AND}}_n$ and $d = c\sqrt{n}$ in Theorem 7.5 gives a bound equivalent to the result we want. □

8 Relation to the Generalized Inner Product

In the introduction, we defined $G^g_{n,k}$ as a function g on the bits where all players receive a 1. We noted that $\text{DISJ}_{n,k} = G^{\text{NOR}}_{n,k}$ and that the generalized inner product $\text{GIP}_{n,k} = G^{\text{PARITY}}_{n,k}$. Both of these are derivatives on the number of shared 1-bits among all players, with GIP being the result mod 2 and DISJ being whether the result is 0. Despite their similarities in form, it has been much easier to prove stronger bounds for GIP than DISJ.

In 1992, Babai, Nisan, and Szegedy [1] showed that $R_{1/3}(\text{GIP}_{n,k}) = \Omega(\frac{n}{4^k})$. This was soon shown to be close to optimal, as the original proof by Grolmusz [7] in 1994 showed $D(\text{GIP}_{n,k}) = O(\frac{kn}{2^k})$. As we showed in Section 2, that protocol can be extended to a $O(\frac{k^2 n}{2^k})$ protocol for DISJ (or any $G^g_{n,k}$ where g depends only on the number of 1s).

Before [13], the best lower bounds for DISJ for $k \geq 3$ players were weaker than $\Omega((\frac{n}{2^{k^3}})^{1/(k+1)})$ and [13] improved this only to $\Omega((\frac{n}{4^k})^{1/4})$. The best lower bound on randomized complexity currently available is $\Omega(\frac{\sqrt{n}}{2^k k})$. Thus twenty years after [7] showed GIP is $\Theta(n)$ for constant k, we still have an $\Omega(n^{1/2})$, $O(n)$ gap for DISJ.

Intuitively, this difference in difficulty in placing lower bounds comes from the fact that the value of PARITY depends on all of its inputs while on all but the 0 input, NOR can depend on a single bit. Informally, this makes PARITY harder to approximate.

In fact, Sherstov [13] in achieving the bound for DISJ uses the hardness of PARITY, as we saw in Section 4 in the definition of $\Gamma = \text{disc}(\bigoplus \text{UDISJ})$.

9 Discussion

It is natural to ask how far the methods used in proving lower bounds so far can take us. To break the longstanding barrier and get the first $\Omega(n^c/2^{O(k)})$ bound on the randomized complexity of $\text{DISJ}_{n,k}$ for a constant c, Sherstov [13] needed to modify the usual approach of applying the generalized discrepancy method and the pattern matrix method, forming bounds on the individual terms of the Fourier decomposition, and exploiting properties like conditional independence in the probability distributions. To improve this bound to $\Omega(n^{1/2}/2^{O(k)})$, Sherstov [15] again had to introduce new techniques, this time including more analytic tools like directional derivatives.

It remains open whether $R_{1/3}(\text{DISJ}_{n,k}) \geq \Omega(n/2^{O(k)})$, and whether such a bound can be proven using the kinds of approaches that have been used so far like using cylinder intersections, discrepancies, and polynomial approximations.

In addition, although essentially tight bounds have been proven for the deterministic and quantum communication complexity of $\mathrm{UDISJ}_{n,k}$ using adaptations of Sherstov's methods in [13], there are still many other protocol settings to consider for which the sharp bounds have not yet been determined. For example, nondeterministic and Merlin-Arthur protocols are two types that are addressed in [13] and [15], but for which polynomial gaps between upper and lower bounds still remain.

Finally, results like Theorem 6.1 allow us to generalize our lower bounds to compositions of disjointness with other functions f, assuming corresponding polynomial lower bounds on f. In fact, Sherstov [13] also shows that similar bounds on compositions hold with $\mathrm{UDISJ}_{n,k}$ replaced by different expressions like $\mathrm{OR}_k \vee \mathrm{AND}_k$. It is open to determine how much these results can be generalized – how complicated the functions we compose with in place of $\mathrm{UDISJ}_{n,k}$ can get – while still allowing us to prove strong lower bounds on randomized communication complexity using an essentially similar method.

References

[1] Babai, L., Nisan, N., and Szegedy, M. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. Syst. Sci. 45*, 2 (Oct. 1992), 204–232.

[2] Beame, P., and Huynh-Ngoc, D.-T. Multiparty communication complexity and threshold circuit size of AC^0. In *Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science* (Washington, DC, USA, 2009), FOCS '09, IEEE Computer Society, pp. 53–62.

[3] Beame, P., Pitassi, T., Segerlind, N., and Wigderson, A. A strong direct product theorem for corruption and the multiparty communication complexity of disjointness. *Comput. Complex. 15*, 4 (Dec. 2006), 391–432.

[4] Chandra, A. K., Furst, M. L., and Lipton, R. J. Multi-party protocols. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing* (New York, NY, USA, 1983), STOC '83, ACM, pp. 94–99.

[5] Chattopadhyay, A., and Ada, A. Multiparty communication complexity of disjointness. *CoRR* (2008).

[6] Chor, B., and Goldreich, O. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM J. Comput. 17*, 2 (Apr. 1988), 230–261.

[7] Grolmusz, V. The bns lower-bound for multiparty protocols is nearly optimal. *Inf. Comput. 112*, 1 (July 1994), 51–54.

[8] Kalyanasundaram, B., and Schintger, G. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics 5*, 4 (1992), 545–557.

[9] Klauck, H. A strong direct product theorem for disjointness. In *Proceedings of the forty-second ACM symposium on Theory of computing* (2010), ACM, pp. 77–86.

[10] Lee, T., and Shraibman, A. Disjointness is hard in the multi-party number on the forehead model. *CoRR* (2007).

[11] Nisan, N., and Szegedy, M. On the degree of boolean functions as real polynomials. In *Proceedings of the Twenty-fourth Annual ACM Symposium on Theory of Computing* (New York, NY, USA, 1992), STOC '92, ACM, pp. 462–467.

[12] Sherstov, A. A. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the Fortieth Annual ACM Symposium on Theory of Computing* (New York, NY, USA, 2008), STOC '08, ACM, pp. 85–94.

[13] Sherstov, A. A. The multiparty communication complexity of set disjointness. In *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing* (New York, NY, USA, 2012), STOC '12, ACM, pp. 525–548.

[14] Sherstov, A. A. Strong direct product theorems for quantum communication and query complexity. *SIAM Journal on Computing 41*, 5 (2012), 1122–1165.

[15] Sherstov, A. A. Communication lower bounds using directional derivatives. *Journal of the ACM (JACM) 61*, 6 (2014), 34.

[16] Tesson, P. *Computational Complexity Questions Related to Finite Monoids and Semigroups.* PhD thesis, Montreal, Que., Canada, Canada, 2003.

18.405J / 6.841J Advanced Complexity Theory
Spring 2016