The Relationship between Inner Product and Counting Cycles

Xiaoming Sun
ITCS, IIIS, Tsinghua
xiaomings@tsinghua.edu.cn

Chengu Wang
ITCS, IIIS, Tsinghua
wangchengu@gmail.com

Wei Yu
ITCS, IIIS, Tsinghua
zig.wei@gmail.com

Abstract

CYCLE-COUNTING is the following communication complexity problem: Alice and Bob each holds a permutation of size n with the promise that there will be either a cycles or b cycles in the product of these two permutations. They want to distinguish the two cases. We show a quantum/nondeterministic lower bound of $\Omega(\frac{n-a}{b-a}\log p(b-a+1) - \log(b-a+1))$ when $a \equiv b \pmod 2$, where $p(b-a+1)$ is the smallest prime factor of $b-a+1$. The reduction we are using here is simple but elegant, constructing a bridge for various problems, including IN-SAME-CYCLE introduced by Harvey [9], ONE-CYCLE introduced by Raz and Spiekler [15], and BIPARTITENESS on constant degree graph introduced by Hajnal et al.[8]. We also give the space lower bounds in the streaming model for the CONNECTIVITY, BIPARTITENESS and GIRTH problems discussed by Feigenbaum et al. [6].

We prove them by reduction from a variant of the inner product problem over $\mathbb{Z}_m$, which has a quantum lower bound of $\Omega(n \log p(m))$. It implies that our lower bound for cycle counting and related problems have same quantum communication lower bounds, which was not known before this paper. Our technique of using multicolor discrepancy to bound binary discrepancy, might also be useful for lower bounding other promised functions defined over the finite field.

1 Introduction

The model of communication complexity was first introduced by Yao [20], and then was studied extensively. The communication complexity model deals with the following game between Alice and Bob. Given a function $f : X \times Y \mapsto Z$, Alice holds $x \in X$, and Bob holds $y \in Y$. They will follow a protocol to let both of them know the value of $f(x,y)$ by sending and receiving bits from each other. We call the least number of bits transmitted in the protocol $D(f)$, the deterministic communication complexity for computing f.

The model could also be extended to the case with randomization. In randomized communication complexity, Alice and Bob have shared random coins, and they can send and receive message by using these random coins. At the end of the communication, Alice and Bob will decide an output for the protocol, and we call this $P(x,y)$. We say P is a randomized protocol of f with error ϵ if for any input (x,y), $\Pr[P(x,y) = f(x,y)] \geq 1-\epsilon$. We call the least number of bits transmitted for the worst input (x,y) and the best protocol P $R_\epsilon(f)$. We can also investigate nondeterministic protocols, where there exists a powerful man who wants to convince Alice and Bob the answer. For $z \in \{0,1\}$, we define $N^z(f)$ to be the amount of communication to convince Alice and Bob $f(x,y) = z$, including both the proof and the bits exchanged by Alice and Bob in order to verify the proof in the most efficient proof system. Since a deterministic protocol is both a randomized protocol and a nondeterministic protocol, we have $R_{1/3}(f) \leq D(f)$, $N^0(f) \leq D(f)$, and $N^1(f) \leq D(f)$. For comprehensive explanations on communication complexity, we refer the reader to [14].

The key problem we are going to talk about in this paper is the CYCLE-COUNTING problem introduced in [19]. The problem could be stated as Alice and Bob each holds a permutation, and they want to decide the number of cycles in the product of the permutations, given the promise on the input that there are either a cycles or b cycles in the product permutation.

There are other problems related to the Cycle-Counting problem. For example, we show that a nondeterministic/randomized lower bound of $\Omega(n)$ for the In-Same-Cycle problem from [9] could be obtained by a reduction from a special instance of the cycle counting problem (say, separating 1 cycle and 3 cycles). The lower bound nearly matches the lower bound in [9] up to a multiplicative constant, but conceptually easier to get. Furthermore, the same lower bound of $\Omega(n)$ could be obtained for the One-Cycle problem and the Bipartiteness problem as well. The One-Cycle problem is to decide if the product of two permutations is one cycle or more than one cycle. It was used by Raz and Spieker [15] to show a separation between log-rank and nondeterministic lower bound, by showing a nondeterministic lower bound of $\Omega(n \log\log n)$. Our lower bound is only $\Omega(n)$ but the proof is much easier. The Bipartiteness problem is to decide if a graph split into Alice and Bob's hand is bipartite or not. A deterministic bound of $\Theta(n \log n)$ was proved for general graphs by Hajnal et al. [8]. Here we show that even for graphs of maximum degree 3, a lower bound of $\Omega(n)$ could also be proved for nondeterministic/randomized protocols.

Besides communication complexity, we consider the streaming model as well. In streaming model, the input of a graph is represented by a sequence of edges in arbitrary order. The streaming complexity is the minimal amount of memory used by the algorithm if the algorithm only reads the input once sequentially. A lot of graph properties are studied in the streaming model. For example, Bar-Yossef et al. [4] counted triangles in a graph; Feigenbaum et al. [7] gave approximation algorithms for matching, diameter and distance problems; and Feigenbaum et al. [6] discussed lots of graph properties including connectivity, bipartiteness, diameter and girth. For every problem discussed in this paper, the lower bound of the communication complexity implies the same lower bounds on the streaming complexity, because Alice can run the streaming algorithm on the first half on input and send the configuration of the machine to Bob, after that Bob can continue the execution on the second half and output the result. Our lower bound of approximating the girth in the streaming model improves the result in [6] when the girth is large. And we also prove the linear lower bound again for the connectivity and bipartiteness problems.

The lower bound for Cycle-Counting is obtained by reduction from a variant of the inner product modulo m problem. The problem could be briefly described as computing the inner product modulo m of two vectors in $\mathbb{Z}_m^n$, where Alice holds one of them, and Bob holds the other. The $m = 2$ case for this problem is well studied, and a lower bound of $\Omega(n)$ is known [14]. We are here to show a $\Omega(n \log p(m))$ nondeterministic/randomized lower bound for general m, where $p(m)$ is the smallest prime factor of m. And this bound is tight for the case when m is prime ($p(m) = m$). We will use the discrepancy method to prove its randomized lower bound, and by investigating the relationship between discrepancy and largest monochromatic rectangles, the nondeterministic lower bound could also be obtained in the same way.

Furthermore, we know that the discrepancy method could also imply quantum communication complexity lower bounds [13]. In quantum settings, Alice and Bob have quantum computers and infinite shared entangled pairs of qubits, and they want to compute the function f with error ϵ by exchanging quantum bits. And we denote the quantum communication complexity of f (the minimum amount of qubits exchanged) by $Q_\epsilon^*(f)$. Since we can use quantum bits to generate random bits, $Q_{1/3}^*(f) \leq O(R_{1/3}(f))$ [13] and $R_{1/3}(f) \leq D(f)$, which means that we can get randomized/deterministic lower bounds by quantum lower bounds. Thus in the rest of the paper, we will only talk about quantum and nondeterministic lower bounds.

1.1 Results

In this section, we formally define all the problems, and state all the theorems only in the communication complexity model although each lower bound implies a same result in the streaming model. The central problem is the following Cycle-Counting problem.

Definition 1 ($\mathrm{CC}_{n,a,b}$). *Let π, σ be permutations in symmetric group S_n with the promise that $\sigma \circ \pi$ has either a cycles or b cycles ($a < b$). The* Cycle-Counting *problem is a communication* *complexity problem that Alice holds π and Bob holds σ, and they want to return* 0 *for a cycles case or return* 1 *for b cycles case.*

We prove the following lower bound for $\mathrm{CC}_{n,a,b}$. It is almost tight (up to a $\log n$ factor, see Lemma 26) because of the upper bound for $\mathrm{CC}_{n,1,m}$.

Theorem 2. *The quantum/nondeterministic lower bound of* $\mathrm{CC}_{n,a,b}$ *is* $\Omega(\frac{n-a}{b-a}\log p(b-a+1)-\log(b-a+1))$ *when* $a \equiv b \pmod 2$*, where* $p(b-a+1)$ *is the smallest prime factor of* $b-a+1$.

Since the length of cycles are all the same in the hard case of $\mathrm{CC}_{n,1,m}$, to distinguish 1 cycles and m cycles is as hard as the to distinguish girth n and girth n/m.

Corollary 3. $\Omega(\frac{n}{m}\log p(m))$ *communication is needed to determine whether the girth of a graph G is either n of n/m for quantum/nondeterministic protocols, if the edges of G is distributed to Alice and Bob, and m is odd.*

The streaming version of Corollary 3 improves the result in [6] when $m = O(n^{1/2-\epsilon})$.

Then, we show a similar lower bound holds for the In-Same-Cycle problem defined in [9].

Definition 4 (In-Same-Cycle). *Let π, σ be permutations in symmetric group S_n.* In-Same-Cycle$_n$ *is a communication complexity problem that Alice holds π and Bob holds σ, and they want to return* 1 *if elements* 1 *and* 2 *are in the same cycle of $\sigma \circ \pi$, or return* 0 *otherwise.*

As stated by Harvey [9], the In-Same-Cycle problem is a special case of the matroid intersection problem (abbr. MAT$-\cap$). So our lower bound holds for MAT$-\cap$ as well. Note that in [9] only nondeterministic lower bounds were discussed, here we talk about quantum lower bound as well. We can show that by an easy argument that In-Same-Cycle is also hard in our hard case for $\mathrm{CC}_{n,1,3}$, thus we have the following corollary.

Corollary 5. *The quantum/nondeterministic lower bound of* In-Same-Cycle *is* $\Omega(n)$.

We also show the same lower bound holds for the following two problems, where the former was defined in [15] and the latter was defined in [8].

Definition 6 (One-Cycle). *Let π, σ be permutations in symmetric group S_n.* One-Cycle$_n$ *is a communication complexity problem that Alice holds π and Bob holds σ, and they want to return* 1 *if $\sigma \circ \pi$ is a Hamiltonian cycle, or return* 0 *otherwise.*

Definition 7 (Bipartiteness). *Let $G_A = \langle V, E_A\rangle, G_B = \langle V, E_B\rangle$ be two graphs on the same vertex set.* Bipartiteness *is a communication complexity problem that Alice holds G_A and Bob holds G_B, and they want to return* 1 *if $G_A \cup G_B = \langle V, E_A \cup E_B\rangle$ is a bipartite graph, or return* 0 *otherwise.*

We show the hard case for $\mathrm{CC}_{n,1,3}$ is also a hard case for both of them, implying the following quantum/nondetereministic lower bound. A similar argument could be proved for the Bipartiteness problem.

Corollary 8. *The quantum/nondeterministic lower bound of* Bipartiteness *is* $\Omega(n)$ *even for graphs with maximum degree* 3*, and the quantum/nondeterministic for* One-Cycle *is* $\Omega(n)$.

Unlike the previous proof in [9] which directly investigated the properties of the cycle counting type problem, we prove the lower bound of $\mathrm{CC}_{n,a,b}$ by reducing from the Inner Product Modular m problem, which is defined below.

Definition 9 ($\mathrm{IP}_{m,n}$, $\mathrm{IP}^{01}_{m,n}$ and $\mathrm{IP}^{01*}_{m,n}$). *The inner product problem ($\mathrm{IP}_{m,n}$) is a communication complexity problem that Alice holds $x \in \mathbb{Z}_m^n$ and Bob holds $y \in \mathbb{Z}_m^n$, and they want to return the value of the inner product $\langle x, y\rangle = \sum_{i=1}^n x_i y_i \mod m$.*

In the reduction we need two promised variants of $\mathrm{IP}_{m,n}$: $\mathrm{IP}^{01}_{m,n}$ is the $\mathrm{IP}_{m,n}$ problem with the promise that $\langle x, y\rangle$ is either 0 or 1; and $\mathrm{IP}^{01*}_{m,n}$ is the $\mathrm{IP}^{01}_{m,n}$ problem with the promise that $y \in (\mathbb{Z}_m^*)^n$, where $\mathbb{Z}_m^*$ the the primitive residue class modulo m (the set of integers relatively prime to m).

The Inner Product problem on the binary field ($m = 2$ case) is well studied. It is known that $Q^*_{1/3}(\mathrm{IP}_{2,n}) = \Omega(n)$ [13], and $D(\mathrm{IP}_{p,n}) = \Omega(n \log p)$ for prime p [5], where $\mathrm{IP}_{p,n}(x, y) = \sum_{i=1}^{n} x_i y_i \mod p$. However, what we actually need for this paper is the $\mathrm{IP}^{01*}_{m,n}$ problem. The problem looks classic but the authors of the paper failed to find a reference for the lower bound. So the proof for the following theorem is claimed in the paper to be "new" with conservation.

Theorem 10. *The quantum/nondeterministic lower bound of* $\mathrm{IP}^{01}_{m,n}$ *is* $\Omega(n \log p(m) - \log(m))$, *and the lower bound of* $\mathrm{IP}^{01*}_{m,n}$ *is* $\Omega(n \log(p(m) - 1) - \log m)$, *where* $p(m)$ *the smallest prime factor of* m.

Since $\mathrm{IP}^{01}_{m,n}$ is a special case of $\mathrm{IP}_{m,n}$, so the lower bound of $\mathrm{IP}^{01}_{m,n}$ also holds for $\mathrm{IP}_{m,n}$.

Reduction from Set-Disjointess [3, 12, 16] is a powerful tool to prove a strong communication lower bound in the random world. However, Q(Set-Disjointess) is much lower for quantum protocols. We hope that the strong quantum lower bounds of more problems can be proved by reduction from Theorem 10.

2 The Cycle Counting Problem and Its Variants

In this section we show the reduction from the inner product problem to the cycle counting problem, and its variants.

2.1 The Cycle Counting Problem

We are going to prove the following theorem in this subsection.

Theorem 11 (Theorem 2 Restated). *Let* $p(x)$ *denote the smallest prime factor of* x, *the following statements hold for the communication complexity of* Cycle-Counting,

1. $Q^*_{1/3}(\mathrm{CC}_{n,1,m}) = \Omega(n/m \cdot \log p(m) - \log m)$, *if* m *is odd;*
2. $Q^*_{1/3}(\mathrm{CC}_{n,a,b}) = \Omega((n-a)/(b-a) \cdot \log p(b-a+1) - \log(b-a+1))$, *if* $a \equiv b \pmod 2$;
3. $D(\mathrm{CC}_{n,a,b}) = 1$, *if* $a \not\equiv b \pmod 2$;
4. $R_{1/3}(\mathrm{CC}_{n,1,m}) = \min\{O(n \log n), O(n/m \cdot \log n \cdot \log(n/m))\}$.

Proof. We here use the reduction from $\mathrm{IP}^{01*}_{m,n}$ to $\mathrm{CC}_{m(n+1),1,m}$ to show the lower bound.

Let (x, y) be an input of the $\mathrm{IP}^{01*}_{m,n}$ problem where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$, $x_i, y_i \in \mathbb{Z}_m$ for $i \in [n]$. According to the definition of $\mathrm{IP}^{01*}_{m,n}$, we have y_i is relative prime to m for $i \in [n]$. Thus by Euclid algorithm we know that there is a y_i^{-1} exists for each y_i such that $y_i y_i^{-1} \equiv 1 \pmod m$. Let $y' = (y'_0, y'_1, \ldots, y'_{n-1}, y'_n) = (y_1^{-1}, y_1 y_2^{-1}, \ldots, y_{n-1} y_n^{-1}, y_n)$ and $x' = (x'_0, x'_1, \ldots, x'_n) = (0, x_1, \ldots, x_n)$.

We are going to construct a bipartite (black vertices on one side and white ones on another) graph $G = \langle V, E \rangle$ as shown in Figure 1, where $V = \{v_{i,j} | 0 \le i \le 2n+1, 0 \le j \le m-1\}$. Alice holds the edges from black vertices to white vertices, and Bob holds the edges from white vertices to black vertices. That is, the edge set Alice holds is $\{(v_{2i,j}, v_{2i+1,(j+x'_i) \mod m})\}$, and the edge set Bob holds is $\{(v_{2i+1,j}, v_{(2i+2) \mod (2n+2),(j \times y_i) \mod m})\}$. Each row represents an element of $\mathbb{Z}_m$. The in-degree and out-degree of each vertex are both exactly 1, thus this bipartite graph is a union of two permutations.

Imagining that we traverse the graph starting from vertex $v_{0,t}$, we will reach the 0-th layer again after following $2(n+1)$ edges. And the row reached will be

$$\begin{aligned}
&(((((t + x'_0) \times y'_0) + x'_1) \times y'_1) + ... + x'_n) \times y'_n \mod m \\
&= (y'_0 y'_1 ... y'_n t + x'_0 y'_0 y'_1 ... y'_n + x'_1 y'_1 y'_2 ... y'_n + ... + x'_n y'_n) \mod m \\
&= (t + x_1 y_1 + x_2 y_2 + ... + x_n y_n) \mod m.
\end{aligned}$$

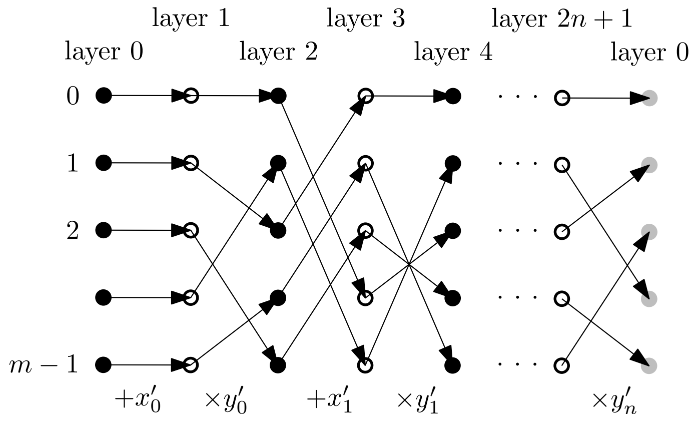

Figure 1: The construction of a $\mathrm{CC}_{m(n+1),1,m}$ instance from an $\mathrm{IP}^{01*}_{n,m}$ instance. In this example, $m = 5$, $x'_0 = 0$, $y'_0 = 2$, $x'_1 = 3$, $y'_1 = 4$ and $y'_n = 3$. The gray vertices in the last layer are the shadows of the first layer. The graph is actually undirected, and the direction only helps for presentation.

Since $x_1y_1+x_2y_2+...+x_ny_n$ is promised to be 0 or 1 modulo m, we know that we will reach either $v_{0,t}$ or $v_{0,(t+1) \bmod m}$. That is, there will be either m cycles or a single cycle. By distinguishing these two cases, we can know the answer for $\mathrm{IP}^{01*}_{m,n}$, so $Q^*_{1/3}(\mathrm{CC}_{n,1,m}) \geq Q^*_{1/3}(\mathrm{IP}^{01*}_{n/(m-1),m}) = \Omega(n/m)$.

Having proved a lower bound for $\mathrm{CC}_{n,1,m}$, we reduce $\mathrm{CC}_{n,1,b-a+1}$ to $\mathrm{CC}_{n+a-1,a,b}$.

Let (π, σ) be an input of the $\mathrm{CC}_{n,1,b-a+1}$ problem. We are going to construct an input (π', σ') for $\mathrm{CC}_{n+a-1,a,b}$. We define $\pi', \sigma' \in S_{n+a-1}$, s.t. $\pi'(i) = \pi(i)$ and $\sigma'(i) = \sigma(i)$ for $1 \leq i \leq n$, and $\pi'(i) = i$ and $\sigma'(i) = i$ for $n+1 \leq i \leq n+a-1$. So permutation $\sigma' \circ \pi'$ is just $\sigma \circ \pi$ after adding $a-1$ small cycle. So in $\sigma' \circ \pi'$, there are either a or b cycles. Precisely speaking, $\sigma \circ \pi$ has 1 cycle if $\sigma' \circ \pi'$ has a cycles, and $\sigma \circ \pi$ has $b-a+1$ cycles if $\sigma' \circ \pi'$ has b cycles.

Thus, $Q^*_{1/3}(\mathrm{CC}_{n,a,b}) \geq Q^*_{1/3}(\mathrm{CC}_{n-a+1,1,b-a+1}) = \Omega((n-a)/(b-a) \cdot \log p(b-a+1) - \log(b-a+1))$.

We leave the proof for the upper bounds in Appendix A. □

2.2 Other Variants

For the In-Same-Cycle problem and the One-Cycle problem, one can easily observe that the reduction we used to get a lower bound of $\mathrm{CC}_{n,1,3}$ is also a reduction for both In-Same-Cycle and One-Cycle.

For the Bipartiteness problem, the proof is almost the same as the proof of the lower bound of $\text{Cycle-Counting}_{n,1,3}$, but we add an edge between $(0,0)$ and $(0,1)$ (the bold edge in the Figure 2). We know that a graph is bipartite if and only if there are no odd cycles in the graph. If the inner product is 0, the graph has of 3 even cycles, and the bold edge does not contribute to Bipartiteness. If the inner product is 1, after walking $2(n+1)$ steps from $(0,0)$ we reach $(0,1)$, then we go back to $(0,0)$ by the bold edge, so it contains an odd cycle of length $2n+3$, which means the graph is not bipartite.

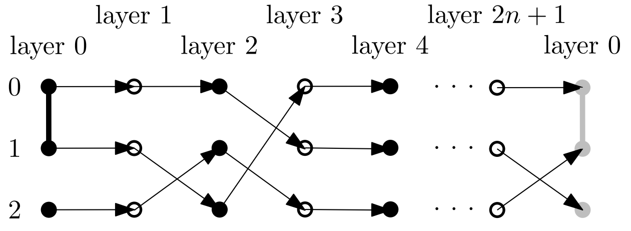

Figure 2: Reduction from Inner Product to Bipartiteness

Therefore, the quantum/nondeterministic communication complexities of In-Same-Cycle, Bipartiteness and One-Cycle are all $\Omega(n)$. Thus Corollary 5 and 8 follow.

3 The Lower Bound for Inner Product over $\mathbb{Z}_m$

In this section we prove an $\Omega(n \log p(m))$ lower bound for $\mathrm{IP}^{01}_{m,n}$, and an $\Omega(n \log(p(m) - 1))$ for $\mathrm{IP}^{01*}_{m,n}$. The main idea of the proof is to give an upper bound on the discrepancy of the two problems. This could be done by first upper bounding the discrepancy by the sum of the norms of several matrices formed by applying characters of $\mathbb{Z}_m$ on the communication matrix. Then, we show that the norm of these matrices could be computed by hand in a nice form, thus implying a communication lower bound by the relation between discrepancy of the communication matrix and quantum communication complexity. We also show that, by the relation of largest monochromatic rectangle and discrepancy, we can have the same bound for nondeterministic communication complexity. The basics of representation theory and matrix analysis are used in this section.

It is worth noting that we use "excess count", a quantity used in multi-color discrepancy, to bound the binary discrepancy here. And we just used the idea, but not the multi-color discrepancy itself. The reason we use this "excess count" but not to bound binary discrepancy directly is because the distribution we use here is not uniform on the result, but uniform on each each non-star entry in the $\mathrm{IP}^{01}_{m,n}$ problem (i.e. the numbers of 0's and 1's in the communication matrix are not the same), thus the binary discrepancy is hard to compute without the help of this quantity. In other words, we are proposing here a hard distribution and an easy way to compute discrepancy under this very distribution for the promised problem $\mathrm{IP}^{01}_{m,n}$.

3.1 Preliminaries

Notations. In the next subsections, we denote the multiplicative group of nonzero complex numbers by $\mathbb{C}^\times$. G is always a finite Abelian group (e.g. $\mathbb{Z}_m$). We denote $G_{X\times Y}$ (or $\mathbb{C}_{X\times Y}$) the set of matrices on G (or $\mathbb{C}$) coordinated by $X \times Y$. We use $\langle x, y\rangle$ to denote the inner product over $\mathbb{Z}_m$ for $x, y \in \mathbb{Z}_m^n$.

Group and Representation Theory. We define a *character* of G to be a homomorphism $\chi : G \to \mathbb{C}^\times$. Thus we know that for $a, b \in G$, $\chi(a + b) = \chi(a)\chi(b)$. And clearly that $\chi(a)^m = \chi(ma) = \chi(0) = 1$. So the values of χ are the m-th roots of unity. In particular, if $G = \mathbb{Z}_m$, we have $\chi_i(a) = e^{\frac{2\pi i}{m}\cdot a}$ for $0 \le i < m$ are the characters of $\mathbb{Z}_m$. The *principle character* χ_0 of G is the character such that $\forall a, \chi_0(a) = 1$.

The following properties could be found on any algebra book, or in particular in [1].

Proposition 12. *The following properties holds for Abelian group G of order m:*

1. *All the characters of G form a group $\hat{G}$, and $\hat{G}$ is an isomorphism of G.*
2. *Assuming the order of χ is d in $\hat{G}$, then we know $\forall a, \chi(da) = \chi(a)^d = \chi_0(a) = 1$.*
3. *For any $\chi \neq \chi_0$, $\sum_{a\in G} \chi(a) = 0$.*
4. *$\overline{\chi(a)} = \chi(-a)$, where $\overline{\chi(a)}$ is the conjugate of $\chi(a)$.*

Matrix Analysis. For an n dimensional vector $x = (x_1, x_2, \cdots, x_n)^T$, we define its ℓ_2-norm $\|x\|_2 = \sqrt{\sum_{k=1}^n x_k^2}$.

For a matrix M, we use $M^\dagger$ to denote the conjugate transpose of M. For a function $\chi : G \to \mathbb{C}$ and a matrix $M \in G_{X\times Y}$, we use $\chi(M)$ to denote the matrix formed by $[\chi(M(x, y))]$, which is an element of $\mathbb{C}_{X\times Y}$.

We use the standard definition of spectral norm $\|\cdot\|$ for a matrix M to be $\|M\| = \max_{x\neq 0} \frac{\|Mx\|_2}{\|x\|_2}$, which is the largest singular value of M [10, Theorem 5.6.6].

The Kronecker product (or tensor product) of two matrices $A = [a_{i,j}]$ and B is denoted by $A \otimes B$, and is defined to be the following block matrix

$$\begin{pmatrix} a_{1,1}B & \cdots & a_{1,n}B \\ \vdots & \ddots & \vdots \\ a_{m,1}B & \cdots & a_{m,n}B \end{pmatrix}.$$

It satisfies the following lemma from [10, Theorem 4.2.12, 4.2.15].

Lemma 13. *Assume that the nonzero singular values of two matrices A and B are $\{\mu_i | 1 \le i \le m\}$ and $\{\lambda_j | 1 \le j \le n\}$ respectively, then the singular values of $A \otimes B$ are $\{\mu_i\lambda_j | 1 \le i \le m, 1 \le j \le n\}$.*

Number Theory. We will use $\varphi(m)$ to denote the Euler function of m, which is defined to be the number of positive integers less than or equal to m that are co-prime to m. For integer $m = p_1^{\alpha_1} \cdots p_k^{\alpha_k}$, we know $\varphi(m) = m \cdot \prod_{i=1}^{k} (1 - 1/p_k)$ from [17, Theorem 7.5].

3.2 The Discrepancy Method

The discrepancy method is a method to derive the communication complexity lower bound by giving an upper bound for a value called discrepancy defined below.

Definition 14 (Discrepancy)**.** *Let $f : X \times Y \mapsto \{0, 1\}$ be a function, R be a rectangle, and μ be a probability distribution on $X \times Y$. Denote $\mathrm{disc}_\mu(R, f) = |\sum_{(x,y)\in R} \mu(x, y)(-1)^{f(x,y)}|$, and $\mathrm{disc}_\mu(f) = \max_R \mathrm{disc}_\mu(R, f)$.*

The discrepancy is widely used in proving communication complexity lower bound [3, 21, 14], with many many applications. It was also used to prove the quantum lower bound [13, 18], and could be phrased in the following theorem.

Theorem 15 ([13])**.** *For any function f and any distribution μ, we have*

$$Q_\epsilon^*(f) = \Omega\left(\log \frac{1 - 2\epsilon}{\mathrm{disc}_\mu(f)}\right).$$

If the communication complexity problem is with promise, the discrepancy method still works if $\mu(x, y) = 0$ on (x, y) which is not in the promise.

We can use discrepancy to bound the quantum lower bound. Similarly, we can use the weight of the largest monochromatic rectangle to bound the nondeterministic lower bounds.

Definition 16. *Let $f : X \times Y \mapsto \{0, 1\}$ be a function, and μ be a probability distribution on $X \times Y$. We define the weight of the largest monochromatic rectangle filled with b to be*

$$\mathrm{mono}_\mu^b(f) = \max_{S\subseteq X, T\subseteq Y} \left\{\mu(S \times T) \middle| S \times T \subseteq f^{-1}(b)\right\}.$$

The following theorem will relate nondeterministic communication complexity and the size of largest monochromatic rectangle [14, Proposition 2.15].

Lemma 17. *For any $b \in \{0, 1\}$, we have the nondeterministic communication complexity of $f : X \times Y \to \{0, 1\}$ satisfies*

$$N^b(f) \ge \log_2 \frac{\mu(f^{-1}(b))}{\mathrm{mono}_\mu^b(f)}.$$

In this paper, we are going to use the discrepancy of a function to bound the weight of largest monochromatic rectangle.

Lemma 18. *For any function $f : X \times Y \to \{0,1\}$, distribution μ on $X \times Y$, and $b \in \{0,1\}$,*

$$\mathrm{mono}_\mu^b(f) \le \mathrm{disc}_\mu(f).$$

Proof.

$$\begin{aligned}\mathrm{disc}_\mu(f) &= \max_{S\times T\subseteq X\times Y} \mathrm{disc}_\mu(S\times T, f)\\ &\ge \max_{S\times T\subseteq f^{-1}(b)} \mathrm{disc}_\mu(S\times T, f) = \max_{S\times T\subseteq f^{-1}(b)} \mu(S\times T) = \mathrm{mono}_\mu^b(f)\end{aligned}$$

□

In this paper, we also need some tools from discrepancy on non-binary functions. The following concept of *excess count* has been used in [2] to give a definition of a value called strong multi-color discrepancy, which gives a lower bound for randomized communication complexity for multi-valued functions.

Definition 19 (Excess Count)**.** *Let $M \in G_{X\times Y}$ be a matrix. We define the* excess count *for an element $g \in G$ in a rectangle $S \times T \subseteq X \times Y$ as*

$$\mathrm{excess}_M(g, S\times T) = \big|\{(x,y)\in S\times T | M(x,y) = g\}\big| - \frac{|S||T|}{|G|}.$$

And the excess count for an element g is defined as the maximum value among all possible rectangles $S \times T$,

$$\mathrm{excess}_M(g) = \max_{S\times T\subseteq X\times Y} \mathrm{excess}_M(g, S\times T).$$

Furthermore, the strong multi-color discrepancy is upper bounded by another value called weak multi-color discrepancy. We phrase the relationship between strong and weak multi-color discrepancy in terms of excess count in the following lemma.

Lemma 20 (Lemma 2.9 of [2])**.** *For matrix $M \in G_{X\times Y}$ and any $S \times T \subseteq X \times Y$, we have*

$$\max_{g\in G} |\mathrm{excess}_M(g, S\times T)| \le \frac{1}{|G|} \sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}} \left| \sum_{(x,y)\in S\times T} \chi(M(x,y)) \right|.$$

We are going to phrase the above lemma in terms of matrix norms.

Lemma 21. *For matrix $M \in G_{X\times Y}$, we have*

$$\max_{g\in G}\{\mathrm{excess}_M(g)\} \le \frac{\sqrt{|X||Y|}}{|G|} \sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}} \|\chi(M)\|.$$

Lemma 21 can be proved by Lemma 20 and $\left|\sum_{(x,y)\in S\times T} \chi(M(x,y))\right| \le \|\vec{1_S}\|_2 \cdot \|\chi(M)\| \cdot \|\vec{1_T}\|_2$. We put the proof details in Appendix B.

3.3 Lower Bound for $\mathbf{IP}^{01}_{m,n}$ and $\mathbf{IP}^{01*}_{m,n}$

We define matrices $\Phi \in (\mathbb{Z}_m)_{m^n \times m^n}$ by $\Phi(x,y) = \langle x,y\rangle$ and $\Phi^* \in (\mathbb{Z}_m)_{m^n \times \varphi(m)^n}$ by $\Phi^*(x,y^*) = \langle x,y^*\rangle$ to be the communication matrices of $\mathrm{IP}_{m,n}$ on $\mathbb{Z}_m^n \times \mathbb{Z}_m^n$ and $\mathbb{Z}_m^n \times (\mathbb{Z}_m^*)^n$, respectively, where $x,y \in \mathbb{Z}_m^n$ and $y^* \in (\mathbb{Z}_m^*)^n$.

We first state the lemmas we need to get a lower bound of $\mathrm{IP}^{01}_{m,n}$ and $\mathrm{IP}^{01*}_{m,n}$ with the proof delayed to Appendix B.

Lemma 22. *Let $\chi \in \widehat{\mathbb{Z}_m}$, $\chi \neq \chi_0$ be an order d character of $\mathbb{Z}_m$, we have*

$$\|\chi(\Phi)\| = \left(\frac{m^2}{d}\right)^{n/2} \text{ and } \|\chi(\Phi^*)\| = \left(m \cdot \frac{\varphi(m)}{\varphi(d)}\right)^{n/2}.$$

Lemma 23. *In Φ, the number of 0's is at least m^{2n-1} and the number of 1's is at least $\varphi(m)\cdot m^{2n-2}$. In Φ^*, the number of k's is $m^{n-1}\cdot\varphi(m)^n$ for $k = 0,1,\cdots,m-1$.*

By combining the above lemmas with Lemma 21, we have the following theorem.

Theorem 24 (Theorem 10 Restated)**.** *The quantum communication complexity of $\mathrm{IP}^{01}_{m,n}$ and $\mathrm{IP}^{01*}_{m,n}$ satisfy*

$$Q^*_{1/3}(\mathrm{IP}^{01}_{m,n}) = \Omega(n \log p(m) - \log m)$$

and

$$Q^*_{1/3}(\mathrm{IP}^{01*}_{m,n}) = \Omega(n \log(p(m)-1) - \log m),$$

where $p(m)$ is the smallest prime factor of m. And the nondeterministic communication complexity of $\mathrm{IP}^{01}_{m,n}$ and $\mathrm{IP}^{01*}_{m,n}$ satisfy

$$N^b(\mathrm{IP}^{01}_{m,n}) = \Omega(n \log p(m) - \log m)$$

and

$$N^b(\mathrm{IP}^{01*}_{m,n}) = \Omega(n \log(p(m)-1) - \log m)$$

for any $b \in \{0,1\}$.

Proof. Let μ be the distribution uniformly distributed on the coordinates $(x,y) \in \mathbb{Z}_m^n \times \mathbb{Z}_m^n$ where $\langle x,y\rangle \in \{0,1\}$, and let μ^* be the distribution uniformly distributed on the coordinates $(x,y^*) \in \mathbb{Z}_m^n \times (\mathbb{Z}_m^*)^n$ where $\langle x,y^*\rangle \in \{0,1\}$. We are going to give upper bounds for $\mathrm{disc}_\mu(\mathrm{IP}^{01}_{m,n})$ and $\mathrm{disc}_{\mu^*}(\mathrm{IP}^{01*}_{m,n})$ to obtain lower bounds for their communication complexity.

We know $\mu(x,y) = \alpha$ is the same for all (x,y) satisfying $\langle x,y\rangle \in \{0,1\}$. So we can bound the discrepancy of $\mathrm{IP}^{01}_{m,n}$ by the excess of Φ in the following way:

$$\begin{aligned}
\mathrm{disc}_\mu(\mathrm{IP}^{01}_{m,n}) &= \max_{S\times T\subseteq X\times Y} \mathrm{disc}_\mu(\mathrm{IP}^{01}_{m,n}, S\times T)\\
&= \max_{S\times T\subseteq X\times Y} \left|\sum_{(x,y)\in S\times T} \mu(x,y)\cdot(-1)^{\Phi(x,y)}\right|\\
&= \max_{S\times T\subseteq X\times Y} \alpha\cdot\big|\,|\{(x,y)\in S\times T|\Phi(x,y)=0\}| - |\{(x,y)\in S\times T|\Phi(x,y)=1\}|\,\big|\\
&= \max_{S\times T\subseteq X\times Y} \alpha\,|\mathrm{excess}_\Phi(0,S\times T) - \mathrm{excess}_\Phi(1,S\times T)|\\
&\le \alpha\cdot \max_{S\times T\subseteq X\times Y} 2\max_{g\in\mathbb{Z}_m} |\mathrm{excess}_\Phi(g,S\times T)| && \text{(Triangle Eq.)}\\
&\le \alpha\cdot\frac{2\sqrt{m^n\cdot m^n}}{m}\sum_{\substack{\chi\in\widehat{\mathbb{Z}_m}\\ \chi\neq\chi_0}} \|\chi(\Phi)\|. && \text{(Lemma 21)}
\end{aligned}$$

By Lemma 22 we know that for χ with order d the norm of $\chi(\Phi)$ is $\left(\frac{m^2}{d}\right)^{n/2}$. Since d is an order and $\chi \neq \chi_0$ we know $d|m$ and $d \neq 1$. So we have the norm of $\chi(\Phi)$ satisfies $\left(\frac{m^2}{d}\right)^{n/2} \leq \left(\frac{m^2}{p(m)}\right)^{n/2}$. And by Lemma 23 we know that $\alpha \leq 1/m^{2n-1}$, thus

$$\mathrm{disc}_\mu(\mathrm{IP}^{01}_{m,n}) \leq \alpha \cdot 2m^{n-1} \cdot (m-1) \left(\frac{m^2}{p(m)}\right)^{n/2} \leq \frac{2m}{p(m)^{n/2}}.$$

By Theorem 15,

$$Q^*_\epsilon(\mathrm{IP}^{01}_{m,n}) \geq \log \frac{1-2\epsilon}{\mathrm{disc}_\mu(\mathrm{IP}^{01}_{m,n})} = \Omega(n \log p(m) - \log m + \log(1-2\epsilon)).$$

For $\mathrm{IP}^{01*}_{m,n}$, we can also bound $\mathrm{disc}_{\mu^*}(\mathrm{IP}^{01*}_{m,n})$ by $\chi(\Phi^*)$ in the same way, yielding a bound of

$$\mathrm{disc}_{\mu^*}(\mathrm{IP}^{01*}_{m,n}) \leq 2m(\varphi(p(m)))^{-n/2},$$

which in turn means

$$Q^*_\epsilon(\mathrm{IP}^{01*}_{m,n}) \geq \log \frac{1-2\epsilon}{\mathrm{disc}_\mu(\mathrm{IP}^{01*}_{m,n})} = \Omega(n \log(p(m)-1) - \log m + \log(1-2\epsilon)).$$

For the nondeterministic lower bound, we first show that by Lemma 18, we have

$$\mathrm{mono}^b_\mu(\mathrm{IP}^{01}_{m,n}) \leq \mathrm{disc}_\mu(\mathrm{IP}^{01}_{m,n}) \leq m(p(m))^{-n/2}$$

and

$$\mathrm{mono}^b_{\mu^*}(\mathrm{IP}^{01*}_{m,n}) \leq \mathrm{disc}_{\mu^*}(\mathrm{IP}^{01*}_{m,n}) \leq 2m(\varphi(p(m)))^{-n/2}.$$

Theorem 23 claims that the number of 0's and the number of 1's in Φ^* is the same, implying $\mu^*({\mathrm{IP}^{01*}_{m,n}}^{-1}(0)) = \mu^*({\mathrm{IP}^{01*}_{m,n}}^{-1}(1)) = 1/2$. And the number of 0's and 1's in Φ are at least m^{2n-1} and $\varphi(m)m^{2n-2}$, respectively. So we have

$$\mu({\mathrm{IP}^{01}_{m,n}}^{-1}(0)) \geq m^{2n-1}/m^{2n} = 1/m$$

and

$$\mu({\mathrm{IP}^{01}_{m,n}}^{-1}(1)) \geq \varphi(m)m^{2n-2}/m^{2n} = \varphi(m)/m^2.$$

By Lemma 17, we have the nondeterministic communication complexity of $\mathrm{IP}^{01}_{m,n}$ and $\mathrm{IP}^{01*}_{m,n}$ for any $b \in \{0,1\}$,

$$\begin{aligned}
N^0(\mathrm{IP}^{01}_{m,n}) &\geq \log_2 \frac{\mu({\mathrm{IP}^{01}_{m,n}}^{-1}(0))}{\mathrm{mono}^0_\mu(\mathrm{IP}^{01}_{m,n})} \\
&\geq \log_2 \frac{1/m}{2m(p(m))^{-n/2}} \geq \frac{n}{2}\log_2 p(m) - 2\log_2 m - 1, \\
N^1(\mathrm{IP}^{01}_{m,n}) &\geq \log_2 \frac{\mu({\mathrm{IP}^{01}_{m,n}}^{-1}(1))}{\mathrm{mono}^1_\mu(\mathrm{IP}^{01}_{m,n})} \\
&\geq \log_2 \frac{\varphi(m)/m^2}{2m(p(m))^{-n/2}} \geq \frac{n}{2}\log_2 p(m) - 3\log_2 m + \log_2 \varphi(m) - 1, \\
N^b(\mathrm{IP}^{01*}_{m,n}) &\geq \log_2 \frac{\mu^*({\mathrm{IP}^{01*}_{m,n}}^{-1}(b))}{\mathrm{mono}^b_{\mu^*}(\mathrm{IP}^{01*}_{m,n})} \\
&\geq \log_2 \frac{1/2}{2m(\varphi(p(m)))^{-n/2}} \geq \frac{n}{2}\log_2(p(m)-1) - \log_2 m - 2.
\end{aligned}$$

□

References

[1] L. Babai. The Fourier transform and equations over finite abelian groups. *Lecture Notes, version 1.3*, 1, 1989.

[2] L. Babai, T.P. Hayes, and P.G. Kimmel. The cost of the missing bit: Communication complexity with help. *Combinatorica*, 21(4):455–488, 2001.

[3] László Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory (preliminary version). In *FOCS*, pages 337–347. IEEE, 1986.

[4] Ziv Bar-Yossef, Ravi Kumar, and D. Sivakumar. Reductions in streaming algorithms, with an application to counting triangles in graphs. In *SODA*, pages 623–632, 2002.

[5] Jeff I. Chu and Georg Schnitger. Communication complexity of matrix computation over finite fields. *Mathematical Systems Theory*, 28(3):215–228, 1995.

[6] Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. Graph distances in the streaming model: the value of space. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, SODA '05, pages 745–754, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[7] Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. On graph problems in a semi-streaming model. *Theoretical Computer Science*, 348(2-3):207–216, 2005. Automata, Languages and Programming: Algorithms and Complexity (ICALP-A 2004).

[8] András Hajnal, Wolfgang Maass, and György Turán. On the communication complexity of graph properties. In *STOC*, pages 186–191. ACM, 1988.

[9] Nicholas J. A. Harvey. Matroid intersection, pointer chasing, and young's seminormal representation of sn. In Shang-Hua Teng, editor, *SODA*, pages 542–549. SIAM, 2008.

[10] R.A. Horn and C.R. Johnson. *Matrix analysis*, volume 2. Cambridge Univ Pr, 1990.

[11] Nathan Jacobson. *Basic Algebra I: Second Edition.* W. H. Freeman and Company, 2009.

[12] Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992.

[13] I. Kremer. Quantum Communication. *Master's thesis, The Hebrew University of Jerusalem*, 1995.

[14] E. Kushilevitz and N. Nisan. *Communication Complexity.* Cambridge Univ Pr, 1997.

[15] Ran Raz and Boris Spieker. On the "log rank"-conjecture in communication complexity. In *FOCS*, pages 168–176. IEEE, 1993.

[16] Alexander A. Razborov. On the distributional complexity of disjontness. In Mike Paterson, editor, *ICALP*, volume 443 of *Lecture Notes in Computer Science*, pages 249–253. Springer, 1990.

[17] Kenneth H. Rosen. *Elementary Number Theory and Its Applications.* Addison-Wesley, 3 sub edition edition, 1992.

[18] Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In Cynthia Dwork, editor, *STOC*, pages 85–94. ACM, 2008.

[19] E. Verbin and W. Yu. The Streaming Complexity of Cycle Counting, Sorting By Reversals, and Other Problems. In *SODA*, 2011.

[20] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *STOC*, pages 209–213. ACM, 1979.

[21] Andrew Chi-Chih Yao. Lower bounds by probabilistic arguments (extended abstract). In *FOCS*, pages 420–428. IEEE, 1983.

A Upper Bound for the Cycle Counting Problem

Lemma 25. *There is a* 1*-bit deterministic protocol for* $\mathrm{CC}_{n,a,b}$ *if* $a \not\equiv b \pmod 2$*.*

Proof. First, we show that for the $\mathrm{CC}_{n,a,b}$ problem, if $a \not\equiv b \pmod 2$, we will have a 1 bit deterministic protocol to solve the problem.

The parity (oddness or evenness) of a permutation is defined as the parity of the number of transpositions in factorization. According to [11, Section 1.6] we know that the parity of $\sigma \circ \pi$ is same to the parity of σ plus the parity of π. So Alice can send the parity of σ to Bob by using only 1 bit. And after that, since they have different parity, Bob could know if there are a cycles or b cycles according to the parity. □

The following theorem give an upper bound for $R(\mathrm{CC}_{n,1,m})$. It is tight up to a poly-logarithm factor compare with Theorem 11.

Lemma 26. *There is a randomized protocol for* $\mathrm{CC}_{n,1,m}$ *for odd* m *with communication complexity*

$$\min\left\{O(n\log n), O\left(n/m \cdot \log n \cdot \log(n/m)\right)\right\}.$$

Proof. $O(n\log n)$ is a trivial upper bound since Alice can send π in $O(n\log n)$ bits.

Let $l = \lceil \log_2(n/m) \rceil$, we have the following protocol that will output 1 for the m cycles case with probability at least 2/3; and 0 for the 1 cycle case.

1. Repeat the following $16(l+2)$ times:
   (a) Pick r uniformly random in $\{0, 1, ..., l+1\}$.
   (b) Repeat the following $\frac{n}{2^r m}$ times:
      i. Pick $x \in [n]$ uniformly random, and follow $\sigma \circ \pi$ 2^r times by communicating with each other to check if there is a cycle in $x, \sigma \circ \pi(x), \dots, (\sigma \circ \pi)^{2^r-1}(x)$. If there is a cycle with length less than n, return 1; otherwise continue.
2. Return 0.

The amount of bits sent is

$$16(l+2) \cdot \frac{n}{2^r m} \cdot 2^r \log n = O\left(\log n \cdot \log\frac{n}{m} \cdot \frac{n}{m}\right).$$

It is easy to see that the protocol will always return 0 for the 1 cycle case. We are going to lower bound the probability that the protocol returns 0 for the m cycle case.

If the $\sigma \circ \pi$ consists of m cycles, we assume that the m cycles are of length $s_1, s_2, ..., s_m$, which satisfy $0 < s_1 \le s_2 \le ... \le s_m \le n$ and $s_1 + s_2 + ... + s_m = n$. Let $f(r)$ to be the maximal index i which satisfies $s_i \le 2^r$, we can see that $f(l+1) \ge m/2$.

For a fixed r, the probability the protocol detected a cycle (Step 1(b)i) is

$$\begin{aligned}\Pr[\text{Step 1(b)i returns 1}] &= (s_1 + s_2 + \ldots + s_{f(r)})/n \\ &\ge (s_{f(r-1)+1} + \ldots + s_{f(r)})/n \\ &= (f(r) - f(r-1)) \cdot 2^{r-1}/n.\end{aligned}$$

We repeat it $\frac{n}{2^r m}$ times, so the probability of discovering at least one circle is

$$\begin{aligned}
\Pr[\text{Step 1b returns 1}] &= 1-(1-\Pr[\text{Step 1(b)i returns 1}])^{\frac{n}{2^r m}} \\
&> 1-\exp\left(-\frac{2^{r-1}(f(r)-f(r-1))}{n}\cdot\frac{n}{2^r m}\right) \\
&> \frac{(f(r)-f(r-1))/(2m)}{1+(f(r)-f(r-1))/(2m)} \\
&> \frac{(f(r)-f(r-1))/(2m)}{1+1} \\
&= \frac{f(r)-f(r-1)}{4m}.
\end{aligned}$$

This is the probability for a fixed r, we need to average over all the possible r,

$$\begin{aligned}
\Pr[\text{Step 1a returns 1}] &= \frac{1}{l+2}\sum_{r=0}^{l+1}\Pr[\text{Step 1b returns 1}] \\
&> \frac{1}{l+2}\sum_{r=0}^{l+1}\frac{f(r)-f(r-1)}{2m} \\
&= \frac{1}{l+2}\cdot\frac{f(l+1)}{4m} \\
&\geq \frac{1}{8(l+2)}.
\end{aligned}$$

Here the last step is because $f(l+1)\geq m/2$ as discussed.

If we repeat the whole procedure $16(l+2)$ times, we will see that

$$\begin{aligned}
\Pr[\text{Step 1 return 1}] &> 1-(1-\Pr[\text{Step 1a returns 1}])^{16(l+2)} \\
&> 1-\exp\left(-\frac{1}{8(l+2)}\cdot 16(l+2)\right) \\
&> \frac{2}{3}.
\end{aligned}$$

Thus it completes the proof. □

B Proof of Lemma 21, Lemma 22 and Lemma 23

Proof of Lemma 21. By using $\vec{1_S}$ and $\vec{1_T}$ to denote the indicator vector of S and T (i.e. $\vec{1_S}(x)=1 \iff x\in S$), we have

$$\begin{aligned}
\left|\sum_{(x,y)\in S\times T}\chi(M(x,y))\right| &= \left|\vec{1_S}^{\dagger}\cdot\chi(M)\cdot\vec{1_T}\right| && (1)\\
&\leq \|\vec{1_S}\|_2\cdot\|\chi(M)\cdot\vec{1_T}\|_2 && \text{(Cauchy-Schwarz)}\\
&\leq \|\vec{1_S}\|_2\cdot\|\chi(M)\|\cdot\|\vec{1_T}\|_2. && \text{(Definition of Matrix Norm)}
\end{aligned}$$

Combine the above inequality with Lemma 20 we know

$$\begin{aligned}
\max_{g\in G}\{\mathrm{excess}_M(g)\} &= \max_{g\in G}\left\{\max_{S\times T\subseteq X\times Y} \mathrm{excess}_M(g, S\times T)\right\} \\
&\le \max_{S\times T\subseteq X\times Y}\left\{\frac{1}{|G|}\sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}}\left|\sum_{(x,y)\in S\times T}\chi(M(x,y))\right|\right\} && \text{(Lemma 20)} \\
&\le \max_{S\times T\subseteq X\times Y}\left\{\frac{1}{|G|}\sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}}\|\vec{1_S}\|_2\cdot\|\chi(M)\|\cdot\|\vec{1_T}\|_2\right\} \\
&\le \frac{1}{|G|}\sum_{\substack{\chi\in\hat{G}\\ \chi\neq\chi_0}}\sqrt{|X|}\cdot\|\chi(M)\|\cdot\sqrt{|Y|}.
\end{aligned}$$

□

Proof of Lemma 22. Actually, we can write out all the singular values of $\chi(\Phi)$.

First, let $H \in (\mathbb{Z}_m)_{m\times m}$ be the matrix where $H(u,v) = \chi(uv)$ for $u, v \in \mathbb{Z}_m$. We want to show that $\chi(\Phi) = H\otimes H\otimes\cdots\otimes H = H^{\otimes n}$. This is because for $x, y \in \mathbb{Z}_m^n$, we have

$$\begin{aligned}
\chi(\Phi)(x,y) &= \chi\left(\sum_{1\le i\le n} x_i y_i\right) \\
&= \prod_{1\le i\le n}\chi(x_i y_i) && \text{(Homomorphism)} \\
&= \prod_{1\le i\le n} H(x_i, y_i) = H^{\otimes n}(x,y).
\end{aligned}$$

Second, we want to know the singular values of H. So we examine $H^\dagger H$ as following,

$$\begin{aligned}
H^\dagger H(u,v) &= \sum_{w\in\mathbb{Z}_m}\overline{\chi(uw)}\chi(wv) \\
&= \sum_{w\in\mathbb{Z}_m}\chi(w(v-u)).
\end{aligned}$$

If $d|v-u$, we know that $\chi(w(v-u)) = \chi^{v-u}(w) = \chi_0(w) = 1$, thus $\sum_{w\in\mathbb{Z}_m}\chi(w(v-u)) = m$. Otherwise $\sum_{w\in\mathbb{Z}_m}\chi(w(v-u)) = \sum_{w\in\mathbb{Z}_m}\chi^{v-u}(w) = 0$ by Proposition 12. So it is not difficult to write out $H^\dagger H = m\cdot\mathbf{1}_{m/d\times m/d}\otimes I_{d\times d}$, where $\mathbf{1}_{m/d\times m/d}$ is the all 1 matrix. The singular values of H are the square root of the eigenvalues of $H^\dagger H$, which are the product of m/d (eigenvalue of $\mathbf{1}_{m/d\times m/d}$) and 1 (the eigenvalue of I) by Lemma 13 without counting multiplicities.

So by Lemma 13, the singular values of $\chi(\Phi)$ are $\left(\frac{m^2}{d}\right)^n$ without counting multiplicities.

For $\chi(\Phi^*)$, let H^* to be the submatrix of H restricted on $\mathbb{Z}_m\times\mathbb{Z}_m^*$, i.e., only with the columns restricted to $\mathbb{Z}_m^*$. It is easy to see that $\chi(\Phi^*) = (H^*)^{\otimes n}$, because $\chi(\Phi^*)$ is just $\chi(\Phi^*)$ with columns restricted on $(\mathbb{Z}_m^*)^n$. Moreover, we can see that $(H^*)^\dagger H^*$ is just $H^\dagger H$ restricted on $(\mathbb{Z}_m^*)^n\times(\mathbb{Z}_m^*)^n$, implying $(H^*)^\dagger H^* = m\cdot\mathbf{1}_{\frac{\varphi(m)}{\varphi(d)}\times\frac{\varphi(m)}{\varphi(d)}}\otimes I_{\varphi(d)\times\varphi(d)}$. Thus by Lemma 13, the singular values of $\chi(\Phi^*)$ are $\left(m\cdot\frac{\varphi(m)}{\varphi(d)})\right)^n$ without counting multiplicities. □

Proof of Lemma 23. First, we count the number of 0's in Φ.

When m is a prime power, assume that $m = p^\alpha$ for some prime p, then for a fixed y, let $d = \gcd(m, y_1, y_2, \ldots, y_n)$. Since $m = p^\alpha$, we know $d | y_i$ for all i. w.l.o.g., we assume $\gcd(m, y_1) = d$. For a fixed tuple $x_2, x_3, \cdots, x_n$, we want to know how many different values of x_1 could satisfy the following equation.

$$x_1 y_1 + \sum_{i=2}^{n} x_i y_i \equiv 0 \pmod{m}.$$

We know the solutions to the above equation is the same as the following one.

$$x_1 \cdot \frac{y_1}{d} + \sum_{i=2}^{n} x_i \cdot \frac{y_i}{d} \equiv 0 \pmod{\frac{m}{d}}.$$

Since $\frac{y_1}{d}$ is relatively prime to $\frac{m}{d}$, we know there is unique x_1 satisfying the above equation. Thus there are at least $d/m \geq 1/m$ fraction of 1's in column y for any y.

When $m = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_l^{\alpha_l}$, where p_j's are distinct primes and α_j's are positive integers. For $1 \leq j \leq l$ we let $x_i^{(j)} = x_i \mod p_j^{\alpha_j}$ and $y_i^{(j)} = y_i \mod p_j^{\alpha_j}$. By Chinese Remainder Theorem, we know that $\sum_{1 \leq i \leq n} x_i y_i \equiv 0 \pmod{m}$ iff $\sum_{1 \leq i \leq n} x_i^{(j)} y_i^{(j)} \equiv 0 \pmod{p_j^{\alpha_j}}$ for any $1 \leq j \leq l$. Thus,

$$\Pr_{x,y \in \mathbb{Z}_m^n} \left[\sum_i x_i y_i = 0 \right] = \prod_j \Pr_{x^{(j)}, y^{(j)} \in \left(\mathbb{Z}_{p_j^{\alpha_j}}\right)^n} \left[\sum_i x_i^{(j)} y_i^{(j)} = 0 \right] \geq \prod_j \frac{1}{p_j^{\alpha_j}} = \frac{1}{m}.$$

Second, we count the number of 1's in Φ. For any y such that y_1 is relatively prime to m, $\langle x, y \rangle = 1$ if $x_1 = y_1^{-1}(1 - \sum_{i=2}^{n} x_i y_i)$. We have $\varphi(m) \cdot m^{n-1}$ such y's, and for each y we have m^{n-1} choices of x. So, the number of 1's is at least $\varphi(m) \cdot m^{2n-2}$.

At last, counting the number of k's in Φ^* is the same as counting the number of 1's in Φ. We have $\varphi(m)^n$ y's, and each y corresponds to m^{n-1} x's. □