Randomized Communication Complexity for Linear Algebra Problems over Finite Fields *

Xiaoming Sun[1] and Chengu Wang[2]

1 **Institute of Computing Technology, Chinese Academy of Sciences**
Beijing, China
sunxiaoming@ict.ac.cn

2 **IIIS, Tsinghua University**
Beijing, China
wangchengu@gmail.com

Abstract

Finding the singularity of a matrix is a basic problem in linear algebra. Chu and Schnitger [3] first considered this problem in the communication complexity model, in which Alice holds the first half of the matrix and Bob holds the other half. They proved that the deterministic communication complexity is $\Omega(n^2 \log p)$ for an $n \times n$ matrix over the finite field $\mathbb{F}_p$. Then, Clarkson and Woodruff [4] introduced the singularity problem to the streaming model. They proposed a randomized one pass streaming algorithm that uses $O(k^2 \log n)$ space to decide if the rank of a matrix is k, and proved an $\Omega(k^2)$ lower bound for randomized one-way protocols in the communication complexity model.

We prove that the randomized/quantum communication complexity of the singularity problem over $\mathbb{F}_p$ is $\Omega(n^2 \log p)$, which implies the same space lower bound for randomized streaming algorithms, even for a constant number of passes. The proof uses the framework by Lee and Shraibman [8], but we choose Fourier coefficients as the witness for the dual approximate norm of the communication matrix. Moreover, we use Fourier analysis to show the same randomized/quantum lower bound when deciding if the determinant of a non-singular matrix is a or b for non-zero a and b.

1998 ACM Subject Classification F.2.1 Numerical Algorithms and Problems; G.1.3 Numerical Linear Algebra

Keywords and phrases communication complexity, streaming, matrix, singularity, determinant

1 Introduction

Communication complexity, introduced by Yao [16], is a powerful tool to solve a variety of problems in areas as disparate as VLSI design, decision trees, data structures and circuit complexity [7]. It is a game between two parties, Alice and Bob, with unlimited computing power, that want to compute the value of a function $f : X \times Y \mapsto \{0, 1\}$, but Alice only knows $x \in X$ while Bob only knows $y \in Y$. The communication complexity is the minimal amount of bits they transfer. We denote the randomized and quantum communication complexity which succeeds with probability at least $1 - \epsilon$ by $R_\epsilon(f)$ and $Q_\epsilon(f)$ (or $Q^*_\epsilon(f)$) respectively, where $R_\epsilon(f)$ is with private coin, $Q_\epsilon(f)$ is without entanglement and $Q^*_\epsilon(f)$ is with entanglement. The three functions have the following relationship: $Q^*_\epsilon(f) \leq Q_\epsilon(f) \leq O(R_\epsilon(f))$ [6].

* This work was supported in part by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61033001, 61061130540, 61073174. The first author was also supported in part by the National Natural Science Foundation of China Grant 61170062 and Tsinghua University Initiative Scientific Research Program 2009THZ02120.

In the streaming model, the input is presented as a sequence and can be examined by the algorithm in only a few passes (typically just one). We are interested in the size of memory the algorithm uses. It can be proved by a reduction that the size of memory times the number of passes can be bounded by the communication complexity if Alice holds the first part of the stream and Bob holds the remaining.

We focus on linear algebra problems, because they are fundamental problems in mathematics, and the matrix computation is used everywhere. Furthermore, the singularity problem is the most basic problem, because it can be reduced to many linear algebra problems, e.g. to determine whether linear equations have a solution, to compute the diagonal of the LU decomposition or QR decomposition and to determine whether two subspaces intersect. Formally speaking, for an $n \times n$ matrix x whose entries are in the the finite field $\mathbb{F}_p$, the singularity problem is to decide whether x is singular over $\mathbb{F}_p$, and the determinant problem is to compute the determinant of x over $\mathbb{F}_p$ for non-singular x. In the streaming model, the input x comes row by row sequentially, while in the communication complexity model, Alice holds the first $n/2$ rows, and Bob holds the remaining $n/2$ rows.

These problems have a trivial deterministic algorithm that uses $O(n^2 \log p)$ spaces in one pass or the same number of communications in one-way. Chu and Schnitger [3] proved an $\Omega(n^2 \log p)$ communication complexity for deterministic protocols of the singularity problem. Luo and Tsitsiklis [9] proved that a deterministic protocol must transfer $\Omega(n^2)$ real numbers for the matrix inversion problem. Clarkson and Woodruff [4] proposed a randomized one pass streaming algorithm that uses $O(k^2 \log n)$ space to decide if the rank of a matrix is k, and proved an $\Omega(k^2)$ lower bound for randomized one-way protocol in the communication complexity model by reducing from the index function, which implies an $\Omega(n^2)$ space lower bound in the streaming model with one pass. Deciding the disjointness of two $n/2$ dimensional subspaces is actually the singularity problem. Miltersen et al. [10] showed a tight lower bound when deciding whether a vector is in a subspace of $\mathbb{F}_2^n$ in the one-sided error randomized asymmetric communication complexity model, by using the Richness Lemma.

In the communication model, there is another way to distribute the input: Alice and Bob each holds an $n \times n$ matrix x and y, respectively, and they want to compute the singularity or determinant of $x + y$. The two ways are equivalent up to a constant factor, because

$$\det(x+y) = \det \begin{pmatrix} x+y & \mathbf{0}_{n\times n} \\ y & \mathbf{I}_{n\times n} \end{pmatrix} = \det \begin{pmatrix} x & -\mathbf{I}_{n\times n} \\ y & \mathbf{I}_{n\times n} \end{pmatrix}.$$

This way is more beautiful and symmetric. When $p = 2$, $f(x + y) = f(x \oplus y)$ is a important block-composed function with good properties and attracted lots of attentions recently [11, 18, 14, 17]. Here, we formally define the problems in this way.

▶ **Problem 1** (Singularity)**.** *Alice and Bob hold two $n \times n$ matrices x and y over $\mathbb{F}_p$, separately. They want to determine whether $x + y$ is singular over $\mathbb{F}_p$.*

▶ **Problem 2** ($\text{Det}_{a,b}$)**.** *Alice and Bob hold two $n \times n$ matrices x and y over $\mathbb{F}_p$, separately. For $a, b \in \mathbb{F}_p \setminus \{0\}$, we promise that $\det_p(x+y)$ is either a or b, where $\det_p$ is the determinant over $\mathbb{F}_p$. They want to compute $\det_p(x + y)$.*

1.1 Our Results

▶ **Theorem 3.** *The randomized/quantum communication complexity of* Singularity *is $\Omega(n^2 \log p)$.*

We prove it using the duality of the approximate norm [8]. We compute all the Fourier coefficients of the singularity function and use them as a witness in the Duality Theorem.

This result implies the same lower bound when deciding if two subspaces of $\mathbb{F}_p^n$ intersect. Also, it implies that $\Omega(k^2)$ communication is required to decide if the rank of $x+y$ is k, by padding n^2-k^2 zeros, which improves the determinant result in [4].

▶ **Theorem 4.** *The randomized/quantum communication complexity of* $\textsc{Det}_{a,b}$ *is* $\Omega(n^2 \log p)$*, for all non-zero* a *and* b*.*

For this problem, we prove a small spectral norm of the matrix representing the problem, by decomposing the matrix onto Fourier basis which has good properties.

All these results imply the same space lower bound for randomized streaming algorithms, even if the algorithm reads the stream a constant number of passes.

1.2 Outline

In Section 2, we define the basic notations. We prove Theorem 3 in Section 3 and Theorem 4 in Section 4. Finally, we discuss the open problem in Section 5.

2 Preliminaries

For prime p, $\mathbb{F}_p$ is a finite field. For a function $f : \mathbb{F}_p^N \mapsto \mathbb{R}$, the Fourier coefficient of f is

$$\hat{f}(s) = \frac{1}{p^N} \sum_{x \in \mathbb{F}_p^N} \omega^{-\langle s,x \rangle} f(x),$$

where $\omega = e^{2\pi i/p}$. The inverse transform is

$$f(x) = \sum_{s \in \mathbb{F}_p^N} \omega^{\langle s,x \rangle} \hat{f}(s).$$

The Kronecker delta, denoted by $\delta_{i,j}$, is 1 if $i = j$ and 0 otherwise.

▶ **Fact 5.** *The number of* $n \times n$ *matrices of rank* r *over* $\mathbb{F}_p$ *is* $p^{r(r-1)/2} \binom{n}{r}_p \prod_{k=n-r+1}^{n} (p^k - 1)$*. Especially, the number of non-singular* $n \times n$ *matrices over* $\mathbb{F}_p$ *is* $\prod_{k=0}^{n-1} (p^n - p^k)$*.*

In this paper, we don't distinguish vectors (matrices) from discrete unary (bivariate) functions. For example, for vector v, $v(x)$ means the x-th element of v, and for matrix A, $A(x, y)$ means the entry at x-th row and y-th column.

For a vector x, we define the ℓ_1-norm $\|x\|_1 = \sum_i |x_i|$, and the ℓ_∞-norm $\|x\|_\infty = \max_i |x_i|$.

For a matrix A, we also define the ℓ_1-norm $\|A\|_1 = \sum_{i,j} |x_{i,j}|$, and the ℓ_∞-norm $\|A\|_\infty = \max_{i,j} |A_{i,j}|$. Let $\sigma = (\sigma_1, \cdots, \sigma_{\text{rank}(A)})$ be the vector of nonzero singular values of A. The trace norm of A is $\|A\|_{\text{tr}} = \|\sigma\|_1$. The spectral norm is $\|A\| = \|\sigma\|_\infty$, which is also the square root of the largest eigenvalue of the positive-semidefinite matrix $A^\dagger A$ [5], where $A^\dagger$ is the conjugate transpose of A.

3 Lower Bound for Singularity Problem

We first introduce the approximate rank and norm. Then, for XOR composed function $g(x \oplus y)$, the trace norm is equal to the ℓ_1 norm of the Fourier coefficients. This property still holds for approximate norm and for $g(x+y)$ in $\mathbb{F}_p$. After that, we present the Duality Theorem, which converts the definition of the approximate norm from min to max. Finally, we compute Fourier coefficients of the singularity function, and choose it as the witness.

3.1 Approximation Norms

The matrix rank and matrix norm can give a lower bound for deterministic communication complexity. Similarly, the approximate rank and norm can prove a lower bound for randomized/quantum protocols.

For $\alpha \geq 1$ and a sign matrix A, we define the approximate trace norm by

$$\|A\|_{\mathrm{tr}}^{\alpha} = \min_{B:1\leq A_{i,j}B_{i,j}\leq\alpha} \|B\|_{\mathrm{tr}},$$

and the approximate rank by

$$\mathrm{rank}^{\alpha}(A) = \min_{B:1\leq A_{i,j}B_{i,j}\leq\alpha} \mathrm{rank}(B).$$

For $\alpha \geq 1$ and a sign vector x, we define the approximate Fourier ℓ_1-norm by

$$\|\hat{x}\|_1^{\alpha} = \min_{y:1\leq x_i y_i\leq\alpha} \|\hat{y}\|_1.$$

▶ **Theorem 6.** *[2] Let A be a sign matrix and $0 < \epsilon < 1/2$, and $\alpha = 1/(1-2\epsilon)$, then $Q_\epsilon(A) \geq \frac{1}{2}\log \mathrm{rank}^{\alpha}(A)$.*

Because the approximate rank can be bounded by the approximate trace norm [8]:

$$\mathrm{rank}^{\alpha}(A) \geq \frac{(\|A\|_{\mathrm{tr}}^{\alpha})^2}{\alpha^2 \cdot \mathrm{size}(A)},$$

the approximate trace norm can give a lower bound for the communication complexity. This result can also be found in [12].

▶ **Lemma 7.** *Let $g : \mathbb{F}_p^N \mapsto \{-1,1\}$ be a sign function, and $f = g\circ+^{\otimes N}$, i.e. $f(x,y) = g(x+y)$. Let A be the sign matrix representing f. Then, $\|A\|_{\mathrm{tr}}^{\alpha} \geq \|\hat{g}\|_1^{\alpha} \cdot p^N$.*

The $p = 2$ case of Lemma 7 can be found in [8, Theorem 85]. It can be easily generalized to $\mathbb{F}_p$ with little changes. As a result, we omit this proof.

Combining them together, the randomized/quantum communication complexity can be bounded by the $\|\hat{\cdot}\|_1^{\alpha}$ norm.

▶ **Corollary 8.** *For $g : \mathbb{F}_p^N \mapsto \{-1,1\}$ and $f = g \circ +^{\otimes N}$, $Q_\epsilon(f) \geq \log \|\hat{g}\|_1^{\alpha} - 2\alpha$, where $\alpha = 1/(1-2\epsilon)$.*

3.2 Duality

The definition of the approximate Fourier ℓ_1 norm begins with $\min_y$. In such a definition, we have to check every y if we want to prove a lower bound. However, the Duality Theorem converts $\min_y$ to $\max_y$. As a result, a particular y, called the witness, is enough to prove a lower bound.

▶ **Definition 9.** For a general norm $\|\cdot\|$ on $\mathbb{R}^N$, the dual norm on $\mathbb{R}^N$, denoted by $\|\cdot\|^*$, is defined by

$$\|x\|^* = \max_{y\in\mathbb{R}^N:\|y\|\leq 1} \langle x, y\rangle.$$

▶ **Theorem 10** (Duality Theorem)**.** *[8, Theorem 64] For a general norm $\|\cdot\|$ on $\mathbb{R}^N$,*

$$\|x\|^{\alpha} = \max_{y:\|y\|^*\leq 1} \frac{1+\alpha}{2}\langle x, y\rangle + \frac{1-\alpha}{2}\|y\|_1.$$

3.3 Choosing Fourier Coefficients as Witness

It is difficult to find a useful witness. The first choice that comes to mind is to choose $h = g$, which can be used to prove the inner product problem. The discrepancy method is the special case of taking $h = \mu \circ g$ for a distribution μ [8]. Here we propose a new choice: taking $h = \hat{g}$. Now we calculate $\hat{g}$ first.

We define a sign function $g : \mathbb{F}_p^{n\times n} \mapsto \{-1, 1\}$, where $g(x) = -1$ if x is full rank over $\mathbb{F}_p$ and $g(x) = 1$ otherwise. Then, we define $f(x, y) = g(x + y)$. In such a definition, f is the function representing the SINGULARITY problem.

We change the value of g from $\{-1, 1\}$ to $\{0, 1\}$, by defining $g_{01} = (1 - g)/2$. In other words, $g_{01}(x) = 1$ if x is full rank, and $g_{01}(x) = 0$ otherwise.

$\widehat{g_{01}}$ has a good property that a same rank results in a same value.

▶ **Lemma 11.** *For $s, t \in \mathbb{F}_p^{n\times n}$, $\widehat{g_{01}}(s) = \widehat{g_{01}}(t)$ if $\mathrm{rank}_p(s) = \mathrm{rank}_p(t)$, where rank_p is the matrix rank over $\mathbb{F}_p$.*

Proof. Because $\mathrm{rank}_p(s) = \mathrm{rank}_p(t)$, there are full rank matrices u and v, such that $s = vtu$. Since v^{T} and u^{T} are full rank, $y = v^{\mathrm{T}}xu^{\mathrm{T}}$ is a bijection between matrices x and y.

$$\widehat{g_{01}}(s) = \frac{\sum_x \omega^{-\mathrm{tr}(s^{\mathrm{T}}x)} g_{01}(x)}{p^{n^2}} = \frac{\sum_x \omega^{-\mathrm{tr}(t^{\mathrm{T}}v^{\mathrm{T}}xu^{\mathrm{T}})} g_{01}(x)}{p^{n^2}} = \frac{\sum_y \omega^{-\mathrm{tr}(t^{\mathrm{T}}y)} g_{01}(y)}{p^{n^2}} = \widehat{g_{01}}(t)$$

◀

▶ **Lemma 12.** *Let $r = \mathrm{rank}_p(s)$, then*

$$\widehat{g_{01}}(s) = (-1)^r p^{-n(n+1)/2} \prod_{k=1}^{n-r} (p^k - 1).$$

Proof. By Lemma 11, we only need to choose one s for each rank to prove it. For rank r, we choose $s = \mathrm{diag}(1, \cdots, 1, 0, \cdots, 0)$, which is a diagonal matrix with r 1's in the diagonal.

We start from the most simple case of $\mathrm{rank}_p(s) = 0$, i.e. s is an all-zero matrix.

$$p^{n^2} \cdot \widehat{g_{01}}(s) = \sum_x \omega^{-\langle \mathbf{0}, x\rangle} g_{01}(x) = \sum_x g_{01}(x) = \#\text{ of full rank matrices} = \prod_{k=0}^{n-1} (p^n - p^k)$$

Then, we consider rank-1 matrix $s = \mathrm{diag}(1, 0, 0, \cdots, 0)$, which is a matrix with all zero entries except for the top left one.

For $k = 0, 1, \cdots, n-1$, we denote the k-th row of matrix x by $x(k, \cdot)$, and the submatrix from the k-th row to the last row by $x(k\text{–}, \cdot)$.

$$p^{n^2} \cdot \widehat{g_{01}}(s) = \sum_{x(0,\cdot)} \sum_{x(1\text{–},\cdot)} \omega^{-\langle s, x\rangle} g_{01}(x) = \sum_{x(0,\cdot)} \omega^{-x(0,0)} \sum_{x(1\text{–},\cdot)} g_{01}(x)$$

$\sum_{x(1\text{–},\cdot)} g_{01}(x)$ is the number of the full rank matrices given the first row. It is $\prod_{k=1}^{n-1}(p^n - p^k)$ if the first row is non-zero, and 0 if the first row is zero. Except for the case that the first row is zero, they are all canceled out because $\sum_{x(0,0)} \omega^{-x(0,0)} = 0$. Thus, the remaining is the minus value of the case that the first row is non-zero, i.e. $p^{n^2} \cdot \widehat{g_{01}}(s) = -\prod_{k=1}^{n-1}(p^n - p^k)$.

In general, for the rank-r matrix $s = \mathrm{diag}(1, \cdots, 1, 0, \cdots, 0)$ with r 1's in the diagonal,

$$p^{n^2} \cdot \widehat{g_{01}}(s) = \sum_{x(0,\cdot)} \omega^{-x(0,0)} \sum_{x(1,\cdot)} \omega^{-x(1,1)} \cdots \sum_{x(r-1,\cdot)} \omega^{-x(r-1,r-1)} \sum_{x(r\text{–},\cdot)} g_{01}(x).$$

$\sum_{x(r-,\cdot)} g_{01}(x)$ is the number of full rank matrices given the first r rows. It is $\prod_{k=r}^{n-1}(p^n - p^k)$ if the first r rows is linear independent, and 0 otherwise.

We define ζ as follows: $\zeta(x,k) = 0$ if the first k rows of x are linear dependent and $\zeta(x,k) = 1$ otherwise. Now we have

$$p^{n^2} \cdot \widehat{g_{01}}(s) = \sum_{x(0,\cdot)} \omega^{-x(0,0)} \sum_{x(1,\cdot)} \omega^{-x(1,1)} \cdots \sum_{x(r-1,\cdot)} \omega^{-x(r-1,r-1)} \zeta(x,r) \prod_{k=r}^{n-1} (p^n - p^k)$$

Then we slightly change ζ to $\bar{\zeta}$: $\bar{\zeta}(x,k) = 0$ if the first $k-1$ rows of x are linear independent but the first k rows are linear dependent, and $\bar{\zeta}(x,k) = 1$ otherwise. It is clear that $\zeta(x,r) = \prod_{k=1}^{r} \bar{\zeta}(x,k)$, which gives

$$p^{n^2} \cdot \widehat{g_{01}}(s) = \sum_{x(0,\cdot)} \omega^{-x(0,0)} \bar{\zeta}(x,1) \sum_{x(1,\cdot)} \omega^{-x(1,1)} \bar{\zeta}(x,2) \cdots \sum_{x(r-1,\cdot)} \omega^{-x(r-1,r-1)} \bar{\zeta}(x,r) \prod_{k=r}^{n-1} (p^n - p^k).$$

Since $\sum_{x(k,\cdot)} \omega^{-x(k,k)} = 0$, we have

$$\sum_{x(k,\cdot)} \omega^{-x(k,k)} \bar{\zeta}(x,k+1) = \sum_{x(k,\cdot):\bar{\zeta}(x,k+1)=1} \omega^{-x(k,k)} = -\sum_{x(k,\cdot):\bar{\zeta}(x,k+1)=0} \omega^{-x(k,k)}.$$

$x(k,\cdot)$ goes over the linear combinations of the first k rows. At the same time, $x(k,k)$ goes over the linear combinations of $x(0,k), x(1,k), \cdots, x(k-1,k)$. If $x(0,k), x(1,k), \cdots, x(k-1,k)$ are not all zero, $x(k,k)$ is balanced, so $-\sum \omega^{-x(k,k)} = 0$. If $x(0,k), x(1,k), \cdots, x(k-1,k)$ are all zero, $x(k,k) = 0$, so $-\sum \omega^{-x(k,k)} = -p^k$.

Consequently, we have

$$p^{n^2} \cdot \widehat{g_{01}}(s) = (-1)(-p)(-p^2)\cdots(-p^{r-1}) \prod_{k=r}^{n-1} (p^n - p^k) = (-1)^r p^{-n(n+1)/2} \prod_{k=1}^{n-r} (p^k - 1).$$

◀

▶ **Lemma 13.** $\|\hat{g}(s)\|_1 < 1 + 6 \cdot p^{-n} \prod_{k=1}^{n} (p^k - 1)$

Proof.

$$\begin{aligned}
\|\hat{g}(s)\|_1 &\leq 1 + 2\|\widehat{g_{01}}(s)\|_1 \\
&= 1 + 2\sum_{r=0}^{n} \sum_{s:\mathrm{rank}_p(s)=r} |\widehat{g_{01}}(s)| \\
&= 1 + 2\sum_{r=0}^{n} p^{r(r-1)/2} \binom{n}{r}_p \prod_{k=n-r+1}^{n} (p^k - 1) \cdot p^{-n(n+1)/2} \prod_{k=1}^{n-r} (p^k - 1) \\
&= 1 + 2p^{-n(n+1)/2} \prod_{k=1}^{n} (p^k - 1) \sum_{r=0}^{n} p^{r(r-1)/2} \binom{n}{r}_p \\
&= 1 + 2p^{-n(n+1)/2} \prod_{k=1}^{n} (p^k - 1) \prod_{k=0}^{n-1} (1 + p^k) \\
&= 1 + 2p^{-n} \prod_{k=1}^{n} (p^k - 1) \prod_{k=0}^{n-1} \frac{1 + p^k}{p^k} \\
&< 1 + 2p^{-n} \prod_{k=1}^{n} (p^k - 1) \cdot 3
\end{aligned}$$

◀

Now we can prove that the approximate Fourier ℓ_1-norm of g is large.

▶ **Lemma 14.**

$$\|\hat{g}\|_1^{3/2} = p^{\Omega(n^2)}$$

Proof. In the proof of Lemma 7, we know $\|\hat{h}\|_1^* = p^{n^2}\|\hat{h}\|_\infty$. We rewrite Theorem 10:

$$\|\hat{g}\|_1^\alpha = \max_{h:p^{n^2}\|\hat{h}\|_\infty \le 1} \frac{1+\alpha}{2}\langle g, h\rangle + \frac{1-\alpha}{2}\|h\|_1.$$

We choose $h = (-1)^{n+1}\widehat{g_{01}}$. So, $\hat{h} = (-1)^{n+1}g_{01}/p^{n^2}$, and $\|\hat{h}\|_\infty = 1/p^{n^2}$.

$$\begin{aligned}
\langle g_{01}, h\rangle &= \langle g_{01}, (-1)^{n+1}\widehat{g_{01}}\rangle \\
&= (-1)^{n+1}\sum_s g_{01}(s)\widehat{g_{01}}(s) \\
&= (-1)^{n+1}\sum_{s:\mathrm{rank}_p(s)=n} \widehat{g_{01}}(s) \\
&= (-1)^{n+1}\prod_{k=0}^{n-1}(p^n - p^k)\cdot(-1)^n p^{-n(n+1)/2} \\
&= -p^{-n}\prod_{k=1}^{n}(p^k - 1)
\end{aligned}$$

$$\langle g, h\rangle = -2\langle g_{01}, h\rangle + \sum_x h(x) = -2\cdot -p^{-n}\prod_{k=1}^{n}(p^k-1) + 0 = 2p^{-n}\prod_{k=1}^{n}(p^k-1)$$

$$\begin{aligned}
\|\hat{g}\|_1^{3/2} &\ge \frac{1+3/2}{2}\langle g, h\rangle + \frac{1-3/2}{2}\|h\|_1 \\
&\ge \frac{5}{4}\cdot 2p^{-n}\prod_{k=1}^{n}(p^k-1) - \frac{1}{4}\left(6p^{-n}\prod_{k=1}^{n}(p^k-1) + 1\right) \\
&= p^{-n}\prod_{k=1}^{n}(p^k-1) - \frac{1}{4} \\
&\ge p^{-n}\prod_{k=1}^{n}p^{k-1} - \frac{1}{4} \\
&= p^{n(n-3)/2} - \frac{1}{4}
\end{aligned}$$

◀

▶ **Theorem 15** (Theorem 3 Restated)**.** *The randomized/quantum communication complexity of the* Singularity *problem is* $\Omega(n^2\log p)$.

Proof. By Corollary 8, $Q_{1/6}(f) \ge \log\|\hat{g}\|_1^{3/2} - 3 = \Omega(n^2\log p)$. $R_{1/6}(f) = \Omega(Q_{1/6}(f))$. ◀

4 Lower Bound for Determinant of Non-singular Matrix

We use a small spectral norm of the matrix representing the $\mathrm{Det}_{a,b}$ problem to prove the communication complexity lower bound. To prove the lower bound of spectral norm, we decompose the matrix onto Fourier basis, because the Fourier basis has good properties and a small spectral norm.

4.1 Spectral Norm Method

In the previous section, a large trace norm implies a large communication complexity lower bound. However, here we use a small spectral norm to prove a large communication complexity lower bound.

The spectral norm method is based on the discrepancy method, which can derive the communication complexity lower bound by giving an upper bound for a value called discrepancy defined below.

▶ **Definition 16** (Discrepancy). Let $f : \mathbb{F}_p^N \times \mathbb{F}_p^N \mapsto \{0,1\}$ be a function, $S \times T$ be a rectangle, and μ be a probability distribution on $\mathbb{F}_p^N \times \mathbb{F}_p^N$. Denote $\mathrm{disc}_\mu(S \times T, f) = |\sum_{(x,y)\in S\times T} \mu(x,y)(-1)^{f(x,y)}|$, and $\mathrm{disc}_\mu(f) = \max_{S,T\subseteq\mathbb{F}_p^N} \mathrm{disc}_\mu(S \times T, f)$.

The discrepancy is widely used in proving communication complexity lower bound [1, 15, 7], with many applications. It was also used to prove the quantum lower bound [6, 13], and could be phrased in the following theorem.

▶ **Theorem 17.** *[6] For any function f and any distribution μ, we have*

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

Furthermore, the discrepancy can be bounded by the spectral norm.

▶ **Theorem 18.** *[7, Example 3.29] Let $f : \mathbb{F}_p^N \times \mathbb{F}_p^N \mapsto \{0,1,\perp\}$ be a partial Boolean function. We define the corresponding partial sign matrix F by its entries*

$$F(x,y) = \begin{cases} 1 & \text{if } f(x,y) = 0, \\ -1 & \text{if } f(x,y) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

For the uniform distribution μ on the defined inputs of f, we have $\mathrm{disc}_\mu(f) \le p^N \cdot \|F\|/\|F\|_1$.

4.2 Fourier Basis Matrix

For the determinant problem, it is difficult to compute $\|F\|$ directly. We will decompose F into Fourier basis: $F = \sum_k \lambda_k H_k$. The spectral norm of the Fourier basis H_k is easier to compute. At last, we will use the triangle inequality to bound $\|F\|$.

▶ **Definition 19** (Discrete logarithm). $\mathbb{F}_p^* = \mathbb{F}_p \setminus \{0\}$ is a cyclic multiplicative group, in which 2 is a primitive element (the generator for the multiplicative group). For $a \in \mathbb{F}_p^*$, we say $k = \log_2 a$ if $2^k = a$.

Let $\eta = e^{2\pi i/(p-1)}$. For $k = 0, 1, \cdots, p-2$ and $a \in \mathbb{F}_p$, we define a family of τ as below.

$$\tau_k(a) = \begin{cases} 0 & \text{if } a = 0, \\ \eta^{k \log_2 a} & \text{otherwise.} \end{cases}$$

▶ **Lemma 20.** *For $a, b \in \mathbb{F}_p$,*

1. $\tau_k(a)\tau_k(b) = \tau_k(ab)$;
2. $\sum_{a=0}^{p-1} \tau_k(a) = 0$, *if* $k \neq 0$;
3. $\tau_k(a)\tau_k(a)^*\tau_k(a) = \tau_k(a)$.

Proof. 1. If $a = 0$ or $b = 0$, both sides are 0. If $a \neq 0$ and $b \neq 0$, $\tau_k(a)\tau_k(b) = \eta^{k\log_2 a}\eta^{k\log_2 b} = \eta^{k\log_2(ab)} = \tau_k(ab)$.

2.
$$\sum_{a=0}^{p-1}\tau_k(a) = \sum_{a=1}^{p-1}\tau_k(a) = \sum_{a=1}^{p-1}\eta^{k\log_2 a} = \sum_{l=0}^{p-2}\eta^{kl} = 0$$

3. If $a = 0$, both sides are 0. If $a \neq 0$, $\tau_k(a)^*\tau_k(a) = |\tau_k(a)|^2 = 1$. ◀

Then, we define a family of h_k. For $x \in \mathbb{F}_p^{n\times n}$, we define $h_k(x) = \tau_k(\det_p(x))$.

▶ **Lemma 21.** *For $x, y \in \mathbb{F}_p^{n\times n}$,*
1. *$h_k(x)h_k(y) = h_k(xy)$;*
2. *$h_k(x)h_k(x)^*h_k(x) = h_k(x)$.*

Proof.
1. $h_k(x)h_k(y) = \tau_k(\det_p(x))\tau_k(\det_p(y)) = \tau_k(\det_p(x)\det_p(y)) = \tau_k(\det_p(xy)) = h_k(xy)$;
2. $h_k(x)h_k(x)^*h_k(x) = \tau_k(\det_p(x))\tau_k(\det_p(x))^*\tau_k(\det_p(x)) = \tau_k(\det_p(x)) = h_k(x)$. ◀

▶ **Lemma 22.** *For $t \in \mathbb{F}_p^{n\times n}$, $\hat{h_k}(t) = 0$ if the first row of t is all-zero and $k \neq 0$.*

Proof.
$$\hat{h_k}(t) = \sum_x h(x)\omega^{\langle x,t\rangle} = \sum_{x(1-,\cdot)}\omega^{\langle x(1-,\cdot),t(1-,\cdot)\rangle}\sum_{x(0,\cdot)}h(x)$$

If $x(1-,\cdot)$ is fixed and $x(0,\cdot)$ goes over $\mathbb{F}_p^n$, $\det_p(x)$ is balanced on all non-zero values. Thus, $\sum_{x(0,\cdot)} h(x) = 0$. ◀

▶ **Lemma 23.** *For $w, t \in \mathbb{F}_p^{n\times n}$, if $\det(w) \neq 0$, then $\hat{h_k}(wt) = h_k(w)^*\hat{h_k}(t)$.*

Proof.
$$\begin{aligned}
\hat{h_k}(wt) &= h_k(w)^*h_k(w)\hat{h_k}(wt)\\
&= h_k(w)^*h_k(w)\frac{1}{p^{n^2}}\sum_x h_k(x)\omega^{-\langle x,wt\rangle}\\
&= h_k(w)^*h_k(w^{\mathrm{T}})\frac{1}{p^{n^2}}\sum_x h_k(x)\omega^{-\langle w^{\mathrm{T}}x,t\rangle}\\
&= h_k(w)^*\frac{1}{p^{n^2}}\sum_x h_k(w^{\mathrm{T}}x)\omega^{-\langle w^{\mathrm{T}}x,t\rangle}\\
&= h_k(w)^*\hat{h_k}(t)
\end{aligned}$$

◀

▶ **Lemma 24.** *For $t \in \mathbb{F}_p^{n\times n}$, if $\det(t) = 0$ and $k \neq 0$, then $\hat{h_k}(t) = 0$.*

Proof. Because t is singular, we can find an invertible matrix w such that the first row of wt is all zero. By Lemma 22 and Lemma 23, $\hat{h_k}(t) = \hat{h_k}(wt)/h_k(w)^* = 0/h_k(w)^* = 0$. ◀

▶ **Lemma 25.** *For $k \neq 0$ and $t \in \mathbb{F}_p^{n\times n}$, $\hat{h_k}(t) = \hat{h_k}(I)h_k(t)^*$, where I is the identity matrix of size $n \times n$.*

Proof. If $\det(t) = 0$, $\hat{h_k}(t) = 0 = \hat{h_k}(I)\cdot 0 = \hat{h_k}(I)h_k(t)^*$.
If $\det(t) \neq 0$, $\hat{h_k}(t) = \hat{h_k}(t\cdot I) = h_k(t)^*\hat{h_k}(I)$. ◀

▶ **Lemma 26.** *For $k \neq 0$, $\hat{h_k}(I)^*\hat{h_k}(I) = p^{-n^2}$.*

Proof.

$$\begin{aligned}
\langle \hat{h_k}, \hat{h_k} \rangle &= \sum_x \hat{h_k}(x)^* \hat{h_k}(x) \\
&= \sum_x \left(\hat{h_k}(I) h_k(x) \right)^* \left(\hat{h_k}(I) h_k(x) \right) \\
&= \hat{h_k}(I)^* \hat{h_k}(I) \sum_x h_k(x)^* h_k(x) \\
&= \hat{h_k}(I)^* \hat{h_k}(I) \langle h_k, h_k \rangle \\
&= \hat{h_k}(I)^* \hat{h_k}(I) \cdot p^{n^2} \langle \hat{h_k}, \hat{h_k} \rangle
\end{aligned}$$

Therefore, $\hat{h_k}(I)^* \hat{h_k}(I) = p^{-n^2}$. ◀

At last, we define a family of matrix H_k. For $x, y \in \mathbb{F}_p^{n \times n}$, we define $H_k(x, y) = h_k(x + y)$.

▶ **Lemma 27.** *For $k \neq 0$, $H_k H_k^\dagger H_k = p^{n^2} \cdot H_k$.*

Proof. For $w, x, y, z, r, s, t \in \mathbb{F}_p^{n \times n}$,

$$\begin{aligned}
& (H_k H_k^\dagger H_k)(w, z) \\
= & \sum_x \sum_y H_k(w, x) H_k(y, x)^* H_k(y, z) \\
= & \sum_x \sum_y h_k(w + x) h_k(x + y)^* h_k(y + z) \\
= & \sum_x \sum_y \left(\sum_r \hat{h_k}(r) \omega^{\langle r, w+x \rangle} \right) \left(\sum_s \hat{h_k}(s) \omega^{\langle s, x+y \rangle} \right)^* \left(\sum_t \hat{h_k}(t) \omega^{\langle t, y+z \rangle} \right) \\
= & \sum_r \sum_s \sum_t \hat{h_k}(r) \hat{h_k}(s)^* \hat{h_k}(t) \omega^{\langle r, w \rangle} \left(\sum_x \omega^{\langle r-s, x \rangle} \right) \left(\sum_y \omega^{\langle -s+t, y \rangle} \right) \omega^{\langle t, z \rangle} \\
= & \sum_r \sum_s \sum_t \hat{h_k}(r) \hat{h_k}(s)^* \hat{h_k}(t) \omega^{\langle r, w \rangle} \left(p^{n^2} \delta_{r,s} \right) \left(p^{n^2} \delta_{s,t} \right) \omega^{\langle t, z \rangle} \\
= & p^{2n^2} \sum_r \hat{h_k}(r) \hat{h_k}(r)^* \hat{h_k}(r) \omega^{\langle r, w \rangle} \omega^{\langle r, z \rangle} \\
= & p^{2n^2} \sum_r (\hat{h_k}(I) h_k(r)^*)(\hat{h_k}(I)^* h_k(r))(\hat{h_k}(I) h_k(r)^*) \omega^{\langle r, w+z \rangle} \\
= & p^{2n^2} \sum_r \hat{h_k}(I) \hat{h_k}(I)^* \hat{h_k}(I) h_k(r)^* h_k(r) h_k(r)^* \omega^{\langle r, w+z \rangle} \\
= & p^{2n^2} \sum_r p^{-n^2} \hat{h_k}(I) h_k(r)^* \omega^{\langle r, w+z \rangle} \\
= & p^{n^2} \sum_r \hat{h_k}(r) \omega^{\langle r, w+z \rangle} \\
= & p^{n^2} h_k(w + z) \\
= & p^{n^2} H_k(w, z)
\end{aligned}$$

◀

▶ **Lemma 28.** *For $k \neq 0$,*

$$\|H_k\| = p^{n^2/2}.$$

Proof. We denote the largest eigenvalue of a semi-definite matrix A by $\max\operatorname{eval}(A)$. Thus, $\|A\| = \sqrt{\max\operatorname{eval}(A^\dagger A)}$.

$$\begin{aligned}\|H_k H_k^\dagger H_k\| &= \sqrt{\max\operatorname{eval}((H_k H_k^\dagger H_k)^\dagger \cdot H_k H_k^\dagger H_k)} \\ &= \sqrt{\max\operatorname{eval}((H_k^\dagger H_k)^3)} \\ &= \left(\sqrt{\max\operatorname{eval}(H_k^\dagger H_k)}\right)^3 \\ &= \|H_k\|^3\end{aligned}$$

Comparing to Lemma 27, we have $\|H_k\|^3 = p^{n^2}\|H_k\|$. Therefore, $\|H_k\| = p^{n^2/2}$. ◀

4.3 Lower Bound for Det

For $\textsc{Det}_{a,b}$, $f(x,y)$ is 0 if $\det_p(x+y) = a$, 1 if $\det_p(x+y) = b$ and undefined for other cases. We define the corresponding partial sign matrix F as below.

$$F(x,y) = \begin{cases} 1 & \text{if } \det_p(x+y) = a, \\ -1 & \text{if } \det_p(x+y) = b, \\ 0 & \text{otherwise.} \end{cases}$$

▶ **Lemma 29.** $\|F\|_1 = \Omega(p^{2n^2-1})$.

Proof. In the uniform distribution, $\Pr_{x\in\mathbb{F}_p^{n\times n}}[\det_p(x) \neq 0] = \prod_{k=1}^{n}(1-p^{-k}) > \phi(p^{-1}) > 0$, where ϕ is the Euler function. This means that the density of the non-singular matrix is greater than a constant. Furthermore, the determinant is balanced on all non-zero values. Thus, $\Pr_{x\in\mathbb{F}_p^{n\times n}}[\det_p(x) = a \text{ or } b] > \frac{2}{p-1}\phi(p^{-1})$. Finally, $\|F\|_1 \geq \frac{2}{p-1}\phi(p^{-1})\cdot p^{2n^2} = \Omega(p^{2n^2-1})$. ◀

▶ **Lemma 30.** $\|F\| \leq 2p^{n^2/2}$.

Proof. It is easy to check that we can decompose F to the Fourier basis matrices:

$$F = \frac{1}{p}\cdot\sum_{k=1}^{p-2}\left(\eta^{-k\log_2 a} - \eta^{-k\log_2 b}\right)H_k$$

. Then, we use the triangle inequality of matrix norm to bound the norm of F:

$$\|F\| \leq \frac{1}{p}\cdot\sum_{k=1}^{p-2}\left|\eta^{-k\log_2 a} - \eta^{-k\log_2 b}\right|\|H_k\| \leq \frac{1}{p}\cdot(p-2)\cdot 2\cdot p^{n^2/2} < 2p^{n^2/2}.$$

◀

▶ **Theorem 31** (Theorem 4 Restated)**.** *The randomized/quantum communication complexity of* $\textsc{Det}_{a,b}$ *is* $\Omega(n^2\log p)$.

Proof.

$$\operatorname{disc}_\mu(f) \leq p^{n^2}\frac{\|F\|}{\|F\|_1} \leq p^{n^2}\cdot\frac{2p^{n^2/2}}{\Omega(p^{2n^2-1})} = O(p^{-n^2/2+1})$$

$$R_{1/3}(f) = \Omega(Q^*_{1/3}(f)),\ Q^*_{1/3}(f) = \Omega\left(\log\frac{1}{\operatorname{disc}_\mu(f)}\right) = \Omega(n^2\log p).$$

◀

5 Open Problems

One open problem is to distinguish between $\det_p(x) = 0$ and $\det_p(x) = a$. We guess it has an $\Omega(n^2 \log p)$ lower bound even for quantum protocols. The proof could be similar to Section 3.

The other one is to compute the (i, j)-th element of the inverse of matrix x. We conjecture the quantum communication $\Omega(n^2 \log p)$ as well. Actually, this problem is as hard as solving linear equations.

We discuss all these problems over $\mathbb{F}_p$, but we think they are still hard over integers.

References

1 László Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, pages 337–347, Washington, DC, USA, 1986. IEEE Computer Society.

2 Harry Buhrman and Ronald de Wolf. Communication complexity lower bounds by polynomials. In *IEEE Conference on Computational Complexity*, pages 120–130, 2001.

3 J. I. Chu and G. Schnitger. Communication complexity of matrix computation over finite fields. *Theory of Computing Systems*, 28:215–228, 1995. 10.1007/BF01303056.

4 Kenneth L. Clarkson and David P. Woodruff. Numerical linear algebra in the streaming model. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, STOC '09, pages 205–214, New York, NY, USA, 2009. ACM.

5 Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge Univ Pr, 1990.

6 I. Kremer. *Quantum communication*. PhD thesis, Citeseer, 1995.

7 E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge Univ Pr, 1997.

8 Troy Lee and Adi Shraibman. Lower bounds in communication complexity: A survey.

9 Zhi-Quan Luo and John N. Tsitsiklis. On the communication complexity of distributed algebraic computation. *J. ACM*, 40:1019–1047, November 1993.

10 Peter Bro Miltersen, Noam Nisan, Shmuel Safra, and Avi Wigderson. On data structures and asymmetric communication complexity. In *Proceedings of the 27th annual ACM symposium on Theory of computing*, STOC '95, pages 103–111, New York, NY, USA, 1995. ACM.

11 Ran Raz. Fourier analysis for probabilistic communication complexity. *Computational Complexity*, 5:205–221, 1995. 10.1007/BF01206318.

12 AA Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya: Mathematics*, 67:145–159, 2003.

13 Alexander A. Sherstov. The pattern matrix method for lower bounds on quantum communication. In *Proceedings of the 40th annual ACM symposium on Theory of computing*, STOC '08, pages 85–94, New York, NY, USA, 2008. ACM.

14 Yaoyun Shi and Yufan Zhu. Quantum communication complexity of block-composed functions. *Quantum Information and Computation*, 9:444–460, May 2009.

15 Andrew C. Yao. Lower bounds by probabilistic arguments. In *Proceedings of the 24th Annual Symposium on Foundations of Computer Science*, pages 420–428, Washington, DC, USA, 1983. IEEE Computer Society.

16 Andrew Chi-Chih Yao. Some complexity questions related to distributive computing(preliminary report). In *Proceedings of the 7th annual ACM symposium on Theory of computing*, STOC '79, pages 209–213, New York, NY, USA, 1979. ACM.

17 Zhiqiang Zhang and Yaoyun Shi. Communication complexities of symmetric xor functions. *Quantum Information and Computation*, 9:255–263, March 2009.

18 Zhiqiang Zhang and Yaoyun Shi. On the parity complexity measures of boolean functions. *Theoretical Computer Science*, 411(26-28):2612–2618, 2010.