Quantum communication complexity and nonlocality of bipartite quantum operations

by

Yufan Zhu

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Computer Science and Engineering) in The University of Michigan 2008

Doctoral Committee:
Assistant Professor Yaoyun Shi, Chair
Professor John P. Hayes
Associate Professor Kevin J. Compton
Associate Professor Luming Duan
Associate Professor Igor L. Markov
ACKNOWLEDGEMENTS

This dissertation would not have been written without the support and guidance of my advisor, Prof. Yaoyun Shi, who gave me great freedom on research and always encouraged me to work on any problems that I am interested in. He has provided me extensive personal and professional guidance, and taught me a great deal about both scientific research and life in general. He and other faculty members that served on my dissertation committee, Prof. Kevin Compton, Prof. Luming Duan, Prof. John Hayes and Prof. Igor Markov patiently guided me through the dissertation process, never accepting less than my best efforts. I thank them all.

I’m grateful to all of those whom I had the pleasure to work with during my PhD program. Special thanks to co-authors Jianxin Chen, Wei Huang, and Shengyu Zhang for thoughtful discussions, and my officemates Johnathan Brown, Ye Du, Sindhu Kuttty, Julia Lipman, Xiaolin Shi, Denny VandenBerg, and Xuan Zheng etc. for discussions about research, theory, and other topics.

My life would not be so colorful without my friends in Ann Arbor, to name a few, Sang-Jung Han, Sijian Wang, Liang Zhang etc. Thanks to my roommate, Yin Wang, who organized the numerous ski trips in the winter, from which I got refreshed and become energetic towards my research. I also wish to thank my good friends from my undergraduate days at Beijing University, especially Haobo Peng and Meng Yang, who provide unending support and inspirations over the years.

Nobody has been more important to me in the pursuit of my degree than my
brother and my parents. Their love and support are with me whatever I pursue.

Lastly, I should thank many individuals, friends and colleagues who have not been mentioned here personally in making this Ph.D process a success. I could not have made it without your supports.

This dissertation is supported by NSF Grant No. 0323555, 0347078, 0622033.
# TABLE OF CONTENTS

ACKNOWLEDGEMENTS .......................................................... ii
LIST OF FIGURES .............................................................. vi
LIST OF SYMBOLS ............................................................. vii
ABSTRACT .............................................................................. ix

CHAPTER

I. Introduction ........................................................................ 1
   1.1 Backgrounds ............................................................... 1
   1.2 Quantum mechanics .................................................... 3
   1.2.1 Quantum states ....................................................... 3
   1.2.2 Tensor products ...................................................... 3
   1.2.3 Quantum bits ......................................................... 4
   1.2.4 Quantum measurements .......................................... 5
   1.2.5 Unitary transformations .......................................... 6
   1.2.6 Density operators ................................................... 8
   1.2.7 Superoperators ...................................................... 9
   1.3 Communication complexity ........................................ 10
   1.3.1 Classical communication complexity ......................... 10
   1.3.2 Quantum communication complexity ....................... 12
   1.4 Quantum entanglement .............................................. 13
   1.5 Nonlocality of quantum operations .............................. 15
   1.5.1 Classical simulations of bipartite quantum measurements. 15
   1.5.2 The maximum tensor norm of bipartite superoperators. .... 17
   1.6 Organization ............................................................. 18

II. The communication complexity of the Hamming Distance problem .... 19
   2.1 Introduction and Summary of results ............................ 19
   2.2 Lower bound for the quantum communication complexity of the HAMMING DISTANCE problem ............................................. 22
   2.3 Upper bound for the classical communication complexity of the HAMMING DISTANCE problem ............................................. 23

III. The communication complexity of block-composed functions ........ 27
   3.1 Introduction and summary of results ............................. 27
   3.2 Preliminaries ............................................................ 32
   3.2.1 Communication complexities and quantum lower bound by approximate trace norm ................................................. 32
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2.2</td>
<td>Approximate polynomial degree</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>The Main Lemma</td>
<td>34</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Witness of high approximate degree</td>
<td>34</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Witness of large approximate trace norm</td>
<td>36</td>
</tr>
<tr>
<td>3.4</td>
<td>Applications</td>
<td>38</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Composition with hard $g_k$</td>
<td>39</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Composition with SET DISJOINTNESS</td>
<td>40</td>
</tr>
<tr>
<td>IV.</td>
<td>Classical simulations of nonlocal quantum measurements</td>
<td>46</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of results</td>
<td>46</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Applications on quantum communication complexity</td>
<td>47</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Applications on simulating quantum correlations</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>A simulation framework</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>The main theorem</td>
<td>51</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Classical simulation of quantum measurements</td>
<td>52</td>
</tr>
<tr>
<td>4.3.2</td>
<td>The diamond norm on bipartite operators</td>
<td>53</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Upper bounding Com($Q$) by the diamond norm</td>
<td>56</td>
</tr>
<tr>
<td>4.4</td>
<td>Applications</td>
<td>58</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Quantum SMP with shared entanglement</td>
<td>58</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Two way interactive quantum communication with shared entanglement</td>
<td>59</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Simulating quantum correlations</td>
<td>60</td>
</tr>
<tr>
<td>V.</td>
<td>The maximum tensor norm of bipartite superoperators</td>
<td>63</td>
</tr>
<tr>
<td>5.1</td>
<td>Summary of results</td>
<td>63</td>
</tr>
<tr>
<td>5.2</td>
<td>Nonlocality criteria for superoperators</td>
<td>65</td>
</tr>
<tr>
<td>5.3</td>
<td>Maximum tensor norm of elementary superoperators</td>
<td>67</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Maximum tensor norm of CNOT</td>
<td>67</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Maximum tensor norm of SWAP</td>
<td>69</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Maximum tensor norm of measuring one qubit and sending the result</td>
<td>72</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Maximum tensor norm of sending one quantum bit</td>
<td>73</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Maximum tensor norm of a measurement operator</td>
<td>75</td>
</tr>
<tr>
<td>5.4</td>
<td>Connections with communication complexity</td>
<td>76</td>
</tr>
<tr>
<td>VI.</td>
<td>Conclusions</td>
<td>78</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary of this dissertation</td>
<td>78</td>
</tr>
<tr>
<td>6.2</td>
<td>Future directions</td>
<td>80</td>
</tr>
<tr>
<td>APPENDIX</td>
<td></td>
<td>83</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td></td>
<td>89</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The model of classical communication complexity</td>
<td>10</td>
</tr>
<tr>
<td>1.2</td>
<td>The model of quantum communication complexity</td>
<td>12</td>
</tr>
<tr>
<td>1.3</td>
<td>Bipartite quantum measurements</td>
<td>16</td>
</tr>
<tr>
<td>1.4</td>
<td>Classical simulations of bipartite quantum measurements</td>
<td>16</td>
</tr>
<tr>
<td>2.1</td>
<td>The Simultaneously Message Passing (SMP) model</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>The classical decision tree model</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>Symmetric predicates</td>
<td>29</td>
</tr>
<tr>
<td>3.3</td>
<td>Block-composed functions</td>
<td>30</td>
</tr>
</tbody>
</table>
### LIST OF SYMBOLS

- \( f(n) = O(g(n)) \)  
  - there exist some numbers \( n_0, c > 0 \) such that for all \( n > n_0, f(n) < cg(n) \)

- \( f(n) = \Theta(g(n)) \)  
  - there exist some numbers \( n_0, c_1, c_2 > 0 \) such that for all \( n > n_0, c_1g(n) < f(n) < c_2g(n) \)

- \( f(n) = \Omega(g(n)) \)  
  - there exist some numbers \( n_0, c > 0 \) such that for all \( n > n_0, f(n) > cg(n) \)

- \( \{0, 1\}^n \)  
  - the set of all binary strings with length \( n \)

- \( \wedge \)  
  - set intersection

- \( \cap \)  
  - tensor products (p. 4)

- \( D(f) \)  
  - the deterministic communication complexity of \( f \) (p. 11)

- \( R(f) \)  
  - the randomized communication complexity of \( f \), with private coin (p. 11)

- \( R^{\text{pub}}(f) \)  
  - the randomized communication complexity of \( f \), with private coin (p. 11)

- \( Q(f) \)  
  - the quantum communication complexity of \( f \) (p. 12)

- \( \text{Com}(Q) \)  
  - the classical communication complexity of measurement \( Q \) (p. 16)

- \( R(f) \)  
  - the randomized communication complexity of \( f \) in the Simultaneous Message Passing model (p. 19)

- \( R^{\text{pub}}(f) \)  
  - the randomized communication complexity of \( f \) in the Simultaneous Message Passing model, with public coins (p. 19)

- \( Q(f) \)  
  - the quantum communication complexity of \( f \) in the SMP model (p. 20)

- \( Q^{\parallel}(f) \)  
  - the quantum communication complexity of \( f \) in the SMP model with unlimited entanglement (p. 20)

- \( \text{HAM}_{n,d} \)  
  - the Hamming Distance problem (p. 20)

- \( \text{DISJ}_n \)  
  - the Set Disjointness problem (p. 22)

- \( f_n \sqcap g_k \)  
  - the block composition of \( f_n \) and \( g_k \) (p. 29)

- \( \text{IP}_n \)  
  - the Inner Product function (p. 30)

- \( \|A\|_{\text{tr}} \)  
  - the approximate trace norm of \( A \) (p. 32)

- \( \deg(f) \)  
  - the approximate degree of \( f \) (p. 33)

- \( |\phi\rangle \)  
  - column vector (p. 84)

- \( \langle \phi | \)  
  - row vector (p. 84)

- \( \langle \phi | \psi \rangle \)  
  - the inner product of \( |\phi\rangle \) and \( |\psi\rangle \) (p. 84)

- \( \mathbb{L}(\mathcal{N}) \)  
  - the space of linear operators on \( \mathcal{N} \)

- \( \mathbb{L}(\mathcal{N}, \mathcal{M}) \)  
  - the space of linear operators from \( \mathcal{N} \) to \( \mathcal{M} \)

- \( A^\dagger \)  
  - the adjoint of operator \( A \) (p. 85)

- \( \text{tr}(A) \)  
  - the trace of matrix \( A \) (p. 85)

- \( \|A\| \)  
  - the operator norm of \( A \) (p. 86)
<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$|A|_{tr}$</td>
<td>the trace norm of $A$ (p. 86)</td>
</tr>
<tr>
<td>$tr_\mathcal{B}(\cdot)$</td>
<td>the partial trace over space $\mathcal{B}$ (p. 86)</td>
</tr>
<tr>
<td>$|\cdot|_\diamond$</td>
<td>the diamond norm (p. 86)</td>
</tr>
<tr>
<td>$|\cdot|_\gamma$</td>
<td>the maximum tensor norm (p. 88)</td>
</tr>
</tbody>
</table>
ABSTRACT

This dissertation is motivated by the following fundamental questions: (a) are there any exponential gaps between quantum and classical communication complexities? (b) what is the role of entanglement in assisting quantum communications? (c) how to characterize the nonlocality of quantum operations? We study four specific problems below.

The communication complexity of the Hamming Distance problem. The Hamming Distance problem is for two parties to determine whether or not the Hamming distance between two $n$-bit strings is more than a given threshold. We prove tighter quantum lower bounds in the general two-party, interactive communication model. We also construct an efficient classical protocol in the more restricted Simultaneous Message Passing model, improving previous results.

The Log-Equivalence Conjecture. A major open problem in communication complexity is whether or not quantum protocols can be exponentially more efficient than classical ones for computing a total Boolean function in the two-party, interactive model. The answer is believed to be “No”. Razborov proved this conjecture for the most general class of functions so far. We prove this conjecture for a broader class of functions that we called block-composed functions. Our proof appears to be the first demonstration of the dual approach of the polynomial method in proving new results.

Classical simulations of bipartite quantum measurement. We define a new
concept that measures the nonlocality of bipartite quantum operations. From this measure, we derive an upper bound that shows the limitation of entanglement in reducing communication costs. As applications, we show that (a) if the amount of communication is constant, quantum and classical communication protocols with an unlimited amount of shared entanglement or shared randomness compute the same set of functions; (b) a local hidden variable model needs only a constant amount of communication to create, within an arbitrarily small statistical distance, a distribution resulting from local measurements of an entangled quantum state, as long as the number of measurement outcomes is constant.

**The maximum tensor norm of bipartite superoperators.** We define a maximum tensor norm to quantify the nonlocality of bipartite superoperators. We show that a bipartite physically realizable superoperator is bi-local if and only if its maximum tensor norm is 1. Furthermore, the estimation of the maximum tensor norm can also be used to prove quantum lower bounds on communication complexities.
CHAPTER I

Introduction

1.1 Backgrounds

The modern theory of quantum mechanics, discovered at the beginning of the twentieth century, describes a different universe from that of classical physics. One counter-intuitive postulate in the quantum world is that a quantum system can be in a superposition of many different classical states, and may exhibit interference during its evolutions. Another marvellous property is that spatially separated quantum systems can share entanglement, which could display “nonlocal” effects.

Quantum computation and quantum information is the field that investigates the information processing power of systems built upon quantum physics (e.g.,[70, 74]). One important objective of the field is to find problems that quantum computers can solve significantly faster than classical computers. Benioff [13] and Manin [65] are probably the first to introduce the idea of building quantum computers based on quantum mechanics in 1980. Two years later, Richard Feynman [46, 47] suggested developing quantum computers to simulate quantum mechanical systems, since there seems to be forbidding difficulties to do the simulations on classical computers. A formal model of the universal quantum Turing machine was soon defined in 1985 by David Deutsch [40]. In the same paper, Deutsch showed that a quantum computer
can solve a problem faster than classical computers. The results of Deutsch were improved in the subsequent decade. Among them, Bernstein and Vazirani [18] formalized quantum complexity theory; Deutsch and Jozsa [41] showed that a quantum computer can solve some problems exponentially faster than a classical computer. The most important discovery in the field so far is probably Peter Shor’s demonstration that two important problems — the problem of integer factorization, and the problem of finding discrete logarithms — could be solved efficiently on a quantum computer [87]. These two problems are believed, though not proved yet, to have no efficient solutions.

Another important objective of the field is to understand the power of quantum communication between spatially separated parties. One direction is to study the amount of information that can be transmitted over a certain quantum communication channel. In 1973, Holevo [53] proved that, by sending a single quantum bit, one party can transmit only one bit of information to the other party. However, if two parties have shared entanglement, two bits of information can be transmitted, using superdense encoding discovered by Bennett and Wiesner [17] in 1992. Moreover, superdense encoding was proved to be optimal by [34]. On the other hand, Bennett et al. [15] discovered a mechanism, now called quantum teleportation, to transfer one quantum bit using shared entanglement and transmission of two classical bits. This mechanism has been verified by many different physical experiments, e.g., [20].

Another direction is to investigate the quantum communication complexity (or distributed quantum computation), which is the minimum amount of communication that is required for two spatially separated quantum computers to solve a particular problem together. Quantum communication complexities are the main subjects of this dissertation. We shall explain related concepts rigorously in subsequent sections:
basic notations of quantum mechanics in section 1.2, the quantum communication complexity models in section 1.3, the concept of quantum entanglement in section 1.4, and nonlocality of quantum operations in 1.5.

1.2 Quantum mechanics

We introduce quantum mechanics that will be used in this dissertation. We follow notations in [70, 59]. For linear algebra, we refer to the Appendix.

1.2.1 Quantum states

Consider a system with $k$ different classical states, for example, an electron with a ground state and $k-1$ different excited states. Classical physics asserts that the electron must be in one of these $k$ states. However, in quantum mechanics, the electron can also exist in a superposition of these $k$ classical states, i.e., the electron can exist in these $k$ different classical states simultaneously!

Let $\mathbb{C}^k$ denote the $k$ dimensional complex Hilbert space. Mathematically, we represent the above $k$ classical states as vectors $|0\rangle, |1\rangle, \ldots, |k-1\rangle$. Then according to quantum mechanics, any quantum state (usually called state) of the above system can be represented as a unit vector $|\phi\rangle \in \mathbb{C}^k$,

$$|\phi\rangle = \alpha_0|0\rangle + \alpha_1|1\rangle + \ldots + \alpha_{k-1}|k-1\rangle,$$

where $\alpha_0, \alpha_1, \ldots, \alpha_{k-1}$ are complex numbers and $\sum_i |\alpha_i|^2 = 1$. The complex number $\alpha_i$ is also called the amplitude for $|i\rangle$.

1.2.2 Tensor products

Suppose we have two quantum systems $A$ and $B$. Then what is their composite system? We will use the concept of tensor products to describe it. Suppose vector space $\mathcal{N}_A$ has dimension $k$ with basis $\{|i\rangle_A : 0 \leq i \leq k-1\}$, and vector space $\mathcal{N}_B$
has dimension \(l\) with basis \(\{|j\}_B : 0 \leq j \leq l - 1\). The tensor product of \(\mathcal{N}_A\) and \(\mathcal{N}_B\) is a \(kl\) dimensional vector space \(\mathcal{N} = \mathcal{N}_A \otimes \mathcal{N}_B\) with basis \(\{|i\}_A \otimes |j\}_B : 0 \leq i \leq k - 1, 0 \leq j \leq l - 1\). The elements of \(\mathcal{N}\) are linear combinations of \(\{|i\}_A \otimes |j\}_B\).

Assume system \(A\) has \(k\) and system \(B\) has \(l\) different classical states respectively. Then the state space of system \(A\) can be represented as \(\mathbb{C}^k\), and the state space of system \(B\) can be represented as \(\mathbb{C}^l\). The state space of their composite system can be represented as the Hilbert space \(\mathbb{C}^{kl} = \mathbb{C}^k \otimes \mathbb{C}^l\). For example, if the state of system \(A\) is \(|\phi\rangle = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{k-1} \end{pmatrix} \in \mathbb{C}^k\) and the state of system \(B\) is \(|\psi\rangle = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{l-1} \end{pmatrix} \in \mathbb{C}^l\), then the state of the composite system is the \(kl\) dimensional vector \(|\phi\rangle \otimes |\psi\rangle = \begin{pmatrix} \alpha_0 \beta_0 \\ \alpha_0 \beta_1 \\ \vdots \\ \alpha_{k-1} \beta_{l-1} \end{pmatrix} \in \mathbb{C}^k \otimes \mathbb{C}^l\), where the \((i \times l + j)\)th entry of \(|\phi\rangle \otimes |\psi\rangle\) is \(\alpha_i \beta_j\).

A quantum state on system \(AB\) is said to be separable if and only if it can be written as the tensor product of a state on \(A\) and a state on \(B\). Otherwise, the quantum state is entangled. One example of entangled states is the Bell state (or called EPR pairs) \(\frac{1}{\sqrt{2}}(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B\)). Quantum entanglement plays a key role in quantum information processing. We shall discuss it further in section 1.4.

1.2.3 Quantum bits

In classical computation, the unit of information is a bit, which can be either 0 or 1. In quantum computation, the unit is a quantum bit, usually called qubit. The
The state of a single qubit system can be represented as a unit vector $|\phi\rangle \in \mathbb{C}^2$,

$$|\phi\rangle = \alpha|0\rangle + \beta|1\rangle.$$  

In general, an $n$-qubit system can be regarded as the composition of $n$ single-qubit systems, represented by Hilbert space $\mathbb{C}^{2^n} = \mathbb{C}^2 \otimes \mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2$. The state of such a composite system can be represented as a unit vector $|\phi\rangle \in \mathbb{C}^{2^n}$,

$$|\phi\rangle = \alpha_x|x\rangle, \quad x \in \{0, 1\}^n,$$

where $\alpha_x$ are complex numbers and $\sum_x |\alpha_x|^2 = 1$.

### 1.2.4 Quantum measurements

Notice that the state of an $n$-qubit quantum system has $2^n$ complex amplitudes. So it “appears” to contain exponentially many bits of information. However, we can not “observe” those $2^n$ complex amplitudes directly. Quantum mechanics only allows two kinds of operations on quantum states: quantum measurement and unitary operations. We describe quantum measurements in this section and unitary operations in the next one.

The simplest form of quantum measurement is the projective measurement performed in the computational basis. For a single qubit system $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$, a measurement in the basis $\{|0\rangle, |1\rangle\}$ will yield state $|0\rangle$ with probability $|\alpha|^2$ and state $|1\rangle$ with probability $|\beta|^2$, i.e., after the measurement, the superposition disappears and the quantum state “collapses” to one of the basis vectors.

Similarly, for an $n$-qubit quantum state $|\phi\rangle = \alpha_x|x\rangle, \quad x \in \{0, 1\}^n$, if we perform a measurement in the computational basis $\{|x\rangle, x \in \{0, 1\}^n\}$, we will obtain state $|x\rangle$ with probability $|\alpha_x|^2$ and the superposition $|\phi\rangle$ will be disappear.

Measurements can also be performed on part of a system. Take a two qubit state

$$|\phi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle$$
as an example. Suppose we measure the first qubit in the basis \{\ket{0}, \ket{1}\}. Then we will observe outcome 0 with a probability \(|\alpha_{00}|^2 + |\alpha_{01}|^2\) and outcome 1 with a probability \(|\alpha_{10}|^2 + |\alpha_{11}|^2\). If the outcome is 0, then \(\ket{\phi}\) collapse to the unit vector
\[
\alpha_{00}\ket{00} + \alpha_{01}\ket{01} / \sqrt{|\alpha_{00}|^2 + |\alpha_{01}|^2}.
\]

Measurements can also be performed in other bases. For example, the Hadamard basis for Hilbert space \(\mathbb{C}^2\) is defined as

\[
\ket{+} = \frac{1}{\sqrt{2}} (\ket{0} + \ket{1}),
\]
\[
\ket{-} = \frac{1}{\sqrt{2}} (\ket{0} - \ket{1}).
\]

A single qubit state \(\ket{\phi} = \alpha\ket{0} + \beta\ket{1}\) can be written as
\[
\ket{\phi} = \frac{\alpha + \beta}{\sqrt{2}} \ket{+} + \frac{\alpha - \beta}{\sqrt{2}} \ket{-}.
\]

A measurement of \(\ket{\phi}\) in the basis \{\ket{+}, \ket{-}\} will yield the state \(\ket{+}\) with a probability \((|\alpha + \beta|^2)/2\) and the state \(\ket{-}\) with probability \((|\alpha - \beta|^2)/2\).

In general, measurements are described by a set of measurement operators \(\{M_m\}\) satisfying \(M_m^\dagger M_m = I\). The index \(m\) refers to the measurement outcome. After applying the measurement \(\{M_m\}\) on state \(\ket{\phi}\), we will get outcome \(m\) with probability \(\bra{\phi} M_m^\dagger M_m \ket{\phi}\). If outcome \(m\) occurs, the state of the system will become
\[
M_m \ket{\phi} / \sqrt{\bra{\phi} M_m^\dagger M_m \ket{\phi}}.
\]

### 1.2.5 Unitary transformations

Quantum mechanics also allows unitary operators to be applied to quantum states. For an \(n\)-qubit system with state vector \(\ket{\phi}\), after applying a unitary operator \(U\), the new quantum state will be \(U\ket{\phi}\). Since unitary operators keep the length of a
vector, $U|\phi\rangle$ is still a unit vector. Unitary operators can be represented as matrices after choosing a basis. For example, the NOT transform, which maps $|0\rangle \to |1\rangle$, and $|1\rangle \to |0\rangle$, can be represented as the following matrix $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ in the standard basis. The new quantum state after applying $X$ on $|\phi\rangle$ is

$$X|\phi\rangle = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}.$$ 

If we have a unitary operator $U_A$ on system $A$ and a unitary operator $U_B$ on system $B$, what is their joint operator on system $AB$? To describe this, we need the concept of tensor product of operators. Let the state spaces of system $A$ and $B$ be represented as vector spaces $\mathcal{N}_A$ and $\mathcal{N}_B$, respectively. Let $|\phi\rangle \in \mathcal{N}_A$ and $|\psi\rangle \in \mathcal{N}_B$, we define a linear operator $U_A \otimes U_B \in \mathcal{L}(\mathcal{N}_A \otimes \mathcal{N}_B)$ such that

$$(U_A \otimes U_B)(|\phi\rangle \otimes |\psi\rangle) = U_A|\phi\rangle \otimes U_B|\psi\rangle,$$

for any $|\phi\rangle$, $|\psi\rangle$.

It can be verified that $U_A \otimes U_B$ is well defined and it is a unitary operator. Suppose $U_A$ is represented as $m \times n$ dimensional matrix $S$ and $U_B$ is represented as $k \times l$ dimensional matrix $T$. Let $S_{i,j}$ denote the $(i,j)$th entry of matrix $S$. Then $U_A \otimes U_B$ is represented as a $mk \times nl$ matrix $S \otimes T$,

$$S \otimes T = \begin{pmatrix} S_{11}T & S_{12}T & \cdots & S_{1n}T \\ S_{21}T & S_{22}T & \cdots & S_{2n}T \\ \vdots & \vdots & \ddots & \vdots \\ S_{mn}T & S_{m2}T & \cdots & S_{mn}T \end{pmatrix}.$$ 

Notice that, operations on quantum states must either be unitary transformations or measurements. A consequence is that unknown quantum states cannot be cloned. This is now called the no-cloning theorem, discovered by Dieks [42], and by Wootters.
and Zurek [94]. Otherwise, suppose we have a quantum operator $U$ that can copy arbitrary one qubit quantum states, then

$$
U|0\rangle|0\rangle = |0\rangle|0\rangle
$$
$$
U|1\rangle|0\rangle = |1\rangle|1\rangle
$$
$$
U\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle\right) = \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle).
$$

On the other hand, by the linearity of $U$,

$$
U\left(\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|0\rangle\right) = \frac{1}{\sqrt{2}}(U|0\rangle|0\rangle + U|1\rangle|0\rangle) = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \neq \frac{1}{2}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle).
$$

We have a contradiction.

### 1.2.6 Density operators

Recall that we can represent the state of an $n$-qubit system as a vector in Hilbert space $\mathbb{C}^{2^n}$. These states are called pure states. Now suppose we prepare a quantum system to be in a mixture of pure quantum states. More precisely, with probability $p_i$, the system is in pure state $\eta_i$. Then the state of such a system can be represented as a density operator:

$$
\rho = p_1|\eta_1\rangle\langle\eta_1| + \ldots + p_n|\eta_n\rangle\langle\eta_n|.
$$

Using the density operator representation, applying a unitary operator $U$ on system $A$ will change its state from $\rho$ to $U\rho U^\dagger$. On the other hand, if measurement $\{M_m\}$ is applied to system $A$, then outcome $m$ occurs with probability $tr\left(M_m\rho M_m^\dagger\right)$. If outcome $m$ occurs, the state of system $A$ will become

$$
\frac{M_m\rho M_m^\dagger}{tr\left(M_m\rho M_m^\dagger\right)}.
$$
If the state of system $A$ is prepared as a density operator $ρ_A$ and state of system $B$ is prepared as a density operator $ρ_B$, then the state of their composite system is the density operator $ρ_{AB} = ρ_A ⊗ ρ_B$.

For a density operator $ρ$, it is easy to see $\text{tr} (ρ) = 1$. In fact, if $\text{tr} (ρ^2) = 1$, we know $ρ$ always represents a pure state. If $\text{tr} (ρ^2) < 1$, we call the corresponding quantum state a mixed state.

The density operator formulation is mathematically equivalent to the approach of state vectors, but sometimes it is more convenient for describing a subsystem. E.g, let $ρ_{AB}$ be the density operator of system $AB$, then the state of system $A$ is just $ρ_A = \text{tr}_B(ρ_{AB})$, where $\text{tr}_B(·)$ denotes partial trace over $B$. (Ref. Equation A.2 for the definition of partial trace). In this case, $ρ_A$ is also called the reduced density operator of $ρ_{AB}$.

1.2.7 Superoperators

So far we dealt with only closed quantum systems that do not interact with the environment. The evolution of a quantum system interacting with its environment can be described by physically realizable superoperators [70]. A physically realizable superoperator is a superoperator that has the following form: $T = \text{tr}_F(V \cdot V^\dagger) : ρ \to \text{tr}_F(VρV^\dagger)$, where $V \in \mathcal{L}(\mathcal{N},\mathcal{N} ⊗ \mathcal{F})$ is an isometric embedding. An equivalent formulation, called operator-sum representation, is that any physically realizable superoperator $T : \mathcal{L}(\mathcal{N}) → \mathcal{L}(\mathcal{M})$ can be represented as

\begin{equation}
T = \sum_k E_k \cdot E_k^\dagger.
\end{equation}

where $E_k \in \mathcal{L}(\mathcal{N},\mathcal{M})$ and $\sum_k E_k^\dagger E_k = I$ [59].

In this dissertation, we shall investigate the nonlocality of physically realizable superoperators. The motivations are discussed in Section 1.5.
1.3 Communication complexity

1.3.1 Classical communication complexity

Consider the following scenario: there are two spatially separated parties, Alice and Bob. Alice is given a binary string $x$ and Bob a binary string $y$. Both of them also receive a description of a function $f$ and an unlimited amount of computational resources (CPUs, memories, etc.). Their task is to compute the value of $f(x, y)$. Assume that the function $f$ depends on both $x$ and $y$, otherwise the solution is trivial. Neither party has sufficient information to finish the task unless they communicate to each other. A simple solution would be for Alice to send the whole string $x$ to Bob, who then computes $f(x, y)$ and sends the result back. This leads to the following question: are there solutions where they exchange fewer bits?

The minimum amount of information that Alice and Bob have to exchange, in order to compute $f$, is defined as the communication complexity of $f$ (as shown in Figure 1.1, contrast this with information theory, where the amount of bits to communicate is given and the task is how to send them over to the other party). This was first studied by Yao in his seminar paper [95]. Since then, classical communication complexity has now become a major branch of complexity theory, with a wide range of applications such as in VLSI design, time-space tradeoff, derandomization, and circuit complexity. The excellent monograph of Kushilevitz and Nisan [64] surveys results up to 1997.
There are several variants of communication complexities: each of which corresponds to different types of interactions allowed and whether or not small error probabilities are allowed. Informally, the deterministic communication complexity of $f$, denoted by $D(f)$, is defined to be the minimum amount of bits that Alice and Bob have to exchange to compute $f$ correctly for any pair of inputs. The randomized communication complexity of $f$, denoted by $R(f)$, is similarly defined, with the exception that Alice and Bob have access to their own private and independent random sources and that they are only required to compute $f(x,y)$ correctly with a probability of at least $2/3$. If Alice and Bob are allowed to use publicly announced random bits instead, the complexity is called randomized communication complexity with public coins, denoted by $R^{\text{pub}}(f)$.

One of the central themes in the study of classical communication complexity is to understand how randomness helps save the communication cost. A basic finding of Yao [95] is that there are functions $f$ such that the cost in the randomized model is exponentially smaller than that in the deterministic model. One example of such functions is checking the equality of binary strings $x \in \{0,1\}^n$ and $y \in \{0,1\}^n$. Its deterministic communication complexity is $\Theta(n)$, while its randomized communication complexity is only $\Theta(\log n)$ (assuming private randomness).

Different ways of using randomness also result in subtle changes on communication complexities. A basic finding in this regard, due to Newman [68], is that public-coin protocols can save at most $O(\log n)$ bits over private-coin protocols. However, the situation is dramatically different in the Simultaneous Message Passing (SMP) model, also introduced by Yao [95], where Alice and Bob each sends a message to a third person, who then outputs the outcome of the protocol. Apparently, this is a more restricted model, and for any function, the communication complexity in
Goal: compute $f(x, y)$

Figure 1.2: The model of quantum communication complexity

this model is at least that in the general interactive communication model. Denote by $R^\parallel(f)$ and $R^\parallel,\text{pub}(f)$ the communication complexities in the SMP model with private and public random coins respectively. It is interesting to note that $R^\parallel,\text{pub}(\text{Equality}) = O(1)$ but $R^\parallel(\text{Equality}) = \Theta(\sqrt{n})$ [4, 69, 7].

1.3.2 Quantum communication complexity

Now suppose Alice and Bob are equipped with quantum computers and they can exchange qubits rather than classical bits (shown in Figure 1.2). As in the randomized communication complexity model, they are only required to compute $f(x, y)$ correctly for any $(x, y)$ with a probability of at least $2/3$. The minimum amount of qubits that they need to exchange is called quantum communication complexity, denoted by $Q(f)$. This was also introduced by Yao [96] in 1993 to prove lower bounds on the size of quantum formula. Since then, it has developed into a rich field, both for proving strong quantum lower bounds and for its own sake. [37, 36] survey results up to 2001.

Like any other directions involving quantum information processing, the central problem in this area is to identify problems that have an exponential gap between quantum and classical communication complexities, or to prove that such a problem does not exist.

Indeed, exponential gaps are found for several communication tasks [6, 75, 9,
51, 50]. However, those tasks are either sampling, or computing a partially defined Boolean function or a relation. An exponential gap is known for a total Boolean function$^1$ — checking equality, but in a restricted model involving a third party [24]. It remains open today if super-polynomial gaps are possible for computing a total Boolean function in the more commonly studied model of two-party, interactive communication. This is perhaps the most significant problem in quantum communication complexity.

In this dissertation, we first consider the classical and quantum communication complexities of a specific problem — the HAMMING DISTANCE problem, which is for two separated parties to determine whether or not the Hamming distance between their private strings is above a given threshold — in Chapter II. Then we proceeds to deal with the conjecture that there is no exponential gap between classical and quantum communication complexities for all total Boolean functions in the two-party, interactive communication models in Chapter III.

1.4 Quantum entanglement

The famous EPR “paradox”, named after Einstein, Podolsky and Rosen [44], is a thought experiment to challenge the “completeness” of quantum mechanics. It is essentially as follows [19]: two “quantum coins” (e.g., polarized photons), possessed by two spatially separated parties Alice and Bob, may be correlated in a state that can be schematically represented as

$$\frac{1}{\sqrt{2}} (|0\rangle_A |1\rangle_B - |1\rangle_A |0\rangle_B) .$$

If each party measures his/her own coin, he/she will observe two outcomes (0 and 1) with equal chances. However, once a measurement is made by either party, say,

$^1$For a Boolean function $f : D \rightarrow \{0,1\}$, where $D \subseteq X \times Y$, if $D = X \times Y$, then $f$ is called a total function; otherwise, it is called a partially defined function.
Alice, then Bob will always observe the opposite outcome with certainty. A unique property of the above state is that no matter what property of the coins is measured – be it determining their positions or the velocities – Bob’s outcome is always opposite to that of Alice with certainty. Since intuitively, what Alice does locally should not affect Bob’s world, this allows any pair of properties of a coin (e.g., positions and velocities) to be determined precisely. This is at odds with the “uncertainty” principle of quantum mechanics that not all pairs of properties can be determined with certainty.

The EPR paradox did not reduce quantum mechanics to contradictions. Instead, it revealed the essence — quantum entanglement — that underlies the many counter-intuitive properties and marvellous capabilities of quantum information. For example, John Bell formulated a set of inequalities, referred as Bell Inequalities [12] now, which must be satisfied by the correlations produced by any classical hidden variable model, but would nevertheless be violated by some quantum correlations. The violations has been confirmed by several physical experiments (e.g., [92]).

Given its importance, quantum entanglement has been the subject of numerous studies (see, e.g., the books [70, 74]). The focus of these studies has been on understanding the inherent quantitative tradeoffs among various resources involved in the creation and conversion of entangled states. In the context of quantum communication, a basic question is: what is the role of shared entanglement in assisting quantum communication?

In fact, this question has puzzled many researchers [33, 23, 60, 67]. It is known that shared entanglement could save a constant number of bits over shared randomness [33, 23] or even a half of the communication [17, 34]. However, little is known on the limit of the savings. This is in sharp contrast with the classical case of shared
randomness, where we know that it can only save at most a logarithmic additive term [68]. If there is a quantum protocol that exchanges $q$ qubits with $m$ qubits of shared entanglement, then the best classical simulation we know is $\exp(\Omega(q + m))$.

In this dissertation, we study the limit of the benefits of shared entanglement in assisting quantum communication. Our approach is discussed in the next section.

1.5 Nonlocality of quantum operations

Since entanglement is the result of nonlocal quantum interactions, understanding nonlocality of quantum operations is also of fundamental importance. A basic question is: how to quantify nonlocality?

A natural nonlocality measure of a quantum operation is its *generating capacity*, which is the maximum increase of entanglement that it could create (see e.g., [16]). Another approach, more from a computational point of view, is to consider the amount of resources, such as the time in the case of using elementary Hamiltonians, or the number of elementary gates, required to simulate the operator (e.g., [30, 31]).

In this dissertation, we take two different approaches: classical simulation and maximum tensor norms. The frameworks of these two approaches are described below.

1.5.1 Classical simulations of bipartite quantum measurement.

Our first approach follows intuitions from the subject of communication complexity. Consider the following quantum process as shown in Figure 1.3. Alice and Bob share a bipartite state $|E\rangle$. They apply local operations $R_A$ and $R_B$ to his/her system respectively. Then they perform a measurement $Q$ to the joint system, producing a distribution of measurement outcomes, denoted by $\mu$.

Imagine that Alice and Bob have lost their quantum power. They both know
classical descriptions of $Q$, $|E\rangle$, and their local operations, but do not know the other party’s local operation. They are also given an unlimited supply of common random bits. From this classical information, they hope to simulate the quantum process, by producing an output whose distribution is close to $\mu$ (as shown in Figure 1.4). We define the classical communication complexity of $Q$, denoted by $\text{Com}(Q)$, to be the minimum number of bits that need to be exchanged by the simulating process.

Intuitively, $\text{Com}(Q)$ reflects how nonlocal $Q$ is. Consider, for example, the simple case that $Q$ consists of local operations. If there is no quantum correlation in the initial state, it is clear that Alice and Bob could simulate the quantum process without interaction. On the other hand, $\text{Com}(Q)$ could be much larger. Let $n \geq 1$ be an integer. For $x, y \in \{0, 1\}^n$, let $x \cdot y \overset{\text{def}}{=} x_1y_1 + x_2y_2 + \cdots + x_ny_n \mod 2$. The problem of determining whether $x \cdot y = 1$ is called the Inner Product Problem in the communication complexity literature. It is well known that any classical communication
protocol for determine whether $x \cdot y = 1$ requires $\Omega(n)$ bits of communication. In fact, Cleve, van Dam, Nielsen, and Tapp [34] proved that $\Omega(n)$ quantum bits are necessary, too. Consider the following measurement operator,

$$IP_n \overset{\text{def}}{=} \sum_{x,y \in \{0,1\}^n, x \cdot y = 1} |x\rangle \langle x| \otimes |y\rangle \langle y|.$$  

When $R_A$ is to create a state $|x\rangle$, $x \in \{0,1\}^n$, and $R_B$ is to create a state $|y\rangle$, $y \in \{0,1\}^n$, the outcome of measurement $IP_n$ reveals whether $x \cdot y = 1$. Thus $\text{Com}(IP_n) = \Omega(n)$. We do not know if this bound for $\text{Com}(IP_n)$ is tight.

We investigate how $\text{Com}(Q)$ is determined in general. It is not immediately clear if $\text{Com}(Q)$ can be bounded from above for all $Q$, as the dimension of the initial state $|E\rangle$ could be arbitrarily large.

### 1.5.2 The maximum tensor norm of bipartite superoperators.

Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be two Banach spaces endowed with norm $\| \cdot \|$. A norm $\| \cdot \|_\alpha$ on $\mathcal{H}_A \otimes \mathcal{H}_B$ is called a tensor norm (also called a crossnorm), if for any $a \in \mathcal{H}_A$ and $b \in \mathcal{H}_B$, it follows $\|a \otimes b\|_\alpha = \|a\| \|b\|$. Tensor norms are powerful tools for the study of tensor product spaces. Their study was pioneered by Robert Schatten [83], and has since then developed into a subject with rich and deep results (e.g., [38, 82]).

Informally, tensor norms quantify how different an element is from a product element; hence it may be useful for the study of nonlocality. Surprisingly, only a few explorations have been done in this direction. In a pioneering work, Rudolph [78] proves that a mixed state is separable if and only if its maximum tensor norm with respect to the trace norm is precisely 1. The criterion is further explored in [80, 81, 54, 73, 3] and the tensor norm is also used as an entanglement measure in [79, 43, 29].

This motivated us to use the maximum tensor norm of bipartite superoperators
to quantify its nonlocality. The maximum tensor norm we define is with respect to the diamond norm [59], and it does not appear to have been studied before.

1.6 Organization

The rest of this dissertation is organized as follows. We start by studying the classical and quantum communication complexities of a specific problem, the Hamming Distance problem, in Chapter II. There is a gap between the communication protocols and the lower bounds. We prove a stronger quantum lower bound and also construct better classical protocols.

Next, we proceed to deal with the conjecture that there is no exponential gap between quantum and classical communication complexities in Chapter III. Razborov proved this conjecture for so far the most general class of functions. We prove the conjecture for a broader class of functions.

After that, we switch to the direction of measuring nonlocality of quantum operations. In Chapter IV, we define a certain tensor norm to measure nonlocality of bipartite quantum measurements. The tensor norm turns out to also imply the limitation of quantum entanglement in reducing communication costs. In Chapter V, we define a maximum tensor norm on superoperators. We prove that the value of this maximum tensor norm is a criterion for deciding whether a bipartite superoperator is bi-local. Furthermore, we show that estimates of the maximum tensor norm can be used to derive strong quantum lower bound in communication complexities.

We summarize the contribution of this dissertation and discuss future directions in Chapter VI.
CHAPTER II

The communication complexity of the Hamming Distance problem

This chapter is based on [55]. We investigate the randomized and quantum communication complexity of the HAMMING DISTANCE problem, which is to determine if the Hamming distance between two n-bit strings is no less than a threshold \(d\). We construct better classical protocols and also proved a stronger quantum lower bound for this problem.

2.1 Introduction and Summary of results

We discussed the two-party communication model in Chapter I. Apart from the two-party communication model, Yao also introduced the Simultaneous Message Passing (SMP) model [95], where Alice and Bob each sends a message to a third person, who then outputs the outcome of the protocol (as shown in Figure 2.1). Apparently, this is a more restricted model and for any function, the communication complexity in this model is at least that in the general interactive communication model, as shown in Figure 2.1. Denote by \(R^\parallel(f)\) and \(R^\parallel,\text{pub}(f)\) the communication complexities in the SMP model with private and public random coins, respectively. It is interesting to note that \(R^\parallel,\text{pub}({\text{EQUALITY}}) = O(1)\) but \(R^\parallel({\text{EQUALITY}}) = \Theta(\sqrt{n})\) [4, 69, 7].
Similarly, the quantum communication complexities in the SMP model are denoted by $Q^\parallel(f)$ and $Q^{\parallel,*}$, depending on whether or not shared entanglement is allowed. The following relations among the measures are easy to observe.

\[ Q^*(f) \leq R_{\text{pub}}^\parallel(f) \leq R_{\text{pub}}^{\parallel,*}(f) \quad (2.1) \]

An interesting problem in both quantum and classical communication models is to determine the biggest gap between quantum and randomized communication complexities. Buhrman, Cleve, Watrous and de Wolf [24] proved that $Q^\parallel(\text{Equality}) = O(\log n)$, an exponential saving compared to the randomized counterpart result $R^\parallel(\text{Equality}) = \Theta(\sqrt{n})$ mentioned above. This exponential separation is generalized by Yao [97], showing that $R_{\text{pub}}^{\parallel,*}(f) = \text{constant}$ implies $Q^\parallel(f) = O(\log n)$. As an application, Yao considered the \textsc{Hamming Distance} problem defined below.

For any $x, y \in \{0,1\}^n$, the Hamming weight of $x$, denoted by $|x|$, is the number of 1’s in $x$, and the Hamming distance of $x$ and $y$ is $|x \oplus y|$, with “$\oplus$” being bit-wise XOR.

**Definition 2.1.** For $1 \leq d \leq n$, the $d$-\textsc{Hamming Distance} problem is to compute the following Boolean function $\text{HAM}_{n,d} : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, with $\text{HAM}(x, y) = 1$ if and only if $|x \oplus y| > d$. 

![Figure 2.1: The Simultaneously Message Passing (SMP) model](image-url)
Lemma 2.2 (Yao). $R_{||,\text{pub}}^\bullet (\text{HAM}_{n,d}) = O(d^2)$.

In a recent paper [49], Gavinsky, Kempe and de Wolf gave another classical protocol, which is an improvement over Yao’s when $d \gg \log n$.

Lemma 2.3 (GKW). $R_{||,\text{pub}}^\bullet (\text{HAM}_{n,d}) = O(d \log n)$.

The best known lower bound $\Omega(d/\log d)$ is proved in the quantum two party, interactive model [60]. We observe a lower bound for $Q^*(\text{HAM}_{n,d})$, which is also a lower bound for $R_{||,\text{pub}}^\bullet (\text{HAM}_{n,d})$ according to Equation (2.1).

Notice that $\text{HAM}(x,y) = n - \text{HAM}(x,\bar{y})$, where $\bar{y} \stackrel{\text{def}}{=} 11 \cdots 1 \oplus y$. Therefore $Q^*(\text{HAM}_{n,d}) = Q^*(\text{HAM}_{n,n-d})$, and we need only consider the case $d \leq n/2$.

Proposition 2.4. For any $d \leq n/2$, $Q^*(\text{HAM}_{n,d}) = \Omega(d)$.

We then construct a public-coin randomized SMP protocol that almost matches the lower bound and improves both of the above protocols.

Theorem 2.5. $R_{||,\text{pub}}^\bullet (\text{HAM}_{n,d}) = O(d \log d)$.

We shall prove the above two results in the following two sections.

Other related work: Ambainis, Gasarch, Srinavasan, and Utis [5] considered the error-free communication complexity, and proved that any error-free quantum protocol for the HAMMING Distance problem requires at least $n - 2$ qubits of communication in the interactive model, for any $d \leq n - 1$. Feigenbaum et al. [45] studied the secure multiparty approximate computation of the Hamming distance.
2.2 Lower bound for the quantum communication complexity of the Hamming Distance problem

For proving the lower bound, we restrict $\text{HAM}_{n,d}$ on those pairs of inputs with equal Hamming distances. More specifically, for an integer $k$, $1 \leq k \leq n$, define $X_k = Y_k \overset{\text{def}}{=} \{ x : x \in \{0,1\}^n, |x| = k \}$. Let $\text{HAM}_{n,k,d} : X_k \times Y_k \to \{0,1\}$ be the restriction of $\text{HAM}_{n,d}$ on $X_k \times Y_k$.

Before proving Proposition 2.4, we briefly introduce some related results. Let $x, y \in \{0,1\}^n$. The Set Disjointness problem is to compute the following Boolean function $\text{DISJ}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, $\text{DISJ}_n(x,y) = 1$ if and only if there exists an integer $i$, $1 \leq i \leq n$, so that $x_i = y_i = 1$. It is known that $R(\text{DISJ}_n) = \Theta(n)$ [57, 76, 10], and $Q^*(\text{DISJ}_n) = \Theta(\sqrt{n})$ [77, 1].

We shall use an important lemma in Razborov [77], which is more general than his remarkable lower bound on quantum communication complexity of Disjointness.

Here we may abuse the notation by viewing $x \in \{0,1\}^n$ as the set $\{ i \in [n] : x_i = 1 \}$.

**Lemma 2.6 (Razborov).** Suppose $k \leq n/4$ and $l \leq k/4$. Let $D : [k] \to \{0,1\}$ be any Boolean predicate such that $D(l) \neq D(l-1)$. Let $f_{n,k,D} : X_k \times Y_k \to \{0,1\}$ be such that $f_{n,k,D}(x,y) \overset{\text{def}}{=} D(|x \cap y|)$. Then $Q^*(f_{n,k,D}) = \Omega(\sqrt{kl})$.

**Proof of Proposition 2.4.** Consider $D$ in Lemma 2.6 such that $D(t) = 1$ if and only if $t < l$. For any $x, y \in X_k$, we have $|x \cap y| = k - \text{HAM}(x,y)/2$. Let $l = k - d/2$, then $k - \text{HAM}(x,y)/2 < l$ if and only if $\text{HAM}(x,y) > d$. Therefore, $D(|x \cap y|) = 1$ if and only if $\text{HAM}(x,y) > d$. This implies that $f_{n,k,D}$ and $\text{HAM}_{n,k,d}$ are actually the same function, and thus $Q^*(f_{n,k,D}) = Q^*(\text{HAM}_{n,k,d})$.

To use lemma 2.6, the following two constraints on $k$ and $l$ need to be satisfied: $k \leq n/4$ and $l \leq k/4$. When $d \leq 3n/8$, let $k = 2d/3 \leq n/4$, then $l = 2d/3 - d/2 =$
Both requirements for $k$ and $l$ are satisfied. So applying lemma 2.6, we get $Q^*(\text{HAM}_{n,k,d}) = Q^*(f_{n,k,d}) = \Omega(\sqrt{k}) = \Omega(d)$.

For $3n/8 < d \leq n/2$, it is reduced to the above case ($d \leq 3n/8$) rather than lemma 2.6. Let $m = \lceil 8d/5 - 3n/5 \rceil$. Fix first $m$ bits in $x$ to be all 1’s, and use $x'$ to denote $x_{m+1} \ldots x_n$. Similarly, fix first $m$ bits of $y$ to be all 0’s, and use $y'$ to denote $y_{m+1} \ldots y_n$. Put $n' = n - m$, $k' = n'/4$, and $d' = d - m$. Then $\text{HAM}(x,y) = \text{HAM}(x',y') + m$ and $Q^*(\text{HAM}_{n,d})(x,y) \geq Q^*(\text{HAM}_{n',k',d'})(x',y')$. It is easy to verify that $d' \leq 3n'/8$ and $d' = \Omega(d)$. Employing the result of the case that $d \leq 3n/8$, we have $Q^*(\text{HAM}_{n',k',d'}) = \Omega(d')$. Thus $Q^*(\text{HAM}_{n,d}) \geq Q^*(\text{HAM}_{n',k',d'}) = \Omega(d') = \Omega(d)$.

2.3 Upper bound for the classical communication complexity of the Hamming Distance problem

To prove theorem 2.5, we reduce the $\text{HAM}_{n,d}$ problem to $\text{HAM}_{16d^2,d}$ problem by the following lemma.

Lemma 2.7.

$$R^{\parallel,\text{pub}}(\text{HAM}_{n,d}) = O(R^{\parallel,\text{pub}}(\text{HAM}_{16d^2,d}))$$

Note that Theorem 2.5 immediately follows from Lemma 2.7 because by Lemma 2.3, $R^{\parallel,\text{pub}}(\text{HAM}_{n,d}) = O(d \log n)$, thus $R^{\parallel,\text{pub}}(\text{HAM}_{16d^2,d}) = O(d \log d^2) = O(d \log d)$. Now by Lemma 2.7, we have $R^{\parallel,\text{pub}}(\text{HAM}_{n,d}) = O(d \log d)$. So in what follows, we shall prove Lemma 2.7. Define a partial function $\text{HAM}_{n,d|2d}(x,y)$ with domain $\{(x,y) : x, y \in \{0, 1\}^n, |x \oplus y| \text{ is either less than } d \text{ or at least } 2d\}$ as follows.

$$\text{HAM}_{n,d|2d}(x,y) = \begin{cases} 0 & \text{if } \text{HAM}(x,y) \leq d \\ 1 & \text{if } \text{HAM}(x,y) > 2d \end{cases}$$

Then
Lemma 2.8.

\[ R_X^{||,\text{pub}}(\text{HAM}_{n,d|2d}) = O(1) \]

Proof of Lemma 2.8. We revise Yao’s protocol [97] to design an \( O(1) \) protocol for \( \text{HAM}_{n,d|2d} \). Assume the Hamming distance between \( x \) and \( y \) is \( k \). Alice and Bob share some random public string, which consists of a sequence of \( \gamma n (\gamma \) is some constant to be determined later) random bits, each of which is generated independently with probability \( p = 1/(2d) \) of being 1. Denote this string by \( z_1, z_2, \ldots, z_\gamma \), each of length \( n \). Party \( A \) sends the string \( a = a_1a_2 \cdots a_\gamma \) to the referee, where \( a_i = x \cdot z_i \mod 2 \). Party \( B \) sends the string \( b = b_1b_2 \cdots b_\gamma \) to the referee, where \( b_i = y \cdot z_i \mod 2 \). The referee announces \( \text{HAM}_{n,d}(x,y) = 1 \) if and only if the Hamming distance between \( a \) and \( b \) is more than \( m = (1/2 - q)\gamma \) where \( q = ((1 - 1/d)^d + (1 - 1/d)^{2d})/4 \).

Now we prove the above protocol is correct with probability at least \( 49/50 \). Let \( c_i = a_i \oplus b_i \). Notice that the Hamming distance between \( a \) and \( b \) is the number of 1’s in \( c = c_1c_2 \cdots c_\gamma \). We need the following Lemma by Yao [97].

Lemma 2.9. Assume that the Hamming distance between \( x \) and \( y \) is \( k \). Given \( c \) as defined above, each \( c_i \) is an independent random variable with probability \( \alpha_k \) of being 1, where \( \alpha_k = 1/2 - 1/2(1 - 1/d)^k \).

Since \( \alpha_k \) is an increasing function over \( k \), to separate \( k \leq d \) from \( k > 2d \), it would be sufficient to discriminate the two cases that \( k = d \) and \( k = 2d \). Let \( N_k \) be a random variable denoting the number of 1’s in \( c \), and \( E(N_k) \) and \( \sigma(N_k) \) denote corresponding expectation and standard deviation, respectively. Then we have \( E(N_k) = \alpha_k \gamma \), and \( \sigma(N_k) \leq (\alpha_k \gamma)^{1/2} \). Thus \( E(N_{2d}) - E(N_d) = \gamma(\alpha_{2d} - \alpha_d) = \frac{1}{2}\gamma(1 - \frac{1}{d})^d(1 - (1 - \frac{1}{2})^d) \geq \frac{1}{3}\gamma \). Let \( \gamma = 20000 \), then \( E(N_{2d}) - E(N_d) \geq 2500 \), while \( \sigma(N_d), \sigma(N_{2d}) < (\frac{1}{2}\gamma)^{1/2} = 100 \). The cutoff point in the protocol is the middle
of \( E(N_d) \) and \( E(N_{2d}) \). By the Chebyshev Inequality, with probability of at most \( 1/100 \), \(|N_d - E(N_d)| > 10\sigma(N_d) = 1000 \). So does \( N_{2d} \). Thus with probability of at least \( 49/50 \), the number of 1’s in \( c \) being more than cutoff point implies \( k > 2d \) and vice versa. Therefore, \( O(\gamma) \) communication is sufficient to discriminate the case \( \text{HAM}(x, y) > 2d \) and \( \text{HAM}(x, y) \leq d \) with error probability of at most \( 1/50 \).

The following fact is also useful.

**Fact 1.** If \( 2d \) balls are randomly thrown into \( 16d^2 \) buckets, then with probability of at least \( 7/8 \), each bucket has at most one ball.

**Proof of Fact 1.** There are \( \binom{2d}{2} \) pairs of balls. The probability of one specific pair of balls falling into the same bucket is \( \frac{1}{16d^2} \cdot \frac{1}{16d^2} \cdot 16d^2 = \frac{1}{16d^2} \). Thus the probability of having a pair of balls in the same bucket is upper bounded by \( \frac{1}{16d^2} \cdot \binom{2d}{2} < 1/8 \). Thus Fact 1 holds.

Now we are ready to prove Lemma 2.7.

**Proof of Lemma 2.7.** If \( 16d^2 \geq n \), the lemma is obviously true by appending 0’s to \( x \) and \( y \).

If \( 16d^2 < n \), suppose we already have a protocol \( P_1 \) of \( C \) communication to distinguish the cases \( |x \oplus y| \leq d \) and \( d < |x \oplus y| \leq 2d \) with error probability at most \( 1/8 \). Then we can have a protocol of \( C + O(1) \) communication for \( \text{HAM}_{n,d} \) with error probability at most \( 1/4 \). Actually, by repeating the protocol for \( \text{HAM}_{n,d/2d}(x, y) \) several times, we can have a protocol \( P_2 \) of \( O(1) \) communication to distinguish the cases \( |x \oplus y| \leq d \) and \( |x \oplus y| > 2d \) with error probability at most \( 1/8 \). Now the whole protocol \( P \) is as follows. Alice sends the concatenation of \( m_{A,1} \) and \( m_{A,2} \), which are her messages when she runs \( P_1 \) and \( P_2 \), respectively. So does Bob send the concatenation of his two corresponding messages \( m_{B,1} \) and \( m_{B,2} \). The referee then
runs protocol $P_i$ on $(m_{A,i}, m_{B,i})$ and gets the results $r_i$. The referee now announces $|x \oplus y| \leq d$ if and only if both $r_1$ and $r_2$ say $|x \oplus y| \leq d$.

It is easy to see that the protocol is correct. If $|x \oplus y| \leq d$, then both protocols announces so with probability at least $7/8$, and thus $P$ is correct with probability at least $3/4$. If $|x \oplus y| > d$, then one of the protocols gets the correct range of $|x \oplus y|$ with probability at least $7/8$, and thus $P$ announces $|x \oplus y| > d$ with probability at least $7/8$ too.

Now it remains to design a protocol of $O(R^{\|\cdot\|_{\text{pub}}}(\text{HAM}_{16d^2, d}))$ communication to distinguish $|x \oplus y| \leq d$ and $d < |x \oplus y| \leq 2d$. First we assume that $n$ is divisible by $16d^2$, otherwise we pad some 0’s to the end of $x$ and $y$. Using the public random bits, Alice divides $x$ randomly into $16d^2$ parts evenly, Bob also divides $y$ correspondingly. Let $A_i, B_i (1 \leq i \leq 16d^2)$ denote corresponding parts of $x, y$. By Fact 1, with probability at least $7/8$, each pair $A_i, B_i$ would contain at most one bit on which $x$ and $y$ differ. Therefore, the Hamming distance of $A_i$ and $B_i$ would be either 0 or 1, i.e, the Hamming distance between $A_i$ and $B_i$ equals the parity of $A_i \oplus B_i$, which is further equal to $\text{PARITY}(A_i) \oplus \text{PARITY}(B_i)$. Let $a_i$ denote the parity of $A_i$, $b_i$ denote the parity bit of $B_i$, and let $a = a_1a_2 \cdots a_{16d^2}$, $b = b_1b_2 \cdots b_{16d^2}$. Then $\text{HAM}_{16d^2,d}(a, b) = \text{HAM}_{n,d}(x, y)$ with probability at least $7/8$. So we run the best protocol for $\text{Ham}_{16d^2,d}$ on the input $(a, b)$, and use the answer to distinguish $|x \oplus y| \leq d$ and $d < |x \oplus y| \leq 2d$. ■
CHAPTER III

The communication complexity of block-composed functions

This chapter is based on [86]. We investigate the conjecture that there is no super-polynomial gaps between quantum and classical communication complexities in the two-party, interactive model.

3.1 Introduction and summary of results

It remains open today if super-polynomial gaps are possible for computing a total Boolean function in the more commonly studied model of two-party interactive communication. This is one of the most significant problems in quantum communication complexity. The answer is widely conjectured to be “No”. We shall refer to it the Log-Equivalence Conjecture. Besides the lack of a natural candidate for a super-polynomial gap, two other intuitions support this conjecture. The first relates to the well known Log-Rank Conjecture, which states that the randomized communication complexity of any function $F : X \times Y \rightarrow \{0,1\}$ is polynomially related to $\widetilde{\text{Logrank}}(F)$: the logarithm of the smallest rank of a matrix $[\tilde{F}(x,y)]_{x,y}$ with $|\tilde{F}(x,y) - F(x,y)| \leq 1/3$, for all $(x,y) \in X \times Y$. Since $\frac{1}{2} \widetilde{\text{Logrank}}(F) \leq Q(F) \leq R(F)$, the Log-Equivalence Conjecture follows from the Log-Rank Conjecture.

\footnote{This is known to be true only for the case with no prior entanglement.}
The second intuition supporting the Log-Equivalence Conjecture is the fact that the similar conjecture is true for the closely related decision tree complexity. Recall that a decision tree algorithm computes a function \( f_n : \{0, 1\}^n \rightarrow \{0, 1\} \) by making queries of the type “what is the \( i \)’th bit of the input?” The decision tree complexity of \( f_n \) is the minimum number of queries required to compute \( f_n \) correctly for any input, as shown in Figure 3.1. Making use earlier results of Nisan and Szegedy [71] and Paturi [72], Beals, Buhrman, Cleve, Mosca, and de Wolf [11] proved that the quantum and the deterministic decision tree complexities are polynomially related. This is in sharp contrast with the exponential quantum speedups [89, 88, 32] on partial functions achieved by the quantum algorithms of Simon’s and Shor’s.

Razborov’s work [77] is a significant progress for the Log-Equivalence Conjecture. He defined the following notion of symmetric predicates. Let \( f_n : \{0, 1\}^n \rightarrow \{0, 1\} \) be a symmetric function, i.e., \( f_n(x) \) depends only on the Hamming weight of \( x \). A function \( F : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\} \) is called a symmetric predicate if \( F(x, y) = f(x_1 \land y_1, x_2 \land y_2, \ldots, x_n \land y_n) \). The DISJOINTNESS function \( \text{DISJ}_n \) is an important symmetric predicate that has been widely studied:

\[
\text{DISJ}_n(x, y) \overset{\text{def}}{=} \begin{cases} 
1 & \exists i, \ x_i = y_i = 1, \\
0 & \text{otherwise}.
\end{cases}
\]

**Theorem 3.1 (Razborov [77]).** For any symmetric predicate \( F : \{0, 1\}^n \times \{0, 1\}^n \rightarrow \{0, 1\} \), \( D(F) = O(\max\{Q(F)\}^2, Q(F) \log n) \).
Figure 3.2: Symmetric predicates

Combined with the $O(d \log d)$-bit classical protocol for deciding if $x, y \in \{0, 1\}^n$ has Hamming distance $|x \oplus y| \geq d$ (in Theorem 2.5), Razborov’s lower bound implies the following.

**Proposition 3.2.** For any symmetric predicate $F : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}$, $R(F) = O((Q(F))^2)$.

This bound is tight on DISJ$_n$, which admits the largest known quantum-classical gap for total Boolean functions. The class of symmetric predicates is also the most general class of functions on which the Log-Equivalence Conjecture is known to hold.

Notice that Razborov’s lower bound method relies on the symmetry of $f_n$. Thus we aim to develop lower-bound techniques for an arbitrary $f_n$, and to derive new quantum lower bounds. To this end, we consider the following class of functions.

**Definition 3.3.** Let $k, n \geq 1$ be integers. Given $f_n : \{0, 1\}^n \to \{0, 1\}$, and $g_k : \{0, 1\}^k \times \{0, 1\}^k \to \{0, 1\}$, the block-composition of $f_n$ and $g_k$ is the function $f_n \Box g_k : \{0, 1\}^{nk} \times \{0, 1\}^{nk} \to \{0, 1\}$ such that on $x, y \in \{0, 1\}^{nk}$, with $x = x_1 x_2 \cdots x_n$, and $y = y_1 y_2 \cdots y_n$, where $x_i, y_i \in \{0, 1\}^k$,

$$f_n \Box g_k(x, y) = f_n(g_k(x_1, y_1), g_k(x_2, y_2), \cdots, g_k(x_n, y_n)).$$

Note that a symmetric predicate based on a symmetric $f_n : \{0, 1\}^n \to \{0, 1\}$ is
the block composition $f_n \Box \wedge$, where $\wedge$ denotes the binary AND function. In our Main Lemma, stated and proved in Section 3.3, we derive a sufficient condition for $Q(f_n \Box g_k)$ to have a strong lower bound. An application of this Main Lemma is the following.

**Theorem 3.4 (Informal).** For any integer $n \geq 1$ and any function $f_n : \{0, 1\}^n \to \{0, 1\}$, the block composition of $f_n$ with a $g_k : \{0, 1\}^k \to \{0, 1\}$ has polynomially related quantum and randomized communication complexities, if $Q(g_k)$ and $R(g_k)$ are polynomially related, and $k$ is sufficiently large.

We state below an incarnation of the above theorem. Let $\text{IP}_k : \{0, 1\}^k \times \{0, 1\}^k \to \{0, 1\}$ be the widely studied INNER PRODUCT function

$$\text{IP}_k(x, y) \overset{\text{def}}{=} \sum_i x_i y_i \mod 2, \quad \forall x, y \in \{0, 1\}^k.$$  

**Corollary 3.5.** For any integers $k$ and $n$ with $k \geq 2 \log_2 n + 5$, and for an arbitrary $f_n : \{0, 1\}^n \to \{0, 1\}$, $D(f_n \Box \text{IP}_k) = O((Q(f_n \Box \text{IP}_k))^7)$.

The above corollary also holds for a random $g_k$ with high probability. Our technique can also be applied to symmetric predicates, thus giving an alternative proof to Razborov’s result, albeit with a weaker parameter.

**Theorem 3.6.** For any symmetric $f_n : \{0, 1\}^n \to \{0, 1\}$, $D(f_n \Box \wedge) = O((Q(f_n \Box \wedge))^3)$. 

--- 

**Figure 3.3: Block-composed functions**
Our approach is inspired by how the Log-Equivalence result in decision tree complexity was proved: for any \( f_n : \{0, 1\}^n \to \{0, 1\} \), both the quantum and the deterministic decision tree complexities were shown [71, 11] to be polynomially related to the \textit{approximate polynomial degree} \( \tilde{\deg}(f_n) \), which is the smallest degree of a real polynomial that approximate \( f_n \) to be within 1/3 on any 0/1 inputs. In our Main Lemma, we derive a sufficient condition on \( n \) and \( k \), and \( g_k \) such that \( Q(f_n \Box g_k) = \Omega(\tilde{\deg}(f_n)) \), for any \( f_n \). The randomized upper bound is obtained by simulating a decision tree algorithm for \( f_n \), and whenever one input bit of \( f_n \) is needed, the protocol calls a sub-protocol for computing \( g_k \) on the corresponding block.

One may consider Razborov’s lower bound on \textsc{DISJ} an application of the polynomial method as well. This is because, he showed that if there is a \( q \)-qubit protocol for \textsc{DISJ}_n, then there is a \( O(q) \)-degree polynomial approximating \( \textsc{OR}_n \). Thus the quantum lower bound of \( \Omega(\sqrt{n}) \) follows from the same lower bound on \( \tilde{\deg}(\textsc{OR}_n) \) due to Nisan and Szegedy [71] and Paturi [72]. We emphasize this connection of approximating polynomial and quantum protocol is not obvious at all and it makes use the symmetric of \textsc{DISJ} critically.

We avoid the dependence of Razborov’s proof on the symmetry property of \( f_n \) by taking the \textit{dual} approach of the polynomial method. We show that from the linear programming formulation of polynomial approximation, we can obtain a “witness” for \( f_n \) requiring a high approximate degree. This witness is then turned into a “witness” for the hardness of \( f_n \Box g_k \), under certain assumptions. While the approximate polynomial degree has been used to prove lower bounds, and its dual formulation has been known to several researchers\(^2\), our application of the dual form appears to be the first demonstration of its usefulness in proving new results\(^3\).

\(^2\)from Y. Shi’s personal communications with A. A. Razborov and M. Szegedy, respectively
\(^3\)During the writing of this dissertation, Sherstov [84] uses the same approach, dual of polynomial method, to prove a similar result independently. The main difference is that the \( \lambda \) matrix in his paper has a simple operator
Before we proceed to the proofs, we briefly review some other closely related works. Buhrman and de Wolf [26] are probably the first to systematically study the relationship of polynomial representations and communication complexity. However, their result applies to error-free quantum protocols, while we consider bounded-error case. Klauck [60] proved strong lower bounds for some symmetric predicates such as MAJORITY based on the properties of their Fourier coefficients. The same author formulated a lower bound framework that includes several known lower bound methods [61]. It would be interesting to investigate the limitations of our polynomial method in this framework.

3.2 Preliminaries

3.2.1 Communication complexities and quantum lower bound by approximate trace norm

Denote the domain of a function by \( \text{dom}(\cdot) \). For a positive integer \( n \), denote by \( F_n \defeq \{ f_n : \{0,1\}^n \to \{0,1\} \} \), and by \( G_n \defeq \{ g_k : \text{dom}(g_k) \to \{0,1\}, \text{dom}(g_k) \subseteq \{0,1\}^k \times \{0,1\}^k \} \). For the rest of this article \( f_n \in F_n \) and \( g_k \in G_k \), for some integers \( n,k \geq 1 \). If \( F \in G_n \) is a total function, we also denote by \( F \) the \( \{0,1\}^{2^n \times 2^n} \) matrix \( [F(x,y)]_{x,y \in \{0,1\}^n} \).

A powerful method for proving quantum communication complexity lower bounds is the following lemma, which was obtained by Razborov [77], extending a lemma of Yao [96]. Recall that the trace norm of a matrix \( A \in \mathbb{R}^{N \times M} \) is \( \|A\|_{tr} \defeq \text{trace} \sqrt{A^\dagger A} = \text{trace} \sqrt{AA^\dagger} \). Let \( F \) be a partial Boolean function defined on a subset \( \text{dom}(F) \subseteq X \times Y \). The approximate trace norm of \( F \) with error \( \epsilon \), \( 0 \leq \epsilon < 1/2 \), is

\[
\|F\|_{\epsilon, tr} \defeq \min\{\|\tilde{F}\|_{tr} : \tilde{F} \in \mathbb{R}^{N \times M}, \forall (x,y) \in \text{dom}(F), |\tilde{F}(x,y) - F(x,y)| \leq \epsilon \}.
\]

norm, while in our result, we need to apply triangle inequalities to upper bound the operator norm of \( h \).
Lemma 3.7 (Razborov-Yao[77, 96]). For any partial Boolean function $F$ whose domain is a subset of $X \times Y$, $Q_{\epsilon}(F) = \Omega(\log \frac{\|F\|_{\ell_{\infty}}}{\sqrt{|X|\cdot|Y|}})$.

3.2.2 Approximate polynomial degree

The study of low degree polynomial approximations of Boolean function under the $\ell_{\infty}$ norm was pioneered by Nisan and Szegedy [71] and Paturi [72], and has since then been a powerful tool in studying concrete complexities, including the quantum decision tree complexity (c.f. the survey by Buhrman and de Wolf [27]).

Let $f \in \mathcal{F}_n$. A real polynomial $\tilde{f} : \mathbb{R}^n \to \mathbb{R}$ is said to approximate $f$ with an error $\epsilon$, $0 < \epsilon < 1/2$, if

$$|f(x) - \tilde{f}(x)| \leq \epsilon, \quad \forall x \in \{0,1\}^n.$$ 

The approximate degree of $f$, denoted by $\tilde{\deg}(f)$ is smallest degree of a polynomial approximating $f$ with an error $\epsilon$. Difference choices for $\epsilon$ only result in a constant factor difference in the approximate degrees. Thus we omit the subscript $\epsilon$ for asymptotic estimations.

While the approximate degree of symmetric functions has a simple characterization [71, 72], it is difficult to determine in general. For example, the approximate degree of the two level AND-OR trees is still unknown. On the other hand, $\tilde{\deg}(f)$ is polynomially related to the deterministic decision tree complexities $T(f)$. Formally, $T(f)$ is defined to be the minimum integer $k$ such that there is an ordered full binary tree $T$ of depth $k$ satisfying the following properties: (a) each non-leaf vertex is labelled by a variable $x_i$, and each leaf is labelled by either 0 or 1 (but not both); (b) for any $x \in \{0,1\}^n$, the following walk leads to a leaf labelled with $f(x)$: start from the root, at each non-leaf vertex labeled with $x_i$, take the left edge if $x_i = 0$, and take the right edge otherwise.
Theorem 3.8 (Nisan and Szegedy [71], Beals et al. [11]). For any Boolean function $f_n$, there are constants $c_1$ and $c_2$ such that

$$c_1 T^{1/6}(f) \leq \tilde{\deg}(f) \leq c_2 T(f).$$

The exponent $1/6$ is not known to be optimal. The conjectured value is $1/2$.

As observed by Buhrman, Cleve, and Wigderson [25], a decision tree algorithm can be turned into a communication protocol for a related problem. In such a protocol for $f_n \square g_k$, one party simulates the decision tree algorithm for $f_n$, and initiates a sub-protocol for computing $g_k$ whenever one input bit of $f_n$ is needed. The sub-protocol repeats an optimal protocol for $g_k$ for $O(\log \tilde{\deg}(f_n))$ times, ensuring that the error probability is $\leq \frac{1}{3c_1 \deg(f_n)}$. Thus the larger protocol computes $f_n \square g_k$ with error probability $\leq 1/3$, and exchanges $O(R(g_k) \tilde{\deg}(f_n) \log \tilde{\deg}(f_n))$ bits.

Proposition 3.9 ([25, 11]). For any function $f_n \in \mathcal{F}_n$ with $\tilde{\deg}(f_n) = d$, and any $g_k \in \mathcal{G}_k$, $R(f_n \square g_k) = O(R(g_k) d^6 \log d)$.

3.3 The Main Lemma

In this section, we prove that under some assumptions, $Q(f_n \square g_k) = \Omega(\tilde{\deg}(f_n))$. This is shown by turning a “witness” for $f_n$ requiring a high approximate degree into a “witness” for the hardness of $f_n \square g_k$.

3.3.1 Witness of high approximate degree

We now fix a function $f_n \in \mathcal{F}_n$ with $\tilde{\deg}_\epsilon(f_n) = d$. For $w \in \{0,1\}^n$, denote by $\chi_w \in \mathcal{F}_n$ the function $\chi_w(x) = (-1)^{w \cdot x}$. Then there is no feasible solution to the following linear system, where the unknowns are $\alpha_w$:

$$-\epsilon + f(x) \leq \sum_{w:|w|<d} (-1)^{w \cdot x} \alpha_w \leq f(x) + \epsilon, \quad \forall x \in \{0,1\}^n.$$
By the duality of linear programming, there exist $q_x^+ \geq 0$ and $q_x^- \geq 0$, $x \in \{0, 1\}^n$, such that

$$\sum_x (q_x^+ - q_x^-) \cdot x = 0, \quad \forall w, |w| < d, \quad \text{and},$$

(3.2) $$\sum_x (q_x^+ - q_x^-) f(x) + \epsilon (q_x^+ + q_x^-) < 0.$$

Define $q : \{0, 1\}^n \to \mathbb{R}$ as $q(x) = q_x^- - q_x^+$. Then

$$q^T \chi_w = 0, \quad \text{and}, \quad \|q\|_1 < \frac{1}{\epsilon} q^T f.$$

Without loss of generality, assume that $q^T f = 1$ (otherwise this will hold after multiplying $q$ with an appropriate positive number). Then $\|q\|_1 < 1/\epsilon$.

Since $q$ is orthogonal to all polynomials of degree less than $d$, it has non-zero Fourier coefficients only on higher frequencies:

$$q = \sum_{w:|w| \geq d} \hat{q}_w \chi_w,$$

where

$$\hat{q}_w = \frac{1}{N} \sum_x q(x) \chi_w(x).$$

Since $\|q\|_1 < 1/\epsilon$, those Fourier coefficients must be small:

(3.3) $$|\hat{q}_w| < \frac{1}{N \epsilon}, \quad \forall w : |w| \geq d.$$

We summarize the above discussion in the following lemma.

**Lemma 3.10.** Let $\epsilon \in \mathbb{R}$, $0 \leq \epsilon < 1/2$. For any $f \in \mathcal{F}_n$, there exists a function $q : \{0, 1\}^n \to \mathbb{R}$ such that: (a) $q^T f = 1$, (b) $\|q\|_1 < 1/\epsilon$, (c) $|\hat{q}_w| \leq \frac{1}{N \epsilon}$, for all $w \in \{0, 1\}^n$, and (d) $\hat{q}_w = 0$ whenever $|w| < \tilde{\deg}_t(f_n)$. 
3.3.2 Witness of large approximate trace norm

In order to convert a witness of high approximate degree for \( f_n \) to that of large approximate trace norm for \( f_n \Box g_k \), we need to require that \( g_k \) satisfies certain property, which we now formulate. Let \( I_A, I_B \subseteq \{0, 1\}^k \). For \( b \in \{0, 1\} \), a matrix \( \mu \in \mathbb{R}^{I_A \times I_B} \) is said to be a \( b \)-distribution for \( g_k \) if

1. \( \mu(x, y) \geq 0, \forall (x, y) \in I_A \times I_B \),
2. \( \sum_{(x, y) \in I_A \times I_B} = 1 \), and,
3. \( \mu(x, y) = 0, \forall (x, y) \in I_A \times I_B \cap g_k^{-1}(1 - b) \).

**Definition 3.11.** The strong discrepancy of \( g_k \in \mathcal{G}_k \), denoted by \( \rho(g_k) \), is the minimum \( r \in \mathbb{R} \) such that there exist \( I_A, I_B \subseteq \{0, 1\}^k \), and \( b \)-distributions \( \mu_b \in \mathbb{R}^{I_A \times I_B} \) for \( g_k, b \in \{0, 1\} \), satisfying the following conditions.

1. \( \sqrt{|I_A| \cdot |I_B|} \cdot \|\mu_0 + \mu_1\| \leq 1 + r \), and,
2. \( \sqrt{|I_A| \cdot |I_B|} \cdot \|\mu_0 - \mu_1\| \leq r \).

It follows from the definition of discrepancy of a Boolean matrix (c.f. pp. 38 [64]) that the strong discrepancy is at least as large as the discrepancy. Thus, the following proposition follows from the discrepancy lower bound for quantum communication complexity.

**Proposition 3.12.** For any \( g_k \in \mathcal{G}_k \), \( Q(g_k) = \Omega(\log \frac{1}{\rho(g_k)}) \).

We are now ready to state and prove our Main Lemma.

**Lemma 3.13 (Main Lemma).** Let \( n, k \geq 1 \) be integers, \( g_k \in \mathcal{G}_k \), and \( f_n \in \mathcal{F}_n \). If \( \rho(g_k) \leq \frac{\deg(f_n)}{2n} \), then \( Q(f_n \Box g_k) = \Omega(\tilde{\deg}(f_n)) \).

**Proof.** Let \( d \overset{\text{def}}{=} \tilde{\deg}(f_n) \), and \( F \overset{\text{def}}{=} f_n \Box g_k \). Suppose \( \rho \overset{\text{def}}{=} \rho(g_k) \) is achieved with \( I_A, I_B \subseteq \{0, 1\}^k \), and \( \mu_b, b \in \{0, 1\} \). Denote \( K_A \overset{\text{def}}{=} |I_A|, K_B \overset{\text{def}}{=} |I_B| \). Let \( F_1 \) be the
restriction of $f_n \Box g_k$ on $(I_A \times I_B)^{\otimes n} \cap \text{dom}(F)$. We shall prove the desired lower bound on $F_1$. By Lemma 3.7, it suffices to prove a lower bound on $\|F_1\|_{\epsilon', \text{tr}}$ for $\epsilon' = 1/6$.

Let $q$ be the function that exists by Lemma 3.10 with respect to $f_n$ and $\epsilon = 1/3$.

For a set $w \subseteq [n]$, and a $K_A \times K_B$ matrix $A$, by $A^{\otimes w}$ we mean putting $A$ in each component $i \in w$ in the tensor product space $(\mathbb{R}^{K_A \times K_B})^{\otimes n}$. Denote by $\bar{w}$ the complement of $w$. Define $h \in (\mathbb{R}^{K_A \times K_B})^{\otimes n}$ as follows

$$h \overset{\text{def}}{=} \sum_{z \in \{0, 1\}^n} q(z) \cdot \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}}.$$  

Then $\|h\|_1 = \sum_{x,y} |h_{x,y}| = \|q\|_1 \leq 1/\epsilon$, and $\text{tr}(h^T F) = q^T f_n = 1$. Fix an $\tilde{F} \in (\mathbb{R}^{K_A \times K_B})^{\otimes n}$ with $|F_1(x, y) - \tilde{F}(x, y)| \leq \epsilon'$, $\forall (x, y) \in \text{dom}(F_1)$. Then,

$$|\text{tr}(h^T \tilde{F})| = \left| \sum_{(x,y) \in \text{dom}(F_1)} h(x, y) \tilde{F}(x, y) \right|$$

$$\geq \left| \sum_{(x,y) \in \text{dom}(F)} h(x, y) F(x, y) \right| - \epsilon' \|h\|_1$$

$$\geq 1 - \epsilon'/\epsilon$$

$$\geq 1/2.$$  

Therefore,

$$(3.4) \quad \|\tilde{F}\|_{\text{tr}} \geq \frac{|\text{tr}(h^T \tilde{F})|}{\|h\|} \geq \frac{1}{2\|h\|}. $$

Hence we need only to prove that $\|h\|$ is very small. To this end we first express $h$ using the Fourier representation of $q$:  

$$h = \sum_{w:|w| \geq d} \hat{q}_w (-1)^{w \cdot z} \cdot \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}}$$

$$= \sum_{w:|w| \geq d} \hat{q}_w \cdot \sum_{z \in \{0, 1\}^n} (-1)^{w \cdot z} \cdot \bigotimes_{i=1}^n \mu_{z_i}^{\otimes \{i\}}$$

$$= \sum_{w:|w| \geq d} \hat{q}_w \cdot ((\mu_0 + \mu_1)^{\otimes \bar{w}}) \otimes ((\mu_0 - \mu_1)^{\otimes w}).$$
Using \( \hat{q}_w \leq 1/\epsilon N \),

\[
\|h\| \leq \sum_{w:|w|\geq d} |\hat{q}_w| \|\mu_0 + \mu_1\|^{n-|w|} \cdot \|\mu_0 - \mu_1\|^{|w|}
\]

\[
\leq \frac{1}{\epsilon} \sum_{\ell, \ell \geq d} \binom{n}{\ell} \cdot \left( \frac{\mu_0 + \mu_1}{2} \right)^{n-|w|} \cdot \left( \frac{\mu_0 - \mu_1}{2} \right)^{|w|}.
\]

(3.5)

By the choice of \( \mu_0 \) and \( \mu_1 \), \( \|\mu_0 + \mu_1\| \leq \frac{1+\rho}{\sqrt{K_A K_B}} \), and \( \|\mu_0 - \mu_1\| \leq \frac{\rho}{\sqrt{K_A K_B}} \). Thus

\[
\|h\| \leq \frac{(1+\rho)^n}{\epsilon (K_A K_B)^{n/2}} \sum_{\ell, \ell \geq d} \binom{n}{\ell} \rho^\ell.
\]

(3.6)

If \( \rho \leq \frac{d}{2e^{\ell}} \), using \( \binom{n}{\ell} \leq \left( \frac{en}{\ell} \right)^\ell \), and \( (1+\rho)^n \leq e^{\ell m} \), we have

\[
\|h\| \leq \frac{e^{\ell m}}{\epsilon (K_A K_B)^{n/2}} \sum_{\ell, \ell \geq d} \left( \frac{en}{\ell} \right)^\ell \rho^\ell \leq \frac{e^{\ell m}}{\epsilon (K_A K_B)^{n/2}} \sum_{\ell, \ell \geq d} \left( \frac{d}{2\ell} \right)^\ell \leq \frac{2}{\epsilon (K_A K_B)^{n/2}} \cdot e^{-(\ln 2 - 1/(2e))d} \leq \frac{2}{\epsilon (K_A K_B)^{n/2}} e^{-0.5d}.
\]

(3.8)

Together with Equation 3.4, this implies

\[
\|\hat{F}\| \geq \frac{\epsilon}{4} \cdot (K_A K_B)^{n/2} \cdot e^{-0.5d}.
\]

Thus \( \|F_1\|_{1/\epsilon, \text{tr}} \geq \frac{1}{24} \cdot (K_A K_B)^{n/2} \cdot e^{-0.5d} \). Plugging this inequality to the Razborov-Yao Lemma, we have \( Q(F) \geq Q(F_1) = \Omega(d) \).

\[
\Box
\]

3.4 Applications

We now apply the Main Lemma to derive two quantum lower bounds. The first deals with those \( g_k \) that have polynomially related quantum and randomized communication complexities. As a concrete example we consider \( g_k \) being the INNER PRODUCT function. The second result shows that without this knowledge on \( g_k \),
we may still be able to obtain strong quantum lower bounds. This is done through a “hardness amplification” technique that makes use of the self-similarity of the function considered. We demonstrate this technique by giving an alternative proof of Theorem 3.6 with a weaker parameter.

### 3.4.1 Composition with hard \( g_k \)

We now restate Theorem 3.4 rigorously.

**Theorem 3.14.** Let \( n, k \geq 1 \) be integers and \( g_k \in G_k \). If \( Q(g_k) \) and \( R(g_k) \) are polynomially related, so is \( Q(f_n \circ g_k) \) and \( R(f_n \circ g_k) \) for any \( f_n \in F_n \) and for \( \rho(g_k) \leq \frac{1}{2en} \).

Proof. If \( f_n \) or \( g_k \) is a constant function, \( Q(f_n \circ g_k) = R(f_n \circ g_k) = 0 \), hence the statement holds. Otherwise, one can fix the value of all but one input block so that \( f_n \circ g_k \) computes \( g_k \) on the remaining block. Thus \( Q(f_n \circ g_k) \geq Q(g_k) \). By Main Lemma, under the assumption that \( \rho(g_k) \leq \frac{1}{2en} \), \( Q(f_n \circ g_k) = \Omega(\widetilde{\deg}(f_n)) \). Thus \( Q(f_n \circ g_k) = \Omega(\widetilde{\deg}(f_n) + Q(g_k)) \). On the other hand \( R(f_n \circ g_k) = O(R(g_k)\widetilde{\deg}(f_n) \log \widetilde{\deg}(f_n)) \), by Proposition 3.9. Thus, under the assumption that \( R(g_k) \) and \( Q(g_k) \) are polynomially related, so are \( Q(f_n \circ g_k) \) and \( R(f_n \circ g_k) \). \( \square \)

Similarly, the same statement holds with \( R(f_n \circ g_k) \) and \( R(g_k) \) replaced by \( D(f_n \circ g_k) \) and \( D(g_k) \), respectively. Estimating \( \rho(g_k) \) is unfortunately difficult in general. However, if we can show \( \rho(g_k) = \exp(-\Omega(k^c)) \) for some constant \( c \), it implies \( R(g_k) \) and \( Q(g_k) \) are polynomially related, by Proposition 3.12. Thus \( Q(f_n \circ g_k) \) and \( R(f_n \circ g_k) \) are polynomially related for \( k \geq \log_2^{1/c}(2en) \).

We now prove Corollary 3.5.

**Proof of Corollary 3.5.** We need only to consider the case that \( f_n \) is not a constant function. Then \( Q(f_n \circ g_k) = \Omega(\text{IP}_k) \). It is known that \( Q(\text{IP}_k) = \Omega(k) \) [34]. Thus
$Q(f_n □ g_k) = \Omega(k)$. Let $K \stackrel{\text{def}}{=} 2^k$, $I_A \stackrel{\text{def}}{=} \{0,1\}^k - \{0^k\}$, and $I_B \stackrel{\text{def}}{=} \{0,1\}^k$. For $b \in \{0,1\}$, let $\mu_b$ be the uniform distribution on $\{(x,y) : \text{IP}(x,y) = b, x \neq 0\}$. Then $\|\frac{\mu_0 + \mu_1}{2}\| = 1/\sqrt{K(K-1)}$, and $\|\frac{\mu_0 - \mu_1}{2}\| = 1/((K-1)\sqrt{K})$. Thus $\rho(\text{IP}_k) \leq 1/\sqrt{K-1}$. When $k \geq 2 \log_2 n + 5 > \log_2(4e^2n^2+1)$, $\rho(\text{IP}_k) \leq 1/2en \leq \tilde{\deg}(f_n)/(2en)$.

By Main Lemma 3.13, this implies $Q(f_n □ \text{IP}_k) = \Omega(\tilde{\deg}(f_n))$. Therefore, $Q(f_n □ \text{IP}_k) = \Omega(k + \tilde{\deg}(f_n))$. On the other hand, $D(f_n □ \text{IP}_k) = O(k \tilde{\deg}(f_n))$. Thus $D(f_n □ \text{IP}_k) = O(Q7(f_n □ \text{IP}_k))$.

We remark that since for a random $g_k$, $\rho(g_k) = \exp(-\Omega(k))$, the above corollary holds for most $g_k$ up to a constant additive difference in the bound for $k$.

### 3.4.2 Composition with Set Disjointness

In this section we prove Theorem 3.6. We introduce some notions following [77].

For an integer $k \geq 1$, let $[k] \stackrel{\text{def}}{=} \{1,2,\cdots,k\}$. For an integer $p$, $0 \leq p \leq k$, denote by $[k]^p$ the set of $p$-element subsets of $[k]$. For integers $s$ and $p$ with $0 \leq s \leq p \leq k/2$, denote by $J_{k,p,s} \in \{0,1\}^{[k]^p \times [k]^p}$ the indicator function for $|x \cap y| = s$. That is, for any $(x,y) \in [k]^p \times [k]^p$,

$$(J_{k,p,s})_{x,y} \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } |x \cap y| = s, \\ 0 & \text{otherwise}. \end{cases}$$

The spectrum of these combinatorial matrices are described by Hahn polynomials [39]. We will use a formula given by Knuth [62].

**Proposition 3.15 (Knuth).** Let $p \leq k/2$. Then the matrices $J_{k,p,s}$, $0 \leq s \leq p$, share the same eigenspaces $E_0$, $E_1$, $\ldots$, $E_p$, and the eigenvalue corresponding to the eigenspace $E_t$, $0 \leq t \leq p$, is given by

$$\sum_{i=\max\{0,s+t-p\}}^{\min\{s,t\}} (-1)^{t-i} \binom{t}{i} \binom{p-t+i}{s-i} \binom{k-p-t+i}{p-s-t+i}. \quad (3.9)$$
Lemma 3.16. Let $n, k \geq 1$ be integers, $f_n \in \mathcal{F}_n$, and $k \geq \frac{6\text{en}}{\text{deg}(f_n)}$. Then $Q(f_n \square \text{DISJ}_k^{\leq 1}) = \Omega(\text{deg}(f_n))$.

Proof. Let $p \overset{\text{def}}{=} k/3$ and $M \overset{\text{def}}{=} \binom{k}{p}$. Let $w_s \overset{\text{def}}{=} |(\text{DISJ}_k^{\leq 1})^{-1}(s)|$, $s \in \{0, 1\}$. That is,

$$w_0 = \binom{k}{p} \binom{k-p}{p} = M \binom{k-p}{p}, \quad \text{and,} \quad w_1 = \binom{k}{p} \binom{p}{1} \binom{k-p}{p-1} = M \binom{p}{1} \binom{k-p}{p-1}.$$ 

Let $\mu_s$, $s \in \{0, 1\}$, be the distribution matrix for the uniform distribution on the $s$-inputs of $\text{DISJ}_k^{\leq 1}$. That is,

$$\mu_0 \overset{\text{def}}{=} \frac{1}{w_0} J_{k,p,0}, \quad \text{and,} \quad \mu_1 \overset{\text{def}}{=} \frac{1}{w_1} J_{k,p,1}.$$ 

By Proposition 3.15, $\mu_0$ and $\mu_1$ have the same eigenspaces. Furthermore, if $\lambda_{s,t}$, $s \in \{0, 1\}$ and $0 \leq t \leq p$, is the eigenvalue of $\mu_s$ for the eigenspace $E_t$,

$$\lambda_{s,t} = \frac{1}{w_s} \min_{i=\max\{0,s+t-p\}} \sum_{i} (-1)^{t-i} \binom{t}{i} \binom{p-i}{s-i} \binom{k-p-t+i}{p-s-t+i},$$

and

$$\|\mu_0 - \mu_1\| = \max_{0 \leq t \leq p} |\lambda_{0,t} - \lambda_{1,t}|.$$ 

After simplification,

$$\lambda_{0,t} = \frac{(-1)^t \binom{k-p-t}{p-t}}{M},$$

and,

$$\lambda_{1,t} = \frac{(-1)^t \binom{k-p-t}{p-t} - t \binom{k-p-t+1}{p-t+1}}{M}.$$ 

Since $\lambda_{0,0} = \lambda_{1,0} = 1$, we only need to bound $\max_t |\lambda_{0,t} - \lambda_{1,t}|$ for $t \geq 1$.

$$\lambda_{0,t} - \lambda_{1,t} = \frac{(-1)^t \binom{k-p-t}{p-t}}{M} (1 - \frac{p-t}{p} + \frac{t(k-p-t+1)}{p^2}) = \frac{(-1)^t \frac{1}{p} \binom{k-p-t}{p-t}}{M} \frac{t(k-t+1)}{p^2}.$$ 

We actually need only to consider $s \in \{0, 1\}$. Effectively, we are restricting $\text{DISJ}_k$ on $\{(u, v) : u, v \in [k]^p, |u \cap v| \leq 1\}$. Denote this restriction by $\text{DISJ}_k^{\leq 1}$.
Using $k = 3p$,

$$t\binom{k-p-t}{p-t} = \frac{t \cdot p \cdot (p-1) \ldots (p-t+1)}{(k-p) \cdot (k-p-1) \ldots (k-p-t+1)} \leq \left( \frac{p}{k-p} \right)^t \cdot t \leq \left( \frac{1}{2} \right)^t \leq \frac{1}{2}. $$

Hence

$$|\lambda_{0,t} - \lambda_{1,t}| \leq \frac{1}{2} \cdot \frac{k-t+1}{Mp^2} = \frac{1}{2} \cdot \frac{k}{M\left(\frac{k}{3}\right)^2} \leq \frac{6}{Mk}. $$

Combining Equations 3.11 and 3.13, we have

$$M\left\| \frac{\mu_0 - \mu_1}{2} \right\| \leq \frac{3}{k}. $$

Since $\frac{\mu_0 + \mu_1}{2}$ is doubly stochastic,

$$\left\| \frac{\mu_0 + \mu_1}{2} \right\| = 1. $$

Thus $\rho(g_k) \leq 3/k$. Therefore, when $k \geq 6en/d$, we have $\rho(g_k) \leq d/(2en)$. By Main Lemma 3.13, this implies $Q(f_n \Box \text{DISJ}^<_k) = \Omega(\widetilde{\deg}(f_n))$. \hfill \Box

Let $f_n \in \mathcal{F}_n$ be a symmetric function. Following [77], define

$$\ell_0(f_n) \overset{\text{def}}{=} \max\{m : 1 \leq m \leq n/2, f_n(1^m 0^{n-m}) \neq f_n(1^{m-1} 0^{n-m+1})\} \cup \{0\},$$

and

$$\ell_1(f_n) \overset{\text{def}}{=} \max\{n-m : n/2 \leq m \leq n, f_n(1^m 0^{n-m}) \neq f_n(1^{m+1} 0^{n-m-1})\} \cup \{0\}.$$

We will use the following result in proving quantum lower bounds on $f_n \Box \wedge$.

**Theorem 3.17 (Paturi [72]).** Let $f_n \in \mathcal{F}_n$ be symmetric. Then for some universal constant $c$, $\widetilde{\deg}(f_n) \geq c\sqrt{n(\ell_0(f_n) + \ell_1(f_n))}$. 

Theorem 3.18. For any symmetric \( f_n \in \mathcal{F}_n \), \( Q(f_n \land) = \Omega(n^{1/3} \ell_0^{2/3} f_n + \ell_1(f_n)) \).

Proof. Let \( c \) be the constant in Theorem 3.17, \( \beta \overset{\text{def}}{=} \min\{\sqrt{23}, (\frac{\beta}{12e})^{2/3}\} \), and \( \alpha \overset{\text{def}}{=} (\beta/2)^{2/3} \). Suppose that \( \ell_0 \overset{\text{def}}{=} \ell_0(f_n) \leq \alpha n \). Let \( n' \overset{\text{def}}{=} \beta n^{2/3} \ell_0^{1/3} \), and \( f_{n'} \in \mathcal{F}_{n'} \) be such that \( f_n(x) = f_{n'}(x0^n-n') \), \( \forall x \in \{0,1\}^n \). By direct inspection, \( n' \leq n \), thus \( f_{n'} \) is well-defined. Since

\[
\forall x,y \in \{0,1\}^n, \quad f_{n'}(x0^n-n') = f_{n'}(x0^n-n'),
\]

and by direct inspection, \( \ell_0 \leq n'/2 \), we have \( \ell_0(f_{n'}) \geq \ell_0 \). By Theorem 3.17,

\[
\overline{\deg}(f_{n'}) \geq c\sqrt{n'(\ell_0(f_{n'}) + \ell_1(f_{n'}))} \geq c\sqrt{n'\ell_0}.
\]

Set \( k \overset{\text{def}}{=} \left\lceil \frac{6cn'}{\deg(f_{n'})} \right\rceil \). By Lemma 3.16, \( Q(f_{n'} \Box \text{DISJ}_k^{\leq 1}) = \Omega(\overline{\deg}(f_{n'})) = \Omega(n^{1/3} \ell_0^{2/3}) \).

Note that

\[
n'k \leq \beta n^{2/3} \ell_0^{1/3} \cdot \frac{12\sqrt{3}}{c} \left( \frac{n}{\ell_0} \right)^{1/3} = \beta \cdot 3/2 \cdot \frac{12\sqrt{3}}{c} n \leq n.
\]

Therefore, \( \forall (x, y) \in \text{dom}(f_{n'} \Box \text{DISJ}_k^{\leq 1}), (f_{n'} \Box \text{DISJ}_k^{\leq 1})(x, y) = (f_n \land)(x0^n-n'k, y0^n-n'k) \).

Thus \( Q(f_n \land) \geq Q(f_{n'} \Box \text{DISJ}_k^{\leq 1}) = \Omega(n^{1/3} \ell_0^{2/3}) \).

Now consider the case that \( \alpha n < \ell_0 \leq n/2 \). Set \( k \overset{\text{def}}{=} \left\lceil \frac{6\sqrt{2e}}{c} \right\rceil \), and \( n' \overset{\text{def}}{=} \min\{\frac{n-\ell_0+1}{2k-1}, \ell_0 - 1\} \). Then \( n' = \Theta(n) = \Theta(\ell_0) \). Define \( f_{n'} \in \mathcal{F}_{2n'} \) as follows:

\[
f_{n'}(x) = f_n(x1^{\ell_0-1-n'0^n-2n'-(\ell_0-1-n')}), \quad \forall x \in \{0,1\}^{2n'}.
\]

By direction inspection, \( f_{n'} \) is well-defined. Then

\[
f_{n'}(1^{n'}0^n) = f_n(1^{\ell_0-1-0^n-\ell_0+1}) \neq f_n(1^{\ell_00^n-\ell_0}) = f_{n'}(1^{n'+10^n-1}).
\]

Therefore, \( \ell_1(f_{n'}) = n' \), and \( \overline{\deg}(f_{n'}) \geq \sqrt{2cn'} \), by Theorem 3.17. By direct inspection, \( k \geq \frac{6e(2n')}{\deg(f_{n'})} \), thus \( Q(f_{n'} \Box \text{DISJ}_k^{\leq 1}) = \Omega(\overline{\deg}(f_{n'})) = \Omega(n') \). Note that for all \( (x, y) \in \text{dom}(f_{n'} \Box \text{DISJ}_k^{\leq 1}), (f_{n'} \Box \text{DISJ}_k^{\leq 1})(x, y) = (f_n \land)(x1^{\ell_0-1-n'0^n-(\ell_0-1-n')-2kn'}, y1^{\ell_0-1-n'0^n-(\ell_0-1-n')-2kn'}). \)
By direct inspection, the number of 0’s and 1’s padded in the above equation is non-negative. Thus
\[ Q(f_n \square \land) = \Omega(Q(f_n \square \text{DISJ}^{\leq 1}_k)) = \Omega(n^\prime) = \Omega(\ell_1) = \Omega(n^{1/3} \ell_0^{2/3}). \]

We use a similar reduction to prove \( Q(f_n \square \land) = \Omega(\ell_1) \). Let \( k \) be the same as above. Set \( n^\prime \overset{\text{def}}{=} \left\lfloor \frac{\ell_1}{2k-1} \right\rfloor \), and define \( f_n^\prime \in \mathcal{F}_{2n^\prime} \) as follows
\[ f_n^\prime(x) = f_n(x^{1n-\ell_1-n^\prime}0^{n-2n^\prime-(n-\ell_1-n^\prime)}) \quad \forall x \in \{0, 1\}^{2n^\prime}. \]

By direct inspection, the numbers of padded 0’s and 1’s are non-negative, thus \( f_n^\prime \) is well-defined. Since
\[ f_n^\prime(1n^\prime0^\prime) = f_n(1n-\ell_10^{\ell_1}) \neq f_n(1n-\ell_1+10^{n-\ell_1-1}) = f_n^\prime(1n^\prime+10^\prime-1), \]
we have \( \ell_1(f_n^\prime) = n^\prime \). Thus \( \tilde{\deg}(f_n^\prime) \geq \sqrt{2cn^\prime} \) by Theorem 3.17, and \( Q(f_n^\prime \square \text{DISJ}^{\leq 1}_k) = \Omega(\tilde{\deg}(f_n^\prime)) = \Omega(\ell_1) \) by Lemma 3.16. For all \((x, y) \in \text{dom}(f_n^\prime \square \text{DISJ}^{\leq 1}_k), \)
\[ (f_n^\prime \square \text{DISJ}^{\leq 1}_k)(x, y) = (f_n \square \land)(x^{1n-\ell_1-n^\prime}0^{n-2kn^\prime-(n-\ell_1-n^\prime)}, y^{1n-\ell_1-n^\prime}0^{n-2kn^\prime-(n-\ell_1-n^\prime)}). \]

By direct inspection again, the numbers of the padded digits in the above are non-negative. Thus \( Q(f_n \square \land) \geq Q(f_n^\prime \square \text{DISJ}^{\leq 1}_k) = \Omega(\ell_1). \)

Next, we establish a classical upper bound on the randomized complexity of symmetric predicates. We will use the protocol for the Hamming Distance problem from Chapter II.

**Proposition 3.19.** Let \( f_n \in \mathcal{F}_n \) be symmetric with \( \ell_0(f_n) = 0 \). Then
\[ R(f_n \square \land) = O(\ell_1 \log^2 \ell_1 \log \log \ell_1). \]

**Proof of Proposition 3.19.** Without loss of generality, assume \( f_n(1^m0^{n-m}) = 0 \) for all \( m, 0 \leq m \leq n - \ell_1 \). The following randomized protocol computes \( f_n \square \land \) with
(ℓ_1 \log \ell_1 \log \log \ell_1$) bits of communication. Fix an input \((x, y)\), and let \(z_A \overset{\text{def}}{=} n - |x|\) and \(z_B \overset{\text{def}}{=} n - |y|\). Alice and Bob first check if \(z_A \geq \ell_1\) or \(z_B \geq \ell_1\). If yes, they output 0 and terminate the protocol. Otherwise, Alice sends \(z_A\) to Bob using \(\lceil \log_2 (\ell_1 - 1) \rceil\) bits, and they compute \(\delta \overset{\text{def}}{=} |x \oplus y|\). Knowing \(z_A\) and \(\delta\), Bob is able to compute \(f(|x \cap y|) = f((|x| + |y| - |x \oplus y|)/2)\). Note that \(\Delta \overset{\text{def}}{=} 2(\ell_1 - 1) \geq \delta \geq 0\). Thus Alice and Bob can perform a binary search to determine \(\delta\) with \(\log_2 (\Delta + 1)\) sub-protocols for the \textsc{Hamming Distance Problem}. For each candidate value \(d\) of \(\delta\), they repeat the randomized protocol in Theorem 2.5 for \(\text{HAM}_{n,d}\) for \(\Theta(\log \log \Delta)\) times so that the error probability is \(\leq \frac{1}{3(\log_2 \Delta + 1)}\). Thus the total number of bits exchanged is \(O(\Delta \log^2 \Delta \log \log \Delta) = O(\ell_1 \log^2 \ell_1 \log \log \ell_1)\), and the error probability of the complete protocol is \(\leq 1/3\).

Theorem 3.6 follows from Theorem 3.18 and Proposition 3.19 straightforwardly.

**Remark 3.20.** While both Razborov’s proof and the above use the spectrum decompositions of the matrix \(J_{k,p,s}\), we emphasize their difference: we only need to analyze \(\|\mu_0 - \mu_1\|\), which corresponds to \(s = 0, 1\). In contrast, Razborov’s proof needs much more details of the spectrum decompositions, in particular, it needs to consider \(s = 0, 1, \cdots, \Theta(n)\).

As a result of considering only \(s = 0\) and \(s = 1\), our estimation of \(\rho(\text{DISJ}_k)\) only gives a \(\Omega(\log k)\) lower bound on \(\text{DISJ}_k\). This very weak bound \((\Omega(\log n)\) when \(k = \Theta(\sqrt{n})\), can be, surprisingly, amplified to \(\Omega(n^{1/3})\) through the duality machinery of the polynomial method. Finding more examples of such “hardness amplification” would be very interesting.
CHAPTER IV

Classical simulations of nonlocal quantum measurements

This chapter is based on [85]. We quantify nonlocalness of a bipartite measurement by the minimum amount of classical communication required to simulate the measurement. We derive general upper bounds, which are expressed in terms of certain tensor norms of the measurement operator. As applications, we show that (a) if the amount of communication is constant, quantum and classical communication protocols with an unlimited amount of shared entanglement or shared randomness compute the same set of functions; (b) a local hidden variable model needs only a constant amount of communication to create, within an arbitrarily small statistical distance, a distribution resulting from local measurements of an entangled quantum state, as long as the number of measurement outcomes is constant.

4.1 Summary of results

Recall that \( \text{Com}(Q) \) is the minimum number of bits that need to be exchanged by the simulating communication process. Our main result is to derive a general upper bound on \( \text{Com}(Q) \) in terms of a certain operator norm \( \|Q\|_{\circ} \) on \( Q \), which is bounded from above polynomially in \( Q \)'s dimension.

**Theorem 4.1 (Informally).** For any bipartite quantum measurement \( Q \), \( \text{Com}(Q) = O\left(\|Q\|_2^2\right) \). In particular, if \( K \) is the dimension of the space that \( Q \) acts on, \( \text{Com}(Q) = \)

46
The diamond norm \( \|Q\|_\diamond \) is originally defined on superoperators, and has been a powerful tool in the study of quantum interactive proof systems [58] and quantum circuits on mixed states [2]. We make use a natural mapping from bipartite operators to superoperators to define norms on the former based on norms on the latter.

The approach in proving Theorem 4.1 can be extended to obtain general upper bounds on \( \text{Com}(Q) \) in terms of other operators norms. Those norms belong to so called tensor norms, i.e., norms \( \| \cdot \|_\alpha \) that satisfies \( \|P\|_\alpha = \|A\| \cdot \|B\| \), whenever \( P = A \otimes B \). Tensor norms have been studied for decades with a great deal of rich concepts and deep results (see, e.g., [38]). In recent years, they have been applied to quantum information theory to characterize and quantify the nonlocality of quantum states [78, 81]. The tensor norms that appear in our upper bounds capture the nonlocality of bipartite operators in their own way, and may have further applications.

### 4.1.1 Applications on quantum communication complexity

After obtaining those general upper bounds, we show that they in turn have useful applications on quantum communication complexity. Recall that in the setting of communication complexity [95, 96], Alice and Bob wish to compute a function \( f(x, y) \), where \( x \) is known to Alice only, and \( y \) is known only to Bob. The communication complexity of \( f \) is the minimum amount of information that Alice and Bob need to exchange in order to compute \( f \) correctly for any input. Communication complexity has been a major research field (see, e.g., the book [64]), with many problems of rich structures and deep connections to other aspects of complexity theory.

A concrete application of our result is on the advantage of sharing entanglement
in quantum protocols. If there is a quantum protocol that exchanges \( q \) qubits with \( m \) qubits of prior entanglement, then the best classical simulation we know is \( \exp(\Omega(q + m)) \). This is embarrassingly large, especially when \( q << m \). Using our upper bound on the classical communication complexity of nonlocal operators, we prove the following result. Note that in the *Simultaneous Message Passing (SMP)* model with shared randomness, the two parties holding the inputs share an arbitrarily long random string, and each send a single message to a third party, who is required to determine the outcome correctly with high probability.

**Theorem 4.2.** If a twoway quantum protocol uses \( q \) qubits of communication and \( m \) qubits of share entanglement, then it can be simulated by a classical protocol using \( \exp(O(q)) \) bits with shared randomness. The simulation does not depend on \( m \). Furthermore, it can be carried out in the SMP model with shared randomness.

Notice that the exponential dependence on \( q \) can not be improved, because of the existence of an exponential separation of quantum and classical communication complexities for some partial function, discovered by Raz [75]. As a consequence of the above theorem,

**Corollary 4.3.** If a communication complexity problem has a constant cost quantum communication protocol with shared entanglement, it also has a constant cost classical SMP protocol with shared randomness.

It is interesting to contrast the above with a recent result by Yao [97], which is of a similar type but of the opposite direction.

**Theorem 4.4 ([97]).** If a communication complexity problem of input size \( n \) has a constant cost classical SMP protocol with shared randomness, it has an \( O(\log n) \) cost quantum SMP protocol without shared entanglement.
Combining this result with ours, we have

**Corollary 4.5.** If a communication complexity problem of input size $n$ has a constant cost twoway quantum protocol with shared entanglement, it has an $O(\log n)$ cost quantum SMP protocol without shared entanglement.

4.1.2 Applications on simulating quantum correlations

Yet another application of our classical simulation of quantum measurements is to give efficient simulations of quantum correlations by the hidden variable model assisted with classical communication. The scenario is as follows. Suppose Alice and Bob are given an entangled quantum state. Then each of them, without any communication, applies to their portion of the state some local measurement not known to the other party. The result is a correlated joint distribution on both measurement outcomes. There are such correlations that violate the Bell Inequalities, hence impossible to generate by any reasonable classical procedure in which Alice and Bob do not communicate.

A natural next step to extend the above work of Bell is to investigate the minimum amount of classical communication required to simulate a quantum correlation. Most of the works addressing this question focus on the exact simulation and on measuring a constant number of qubits [93, 8, 35, 91, 21, 66]. We study the approximate and asymptotic simulation of quantum correlations, where the joint random variables take a constant number of possible values but are nevertheless produced from (the two party) sharing an entangled state of an arbitrary dimension and applying arbitrary local measurements.

**Theorem 4.6 (Informally).** In the above scenario, a $O(\ln \frac{1}{\epsilon} / \delta^2)$ number of classical bits is sufficient to approximate the quantum correlation with a $\delta$ statistical
distance and 1 − ε probability.

The rest of the chapter is organized as follows. We start with a general framework for classical simulations of quantum protocols. Then we optimize the cost parameter of this framework is then optimized and give the main theorem. In the section that follows we give applications of the main theorem.

4.2 A simulation framework

Our classical simulation of quantum protocols falls into the following framework. Let \( p \) be the acceptance probability (i.e., the probability of outputting 1) of a given quantum protocol (which arises either from a communication task or from a bipartite measurement). We express \( p = \langle \psi_A | \psi_B \rangle \), for two vectors \(|\psi_A\rangle\) and \(|\psi_B\rangle\) that can be prepared by Alice and Bob by herself/himself. Note that the lengths of the two vectors may be very large, in general. Indeed the shorter their lengths are, the better our simulation is.

More precisely, if for some number \( C \), \(||\psi_A||| \leq C \) and \(||\psi_B||| \leq C \), then the following simulation uses \( O(C^4) \) bits. Alice and Bob send Charlie \(||\psi_A|||\) and \(||\psi_B|||\), respectively, up to \( O(1/C) \) precision. This requires \( O(\log C) \) bits. They then proceed to estimate \( \cos \theta \), for the angle \( \theta \) between \(|\psi_A\rangle\) and \(|\psi_B\rangle\) up to a precision of \( O(1/C^2) \). The protocol in Kremer, Nisan and Ron [63], which is based on the following observation of Goemans and Williamson [52], gives a protocol that accomplishes the latter task using \( O(C^4) \) bits.

Assume for simplicity that all vectors are real (the complex number case can be easily reduced to the real case). If \(|\psi\rangle\) is a random unit vector in the same space of \(|\phi_A\rangle\) and \(|\phi_B\rangle\), then

\[
\text{Prob} \left[ \text{sign}(\langle \psi | \psi_A \rangle) \neq \text{sign}(\langle \psi | \psi_B \rangle) \right] = \frac{\theta}{\pi}.
\]
Hence, in order to estimate $\cos \theta$ with error term $\delta'$, it suffices to estimate $\theta/\pi$ to some error term $O(\delta')$ using the above equality checking of signs. Obviously this can be done by a SMP protocol, and by a simple application of Chernoff Bound, requires $O \left( \ln \frac{1}{\epsilon} / \delta'^2 \right)$ repetitions, where $\epsilon$ is the failure probability. With $\delta' = O(\delta/C^2)$, this is $O \left( C^4 \ln \frac{1}{\epsilon} / \delta^2 \right)$ bits.

We note that [93] gives a procedure along the lines of checking equality of signs but it produces a random $\pm 1$ variable whose expectation is precisely $\cos \theta$, though this is not asymptotically advantageous.

We summarize the above discussion as the basis for our future discussions.

**Theorem 4.7 ([63, 52]).** Suppose the acceptance probability of a quantum protocol can be expressed as $\langle \psi_A | \psi_B \rangle$, where $|\psi_A\rangle$ and $|\psi_B\rangle$ can be prepared by each party individually. Furthermore, for some nonnegative number $C$, $\| |\psi_A\rangle\| \leq C$, and $\| |\psi_B\rangle\| \leq C$. Then there is a classical SMP protocol with shared coins that uses $O \left( C^4 \ln \frac{1}{\epsilon} / \delta^2 \right)$ bits and whose acceptance probability deviates from that of the quantum protocol by at most $\delta$ with probability at least $1 - \epsilon$.

### 4.3 The main theorem

In this section, we formally define the classical communication complexity and the diamond norm of bipartite quantum operators, and derive an upper bound on the former in terms of the latter. We shall focus on the following case: that the measurement gives two outcomes, and that the dimensions of the two systems are the same. Our results can be extended trivially to more general cases.

We use script letters $\mathcal{N}$, $\mathcal{M}$, $\mathcal{F}$, $\cdots$, to denote Hilbert spaces, and $\text{L}(\mathcal{N})$ to denote the space of operators on $\mathcal{N}$. The identity operator on $\mathcal{N}$ is denoted by $I_\mathcal{N}$, and the identity superoperator on $\text{L}(\mathcal{N})$ is denoted by $I_{\mathcal{N}}$. Recall that a positive-operator-
valued measurement (POVM) on a Hilbert space $\mathcal{H}$ is a set of positive semidefinite operators $\{Q_1, Q_2, \cdots, Q_m\}$ on $\mathcal{H}$, such that $\sum_{i=1}^{m} Q_i = I_{\mathcal{H}}$. Each $Q_i$ is called a measurement element, and corresponds to the measurement outcome $i$. We may refer to a semidefinite operator $Q$, $0 \leq Q \leq 1$, as a measurement element of the implicit binary POVM $\{Q, I - Q\}$. For more details on the foundations of quantum information processing, refer to the textbook [70].

### 4.3.1 Classical simulation of quantum measurements

In this subsection we define the central concept of this chapter: the classical communication complexity of quantum measurements.

Let $Q$ be measurement element acting on a bipartite system $AB$. Let $|E\rangle_{A'B'}$ be a bipartite state, where $A'$ ($B'$) includes $A$ ($B$) as a subsystem. Let $R_A$ and $R_B$ be physically realizable operators acting on system $A'$ and $B'$, respectively. Denote by $\mu(Q, |E\rangle, R_A, R_B)$ the probability

$$\mu(Q, |E\rangle, R_A, R_B) \overset{\text{def}}{=} \text{tr}(QR_A \otimes R_B (|E\rangle\langle E|)).$$

**Definition 4.8.** Let $\delta, \epsilon \in [0, 1/2)$, and $Q$ be a measurement elements. The classical communication complexity of $Q$ with precision $\delta$ and success probability $1 - \epsilon$, denoted by $\text{Com}_{\delta,\epsilon}(Q)$, is the minimum number $k$ such that for any $|E\rangle$, $R_A$ and $R_B$ described above, there is a classical communication protocol between two parties Alice and Bob that satisfies the following conditions:

1. The input of Alice (Bob) is a classical description of $|E\rangle$, and a classical description of $R_A$ ($R_B$);

2. The protocol exchanges $\leq k$ bits and is allowed to use an unlimited amount of shared randomness.
(3). The output $p$ satisfies

$$|p - \mu(Q, |E\rangle, R_A, R_B)| \leq \delta$$

with probability at least $1 - \epsilon$. The probability is over the shared randomness.

### 4.3.2 The diamond norm on bipartite operators

Let $\mathcal{N}$ be a Hilbert space and $T : \mathbf{L}(\mathcal{N}) \rightarrow \mathbf{L}(\mathcal{N})$ be a superoperator. The diamond norm on superoperators is defined in the Appendix (Equation A.3). For our application, the following alternative characterization of the diamond norm is more convenient.

**Lemma 4.9 (e.g., [59]).** For any superoperator $T$,

$$\|T\|_\diamond = \min \left\{ \sqrt{\| \sum_t A_t^\dagger A_t \|} \cdot \sqrt{\| \sum_t B_t^\dagger B_t \|} : A_t, B_t \in \mathbf{L}(\mathcal{N}), \ T = \sum_t A_t \cdot B_t^\dagger \right\}.$$ 

Let $\mathcal{N}_A$, $\mathcal{N}_B$, and $\mathcal{N}$ be Hilbert spaces of the same dimension. We fix an isomorphism between any two of them. For an operator in one space, we use the same notation for its images and preimages, under the isomorphisms, in the other spaces.

Let $Q \in \mathbf{L}(\mathcal{N}_A \otimes \mathcal{N}_B)$ be a bipartite operator and $Q = \sum_t A_t \otimes B_t^\dagger$, for some $A_t \in \mathbf{L}(\mathcal{N}_A)$, and $B_t \in \mathbf{L}(\mathcal{N}_B)$. Define a mapping $T$ from bipartite operators on $\mathcal{N}_A \otimes \mathcal{N}_B$ to superoperators $\mathbf{L}(\mathcal{N}) \rightarrow \mathbf{L}(\mathcal{N})$ by mapping $Q \mapsto T(Q) \overset{\text{def}}{=} \sum_t A_t \cdot B_t^\dagger$. It can be easily verified that the mapping is independent of the choice of the decomposition of $Q$ and is indeed an isomorphism.

**Definition 4.10.** Let $Q \in \mathbf{L}(\mathcal{N}_A \otimes \mathcal{N}_B)$ be a bipartite operator. The diamond norm of $Q$, denoted by $\|Q\|_\diamond$, is $\|Q\|_\diamond \overset{\text{def}}{=} \|T(Q)\|_\diamond$.

By Lemma 4.9, for any $Q$,

$$\|Q\|_\diamond = \min \left\{ \sqrt{\| \sum_t A_t^\dagger A_t \|} \cdot \sqrt{\| \sum_t B_t^\dagger B_t \|} : A_t \in \mathbf{L}(\mathcal{N}_A), B_t \in \mathbf{L}(\mathcal{N}_B), \ Q = \sum_t A_t \otimes B_t^\dagger \right\}.$$
Note that if a superoperator $T = A \cdot B$ for some $A, B \in L(\mathcal{N})$, $\|T\|_\diamond = \|A\| \cdot \|B\|$. Therefore the diamond norm on bipartite operators is a tensor norm:

**Lemma 4.11.** If $K = A \otimes B$, $\|K\|_\diamond = \|A\| \cdot \|B\|$.

A nice property of the superoperator diamond norm is that it is “stable”, i.e., it remains unchanged when tensor with the identity operator on an additional space (Proposition A.1). This stability property carries over to our diamond norm and is important for our applications. Let $\mathcal{F}_A$ and $\mathcal{F}_B$ be Hilbert spaces of the same dimension, and $Q \in L(\mathcal{N}_A \otimes \mathcal{N}_B)$. Denote by $Q_{\mathcal{F}_A,\mathcal{F}_B}$ the bipartite operator $Q \otimes I_{\mathcal{F}_A \otimes \mathcal{F}_B}$, where the two subsystems are $\mathcal{N}_A \otimes \mathcal{F}_A$ and $\mathcal{N}_B \otimes \mathcal{F}_B$.

**Lemma 4.12.** For any $Q$, $\|Q_{\mathcal{F}_A,\mathcal{F}_B}\|_\diamond = \|Q\|_\diamond$.

If $Q$ is a measurement element of a POVM acting on a Hilbert space of dimension $K$, then we have the following upper bound on $\|Q\|_\diamond$.

**Proposition 4.13.** If a bipartite operator $Q$ is measurement element of a POVM acting on a Hilbert space of dimension $K$, then $\|Q\|_\diamond \leq K$.

**Proof.** For any bipartite pure state $|u\rangle$, let $|u\rangle = \sum_i \sqrt{p_i} |i\rangle_A |i\rangle_B$ for some $p_i \geq 0$, $\sum p_i = 1$ and orthonormal basis $\{|i\rangle\}$ by Schmidt decomposition. Then

$$|u\rangle\langle u| = \sum_{i,j} \sqrt{p_i p_j} |i\rangle_A \langle j| \otimes |i\rangle_B \langle j|_B$$

Let $A_{i,j} = \sqrt{p_i} |i\rangle_A \langle j|_A$, $B_{i,j} = \sqrt{p_j} |j\rangle_B \langle i|_B$. We have

$$\| \sum_{i,j} \tilde{A}_{i,j}^\dagger A_{i,j} \| = \| \sum_{i,j} p_i |i\rangle_A \langle i|_A \langle j|_A \langle j|_B \| = \|I\| = 1,$n

similarly, $\| \sum_{i,j} \tilde{B}_{i,j}^\dagger B_{i,j} \| = 1$. Since $T(|u\rangle\langle u|) = \sum_{i,j} A_{i,j} \cdot B_{i,j}^\dagger$, according to Lemma 4.9, diamond norm of $T(|u\rangle\langle u|)$ is upper bounded by 1. Thus $\| |u\rangle\langle u| \|_\diamond \leq 1$. 

Let positive operator \( Q = \sum c_i |u_i \rangle \langle u_i | \) for some \( 0 \leq c_i \leq 1 \) and orthonormal basis \( \{u_i\} \). By the triangle inequality, the diamond norm of \( Q \) is upper bounded by the dimension \( K \).

This bound is not far from being optimal for \( \text{IP}_n \), in which case \( K = 2^{2n} \).

**Proposition 4.14.** For the \( \text{IP}_n \) operator defined in Equation 1.2, \( \| \text{IP}_n \|_\diamond \geq 2^{n/2} - \frac{1}{2} \).

**Proof.** By definition,

\[
\mathcal{T}(\text{IP}_n) = \sum_{x,y \in \{0,1\}^n, x \cdot y = 1} |x \rangle \langle x| \cdot |y \rangle \langle y|.
\]

To prove a lower bound on \( \| \text{IP}_n \|_\diamond \), we use a dual characterization of the diamond norm as in Equation A.4. We set \( \rho = \sum_{x,y} |x \rangle \langle y| \otimes I_G \), resulting in

\[
\| \mathcal{T}(\text{IP}_n) \|_\diamond \geq \frac{1}{2^n} \left\| \sum_{x,y \in \{0,1\}^n, x \cdot y = 1} |x \rangle \langle y| \right\|_{\text{tr}}.
\]

Let \( A = \sum_{x,y \in \{0,1\}^n, x \cdot y = 1} |x \rangle \langle y| \), \( J \) be all-one matrix and \( H \) be the Hadamard matrix (i.e., \( \sum_{x,y \in \{0,1\}^n} (-1)^{x \cdot y} |x \rangle \langle y| \)). Then \( A = \frac{1}{2} (J - I) \). The largest eigenvalue of \( J \) is \( 2^n \) and all other eigenvalues are 0, thus \( \|J\|_{\text{tr}} = 2^n \). All the eigenvalues of \( H \) are \( \pm 2^{n/2} \), thus \( \|H\|_{\text{tr}} = 2^{3n/2} \). Therefore,

\[
\|A\|_{\text{tr}} \geq \frac{1}{2} (\|H\|_{\text{tr}} - \|J\|_{\text{tr}}) = \frac{1}{2} (2^{3n/2} - 2^n).
\]

Thus \( \| \text{IP}_n \|_\diamond \geq 2^{n/2} - \frac{1}{2} \). \( \square \)

We conclude this subsection by noting that our diamond norm on bipartite operators appears natural in connection with the following matrix analogy of the Cauchy Schwartz Inequality.
Theorem 4.15 (Jocić [56]). For any operators $A_t$ and $B_t$,

\begin{equation}
\|\sum_t A_t \otimes B_t^\dagger\| \leq \sqrt{\|\sum_t A_t^\dagger A_t\| \cdot \|\sum_t B_t^\dagger B_t\|}.
\end{equation}

The above inequality (4.2) may actually be proved by the same approach that we use to prove Theorem 4.16 below.

4.3.3 Upper bounding $\text{Com}(Q)$ by the diamond norm

We now use the diamond norm to derive an upper bound on $\text{Com}_{\delta,\epsilon}(Q)$. Recall that if $\mathcal{M}$ and $\mathcal{N}$ are two Hilbert spaces, an isometric embedding $U : \mathcal{M} \to \mathcal{N}$ is a linear map that satisfies $U^\dagger U = I_\mathcal{M}$.

Theorem 4.16. For any bipartite positive semidefinite operator $Q$ acting on a Hilbert space of dimension $K$,

\begin{equation}
\text{Com}_{\delta,\epsilon}(Q) = O\left(\|Q\|_\diamond^2 \cdot \ln \frac{1}{\epsilon}/\delta^2\right).
\end{equation}

In particular $\text{Com}_{\delta,\epsilon}(Q) = O(K^2 \log \ln \frac{1}{\epsilon}/\delta^2)$. Furthermore, the upper bound (4.3) can be achieved by a SMP protocol with shared randomness.

Proof. Without loss of generality, assume that on receiving their portions of $|E\rangle$, Alice and Bob apply an isometric embedding $U : \mathcal{M}_A \to \mathcal{N}_A \otimes \mathcal{F}_A$, and $V : \mathcal{M}_B \to \mathcal{N}_B \otimes \mathcal{F}_B$, respectively, for some Hilbert spaces $\mathcal{F}_A$ and $\mathcal{F}_B$ with an equal dimension. The distribution resulted from Charlie’s measuring $Q$ on $\text{Tr}_{\mathcal{F}_A,\mathcal{F}_B}((U \otimes V)|E\rangle\langle E|(U \otimes V)^\dagger)$ is the same as that of Charlie applying $Q_{\mathcal{F}_A,\mathcal{F}_B}$ on the larger state $(U \otimes V)|E\rangle\langle E|(U \otimes V)^\dagger$. By Lemma 4.12, $\|Q_{\mathcal{F}_A,\mathcal{F}_B}\|_\diamond = \|Q\|_\diamond$. Therefore, to prove the theorem we need only to consider isometric embeddings $U : \mathcal{M}_A \to \mathcal{N}_A$ and $V : \mathcal{M}_A \to \mathcal{N}_B$.

Without loss of generality, we assume that Alice and Bob have agreed on a Schmidt decomposition $|E\rangle = \sum_i \sqrt{p_i} |i\rangle_A \otimes |i\rangle_B$, for some $p_i \geq 0$, $\sum_i p_i = 1$, and for an
orthonormal basis \( \{|i\} \). Denote by \(|i_A\rangle \overset{\text{def}}{=} U|i\rangle\), and \(|i_B\rangle \overset{\text{def}}{=} V|i\rangle\). Then the message that Charlie receives is \(|\bar{E}\rangle \overset{\text{def}}{=} (U \otimes V)|E\rangle = \sum_i \sqrt{p_i}|i_A\rangle \otimes |i_B\rangle\).

Suppose \(\|Q\|_o\) is achieved under the decomposition \(Q = \sum_t A_t \otimes B_t^\dagger\), with which if \(Q_A \overset{\text{def}}{=} \sum_t A_t^\dagger A_t\), and, \(Q_B \overset{\text{def}}{=} \sum_t B_t^\dagger B_t\), we have \(\|Q_A\| = \|Q_B\| = \|Q\|_{o,\alpha}^E\). With those definitions, we have

\[
p = \langle \bar{E}|Q|\bar{E}\rangle = \sum_{i,j,t} \sqrt{p_i p_j} \langle i_A|A_t^\dagger j_A\rangle \cdot \langle i_B|B_t^\dagger j_B\rangle.
\]

Define two vectors

\[
(4.4) \quad |\psi_A\rangle = \sum_{i,j,t} \sqrt{p_j} \langle j_A|A_t^\dagger|i_A\rangle \; |i,j,t\rangle, \quad \text{and,}
\]

\[
(4.5) \quad |\psi_B\rangle = \sum_{i,j,t} \sqrt{p_i} \langle i_B|B_t^\dagger|j_B\rangle \; |i,j,t\rangle.
\]

Then \(p = \langle \psi_A|\psi_B\rangle\). Further, with \(\rho_A \overset{\text{def}}{=} \sum_j p_j|j_A\rangle\langle j_A|\),

\[
\langle \psi_A|\psi_A\rangle = \sum_{i,j,t} p_j |\langle j_A|A_t^\dagger|i_A\rangle|^2 = \text{tr}(\rho_A Q_A) \leq \|Q_A\| = \|Q\|_{o,\alpha}^E.
\]

Similarly, \(\langle \psi_B|\psi_B\rangle \leq \|Q_B\| = \|Q\|_o\). Therefore, by Theorem 4.7, the measurement scenario can be approximated by a classical SMP with shared coins to be within an \(\epsilon\) precision using \(O(\|Q\|^2_o \ln \frac{1}{\epsilon}/\epsilon^2)\) bits. This bound is \(O(K^2 \log \ln \frac{1}{\epsilon}/\epsilon^2)\) as \(\|Q\|_o = O(K)\) by Proposition 4.13.

\(\Box\)

**Remark 4.17.** One may improve the above upper bound on \(\text{Com}_{\delta,\epsilon}(Q)\) by a more carefully chosen \(|\psi_A\rangle\) and \(|\psi_B\rangle\) in Equation 4.4 and 4.5. More specifically, let \(\alpha \in [0,1]\), define

\[
|\psi_A^\alpha\rangle = \sum_{i,j,t} \sqrt{p_i^\alpha p_j^{1-\alpha}} \langle j_A|A_t^\dagger|i_A\rangle \; |i,j,t\rangle, \quad \text{and,}
\]

\[
|\psi_B^\alpha\rangle = \sum_{i,j,t} \sqrt{p_i^{1-\alpha} p_j^{\alpha}} \langle i_B|B_t^\dagger|j_B\rangle \; |i,j,t\rangle.
\]
One can verify that minimizing $\|\psi_A\| \cdot \|\psi_B\|$ over all decompositions of $Q$ gives rise to a tensor norm, which we do not know if is stable under tensoring with identity superoperators. Although we have not found any useful application of an $\alpha \neq 0$, we cannot rule out the possibility that a carefully chosen $\alpha$ may give a better bound.

**Remark 4.18.** In the case that $|E\rangle$ is not entangled, the same approach in Theorem 4.16 can be used to derive a systematic classical simulation. More specifically, in this context we would like to estimate $p = \langle \phi_A \otimes \phi_B | Q | \phi_A \otimes \phi_B \rangle$, for a state $|\phi_A\rangle$ known to Alice only and a state $|\phi_B\rangle$ known to Bob only. For a decomposition of $Q = \sum_t A_t \otimes B_t^\dagger$, we define

$$|\psi_A\rangle = \sum_t \langle \phi_A | A_t^\dagger | \phi_A \rangle |t\rangle,$$

and,

$$|\psi_B\rangle = \sum_t \langle \phi_B | B_t^\dagger | \phi_B \rangle |t\rangle.$$

Then $p = \langle \psi_A | \psi_B \rangle$. It can be verified that

$$\|Q\|_\otimes \overset{\text{def}}{=} \inf \{ \|\psi_A\| \cdot \|\psi_B\| : Q = \sum_t A_t \otimes B_t^\dagger \}$$

defines a tensor norm and $\|Q\|_\otimes \leq \|Q\|_\diamond$. This approach gives a constant cost simulation of the elegant quantum fingerprint protocol of Buhrman, Cleve, Watrous, and de Wolf [24] for testing equality of two input strings.

### 4.4 Applications

We now apply the above to derive classical upper bounds on quantum communication complexity.

#### 4.4.1 Quantum SMP with shared entanglement

If the quantum protocol is in the SMP model with shared entanglement, we immediately have,
Corollary 4.19 (of Theorem 4.16). If in a quantum SMP protocol, Charlie applies the measurement $P$, then the protocol can be simulated by a classical SMP protocol with shared coins and using $O(\|P\|_2^2)$ bits.

4.4.2 Twoway interactive quantum communication with shared entanglement

Now consider the general twoway interactive quantum communication. We need the following lemma due to Yao [96], and the following formulation is from [77]:

Lemma 4.20 ([96, 77]). Let $\mathcal{P}$ be a two-party interactive quantum communication protocol that uses $q$ qubits. Let $\mathcal{H}_A$ and $\mathcal{H}_B$ be the state spaces of Alice and Bob, respectively. For an input $(x,y)$, denote by $|\Phi_{x,y}\rangle_{AB}$ the joint state of Alice, Bob before the protocol starts. Then there exist linear operators $A_h \in L(\mathcal{H}_A)$, and $B_h \in L(\mathcal{H}_B)$, for each $h \in \{0,1\}^{q-1}$, such that

(a) $\|A_h\| \leq 1$ and $\|B_h\| \leq 1$ for all $h \in \{0,1\}^{q-1}$;
(b) the acceptance probability of $\mathcal{P}$ on input $x$ and $y$ is $\|P|\Phi_{x,y}\rangle\|^2$, where $P \overset{\text{def}}{=} \sum_{h \in \{0,1\}^{q-1}} A_h \otimes B_h$.

We are now ready to prove Theorem 4.2.

Proof of Theorem 4.2. Let $|E\rangle_{AB}$ be the shared entanglement, For an $n$-bit binary string $x$, denote by $U_x$ the isometric embedding from $\mathbb{C}$ to $\mathbb{C}^{2^n}$ that maps $c \mapsto c|x\rangle$. Let $P$, $A_h$, and $B_h$ be those in Lemma 4.20. Then the quantum protocol gives rise to a measurement scenario in which the measurement is $P^\dagger P$, the shared entanglement is $|E\rangle$, and on an input pair $(x,y)$, Alice’s private operator is $U_x$ and that of Bob is $U_y$.

By Theorem 4.16, the acceptance probability can be estimated with $O(\|P^\dagger P\|_2^2)$ bits of communication in the SMP model with shared randomness. Since $\| \cdot \|_\circ$ is a
tensor norm, we have
\[ \| P^\dagger P \|_\diamond \leq \sum_{h,h'} \| (A_{h'}^\dagger A_h) \otimes (B_{h'}^\dagger B_h) \|_\diamond = \sum_{h,h'} \| A_h \| \| A_{h'} \| \| B_h \| \| B_{h'} \| \leq 2^{2(q-1)}. \]

The last inequality is because \( \| A_h \| \leq 1 \) and \( \| B_h \| \leq 1 \) for all \( h \). Hence the acceptance probability can be estimated by a classical SMP protocol using \( \exp(O(q)) \) bits.

Corollary 4.3 follows trivially from the above by setting \( q \) to be a constant. Corollary 4.5 follows immediately from Theorem 4.4 and Corollary 4.3.

### 4.4.3 Simulating quantum correlations

We shall define precisely what we mean by simulating quantum correlations.

We define a quantum measurement game as a triple \( G = (|E\rangle_{AB}, P_A, P_B) \), where \( |E\rangle_{AB} \) is a bipartite quantum state, \( P_A, P_B \) are sets of possible measurements on the system \( A \) and the system \( B \), respectively. Let \( \mathcal{V}_A (\mathcal{V}_B, \text{respectively}) \) be the set of possible measurement outcomes of \( P_A (P_B, \text{respectively}) \). For \( P_A \in P_A \) and \( P_B \in P_B \), denote by \( \omega_G(P_A, P_B) \) the distribution of the measurement outcomes when \( P_A \otimes P_B \) is applied to \( |E\rangle \).

A classical simulation of a quantum measurement game \( G = (|E\rangle_{AB}, P_A, P_B) \) is a classical communication protocol between two parties Alice and Bob, who start with an unlimited mount of shared randomness, and Alice has the classical description of an element \( P_A \in P_A \), while Bob has the classical description of an element \( P_B \in P_B \). At the end of the protocol, Alice (and Bob) outputs an element from \( \mathcal{V}_A (\mathcal{V}_B, \text{respectively}) \), resulting in a distribution \( \tilde{\omega}(P_A, P_B) \).

We are now able to rigorously state Theorem 4.6. Recall that the statistical distance between two distributions \( \pi = (p_1, \cdots, p_n) \) and \( \bar{\pi} = (\bar{p}_1, \cdots, \bar{p}_n) \) is \( \| \pi - \bar{\pi} \|_1 \overset{\text{def}}{=} \sum_i |p_i - \bar{p}_i| \).
Theorem 4.21. Let $G = (|E\rangle_{AB}, \mathcal{P}_A, \mathcal{P}_B)$ be a quantum measurement game, $m = |\mathcal{V}_A| \cdot |\mathcal{V}_B|$, and $\epsilon, \delta \in \mathbb{R}$, $0 \leq \epsilon, \delta < 1$. There is a classical simulation of $G$ that exchanges $O\left(\frac{m^3}{\epsilon^2} \cdot \ln \frac{m}{\epsilon}\right)$ number of bits and the output distribution $\tilde{\omega}(P_A, P_B)$ for any $P_A \in \mathcal{P}_A$ and $P_B \in \mathcal{P}_B$ satisfies

$$\|\tilde{\omega}(P_A, P_B) - \omega_G(P_A, P_B)\|_1 \leq \delta$$

with probability at least $1 - \epsilon$. In particular, the simulation cost is $O(\ln \frac{1}{\epsilon}/\delta^2)$ if $m = O(1)$.

Proof. Recall that a POVM measurement can be expressed as a physically realizable operator followed by a projective measurement (see, e.g., [59]). Thus we can assume without loss of generality that there exist projections $P^v_A$, $v \in \mathcal{V}_A$, and $P^{v'}_B$, $v' \in \mathcal{V}_B$, such that for each $P_A \in \mathcal{P}_A$ ($P_B \in \mathcal{P}_B$), there is an isometric embedding $U_A$ ($U_B$) so that $P_A$ ($P_B$) consists of the measurement elements $\{U_A^\dagger P^v_A U_A : v \in \mathcal{V}_A\}$ ($\{U_B^\dagger P^{v'}_B U_B : v' \in \mathcal{V}_B\}$).

Fix a pair of measurements $(P_A, P_B)$. In the classical simulation protocol, Alice and Bob first compute the probability of outputting $(v, v')$ to be within $\delta/m$ deviation with probability at least $1 - \epsilon/m$, for each $v \in \mathcal{V}_A$ and $v' \in \mathcal{V}_B$. They then output $(v, v')$ according to the probabilities computed. Thus $\tilde{\omega}(P_A, P_B)$ is within $\delta$ statistical distance to $\omega(P_A, P_B)$ with probability at least $1 - \epsilon$.

Fix a pair of possible outcome $(v, v')$. Let $P^{v,v'} \overset{\text{def}}{=} P^v_A \otimes P^{v'}_B$. Then by Lemma 4.11, $\|P^{v,v'}\|_\diamond = \|P^v_A\| \cdot \|P^{v'}_B\| \leq 1$. The estimation of $\omega_G(P_A, P_B)$ now becomes the simulation of the measurement element $P^{v,v'}$ with the initial state being $|E\rangle$, and the local physically realizable operators being $U_A^\dagger \cdot U_A$ and $U_B^\dagger \cdot U_B$.

Hence by Theorem 4.16, the probability of observing outcome $(v, v')$ can be calculated to be within precision $O(\delta/m)$ and with probability at least $1 - \epsilon/m$ by a
classical protocol using $O(m^2 \ln(m/\epsilon)/\delta^2)$ bits. Thus the overall simulation cost is $O(m^3 \ln(m/\epsilon)/\delta^2)$ bits, which is $O(\ln \frac{1}{\epsilon}/\delta^2)$ when $m = O(1)$. \qed
CHAPTER V

The maximum tensor norm of bipartite superoperators

This chapter is based on [28]. We study the maximum tensor norm of bipartite physically realizable superoperators, with respect to the diamond norm, as a measure of their nonlocality. We show that a bipartite physically realizable superoperator is bi-local if and only if its maximum tensor norm is exactly 1. With the help of the dual characterization, we are able to calculate the exact maximum tensor norm of several elementary superoperators. As an application of the maximum tensor norm, we show that estimations of the norm can be used to prove lower bounds on the amount of quantum communication required to realize the superoperator, and this connection to quantum communication complexity could be used to prove quantum lower bounds.

5.1 Summary of results

In this chapter, we focus on the maximum tensor norm of superoperators endowed with the diamond norm. We refer to the Appendix for definitions of the diamond norm and the corresponding maximum tensor norm.

A bipartite physically realizable superoperator is bi-local if \( T = \sum_i p_i T_i^A \otimes T_i^B \), where \([p_i]_i\) is a probability distribution, and \(T_i^A, T_i^B\) are local physically realizable superoperators. Our first result is analogous to that of Rudolph [78], but requires a
Theorem 5.1. For any physically realizable bipartite superoperator $T$, the maximum tensor norm $\|T\|_\gamma \geq 1$. Furthermore, $\|T\|_\gamma = 1$ if and only if $T$ is bi-local (we also have a similar result for unitary operators).

It is usually difficult to determine maximum tensor norm due to the infimum. Fortunately, being a maximum tensor norm, $\| \cdot \|_\gamma$ has a dual characterization that allows us to prove a lower bound by finding an appropriate bilinear operator.

Proposition 5.2. Let $T_A$ and $T_B$ be any superoperators on system $A$ and $B$ respectively. Let $h$ be a bilinear operator such that $h(T_A, T_B) \leq 1$ for any $\|T_A\|_\diamond, \|T_B\|_\diamond \leq 1$. For any bipartite superoperator $T$ on system $AB$, let $T = \sum_i T^A_i \otimes T^B_i$ and $h(T) = \sum_i h(T^A_i, T^B_i)$. Then the maximum tensor norm $\|T\|_\gamma$ satisfies:

\begin{equation}
\|T\|_\gamma = \sup_{\|T_A, T_B\| \leq 1} |h(T)|.
\end{equation}

A proof of the above proposition (for any maximum tensor norm) can be found in [82]. Using this dual characterization, we are able to give a simple proof for the nonlocality of a superoperator found by Bennett et al. [14], which is a projective measurement to a set of tensor product states. We are also able to compute the exact maximum tensor norm of several elementary superoperators. Denote by CNOT, SWAP, CC, and QC the superoperators for the Controlled-NOT gate, the SWAP gate, measuring one qubit and sending the measurement result (i.e. $\langle 0|_A \cdot |0\rangle_A \otimes |0\rangle_B \langle 0|_B + \langle 1|_A \cdot |1\rangle_A \otimes |1\rangle_B \langle 1|_B$), sending one quantum bit (i.e. $\sum_{i,j \in \{0, 1\}} \langle i|_A \cdot |j\rangle_A \otimes |i\rangle_B \langle j|_B$), respectively.

Theorem 5.3. $\|\text{CNOT}\|_\gamma = \|\text{CC}\|_\gamma = 2$, $\|\text{SWAP}\|_\gamma = \|\text{QC}\|_\gamma = 4$.

We also have a connection between communication complexity and the maximum tensor norm.
Theorem 5.4. If there exists a communication protocol that realizes a bipartite superoperator $T$ with $c$ classical bits and $q$ qubits, then the maximum tensor norm $\|T\|_\gamma \leq 2^{c+2q}$.

Another result is Razborov’s lower bound on the quantum communication complexity of the Set Disjointness Problem [77] can be extended as follows.

Theorem 5.5. Any superoperator for computing the Set Disjointness Problem must have $\exp(\Omega(\sqrt{n}))$ gamma norm.

The rest of this chapter is organized as the following. We start by proving the criteria for bi-local superoperators. Next, we use the dual characterization to calculate the maximum tensor norm of several elementary superoperators. Then we show the connections with communication complexity.

5.2 Nonlocality criteria for superoperators

We prove Theorem 5.1 in this section.

Proof of Theorem 5.1. Let $T^{AB} : L(N_A \otimes N_B) \to L(M_A \otimes M_B)$ be a bipartite physically realizable superoperator. Since $T^{AB}$ is physical realizable, $\|T^{AB}\|_\circ = 1$.

Let $T^{AB} = \sum_i T^A_i \otimes T^B_i$ be any decomposition, where $T^A_i : L(N_A) \to L(M_A)$ and $T^B_i : L(N_B) \to L(M_B)$. By the triangle inequality, we have $\sum_i \|T^A_i\|_\circ \|T^B_i\|_\circ \geq \|\sum_i T^A_i \otimes T^B_i\|_\circ = \|T^{AB}\|_\circ = 1$. So $\|T^{AB}\|_\gamma \geq 1$.

Now we prove that $\|T^{AB}\|_\gamma = 1$ if and only if $T^{AB}$ is bi-local. When $T^{AB}$ is bi-local, let $T^{AB} = \sum_i p_i T^A_i \otimes T^B_i$ be a decomposition with $\sum_i p_i = 1$ and $T^A_i, T^B_i$ be physically realizable superoperators. Then $\|T^A_i\|_\circ = \|T^B_i\|_\circ = 1$ by Proposition A.2. By the triangle inequality, $\|T^{AB}\|_\gamma \leq \sum_i p_i \|T^A_i\|_\circ \|T^B_i\|_\circ = \sum_i p_i = 1$. Thus $\|T^{AB}\|_\gamma = 1$. 
When $\|T^{AB}\|_\gamma = 1$, let $T^{AB} = \sum_i p_i T^A_i \otimes T^B_i$ be a decomposition that achieves the minimum $\sum_i p_i T^A_i \|_\diamond T^B_i \|_\diamond$, and local superoperators $T^A_i, T^B_i$ be normalized such that $\|T^A_i\|_\diamond = \|T^B_i\|_\diamond = 1$. Then $\sum_i p_i = \sum_i p_i \|T^A_i\|_\diamond \|T^B_i\|_\diamond = \|T^{AB}\|_\gamma = 1$. We now show that $T^A_i$ and $T^B_i$ are physically realizable.

Let $\rho^A \in \mathcal{N}_A, \rho^B \in \mathcal{N}_B$ be two density operators on system $A, B$ respectively. Then $\rho^A \otimes \rho^B$ is a density operator on the system $\mathcal{N}_A \otimes \mathcal{N}_B$. Since superoperator $T^{AB}$ is trace preserving, $\text{tr}(T^{AB}(\rho^A \otimes \rho^B)) = 1 = \|T^{AB}(\rho^A \otimes \rho^B)\|_\text{tr}$. We have the following inequalities,

$$\|T^{AB}(\rho^A \otimes \rho^B)\|_\text{tr} \leq \sum_i p_i \|T^A_i(\rho^A)\|_\text{tr} \|T^B_i(\rho^B)\|_\text{tr} \leq \sum_i p_i \|T^A_i\|_1 \|T^B_i\|_1 \leq \sum_i p_i \|T^A_i\|_\diamond \|T^B_i\|_\diamond \leq \sum_i p_i = 1$$

where the first inequality follows from triangle inequalities of trace norms, the second from definition of $\| \cdot \|_1$, and the third from the fact that $\| \cdot \|_1$ is no more than $\| \cdot \|_\diamond$. Thus all the inequalities become equalities and $\|T^A_i(\rho^A)\|_\text{tr} = \|T^B_i(\rho_B)\|_\text{tr} = 1$.

Notice that $\rho^A$ is arbitrary, thus $\|T^A_i(\rho^A)\|_\text{tr} = 1$ for every density operator on $\mathcal{N}_A$. Since $\|T^A_i\|_\diamond = 1$, let $T^A_i = \text{tr}_\mathcal{F}(V \cdot W^\dagger)$ be a decomposition that achieves minimum $\|V\| \cdot \|W\|$ and $\|V\| = \|W\|$. Then $\|V\| \cdot \|W\| = \|T^A_i\|_\diamond = 1$, and $\|V\| = \|W\| = 1$. Let $\rho^A = |\eta\rangle\langle \eta|$ be a pure state. Then $1 = \|\text{tr}_\mathcal{F}(V|\eta\rangle\langle \eta|W^\dagger)\|_\text{tr} \leq \|V|\eta\rangle\langle \eta|W^\dagger\|_\text{tr} \leq \|V|\eta\rangle\| \cdot \|W|\eta\rangle\| \leq 1$. Hence all inequalities become equalities and $\|V|\eta\rangle\| = \|W|\eta\rangle\| = 1$. This holds for any pure state $|\eta\rangle$, so $V$ and $W$ are isometric embeddings, and it is not hard to see $V = W$. Thus $T^A_i = \text{tr}_\mathcal{F}(V \cdot V^\dagger)$ is a physically realizable superoperator. Similarly we can show that $T^B_i$ is a physically realizable superoperator. This completes the proof. \[\blacksquare\]
5.3 Maximum tensor norm of elementary superoperators

We calculate the exact maximum tensor norm for CNOT, SWAP, CC (superoperator for sending on classical bit) and QC (superoperator for sending one quantum bit) in this section. We also give lower bound of the maximum tensor norm of a superoperator defined in Bennett et al. [14].

To show the upper bound of maximum tensor norm for a bipartite superoperator $T$, we give explicit decompositions for $T$; to show the lower bound, we apply the dual characterization by constructing bilinear maps of the following form $h(T_A, T_B) = \text{tr}((T_A \otimes T_B)(\rho)M)$, and show that $h(T_A, T_B) \leq 1$, where $c$ is a numerical constant, $\rho$ is a bipartite density operator and $M$ is a projective measurement.

5.3.1 Maximum tensor norm of CNOT

First we prove $\|\text{CNOT}\|_\gamma \geq 2$. For any local superoperators $T_A \in \mathcal{L}(\mathcal{N}_A, \mathcal{M}_A \otimes \mathcal{F})$ and $T_B \in \mathcal{L}(\mathcal{N}_B, \mathcal{M}_B \otimes \mathcal{G})$, let

$$ h(T_A, T_B) \overset{\text{def}}{=} 2\text{tr}((T_A \otimes T_B)(\rho)M), $$

where $\rho$ is the density operator of the state $\frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A) \otimes |0\rangle_B$ and $M$ is the projection to state $\frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$. It is easy to verify $h(\text{CNOT}) = 2$.

Now we proceed to prove $|h(T_A, T_B)| \leq 1$ for $\|T_A\|_\diamond = \|T_B\|_\diamond = 1$. let $\rho_A = T_A(\frac{1}{2}(|0\rangle_A + |1\rangle_A)(|0\rangle_A + |1\rangle_A))$ and $\rho_B = T_B(|0\rangle_B\langle 0|_B)$ Then $\|\rho_A\|_\text{tr} \leq \|T_A\|_\diamond \|\frac{1}{2}(|0\rangle_A + |1\rangle_A)(|0\rangle_A + |1\rangle_A)\|_\text{tr} \leq 1$. Then $\text{tr}(\rho_A^2) \leq \|\rho_A\|^2_\text{tr} \leq 1$. Similarly, we have $\text{tr}(\rho_B^2) \leq 1$. 
Therefore, we have

\[
| h(T_A, T_B) | = | 2 \text{tr}((T_A(\frac{1}{2})(|0\rangle_A + |1\rangle_A)(|0\rangle_A + \langle 1|A\rangle) \otimes T_B(|0\rangle_B\langle 0|_B))M)| \\
= | 2 \times \frac{1}{2} \text{tr}((|0\rangle_B\langle 0| + |1\rangle_B\langle 1|)M)| \\
= | \sum_{i,j \in \{0,1\}} \langle i|A\rho_A|j\rangle_A \langle i|B\rho_B|j\rangle_B | \\
\leq \sqrt{\sum_{i,j}(\langle i|A\rho_A|j\rangle_A)^2} \cdot \sqrt{\sum_{i,j}(\langle i|B\rho_B|j\rangle_B)^2} \\
= \sqrt{\text{tr}(\rho_A^2)} \cdot \sqrt{\text{tr}(\rho_B^2)} \leq 1
\]

(5.2)

Equation 5.2 is due to the Cauchy-Schwartz Inequality. Since \( h(\text{CNOT}) = 2 \), this proves that \( \|\text{CNOT}\|_\gamma \geq 2 \).

Then we show a decomposition of CNOT that achieves minimum. Denote Pauli operation as follows:

\[
I = |0\rangle\langle 0| + |1\rangle\langle 1| \\
X = |0\rangle\langle 1| + |1\rangle\langle 0| \\
Y = -i|0\rangle\langle 1| + i|1\rangle\langle 0| \\
Z = |0\rangle\langle 0| - |1\rangle\langle 1|,
\]

and let operator

\[
A = I + iZ \\
B = I - iZ \\
C = I + iX \\
D = I - iX.
\]

Then operator norm \( \|A\| = \|B\| = \|C\| = \|D\| = \sqrt{2} \). We decompose CNOT as the
\[ CNOT = (|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 11| + |11\rangle\langle 10|) \]
\[ \cdot (|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 11| + |11\rangle\langle 10|) \]
\[ = \frac{1}{2}(I \otimes I + Z \otimes I + I \otimes X - Z \otimes X) \cdot \frac{1}{2}(I \otimes I + Z \otimes I + I \otimes X - Z \otimes X) \]
\[ = \left( \frac{1-i}{4}(A \otimes C) + \frac{1+i}{4}(B \otimes D) \right) \cdot \left( \frac{1-i}{4}(A \otimes C) + \frac{1+i}{4}(B \otimes D) \right) \]

Then for the maximum tensor norm of CNOT,
\[ \|CNOT\|_{\gamma} \leq \frac{1}{8}(\|A \otimes C\|\|A \otimes C\|_{\diamond} + \|A \otimes C\|\|B \otimes D\|_{\diamond}) \]
\[ + \|A \otimes C\|\|B \otimes D\|_{\diamond} \]
\[ \leq \frac{1}{8}(\|A^2\|C\|C\|_{\diamond} + \|A\|\|B\|\|C\|D\|_{\diamond} + \|B\|\|A\|\|D\|\|C\|_{\diamond} + \|B\|\|D\|_{\diamond}^2) \]
\[ = \frac{1}{8}(4 \times 4) = 2. \]

This completes the proof of \(\|CNOT\|_{\gamma} = 2\).

5.3.2 Maximum tensor norm of SWAP

First we prove \(\|SWAP\|_{\gamma} \geq 4\). For any local superoperators \(T_A \in L(N_A, M_A \otimes F)\) and \(T_B \in L(N_B, M_B \otimes G)\), let
\[ h(T_A, T_B) \overset{\text{def}}{=} 4\text{tr}((T_A \otimes T_B \otimes I)(\rho)M), \]
where \(\rho\) is a density operator and \(M\) is a projective measurement to be specified later. For local superoperators \(T_A, T_B\) with \(\|T_A\|_{\diamond} = \|T_B\|_{\diamond} = 1\), let \(T_A = \text{tr}_F(U_1 V_1^\dagger)\),
\[ T_B = \text{tr}_G(U_2 \cdot V_2^\dagger), \] where \( \|U_1\| = \|V_1\| = \|U_2\| = \|V_2\| = 1. \]

\[
|h(T_A, T_B)| = 4|\text{tr}_\mathcal{G}(((U_1 \otimes U_2) \cdot (V_1^\dagger \otimes V_2^\dagger)) \otimes I)\rho)|M| \\
= 4|\text{tr}_\mathcal{G}((U_1 \otimes U_2 \otimes I)\rho(V_1^\dagger \otimes V_2^\dagger \otimes I))M| \\
\leq 4\sqrt{\text{tr}((M \otimes I)(U_1 \otimes U_2 \otimes I)\rho(U_1 \otimes U_2 \otimes I)^\dagger)} \\
\cdot \sqrt{\text{tr}((M \otimes I)(V_1 \otimes V_2 \otimes I)\rho(V_1 \otimes V_2 \otimes I)^\dagger)},
\]

where Equation 5.3 is because \( \text{tr}(\text{tr}_\mathcal{X}'(\rho)M) = \text{tr}(\rho(M \otimes I_N^\dagger)) \) and Equation 5.4 is from the Cauchy-Schwartz Inequality of the form \( |\text{tr}(AB^\dagger)|^2 \leq \text{tr}(AA^\dagger)\text{tr}(BB^\dagger) \) and \( M^\dagger = M \).

Let \( \rho = |\psi\rangle \langle \psi|, M = |\phi\rangle \langle \phi|, f \overset{\text{def}}{=} \text{tr}((M \otimes I)(U_1 \otimes U_2 \otimes I)\rho(U_1 \otimes U_2 \otimes I)^\dagger). \) Then

\[
f = \text{tr}((|\phi\rangle \langle \phi| \otimes I)(U_1 \otimes U_2 \otimes I)(|\psi\rangle \langle \psi|)(U_1 \otimes U_2 \otimes I)^\dagger) \\
(5.5) = \langle \psi|((U_1 \otimes U_2 \otimes I)^\dagger(|\phi\rangle \otimes I)((\phi| \otimes I)(U_1 \otimes U_2 \otimes I)|\psi). \]

Let \( g \overset{\text{def}}{=} \|\langle \psi|((U_1 \otimes U_2 \otimes I)^\dagger(|\phi\rangle \otimes I)\|. \) Due to the symmetry of Equation 5.4 and Equation 5.5, to show \( h(T_A, T_B) \leq 1 \), it is sufficient to show \( g \leq 1/2 \) for some states \( \psi \) and \( \phi \).

Let \( |\psi\rangle = \frac{1}{2}(|0000\rangle + |0101\rangle + |1010\rangle + |1111\rangle) \) and \( |\phi\rangle = \frac{1}{2}(|0000\rangle + |1001\rangle + |0110\rangle + |1111\rangle) \), we have

\[
g = \frac{1}{4}\|((0000) + (0101) + (1010) + (1111))(U_1 \otimes U_2 \otimes I)^\dagger \\
\hspace{2cm}((|0000\rangle + |1001\rangle + |0110\rangle + |1111\rangle) \otimes I)\| \\
= \frac{1}{4}\|((00)(U_1 \otimes U_2)^\dagger(|00\rangle \otimes I) + (01)(U_1 \otimes U_2)^\dagger(|10\rangle \otimes I) \\
+ (10)(U_1 \otimes U_2)^\dagger(|01\rangle \otimes I) + (11)(U_1 \otimes U_2)^\dagger(|11\rangle \otimes I)\|. \]
Let \( p \overset{\text{def}}{=} \| \langle 00 \rangle (U_1 \otimes U_2)\dagger (|00\rangle \otimes I) + \langle 01 \rangle (U_1 \otimes U_2)\dagger (|10\rangle \otimes I) \| \), Then

\[
p = \sqrt{\langle 0|U_1\dagger (|0\rangle \otimes I) \otimes \langle 0|U_2\dagger (|0\rangle \otimes I) + \langle 0|U_1\dagger (|1\rangle \otimes I) \otimes \langle 1|U_2\dagger (|0\rangle \otimes I) \|}
\]

(5.6)

\[
\leq \sqrt{\langle 0|U_1\dagger (|0\rangle \otimes I) U_1|0\rangle + \langle 0|U_1\dagger (|1\rangle \otimes I) U_1|0\rangle} + \langle 1|U_2\dagger (|0\rangle \otimes I) U_2|1\rangle
\]

(5.7)

\[
\leq \sqrt{\langle 0|U_1 U_1|0\rangle \cdot \sqrt{\text{tr}(U_2\dagger (|0\rangle \otimes I) U_2)} \leq 1,}
\]

where Equation 5.6 is from Cauchy-Schwartz and Equation 5.7 is because \( \|U_1\|, \|U_2\| \leq 1 \). Similarly, we can prove \( \|\langle 10\rangle (U_1 \otimes U_2)\dagger (|01\rangle \otimes I) + \langle 11\rangle (U_1 \otimes U_2) (|11\rangle \otimes I) \| \leq 1 \).

Thus \( g \leq 1/2 \). It follows that \( f \leq 1/4 \) and \( |h(T_A, T_B)| \leq 1 \). On the other hand, it is easy to verify \( h(\text{SWAP}) = 4 \). Thus \( \|\text{SWAP}\|_\gamma \geq 4 \).

To prove \( \|\text{SWAP}\|_\gamma \leq 4 \), we decompose SWAP as the following,

\[
\text{SWAP} = \langle 00\rangle \langle 00\rangle + \langle 01\rangle \langle 10\rangle + \langle 10\rangle \langle 01\rangle + \langle 11\rangle \langle 11\rangle
\]

\[
\cdot (\langle 00\rangle \langle 00\rangle + \langle 01\rangle \langle 10\rangle + \langle 10\rangle \langle 01\rangle + \langle 11\rangle \langle 11\rangle)
\]

\[
= \frac{1}{2} (I \otimes I + Z \otimes Z + X \otimes X - Y \otimes Y)
\]

\[
\cdot \frac{1}{2} (I \otimes I + Z \otimes Z + X \otimes X - Y \otimes Y)
\]

\[
= \frac{1}{4} ((I \cdot I) \otimes (I \cdot I) + (I \cdot Z) \otimes (I \cdot Z)
\]

\[
+ (I \cdot X) \otimes (I \cdot X) - (I \cdot Y) \otimes (I \cdot Y)
\]

\[
+ (Z \cdot I) \otimes (Z \cdot I) + (Z \cdot Z) \otimes (Z \cdot Z)
\]

\[
+ (Z \cdot X) \otimes (Z \cdot X) - (Z \cdot Y) \otimes (Z \cdot Y)
\]

\[
+ (X \cdot I) \otimes (X \cdot I) + (X \cdot Z) \otimes (X \cdot Z)
\]

\[
+ (X \cdot X) \otimes (X \cdot X) - (X \cdot Y) \otimes (X \cdot Y)
\]

\[
- (Y \cdot I) \otimes (Y \cdot I) - (Y \cdot Z) \otimes (Y \cdot Z)
\]

\[
- (Y \cdot X) \otimes (Y \cdot X) + (Y \cdot Y) \otimes (Y \cdot Y)).
\]
Since operator norm $\|I\| = \|X\| = \|Y\| = \|Z\| = 1$, then $(I \cdot I) \otimes (I \cdot I) \leq I \cdot I \leq I^4 = 1$. Similarly the maximum tensor norm of all other 15 term in the above equation is no more than 1. Thus $\|\text{SWAP}\| \leq 4$.

This completes the proof of $\|\text{SWAP}\| = 4$.

5.3.3 Maximum tensor norm of measuring one qubit and sending the result

First we prove $\|\text{CC}\| \leq 2$. Recall that

$$\text{CC} \overset{\text{def}}{=} (\langle 0 | A \cdot | 0 \rangle_A \otimes | 0 \rangle_B \langle 0 | B) + (\langle 1 | A \cdot | 1 \rangle_A \otimes | 1 \rangle_B \langle 1 | B).$$

Observe that diamond norm $\|\langle 0 | A \cdot | 0 \rangle_A \| \leq \|\langle 0 | A \| \| | 0 \rangle_A \| \leq \| | 0 \rangle_B \langle 0 | B \| = 1$. Similarly, diamond norm $\|\langle 1 | A \cdot | 1 \rangle_A \| \leq 1$ and $\| | 1 \rangle_B \langle 1 | B \| \leq 1$.

Therefore, the maximum tensor norm $\|\text{CC}\| \leq 2$.

Then we show $\|\text{CC}\| \geq 2$. For any local superoperators $T_A \in L(N_A, M_A \otimes \mathcal{F})$ and $T_B \in L(N_B, M_B \otimes \mathcal{G})$, let

$$h(T_A, T_B) \overset{\text{def}}{=} 2\text{tr}((T_A \otimes T_B \otimes I_C)(\rho)M),$$

where $\rho = \frac{1}{2}(\langle 0 | A \otimes | 0 \rangle_C \langle 0 | C + \langle 0 | A \otimes | 1 \rangle_C \langle 1 | C) + \langle 1 | A \otimes | 0 \rangle_C \langle 0 | C + \langle 1 | A \otimes | 1 \rangle_C \langle 1 | C)$. Then $h(\text{CC}) = 2$. For superoperators $T_A$ and $T_B$ with $\|T_A\| = \|T_B\| = 1$, let $\alpha_0 = T_A(|0\rangle_A \langle 0|) \leq \|T_A\| \|0\rangle_A \langle 0| \|_{\text{tr}} \leq 1$, $\alpha_1 = T_A(|1\rangle_A \langle 1|) \leq 1$,
\[ T_B = \text{tr}_G(|v\rangle\langle w|), \text{ where } ||v||, ||w|| \leq 1. \] Substitute \( \rho \) and \( M \), we have

\[
|h(T_A, T_B)| = |\text{tr}((T_A|0\rangle_A\langle 0|_A)\text{tr}_G(|v\rangle\langle w|))|0\rangle_B\langle 0|_B + (T_A|1\rangle_A\langle 1|_A)\text{tr}_G(|v\rangle\langle w|))|1\rangle_B\langle 1|_B|
\]

\[
= |\text{tr}(\alpha_0(|0\rangle_B \otimes I_G)|v\rangle\langle w|(|0\rangle_B \otimes I_G) + \alpha_1(|1\rangle_B \otimes I_G)|v\rangle\langle w|(|1\rangle_B \otimes I_G)|
\]

\[
\leq |(|0\rangle_B \otimes I_G)|v\rangle| \times |(|w|(|0\rangle_B \otimes I_G)| + |(|1\rangle_B \otimes I_G)|v\rangle| \times |(|w|(|1\rangle_B \otimes I_G)|
\]

\[
\leq \sqrt{((|0\rangle_B \otimes I_G)|v\rangle)^2 + ((|1\rangle_B \otimes I_G)|v\rangle)^2}
\]

\[
= \sqrt{\text{tr}(|v\rangle\langle v|) \cdot \text{tr}(|w\rangle\langle w|)} = 1,
\]

where Equation 5.8 is because \( \alpha_0 \leq 1, \alpha_1 \leq 1 \) and Equation 5.9 is due to Cauchy-Schwartz. This completes the proof that \( ||CC||_\gamma \geq 2 \). Thus \( ||CC||_\gamma = 2 \).

### 5.3.4 Maximum tensor norm of sending one quantum bit

First we prove \( ||QC||_\gamma \leq 4 \). Recall that

\[
QC \overset{\text{def}}{=} \sum_{i,j \in \{0,1\}} (|i\rangle_A \cdot |j\rangle_A) \otimes |i\rangle_B \langle j|_B.
\]

Observe that \( |||i\rangle_A \cdot |j\rangle_A||_\diamond \leq |||i\rangle_A|| |||j\rangle_A|| = 1 \) and similarly \( |||i\rangle_B \langle j|_B||_\diamond \leq 1 \). Therefore, the maximum tensor norm \( ||QC||_\gamma \leq 4 \).

Then we show \( ||QC||_\gamma \geq 4 \). Let \( |\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_B|0\rangle_C + |1\rangle_B|1\rangle_C), |\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_C + |1\rangle_A|1\rangle_C), M = |\phi\rangle\langle \phi| = \frac{1}{2} \sum_{i,j} |i\rangle_B\langle j|_B \otimes |i\rangle_C\langle j|_C, \rho = |\psi\rangle\langle \psi| = \frac{1}{2} \sum_{i,j} |i\rangle_A\langle j|_A \otimes |i\rangle_C\langle j|_C. \) For any local superoperators \( T_A \in L(N_A, M_A \otimes F) \) and \( T_B \in L(N_B, M_B \otimes G) \), let

\[
h(T_A, T_B) \overset{\text{def}}{=} 4\text{tr}((T_A \otimes T_B \otimes I_C)(\rho)M).
\]
Thus \( h(QC) = 4 \). For \( \|T_A\|_\diamond = \|T_B\|_\diamond = 1 \), substitute \( \rho \) and \( M \), we have

\[
|h(T_A, T_B)| = \sum_{i,j} \left| \sum_{k,l} \text{tr}((T_A(|i\rangle_A\langle j|_A) \otimes |j\rangle_C\langle k|_C)(|k\rangle_B\langle l|_B \otimes |k\rangle_C\langle l|_C)) \right|
\]

\[
= \sum_{i,j} |\text{tr}((T_A(|i\rangle_A\langle j|_A) \otimes T_B)\langle j|_B\langle i|_B))|. \]

Let \( T_A = \text{tr}_F(P \cdot Q^\dagger) \), \( T_B = \text{tr}_G(|v\rangle\langle w|) \), where \( \|P\|, \|Q\|, \|v\|, \|w\| \leq 1 \).

\[
|h(T_A, T_B)| = \sum_{i,j} |\text{tr}((\text{tr}_F(|j\rangle_AQ^\dagger P|i\rangle_A)\text{tr}_G(|v\rangle\langle w|))\langle j|_B\langle i|_B)|
\]

\[
= \sum_{i,j} |\text{tr}(|j\rangle_AQ^\dagger P|i\rangle_A\langle i|_B\text{tr}_G(|v\rangle\langle w|)\langle j|_B)|.
\]

Since \( |j\rangle_AQ^\dagger P|i\rangle_A \) is a number, substitute \( |j\rangle_AQ^\dagger P|i\rangle_A \) by \( |j|_BQ^\dagger P|i|_B \),

\[
|h(T_A, T_B)| = \sum_{i,j} |\text{tr}(|j|_BQ^\dagger P|i|_B\langle i|_B\text{tr}_G(|v\rangle\langle w|)\langle j|_B)|
\]

\[
= |\text{tr}(Q^\dagger P\text{tr}_G(|v\rangle\langle w|))| 
\]

\[
(5.10) \leq \|Q^\dagger\|\|P\|\|(|v\rangle\langle w|)\|_{tr} \leq 1. 
\]

This completes the proof of \( \|QC\|_\gamma \geq 4 \). Thus \( \|QC\|_\gamma = 4 \).
5.3.5 Maximum tensor norm of a measurement operator

We now look at the more tricky example in Bennett et al. [14]. We use the same notation as the above paper for the basis:

\[ |\alpha_i\rangle (\text{Alice}) \quad |\beta_i\rangle (\text{Bob}) \]
\[ |\psi_1\rangle = |1\rangle \quad |1\rangle \]
\[ |\psi_2\rangle = |0\rangle \quad \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \]
\[ |\psi_3\rangle = |0\rangle \quad \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \]
\[ |\psi_4\rangle = |2\rangle \quad \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle) \]
\[ |\psi_5\rangle = |2\rangle \quad \frac{1}{\sqrt{2}} (|1\rangle - |2\rangle) \]
\[ |\psi_6\rangle = \frac{1}{\sqrt{2}} (|1\rangle + |2\rangle) \quad |0\rangle \]
\[ |\psi_7\rangle = \frac{1}{\sqrt{2}} (|1\rangle - |2\rangle) \quad |0\rangle \]
\[ |\psi_8\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \quad |2\rangle \]
\[ |\psi_9\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) \quad |2\rangle. \]

For any bipartite superoperator \( T \), let \( h(T) = \text{tr}(T(\rho)M) \), where \( \rho \) and \( M \) are as follows: \( \rho \) is the density operator for a tri-partite state

\[ \frac{1}{\sqrt{8}} \sum_{i=2}^{9} |\alpha_i\rangle_A \otimes |\beta_i\rangle_B \otimes |i\rangle_C, \]

and \( M \) is the measurement operator

\[ M = \sum_{i=1}^{9} |i\rangle_A \langle i|_A \otimes |i\rangle_B \langle i|_B \otimes |i\rangle_C \langle i|_C. \]

Let measurement superoperator \( S \) be define as in Bennett et. al [14], i.e.,

\[ S = \sum_i (|i\rangle_A \otimes |i\rangle_B) \langle \psi_i| \cdot |\psi_i\rangle \otimes (|i\rangle_A \otimes \langle i|_B). \]

Then \( h(S) = 1 \).

Now fix a pair of local superoperators \( T_A \) and \( T_B \) with

\[ T_A = \text{tr}_x(A_1 \cdot A_2^\dagger), \quad \|A_1\| = \|A_2\| = 1, \quad \text{and,} \]
\[ T_B = \text{tr}_G(B_1 \cdot B_2^\dagger), \quad \|B_1\| = \|B_2\| = 1. \]

Without loss of generality, we assume that \( \dim(\mathcal{F}) = \dim(\mathcal{G}) \). We have

\[ |h(T_A, T_B)| = \frac{1}{8} \left| \sum_{i=2}^9 \text{tr}(\langle i | A \otimes I_F | A_i \rangle \langle A_i | A_2^\dagger (i | A \otimes I_F) \rangle \cdot \text{tr}(\langle i | B \otimes I_G | B_1 \rangle \langle B_1 | B_2^\dagger (i | B \otimes I_G) \rangle) \right|. \]

Applying Cauchy-Schwartz, we can upper-bound the above by

\[ \frac{1}{8} \sup_{A, \|A\|=1} \left| \sum_{i=2}^9 \text{tr}(\langle i | A \otimes I_F | A_i \rangle \langle A_i | A_2^\dagger (i | A \otimes I_F) \rangle) \right|. \]

Since \( |\alpha_2\rangle = |\alpha_3\rangle \) and \( |\alpha_4\rangle = |\alpha_5\rangle \), the above equation is upper bounded by 6 (This is not the optimal bound, which can be computed from a semi-definite programming).

Hence \( |h(T_A, T_B)| \leq 3/4 \). This concludes \( \|S\|_{\gamma} \geq 4/3 \). Therefore the superoperator \( S \) is not bi-local.

### 5.4 Connections with communication complexity

First we show that the maximum tensor norm is upper bounded by the amount of classical and quantum communication to realize the superoperator.

**Proof of Theorem 5.4.** Let \( \mathcal{N}_A \) and \( \mathcal{N}_B \) be the Hilbert spaces of the input qubits for Alice and Bob, respectively. Fix a quantum protocol and let \( T \) be the superoperator composed from the communication, the final measurement, and discarding all qubits. Hence \( T : \mathcal{L}(\mathcal{N}) \otimes \mathcal{L}(\mathcal{N}) \rightarrow \mathbb{C} \) maps a density operator \( \rho_{x,y} = |x\rangle \langle x| \otimes |y\rangle \langle y| \) to the acceptance probability \( p_{x,y} = T(\rho_{x,y}) \). By applying Yao’s Lemma (referenced as Lemma 4.20) on quantum communication, if \( q \) is the number of qubits communicated, for some auxiliary systems \( \mathcal{M}_A \) and \( \mathcal{M}_B \), and operators \( A_h \in \mathcal{L}(\mathcal{N}_A, \mathcal{M}_A), \quad B_h \in \mathcal{L}(\mathcal{N}_B, \mathcal{M}_B) \), with \( \|A_h\| \leq 1 \) and \( \|B_h\| \leq 1 \), for all \( h \in \{0,1\}^q \), superoperator
$T$ can be written as

$$\sum_{h, h' \in \{0,1\}^q} \text{tr}_{M_A}(A_h \cdot A_{h'}^\dagger) \otimes \text{tr}_{M_B}(B_h \cdot B_{h'}^\dagger).$$

Hence $\|T\|_\gamma \leq 2^{2q}$. ■

Then we prove that Razborov's lower bound on the \textit{quantum communication complexity} of the Disjointness Problem can be extended to show that any superoperator for computing the Disjointness Problem must have $\exp(\Omega(\sqrt{n}))$ gamma norm.

\textbf{Proof of Theorem 5.5.} Let $P = [p_{x,y}]$ be the acceptance probability matrix for Set Disjointness Problem. Razborov proved $\|P\|_\text{tr} = N \exp(\Omega(\sqrt{n}))$ ([77]).

Let $U$ be a unitary operator so that $\text{tr}(PU) = \|P\|_\text{tr}$. Let $\text{DISJ} : \mathcal{L(N)} \otimes \mathcal{L(N)} \rightarrow \mathbb{C}$ maps a density operator $\rho_{x,y} = |x\rangle\langle x| \otimes |y\rangle\langle y|$ to the acceptance probability $p_{x,y} = \text{DISJ}(\rho_{x,y})$. Define a bilinear mapping

$$h(T) \overset{\text{def}}{=} \sum_{x,y} \text{tr}(T(\rho_{x,y})|x\rangle\langle y|U).$$

Then $h(\text{DISJ}) = \sum_{x,y} \text{tr}(p_{x,y}|x\rangle\langle y|U) = \text{tr}(PU) = N \exp(\Omega(\sqrt{n}))$. For any local superoperators $T_A$ and $T_B$ of diamond norm 1, $|T_A(|x\rangle\langle x|)| \leq 1$, $|T_B(|y\rangle\langle y|)| \leq 1$, hence

$$|h(T_A, T_B)| = |\text{tr}(\sum_{x,y} T_A(|x\rangle\langle x|)T_B(|y\rangle\langle y|)|x\rangle\langle y|U)|$$

$$\leq \|\sum_x T_A(|x\rangle\langle x|)|x\rangle\| \cdot \|\sum_y T_B(|y\rangle\langle y|)|y\rangle\| \leq N.$$

This proves $\|\text{DISJ}\|_\gamma \geq \exp(\sqrt{n})$. ■
CHAPTER VI

Conclusions

In this chapter, we summarize our results in this dissertation. We also discuss future directions and provide some clues.

6.1 Summary of this dissertation

We investigate quantum communication complexities and nonlocality of quantum operations in this dissertation.

For quantum communication complexity, we first study a specific problem — the Hamming Distance problem. We prove a quantum lower bound of $\Omega(d)$, improving over the previous bound of $\Omega(d/\log d)$. Moreover, our lower bound is established in the general two-party model with shared entanglement, while the previous lower bound is proved in the model without entanglement. We also construct a public-coin randomized SMP protocol of $\Omega(d \log d)$, which almost matches the lower bound. This is an improvement over the previous protocols of $O(d^2)$ and $O(d \log n)$ in the same model.

Then we deal with the Log-Equivalence Conjecture. We prove that the Log-Equivalence Conjecture is true for certain block-composed functions $f \square g$. Specifically, when the basic building block $g$ is “hard” enough, there is no exponential gap between quantum and classical communication complexities. We obtain our result
by proving that, if $f$ has no polynomial approximation of degree $d$, for certain class of $g$, the quantum communication complexity of $f \Box g$ is at least $\Omega(d)$ in the general two-party model with shared entanglement. Unlike Razborov’s method that depends on the $f$ being symmetric, we avoid this reliance on the symmetry by taking the dual approach of the polynomial method. We show that a “witness” for $f_n$ requiring a high approximation degree can be turned into a “witness” for the hardness of $f_n \Box g_k$. Our application of this dual approach appears to be the first demonstrations of its usefulness to our best knowledge.

For nonlocality of quantum operations, we take two different approaches to quantify it. Our first approach is the minimum amount of classical communication, denoted by $\text{Com}(Q)$, required to simulate the quantum measurement $Q$. The main result is a general upper bound on $\text{Com}(Q)$ in terms of a certain tensor norm on $Q$. In particular, if $K$ is the dimension of the space that $Q$ acts on, then $\text{Com}(Q) = O(K)$.

We apply the above result on the role of shared entanglement in assisting communications. It implies that, if a two-party, interactive quantum protocol uses $q$ qubits of communication and $m$ qubits of shared entanglement, then it can be simulated by a classical protocol using $\exp(O(q))$ bits with shared randomness. The simulation does not depend on $m$ and it can be carried out in the SMP model. Setting $q$ to a constant, this implies that constant cost quantum protocols with unlimited shared entanglement and constant cost classical protocols with unlimited shared randomness compute the same set of functions.

The above result also implies that local measurements of an entangled state can be simulated by a local hidden variable model with a constant amount of communication, as long as the number of measurement outcomes is constant.

Our second approach is to define a maximum tensor norm on superoperators with
respect to the diamond norm. We prove that the value of this maximum tensor norm is a criterion for deciding whether a bipartite superoperator is bi-local. By using the dual characterization of the maximum tensor norms, we are also able to compute the exact maximum tensor norms of the following elementary superoperators: the superoperators for the Controlled-NOT gate, the SWAP gate, measuring one qubit and sending the measurement result, and sending one quantum bit.

Furthermore, we have a connection between the maximum tensor norm and communication complexity: if there exists a communication protocol that realizes a bipartite superoperator $T$ with $c$ classical bits and $q$ qubits, then the maximum tensor norm of $T$ is at most $2^{c+2q}$. Thus the maximum tensor norm can be used to prove quantum communication lower bounds. We derive a lower bound method which is at least as powerful as the lower bound method derived independently by Razborov [77].

6.2 Future directions

For the Hamming Distance problem, we conjecture that our quantum lower bound of $\Omega(d)$ is tight. It seems plausible to remove the $O(\log d)$ factor in our upper bound. Recently, Aaronson and Ambainis [1] sharpened the upper bound of the Set Disjointness problem from $O(\sqrt{n \log n})$ to $O(\sqrt{n})$ using quantum local search instead of Grover’s search. In their method, it takes only constant communication of qubits to synchronize two parties and simulate each quantum query. From Yao’s protocol [97], one can easily derive an $O(d \log d)$ two way, interactive quantum communication protocol using quantum counting [22] and the connection between quantum query and communication [25]. Methods similar to [1] might help remove the $O(\log d)$ factor in this upper bound.
The Log-Equivalence Conjecture remains open for general total Boolean functions. One possible way to extend our results is to apply them to asymmetric primitive functions. Moreover, our approach seems to possess the ability of “hardness amplification”, i.e., it can turn a weak lower bound on a primitive function $g_k$ into a strong lower bound on composed function $f_n \Box g_k$. One such example is $g_k$ being the Set Disjointness problem and $f_n$ being the OR function. Finding more examples of such “hardness amplification” would be very interesting.

A recent advance on the usefulness of quantum entanglement was made by Gavin-sky [48], in which he showed that entanglement is responsible for exponential savings for some communication tasks and in some restricted models. Whether or not entanglement could result in exponential savings for the more standard two-party, interactive communication model and for the computation of functions remains unsolved. Can our result on removing the entanglement be strengthened to that one can always use an amount of entanglement linear in size of the messages, with at most a logarithmic additive term?

We demonstrate that the maximum tensor norm allows us to prove communication lower bounds to realize a superoperator. This approach is potentially stronger than Razborov’s approach [77] to prove quantum communication complexity of computing a Boolean function, though we have not been able to give a concrete example showing a gap of those two bounds. Another direction is to extend our result from bi-local to LOCC (local operation and classical communication), i.e., can maximum tensor norm be used to distinguish the set of superoperators realizable by LOCC?

For nonlocality of quantum measurements, it would be interesting to relate $\text{Com}(Q)$ to other measures of nonlocality, such as the entanglement capacity, and the minimum number of elementary gates, or the amount of time for evolving some elementary
Hamiltonian, needed to approximate $Q$. It is conceivable that the comparisons of those measures may lead to a unique and representative measure of nonlocality.
APPENDIX
This appendix collects Dirac notations and various matrix norms that are used in this dissertation. We refer the reader to [70, 59] for more details.

A.1 Dirac notations

We use Dirac notation to represent vectors. Let $|\phi\rangle$ denote a complex vector in a finite dimensional Hilbert space. Given an orthogonal basis $\{|i\rangle, \; 0 \leq i \leq d - 1\}$ for the Hilbert space, the vector $|\phi\rangle$ can be represented as a complex column vector

$$
\begin{pmatrix}
\alpha_0 \\
\alpha_1 \\
\vdots \\
\alpha_{d-1}
\end{pmatrix}
$$

where $\alpha_0, \alpha_1, \ldots, \alpha_{d-1}$ are complex numbers. For example, the basis vector $|i\rangle$ can be represented as the column vector with the $i$th row being 1 and 0 elsewhere. The basis $\{|i\rangle, \; 0 \leq i \leq d - 1\}$ is usually called computational basis.

The complex conjugate of $|\phi\rangle$, denoted by $\langle\phi|$, can be represented as a row vector $(\alpha_0^*, \alpha_1^*, \ldots, \alpha_{d-1}^*)$, where $\alpha^*$ is the complex conjugate of $\alpha$. The inner product of vectors $|\phi\rangle = \begin{pmatrix}
\alpha_0 \\
\alpha_1 \\
\vdots \\
\alpha_{d-1}
\end{pmatrix}$ and $|\psi\rangle = \begin{pmatrix}
\beta_0 \\
\beta_1 \\
\vdots \\
\beta_{d-1}
\end{pmatrix}$ is denoted by $\langle\phi|\psi\rangle \equiv \sum_{i=0}^{d-1} \alpha_i^* \beta_i$. 

For a complex vector $|\phi\rangle$, its length is defined as $\|\phi\| \overset{\text{def}}{=} \sqrt{\langle \phi | \phi \rangle}$.

Let $\mathcal{N}$ and $\mathcal{M}$ denote Hilbert spaces. Let $\mathbf{L}(\mathcal{N}, \mathcal{M})$ denote the space of linear operators $A : \mathcal{N} \rightarrow \mathcal{M}$ and $\mathbf{L}(\mathcal{N})$ be a shorthand for $\mathbf{L}(\mathcal{N}, \mathcal{N})$. For a linear operator $A \in \mathbf{L}(\mathcal{N}, \mathcal{M})$, suppose $|\phi_1\rangle, \ldots, |\phi_n\rangle$ is a basis for $\mathcal{N}$ and $|\psi_1\rangle, \ldots, |\psi_m\rangle$ is a basis for $\mathcal{M}$. Represent $A|\phi_j\rangle$ as follows,

$$A|\phi_j\rangle = \sum_i A_{ij} |\psi_i\rangle.$$ 

Then the $m \times n$ matrix with $(i, j)$th entry being $A_{ij}$ is called the matrix representation of the operator $A$ with respect to basis $\{|\psi_i\rangle, |\phi_j\rangle\}$. Linear operator $A$ can be written as $A = \sum_{i,j} A_{ij} |\psi_i\rangle \langle \phi_j|$.

### A.2 Operator norms and trace norms

For a linear operator $A \in \mathbf{L}(\mathcal{N}, \mathcal{M})$, there exists a unique linear operator $A^\dagger \in \mathbf{L}(\mathcal{M}, \mathcal{N})$ such that

$$\langle \phi | A \psi \rangle = \langle A^\dagger \phi | \psi \rangle,$$

for all $|\psi\rangle \in \mathcal{N}, |\phi\rangle \in \mathcal{M}$. This linear operator $A^\dagger$ is called the adjoint of the operator $A$. The matrix representation of $A^\dagger$ is just the conjugate transpose of the matrix representation of the operator $A$.

A linear operator $A \in \mathbf{L}(\mathcal{N})$ is unitary if it preserves the length of all vectors $|\phi\rangle \in \mathcal{N}$, i.e., $\|A|\phi\| = \||\phi\|\|$ for any $|\phi\rangle \in \mathcal{N}$. Any matrix representation of $A$ is a unitary matrices. A linear operator $A$ is Hermitian if $A = A^\dagger$. A linear operator $A$ is positive if $\langle \phi | A |\phi \rangle \geq 0$ for all $\phi \in \mathcal{N}$.

Let $\{|i\rangle\}$ be a basis for $\mathcal{N}$. For a linear operator $A \in \mathbf{L}(\mathcal{N})$, the trace of $A$, denoted by $\text{tr}(A)$, is defined as

$$\text{tr}(A) = \sum_i \langle i | A | i \rangle.$$
which is the absolute value of the sum of diagonal entries of any matrix representation of \( A \). The \textit{operator norm} (usually called norm) of \( A \) is defined as

\[
\|A\| \overset{\text{def}}{=} \max\{\|A\phi\| : \|\phi\| \leq 1, \phi \in \mathcal{N}\}.
\]

The \textit{trace norm} of \( A \) is defined as

\[
\|A\|_{\text{tr}} \overset{\text{def}}{=} \max\{\|\text{tr}(AB)\| : \|B\| \leq 1, B \in \mathbf{L}(\mathcal{N})\}.
\]

From the above equation, it is easy to observe \( \text{tr}(A) \leq \|A\|_{\text{tr}} \) and \( \|AB\|_{\text{tr}} \leq \|A\|\|B\|_{\text{tr}} \).

\section*{A.3 Superoperators and Diamond norms}

For a bipartite operator \( A \in \mathbf{L}(\mathcal{N}_1 \otimes \mathcal{N}_2) \) with \( A = \sum_i x_i \otimes y_i \), where \( x_i \in \mathbf{L}(\mathcal{N}_1) \) and \( y_i \in \mathbf{L}(\mathcal{N}_2) \), the partial trace of \( A \) over the space \( \mathcal{N}_2 \) is defined as

\[
\text{tr}_{\mathcal{N}_2}(X) \overset{\text{def}}{=} \sum_i x_i(\text{tr}(y_i)).
\]

A \textit{superoperator} is a linear mapping from operators to operators. Any superoperator \( T : \mathbf{L}(\mathcal{N}) \to \mathbf{L}(\mathcal{M}) \) can be represented in the following form: \( T = \text{tr}_{\mathcal{F}}(A \cdot B^\dagger) \), i.e., for \( \rho \in \mathbf{L}(\mathcal{N}) \), \( T(\rho) = \text{tr}_{\mathcal{F}}(A\rho B^\dagger) \), where \( A \) and \( B \) are linear mappings from \( \mathcal{N} \) to \( \mathcal{M} \otimes \mathcal{F} \) and \( \text{tr}_{\mathcal{F}} \) is partial trace.

Consider all representations of the superoperator \( T : \mathbf{L}(\mathcal{N}) \to \mathbf{L}(\mathcal{M}) \) in the form of \( T = \text{tr}_{\mathcal{F}}(A \cdot B^\dagger) \). The \textit{diamond norm} of \( T \) is defined as follows [59],

\[
\|T\|_{\diamond} \overset{\text{def}}{=} \inf\{\|A\|\|B\| : \text{tr}_{\mathcal{F}}(A \cdot B^\dagger) = T, \ A, B \in \mathbf{L}(\mathcal{N}, \mathcal{M} \otimes \mathcal{F})\}.
\]

Diamond norms have the following properties.

\textbf{Proposition A.1.} \emph{Let} \( T \emph{ be a superoperator, then}

\emph{(1).} \( \|kT\|_{\diamond} = k\|T\|_{\diamond} \), \emph{where} \( k \emph{ is a nonnegative number}.\)
(2). The triangle inequality: \(|T_1 + T_2|_\diamond \leq |T_1|_\diamond + |T_2|_\diamond\).

(3). \(|T_1 \otimes T_2|_\diamond = |T_1|_\diamond |T_2|_\diamond\). This implies that diamond norms are stable, i.e.,

\(|T_1 \otimes I|_\diamond = |T_1|_\diamond\).

Diamond norms have a dual characterization (e.g., [59], Theorem 11.1). Let \(I_G\) be the identify superoperator on auxiliary spaces \(G\) with dimension no less than that of \(\mathcal{N}\). Then

\[
(A.4) \quad |T|_\diamond = \sup_{\rho \in L(\mathcal{N} \otimes G), \rho \neq 0} \frac{|(T \otimes I_G)(\rho)|_{tr}}{||\rho||_{tr}}
\]

From the dual characterization, we have

\[
(A.5) \quad |T(\rho)|_{tr} \leq |T|_\diamond ||\rho||_{tr}
\]

A physically realizable superoperator is a superoperator that has the following form: \(T = \text{tr}_\mathcal{F}(V \cdot V^\dagger): \rho \rightarrow \text{tr}_\mathcal{F}(V \rho V^\dagger)\), where \(V \in L(\mathcal{N}, \mathcal{N} \otimes \mathcal{F})\) is an isometric embedding. Physically realizable superoperators have the following properties.

**Proposition A.2.** Let \(T\) be a physically realizable superoperator, then

(1). \(T\) is trace preserving, i.e, \(\text{tr}(T(\rho)) = \text{tr}(\rho)\).

(2). \(|T|_\diamond = 1\).

**A.4 Maximum tensor norms**

For any bipartite operator (or superoperator) \(z \in \mathcal{H}_A \otimes \mathcal{H}_B\), its maximum tensor norm is defined as

\[
(A.6) \quad ||z||_\gamma \overset{\text{def}}{=} \inf \left\{ \sum_i ||x_i|| ||y_i|| : z = \sum_i x_i \otimes y_i, x_i \in \mathcal{H}_A, y_i \in \mathcal{H}_B \right\}.
\]

Any other tensor norm \(|| \cdot ||_\alpha\) satisfies \(||z||_\alpha \leq ||z||_\gamma\). In this dissertation, we focus on the maximum tensor norm of bipartite superoperators with respect to the
diamond norm. More Precisely, for a bipartite physically realizable superoperator $T^{AB} : \mathbf{L}(\mathcal{N}_A \otimes \mathcal{N}_B) \to \mathbf{L}(\mathcal{M}_A \otimes \mathcal{M}_B)$, we define its maximum tensor norm

\begin{equation}
\|T\|_\gamma \overset{\text{def}}{=} \inf \left\{ \sum_i \|T^A_i\|_\diamond \|T^B_i\|_\diamond : T = \sum_i T^A_i \otimes T^B_i \right\},
\end{equation}

where $T^A_i \in \mathbf{L}(\mathcal{N}_A) \to \mathbf{L}(\mathcal{M}_A), T^B_i \in \mathbf{L}(\mathcal{N}_B) \to \mathbf{L}(\mathcal{M}_B)$. This maximum tensor norm does not appear to have been studied before.
BIBLIOGRAPHY


