Adaptive Quantum Tomography of Pure States

and Unitary Gates

by

Carson Yanning Li

A Thesis

presented to

the University of Guelph

in partial fulfillment

for the degree of

Master of Science

in

Physics

Guelph, Ontario, Canada

©Carson Yanning Li, Jan. 2017
ABSTRACT

Adaptive Quantum Tomography of Pure States

and Unitary Gates

Carson Yanning Li
University of Guelph, 2017

Adviser:
Professor Bei Zeng

The successful implementation of complex quantum algorithms depend crucially on our ability to determine unknown quantum states and operations. In this work, we present adaptive and non-adaptive methods to determine a 1-qubit unitary gate with 5 and 6 Pauli measurements. We demonstrate the method on the Bloch sphere to show that studying higher dimensional real space may help in finding tomography methods of multiple qubit unitary gates. Next, we show an adaptive method to uniquely determine a general $d$–dimensional pure state among all quantum states with at most $2d - 1$ measurements. This method is then applied to determine a general $d$–dimensional unitary gate with at most $d^2 + d - 1$ measurements. These methods are applied in tomographing a 2-qubit universal gate set with five unitary gates. On the NMR experimental system, the lowest fidelity achieved was above 97% with 42 Pauli measurements, comparing to 99% using traditional method that requires 240 Pauli measurements.
Acknowledgment

I would like to thank Prof. Bei Zeng for giving me the opportunity to experience graduate school and research. This work is not possible without her exceptional patience, tolerance and guidance, especially at times when I was not working in the right direction. My committee members Prof. David Kribs, Prof. Eric Poisson and Prof. Robert Wickham, and the physics department chair Prof. Paul Garrett also gave me their wisdom on research and life as a graduate student in general, for which I am forever grateful.

On the theoretical side, I am in debt to Dr. Nengkun Yu. His sharp insight is the safety net for keeping my work free from errors. On the experimental side, I would like to thank Dr. Dawei Lu for so many helpful explanation on the capabilities and limitations in laboratories in the real world, which keeps me grounded as a theorist. Dr. Xian Ma, Tyler Jackson and Kent Fisher also contributed many times to my understanding of their work, which is essential in the completion of this thesis.

I greatly appreciate the support from the department’s administrative team, Linda Stadig, Kiley Rider, Steve Kempf, Lorena Barker and Reggi Vallillee and from the Office of Graduate Studies staff Sue Shaw and Kristen Campbell. I thank the Institute for Quantum Computing at the University of Waterloo, which I spent most of my time at, for providing a very stimulating research environment with many high quality seminars.

I was fortunate enough to sit in the same office with my good friend Joel David Klassen. It is rare to have someone with whom your discussion on many things can resonate. It is invaluable when these two people come from two completely different cultures. His companies during the countless trips between Waterloo and Guelph.
are the most cherished of my memories. A special thanks goes to the Quiz room administrator Cindy Wells for accompanying us in the teaching assistance schedule.

Outside the academia, I thank Jiani Wang, Qian Wang, my cousin Xin Sui, my aunt Liping Bai and all my other friends in Waterloo for enriching my life.

Finally, I thank my parents’ unyielding support that allowed me to pursue my passion.
## Contents

1 **Introduction** 1

1.1 Importance of QPT ................................. 1

1.2 Current Approaches of QPT .......................... 2

1.3 Why Adaptive Unitary Process Tomography? ........ 8

1.4 Related Recent Works .............................. 10

1.4.1 Algebraic Framework of QST ................... 10

1.4.2 QST of Pure States with Pauli Measurements .... 13

1.4.3 QPT for Unitary Gates ......................... 17

1.4.4 Summary of Important Results ................. 23

2 **Basic Concepts** 27

2.1 Dirac Notation .................................... 27

2.2 Born’s Rule and Projective Measurements .......... 28

2.3 Hermitian Operators and Expectation Values ...... 29

2.4 Mixed States and Density Matrices ................. 30

2.5 Hilbert-Schmidt Inner Product .................... 32

2.6 Quantum State Tomography ....................... 33

2.7 Global and Relative Phase of Quantum States ...... 35

2.8 Unitary Operators ............................... 36
2.9 Pauli Operators ........................................... 38
2.10 Positive Linear Maps ..................................... 40
2.11 Choi’s Theorem and Channel State Duality .......... 41
2.12 General Quantum Operations ......................... 42
2.13 Standard Quantum Process Tomography ............... 45

3 QPT of 1-Qubit Unitary Gates ............................... 47
3.1 Parametrization and QPT Methods ..................... 47
3.2 Experimental Implementation ........................... 52
3.3 Maximum Likelihood Estimation ....................... 53
3.4 Experimental Plan ........................................ 57
3.5 Optimization and Results ................................ 59
3.6 Visualization in 3D ...................................... 64
3.7 Homomorphism to $SO(3)$ and Euler Angles .......... 72

4 Adaptive Quantum Tomography in Arbitrary Dimensions 79
4.1 Adaptive Pure State Tomography ...................... 79
4.2 Adaptive Unitary Process Tomography ................. 86

5 Application to 2-Qubit NMR Systems .................... 90
5.1 Standard Quantum Process Tomography ............... 90
5.2 Unitary Quantum Process Tomography ............... 97
5.3 Experiments and Results ................................ 98

6 Conclusion ..................................................... 103
List of Tables

1.1  Comparison between recent and our results on tomography ........ 23

3.1  Non-Adaptive Measurements in NMR .................................. 48
3.2  Unitary Gates Tomographed in NMR ................................. 58
3.3  Unitary Gates Tomographed in Optics ............................... 58

5.1  Average State Fidelity: standard and unitary QPT ............... 100
List of Figures

3.5.1 Unitary QPT from NMR data using \( l_1 \) ........................................ 62
3.5.2 Unitary QPT from NMR data using \( l_2 \) ........................................ 63
3.6.1 A Point on the Unit Sphere ....................................................... 65
3.6.2 the Bloch Sphere ................................................................. 65
3.6.3 Tomography of \( |u_0\rangle \) on the Bloch sphere ......................... 68
3.6.4 Tomography of \( |u_+\rangle \) on the Bloch sphere ......................... 70
3.6.5 Special Case in Tomography of \( |u_+\rangle \) ................................. 71
3.7.1 Euler Angles of a general rotation in space ............................. 78
5.3.1 \( \hat{P} \) Matrix of all Gates: Theory, Standar QPT and UQPT .......... 101
Chapter 1

Introduction

The task to characterize the action of a quantum operation on its input states is called Quantum Process Tomography (QPT) [1, 2]. The action of the quantum operation cannot be measured directly and thus must be inferred from the output states of known inputs. So it is essential to be able to fully determine an unknown quantum state by measurements. This is called Quantum State Tomography (QST) [3, 4, 5].

1.1 Importance of QPT

The question of what is the minimum number of measurements required for QST and QPT is far from trivial. However, with the development of quantum information processing, answering this question is becoming increasingly important. The initialization of qubits (quantum “bits”) to designated states and the conduction of specified operations on them are the basic tasks in manipulating quantum information. QST and QPT are the main methods by which we can evaluate our ability to performs these tasks. Unlike a classical bit, a qubit is analog and can be varied continuously.
between 0 and 1. Therefore, it is much more prone to accumulating errors due to inaccuracy in the physical implementation of quantum operations [6, 7, 8]. Another source of error is the decoherence of qubits’ states due to inevitable interactions with the environment [9, 10, 11]. Losing coherence among quantum superpositions disrupts quantum computations like Shor’s algorithm [12, 13]. Fortunately, it has been shown by the quantum threshold theorem that quantum computations can work as long as errors are kept below a certain limit. Therefore, it is necessary to perform error corrections regularly during prolonged quantum computations [6, 14]. QPT and QST will allow us to benchmark the precision of quantum operations and quantify these errors in qubits so that error corrections are performed at the appropriate stage during computation.

1.2 Current Approaches of QPT

The main challenge for complete QPT is that under no prior information, we must assume the quantum operation to be in the most general form. Then the number of measurements required to fully characterize the operation is $O(2^{4N})$, where $N$ is the number of qubits the operation acts on [15, Section 8.4.2]. So the resource needed for complete QPT increases exponentially as we use more qubits in quantum computers.

There has been many efforts to side step this exponential growth. Compressed sensing is a signal processing technique first proposed by T. Tao et al. [16]. It provides a way to reconstruct sparse signals with sampling rate much lower than the Nyquist frequency [17]. Applying this method to QPT, the number of measurements needed for reconstruction is reduced to $O(sN)$, where $s$ characterizes the sparsity of the process matrix describing the quantum operation. However, this method requires the prior knowledge of the basis in which the process matrix is sparse [18].
Randomized benchmarking is first proposed by J. Emerson et al. [19]. It is an efficient scalable way to estimate the average fidelity of quantum operations. The method assumes the existence of a universal noise model associated with some physical system that affects all implemented quantum operations. The average gate fidelity is postulated to be determined solely on this noise. By reversing the action of a randomly chosen unitary operation on an arbitrary pure initial state, the average gate fidelity can be estimated, up to precision $O(2^{-N^2/2})$, from the deviation of the final state from the initial state in a projective measurement [19, 20, 21]. This method provides an overall estimate of the precision of physical systems and cannot estimate the fidelity of any particular quantum operation. To do so, a method called interleaved randomized benchmarking is proposed by [22]. While randomized benchmarking is very efficient and scalable in estimating overall average gate fidelity, it is essentially a black box approach that cannot quantify individual internal components. So it is strictly speaking not a tomography method. Much more detailed error information on specified components of specified gates are often needed in building quantum computers.

Ancilla-assisted QPT is first proposed by D’Ariano et al. for both pure and general quantum operations [23]. The idea is inspired by the Choi-Jamiolkowski isomorphism, also known as the channel-state duality, that establishes a one-to-one correspondence between completely positive trace preserving (CPTP) maps, i.e. general quantum operations, and their corresponding Choi matrix states [24, 25]. By attaching different input states to different basis of a second system, a.k.a. the ancilla, in a bipartite state and performing the operation locally only on the first system, the final state will be the direct sum of the outputs from all the input states. This technique of using the quantum superposition to get multiple outputs from a single quantum operation is called quantum parallelism. If the set of attached
input states are sufficient to fully characterize the operation, then we will be able to
determine the operation uniquely by a single QST of the output bipartite state.
In particular, it has been shown in [26] that the maximally entangled input state
\[ |\tilde{I}\rangle = \frac{1}{\sqrt{2^N}} \sum_{i=0}^{2^N-1} |i\rangle \otimes |i\rangle \]
imprints the complete information of a general quantum operation onto the output state. Many other bipartite input states can also capture full information of the operation as long as their Schmidt rank is at least \(2^N\), but the maximally entangled state yields the lowest experimental errors [27, 28]. Although ancilla-assisted QPT only requires to perform QST on one output state, this output state is in a Hilbert space enlarged by tensor multiplying the ancilla. Generally, QST requires \(O(2^N)\) measurements to fully determine the output state attached to each Schmidt rank. So the total number of measurements required is still \(O(2^4N)\). There has been efforts to reduce this number by global measurements on the bipartite state. However, this type of measurements require controlled interactions in many body systems and are very difficult to implement experimentally [29, 30].

**Direct characterization of quantum dynamics (DCQD)** is first proposed by Lidar et al. [31] for general quantum operations on a single qubit. The method does not require any complete QST but utilizes special effects of the action of local Pauli operators on Bell states. For CPTP maps between 1-qubit states, we can express the \(\chi-\)matrix of standard QPT in basis of the identity and the Pauli operators. That is to say that the image of the map can be expanded as 
\[ \varepsilon(\rho) = \sum_{m,n=I,X,Y,Z} \chi_{mn} m\rho n^\dagger. \]
The standard Bell states 
\[ |\phi^\pm\rangle = \frac{1}{\sqrt{2}} (|00\rangle \pm |11\rangle) \]
and 
\[ |\psi^\pm\rangle = \frac{1}{\sqrt{2}} (|10\rangle \pm |01\rangle) \]
form a complete set of 2-qubit orthonormal basis. There are many ways to transform between these Bell states using local Pauli operators. One way is that 
\[ |\phi^-\rangle, |\phi^+\rangle \]
and 
\[ |\psi^-\rangle \]
can be generated by local error operations of phase flip (\(Z\)), bit flip (\(X\)), and phase flip and then bit flip (\(XZ = -iY\)) on the first qubit of \( |\phi^+\rangle \), respectively. So we have 
\[ |\phi^-\rangle = Z \otimes I |\phi^+\rangle, |\psi^+\rangle = X \otimes I |\phi^+\rangle \]
and 
\[ |\psi^-\rangle = -iY \otimes I |\phi^+\rangle. \]
Therefore, if we perform the operation $\varepsilon(\cdot)$ locally on the first qubit of input $|\phi^+\rangle$ and then a projective measurement on the output $\varepsilon \otimes I(|\phi^+\rangle \langle \phi^+|)$ with respect to the standard Bell state complete orthonormal basis, we will get each diagonal element $\chi_{mm}$, $m = I$, $X$, $Y$, $Z$ as the probability of measuring Bell state $|\phi^+\rangle$, $|\psi^+\rangle$, $|\psi^-\rangle$, $|\phi^-\rangle$. These diagonal elements are called quantum dynamical populations. The other off-diagonal elements of the $\chi-$matrix are called quantum dynamical coherence. To obtain these off-diagonal elements, we observe that $Z \otimes I |\phi^\pm\rangle = |\phi^\mp\rangle$, $X \otimes I |\phi^\pm\rangle = |\psi^\pm\rangle$ and $Y \otimes I |\phi^\pm\rangle = i |\psi^\mp\rangle$. So the action of $Z \otimes I$ keeps a state in the $\phi-$subspace inside and the action of $X \otimes I$ or $Y \otimes I$ sends it entirely into the complement $\psi-$subspace. $Z \otimes I$ and $Y \otimes I$ also swap the Bell basis of each subspace. Input state in the $\phi-$subspace $|\phi_C\rangle = \frac{\alpha + \beta}{\sqrt{2}} |\phi^+\rangle + \frac{\alpha - \beta}{\sqrt{2}} |\phi^-\rangle$ with $|\alpha| \neq |\beta| \neq 0$ will have, in its output $\varepsilon \otimes I(|\phi_C\rangle \langle \phi_C|)$, the cross (coherence) term $\chi_{IZ} I \otimes I \left(C |\phi^+\rangle \langle \phi^-| + C^* |\phi^-\rangle \langle \phi^+| \right)Z \otimes I = \chi_{IZ} \left(C |\phi^+\rangle \langle \phi^+| + C^* |\phi^-\rangle \langle \phi^-| \right)$ and $\chi_{ZI} Z \otimes I \left(C |\phi^+\rangle \langle \phi^-| + C^* |\phi^-\rangle \langle \phi^+| \right)I \otimes I = \chi_{ZI} \left(C |\phi^-\rangle \langle \phi^-| + C^* |\phi^+\rangle \langle \phi^+| \right)$, where $C = \frac{1}{2} (\alpha + \beta)(\alpha - \beta)^*$. Then orthogonal commuting measurements of $P_{\pm \phi} = |\phi^+\rangle \langle \phi^+| \pm |\phi^-\rangle \langle \phi^-|$ within the $\phi-$subspace allow us to determine the real and imaginary parts of $\chi_{IZ} = \chi_{ZI}^*$. In the same way, orthogonal commuting measurements of $P_{\pm \psi} = |\psi^+\rangle \langle \psi^+| \pm |\psi^-\rangle \langle \psi^-|$ in the complement $\psi-$space on the same output allow us to determine $\chi_{XY} = \chi_{YX}^*$. The same mathematical structures can be constructed to obtain coherence $\chi_{IX}$ and $\chi_{YZ}$. In the standard Bell states, the $|0\rangle$ and $|1\rangle$ of the first qubit are $\pm 1$ eigenstates of $Z$. By replacing them with the corresponding $\pm 1$ eigenstates of $X$, we have $|\phi_X^\pm\rangle = \frac{1}{\sqrt{2}} (|+\rangle |0\rangle \pm |-\rangle |1\rangle)$ and $|\psi_X^\pm\rangle = \frac{1}{\sqrt{2}} (|0\rangle \pm |+\rangle |1\rangle)$. Then the new transformation relations by local Pauli operators are $X \otimes I |\phi_X^\pm\rangle = |\phi_X^\mp\rangle$, $Z \otimes I |\phi_X^\pm\rangle = |\psi_X^\pm\rangle$ and $Y \otimes I |\phi_X^\pm\rangle = -i |\psi_X^\mp\rangle$. Using these new Bell basis in $|\phi_C\rangle$ and $P_{\pm \phi, \psi}$ will determine $\chi_{IX}$ and $\chi_{YZ}$. Similarly, to obtain $\chi_{IY}$ and $\chi_{XZ}$, we utilize Bell basis $|\phi_Y^\pm\rangle = \frac{1}{\sqrt{2}} (|+i\rangle |0\rangle \pm |-i\rangle |1\rangle)$,
\[ |\psi^\pm_Y \rangle = \frac{1}{\sqrt{2}} (|i\rangle |0\rangle \pm |i\rangle |1\rangle) \] and their transformation relations \( Y \otimes I |\phi^\pm_Y \rangle = |\phi^\mp_Y \rangle \), \( Z \otimes I |\phi^\pm_Y \rangle = |\psi^\pm_Y \rangle \) and \( X \otimes I |\phi^\pm_Y \rangle = i |\psi^\mp_Y \rangle \).

It is worth noticing that this method can be generalized to \( d \)-level quantum systems, where \( d \) is a prime number. For a non-prime \( d \), the task can be accomplished by embedding it into a larger prime dimensional Hilbert space. The details of this generalization is published in [32].

DCQD has nice resource saving features. In both population and coherence determination, the only required measurement device is a Bell type state analyzer. When determining quantum dynamical populations, only a single preparation of state \( |\phi^+ \rangle \) is required as the input. Since a projective measurement in the Bell state basis are performed on \( \varepsilon \otimes I(|\phi^+ \rangle \langle \phi^+|) \), we are guaranteed to have one of the four Bell states after each measurement. Then applying an appropriate local Pauli operator to the output of each projective channel can transform the residual Bell state to \( |\phi^+ \rangle \) and thus recycle it as a new input state. In determining quantum dynamical coherence, we observe that \( P_+ + P_\psi = \hat{1} \). So by measuring \( \langle P_+ \rangle \), we have also measured \( \langle P_+ \rangle = 1 - \langle P_\psi \rangle \). Since \( P_+ = |\phi^+ \rangle \langle \phi^+| + |\phi^- \rangle \langle \phi^-| \) is the projector into the \( \phi \)-subspace, measurement result of 1 or 0 on each copy of \( \rho = \varepsilon \otimes I(|\phi_C \rangle \langle \phi_C|) \) means the final state has been cut down to its \( \phi \)- or \( \psi \)-subspace diagonal block, ie. \( \rho_{1,0} = Tr(P_{\pm \phi,\psi} \rho) P_{\pm \phi,\psi} \). Each scenario has a probability of \( Tr(P_{\pm \phi,\psi} \rho) \) to occur. We can then measure either \( P_{\pm \phi} \) on \( \rho_1 \) or \( P_{\pm \psi} \) on \( \rho_0 \). So the two measurements in each subspace can be done in sequence and measurements in the two subspaces are complimentary to each other. Since the four measurement operators \( P_{\pm \phi,\psi} \) all commute with each other and are equivalent to a projective measurement in Bell state type basis, they can all be incorporated into one experimental configuration. So in total, we need to prepare one copy of a maximally entangled Bell state and three other non-maximally entangled states to fully determine a most general 1-
qubit quantum operation. For each of these input states, we only need to set up one experimental configuration to simultaneously measure all the required commuting observables.

To scale DCQD to an arbitrary number of $N$ qubits, we must measure the tensor product of the observables in the 1-qubit case. For example, in determining the quantum dynamical coherence $\chi_{I} \otimes X \otimes Z \otimes Y$ of a 2-qubit general quantum operation, we measure $P_{\pm \phi} \otimes P_{\pm \psi}$. Only 4 experimental configurations are required to determine $\chi$ in the 1-qubit case. Therefore, the total number of required experimental configurations scales exponentially as $4^N = 2^{2N}$. This is a quadratic improvement over standard or Ancilla-assisted QPT. For example, in the 3 and 4 qubit case, DCQD reduces the total number of experimental configurations from $\sim 5 \times 10^3$ and $\sim 6.5 \times 10^4$ in standard QPT to 64 and 256. The experimental techniques required to accomplish DCQD are single qubit Pauli rotations, Bell state generation and measurements. These are common practices in quantum information science and have been studied extensively and implemented successfully in almost all physical system currently used in quantum information processing. So the method is fully realizable. For example, on systems with controllable two body interactions, e.g. liquid state NMR and trapped ions, 3 and 4 qubit experiments can be carried out. A detailed comparison of required resources for standard QPT, ancilla-assisted QPT and DCQD have been carried out in [33]. In terms of the overall experimental configurations and elementary quantum operations required, standard QPT is the most efficient when controllable two-body interactions are not available. For trapped ions and liquid NMR with controllable interactions, DCQD requires the least number of elementary quantum operations.

The method of DCQD was overlooked during the initial literature review of this thesis work. The method does not use the usual name of quantum process tomog-
raphy and the three papers cited here are the only ones on the subject before 2013. Further, there is a falsely perpetuated claim in some other literature that DCQD provides only partial characterization of the quantum process. This is probably due to such writings in the abstract of [31], even though it is fully capable to characterize a general quantum operation completely. So future works could be done to apply this method to unitary quantum processes.

1.3 Why Adaptive Unitary Process Tomography?

Despite the challenge of exponential scaling, it can be shown that an arbitrary unitary quantum operation of any dimension can be decomposed into the product of a set of 1-qubit and 2-qubit unitary operations. These sets of 1 and 2-qubit unitary gates are said to be universal [34] [15, Section 4.5] [35, and references therein]. A general non-unitary quantum operation can be embedded in a unitary operation of higher dimension because both its input and output states can be purified in a larger Hilbert space of the same size. Therefore, a universal set of 1 and 2-qubit unitary gates are the fundamental building blocks of quantum computation. For this reason, this thesis focuses on the complete characterization of 1 and 2-qubit unitary operations. We expect that the unitarity condition will significantly reduce the number of observables required to uniquely determine the gate.

Fundamentally, measurements in quantum mechanics are dictated by the Born’s rule. All observable quantities in reality are expressed in terms of the squared magnitude of wavefunction amplitudes. This has two immediate consequences for tomography:

- The global phase of quantum states and operations has no observable effects and thus can be chosen at will;
• All terms in the expression of measured observables are second order in terms of state or gate parameters.

Therefore, we are always free to choose one parameter of our state or gate to be real and positive to obtain it by taking the square root of some measurement result. Then the basic strategy in tomography is to measure cross terms between the parameter we seek and the parameter we already determined. However, if the parameter we determined happens to be 0 then we obtain no information about the parameter we seek. This is the case of measure zero.

So if we were to blindly follow a fixed tomography scheme non-adaptively, we must have extra measurements to cover all possible cases of measure zero to guarantee success in all possible cases of parameter values. Otherwise, we can use less measurements in exchange for some probability of failure. This dilemma is only avoidable in an adaptive scheme, where we are allowed to choose the next observable to measure based on the results of previous measurements. This is why we must use adaptive tomography scheme to guarantee success with less measurements. There has been relatively few works on adaptive quantum tomography. The existing ones are mostly focused on QST instead of QPT [36, 37]. Only Teo et al. has presented an incomplete adaptive QPT scheme [38].

In this thesis, we present an adaptive method to uniquely determine any $d$-dimensional pure state among all possible pure or mixed states with at most $2d - 2$ measurements. Using this adaptive QST method, we then provide a method to determine any $d$-dimensional unitary gate with at most $d^2 + d - 1$. 
1.4 Related Recent Works

We first review a few papers closely related to this work. For convenience, notations consistent with each original paper are used.

1.4.1 Algebraic Framework of QST

Heinosaari et al.\textsuperscript{[39]} established the linear algebraic framework that formulates the task of QST.

Let $\mathcal{H}$ denote the Hilbert space of interest and $\mathcal{L}(\mathcal{H})$ denote the space of all linear operators on $\mathcal{H}$. If $\mathcal{H}$ has dimension $d$, then $\mathcal{L}(\mathcal{H})$ has dimension $d^2$. For an operator $L \in \mathcal{L}(\mathcal{H})$, its operator norm $\|L\|$ is defined as the square root of the largest eigenvalue of the Hermitian operator $L^\dagger L$. $\|L\|$ is simply the largest eigenvalue of $L$ if its Hermitian. A positive semi-definite operator $\rho \in \mathcal{L}(\mathcal{H})$ with its trace equal to 1 is the density matrix of some quantum state. The set of all states on $\mathcal{H}$ is denoted as $\mathcal{S}(\mathcal{H})$. Clearly, $\mathcal{S}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$.

Define the Hermitian conjugate of the linear operator space as $\mathcal{L}^\dagger(\mathcal{H}) = \{L^\dagger, L \in \mathcal{L}(\mathcal{H})\}$. Clearly $\mathcal{L}^\dagger(\mathcal{H}) = \mathcal{L}(\mathcal{H})$ since $\mathcal{L}(\mathcal{H})$ includes all linear operators on $\mathcal{H}$. If $\mathcal{L}(\mathcal{H})$ is partitioned as $\mathcal{L}(\mathcal{H}) = \mathcal{O}(\mathcal{H}) \cup \mathcal{O}^\perp(\mathcal{H})$ such that the identity operator $\hat{I} \in \mathcal{O}(\mathcal{H})$ and $\mathcal{O}^\dagger(\mathcal{H}) = \mathcal{O}(\mathcal{H})$, then $\mathcal{O}^{\perp\dagger}(\mathcal{H}) = \mathcal{O}^{\perp}(\mathcal{H})$ also and all elements in $\mathcal{O}^\perp(\mathcal{H})$ are trace zero since $Tr(M) = Tr(M\hat{I})$. Clearly, $\dim(\mathcal{O}(\mathcal{H})) + \dim(\mathcal{O}(\mathcal{H})^{\perp}) = \dim(\mathcal{L}(\mathcal{H})) = d^2$.

Recall that a general measurement on a quantum mechanical system can be described by positive operator valued measurements (POVMs). A POVM measurement with $n$ possible outcomes consists of $n$ linearly independent positive semi-definite operators $\{A_i\}$ such that $\sum_{i=1}^n A_i = \hat{I}$. A single POVM measurement on a quantum state $\rho$ gives each outcome $A_i$ with probability $P_i = Tr(\rho A_i)$. Only $n - 1$ elements
of \{A_i\} can be chosen freely since the last element is always \(A_n = \hat{I} - \sum_{i=1}^{n-1} A_i\).

\(\mathcal{O}(\mathcal{H})\) is called an operator system because there always exists a set of POVM elements \{A_i\} such that \(\mathcal{O}(\mathcal{H}) = \{\sum_{i=1}^{n} c_i A_i \mid c_i \in \mathbb{C}\}\). Conversely, the span of any set of POVM elements is always an operator system.

This equivalence can be proved by construction. Any element \(X \in \mathcal{O}(\mathcal{H})\) can be expressed as a linear combination of two Hermitian operators

\[
X = \frac{1}{2}(X + X^\dagger) + i\frac{1}{2i}(X - X^\dagger).
\]

Since \(\mathcal{O}^\dagger(\mathcal{H}) = \mathcal{O}(\mathcal{H})\) and \(X \in \mathcal{O}(\mathcal{H}), X^\dagger \in \mathcal{O}(\mathcal{H})\). So both \(\frac{1}{2}(X + X^\dagger), \frac{1}{2i}(X - X^\dagger) \in \mathcal{O}(\mathcal{H})\). This also proves that any non-empty operator system contains Hermitian operators. Further more, any Hermitian operator \(X \in \mathcal{O}(\mathcal{H})\) can be expressed as a linear combination of two positive semi-definite operators

\[
X = \frac{1}{2}(\|X\| \hat{I} + X) - \frac{1}{2}(\|X\| \hat{I} - X).
\]

Since \(\hat{I} \in \mathcal{O}(\mathcal{H})\), both \(\|X\| \hat{I} + X, \|X\| \hat{I} - X \in \mathcal{O}(\mathcal{H})\). Therefore, a basis consisting of positive semi-definite operators and \(\hat{I}, \{E_1, E_2, ..., E_{n-1}, \hat{I}\}\) can fully span \(\mathcal{O}(\mathcal{H})\). Hence, a set of POVM elements \{A_i\} that fully spans \(\mathcal{O}(\mathcal{H})\) can be constructed from this basis as \(A_i = \frac{1}{(n-1)\|E_i\|} E_i\) and \(A_n = \hat{I} - \sum_{i=1}^{n-1} A_i\). Clearly, \(dim(\mathcal{O}(\mathcal{H})) = dim(\{A_i\})\). Therefore, any operator system \(\mathcal{O}(\mathcal{H})\) can be labeled by a spanning set of POVMs as \(\mathcal{O}(A)\).

It is worth to note that since \(\mathcal{O}^\perp(\mathcal{H}) = \mathcal{O}^\perp(\mathcal{H})\), the traceless complement subspace \(\mathcal{O}^\perp(\mathcal{H})\) is also spanned by its Hermitian elements.

Very often in QST, the expectation values of a set of \(n - 1\) Hermitian operators \(\{S_i\}\), instead of a POVM set, is measured. But a simple construction can convert
\{S_i\} to a POVM measurement \{A_i\} with \( n \) outcomes in the following way: \( \forall i = 1, 2, \ldots, n - 1, A_i = \frac{1}{2(n-1)} \left( \hat{I} + \|S_i\|^{-1} S_i \right) \) such that \( A_n = \hat{I} - \sum_{i=1}^{n-1} A_i \) is positive semi-definite. It is worth to note that since \( Tr(\hat{I} \rho) = Tr(\rho) = 1 \), the last element \( A_n \) gives no additional information about any measured state \( \rho \). Also, multiple sets of POVM measurements can be used to determine a quantum state. But the joint span of these multiple sets is an operator system and thus can be spanned by a single POVM set. Therefore, the discussion of measurements required for complete QST can be restricted to a single POVM measurement set without loss of generality.

Formally, let \( \mathcal{P} \in \mathcal{S}(\mathcal{H}) \) be a subset of quantum states. A POVM \( \{A_i\} \) is described as informationally complete with respect to \( \mathcal{P} \) if \( \forall \rho_1, \rho_2 \in \mathcal{P}, Tr(\{A_i\} \rho_1) = Tr(\{A_i\} \rho_2) \Rightarrow \rho_1 = \rho_2. \)

The important idea about QST is that for any two different \( \rho_1 \) and \( \rho_2 \) with \( Tr(\{A_i\} \rho_1) = Tr(\{A_i\} \rho_2), (\rho_1 - \rho_2) \in \mathcal{O}^\perp(A). \)

If \( \rho_1 - \rho_2 \neq 0 \), then measurement results of \( \{A_i\} \) can not distinguish between \( \rho_1 \) and \( \rho_2 \). Therefore, a POVM \( \{A_i\} \) is informationally complete with respect to \( \mathcal{P} \) if and only if \( (\mathcal{P} - \mathcal{P}) \cap \mathcal{O}^\perp(A) = \{0\} \).

A quick application of this idea shows that a POVM set \( \{A_i\} \) that is informationally complete for all states \( \mathcal{S}(\mathbb{C}^d) \) must have at least \( d^2 \) elements. Otherwise, let \( \mathcal{O}^\perp(A) \) be spanned by only one non-zero Hermitian operator \( X \). Then the two states \( \frac{1}{d} \hat{I} \) and \( \frac{1}{d} \hat{I} + \|X\|^{-1} X \) cannot be distinguished by \( \{A_i\} \).

In the same fashion, let there be a Hermitian operator \( T \in \mathcal{O}^\perp(A) \). If \( T \) is rank 1, it must be 0 since all elements in \( \mathcal{O}^\perp(A) \) are trace zero. If \( T \) is rank 2, its spectrum decomposition is \( T = \lambda |\psi_1\rangle \langle \psi_1| - \lambda |\psi_2\rangle \langle \psi_2| \). Then the difference between pure states \( |\psi_1\rangle \langle \psi_1| \) and \( |\psi_2\rangle \langle \psi_2| \) is \( T' = \frac{1}{\lambda} T \in \mathcal{O}^\perp(A). \) Therefore, \( |\psi_1\rangle \langle \psi_1| \) and \( |\psi_2\rangle \langle \psi_2| \) cannot be distinguished by \( \{A_i\} \). A POVM \( \{A_i\} \) is informationally complete with respect to the set of pure states if and only if every Hermitian operator \( T \in \mathcal{O}^\perp(A) \)
has rank greater than or equal to 3.

By a careful construction, they also showed that a POVM \( \{A_i\} \) with \( 4r(d - r) \) elements can be informationally complete for all unknown density matrices with rank \( r \) as long as \( r < d/2 \). So the number of measurements required to uniquely determine a unknown quantum state depends on the rank of the unknown state.

### 1.4.2 QST of Pure States with Pauli Measurements

Unfortunately, the POVM elements carefully designed by theorists with nice properties are often not implementable in real experiments. For example, in NMR systems commonly used for quantum information research, the only conveniently measurable expectation values are the tensor products of Pauli operators.

Any general \( n \)-qubit quantum state’s density matrix can be decomposed into Pauli tensor basis as \( \rho = \frac{1}{2^n} I + \sum_{i=1}^{2^n-1} c_i \sigma_i \), where \( c_i \in \mathbb{R} \). This decomposition has its trace equal to 1 since \( Tr \left[ \frac{1}{2^n} I \right] = \frac{1}{2^n} Tr \left[ I^{2n \times 2n} \right] = \frac{1}{2^n} \times 2^n = 1 \) and \( Tr(\sigma_i) = 0 \) for any Pauli tensor basis \( \sigma_i \). Due to the constraint of positive semi-definiteness, these real coefficients \( c_i \) are not free to vary independently from each other. Only some of them need to be determined by measuring the corresponding \( \sigma_i \) during the QST of \( \rho \).

Recently, Jackson et al.[40] has worked on this problem for case of 2 and 3 qubit states.

They first classified the extend to which a quantum state can be determined via QST. In particular, a pure state is Uniquely Determined among all Pure states (UDP) by the measurement results of a set of observables \( \{A_i\} \) if there does not exist any other pure state that can yield the same measurement results when measured against \( \{A_i\} \). If this set of measurement results cannot be obtained by measuring
\{A_i\} on any other pure or mixed quantum state, then the pure state is said to be
Uniquely Determined among All quantum states (UDA) by \{A_i\}.

Next, they utilized the idea from Heinosaari et al.\[39\] that all quantum states
that give the same measurement results of \{A_i\} can only differ from each other
in the complement subspace \{A_i\}$. However, if there exist a pure state $\rho$ that is
compatible with the measured results of \{A_i\} and \{A_i\} is chosen such that every
non-trivial $H \in \{A_i\}^\perp$ has at least two positive and two negative eigenvalues, then
any difference from $\rho$ in \{A_i\} will make the resulting matrix $\rho \pm H$ non-definite and
hence not a valid quantum state.

**Proof:** Let Hermitian operator $\hat{P}$ be the projector$^1$ into the subspace spanned
by eigen vectors of $\pm H$ with negative eigenvalues. If $\rho$ is a pure state a.k.a. rank
1, its projection into this subspace $\hat{P} \rho \hat{P}$, which is a submatrix of $\rho$, is at most rank
1. Let $\hat{P} \rho \hat{P} = a |\phi \rangle \langle \phi |$. Since this subspace is two dimensional, there always exists
some non-zero $|\tau \rangle = \hat{P} |\tau \rangle$, a.k.a. in the subspace, that is orthogonal to $|\phi \rangle$, ie.
$\langle \tau |\phi \rangle = \langle \phi |\tau \rangle = 0$. Then we have

$$\langle \tau |\rho \pm H |\tau \rangle = \langle \tau |\hat{P} \rho \hat{P} |\tau \rangle \pm \langle \tau |H |\tau \rangle = a \langle \tau |\phi \rangle \langle \phi |\tau \rangle \pm \langle \tau |H |\tau \rangle = 0 + \langle \tau |\pm H |\tau \rangle < 0.$$ 

So $\rho \pm H$ is non-definite. QED

So if the underlying unknown state $\rho$ is pure, there will be no other pure or mixed
valid quantum state that is compatible with the results of measuring \{A_i\} on $\rho$. In
other words, any pure state $\rho$ is UDA by measuring \{A_i\}.

The fact that \{A_i\} is UDA only for pure states instead of all quantum states

$^1$If $\hat{I} = |0 \rangle \langle 0 | + |1 \rangle \langle 1 | + |2 \rangle \langle 2 |$, then projectors are operators like $\hat{P}_0 = |0 \rangle \langle 0 |$, $\hat{P}_{12} = |1 \rangle \langle 1 | + |2 \rangle \langle 2 |$. They are identities of some subspace. $\hat{P} = \hat{P}^\dagger$, $\hat{P}^2 = \hat{P}$.  

14
implies an important fact: quantum states of different ranks differ not only in the values of the same set of parameters but also by the minimal number of parameters required to fully describe them. The hope of non-adaptively measuring a universally minimal set of parameters that determines all quantum states is false. Although all quantum states share the property of being trace-1 and positive semi-definite, fixing \{A_i\} to values compatible with mixed states could leave some degrees of freedom to vary \{A_i\}⊥ coefficients without breaking these constraints. Whether a QST method determines a state uniquely not only depends on what measurements are made, but also depends on what values are obtained for those measurements.\(^2\) Therefore, a QST method that uses the minimal number of measurements required for each underlying quantum state must be adaptive: every measurement result must be consulted immediately after the measurement to decide if an additional measurement is needed to fully determine the measured state.

Conversely, if every pure state \(\rho\) is UDA by measuring \{A_i\}, then every non-trivial \(H \in \{A_i\}^\perp\) has at least two positive and two negative eigenvalues. Otherwise, \(\pm H\) will have at most one negative eigenvalue \(n\). Let the eigenstate of \(\pm H\) associated with this negative eigenvalue \(n\) be \(|n\rangle \langle n|\). Clearly, \(\rho' = |n\rangle \langle n| + \frac{1}{|n|} H\) is trace-1 and positive semi-definite because \(H\) is trace-0.\(^3\) Since \(\frac{1}{|n|} H \in \{A_i\}^\perp\), \(|n\rangle \langle n|\) and \(\rho'\) are indistinguishable by measurement results of \{A_i\}.

Therefore, every pure state \(\rho\) is UDA by measuring \{A_i\} \iff Every non-trivial \(H \in \{A_i\}^\perp\) has at least two positive and two negative eigenvalues.

\(^2\)Whereas if we know the rank or some other information that gives the exact parametrization of the state, we only need to study how to perform measurements to fix these parameters without considering if the measured values of these parameters will require more parameters to be added to fully describe the state.

\(^3\)Any Hermitian matrix \(M\) with a non-zero trace has a component of \(\hat{I}\) with coefficient \(Tr(M) = Tr(M\hat{I})\). But this component gives no useful information about any density matrix \(\rho\) because we already know that its contribution to the final measurement outcome is always \(Tr(\rho Tr(M)\hat{I}) = Tr(M)\). So \(\{A_i\} \cup \{A_i\}^\dagger\) only needs to span the trace zero subspace of Hermitian operators.
Their work applied this result to the case of Pauli operators. For every element in the span of a set of 2-qubit Pauli operators to have two positive and two negative eigenvalues, there cannot exist any commuting pair of Pauli operators in the set. Otherwise, since the commuting Pauli operators share a common set of eigen basis, linear combinations of them with real coefficients can have an eigen value of 0. Since two qubit Hermitian operators have only four dimensions, such a linear combination does not have at least two positive and two negative eigenvalues.

To minimize the number of Pauli measurements required to UDA any pure state, they searched for the largest set of mutually non-commuting 2-qubit Pauli operators and found that such a set contains five elements. \{XX, XY, XZ, YI, ZI\} is an example of such a set. It is then shown by direct computation that the determinant of any non-trivial real linear combination of this five elements is positive. The determinant of a 4-dimensional traceless Hermitian matrix is positive if and only if it has exactly two positive and two negative eigenvalues. Therefore, \{XX, XY, XZ, YI, ZI\} is indeed the required \{A_i\}_\perp. Excluding the identity operator, its complement set with ten Pauli operators \{A_i\} = \{XI, YX, YY, YZ, ZX, ZY, ZZ, IX, IY, IZ\} is the UDA set for all pure states.

Note that this result does not say that the space of pure state density matrices is completely spanned by \{A_i\} and \(T[\psi] \langle \psi | \{A_i\}_\perp] \equiv 0 \forall |\psi\rangle\). The result says that if there exist a pure state that is compatible with the ten measurement results of \{A_i\}, then its difference with any other compatible valid quantum state, pure or mixed, must be zero. Hence it is the unique compatible quantum state. This in turn also fixed its coefficients in \{A_i\}_\perp, which are not necessarily all zeros. For example, \(ZI \in \{A_i\}_\perp\). But clearly, |00⟩⟨00| = 1/4 (I + IZ + ZI + ZZ), ie. |00⟩⟨00| has a coefficient of 1/4 for the ZI component. A general 2-qubit density matrix, pure
or mixed, can have a non-zero coefficient of any 2-qubit Pauli operator.\textsuperscript{4}

### 1.4.3 QPT for Unitary Gates

To determine an unknown quantum operation by QPT, not only do we need to know how to perform QST on the outputs from the interested quantum operation, it is also essential to know which set of input states produces outputs that capture the action of the quantum operation completely. Reich et al.[41] has shown that the outputs of a set of input states \( \{\rho_j^{in}\} \) is sufficient to differentiate any two unitary maps if and only if identity is the only operator that commutes with all \( \rho_j^{in} \)'s in the set. This kind of input state sets are call Unitarily Informationally Complete (UIC). Moreover, they have also shown that these outputs \( \{\hat{U}\rho_j^{in}\hat{U}^\dagger\} \) completely distinguishes \( \hat{U} \) from any other general quantum operations, a.k.a. CPTP maps, if and only if \( \{\rho_j^{in}\} \) is a UIC set.

Recently, Baldwin et al. has worked on the tomography of unitary quantum gates[42]. Their work presented two ways to tomograph a unitary gate.

They recognize that the action of a unitary gate is to perform a change of basis in which a quantum state is expressed. Therefore, a unitary gate acting on a \( d \)--level quantum system can be expressed as

\[
\hat{U} = \sum_{n=0}^{d-1} |u_n\rangle \langle n| ,
\]

where \( \{|u_n\rangle\} \) and \( \{|n\rangle\} \) are two complete sets of orthonormal basis of the Hilbert

\textsuperscript{4}Suppose \( \rho \) is a quantum state, then \( \rho' = \hat{U}\rho\hat{U}^\dagger \) is also a valid quantum state. It can be easily proved that for any Pauli operator there exist some \( \hat{U} \) that rotates it to any other Pauli operator. So we can pick some \( \rho \) with at least one non-zero coefficient for some Pauli operator, then we can make this coefficient belong to any Pauli operator in \( \rho' \) by choosing the right \( \hat{U} \). So no Pauli operator is guaranteed to have zero coefficient in all quantum states. In the same way, no Pauli operator, except identity, is guaranteed to have non-zero coefficient in all quantum states.
The task to tomograph a unitary gate is essentially to determine \(|u_n\rangle\) in relation to \(|n\rangle\) and the relative phases of each \(|u_n\rangle \langle n|\) term in \(\hat{U}\).

They first considered the UIC set of probe states with the minimal number of elements provided by Reich et al. in [41]:

\[
\mathcal{S} = \left\{ \rho_0^{in} = \sum_{n=0}^{d-1} \lambda_n |n\rangle \langle n|, \quad \rho_1^{in} = \frac{1}{d} \sum_{n,m=0}^{d-1} |n\rangle \langle m| \right\},
\]

(1.4.2)

where the eigenvalues \(\{\lambda_n\}\) must be non-degenerate. Note that \(\rho_1^{in}\) is a pure state. So \(\rho_1^{out} = \hat{U} \rho_1^{in} \hat{U}^\dagger\) will also be a pure state.

Then probing \(\hat{U}\) with \(\mathcal{S}\) gives

\[
\rho_0^{out} = \hat{U} \rho_0^{in} \hat{U}^\dagger = \sum_{n=0}^{d-1} \lambda_n |u_n\rangle \langle u_n|,
\]

(1.4.3)

\[
\rho_1^{out} = \hat{U} \rho_1^{in} \hat{U}^\dagger = \frac{1}{d} \sum_{n,m=0}^{d-1} |u_n\rangle \langle u_m|.
\]

(1.4.4)

They claim that performing complete quantum state tomography on \(\rho_0^{out}\) generally requires measuring a POVM with at least \(d^2 - 1\) elements. It is easy to count that a \(d \times d\) Hermitian matrix has \(d^2\) independent real parameters and a density matrix has one less due to the trace-1 constraint.

Diagonalizing the density matrix obtained from quantum state tomography gives a set of eigenstates \(\{|u'_n\rangle\}\) corresponding to the set of non-degenerate eigenvalues \(\{\lambda_n\}\). The importance of non-degeneracy in \(\{\lambda_n\}\) is the following: if \(\{\lambda_n\}\) is degenerate, then there is no way to determine which choice of basis in the degenerate subspace corresponds to elements of \(\{|u_n\rangle\}\).

At this point, the method has not yet established that \(\{|u'_n\rangle\} = \{|u_n\rangle\}\). This is
because when diagonalizing $\rho^\text{out}_0$, solving the eigenvalue problem does not constrain the overall phase on an eigenvector: $\rho^\text{out}_0 |u_n\rangle = \lambda_n |u_n\rangle \Rightarrow \rho^\text{out}_0 e^{i\theta} |u_n\rangle = \lambda_n e^{i\theta} |u_n\rangle$.

So in the process of choosing the representation of $|u_n\rangle$, an overall phase $e^{i\theta}$ might have been imposed without notice to get the result $|u'_n\rangle = e^{i\theta} |u_n\rangle$ instead of $|u_n\rangle$.

This is a subtle point for physicists because the overall phase of eigenstates are usually ignored since they have no directly observable physical consequences. However, it does matter in characterizing $\hat{U}$ because, for example $\hat{U} = |+\rangle \langle 0| + |-\rangle \langle 1| \neq \hat{U}' = |+\rangle \langle 0| + i |-\rangle \langle 1|$, the relative phase is part of the gate’s action on the Hilbert space.

They chose the global phase of $\hat{U}$ such that $|u'_0\rangle = |u_0\rangle$. But all overall phases of the other $|u_n\rangle$’s relative to this $|u_0\rangle$ have to come from $\rho^\text{out}_1$.

The fact that output $\rho^\text{out}_1$ is a pure state reduces the number of measurements required to fully specify it. According to [43], all $d$-dimensional pure states, except those whose first diagonal element is zero, can be fully specified by a minimal number of $2d$ POVM elements as the following:

$$E_0 = a |0\rangle \langle 0|,$$

$$E_n = b(1 + |0\rangle \langle n| + |n\rangle \langle 0|), \quad n = 1, \ldots, d - 1,$$

$$\tilde{E}_n = b[1 + i(|0\rangle \langle n| - |n\rangle \langle 0|)], \quad n = 1, \ldots, d - 1,$$

$$E_{2d} = 1 - \left[E_0 + \sum_{n=1}^{d-1} (E_n + \tilde{E}_n)\right] \quad (1.4.5)$$

with $a$ and $b$ chosen such that $E_{2d}$ is positive semi-definite.

So $\rho^\text{out}_1$ is fixed by measuring these $2d$ POVM elements. Then each $|u_n\rangle$ can be computed using $\rho^\text{out}_1$ and the corresponding $|u'_n\rangle\langle u'_n|$ from the diagonalization of $\rho^\text{out}_0$:

$$|u_n\rangle = d |u_n\rangle \langle u_n| \rho^\text{out}_1 |u_0\rangle = d |u'_n\rangle \langle u'_n| \rho^\text{out}_1 |u_0\rangle.$$  \hspace{1cm} (1.4.6)
\[ |u_n\rangle\langle u_n| = |u'_{n}\rangle\langle u'_{n}| \] because the phase difference cancels out in ket-bra multiplication. Using (1.4.6) and the \[ |u_0\rangle \] chosen in diagonalizing \[ \rho_{0}^{\text{out}} \], the overall phases of the other \[ |u_n\rangle \]’s relative to \[ |u_0\rangle \] contained in \[ \rho_{1}^{\text{out}} \] can be extracted. Upon determining all \[ |u_n\rangle \]’s along with their relative phases, \( \hat{U} \) is fully specified. The above protocol requires the measurement of \( d^2 - 1 + 2d \) POVM elements. Because the input states in \( \mathcal{S} \) form a UIC set, this method uniquely identifies \( \hat{U} \) among all possible CPTP maps.

There has not been any experimental method to generate desired mixed states like \( \rho_{0}^{\text{in}} \) with reasonable precision. But any pure state can be reliably produced at will using quantum control techniques[44]. Therefore, experimentalists require a method with only pure input states. A minimal pure UIC set consists of \( d \) pure states that form a non-orthogonal yet complete basis for the Hilbert space. Baldwin et al. proposed the following method:

Use the set of \( d \) pure input states

\[ |\psi_0\rangle = |0\rangle, \]

\[ |\psi_n\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle + |n\rangle \right) \quad n = 1, 2, \ldots, d - 1. \tag{1.4.7} \]

It can be checked that identity is the only operator that commutes with the density matrices of these \( d \) states. Therefore, this is a minimal set of pure states that is UIC.

Strictly speaking, pure states should also be represented by their density matrices since there is no way to experimentally control their overall phase implied in the ket representation.

The outputs of these input states from \( \hat{U} \) are

\[ \hat{U} |\psi_0\rangle \langle \psi_0| \hat{U}^\dagger = |u_0\rangle \langle u_0|, \]
\[ \hat{U} |\psi_n\rangle \langle \psi_n| \hat{U}^\dagger = \frac{1}{2} (|u_0\rangle + |u_n\rangle) \left( \langle u_0| + \langle u_n| \right), \quad n = 1, 2, \ldots, d - 1. \quad (1.4.8) \]

Since these output states are also pure, their density matrices can be uniquely identified by the minimal set of \(2d\) POVMs from (1.4.5). So the total number of POVM elements required is \(d \times 2d = 2d^2\). Since these input states form a UIC set, if their corresponding outputs are determined to be \(\hat{U} |\psi_n\rangle \langle \psi_n| \hat{U}^\dagger\) with the same \(n\), then the underlying map is uniquely identified as \(\hat{U}\) among all CPTP maps. In reconstructing \(\hat{U} = \sum_{n=0}^{d-1} |u_n\rangle \langle n|\), the overall phase of \(|u_0\rangle\) can be set to 0 as before. The other \(|u_n\rangle\)'s along with their phases relative to \(|u_0\rangle\) can be determined as \(|u_n\rangle = \left(2\hat{U} |\psi_n\rangle \langle \psi_n| \hat{U}^\dagger - \hat{I}\right) |u_0\rangle\).

If given the prior information that the unknown map is unitary, the orthogonality of \(\{|u_n\rangle\}\) can be exploited to reduce the number of POVMs used in the above method.

First input \(|\psi_0\rangle \langle \psi_0|\) and tomograph the output \(|u_0\rangle \langle u_0|\) completely as before using the \(2d\) POVMs from (1.4.5). Set the overall phase of \(|u_0\rangle\) to 0. Writing \(|u_0\rangle = \sum_{n=0}^{d-1} c_{0n} |n\rangle\), each \(c_{0n}\) is related to the corresponding measurement result \(p_{0n} = \langle E_n \rangle = Tr(E_n |u_0\rangle \langle u_0|) = \langle u_0| E_n |u_0\rangle\).

In particular,
\[ p_{00} = \langle u_0| E_0 |u_0\rangle = a |c_{00}|^2. \quad (1.4.9) \]

By the freedom to choose the global phase of \(\hat{U}\), they set
\[ c_{00} = \sqrt{\frac{p_{00}}{a}}. \quad (1.4.10) \]

For the other POVM elements,
\[ p_{0n} = \langle u_0| E_n |u_0\rangle = b [1 + 2c_{00} \text{Re}(c_{0n})], \]
\[ \tilde{p}_{0n} = \langle u_0 | \tilde{E}_n | u_0 \rangle = b[1 + 2c_{00} Im(c_{0n})]. \quad (1.4.11) \]

So each \( c_{0n} \) is uniquely determined by two linear constraints. This completes the QST of \( |u_0\rangle \) with the global phase of \( \hat{U} \) chosen such that \( c_{00} \) is positive. Note that this strategy will fail if \( c_{00} = 0 \).

Next they input \( |\psi_1\rangle \langle \psi_1| \) and tomograph its output \( \frac{1}{2} (|u_0\rangle + |u_1\rangle) (\langle u_0| + \langle u_1|) \). Upon determining the \( d - 1 \) coefficients \( c_{1n} = \langle n|u_1\rangle \) with \( n = 0, 1, \ldots, d - 2 \), the last coefficient \( c_{1(d-1)} \) is fixed by the orthogonality constraint \( \langle u_0|u_1\rangle = \langle u_1|u_0\rangle = 0 \). Note that this is in fact two constraints: both the real and imaginary parts must be zero. So they only measure \( 2(d - 1) \) POVM elements to determine \( |u_1\rangle \). They still use the POVMs from (1.4.5), but only with \( n = 0, 1, \ldots, d - 2 \).

Similarly, \( |u_k\rangle \) is constrained by its orthogonality to \( k \) previously determined states \( \{ |u_m\rangle, m = 0, 1, \ldots, k - 1 \} \) and therefore only requires the measurement of \( 2(d - k) \) POVM elements.

In summary, the total number of POVM elements measured in this method is \( \sum_{k=0}^{d-1} 2d - 2k = d^2 + d \). They claim that because they are using a minimal pure UIC set of probe states and a minimal set of POVM elements to measure each output, this result is, by construction, the minimal number of POVMs that must be measured to fully determine a unitary map among all unitary maps.

The obvious drawback of their approach is the cases of failure if the observable \( E_0 \) was measured to an average value of zero for some state. They have emphasized this flaw in their paper. However, there is another important subtle point that they did not address. Because the global phase of \( \hat{U} \) has already been chosen by setting \( c_{00} = \sqrt{\frac{\rho_{00}}{a}} \) in (1.4.10), this choice must be kept consistently when determining the other outputs from \( \hat{U} \). For example, the 1st element of \( \hat{U} |\psi_1\rangle \langle \psi_1| \hat{U}^\dagger \) cannot
be determined as $c_{10} = \sqrt{\frac{1}{a}} \langle \psi_1 | \hat{U}^\dagger E_0 \hat{U} | \psi_1 \rangle$ because setting $c_{00}$ or $c_{10}$ to positive are generally two different choices of global phase. The correct way to determine the other $|u_n\rangle$ kets and their relative phase to $|u_0\rangle$ is to use the their orthogonality to $|u_0\rangle$ as $\langle u_0 | \hat{U} | \psi_n \rangle \langle \psi_n | \hat{U}^\dagger | u_0 \rangle = \frac{1}{2}$. Along with the other measured $2(d-1)$ POVM values, $\hat{U} | \psi_n \rangle \langle \psi_n | \hat{U}^\dagger$ is determined completely. The $|u_n\rangle$ kets and their phase relative to $|u_0\rangle$ can then be determined as $|u_n\rangle = \left( 2\hat{U} | \psi_n \rangle \langle \psi_n | \hat{U}^\dagger - I \right) |u_0\rangle$.

### 1.4.4 Summary of Important Results

Apart from the three references mentioned in the previous sections, there are a number of other important recent results related to our work. We have summarized all relevant results in Table 1.1.

The column abbreviations stand for:

- **UDP**: Uniquely Determine *a pure state* among all **Pure** states;
- **UDA**: Uniquely Determine *a pure state* among **All** states;
- **UPT**: Unitary Quantum Process **Tomography**.

<table>
<thead>
<tr>
<th></th>
<th>UDP</th>
<th>UDA</th>
<th>UPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (Non-adaptive)</td>
<td>$4d - 5$ [39]</td>
<td>$5d - 7$ [46]</td>
<td>$4d^2 - 2d - 4$ [48]</td>
</tr>
<tr>
<td>All (Adaptive)</td>
<td>$2d - 1$</td>
<td>$2d - 1$</td>
<td>$d^2 + d - 1$</td>
</tr>
</tbody>
</table>

Table 1.1: Comparison between recent and our results on tomography

The row entries are:

- **All (Non-adaptive)**: All pure states or all unitary gates can be determined non-adaptively;
• Generic (Non-adaptive): Only generic pure states or unitary gates can be determined non-adaptively. The method fails for some pure states and unitary gates that measure to zero for some observables in the method. An example would be the case when the $E_0$ measurement is zero in [42];

• All (Adaptive): All pure states or all unitary gates can be determined adaptively.

The result of $4d - 5$ in [39] comes from their more general result: distinguishing a subset of $d$-dimensional quantum states with maximum rank $r$, where $r < d/2$, requires measuring a minimum of $4r(d-r) - 1$ observables. In the case of pure states, i.e. $r = 1$, this reduces to $4d - 5$. Their work proved the general result by construction motivated by [49]. The proof constructed a $(d-2r)^2$-dimensional space of linear operators that is complementary to the span of the POVM elements of the measurement. So there are only $d^2 - (d-2r)^2 - 1 = 4r(d-r) - 1$ independent observables to be measured in the POVM. Further, the construction has every element in the operator space to have rank no less than $2r+1$. Since this is the complement space of the POVM elements, any two states $\rho_1$ and $\rho_2$ that give the same measurement results must only differ in this complement space. However, if all states are restricted to have maximum rank $r$, then this difference has $\text{rank}(\rho_1 - \rho_2) \leq 2r$ instead of $\text{rank}(\rho_1 - \rho_2) \geq 2r + 1$. So the only possibility is that $\rho_1 = \rho_2$, i.e. the state is uniquely determined by the POVM measurements.

The result of $5d - 7$ in [46] follows the same methodology: construct the subspace of linear operators complementary to the span of the measurement set such that any modification from this subspace to a state compatible with the measurement set does not satisfy some criteria. Differently from the UDP case in [39] with bounded rank criteria, we cannot allow a modification from this subspace to achieve any valid
quantum state for the measurement set to be UDA for all pure states. Therefore, we must relax the criteria to the most general criteria of a quantum state, i.e. positive semi-definiteness. Their work uses the fact that if every element of this constructed complement subspace has at least two positive and two negative eigenvalues, then adding or subtracting any element from this complement subspace to a pure state, which has a single positive rank, results in a matrix with at least one negative rank and therefore non-definite, i.e. not a valid quantum state. As in [39], their construction of such a complement subspace uses technique similar to the ones in [49]. They were able to construct such a subspace with \( d^2 - 5d + 6 \) linearly independent Hermitian operators. Therefore, the POVM measurement of this complement subspace has \( d^2 - (d^2 - 5d + 6) - 1 = 5d - 7 \) independent observables.

Following [39] and [46], the work of [48] extended this method of constructing the measurements’ complementary subspace with certain rank and eigenvalue restrictions to the case of quantum process tomography. Such an extension is possible because of the channel state duality: every quantum operation \( \Phi \), i.e. a Completely Positive Trace Preserving (CPTP) linear map, acting on a \( d \)-dimensional quantum system uniquely corresponds to a quantum state \( J(\Phi) \), i.e. a trace-1 positive semi-definite matrix, in the \( d^2 \)-dimensional Hilbert space. This correspondence is also called the Choi-Jamiołkowski isomorphism and \( J(\Phi) \) is called the Choi matrix of \( \Phi \). Therefore, the process tomography of quantum gate \( \Phi \) can be converted into the state tomography of quantum state \( J(\Phi) \) in a larger Hilbert space so that the subspace construction method in [39] and [46] can be used. Apart from specifying observables to be measured as in the case of state tomography, process tomography also needs to specify the input states of \( \Phi \). So the equivalent observables \( H \) measured against the Choi matrix state \( J(\Phi) \) are called interactive observables to distinguish them from regular observables in state tomography. It can be proved by construction
[50, 51] that a joint state $\xi$ of two $d$-dimensional systems and a binary measurement $\{P_{\pm}\}$ can be extracted from every interactive observable $H$ such that the output state $\rho \equiv (\Phi \otimes I)(\xi)$ gives $\text{Tr}(P_+ \rho) - \text{Tr}(P_- \rho) = \text{Tr}(HJ(\Phi))$, just like we would expect for the expectation value of a regular observable $H$ on a regular quantum state $J(\Phi)$. For a quantum operation $\Phi$ with $q$ Kraus operators, its Choi matrix $J(\Phi)$ has rank $q$. By constructing a subspace with minimal rank $2q + 1$ complement to the span of a set of interactive observables $\vec{H} = \{H_1, \ldots, H_m\}$, they determined that a minimum of $4d^2 - 2d - 4$ measurements are required to uniquely determine any unitary gate among all other unitary gates, which correspond to the special case of $q = 1$. It is worth to point out that their work is a completely theoretical approach that disregards practical implementations in the laboratory. It is unclear how these interactive observables can be easily realized on experimental platforms commonly used for quantum information research.

The 2nd row of Table 1.1 summarizes the most recent results for generic state and unitary gate tomography. As shown, comparing to the 1st row, there is a significant reduction of required observables if some rate of failure is allowed. The result of [45] is trivial following the method to obtain the result of [42], as discussed in the previous section. The result of [47] is also proved in this thesis, almost simultaneously with their publication.

The 3rd row of Table 1.1 summarizes the results of this thesis work. By using adaptive measurements, the results in the 2nd row can now be extended to all states and unitary gates instead of generic ones.
Chapter 2

Basic Concepts

This section outlines the very basic knowledge required in quantum mechanics and process tomography for this thesis. Some of the contents merely serve the purpose of completing the thesis as a self-contained document. Any reader who finds these contents trivial can safely move on to the next and refer back to them as needed.

2.1 Dirac Notation

In quantum mechanics, the state of a system can exist in a superposition of many mutually exclusive classical states. All these possible classical states can be represented by a complete set of orthonormal basis. Then the quantum state of the system is a linear combinations of all these basis, each with a complex coefficient called the probability amplitude of the respective classical state. Therefore, the system’s quantum state can be expressed as a column vector with complex elements. In the commonly used Dirac notation, this column vector is represented by a “ket” \(|\psi\rangle\). Its associated row vector obtained by taking the transpose and complex conjugate is called a “bra” \(\langle\psi|\). By normal matrix multiplication, ket bra \(|\psi\rangle\langle\psi|\) is a matrix and
bra ket \( \langle \psi | \psi \rangle \) is a number. The requirement \( \langle \psi | \psi \rangle = 1 \) is called **normalization**. This can always be achieved by a scalar multiple on \( |\psi\rangle \) without affecting its internal configurations. We will always use normalized vectors to represent quantum states since it puts different states on the same metric for comparison.

### 2.2 Born’s Rule and Projective Measurements

The famous Born’s rule of quantum mechanics states that the probability of projecting a state \( |\psi\rangle \) into another one \( |n\rangle \) is

\[
P_n = | \langle \psi | n \rangle |^2 = \langle \psi | n \rangle \langle n | \psi \rangle = \langle n | \psi \rangle \langle \psi | n \rangle.
\]

If we have a **complete set of orthonormal basis**

\[
\left\{ |n\rangle \mid \langle m | n \rangle = \delta_{mn}, \sum_{n=1}^{d} |n\rangle \langle n| = \hat{I} \right\}
\]

in \( d \)-dimensional Hilbert space, then \( |\psi\rangle \) is guaranteed to be projected into one of the \( d \) basis. Because \( \sum_{n=1}^{d} P_n = \sum_{n=1}^{d} | \langle \psi | n \rangle |^2 = \sum_{n=1}^{d} \langle \psi | n \rangle \langle n | \psi \rangle \)

\[
= \langle \psi | \left( \sum_{n=1}^{d} | n \rangle \langle n| \right) | \psi \rangle = \langle \psi | \hat{I} | \psi \rangle = \langle \psi | \psi \rangle = 1.
\]

Each time a system in state \( |\psi\rangle \) will be projected into a particular \( |n\rangle \). But in the statistical limit of measuring a large number of identical systems all prepared in state \( |\psi\rangle \), the count number distribution over all \( |n\rangle \)'s will give us each probability \( P_n \). Measuring a state’s projections against a complete set of orthonormal basis is called a projective measurement.
2.3 Hermitian Operators and Expectation Values

The **Hermitian conjugate** of linear operator \( \hat{A} \) is denoted as \( \hat{A}^\dagger \). We obtain \( \hat{A}^\dagger \) by transposing \( \hat{A} \) and taking the complex conjugate of all the elements. \( \hat{A} \) is a Hermitian operator \( \Leftrightarrow \hat{A}^\dagger = \hat{A} \).

In quantum mechanics, measurable physical quantities are called observables and represented by Hermitian operators. Every Hermitian operator is diagonalizable and has a **spectrum representation**

\[
\hat{A} = \sum_{n=1}^{d} a_n \left| n \right\rangle \left\langle n \right|
\]

where the eigenvalues \( \{a_n\} \) are real numbers and the eigenvectors \( \{|n\rangle\} \) form a complete set of orthonormal basis for the Hilbert space.

When measuring observable \( \hat{A} \) on a system in state \( |\psi\rangle \), the only possible results are \( \hat{A} \)'s eigenvalues \( \{a_n\} \). The exact result cannot be predicted with certainty. But each \( a_n \) has a projective probability of \( P_n = |\langle \psi | n \rangle|^2 \) to be measured.\(^5\) Upon revealing this result, the original state \( |\psi\rangle \) has been projected, or “collapse”, into the corresponding eigenstate \( |n\rangle \)\(^6\).

It is easy to see that

\[
\langle \psi | \hat{A} | \psi \rangle = \sum_{n=1}^{d} a_n \langle \psi | n \rangle \langle n | \psi \rangle = \sum_{n=1}^{d} a_n P_n
\]

\(^5\)As shown before, \( \sum_{n=1}^{d} P_n = 1 \). So \( \{a_n\} \) includes all possible results of the measurement.

\(^6\)This is only correct when eigenvalue \( a_n \) is non-degenerate. Generally, there could be many eigenvectors under the same eigenvalue, so \( \hat{A} = \sum_n a_n \sum_{g=1}^{g_n} |n,g\rangle \langle n,g| = \sum_n a_n \hat{P}_n \), where \( \hat{P}_n \) is the projector into the eigen subspace of \( a_n \). The act of measurement reveals one particular eigenvalue and thus only destroys the coherence outside the eigen subspaces. So the wavefunction only “collapse” into the associated eigen subspace, i.e. \( |\psi_{after}\rangle = \hat{P}_n |\psi\rangle / \langle \psi | \hat{P}_n |\psi\rangle \). The coherence within the eigen subspace is still preserved.

In the case of this thesis work, degeneracies are uncommon. So we will carry on with the non-degenerate case.
is the average value of observable $\hat{A}$. Physically, this is the mean value we get when measuring the physical quantity represented by $\hat{A}$ on a large number of identical systems that are all in state $|\psi\rangle$. This is called the expectation value of $\hat{A}$ in state $|\psi\rangle$.

Projective measurement and expectation value measurement are the two most common types of measurements in quantum mechanics. We also use these two types in our work.

There is a more general way of representing expectation values of observables. By inserting the identity operator $\hat{I} = \sum_{n=1}^{d} |n\rangle \langle n|$, we have

$$\langle \psi| \hat{A} |\psi\rangle = \langle \psi| \hat{I} \hat{A} |\psi\rangle = \sum_{n=1}^{d} \langle \psi|n\rangle \langle n| \hat{A} |\psi\rangle.$$ 

$\langle \psi|n\rangle$ and $\langle n| \hat{A} |\psi\rangle$ can be swapped because they are just complex numbers. Then

$$\langle \psi| \hat{A} |\psi\rangle = \sum_{n=1}^{d} \langle n| \hat{A} |\psi\rangle \langle \psi|n\rangle = Tr(\hat{A} |\psi\rangle \langle \psi|).$$

It is conventional to label $\hat{\rho} = |\psi\rangle \langle \psi|$, $\hat{\rho}$ is called the density matrix of the state. So we have

$$\langle \psi| \hat{A} |\psi\rangle = Tr(\hat{A} \hat{\rho}) = Tr(\hat{\rho} \hat{A}).$$

The expectation value of $\hat{A}$ is frequently denoted simply as $\langle \hat{A} \rangle_\psi$ or simply $\langle \hat{A} \rangle$ for short.

### 2.4 Mixed States and Density Matrices

$Tr(\hat{A} \hat{\rho})$ is more general than $\langle \psi| \hat{A} |\psi\rangle$ because $\hat{\rho}$ is a more general way to represent quantum states than $|\psi\rangle$. Only pure states can be represented as ket $|\psi\rangle$. Sup-
pose that due to our lack of knowledge there are several possible pure states \(\{|\psi_i\rangle\}\) that could be the true state of our system, each with an epistemic probability of \(\{p_i| p_i \geq 0, \sum_i p_i = 1\}\). Then to the best of our knowledge, we can only express the state of the system as a density matrix

\[
\hat{\rho} = \sum_i p_i |\psi_i\rangle \langle \psi_i|.
\]

In another common scenario, we have a large collection of identical systems (for example, 1 mole of \(CO_2\) molecules) that has a certain percentage \(\{p_i\}\) in each pure state \(\{|\psi_i\rangle\}\). Then the state of the collection must also be expressed as \(\hat{\rho}\). Expectation values measured from such a collection is called the ensemble average.

In both scenarios \(\hat{\rho}\) is called a mixed state. Density matrices are the most general way of representing quantum states. In fact, pure states are simply special cases of mixed states when \(\hat{\rho}\) is rank 1. So the properties of density matrices are of great importance to us in studying quantum states, in particular, QST.

- \(Tr(\hat{\rho}) = 1;\)

\[
Tr(\hat{\rho}) = \sum_{n=1}^{d} \langle n| \left( \sum_i p_i |\psi_i\rangle \langle \psi_i| \right) |n\rangle = \sum_i p_i \sum_{n=1}^{d} \langle n|\psi_i\rangle \langle \psi_i|n\rangle = \sum_i p_i \sum_{n=1}^{d} \langle\psi_i|n\rangle \langle n|\psi_i\rangle
\]

\[
= \sum_i p_i \langle \psi_i| \left( \sum_{n=1}^{d} |n\rangle \langle n| \right) |\psi_i\rangle = \sum_i p_i \langle \psi_i| \hat{I}|\psi_i\rangle = \sum_i p_i \langle \psi_i| \psi_i\rangle = \sum_i p_i = 1.
\]

This is the normalization requirement for mixed states. Physically, it means a guarantee to measure one of all the possible outcomes.

- \(\hat{\rho}\) is Hermitian;

\[
\hat{\rho}^\dagger = \sum_i p_i^* |\psi_i\rangle \langle \psi_i| = \hat{\rho} \text{ since } p_i \in \mathcal{R}.
\]
• \( \hat{\rho} \) is **positive semi-definite**: all eigenvalues of \( \hat{\rho} \) are real and non-negative;

\[
\langle \varphi | \hat{\rho} | \varphi \rangle = \sum_i p_i \langle \varphi | \psi_i \rangle \langle \psi_i | \varphi \rangle = \sum_i p_i | \langle \varphi | \psi_i \rangle |^2 \geq 0 \text{ since } p_i \geq 0
\]

and we can choose \( |\varphi\rangle \) to be any eigenstate of \( \hat{\rho} \).

This property is physical since a state must have a non-negative probability to be projected into any state.

### 2.5 Hilbert-Schmidt Inner Product

The set \( \mathbb{C}^{d \times d} \) of linear operators, i.e. \( d \times d \) complex matrices, is closed under addition and scalar multiplication; clearly, there is a zero element. It is easy to verify that \( \mathbb{C}^{d \times d} \) forms a \( d^2 \)-dimensional complex vector space \( \mathbb{C}^{d^2} \). So every \( d \times d \) complex matrix can be considered as a vector in \( \mathbb{C}^{d^2} \), i.e. a linear combination of a complete set of \( d^2 \) orthonormal “matrix” basis. With respect to a fixed set of such basis, the coefficients of this linear decomposition are unique to each matrix. Naturally, “inner product” between this “matrix” vector and each “matrix” basis gives the coefficients of the linear decomposition. But in the case of matrices, the appropriate inner product is the Hilbert-Schmidt inner product defined as

\[
\langle\langle A | B \rangle\rangle \equiv Tr \left( A^\dagger B \right), \quad A, B \in \mathbb{C}^{d \times d}.
\]

A little thought on the matrix multiplication process can reveal that the Hilbert-Schmidt inner product value is simply the sum of regular vector inner product values between each column of \( A \) with its corresponding column of \( B \).

As an example to demonstrate the vector space of matrices, we consider a set of
$d^2$ matrices $\{\hat{E}_{ij}\}$, where $i, j = 1, 2, \ldots, d$. $\hat{E}_{ij}$'s are $d \times d$ matrices with a single 1 as the element at the $i$th row and $j$th column and 0 everywhere else, ie. $(\hat{E}_{ij})_{ab} = \delta_{ia}\delta_{jb}$.

Obviously, these matrices are linearly independent and complete: any matrix in $\mathbb{C}^{d \times d}$ can be expressed as a linear combination of them with generally complex coefficients.

It is easy to show that this set of matrices are also orthonormal under the definition of the Hilbert-Schmidt inner product. We observe that

$$\hat{E}_{mn} \cdot \hat{E}_{pq} = \delta_{np}\hat{E}_{mq}$$

such that

$$\left\langle \langle \hat{E}_{nn} \mid \hat{E}_{pq} \rangle \right\rangle = Tr \left( \hat{E}_{nn}^\dagger \cdot \hat{E}_{pq} \right) = Tr \left( \hat{E}_{mn} \cdot \hat{E}_{pq} \right) = \delta_{np} Tr \left( \hat{E}_{mq} \right) = \delta_{np}\delta_{mq}.$$

Therefore, $\{\hat{E}_{ij}\}$ form a complete set of $d^2$ orthonormal basis of $\mathbb{C}^{d^2}$, ie. the vector space of $d \times d$ complex matrices. Any “matrix” vector $\hat{A}$ in $\mathbb{C}^{d^2}$ can be expressed as a unique linear combination of $\{\hat{E}_{ij}\}$ and the inner product $\left\langle \langle \hat{E}_{ij} \mid \hat{A} \rangle \right\rangle$ gives the coefficient of $\hat{E}_{ij}$ in $\hat{A}$.

### 2.6 Quantum State Tomography

Recall that in quantum mechanics, measurable physical quantities a.k.a. observables have a one-to-one correspondence with Hermitian matrices. The expectation value of Hermitian matrix $\hat{A}$ in a general quantum state $\rho$ takes the form

$$\left\langle \hat{A} \right\rangle = Tr(\hat{A}\rho) = Tr(\hat{A}^\dagger \rho)$$
of the Hilbert-Schmidt inner product. So the mathematical operation of taking inner products, i.e. projecting one “matrix” vector against another, actually describes the physical operation of measuring average values of observables on a quantum state.

From this perspective, the task of QST seems simple. If we can construct a complete set of $d^2$ orthonormal basis of $\mathbb{C}^{d^2}$ that are all Hermitian matrices, then we can measure all the physical quantities represented by them on any state in experiment. These measured average values are the coefficients of the state’s linear decomposition in this set of Hermitian basis. Since these basis are complete and orthonormal, their coefficients uniquely identify the state.

We can certainly construct such a set of $d^2$ Hermitian matrices. For example, it is easy to check that

$$\left\{ \hat{H} \right\} = \left\{ \hat{E}_{ii}, \hat{R}_{jk} = \frac{1}{2}(1 - \delta_{jk}) \left( \hat{E}_{jk} + \hat{E}_{kj} \right), \hat{C}_{jk} = \frac{i}{2}(1 - \delta_{jk}) \left( \hat{E}_{jk} - \hat{E}_{kj} \right) \right\},$$

where $i, j, k = 1, ..., d$, are complete and orthonormal with respect to the Hilbert-Schmidt inner product. The trace-1 constraint on the diagonal elements of density matrices will save one $\hat{E}_{ii}$ measurement. So we can measure the expectation values of the other $d^2 - 1$ observables in $\left\{ \hat{H} \right\}$ to fully determine any $d \times d$ density matrix.

However, in qubit based quantum computers, $d^2 = 2^{2N}$. So the number of measurements grows exponentially with the number of qubits used. The fact is that we can measure less than $d^2 - 1$ observables to determine any $\rho$. This is due to the fact that in the linear decomposition of $\rho$ in terms of $\left\{ \hat{H} \right\}$, the coefficients are not independent from each other because of the constraints of positive semi-definiteness and sometimes rank (for example, pure state $\rho$’s are rank-1).

\[^7\text{Unlike a general Hilbert-Schmidt inner product value that may be complex, measurement results from experiments are always real. Consistently, density matrices } \rho \text{ are Hermitian and hence always have real coefficients for Hermitian basis like } \hat{A}.\]
Therefore, the real challenge of QST is how we can use these constraints to minimize the number of measurements required to fully determine $\rho$.

2.7 Global and Relative Phase of Quantum States

From Born's rule, we see that all perceivable physical quantities of quantum mechanics come from the squared magnitude of wave functions, which are generally complex valued. So the global phase information of the wave function is absent from physical reality. For example, for pure states $|\varphi\rangle = e^{i\varphi}|\psi\rangle$ we have $P_{\varphi,n} = |\langle \varphi|n \rangle|^2 = \langle \varphi|n \rangle \langle n|\varphi \rangle = e^{-i\varphi}\langle \psi|n \rangle e^{i\varphi}\langle n|\psi \rangle = \langle \psi|n \rangle \langle n|\psi \rangle = |\langle \psi|n \rangle|^2 = P_{\psi,n}$. Similarly, $\langle \varphi|\hat{A}|\varphi \rangle = \langle \psi|\hat{A}|\psi \rangle$. In the same way, the global phases cancel out with their complex conjugates in density matrices: if $\{|\varphi_i\rangle\} = \{e^{i\varphi_i}|\psi_i\rangle\}$ and $\{p_i\}$ are the epistemic probabilities associated with each $i$, then we will have $\hat{\rho}_\varphi = \hat{\rho}_\psi = \sum_i p_i |\psi_i\rangle \langle \psi_i|$.

Therefore, we have the freedom to assign arbitrary global phase to a quantum state without affecting any measurement results, or physical reality. In fact, this property is called gauge symmetry of quantum mechanics and it is a continuous symmetry that plays an important role in quantum field theory. Consequently, when trying to determine a quantum state, we only need to determine it up to an arbitrary global phase. So we also do not care about the global phase of an operation acting on a quantum state. For example, if operator $M' = e^{i\alpha}M$, then $M'|\psi\rangle = e^{i\alpha}M|\psi\rangle = M|\psi\rangle$ up to a global phase. We thus consider $M'$ to be equivalent to $M$.

However, the relative phase between components of a quantum state is of crucial importance. For example, let $|\psi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\theta}|1\rangle)$. Then it is easy to compute that $P_{12} = |\langle \psi_1|\psi_2 \rangle|^2 = \frac{1}{2} + \frac{1}{2}\cos(\theta)$. The relative phase $\theta$ now contributes a sinusoidal term in the projective probability. In fact, this term is the cross term in the squared magnitude and indicates the interference within the
\[ |0 \rangle \text{ and } |1 \rangle \text{ superposition. This interference is the signature of the wave nature of quantum mechanics. In real experiments, } \theta \text{ is often implemented with a time or space dependence and can thus be directly measured.} \]

It is very important to differentiate between relative and global phase, any phase that has the potential of becoming a relative phase must be fully captured in determining quantum states or operations on them.

\section*{2.8 Unitary Operators}

A square matrix \( \hat{U} \) is said to be unitary if

\[ \hat{U} \dagger \hat{U} = \hat{I}. \]

Following this definition, it is easy to show the following properties:

1. \( \hat{U} \) is non-singular, \( \hat{U}^{-1} = \hat{U}^\dagger \);

2. \( \hat{U} \hat{U}^\dagger = \hat{I}, \hat{U}^\dagger \) is also unitary;

3. The rows and columns of \( \hat{U} \) each form a set of orthonormal basis;

4. If \( |\varphi \rangle = \hat{U} |\psi \rangle \), then \( \langle \varphi | \varphi \rangle = \langle \psi | \psi \rangle \);

5. All eigenvalues of \( \hat{U} \) lie on the unit circle of the complex plane.

From the definition and Property 2, we know that

\[ \text{Rows of } \hat{U} \text{ are orthonormal} \iff \text{Columns of } \hat{U} \text{ are orthonormal.} \]

We also know from row or column orthogonality that a unitary matrix has the
following representation
\[ \hat{U} = \sum_{i=1}^{d} |u_i\rangle \langle i| , \]
where \( d \) is the dimension of the Hilbert space, \( \{|i\rangle\} \) can be any complete set of orthonormal basis of the Hilbert space and \( \{|u_i\rangle\} \) is its image basis from \( \hat{U} \). In this form, it is clear that \( \hat{U} \) represents a change of basis a.k.a. a rotation of coordinate system in the Hilbert space. With this understanding, Property 4 and 5 are obvious. Rotations do not change the magnitude of vectors. Since there are no change of magnitude, all eigenvalues must be on the unit circle.

Each matrix element of \( \hat{U} \) is generally complex and must be specified by two real parameters, ie. phase and magnitude or real and imaginary parts. The 1st column of a \( d \times d \) \( \hat{U} \) has \( 2d - 2 \) real parameters to be specified. Because by the freedom to choose global phase, \( \hat{U}_{11} \) can be chosen to be real and positive, thus fully fixed by the normalization condition. The orthogonality condition \( \langle u_1 | u_2 \rangle = 0 \) between the 1st and 2nd column of \( \hat{U} \) adds two more constraints, ie. \( \text{Re}(\langle u_1 | u_2 \rangle) = 0 \) and \( \text{Im}(\langle u_1 | u_2 \rangle) = 0 \) on the 2nd column. However, the first element of the 2nd column may not be real and positive relative to the first element of the 1st column. So the 2nd column must also have an overall phase relative to the 1st column. So only \( (2d - 4) + 1 \) real parameters are free to vary in the 2nd column. The 3rd column must be orthogonal to the first two, so there are \( (2d - 6) + 1 \) free real parameters, etc. Therefore, the total number of real parameters that can be freely varied in a \( d \times d \) unitary operator is \( (2d - 2) + (2d - 4) + \ldots + 0 + (d - 1) = d^2 - 1 \).
2.9 Pauli Operators

Pauli operators have very nice properties. They are frequently used in quantum information processing. For example, in NMR experiments, they are the most convenient observables to measure. The three Pauli operators for a 2-level quantum system are

\[ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \]

In this thesis, we sometimes also use \( X, Y, \) and \( Z \) directly. In this section, we frequently use \( i, j \) and \( k \) to each denote all indices \( x, y \) and \( z \). In a compact form, all three Pauli matrices can be captured in the expression

\[ \sigma_i = \begin{pmatrix} \delta_{iz} & \delta_{ix} - i\delta_{iy} \\ \delta_{ix} + i\delta_{iy} & -\delta_{iz} \end{pmatrix}, \]

where \( \delta_{ij} \) is the Kronecker delta function. \( \delta_{ij} = 1 \) if \( i = j \). \( \delta_{ij} = 0 \) if \( i \neq j \). This general expression is useful in derivation when not any particular Pauli operator is chosen. We can choose any of the three Pauli operators by substituting \( i = x, y, z \).

Pauli matrices are Hermitian and unitary. So

\[ \sigma_i = \sigma_i^\dagger = \sigma_i^{-1} \Rightarrow \sigma_i^2 = I_{2\times2}. \]

They have a nice multiplication property

\[ \sigma_i \sigma_j = i\epsilon_{ijk} \sigma_k + \delta_{ij} I_{2\times2}, \]

where \( \epsilon_{ijk} \) is the 3d Levi-Civita tensor. \( \epsilon_{ijk} = 1 \) if \( ijk \) is an even permutation of 123; \( \epsilon_{ijk} = -1 \) if \( ijk \) is an odd permutation of 123; \( \epsilon_{ijk} = 0 \) if any two of the three indices
are equal.

This also translates to the following commutation and anti-commutation relations:

\[
[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k, \quad \{\sigma_i, \sigma_j\} = 2\delta_{ij}\hat{I}_{2\times2}.
\]

The spectrum of \(\sigma_i\) is \(\pm 1\). Therefore, \(Tr(\sigma_i) = 0\) and \(det(\sigma_i) = -1\).

The positive and negative eigenstates of \(\sigma_x\) are respectively

\[
|+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad
|\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix};
\]

the positive and negative eigenstates of \(\sigma_y\) are respectively

\[
|+i\rangle = \frac{1}{\sqrt{2}} (|0\rangle + i|1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix}, \quad
|\rangle = \frac{1}{\sqrt{2}} (|0\rangle - i|1\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix};
\]

the positive and negative eigenstates of \(\sigma_z\) are respectively

\[
|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\]

From the multiplication property, it is easy to see that Pauli operators, together with \(\hat{I}_{2\times2}\), can form a complete set of orthonormal basis with respect to the Hilbert-Schmidt inner product. If their coefficients are restricted to be real, then their span is the vector space of \(2 \times 2\) Hermitian matrices. If their coefficients are allowed to be complex, their span is the entire vector space of \(2 \times 2\) matrices.
2.10 Positive Linear Maps

A linear map \( \varphi : \mathbb{C}^{n \times n} \to \mathbb{C}^{m \times m} \) is called a positive map if for every positive semi-definite matrix \( N \in \mathbb{C}^{n \times n} \), its image matrix \( \varphi(N) \in \mathbb{C}^{m \times m} \) is also positive semi-definite. For any linear map \( \varphi : \mathbb{C}^{n \times n} \to \mathbb{C}^{m \times m} \), we can construct another map \( \hat{I}_{k \times k} \otimes \varphi : \mathbb{C}^{kn \times kn} \to \mathbb{C}^{km \times km} \) that applies map \( \varphi \) to each \( n \times n \) block of the input state. So

\[
\hat{I}_{k \times k} \otimes \varphi \text{ maps } \begin{pmatrix} N_{11} & \ldots & N_{1k} \\ \vdots & \ddots & \vdots \\ N_{k1} & \ldots & N_{kk} \end{pmatrix} \text{ to } \begin{pmatrix} \varphi(N_{11}) & \ldots & \varphi(N_{1k}) \\ \vdots & \ddots & \vdots \\ \varphi(N_{k1}) & \ldots & \varphi(N_{kk}) \end{pmatrix}.
\]

\( \varphi \) is called \( k \)-positive if \( \hat{I}_{k \times k} \otimes \varphi \) is a positive map. \( \varphi \) is called completely positive if \( \hat{I}_{k \times k} \otimes \varphi \) is a positive map for all \( k \).

Proof of \( \Rightarrow \):

Let’s denote each \( n \times n \) block of a positive \( kn \times kn \) matrix \( \mathcal{N} \) as \( N_{pq} \), where \( p, q = 1, \ldots, k \). So \( \mathcal{N} = [N_{pq}]_{1 \leq p, q \leq k} \). Then for any vector \( |v\rangle \in \mathbb{C}^{km} \), we have

\[
\langle v | \hat{I}_{k \times k} \otimes \varphi(\mathcal{N}) |v\rangle = \langle v | \sum_{i} K_i N_{pq} K_i^\dagger |v\rangle = \langle v | \sum_{i} K_i N_{pq} K_i^\dagger |v\rangle_{1 \leq p, q \leq k}
\]

\[
= \sum_{i} \langle v | K_i N_{pq} K_i^\dagger |v\rangle = \sum_{i} \langle v | \hat{I}_{k \times k} \otimes \sum_{i} K_i (N_{pq})_{1 \leq p, q \leq k} \hat{I}_{k \times k} \otimes K_i^\dagger |v\rangle
\]

\[
= \sum_{i} \langle v | \hat{I}_{k \times k} \otimes K_i \mathcal{N} \hat{I}_{k \times k} \otimes K_i^\dagger |v\rangle.
\]
For $\varphi : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{m \times m}$, $K_i \in \mathbb{C}^{m \times n}$. Therefore, $|w\rangle = \hat{I}_{k \times k} \otimes K_i^\dagger |v\rangle \in \mathbb{C}^{kn}$. Thus,

$$
(\langle v| \hat{I}_{k \times k} \otimes \varphi(N) |v\rangle = \sum_i \langle v| \hat{I}_{k \times k} \otimes K_iN \hat{I}_{k \times k} \otimes K_i^\dagger |v\rangle = \sum_i \langle w|N|w\rangle.
$$

This is clearly positive because $N$ is positive. Since this is true for any $|v\rangle$ and any $k$, $\varphi$ is a completely positive map. The proof in the $\Leftarrow$ direction is done in proving Choi’s theorem.

## 2.11 Choi’s Theorem and Channel State Duality

The famous Choi’s theorem states the following about linear maps:

Let $\varphi : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{m \times m}$ be a positive linear map, then the following statements are equivalent:

- $\varphi$ is $n$-positive;

- The Choi matrix

$$
C_\varphi \equiv \left(\hat{I}_{n \times n} \otimes \varphi\right) \left(\sum_{i,j=1}^{n} \hat{E}_{ij} \otimes \hat{E}_{ij}\right) = \sum_{i,j=1}^{n} \hat{E}_{ij} \otimes \varphi(\hat{E}_{ij}) \in \mathbb{C}^{nm \times nm}
$$

is positive, where $(\hat{E}_{ij})_{ab} = \delta_{ia}\delta_{jb}$;

- $\forall N \in \mathbb{C}^{n \times n}$, $\varphi(N) = \sum_i K_iN K_i^\dagger$ for some set of $\{K_i\} \in \mathbb{C}^{m \times n}$. Therefore, $\varphi$ is a completely positive linear map.

The detailed proof of this theorem can be found in M.D. Choi’s 1975 paper[24]. Because of the positivity of the Choi matrix $C_\varphi$, we can consider it as a quantum state. So there is a one-to-one correspondence between quantum channel $\varphi$ and
its associated Choi matrix state $C_{\varphi}$. This is the so called channel state duality, or sometime referred to as the Choi-Jamiolkowski isomorphism.

It is easy to check by direct computation that $\varphi(\hat{E}_{pq}) = \text{Tr}_{C_{m\times m}} \left( C_{\varphi} \cdot \hat{E}_{pq}^{T} \otimes \hat{I}_{C_{m\times m}} \right)$. Then by the linearity of map $\varphi$, we have its output from any input state $\rho$ as

$$\varphi(\rho) = \text{Tr}_{C_{n\times n}} \left( C_{\varphi} \cdot \rho^{T} \otimes \hat{I}_{C_{m\times m}} \right).$$

### 2.12 General Quantum Operations

In its most general form, the action of quantum operation $\varepsilon(\cdot)$ on density matrix $\rho$ can be described by a completely positive trace-preserving (CPTP) linear map:

$$\varepsilon(\rho) = \sum_{i} K_{i} \rho K_{i}^{\dagger}. \tag{2.12.1}$$

This operation must be positive and trace preserving because the output $\varepsilon(\rho)$ must also be a quantum state. The completely positive requirement is due to the fact that $\varepsilon(\cdot)$ could be a local operation on parts of an arbitrarily large system a.k.a. $\hat{I}_{n\times n} \otimes \varepsilon(\cdot)$ must be positive for any $n$. The trace-preserving condition requires

$$\text{Tr}(\varepsilon(\rho)) = \text{Tr}(\sum_{i} K_{i} \rho K_{i}^{\dagger}) = \sum_{i} \text{Tr}(K_{i} \rho K_{i}^{\dagger}) = \sum_{i} \text{Tr}(\rho K_{i}^{\dagger} K_{i})$$

$$= \text{Tr}(\rho \sum_{i} K_{i}^{\dagger} K_{i}) = \text{Tr}(\rho) = 1 \quad \forall \rho.$$

Notice that $\sum_{i} K_{i}^{\dagger} K_{i}$ is a positive Hermitian operator. So we can choose $\rho$ to be any one of its eigenstates. Hence all eigenvalues of $\sum_{i} K_{i}^{\dagger} K_{i}$ are equal to 1. Therefore,
the trace preserving condition is equivalent to

\[ \sum_i K_i^\dagger K_i = \hat{I}. \tag{2.12.2} \]

The \( K_i \)'s are called **Kraus Operators**. \( \varepsilon (\rho) = \sum_i K_i \rho K_i^\dagger \) is called the **Kraus representation** of the CPTP map. Kraus representations are not unique. The number of orthogonal Kraus operators is the rank of the map.

For a unitary map, \( \varepsilon (\rho) = U \rho U^\dagger \) ie. we have only one Kraus operator \( K = U \).

We can choose a complete set of orthonormal operator basis \( \{ \hat{E}_m \} \)\(^8\) to express each \( K_i \) as

\[ K_i = \sum_m e_{im} \hat{E}_m. \tag{2.12.3} \]

Then our general quantum operation can be written as

\[ \varepsilon (\rho) = \sum_{mn} \chi_{mn} \hat{E}_m \rho \hat{E}_n^\dagger, \tag{2.12.4} \]

where

\[ \chi_{mn} = \sum_i e_{im} e_{in}^*. \tag{2.12.5} \]

All information about \( \varepsilon () \) is now transferred into \( \chi_{mn} \). \( \chi \) is called the **Process Matrix** of \( \varepsilon () \).

\( \chi \) is Hermitian since \( \chi_{mn} = \sum_i e_{im} e_{in}^* = \chi_{nm}^* \).

\( \chi \) is also positive semi-definite, ie. \( \chi \geq 0 \), since \( \chi = \sum_i \chi_i \) and \( \chi_i = |e_i \rangle \langle e_i| \) is clearly positive, where \( |e_i \rangle \) is the vector of decomposition coefficients of \( K_i \) with respect to \( \{ \hat{E}_m \} \), ie. \( \langle \hat{E}_m | e_i \rangle = e_{im} \).

The trace preserving condition (2.12.2) is now \( \sum_{mn} \chi_{mn} \hat{E}_m^\dagger \hat{E}_n = \hat{I} \).

\(^8\)Orthonormal with respect to the Hilbert-Schmidt inner product: \( \text{Tr}(\hat{E}_m^\dagger \hat{E}_n) = \delta_{mn} \).
Because $\chi$ describes a linear transformation acting on operators, it is sometimes called a super-operator. If $\{\hat{E}_m\}$ has a total of $r + 1$ elements, we may express $\varepsilon(\rho)$ as

$$
\varepsilon(\rho) = \begin{pmatrix} \hat{E}_0 & \hat{E}_1 & \ldots & \hat{E}_r \end{pmatrix} \begin{pmatrix} \chi_{00} & \chi_{01} & \ldots & \chi_{0r} \\ \chi_{10} & \chi_{11} & \ldots & \chi_{1r} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{r0} & \chi_{r1} & \ldots & \chi_{rr} \end{pmatrix} \rho \begin{pmatrix} \hat{E}_0^\dagger \\ \hat{E}_1^\dagger \\ \vdots \\ \hat{E}_r^\dagger \end{pmatrix},
$$

or in terms of block matrix multiplications

$$
\varepsilon(\rho) = \begin{pmatrix} \hat{E}_0 & \hat{E}_1 & \ldots & \hat{E}_r \end{pmatrix} \begin{pmatrix} \chi_{00} \hat{I}_{d \times d} & \chi_{01} \hat{I}_{d \times d} & \ldots & \chi_{0r} \hat{I}_{d \times d} \\ \chi_{10} \hat{I}_{d \times d} & \chi_{11} \hat{I}_{d \times d} & \ldots & \chi_{1r} \hat{I}_{d \times d} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{r0} \hat{I}_{d \times d} & \chi_{r1} \hat{I}_{d \times d} & \ldots & \chi_{rr} \hat{I}_{d \times d} \end{pmatrix} \hat{I}_{r \times r} \otimes \rho_{d \times d} \begin{pmatrix} \hat{E}_0^\dagger \\ \hat{E}_1^\dagger \\ \vdots \\ \hat{E}_r^\dagger \end{pmatrix}.
$$

The process matrix $\chi$ is equal to the Choi matrix if it is expanded in basis $\{\hat{E}_{ij}\}$. We can see this by examining the output states from $\{\hat{E}_{ij}\}$ as inputs:

$$
\varepsilon(\hat{E}_{ij}) = \sum_{m,n,p,q=1}^d \chi_{mn,pq} \hat{E}_{mn} \hat{E}_{ij} \hat{E}_{pq}^\dagger = \sum_{m,n,p,q=1}^d \chi_{mn,pq} \hat{E}_{mn} \hat{E}_{ij} \hat{E}_{qp}^\dagger
$$

$$
= \sum_{m,n,p,q=1}^d \chi_{mn,pq} \delta_{ni} \delta_{jq} \hat{E}_{mp}
$$

So if we were to construct the Choi matrix of $\varepsilon(\cdot)$, its $ij$th block will be the coefficients $\chi_{mi,pj}$. 

44
2.13 Standard Quantum Process Tomography

For an unknown quantum operation \( \varepsilon() \), we can determine its action on a given quantum state \( \rho \) by performing quantum state tomography on the output \( \varepsilon(\rho) \). If we determine its actions on a complete set of basis in which all density matrices can be represented, then by the linearity of \( \varepsilon() \) we have fully determined its action on the input’s Hilbert space. This procedure is called Quantum Process Tomography.

For a \( d \)-level system, \( \rho \) is a \( d \times d \) matrix. If \( \varepsilon() \) is a map from \( d \times d \) matrices to \( d \times d \) matrices, then \( \chi \) is a \( d^2 \times d^2 \) Hermitian matrix that contains \( d^4 \) independent real parameters.\(^9\) The trace-preserving completeness relation (2.12.2) will give \( d^2 \) constraints. Therefore, we need to determine a total of \( d^4 - d^2 \) independent real parameters to completely specify such a map.

We choose a complete set of state basis \( \{\rho_j, j = 1, 2, \ldots, d^2\} \) to probe the map:

\[
\varepsilon(\rho_j) = \sum_k \lambda_{jk} \rho_k. \tag{2.13.1}
\]

In particular, for each pair of operator basis acting on the state basis, we have

\[
\hat{E}_m \rho_j \hat{E}_n^\dagger = \sum_k \beta_{jk}^{mn} \rho_k. \tag{2.13.2}
\]

Substitute (2.13.2) into (2.12.4), we have

\[
\varepsilon(\rho_j) = \sum_{mn} \chi_{mn} \hat{E}_m \rho_j \hat{E}_n^\dagger = \sum_{mn} \sum_k \chi_{mn} \beta_{jk}^{mn} \rho_k = \sum_k \lambda_{jk} \rho_k. \tag{2.13.3}
\]

\(^9\)For a general \( d \times d \) Hermitian matrix, the \( d \) diagonal elements are real; there are \( \frac{(d^2-d)}{2} \) elements in the upper triangle, each can have a magnitude and phase; the lower triangle is the complex conjugate of the upper triangle. Therefore, its total number of independent real parameters is \( d + \frac{(d^2-d)}{2} \times 2 = d^2 \).
By linear independence of state basis \( \{ \rho_j \} \), we must have

\[
\sum_{mn} \chi_{mn} \beta_{jk}^{mn} = \lambda_{jk}.
\]

(2.13.4)

From (2.13.2), we see that \( \beta_{jk}^{mn} \) describes the relationship between the state basis \( \{ \rho_j \} \) and the operator basis \( \{ \hat{E}_m \} \). So upon choosing these basis, \( \beta_{jk}^{mn} \) can be determined. Then if we start from (2.13.1), probe \( \varepsilon() \) with a complete set of state basis \( \{ \rho_j \} \) and perform quantum state tomography on each output \( \varepsilon(\rho_j) \) to obtain \( \lambda_{jk} \), we will be able to obtain \( \chi_{mn} \) using (2.13.4) and hence fully characterize \( \varepsilon() \)[15, Section 8.4.2].

The above method guarantees the complete characterization of the most general quantum operation. The operation may have rank greater than 1 and the output may be a mixed state. However, if we know that our quantum operation is unitary, then it is rank 1 and it outputs pure states from pure inputs. In this case, we may be able to fully determine the operation with less measurements.

For a \( d = 2 \) one qubit system, we need to measure \( d^4 - d^2 = 12 \) elements of the \( \chi \)-matrix to fully determine the operation. For a \( d = 4 \) two qubit system, we need to measure 240 elements of \( \chi \). In the next section, we will show that if we know our quantum operation is unitary, then we only need to measure 6 observable values to fully characterize a 1-qubit unitary gate.
Chapter 3

QPT of 1-Qubit Unitary Gates

In this section, we present both adaptive and non-adaptive methods to determine a 1-qubit unitary gate. These methods rely on the special way in which 1-qubit unitary gates can be parametrized and thus does not generalize easily to higher dimensions. Nevertheless, 1-qubit unitary gates are so commonly used in quantum information processing that studying this special case is still of great interest in practice. By visualizing the QPT method on the Bloch sphere, the homomorphism between $SU(2)$ and $SO(3)$ has been shown. This in turn shows that studying higher dimensional real space may help in finding QPT methods for multiple qubit unitary gates.

3.1 Parametrization and QPT Methods

Parametrize a general 1-qubit 2x2 unitary matrix as

$$\hat{U} = \begin{pmatrix} \cos\left(\frac{\theta_1}{2}\right) & \sin\left(\frac{\theta_1}{2}\right) e^{i\theta_2} \\ -\sin\left(\frac{\theta_1}{2}\right) e^{i\theta_3} & \cos\left(\frac{\theta_1}{2}\right) e^{i(\theta_2+\theta_3)} \end{pmatrix},$$

where $\theta_1 \in [0, \pi]$, $\theta_2, \theta_3 \in [0, 2\pi)$.

(3.1.1)
Input the following two pure states:

\[
\left| 0 \right> = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \left| + \right> = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.
\]

Denote their corresponding outputs as

\[
\rho_0 = \hat{U} \left| 0 \right> \left< 0 \right| \hat{U}^\dagger, \quad \rho_+ = \hat{U} \left| + \right> \left< + \right| \hat{U}^\dagger.
\]

Measure Pauli observables

\[
\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.
\]

The analytical expressions of the measurement results are summarized in Table 3.1.

<table>
<thead>
<tr>
<th>State</th>
<th>Observable</th>
<th>Measured Observable Average</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>\rho_0</td>
<td>\sigma_z</td>
<td>\text{Tr}(\rho_0 \sigma_z) = \cos(\theta_1)</td>
<td>\theta_1</td>
</tr>
<tr>
<td>\sigma_x</td>
<td>\text{Tr}(\rho_0 \sigma_x) = -\sin(\theta_1)\cos(\theta_3)</td>
<td>\theta_3</td>
<td></td>
</tr>
<tr>
<td>\sigma_y</td>
<td>\text{Tr}(\rho_0 \sigma_y) = -\sin(\theta_1)\sin(\theta_3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\rho_+</td>
<td>\sigma_z</td>
<td>\text{Tr}(\rho_+ \sigma_z) = \sin(\theta_1)\cos(\theta_2)</td>
<td>\theta_2</td>
</tr>
<tr>
<td>\sigma_x</td>
<td>\text{Tr}(\rho_+ \sigma_x) = \cos(\theta_1)\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\sin(\theta_3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\sigma_y</td>
<td>\text{Tr}(\rho_+ \sigma_y) = \cos(\theta_1)\cos(\theta_2)\sin(\theta_3) + \sin(\theta_2)\cos(\theta_3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Non-Adaptive Measurements in NMR

Table 3.1 summarizes the non-adaptive protocol to fully tomograph $\hat{U}$. We are guaranteed to have determined $\theta_1$, $\theta_2$ and $\theta_3$ from the 6 measurement results.
To be convinced of this claim, we have to navigate through the adaptive protocol, where we are allowed to look at the measurement result after each step and make decision accordingly. In the end, we will see that the table above covers all measurements needed in all circumstances.

- We first measure $Tr(\rho_0\sigma_z) = \cos(\theta_1)$. Since $\theta_1 \in [0, \pi]$ coincides with the range of the $\arccos$ function, who is monotonic, we can inverse $\cos$ to uniquely determine

\[
\theta_1 = \arccos(Tr(\rho_0\sigma_z)) \in [0, \pi] \\
\Rightarrow \sin(\theta_1) \geq 0, \quad \sin(\theta_1) = \sqrt{1 - \cos^2(\theta_1)}.
\]

- If $\theta_1 \neq 0$ or $\pi \Rightarrow \sin(\theta_1) \neq 0$:

First we measure $Tr(\rho_0\sigma_x)$ to get

\[
\cos(\theta_3) = -\frac{1}{\sin(\theta_1)} Tr(\rho_0\sigma_x).
\]

If $\cos(\theta_3) = 1 \Rightarrow \theta_3 = 0$.

If $\cos(\theta_3) = -1 \Rightarrow \theta_3 = \pi$.

Otherwise, we also need to measure $Tr(\rho_0\sigma_y)$ to get

\[
\sin(\theta_3) = -\frac{1}{\sin(\theta_1)} Tr(\rho_0\sigma_y)
\]

to fully determine $\theta_3 \in [0, 2\pi)$.
Next we measure \( \text{Tr}(\rho_+\sigma_z) \) to get
\[
\cos(\theta_2) = \frac{1}{\sin(\theta_1)} \text{Tr}(\rho_+\sigma_z).
\]

If \( \cos(\theta_2) = 1 \Rightarrow \theta_2 = 0. \)

If \( \cos(\theta_2) = -1 \Rightarrow \theta_2 = \pi. \)

Otherwise, we also need to get \( \sin(\theta_2) \) to fully determine \( \theta_2 \in [0, 2\pi). \)

If \( \theta_3 \neq 0 \) or \( \pi \Rightarrow \sin(\theta_3) \neq 0, \) then we measure \( \text{Tr}(\rho_+\sigma_x) \) to get
\[
\sin(\theta_2) = \frac{1}{\sin(\theta_3)} [\cos(\theta_1)\cos(\theta_2)\cos(\theta_3) - \text{Tr}(\rho_+\sigma_x)].
\]

Otherwise, \( \sin(\theta_3) = 0 \Rightarrow \cos(\theta_3) \neq 0, \) we measure \( \text{Tr}(\rho_+\sigma_y) \) to get
\[
\sin(\theta_2) = \frac{1}{\cos(\theta_3)} [\text{Tr}(\rho_+\sigma_y) - \cos(\theta_1)\cos(\theta_2)\sin(\theta_3)] = \frac{1}{\cos(\theta_3)} \text{Tr}(\rho_+\sigma_y).
\]

Under these circumstances, we need at least 3, at most 5 measurements to fully determine \( \hat{U}. \)

- If \( \theta_1 = 0 \Rightarrow \cos(\frac{\theta_1}{2}) = 1, \sin(\frac{\theta_1}{2}) = 0, \) so

\[
\hat{U} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i(\theta_2+\theta_3)} \end{pmatrix}.
\]

Therefore we only need to get \( (\theta_2 + \theta_3) \in [0, 2\pi) \) as a whole instead of \( \theta_2 \) and \( \theta_3 \) individually.
We know that $\cos(\theta) = \cos(0) = 1$, so we measure

$$Tr(\rho + \sigma_x) = \cos(\theta_1) \cos(\theta_2) \cos(\theta_3) - \sin(\theta_2) \sin(\theta_3) = \cos(\theta_2 + \theta_3).$$

If $\cos(\theta_2 + \theta_3) = 1 \Rightarrow (\theta_2 + \theta_3) = 0$.

If $\cos(\theta_2 + \theta_3) = -1 \Rightarrow (\theta_2 + \theta_3) = \pi$.

Otherwise, we also need to measure

$$Tr(\rho + \sigma_y) = \cos(\theta_1) \cos(\theta_2) \sin(\theta_3) + \sin(\theta_2) \cos(\theta_3) = \sin(\theta_2 + \theta_3)$$

to fully determine $(\theta_2 + \theta_3) \in [0, 2\pi)$.

Under these circumstances, we need at least 2, at most 3 measurements to fully determine $\hat{U}$.

- If $\theta_1 = \pi \Rightarrow \cos(\frac{\theta_1}{2}) = 0$, $\sin(\frac{\theta_1}{2}) = 1$, so

$$\hat{U} = \begin{pmatrix} 0 & e^{i\theta_2} \\ -e^{i\theta_3} & 0 \end{pmatrix} = \begin{pmatrix} 0 & -e^{i(\theta_2 - \theta_3)} \\ 1 & 0 \end{pmatrix} \text{ up to global phase } - e^{i\theta_3}. $$

Therefore we only need to get $(\theta_2 - \theta_3) \in [0, 2\pi)$ as a whole instead of $\theta_2$ and $\theta_3$ individually.

We know that $\cos(\theta) = \cos(\pi) = -1$, so we measure

$$Tr(\rho + \sigma_x) = \cos(\theta_1) \cos(\theta_2) \cos(\theta_3) - \sin(\theta_2) \sin(\theta_3) = -\cos(\theta_2 - \theta_3).$$

If $\cos(\theta_2 - \theta_3) = 1 \Rightarrow (\theta_2 - \theta_3) = 0$.

If $\cos(\theta_2 - \theta_3) = -1 \Rightarrow (\theta_2 - \theta_3) = \pi$.
Otherwise, we also need to measure
\[
Tr(\rho_+\sigma_y) = \cos(\theta_1)\cos(\theta_2)\sin(\theta_3) + \sin(\theta_2)\cos(\theta_3) = \sin(\theta_2 - \theta_3)
\]
to fully determine \((\theta_2 - \theta_3) \in [0, 2\pi)\).

Under these circumstances, we need at least 2, at most 3 measurements to fully determine \(\hat{U}\).

To sum up, we need at least 2, at most 5 measurements to fully determine \(\hat{U}\) adaptively. The table summarizes all 6 measurements needed in the non-adaptive approach.

### 3.2 Experimental Implementation

There are two common ways to implement a process tomography experiment in real physical systems.

The first one is the Nuclear Magnetic Resonance (NMR) system. In this system, an ensemble of identical molecules are prepared in some mixed state described by density matrix \(\rho\) and the ensemble average of observable \(\sigma\) is measured. The measurement outcome is \(\langle \sigma \rangle = Tr(\rho \sigma)\).

The second way is using the optics system. In this case, measurements are projective. Photons are measured against a complete set of orthonormal basis. For example, projective measurements of a 2-qubit system require four detectors, each attached to projective basis \(|00\rangle, |01\rangle, |10\rangle\) and \(|11\rangle\). When a single photon passes, only one of the four detectors will register a signal. So a large number of \(N\) photons all prepared in the same state must be passed through the detector setup one by one.
In the end, the four detectors will register count numbers \( n_1, n_2, n_3 \) and \( n_4 \). By the completeness of the orthonormal basis, it must be that \( n_1 + n_2 + n_3 + n_4 = N \). Then \( p_i = \frac{n_i}{N} \), where \( i = 1, 2, 3, 4 \). Of course, for a 1-Qubit system we simply have basis \(|0\rangle\) and \(|1\rangle\). So \( p_1 = 1 - p_0 \).

For 1-Qubit systems, there is a simple conversion between the measurement results of the two systems. This is useful during data analysis. We list them below:

\[
P_{00} = \frac{1}{2}(1 + Tr(\rho_0\sigma_z)), \quad P_{0+} = \frac{1}{2}(1 + Tr(\rho_0\sigma_x)), \quad P_{0+i} = \frac{1}{2}(1 + Tr(\rho_0\sigma_y)); \quad (3.2.1)
\]

\[
P_{+0} = \frac{1}{2}(1 + Tr(\rho_+\sigma_z)), \quad P_{++} = \frac{1}{2}(1 + Tr(\rho_+\sigma_x)), \quad P_{++i} = \frac{1}{2}(1 + Tr(\rho_+\sigma_y)). \quad (3.2.2)
\]

### 3.3 Maximum Likelihood Estimation

For a quantum system, Born’s rule gives the theoretical prediction of each measurement outcome’s probability. In real experiments, the actual values measured will always be fluctuating randomly around the theoretical predictions. So in reality, we very likely will obtain raw data values that are incompatible with the unitary parametrization of our gate. For example, we may obtain a set of data values that imply \( \cos^2(\theta_1) + \sin^2(\theta_1) \neq 1 \). Therefore, linear inversion of the raw data is not a feasible way to determine the gate parameters.

Maximum Likelihood Estimation (MLE) is an useful method to obtain parameters of a statistical model in this situation. In this method, we must presume a statistical distribution that the obtained data follows. The method then tunes the parameters in the statistical model to make the obtained data *most likely*. Then this optimized set of parameters is considered as the best fit of this statistical model to our measured
data.

For NMR experiments, if we assume that the measured observable averages $m_j$ in experiments are independent from each other, each follows a Gaussian distribution centered at the theoretical prediction $M_j(\tilde{\theta})$ calculated by Born’s rule (for example, from unitary parametrization (3.1.1)) and have a common standard deviation $\sigma$, then the likelihood of measuring the set of results $\{m_j\}$ is

$$L = \prod_j L_j, \quad \text{where } L_j = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(M_j(\tilde{\theta}) - m_j)^2}{2\sigma^2}\right).$$

It is helpful to use log likelihood when you have a product form. Since log is a monotonic function, applying log to $L$ does not change its extrema. So instead, we can choose to vary parameters $\tilde{\theta}$ to maximize

$$l = \ln(L) = \sum_j \left[-\ln(\sigma \sqrt{2\pi}) - \frac{(M_j(\tilde{\theta}) - m_j)^2}{2\sigma^2}\right].$$

Since we assumed that standard deviation $\sigma$ is the same for each physical system, we only need to minimize

$$l_1 = \sum_j \left(M_j(\tilde{\theta}) - m_j\right)^2. \quad (3.3.1)$$

This looks like a least square optimization. However, we are varying $\tilde{\theta}$, not $M_j$’s directly.

For optics experiments, we projective measure $N$ photons against a complete

\footnote{Within the same physical system, say optics or NMR, all measurements correspond to the same physical process, only with the measurement devices oriented differently. So it might be reasonable to assume that the outcomes have the same spread a.k.a. standard deviation. But further justification might be required regarding the details of the measurement process.}
set of orthonormal basis \{ |j \rangle \}, each with theoretical probability \( P_j(\vec{\theta}) \) calculated by Born’s rule (for example, from unitary parametrization (3.1.1)) and measured count number \( n_j \). Since the projection of each photon is an independent event, the joint probability of measuring the data set \( \{n_j\} \) is

\[
l = \prod_j P_j(\vec{\theta})^{n_j}.
\]

We would like this likelihood function to be an intensive quantity that does not change with the size of sample space. So we normalize the powers \( n_j \) with the total photon number \( N = \sum_j n_j \) to have

\[
l_2 = \prod_j P_j(\vec{\theta})^{f_j}, \quad \text{where} \quad f_j = \frac{n_j}{N}.
\]

We will be varying \( \vec{\theta} = (\theta_1, \theta_2, \theta_3) \) to minimize \( l_1 \) in (3.3.1) and maximize \( l_2 \) in (3.3.2). Then the corresponding \( \hat{U} \) can be determined from the parametrization (3.1.1).

The goal of extremizing these likelihood functions is consistent with our original intention to find a \( \hat{U} \) that is closest to the given measurement data. This is obvious when minimizing \( l_1 \) since it is the squared difference between theoretical results from \( \hat{U} \) and measured data. The case of maximizing \( l_2 \) is less trivial. The following calculations may provide some intuition that maximizing \( l_2 \) over \( \vec{\theta} \) also gives the minimal difference between \( \left\{ P_j(\vec{\theta}) \right\} \) and \( \{f_j\} \).

\footnote{There are different permutations of the \( j \)'s that give the same combination of \( \{n_j\} \). However, we do not need to take into account of this degeneracy factor. Because we are looking for the probability of that particular sequence in which each photon was projected into different channels labeled by \( j \)'s during experiment, not the total probability of getting a combination of \( \{n_j\} \). We do not record this sequence (or permutation) in our data because we only need the total count number \( n_j \) for each channel \( j \) to calculate the probability.}
Let there be a total of $J$ measured probabilities, i.e. $j = 1, 2, \ldots, J$. Then

$$l_2 = l_2(\vec{\theta}) = l_2(P_1(\vec{\theta}), P_2(\vec{\theta}), \ldots, P_J(\vec{\theta})).$$

We must have

$$\vec{\nabla}_{\vec{\theta}} l_2 = \frac{\partial l_2}{\partial P_1} \vec{\nabla}_{\vec{\theta}} P_1 + \frac{\partial l_2}{\partial P_2} \vec{\nabla}_{\vec{\theta}} P_2 + \ldots + \frac{\partial l_2}{\partial P_J} \vec{\nabla}_{\vec{\theta}} P_J = 0$$

at the extrema of $l_2(\vec{\theta})$. In the case of 1-qubit unitary gates, $\vec{\theta} = (\theta_1, \theta_2, \theta_3)$. So we must have

$$\left(\nabla_{\theta_1} l_2, \nabla_{\theta_2} l_2, \nabla_{\theta_3} l_2\right) = \left(0, 0, 0\right),$$

which means that

$$\left(\frac{\partial l_2}{\partial P_1} \cdot \frac{\partial P}{\partial \theta_1}, \frac{\partial l_2}{\partial P_2} \cdot \frac{\partial P}{\partial \theta_2}, \frac{\partial l_2}{\partial P_3} \cdot \frac{\partial P}{\partial \theta_3}\right) = \frac{\partial l_2}{\partial P} \cdot \left(\frac{\partial P}{\partial \theta_1}, \frac{\partial P}{\partial \theta_2}, \frac{\partial P}{\partial \theta_3}\right) = \left(0, 0, 0\right).$$

Clearly this will hold if

$$\frac{\partial l_2}{\partial P} = \left(\frac{\partial l_2}{\partial P_1}, \frac{\partial l_2}{\partial P_2}, \ldots, \frac{\partial l_2}{\partial P_J}\right) = \left(0, 0, \ldots, 0\right).$$

For projective measurements on 1-qubit systems, every complete set of orthonormal basis has only two components. So $J$ must be an even integer and $\{P_j(\vec{\theta})\}$ and $\{f_j\}$ can both be arranged in pairs such that $P_2 = 1 - P_1, f_2 = 1 - f_1, P_4 = 1 - P_3, f_4 = 1 - f_3$, etc. Therefore,

$$l_2 = P_1^{f_1}(1 - P_1)^{(1-f_1)} \times P_3^{f_3}(1 - P_3)^{(1-f_3)} \times \ldots \times P_{J-1}^{f_{J-1}}(1 - P_{J-1})^{(1-f_{J-1})}.$$
Then the 1st and 2nd derivatives of \( l_2 \) with respect to any \( P_i \) are

\[
\frac{\partial l_2}{\partial P_i} = \frac{P_i^{f_i-1}(f_i - P_i)}{(1 - P_i)^{f_i}} \times \prod_{j \neq i} P_j^{f_j}(1 - P_j)^{(1 - f_j)} = 0 \text{ if } P_i = f_i
\]

and

\[
\frac{\partial^2 l_2}{\partial P_i^2} = \frac{P_i^{f_i-2}f_i(f_i - 1)}{(1 - P_i)^{f_i+1}} \times \prod_{j \neq i} P_j^{f_j}(1 - P_j)^{(1 - f_j)} < 0 \text{ since } f_i < 1.
\]

Therefore, \( \{P_j\} = \{f_j\} \) is the maximum point. So maximizing \( l_2 \) indeed minimizes the difference between theoretical values \( \{P_j(\vec{\theta})\} \) and measured data \( \{f_j\} \).

There probably does not exist any \( \vec{\theta} \) that can satisfy \( \{P_j(\vec{\theta})\} = \{f_j\} \) perfectly. Because the ratios between elements of \( \{P_j(\vec{\theta})\} \) implied by the parametrization of \( \vec{\theta} \) are not perfectly preserved in \( \{f_j\} \) due to experimental errors. But as long as experimental errors are reasonably small, the optimization results will be very close to the true values of \( \vec{\theta} \) because \( \{P_j(\vec{\theta})\} \) and \( l_2(\{P_j\}, \{f_j\}) \) are all continuous differentiable functions.

### 3.4 Experimental Plan

In NMR experiments, we measure the ensemble average of \( \sigma_x \), \( \sigma_y \) and \( \sigma_z \) in state \( \rho_0 \) and \( \rho_+ \). As the conversion relations (3.2.1) and (3.2.2) implies, we do projective measurements of \( \rho_0 \) and \( \rho_+ \) onto basis \( \{|0\rangle, |1\rangle\} \), \( \{|+\rangle, |−\rangle\} \) and \( \{|+i\rangle, |-i\rangle\} \) in the optics experiment.

To test the quality of our tomography protocol, we compare the tomographed \( \hat{U}' \) reconstructed from measurements above with the true gate \( \hat{U} \) by calculating the gate fidelity defined as

\[
f = \frac{1}{d} |Tr(\hat{U}^{\dagger}\hat{U}')|, \quad (3.4.1)
\]
where $d = 2^n$ is the dimension of the Hilbert space. So $d = 2$, 4 for 1 and 2-qubit gates. Obviously $f = 1$ when $\hat{U}' = \hat{U}$. So a fidelity number close to 1 means our reconstruction is close to the true gate.

The true gates $\hat{U}$ we used in the 1 qubit NMR experiment are listed in Table 3.2.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Increment</th>
<th># of Points</th>
<th>Fixed Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>$[0, \pi]$</td>
<td>$\frac{\pi}{20}$</td>
<td>21</td>
<td>$\theta_2 = \frac{\pi}{8}$, $\theta_3 = \frac{\pi}{8}$.</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>$[0, 2\pi]$</td>
<td>$\frac{\pi}{20}$</td>
<td>41</td>
<td>$\theta_1 = \frac{\pi}{6}$, $\theta_3 = \frac{\pi}{8}$.</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>$[0, 2\pi]$</td>
<td>$\frac{\pi}{20}$</td>
<td>41</td>
<td>$\theta_1 = \frac{\pi}{6}$, $\theta_2 = \frac{\pi}{8}$.</td>
</tr>
</tbody>
</table>

Table 3.2: Unitary Gates Tomographed in NMR

The true gates $\hat{U}$ we used in the 1 qubit optics experiment are listed in Table 3.3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Increment</th>
<th># of Points</th>
<th>Fixed Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_1$</td>
<td>$[0, \pi]$</td>
<td>$\frac{\pi}{20}$</td>
<td>21</td>
<td>$\theta_2 = \frac{\pi}{8}$, $\theta_3 = \frac{\pi}{8}$.</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>$[0, 2\pi]$</td>
<td>$\frac{2\pi}{20}$</td>
<td>41</td>
<td>$\theta_1 = \frac{2\pi}{9}$, $\theta_3 = \frac{\pi}{8}$.</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>$[0, 2\pi]$</td>
<td>$\frac{2\pi}{20}$</td>
<td>41</td>
<td>$\theta_1 = \frac{2\pi}{9}$, $\theta_2 = \frac{\pi}{3}$.</td>
</tr>
</tbody>
</table>

Table 3.3: Unitary Gates Tomographed in Optics

The optics experimental set up is highly automated and fast in obtaining data. So for comparison, we also repeated the NMR experiment in Table 3.2 on the optical setup.

We set up these tables to do a full range sweep of one parameter while fixing the other two at some general values.
3.5 Optimization and Results

Due to easy conversions between ensemble averages and probabilities in the 1-qubit case as shown in (3.2.1) and (3.2.2), we can use both likelihood functions (3.3.1) and (3.3.2) in the gate reconstruction from NMR and optics data.

As demonstrated before, we will be able to use less measurements in an adaptive tomography procedure, where we stop and look at the measurement result after each step to determine the next step. Nevertheless, this is impractical to operate for experimentalists. So we can only retreat to the non-adaptive approach where all 6 measurements listed in Table (3.1) are used. Unfortunately, $Tr(\rho_+\sigma_z)$ was missed out during the experiment and we are forced to reconstruct all $U$’s from the other 5 measurements. As we will see, this is mostly okay except in the special case of $\theta_1 = \frac{\pi}{2}$. We first relist the expressions of the three Pauli measurement results on $\rho_+$ for convenience:

$$Tr(\rho_+\sigma_z) = \sin(\theta_1)\cos(\theta_2);$$

$$Tr(\rho_+\sigma_x) = \cos(\theta_1)\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\sin(\theta_3);$$

$$Tr(\rho_+\sigma_y) = \cos(\theta_1)\cos(\theta_2)\sin(\theta_3) + \sin(\theta_2)\cos(\theta_3).$$

We have shown in the adaptive tomography procedure that in the special case of $\theta_1 = 0$ or $\pi$, only $Tr(\rho_+\sigma_x)$ and $Tr(\rho_+\sigma_y)$ are required to fully fix the gate. So missing $Tr(\rho_+\sigma_z)$ is okay.

In all other cases, we would have already determined $\theta_1$ and $\theta_3$ from the three Pauli measurements on $\rho_0$. Then to determine $\theta_2$, we generally need to solve for both $\cos(\theta_2)$ and $\sin(\theta_2)$ from $Tr(\rho_+\sigma_x)$ and $Tr(\rho_+\sigma_y)$. This is a $2 \times 2$ linear system. In
matrix form, we have

\[
\begin{pmatrix}
\cos(\theta_1)\cos(\theta_3) & -\sin(\theta_3) \\
\cos(\theta_1)\sin(\theta_3) & \cos(\theta_3)
\end{pmatrix}
\begin{pmatrix}
\cos(\theta_2) \\
\sin(\theta_2)
\end{pmatrix}
= \begin{pmatrix}
\text{Tr}(\rho_+\sigma_x) \\
\text{Tr}(\rho_+\sigma_y)
\end{pmatrix}.
\]

The determinant of the coefficient matrix is simply \(\cos(\theta_1)\). So it is only when \(\det = \cos(\theta_1) = 0 \Rightarrow \theta_1 = \frac{\pi}{2}\) that the linear system is irreversible. This is when we need to measure \(\text{Tr}(\rho_+\sigma_x) = \sin(\theta_1)\cos(\theta_2) = \sin(\frac{\pi}{2})\cos(\theta_2) = \cos(\theta_2)\). When \(\cos(\theta_1) = 0\), \(\text{Tr}(\rho_+\sigma_x) = -\sin(\theta_2)\sin(\theta_3)\) and \(\text{Tr}(\rho_+\sigma_y) = \sin(\theta_2)\cos(\theta_3)\). Since \(\sin(\theta_3)\) and \(\cos(\theta_3)\) cannot be zeros simultaneously, we are guaranteed to get \(\sin(\theta_2)\).

In this special case, we will need all three Pauli measurements on \(\rho_+\) to guarantee determination of \(\theta_2\) hence \(\hat{U}\). We will see an example of this special case in our results.

Substituting the results of Table (3.1) into (3.3.1) and (3.3.2), we have

\[
l_1 = (\cos(\theta_1) - m_{0z})^2 \\
+ (-\sin(\theta_1)\cos(\theta_3) - m_{0x})^2 \\
+ (-\sin(\theta_1)\sin(\theta_3) - m_{0y})^2 \\
+ (\cos(\theta_1)\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\sin(\theta_3) - m_{+x})^2 \\
+ (\cos(\theta_1)\cos(\theta_2)\sin(\theta_3) + \sin(\theta_2)\cos(\theta_3) - m_{+y})^2.
\]
\[
l_2 = (1 + \cos(\theta_1))^{f_0}(1 - \cos(\theta_1))^{1-f_0} \\
(1 - \sin(\theta_1)\cos(\theta_3))^{f_0+}(1 + \sin(\theta_1)\cos(\theta_3))^{1-f_0+} \\
(1 - \sin(\theta_1)\sin(\theta_3))^{f_0+i}(1 + \sin(\theta_1)\sin(\theta_3))^{1-f_0+i} \\
(1 + \cos(\theta_1)\cos(\theta_2)\cos(\theta_3) - \sin(\theta_2)\sin(\theta_3))^{f_0+i} \\
(1 + \cos(\theta_1)\cos(\theta_2)\cos(\theta_3) + \sin(\theta_2)\sin(\theta_3))^{1-f_0+i} \\
(1 - \cos(\theta_1)\sin(\theta_2)\sin(\theta_3) + \sin(\theta_2)\cos(\theta_3))^{f_0+i} \\
(1 + \cos(\theta_1)\cos(\theta_2)\sin(\theta_3) - \sin(\theta_2)\cos(\theta_3))^{1-f_0+i}.
\]

In \(l_2\), we have used the fact that in the 1-Qubit case \(P_0 + P_1 = P_+ + P_- = P_{+i} + P_{-i} = 1\). We also dropped the factor of \(\frac{1}{2}\) for all the terms.

Since \(l_1\) and \(l_2\) both contain powers of sinusoidal functions, they have many local extrema. To avoid trapping the optimization algorithm within the regions of local extrema, we first estimated the \(\theta\)'s from linear inversions of the measured data. These estimates should be reasonably close to the true global extremum such that there are no other local extrema to fall into in the vicinity. We then used these estimates as the initial points to start off the optimization. This also saves significant computing effort when the initial guess is close to the true value.

The fidelity plot obtained by reconstructing from the NMR data using \(l_1\) and \(l_2\) are shown in Figure 3.5.1 and 3.5.2.

According to Table (3.2), we should only have 103 data points, each for a different gate, instead of 123 as shown on the plots. This is because the range of \(\theta_1\) only goes from 0 to \(\pi\). But the experiment varied \(\theta_1\) from 0 to \(2\pi\), giving 20 extra points. These 20 \(\hat{U}\)'s are the same as the 20 with \(\theta_1\) from 0 to \(\pi\) only with an overall negative sign that can be ignored.

The results are mostly very good with average fidelity \(\bar{f}_{1,NMR} = 0.9946\) using \(l_1\).
Figure 3.5.1: Unitary QPT from NMR data using $l_1$ and $\bar{f}_{2,NMR} = 0.9945$ using $l_2$.

The obvious problem is Gate #11 where the fidelity is very low in both plots. This is the fore-mentioned special case. $\theta_2$ and $\theta_3$ were both fixed at $\frac{\pi}{8}$. At $\frac{\pi}{20}$ increment starting from $\theta_1 = 0$, we will have $\theta_1 = \frac{\pi}{2}$ at Gate # 11, thus $\cos(\theta_1) = 0$. Although in real experiment we do not get this zero exactly, we still get a very small number that almost eliminates the influence of $\cos(\theta_2)$ on the value of $Tr(\rho_+\sigma_x)$ and $Tr(\rho_+\sigma_y)$. Since the experiment missed the $Tr(\rho_+\sigma_z)$ measurement, we do not have a way to obtain information of $\cos(\theta_2)$. Knowing $\sin(\theta_2)$, we can only determine $\theta_2$ up to the y-axis reflection, ie. $\sin(\theta_2) = \sin(\pi - \theta_2)$. Indeed, the optimization results from the program at Gate # 11 are $(\theta_1, \theta_2, \theta_3) = (1.5539, 2.7622, 0.4063)$ using $l_1$ and $(\theta_1, \theta_2, \theta_3) = (1.5584, 2.7630, 0.4161)$ using $l_2$. In both cases $\theta_1$ and $\theta_3$ are close to what they should be. However, $\theta_2 \simeq \frac{7\pi}{8} = \pi - \frac{\pi}{8}$, which caused the low fidelity.
Figure 3.5.2: Unitary QPT from NMR data using $l_2$

Overall, the high fidelity for each gate shows that our tomography method is very successful on an NMR system. Have we not missed the $Tr(\rho_+\sigma_z)$ measurement, the results would have been perfect.

Unfortunately, data from the optics experiment have mistakes. The raw data are simply photon count numbers in each projective channel. These measured values are too far off from the theoretical predictions. Some channels have count numbers that are on the order of $10^0$ while it is supposed to be on the order of $10^5$ according to theoretical calculations. It appears that the photon counters associated with those channels does not have enough registers. For example, the photon counters can record the correct number up to 9999 but then overflows to 0000 when the 10000th photon come in. Therefore, we were unable to reconstruct unitary gates out of these wrong data.
3.6 Visualization in 3D

The tomography of 1-qubit unitary gates can be nicely visualized in 3-dimensional real space. To do so, we first introduce the Bloch sphere and its properties.

The Bloch sphere is commonly used to visualize 1-qubit quantum states. A general 1-qubit pure state can be parametrized as

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle,$$

where $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$.

The coefficient of $|0\rangle$ can always be set to real and non-negative by the freedom to choose any global phase of $|\psi\rangle$. So $\theta$ does not need to cover the full range of $[0, 2\pi)$. The density matrix of $|\psi\rangle$ is

$$|\psi\rangle \langle \psi| = \begin{pmatrix} \cos^2\left(\frac{\theta}{2}\right) & e^{-i\varphi}\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) \\ e^{i\varphi}\cos\left(\frac{\theta}{2}\right)\sin\left(\frac{\theta}{2}\right) & \sin^2\left(\frac{\theta}{2}\right) \end{pmatrix}.$$
bijective maps in the domain $\theta \in [0, \pi]$ and $\varphi \in [0, 2\pi)$, there is an one-to-one correspondence between all 1-qubit pure states $\{ |\psi\rangle \}$ and all points $\{ P \}$ on the surface of the unit sphere. If every $P$ is labeled as the corresponding $|\psi\rangle$, then we obtain the Bloch sphere as shown in Figure 3.6.2.
Notice that

\[ |+\rangle \langle +| = \frac{1}{2}(\hat{I} + \hat{x} \cdot \vec{\sigma}), \quad |−\rangle \langle −| = \frac{1}{2}(\hat{I} - \hat{x} \cdot \vec{\sigma}); \]

\[ |+i\rangle \langle +i| = \frac{1}{2}(\hat{I} + \hat{y} \cdot \vec{\sigma}), \quad |−i\rangle \langle −i| = \frac{1}{2}(\hat{I} - \hat{y} \cdot \vec{\sigma}); \]

\[ |0\rangle \langle 0| = \frac{1}{2}(\hat{I} + \hat{z} \cdot \vec{\sigma}), \quad |1\rangle \langle 1| = \frac{1}{2}(\hat{I} - \hat{z} \cdot \vec{\sigma}), \]

where

\[ \hat{x} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \quad \hat{y} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}, \quad \hat{z} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \]

are the unit vectors on the positive \( x \), \( y \) and \( z \)-axis of the 3D Cartesian coordinate system. Therefore, orthonormal basis \{\( |+\rangle \), \( |−\rangle |\} and \{\( |+i\rangle \), \( |−i\rangle |\} lie on the positive and negative \( x \), \( y \) and \( z \)-axis of the Bloch sphere, respectively. This also demonstrates that the two \( \hat{r} \) vectors of a set of orthonormal basis are collinear and opposite of each other. They form a diameter, or “axis”, of the Bloch sphere. So orthonormal states are “antipodal” to each other on the Bloch sphere. For example, \( |0\rangle \) and \( |1\rangle \) are at the north and south pole.

Any 1-qubit mixed state \( \rho \) must have spectrum decomposition \( \rho = p |b_{0}\rangle \langle b_{0}| + (1 - p) |b_{1}\rangle \langle b_{1}| \), where \( 0 < p < 1 \), \( \langle b_{0}|b_{0}\rangle = \langle b_{1}|b_{1}\rangle = 1 \) and \( \langle b_{0}|b_{1}\rangle = 0 \). Therefore, \( \hat{r}_{b_{0}} = -\hat{r}_{b_{1}} \) and \( \hat{r}_{\rho} = p\hat{r}_{b_{0}} + (1-p)\hat{r}_{b_{1}} = (2p-1)\hat{r}_{b_{0}} \). \( |\hat{r}_{\rho}| = |2p - 1| < 1 \). So \( P_{\rho} \) lies inside the Bloch sphere on its diameter connecting \( P_{b_{0}} \) and \( P_{b_{1}} \). The maximally mixed state \( \rho_{0} = \frac{1}{2}(|0\rangle \langle 0| + |1\rangle \langle 1|) \) is not bias towards any direction. So it must be located at the center of the Bloch sphere. Indeed, \( \rho_{0} = \frac{1}{2}\hat{I} \Rightarrow \hat{r}_{\rho_{0}} = (0 \ 0 \ 0) \). Since \{\( |b_{0}\rangle \), \( |b_{1}\rangle \)} can be any basis and \( p \) can vary continuously in \((0, 1)\), \( P_{\rho} \) can move freely along any diameter of the Bloch sphere. So there is a one-to-one correspondence between all points inside the Bloch sphere and all possible 1-qubit mixed states.

As shown in Equation (3.2.1) and (3.2.2), expectation value and projective mea-
surements are equivalent for 1-qubit systems. Projective measurements are easier to visualize on the Bloch sphere. From Born’s rule and the parametrization of a general 1-qubit pure state, we have the projective probability between $|\psi\rangle$ and $|0\rangle$ as

$$P_{\psi0} = |\langle\psi|0\rangle|^2 = \cos^2\left(\frac{\theta}{2}\right).$$

Since $\theta \in [0, \pi] \Rightarrow \frac{\theta}{2} \in [0, \frac{\pi}{2}] \Rightarrow \cos\left(\frac{\theta}{2}\right) \geq 0$, $\cos\left(\frac{\theta}{2}\right) = \sqrt{P_{\psi0}}$. Therefore,

$$\theta = 2\arccos(\sqrt{P_{\psi0}}).$$

So the measured value of $P_{\psi0}$ fixes $\theta$ uniquely within $[0, \pi]$. Therefore, all possible states $|\psi\rangle$ with this measurement value can only differ from each other in $\varphi$. To help visualize this set of states in Figure 3.6.1, we may consider the red vector $r = 1$ as $\hat{r}_\psi$. As $\varphi$ varies over its full range of $[0, 2\pi)$, the motion of $\hat{r}_\psi$ is a precession about $\hat{z}$ at polar angle $\theta$. This precession sweeps out a cone with the center of the Bloch sphere as its vertex, $\hat{z}$ as its axis and $2\theta$ as its opening angle. The tip of $\hat{r}_\psi$, ie. $P_\psi$, or $P(x, y, z)$ as labeled in Figure 3.6.1, traces out a circle on the surface of the Bloch sphere with radius $\sin(\theta)$ centered at $(0 0 \cos(\theta))$. This circle is the collection of points corresponding to all possible states $|\psi\rangle$ that will give the measured $P_{\psi0}$ value. In other words, every measurement constrains the unknown $|\psi\rangle$ to be on one such circle.

Without loss of generality, we may use any set of orthonormal basis $\{|b_0\rangle, |b_1\rangle\}$ to parametrize the same state: $|\psi\rangle = \cos\left(\frac{\theta_b}{2}\right)|b_0\rangle + e^{i\varphi_b}\sin\left(\frac{\theta_b}{2}\right)|b_1\rangle$, where $\theta_b \in [0, \pi]$ and $\varphi_b \in [0, 2\pi)$. So the projective measurement in this basis constrains $|\psi\rangle$ to be on the circle traced out on the surface of the Bloch sphere by the precession of $\hat{r}_\psi$ around $\hat{r}_{b_0}$ at angle $\theta_b$. 

67
Therefore, different measurements on the same $|\psi\rangle$ constrain it to be at one of the points of intersection that all these circles pass through. If there exists only one such point, then $|\psi\rangle$ has been determined uniquely.

Our 1-qubit QPT method can be visualized this way. The outputs $|u_0\rangle = \hat{U}|0\rangle$ and $|u_+\rangle = \hat{U}|+\rangle$ are pure states represented by points on the surface of the Bloch sphere. The expectation value measurement of $\sigma_x$, $\sigma_y$ and $\sigma_z$ are equivalent to projective measurements against basis $\{|+\rangle, |-\rangle\}$, $\{|+i\rangle, |-i\rangle\}$ and $\{|0\rangle, |1\rangle\}$, respectively. So these measurements constrain $|u_0\rangle$ and $|u_+\rangle$ to be on circles traced out on the surface of the Bloch sphere by precession of $\hat{r}_{u_0}$ and $\hat{r}_{u_+}$ around the $x$, $y$ and $z$–axis. The angle of precession is fixed by the measured values. Figure 3.6.3 illustrates the tomography of $|u_0\rangle$. The $x$, $y$ and $z$–axis of the Bloch sphere are colored in yellow, green and blue.

![Figure 3.6.3: Tomography of $|u_0\rangle$ on the Bloch sphere](image)

The expectation value $Tr(|u_0\rangle\langle u_0|\sigma_z)$ is measured first on the left in Figure 3.6.3. It constrains $|u_0\rangle$ to be on the red circle centered around the $z$–axis in blue. As the measured value varies from 1 to $-1$, this circle will expand from a red dot at the north pole to the equator in the $xy$–plane and then contract back to a red dot at the south pole while it translates vertically on the spherical surface. If we are lucky to have measured the value of 1 or $-1$ corresponding to the red dot at the north or
south pole, then \(|u_0\rangle = |0\rangle\) or \(|1\rangle\). These are special cases discussed in our method.

\(Tr(|u_0\rangle \langle u_0| \sigma_x)\) is measured next in the middle in Figure 3.6.3. It constrains \(|u_0\rangle\) to also be on the second red circle centered around the \(x\)-axis in yellow. So \(|u_0\rangle\) must be at one of the two points of intersection between the two red circles as marked by the white dots.

Since measurement values are obtained from the same \(|u_0\rangle\), these two red circles are guaranteed to share at least one common point, ie. be tangent to each other. Then we know that \(|u_0\rangle\) is the state represented by the tangent point and no more measurements are needed to determine it. This is also a special case discussed in our method. But generally there will be two points of intersection so that a third measurement is required to eliminate one of the two and determine \(|u_0\rangle\) uniquely.

As shown on the right in Figure 3.6.3, \(Tr(|u_0\rangle \langle u_0| \sigma_y)\) is measured in the third step to constrain \(|u_0\rangle\) to be on the red circle centered around the \(y\)-axis in green. Since the underlying state is still \(|u_0\rangle\), this circle is guaranteed to pass through one of the two points of intersection between the other two circles. This common point marked by the white star corresponds to the state \(|u_0\rangle\) we seek. The uniqueness of the white star is guaranteed because the third circle is a circle of fixed \(y\)-value while the two white dots only differ in their \(y\)-values.

So in general, we need to measure all three expectation values \(\sigma_x\), \(\sigma_y\) and \(\sigma_z\) to guarantee the unique determination of \(|u_0\rangle\). By determining \(|u_0\rangle\), we have fixed \(\theta_1\) and \(\theta_3\) in our parametrization of \(\hat{U}\) as shown in (3.1.1). To find \(\theta_2\) and completely fix \(\hat{U}\), we need to determine \(|u_+\rangle = \hat{U} |+\rangle\).

To do so, it is important to utilize the condition that \(\langle u_0|u_+\rangle = \langle 0|\hat{U}^\dagger\hat{U}|+\rangle = \langle 0|\hat{I}|+\rangle = \langle 0|+\rangle = \frac{1}{\sqrt{2}}.\) As shown before, this dictates that the angle between \(\hat{r}_{u_0}\) and \(\hat{r}_{u_+}\) on the Bloch sphere is \(2\arccos(\sqrt{P_{u_0u_+}}) = 2\arccos(\sqrt{\langle u_0|u_+\rangle^2}) = 2\arccos\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{2},\) ie. \(\hat{r}_{u_+}\) is always orthogonal to \(\hat{r}_{u_0}\). So if we consider \(|u_0\rangle\) as the
north pole of the Bloch sphere, then $|u_+⟩$ must be on the equator. This is shown on the left in Figure 3.6.4. $|u_1⟩$ is antipodal to $|u_0⟩$ by orthogonality: $⟨u_0|u_1⟩ = ⟨0|\hat{U}^†\hat{U}|1⟩ = ⟨0|\hat{I}|1⟩ = ⟨0|1⟩ = 0$. So $|u_0⟩$ and $|u_1⟩$ are connected by Bloch sphere’s axis colored in black in the figure. $|u_+⟩$ must be somewhere on the black equator centered around this axis. By determining $|u_0⟩$ in previous steps, we have already constrained $|u_+⟩$ to be on this black circle without any new measurement.

Since we already have the first circle, we can just duplicate the last two steps of the tomography of $|u_0⟩$ to determine $|u_+⟩$. So we measure the expectation values of $\sigma_x$ and $\sigma_y$ in state $|u_+⟩$. The obtained value of $Tr(|u_+⟩⟨u_+|\sigma_x)$ constrains $|u_+⟩$ to be on the purple circle centered around the yellow $x$–axis as shown in the middle in Figure 3.6.4. Generally this purple circle will intercept the black circle twice as marked by the two white dots. In a special case, there is only one tangent point between the two circles. Then we can identify $|u_+⟩$ as the unique state represented by this tangent point. Otherwise, we measure $Tr(|u_+⟩⟨u_+|\sigma_y)$ to constrain $|u_+⟩$ further on the purple circle centered around the green $y$–axis as shown on the right in Figure 3.6.4. This second purple circle is guaranteed to intercept the first two circles at one of the two points of intersection at least. If this common point shared by the three circles is unique, then it represents the state $|u_+⟩$ we seek.
Figure 3.6.5: Special Case in Tomography of $|u_+\rangle$

However, this common point is not unique in a special case. As shown on the left in Figure 3.6.5, when the black axis happens to lie inside the $xy$–plane, then $|u_+\rangle$ is constrained to be on the black circle that coincides with a longitude line connecting $|0\rangle$ and $|1\rangle$ on the blue $z$–axis. If we were to still measure expectation values of $\sigma_x$ and $\sigma_y$ to constrain $|u_+\rangle$ on the purple circles centered around the $x$ and $y$–axis as shown on the left, then we have a configuration that is symmetric with respect to the $xy$–plane. Therefore, these three circles must have two points of intersection that are also symmetric with respect to the $xy$–plane, as marked by the two white dots. So in this special case, measuring $\sigma_x$ and $\sigma_y$ cannot determine a unique $|u_+\rangle$. This symmetry with respect to the $xy$–plane means that these two white dots only differ from each other in their $z$–coordinates. So to eliminate one of the two and determine $|u_+\rangle$ uniquely, we must measure $\sigma_z$ and $\sigma_x$ or $\sigma_z$ and $\sigma_y$. The case of measuring $\sigma_z$ and $\sigma_x$ is shown on the right in Figure 3.6.5. $|u_+\rangle$ is determined to be at the unique point marked by the white star.

The above special case corresponds to the low fidelity data point in Figure 3.5.1 and Figure 3.5.2 where $\theta_1 = \frac{\pi}{2}$. The low fidelity of gates constructed from measured $\sigma_x$ and $\sigma_y$ values confirms the above claim that $\sigma_z$–values must be measured for successful tomography. Indeed, at $\theta_1 = \frac{\pi}{2}$, $|u_0\rangle$ lies inside the $xy$–plane.
There are two cases that are even more special. If the black axis lies on the $x$–axis, then the black circle perpendicular to it must be inside the $yz$–plane. In this case, measuring $\sigma_x$ provides no useful information because this measurement constraint simply redraws the black circle. So we can only determine $|u_+\rangle$ by measuring $\sigma_y$ and $\sigma_z$. Similarly, if the black axis lies on the $y$–axis, we can only determine $|u_+\rangle$ by measuring $\sigma_x$ and $\sigma_z$.

So in general, depending on the $|u_0\rangle$ determined previously, we need to measure $\sigma_x$ and $\sigma_y$, $\sigma_z$ and $\sigma_x$ or $\sigma_z$ and $\sigma_y$ to uniquely determine $|u_+\rangle$. This also fixes $\theta_2$, which is the last required parameter to uniquely identify a general 1-qubit unitary operation $\hat{U}$.

In summary, to guarantee the successful tomography of any 1-qubit unitary operation $\hat{U}$, we need to measure the expectation values of all three Pauli observables $\sigma_x$, $\sigma_y$ and $\sigma_z$ in both $|u_0\rangle$ and $|u_1\rangle$. So a total of six measurements are required in non-adaptive QPT. We may choose to save one Pauli observable measurement for state $|u_+\rangle$. But this choice depends on the $|u_0\rangle$ we obtained previously. So a total of at most five measurements are required in adaptive QPT.

### 3.7 Homomorphism to $SO(3)$ and Euler Angles

It is not a coincidence that QPT of 1-qubit unitary gates can be visualized so nicely in 3D. Recall the general parametrization of a 1-qubit unitary gate $\hat{U}$ in Equation (3.1.1):

$$\hat{U} = \begin{pmatrix}
\cos(\theta_1/2) & \sin(\theta_1/2) e^{i\theta_2} \\
-sin(\theta_1/2) e^{i\theta_3} & \cos(\theta_1/2) e^{i(\theta_2+\theta_3)}
\end{pmatrix}, \quad \text{where } \theta_1 \in [0, \pi], \quad \theta_2, \theta_3 \in [0, 2\pi].$$

Due to the freedom to choose global phase, we may consider
\[ \hat{U}' = \hat{U} e^{-\frac{i}{2}(\theta_2 + \theta_3)} = \begin{pmatrix} \cos\left(\frac{\theta_1}{2}\right) e^{-\frac{i}{2}(\theta_2 + \theta_3)} & \sin\left(\frac{\theta_1}{2}\right) e^{\frac{i}{2}(\theta_2 - \theta_3)} \\ -\sin\left(\frac{\theta_1}{2}\right) e^{-\frac{i}{2}(\theta_2 - \theta_3)} & \cos\left(\frac{\theta_1}{2}\right) e^{\frac{i}{2}(\theta_2 + \theta_3)} \end{pmatrix}. \]

For cleaner notation, we will now denote this new \( \hat{U}' \) as \( \hat{U} \).

Then it is easy to compute that

\[ \det(\hat{U}) = 1 \quad \Rightarrow \quad \hat{U} \in SU(2). \]

There is a well known group homomorphism \( \varphi : SU(2) \to SO(3) \) \(^{12}\) with the kernel \( \left\{ \pm \hat{I}_{2\times2} \right\} \). In other words, the quotient group \( Q_{SU(2)} = SU(2)/\left\{ \pm \hat{I}_{2\times2} \right\} \) is isomorphic to \( SO(3) \). Since \(-\hat{I}_{2\times2}\) is essentially a global phase of \(-1\), we can always ignore it in tomography. So the task of QPT is simply to identify each element in \( Q_{SU(2)} \). We can do this by determining the element’s isomorphic image in \( SO(3) \). Denote this image as \( \hat{R} \). So

\[ \hat{R} = \varphi(\hat{U}) = \varphi(-\hat{U}), \quad \hat{R} \in SO(3). \]

Transformations can only be visualized by its effects on the entity it acts on. For

\(^{12}\)Let \( G \) and \( H \) be two groups. Let \( \cdot \) and \( \circ \) denote the group multiplications of \( G \) and \( H \). Then \( G \) is said to be homomorphic to \( H \) if there exists a map \( \varphi : G \to H \) such that \( \forall g_1, g_2 \in G, \varphi(g_1 \cdot g_2) = \varphi(g_1) \circ \varphi(g_2) \). \( G \) is said to be isomorphic to \( H \) if \( \varphi \) is a bijective map.

\( SU(2) = \left\{ \hat{U} \in \mathbb{C}^{2\times2} \mid \det(\hat{U}) = 1, \hat{U}^\dagger \hat{U} = \hat{I}_{2\times2} \right\} \). \( SU(2) \) is the group of all 2 \( \times \) 2 complex unitary matrices whose determinant is equal to 1. “SU” stands for “Special Unitary” group.

\( SO(3) = \left\{ \hat{R} \in \mathbb{R}^{3\times3} \mid \det(\hat{R}) = 1, \hat{R}^T \hat{R} = \hat{I}_{3\times3} \right\} \). \( SO(3) \) is the group of all 3 \( \times \) 3 real orthogonal matrices whose determinant is equal to 1. “SO” stands for “Special Orthogonal” group.

\(^{13}\)The kernel is the set of all elements in \( SU(2) \) whose image under map \( \varphi \) is the identity of \( SO(3) \), ie. \( \hat{I}_{3\times3} \).

\(^{14}\)The quotient group \( SU(2)/\left\{ \pm \hat{I}_{2\times2} \right\} \) is defined as \( \hat{U} \{ \pm \hat{I}_{2\times2} \} \mid \hat{U} \in SU(2) \} = \{ \{ \hat{U}, -\hat{U} \} \mid \hat{U} \in SU(2) \} \). This set forms a group under multiplication “\( \cdot \)” defined as \( \hat{U}_1 \{ \pm \hat{I}_{2\times2} \} \cdot \hat{U}_2 \{ \pm \hat{I}_{2\times2} \} = \hat{U}_1 \hat{U}_2 \{ \pm \hat{I}_{2\times2} \} \). Quotient groups are also called factor groups.

In general, the quotient group \( G/D \) must have its denominator \( D \) as an invariant subgroup of group \( G \): \( D \) and \( G \) are groups and \( D \subset G \); \( \forall \ g \in G \) and \( \forall \ d \in D \), \( gdg^{-1} \in D \).
\(\hat{U}\), this entity is a 1-qubit state \(\rho\). For \(\hat{R}\), it is a vector \(\vec{r}\) in \(\mathbb{R}^3\). For unambiguous visualization, there must be an one-to-one map from \(\rho\) to \(\vec{r}\). As we have shown in the previous section, such a map exists in the form of

\[
\rho = \frac{1}{2}(\hat{I}_{2\times2} + \hat{r}_\rho \cdot \vec{\sigma}) \text{ where } \vec{\sigma} = \begin{pmatrix} \sigma_x & \sigma_y & \sigma_z \end{pmatrix} \text{ is the vector of Pauli matrices.}
\]

\(|\hat{r}_\rho| = 1\) if \(\rho\) is a pure state. \(0 \leq |\hat{r}_\rho| < 1\) if \(\rho\) is a mixed state. Therefore, in the last section, we can successfully visualize the transformation of 1-qubit pure states \(|\psi\rangle\) by \(\hat{U}\) as the transformation of unit vectors \(\hat{r}_\psi\) by \(\hat{R}\).

Transformations by \(\hat{R}\) and \(\hat{U}\) both preserve inner products:

- Let \(\vec{v}_1, \vec{v}_2 \in \mathbb{R}^3\). Denote \(\vec{w}_1 = \hat{R}\vec{v}_1\) and \(\vec{w}_2 = \hat{R}\vec{v}_2\), then

\[\vec{w}_1 \cdot \vec{w}_2 = \vec{w}_1^T \vec{w}_2 = \vec{v}_1^T \hat{R} \hat{R}^T \vec{v}_2 = \vec{v}_1^T \hat{I}_{3\times3} \vec{v}_2 = \vec{v}_1^T \vec{v}_2 = \vec{v}_1 \cdot \vec{v}_2\]

If \(\vec{v}_1 = \vec{v}_2\), then \(\vec{w}_1 = \vec{w}_2\) and we have

\[\vec{w}_1 \cdot \vec{w}_1 = |\vec{w}_1|^2 = \vec{v}_1 \cdot \vec{v}_1 = |\vec{v}_1|^2\]

\[\Rightarrow |\vec{w}_1| = |\vec{v}_1|.
\]

So \(\hat{R}\) preserves vector length.

Inner products of \(\vec{a}, \vec{b} \in \mathbb{R}^3\) can also be expressed as \(\vec{a} \cdot \vec{b} = |\vec{a}| |\vec{b}| \cos(\theta_{ab})\), where \(\theta_{ab}\) is the angle between \(\vec{a}\) and \(\vec{b}\).

Hence,

\[\vec{w}_1 \cdot \vec{w}_2 = |\vec{w}_1||\vec{w}_2| \cos(\theta_w) = |\vec{v}_1||\vec{v}_2| \cos(\theta_w) = \vec{v}_1 \cdot \vec{v}_2 = |\vec{v}_1||\vec{v}_2| \cos(\theta_v)\]

\[\Rightarrow \cos(\theta_w) = \cos(\theta_v). \text{ Since } \theta_w, \theta_v \in [0, \pi], \text{ we must have } \theta_w = \theta_v.
\]

74
So \( \hat{R} \) also preserves the angle between two vectors.

Therefore, \( \hat{R} \) describes a rotations in 3-dimensional real space, e.g. the rotation of rigid bodies in classical mechanics. This is very easy for us to visualize.

- Let \( |\psi_1\rangle, |\psi_2\rangle \in \mathbb{C}^2 \). Denote \( |u_{\psi_1}\rangle = \hat{U} |\psi_1\rangle \) and \( |u_{\psi_2}\rangle = \hat{U} |\psi_2\rangle \), then

\[
\langle u_{\psi_1}|u_{\psi_2}\rangle = \langle \psi_1|\hat{U}^\dagger\hat{U}|\psi_2\rangle = \langle \psi_1|\hat{I}_{2\times2}|\psi_2\rangle = \langle \psi_1|\psi_2\rangle.
\]

If \( |\psi_1\rangle = |\psi_2\rangle \), then \( |u_{\psi_1}\rangle = |u_{\psi_2}\rangle \) and we have

\[
\langle u_{\psi_1}|u_{\psi_1}\rangle = \langle \psi_1|\psi_1\rangle.
\]

So \( \hat{U} \) preserves the magnitude of \( |\psi_1\rangle \).

As shown before, the angle between \( \hat{r}_\psi \) of \( |\psi\rangle \) and \( \hat{r}_\varphi \) of \( |\varphi\rangle \) on the Bloch sphere is

\[
\theta_{\psi\varphi} = 2\arccos(\sqrt{|\langle \psi|\varphi\rangle|^2}).
\]

Therefore,

\[
\langle u_{\psi_1}|u_{\psi_2}\rangle = \langle \psi_1|\psi_2\rangle \]

\[
\Rightarrow \theta_{u_{\psi_1}u_{\psi_2}} = \theta_{\psi_1\psi_2}.
\]

So \( \hat{U} \) also preserves the angle between \( \hat{r}_{\psi_1} \) and \( \hat{r}_{\psi_2} \) on the Bloch sphere as the angle between \( \hat{r}_{u_{\psi_1}} \) and \( \hat{r}_{u_{\psi_2}} \). This means that the distribution of states on the Bloch sphere is not distorted by the transformation. Applying \( \hat{U} \) simply rotates this distribution to another orientation. Then we know that the relative positions between \( |u_0\rangle, |u_1\rangle \) and \( |u_+\rangle \) on the Bloch sphere remain the same after their transformations from \( |0\rangle, |1\rangle \) and \( |+\rangle \) by \( \hat{U} \). This helped us to tomograph \( |u_+\rangle \).

By the Spectral Theorem of normal matrices\(^\text{15}\), \( \hat{U} \) is diagonalizable by a unitary

---

\(^{15}\) Matrix \( M \) is normal. \( \iff M^\dagger M = MM^\dagger \).
matrix. Since $\hat{U}^\dagger \hat{U} = \hat{I}_{2 \times 2}$,

$$\hat{U} = |b_0\rangle \langle b_0| + e^{i\varphi} |b_1\rangle \langle b_1| \text{ up to a global phase.}$$

$\varphi \in [0, 2\pi)$ and the eigenstates $\{|b_0\rangle, |b_1\rangle\}$ form a set of orthonormal basis. In this spectral form, it is clear that the action of $\hat{U}$ on the Bloch sphere is simply a rotation about the axis formed by $\hat{r}_{b_0}$ and $\hat{r}_{b_1}$ by angle $\varphi$. Any general rotation in 3D can be achieved this way by choosing the proper axis and angle of rotation.

In fact, we can factor the general parametrization of $\hat{U}$ as

$$\hat{U} = \begin{pmatrix} \cos(\frac{\theta_1}{2}) & \sin(\frac{\theta_1}{2}) e^{i\theta_2} \\ -\sin(\frac{\theta_1}{2}) e^{i\theta_3} & \cos(\frac{\theta_1}{2}) e^{i(\theta_2+\theta_3)} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta_3} \end{pmatrix} \begin{pmatrix} \cos(\frac{\theta_1}{2}) & \sin(\frac{\theta_1}{2}) \\ -\sin(\frac{\theta_1}{2}) & \cos(\frac{\theta_1}{2}) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta_2} \end{pmatrix}$$

Let

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta_3} \end{pmatrix} = |0\rangle \langle 0| + e^{i\theta_3} |1\rangle \langle 1| = \hat{U}_z(\theta_3),$$

$$\begin{pmatrix} \cos(\frac{\theta_1}{2}) & \sin(\frac{\theta_1}{2}) \\ -\sin(\frac{\theta_1}{2}) & \cos(\frac{\theta_1}{2}) \end{pmatrix} = e^{i\frac{\theta_1}{2}} (|+i\rangle \langle +i| + e^{-i\theta_1} |−i\rangle \langle −i|) = \hat{U}_y(−\theta_1)$$

and

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta_2} \end{pmatrix} = |0\rangle \langle 0| + e^{i\theta_2} |1\rangle \langle 1| = \hat{U}_z(\theta_2).$$

We have shown before that the $\hat{r}$ vectors of $\{|+\rangle, |−\rangle\}$, $\{|+i\rangle, |−i\rangle\}$ and $\{|0\rangle, |1\rangle\}$ are on the positive and negative $x$, $y$ and $z$–axis of the Bloch sphere, respectively. So clearly, the three matrices above perform rotations about the $y$ and $z$–axis by the angle in the parenthesis.
For aesthetic reasons, we may take $-\theta_1$ as $\theta_1$, then
\[
\hat{U}_y(\theta_1) = e^{-i \frac{\theta_1}{2}} \left( |+i\rangle \langle +i| + e^{i\theta_1} |-i\rangle \langle -i| \right)
\]
and the general parametrization of $\hat{U}$ becomes
\[
\hat{U} = \begin{pmatrix}
\cos\left(\frac{\theta_1}{2}\right) & -\sin\left(\frac{\theta_1}{2}\right) e^{i\theta_2} \\
\sin\left(\frac{\theta_1}{2}\right) e^{i\theta_3} & \cos\left(\frac{\theta_1}{2}\right) e^{i(\theta_2 + \theta_3)}
\end{pmatrix}
\]
where $\theta_1 \in [0, \pi]$ and $\theta_2, \theta_3 \in [0, 2\pi)$.

Then $\hat{U}$ can be factored as
\[
\hat{U} = \hat{U}_z(\theta_3) \hat{U}_y(\theta_1) \hat{U}_z(\theta_2).
\]

Therefore, $\hat{U}$ represents a general rotation in 3D space because any orientation in space can be achieved by the composition of extrinsic rotations $z(\theta_2) \rightarrow y(\theta_1) \rightarrow z(\theta_3)$ with $\theta_1 \in [0, \pi]$ and $\theta_2, \theta_3 \in [0, 2\pi)$. “Extrinsic” means that the axes of rotation are fixed to the observer and do not change during the rotation process. This is harder to visualize than “intrinsic” rotations where the axes of rotation are fixed to the object and rotates with it. As shown in Figure 3.7.1, the extrinsic axes are labeled by lower case $x, y$ and $z$ in blue. Starting from this position, the equivalent intrinsic rotation follows the sequence of $z(\theta_3) \rightarrow N(\theta_1) \rightarrow Z(\theta_2)$. The axes $z, N$ and $Z$ are colored in blue, green and red. The final orientation after rotation is marked by the red $X, Y$ and $Z$-axis.

Intrinsic rotations are easy to visualize. The equivalent extrinsic rotations are easier to compute because the rotation axes are fixed so that no coordinate transformations are required when multiplying, for example, $\hat{R}_y(\theta_1)$ and $\hat{R}_z(\theta_3)$.

$\theta_1 \in [0, \pi]$ and $\theta_2, \theta_3 \in [0, 2\pi)$ are the Euler angles of this rotation. As shown
in our QPT method, determining $|u_0\rangle$ fixes $\theta_1$ and $\theta_3$. In Figure 3.7.1, this means that $|0\rangle$ on the blue $z-$axis has been transformed to $|u_0\rangle$ on the red $Z-$axis. However, there still exists the degree of freedom to rotate about this red $Z-$axis that is not detectable from $|u_0\rangle$. Determining $\theta_2$ by tomographing $|u_+\rangle$ fixes this rotation component from the green $N-$axis to the red $X-$axis.

It is unfortunate that this visualization method cannot be easily extrapolated to unitary gates acting on 2 or more qubits. For example, 2-qubit unitary gates are elements of $SU(4)$, which is homomorphic to $SO(6)$. We are limited to visualize only 3–dimensional space so this is not an easy task any more. However, this does provide the perspective of studying QPT in the space of complex unitary matrices by looking at properties in its homomorphic space of real orthogonal matrices.
Chapter 4

Adaptive Quantum Tomography in Arbitrary Dimensions

In this section, we present an adaptive QST method for pure states and an adaptive QPT method for unitary gates. Both methods only assume the most general parametrization of the unknown pure state and unitary gate. Therefore, these methods are applicable to systems of any size.

4.1 Adaptive Pure State Tomography

Let \{ |n\rangle | n = 0, 1, ..., (d - 1) \} be a complete set of orthonormal basis that span the \( d \)-dimensional Hilbert space of interest \( \mathcal{H} \), i.e. \( \langle m|n\rangle = \delta_{mn} \) and \( \sum_{n=0}^{d-1} |n\rangle \langle n| = \hat{I}_{d \times d} \).

Then any pure state \( |\psi\rangle \) in \( \mathcal{H} \) can be expressed as a linear combination of all the \( |n\rangle \)'s, i.e.

\[
|\psi\rangle = \sum_{n=0}^{d-1} \alpha_n |n\rangle .
\]

The goal in pure state tomography is to determine all the \( \alpha_n \)'s, up to a global phase. This can be achieved by measuring the following Hermitian operators on \( |\psi\rangle \):
For each \( n \) and \( k = 0, 1, \ldots, (d - 1) \),

- \( D_n = \langle n | n \rangle \). The measured expectation value \( \langle D_n \rangle = \text{Tr}(\psi \langle n | D_n | \psi \rangle) = \langle \psi | D_n | \psi \rangle = \langle \psi | n \rangle \langle n | \psi \rangle = \alpha_n^* \alpha_n = |\alpha_n|^2 \) is the squared magnitude of each \( \alpha_n \);

- \( R_{nk} = \langle n | k \rangle + \langle k | n \rangle \). The measured expectation value \( \langle R_{nk} \rangle = \text{Tr}(\psi \langle n | R_{nk} | \psi \rangle) = \langle \psi | R_{nk} | \psi \rangle = \langle \psi | n \rangle \langle k | \psi \rangle + \langle \psi | k \rangle \langle n | \psi \rangle = \alpha_n^* \alpha_k + \alpha_k^* \alpha_n \). If \( \alpha_k \) is real, then \( \langle R_{nk} \rangle = 2 \alpha_k \text{Re}(\alpha_n) \);

- \( C_{nk} = i \left( |n \rangle \langle k | - |k \rangle \langle n | \right) \). The measured expectation value \( \langle C_{nk} \rangle = \text{Tr}(\psi \langle n | C_{nk} | \psi \rangle) = \langle \psi | C_{nk} | \psi \rangle = i \left( \langle \psi | n \rangle \langle k | \psi \rangle - \langle \psi | k \rangle \langle n | \psi \rangle \right) = i \left( \alpha_n^* \alpha_k - \alpha_k^* \alpha_n \right) \). If \( \alpha_k \) is real, then \( \langle C_{nk} \rangle = 2 \alpha_k \text{Im}(\alpha_n) \).

\( D_n \) measures the \((n + 1)\)-th diagonal element of a density matrix; \( R_{nk} \)'s and \( C_{nk} \)'s measure along the \((k + 1)\)-th row and column of a density matrix if we fix \( k \) and vary \( n \).

To use the least number of measurements in determining \( |\psi\rangle \), we arrange the above measurements in the following adaptive steps:

1. Sequentially measure \( D_0, D_1, \ldots \) up to \( D_{d-1} \) on \( |\psi\rangle \) until the first non-zero expectation value \( \langle D_k \rangle = |\alpha_k|^2 \neq 0 \) is obtained for some \( k \). We are guaranteed to find such a \( k \). Otherwise, the state is trivial, i.e. \( |\psi\rangle = 0 \). This gives us the smallest \( k \) such that \( \alpha_k \neq 0 \). By the freedom to choose global phase, we can set this \( \alpha_k \) to be real and positive. Then \( \alpha_k = \sqrt{\langle D_k \rangle} \). This step costs \( k + 1 \) measurements. Upon finding this \( k \), we know our state is now \( |\psi\rangle = \sum_{n=k}^{d-1} \alpha_n |n\rangle \).

2. Measure each pair of \( R_{nk} \) and \( C_{nk} \) on \( |\psi\rangle \) for all \( n \in [k + 1, d - 1] \). Since we know that \( \alpha_k \neq 0 \) from Step 1, each pair of \( \langle R_{nk} \rangle \) and \( \langle C_{nk} \rangle \) values fix each
\( \alpha_n : \text{Re}(\alpha_n) = \langle R_{nk} \rangle / 2\alpha_k \) and \( \text{Im}(\alpha_n) = \langle C_{nk} \rangle / 2\alpha_k \). This step costs a total of \( 2(d - k - 1) \) measurements.

In total, \( 2d - k - 1 \) measurements are used to determine \( |\psi\rangle \). In the worst case, we will measure \( D_0 \neq 0 \) and must use \( 2d - 1 \) measurements to fully fix \( |\psi\rangle \). In terms of density matrices, these measurements fix the \((k + 1)\)-th row and column of \( |\psi\rangle \langle \psi| \). For convenience, we denote this set of measured Hermitian operators as \( \{M_k\} = \{D_l, R_{nk}, C_{nk}\} | l = 0, 1, \ldots, k; n = k + 1, k + 2, \ldots, d - 1 \} \) and the measured expectation values of them on state \( |\psi\rangle \) as \( \langle \{M_k\} \rangle_{\psi} \).

Clearly, this method fixes all complex coefficients \( \alpha_n \)'s, up to a global phase, of any general \( d \)-dimensional pure state \( |\psi\rangle \). \( \langle \{M_k\} \rangle_{\rho} = \langle \{M_k\} \rangle_{\psi} \Rightarrow |\phi\rangle = e^{i\beta} |\psi\rangle \). In other words, the method uniquely identifies a pure state among all pure states, i.e. UDP. It is worth to notice that the method does not utilize the normalization condition \( \langle \psi|\psi\rangle = 1 \). So \( \langle \{M_k\} \rangle_{\psi} \) uniquely identifies a general vector \( |\Psi\rangle \) of any magnitude.

More over, this method is also UDA, i.e. if there exists a pure state density matrix \( |\psi\rangle \langle \psi| \) that yields the measured expectation values of \( \{M_k\} \), then no other pure or mixed density matrices will measure the same expectation values of \( \{M_k\} \).

If \( \langle \{M_k\} \rangle_{\rho} = \langle \{M_k\} \rangle_{\psi} \), then \( \rho = |\psi\rangle \langle \psi| \).

This can be proved by constructing a positive semi-definite partition of a general density matrix \( \rho \). We first outline the major steps in this proof and then go into necessary details of each step.

1. If \( \langle \{D_l\} \rangle_{\rho} = \langle \{D_l\} \rangle_{\psi} \), then \( \rho \) is only non-zero starting from its \((k + 1)\)-th row and column, i.e. in the lower right diagonal block starting from the \((k + 1)\)-th diagonal element \( \langle D_k \rangle_{\rho} = \langle k|\rho|k \rangle \).
2. The \((k+1)\)-th column of \(\rho\) is \(\rho |k\rangle\). The magnitude of this column vector is \(\langle k| \rho^2 |k\rangle^{\frac{1}{2}}\) and its normalized version is \(|\psi_k\rangle = \langle k| \rho^2 |k\rangle^{-\frac{1}{2}} \rho |k\rangle\). Since \(\langle k| \rho |k\rangle \neq 0\), \(\langle k| \psi_k\rangle = \langle k| \rho^2 |k\rangle^{-\frac{1}{2}} \langle k| \rho |k\rangle\) is a positive number and thus can be used as a denominator.

3. Use the \((k+1)\)-th column of \(\rho\) to construct its partition as

\[
\rho = \frac{\langle k| \rho^2 |k\rangle^{\frac{1}{2}}}{\langle k|\psi_k\rangle} |\psi_k\rangle \langle \psi_k| + \delta_k, \quad \text{where} \quad \delta_k = \rho - \frac{\langle k| \rho^2 |k\rangle^{\frac{1}{2}}}{\langle k|\psi_k\rangle} |\psi_k\rangle \langle \psi_k|.
\]

The \((k+1)\)-th row and column of \(\frac{\langle k| \rho^2 |k\rangle^{\frac{1}{2}}}{\langle k|\psi_k\rangle} |\psi_k\rangle \langle \psi_k|\) is equal to that of \(\rho\) so that non-zero contents of \(\delta_k\) starts at its \((k+2)\)-th row and column.

4. Therefore, \(\langle \{M_k\}\rangle_{\delta_k} = 0\). Absorb the positive factor \(\frac{\langle k| \rho^2 |k\rangle^{\frac{1}{2}}}{\langle k|\psi_k\rangle}\) into \(|\psi_k\rangle\) as

\[
|\Psi_k\rangle = \left(\frac{\langle k| \rho^2 |k\rangle^{\frac{1}{2}}}{\langle k|\psi_k\rangle}\right) |\psi_k\rangle \quad \text{to write} \quad \rho = |\Psi_k\rangle \langle \Psi_k| + \delta_k, \quad \text{we have} \quad \langle \{M_k\}\rangle_{\rho} = \langle \{M_k\}\rangle_{\Psi_k} + \langle \{M_k\}\rangle_{\delta_k} = \langle \{M_k\}\rangle_{\Psi_k}.
\]

Then if there exists a pure state \(|\psi\rangle\) such that \(\langle \{M_k\}\rangle_{\rho} = \langle \{M_k\}\rangle_{\Psi_k}\), by the uniqueness of \(\langle \{M_k\}\rangle_{\Psi}\) with respect to \(|\Psi\rangle\) of any magnitude, \(|\Psi_k\rangle = e^{i\beta} |\psi\rangle\). So \(\rho = |\psi\rangle \langle \psi| + \delta_k\).

5. We show that \(\delta_k\) is a positive semi-definite matrix and hence must be expressed in the form of \(\delta_k = \sum_j p_j |\psi_j\rangle \langle \psi_j|\) with \(p_j \geq 0 \forall j\). Then the trace-1 constraint on density matrices will enforce that

\[
Tr(\rho) = Tr(|\psi\rangle \langle \psi|) + Tr(\delta_k) = 1 + \sum_j p_j Tr(|\psi_j\rangle \langle \psi_j|) = 1 + \sum_j p_j = 1
\]

\[
\Rightarrow \sum_j p_j = 0. \quad \text{But} \quad p_j \geq 0 \forall j, \quad \text{so we must have each} \quad p_j = 0.
\]

Therefore, \(\delta_k = 0\) and we finally have \(\rho = |\psi\rangle \langle \psi|\).

Next we present the details in each step:
1. By the positive semi-definiteness of all density matrices, \( \rho = \sum_j p_j |\psi_j\rangle \langle \psi_j| \), \( p_j > 0 \forall j \). Then if diagonal element \( \langle l | \rho | l \rangle = \sum_j p_j \langle l | \psi_j \rangle \langle \psi_j | l \rangle = 0 \), \( \langle \psi_j | l \rangle = 0 \) \( \forall j \). Therefore, \( \langle n | \rho | l \rangle = \sum_j p_j \langle n | \psi_j \rangle \langle \psi_j | l \rangle = 0 = \langle l | \rho | n \rangle \ \forall |n\rangle \), i.e. the row and column containing this diagonal element are all zeros.

2. If \( \langle k | \rho | k \rangle \neq 0 \), \( \langle k | \rho | k \rangle > 0 \) since \( \rho \) is positive semi-definite. \( \langle k | \rho^2 | k \rangle = \langle k | \rho \hat{I}_{d \times d} \rho | k \rangle = \langle k | \rho \sum_{n=0}^{d-1} |n\rangle \langle n | \rho | k \rangle = \langle k | \rho | k \rangle^2 + \sum_{n \neq k} |\langle k | \rho | n \rangle|^2 > 0 \). Hence, \( \langle k | \rho^2 | k \rangle - \frac{1}{2} \langle k | \rho | k \rangle > 0 \).

3. This can be checked by substituting \( |\psi_k\rangle = \langle k | \rho^2 | k \rangle^{-\frac{1}{2}} \rho |k\rangle \) into \( \langle n | \delta_k | k \rangle \):

\[
\langle n | \delta_k | k \rangle = \langle n | \rho | k \rangle - \frac{\langle k | \rho^2 | k \rangle}{\langle k | \psi_k \rangle} \langle n | \psi_k \rangle \langle \psi_k | k \rangle = \langle n | \rho | k \rangle - \langle k | \rho^2 | k \rangle^{\frac{1}{2}} \langle n | \psi_k \rangle = \\
\langle n | \rho | k \rangle - \langle k | \rho^2 | k \rangle^{\frac{1}{2}} \langle k | \rho^2 | k \rangle^{-\frac{1}{2}} \langle n | \rho | k \rangle = \langle n | \rho | k \rangle - \langle n | \rho | k \rangle = 0 = \langle k | \delta_k | n \rangle \ \forall |n\rangle.
\]

4. \( \{D_l\} \) only measures up to the \((k + 1)\)-th diagonal element; \( \{R_{nk}, C_{nk}\} \) only measures along the \((k + 1)\)-th row and column of a matrix. So \( \langle \{M_k\} \rangle_{\delta_k} = \langle \{D_l\} \rangle_{\delta_k} \cup \langle \{R_{nk}, C_{nk}\} \rangle_{\delta_k} = 0 \).

5. We show by direct computation that \( \delta_k \) is indeed positive semi-definite, i.e. \( \langle \varphi | \delta_k | \varphi \rangle \geq 0 \ \forall |\varphi\rangle \).

Clearly, \( \delta_k \) is not negative in any direction orthogonal to \( |\psi_k\rangle \):

\[
\langle \psi_k^\perp | \delta_k | \psi_k^\perp \rangle = \langle \psi_k^\perp | \rho | \psi_k^\perp \rangle - \frac{\langle k | \rho^2 | k \rangle}{\langle k | \psi_k \rangle} \langle \psi_k^\perp | \psi_k \rangle \langle \psi_k | \psi_k^\perp \rangle = \langle \psi_k^\perp | \rho | \psi_k^\perp \rangle \geq 0.
\]

So we only need to show that \( \delta_k \) is not negative in the \( |\psi_k\rangle \) direction:

\[
\langle \psi_k | \delta_k | \psi_k \rangle = \langle \psi_k | \rho | \psi_k \rangle - \frac{\langle k | \rho^2 | k \rangle}{\langle k | \psi_k \rangle} \geq 0.
\]
Substituting in $|\psi_k\rangle = \langle k| \rho^2|k\rangle^{-\frac{1}{2}} \rho |k\rangle$, we must show that

$$\frac{\langle k| \rho^3|k\rangle}{\langle k| \rho^2|k\rangle} - \frac{\langle k| \rho^2|k\rangle}{\langle k| \rho|k\rangle} \geq 0.$$ 

Any density matrix $\rho$ is positive semi-definite and permits a spectrum decomposition

$$\rho = \sum_{s=0}^{d-1} p_s |s\rangle \langle s|, \text{ with each } p_s \geq 0 \text{ and orthonormal } \{|s\}\}.$$ 

Then it is easy to compute that

$$\rho^x = \sum_{s=0}^{d-1} p_s^x |s\rangle \langle s|, \ x \in \mathbb{N}^+.$$ 

So we must show

$$\langle \psi_k| \delta_k |\psi_k\rangle = \frac{\sum_{s=0}^{d-1} p_s^3 \langle k|s\rangle \langle s|k\rangle}{\sum_{s=0}^{d-1} p_s^2 \langle k|s\rangle \langle s|k\rangle} - \frac{\sum_{s=0}^{d-1} p_s^2 \langle k|s\rangle \langle s|k\rangle}{\sum_{s=0}^{d-1} p_s \langle k|s\rangle \langle s|k\rangle} \geq 0.$$ 

We may set $\langle k|s\rangle \langle s|k\rangle = b_s^2$, where $b_s \in \mathbb{R}$. Then we must show that

$$\frac{\sum p_s^3 b_s^2}{\sum p_s^2 b_s^2} - \frac{\sum p_s^2 b_s^2}{\sum p_s b_s^2} \geq 0.$$ 

Cross multiply by the two positive denominators to get

$$\sum_{s,r} p_s^3 p_r b_s^2 b_r^2 - \sum_{s,r} p_s^2 p_r^2 b_s^2 b_r^2 \geq 0.$$ 

$$\sum_{s,r} (p_s^3 p_r - p_s^2 p_r^2) b_s^2 b_r^2 \geq 0.$$ 

\[16\text{Notice that } |k\rangle \in \{|n\}\} \text{ is a basis we choose to measure. } \{|s\}\} \text{ is the eigen basis that diagonalizes some } \rho. \text{ In general, } \langle k|s\rangle \neq \delta_{ks}.\]
Clearly, the summand \((p_s^3 p_r - p_s^2 p_r^2)b_s^2 b_r^2\) = 0 when \(s = r\). So we are left to show

\[
\sum_{s \neq r} (p_s^3 p_r - p_s^2 p_r^2)b_s^2 b_r^2 > 0.
\]

Swapping \(s\) and \(r\) only changes the term \(p_s^3 p_r\) to \(p_r^3 p_s\). So we may only sum over the combinations of \(s\) and \(r\) as

\[
\sum_{s < r} (p_s^3 p_r - 2p_s^2 p_r^2 + p_s^3 p_s)b_s^2 b_r^2 \geq 0
\]

\[
\sum_{s < r} p_s p_r (p_s^2 - 2p_s p_r + p_r^2)b_s^2 b_r^2 \geq 0
\]

\[
\sum_{s < r} p_s p_r (p_s - p_r)^2 b_s^2 b_r^2 \geq 0,
\]

which is clearly true since \(p_s, p_r \geq 0\) \(\forall s\) and \(r\). Therefore, \(\delta_k\) is a positive semi-definite matrix.

It has been brought to the author’s attention during the defense of this thesis that a different proof of this result has been given in [47], almost simultaneously with our submission for publication. The overall idea of the proof is the same as ours: By the trace-1 constraint, the difference \(\rho - |\psi\rangle \langle \psi|\) must be traceless. However, the positive semi-definite constraint of \(\rho\) requires this difference must also be positive semi-definite and therefore can only be zero. They were able to sidestep some calculations by using the properties of the Shur complement [52] and matrix inertia [53].
Adaptive Unitary Process Tomography

Recall that a general $d \times d$ unitary operator $\hat{U}$ can be expressed as a change of basis

$$\hat{U} = \sum_{n=0}^{d-1} |u_n\rangle \langle n|,$$

where $\{|n\rangle\}$ can be any complete set of orthonormal basis of the Hilbert space and $\{|u_n\rangle\}$ is its image basis from $\hat{U}$.

The complete tomography of $\hat{U}$ cannot only determine each $|u_n\rangle$ up to an arbitrary overall phase because this phase factor will become a relative phase between components of the output state if $\hat{U}$ acts on an input state that is in a superposition of the $|n\rangle$'s. For example, suppose $\hat{U} = |0\rangle \langle 0| + i |1\rangle \langle 1| + \sum_{n=2}^{d-1} |u_n\rangle \langle n|$, i.e., $|u_0\rangle = |0\rangle$, $|u_1\rangle = i |1\rangle$. If we were to only fix each $|u_n\rangle$ up to an arbitrary relative phase, we may reconstruct the unitary as $\hat{U}' = |0\rangle \langle 0| + |1\rangle \langle 1| + \sum_{n=2}^{d-1} |u_n\rangle \langle n|$. Clearly, $\hat{U} \neq \hat{U}'$ and $\hat{U}|+\rangle = |+\rangle = \hat{U}'|+\rangle$. But as a starting point, we may choose the global phase of $\hat{U}$ such that $|u_0\rangle = \hat{U} |0\rangle$ has its first non-zero coefficient to be real and positive. This allows us to utilize the adaptive pure state tomography method presented in the previous section to determine $|u_0\rangle$.

In determining all the other $|u_n\rangle$'s, their overall phase relative to $|u_0\rangle$ must also be included. We capture this relative phase between two columns of $\hat{U}$ using superposition input states $\frac{1}{\sqrt{2}}(|0\rangle + |n\rangle)$ for $n = 1, 2, ..., (d-1)$. Then the adaptive pure state tomography method uniquely identifies each output state $\frac{1}{\sqrt{2}}(|u_0\rangle + |u_n\rangle) = \hat{U} \frac{1}{\sqrt{2}}(|0\rangle + |n\rangle)$.

\footnote{Strictly speaking, due to Born's rule of squaring the probability amplitudes, the physically measurable quantity of a pure state's wavefunction $|\psi\rangle$ is its density matrix $|\psi\rangle \langle \psi|$. Here we show that $|\varphi\rangle = e^{i\beta} |\psi\rangle \iff |\varphi\rangle \langle \varphi| = |\psi\rangle \langle \psi|$. Trivially, $|\varphi\rangle = e^{i\beta} |\psi\rangle \Rightarrow |\varphi\rangle \langle \varphi| = |\psi\rangle \langle \psi|$. Conversely, if $|\varphi\rangle \langle \varphi| = |\psi\rangle \langle \psi|$, we can have $|\varphi\rangle \langle \varphi| = |\psi\rangle \langle \psi| \Rightarrow |\varphi\rangle = \frac{\langle \varphi|\psi\rangle}{\langle \psi|\varphi\rangle} |\psi\rangle$; we can also have...}
tive phases, since we already know $|u_0\rangle$.

In determining $|u_1\rangle$, we may choose to measure in basis

$$\{|u_0\rangle, |n\rangle | n = 1, 2, \ldots, (d-1)\}.$$ 

Then by the orthonormality of $\{|u_n\rangle\}$, we are guaranteed to have the first coefficient of $\frac{1}{\sqrt{2}} (|u_0\rangle + |u_1\rangle)$ as

$$\langle u_0 | \frac{1}{\sqrt{2}} (|u_0\rangle + |u_1\rangle) = \frac{1}{\sqrt{2}}$$

and

$$|u_1\rangle \in \text{span} \{|n\rangle | n = 1, 2, \ldots, (d-1)\}$$

only occupies the coefficients of the other $(d-1)$ basis. So in this basis $\frac{1}{\sqrt{2}} (|u_0\rangle + |u_1\rangle)$ partitions in a very neat way. There is no need to measure $\{D_i\}$ up to some non-zero $\langle D_k\rangle$ since we know that $\langle D_0\rangle = \langle D_{u_0}\rangle = \frac{1}{2} \neq 0$. Then we simply measure

$$\left\{ R_{nu_0} = |n\rangle \langle u_0| + |u_0\rangle \langle n|, C_{nu_0} = i \left(|n\rangle \langle u_0| - |u_0\rangle \langle n| \right) | n = 1, 2, \ldots, (d-1) \right\}$$

to determine $|u_1\rangle$ and its relative phase to $|u_0\rangle$.

We continue to capitalize on this orthonormality condition in determining the rest of $\{|u_n\rangle\}$. For example, in determining $|u_2\rangle$ we choose to measure in basis

$$\{|u_0\rangle, |u_1\rangle, |n\rangle | n = 2, 3, \ldots, (d-1)\}.$$ 

\[
\langle \psi | \varphi \rangle \langle \varphi | \psi \rangle = \langle \psi | \psi \rangle \langle \varphi | \varphi \rangle \Rightarrow \langle \varphi | \psi \rangle = e^{-i\beta} \langle \psi | \psi \rangle. \text{ Therefore, } |\varphi\rangle = \frac{\langle \psi | \psi \rangle}{e^{-i\beta} \langle \psi | \psi \rangle} |\psi\rangle = e^{i\beta} |\psi\rangle. \text{ Notice that this proof holds for kets } |\psi\rangle \text{ and } |\varphi\rangle \text{ of any magnitude.}
\]

87
Then we know that

$$|u_2\rangle \in \text{span} \{ |n\rangle \mid n = 2, 3, \ldots, (d - 1) \}$$

and the 1st and 2nd coefficients of $\frac{1}{\sqrt{2}} (|u_0\rangle + |u_2\rangle)$ are $\langle u_0 | \frac{1}{\sqrt{2}} (|u_0\rangle + |u_2\rangle) = \frac{1}{\sqrt{2}}$ and $\langle u_1 | \frac{1}{\sqrt{2}} (|u_0\rangle + |u_2\rangle) = 0$. We still only measure the coherence $\{ R_{nu_0}, C_{nu_0} \}$, but now with $n = 2, 3, \ldots, (d - 1)$.

We summarize the above adaptive unitary process tomography steps below:

1. Use the adaptive pure state tomography method on $|u_0\rangle = \hat{U} |0\rangle$ to fully determine it up to a global phase. This step costs $2d - 1$ measurements at most;

2. For each $j = 1, 2, \ldots, d-1$, obtain output state $\frac{1}{\sqrt{2}} (|u_0\rangle + |u_j\rangle) = \hat{U} \frac{1}{\sqrt{2}} (|0\rangle + |j\rangle)$ and measure

$$\left\{ R_{nu_0} = |n\rangle \langle u_0| + |u_0\rangle \langle n|, C_{nu_0} = i \left( |n\rangle \langle u_0| - |u_0\rangle \langle n| \right) \mid n = j, \ldots, (d - 1) \right\}$$

in an incomplete set of orthonormal basis $|n\rangle \in \text{span} \{ |u_0\rangle, |u_1\rangle, \ldots, |u_{j-1}\rangle \}^{\perp}$ to determine $|u_j\rangle$ and its relative phase to $|u_0\rangle$. This step costs $2d - 2j$ measurements for each $j$.

In total, we need to measure $(2d - 1) + (2d - 2) + (2d - 4) + \ldots + 2 = d^2 + d - 1$ observables to fully determine a $d \times d$ unitary operator $\hat{U}$.

Recall that we have counted before that a $d \times d$ unitary operator $\hat{U}$ has $d^2 - 1$ independent real parameters. The above method uses $d$ extra measurements to determine $\hat{U}$. This is because the $d$ normalization conditions $\langle u_n | u_n \rangle = 1$ are not used for any $|u_n\rangle$. Measuring $R_{nu_0}$ and $C_{nu_0}$ determines the real and imaginary parts of a matrix element. The magnitude and phase information of this element is
encoded within the combination. This makes it inconvenient to use the normalization condition because \( r^2 + c^2 = 1 \) only constrains \( r \) and \( c \) up to an important \( \pm \) sign. Whereas if we were to parametrize the element in terms of magnitude and phase, we can require the magnitude to be non-negative and transfer this \( \pm \) sign to the phase. Although it is natural to think in terms of magnitude and phase in physics because there are physical meanings attached to them, it is more economic to use real and imaginary parts in tomography. As demonstrated, we can obtain the real and imaginary parts of each matrix element with only two measurements. But we must use three measurements if we were to fix the same element by its magnitude and phase. Because we generally need to measure both the cosine and sine values to fix a phase angle. Therefore, sacrificing the normalization constraints to measure real and imaginary parts is still a worthy trade-off. For the same reason, the adaptive pure state tomography method uses \( 2d - 1 \) measurements to fix a \( d \)-dimensional pure state with only \( 2d - 2 \) independent real parameters.\(^{18}\)

\(^{18}\)A \( d \)-dimensional pure state can vary in magnitude and phase of its \( d \) coefficients, subject to the normalization constraint that fixes one magnitude value. We can also choose its global phase such that one of the coefficients has zero phase. So, it contains a total of \( 2d - 2 \) free parameters.
Chapter 5

Application to 2-Qubit NMR Systems

In this section, we apply the method of adaptive unitary process tomography presented previously to completely determine 2-qubit unitary gates. The success also demonstrates the feasibility of the presented method of adaptive pure state tomography since it is the first step in the process tomography method. Our experiments are based on the NMR system.

5.1 Standard Quantum Process Tomography

To establish a baseline of comparison, we first perform the standard quantum process tomography procedure on the same set of 2-qubit unitary gates that will be used in adaptive unitary process tomography.

Often in NMR and many other ensemble systems, it is only convenient to measure Pauli observables. For 2-qubit systems, the Pauli observables are a set of $16 \ 4 \times 4$
Hermitian matrices $\sigma = (XX, XY, ..., IZ, II)$ \(^{19}\). It is easy to check that they are mutually orthogonal with respect to the Hilbert-Schmidt inner product. Therefore, they form a complete set of basis for the 16-dimensional space of $4 \times 4$ complex matrices. Any general $4 \times 4$ matrix $m$ can be expanded as a unique linear combination of these 16 basis $m = p_1 XX + p_2 XY + ... + p_{15} IZ + p_{16} II$. More compactly, $m = V_m \odot \sigma$ \(^{20}\) where the coefficient vector $V_m = (p_1, p_2, ..., p_{15}, p_{16})$ contains all information about $m$ and is called the Pauli vector of $m$. Generally, elements of $V_m$ are complex numbers. If $m$ is Hermitian, then $V_m$ is a real vector. Since the first 15 basis are all trace zero, $Tr(m) = p_{16} Tr(II) = 4p_{16} \Rightarrow p_{16} = \frac{1}{4} Tr(m)$. For example, the Pauli vector of a general 2-qubit density matrix $\rho$ is $V_\rho = (p_1, p_2, ..., p_{15}, \frac{1}{4})$, where $p_1, p_2, ..., p_{15}$ are real numbers.

In standard quantum process tomography, only the most general form of the underlying quantum operation is assumed. So all quantum operations are represented as CPTP linear maps. Let $\Lambda$ label a CPTP linear map. Then by the completeness of $\sigma$ and the linearity of $\Lambda$, determining each output

$$\Lambda(\sigma) = (\Lambda(XX), \Lambda(XY), ..., \Lambda(IZ), \Lambda(II))$$

completely characterizes the action of $\Lambda$ on the matrix space. We use superscripts 1, 2, ..., 15, 16 to differentiate the same elements of the Pauli vectors of

$$\Lambda(XX), \Lambda(XY), ..., \Lambda(IZ), \Lambda(II).$$

\(^{19}\)For example, $XX = \sigma_x \otimes \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$. \(^{20}\)Here, $\odot$ is the usual vector dot product. $(a, b, c) \odot (x, y, z) = ax + by + cz$. 

91
For example, $\Lambda(XX) = p_1^1 XX + p_1^2 XY + \ldots + p_{15}^{15} IZ + p_{16}^{16} II$ and $\Lambda(IZ) = p_1^{15} XX + p_2^{15} XY + \ldots + p_{15}^{15} IZ + p_{16}^{15} II$. \{p_j^i | i, j = 1, 2, \ldots, 15, 16\} are real numbers since a completely positive linear map $\Lambda(A) = \sum_i K_i A K_i^\dagger$ maps Hermitian operators to Hermitian operators. Since $\Lambda$ is a trace preserving map $Tr(\Lambda(\sigma)) = Tr(\sigma)$, $p_{16}^i = 0$ for $i = 1, 2, \ldots, 15$ and $p_{16}^{16} = 1$.

The **Pauli transfer matrix** $\hat{P}_\Lambda$ of map $\Lambda$ is constructed using the 16 Pauli vectors of $\Lambda(\sigma)$ as columns:

$$
\hat{P}_\Lambda = \begin{pmatrix}
p_1^1 & p_1^2 & \ldots & p_{15}^1 & p_{16}^1 \\
p_1^2 & p_2^2 & \ldots & p_{15}^2 & p_{16}^2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
p_{15}^1 & p_{15}^2 & \ldots & p_{15}^{15} & p_{15}^{16} \\
0 & 0 & \ldots & 0 & 1
\end{pmatrix}.
$$

Using $\hat{P}_\Lambda$, computing output $\Lambda(m)$ becomes simple because

$$V_{\Lambda(m)} = \hat{P}_\Lambda V_m$$

and we simply have $\Lambda(m) = V_{\Lambda(m)} \odot \sigma$. Clearly, determining $\hat{P}_\Lambda$ completely characterizes the 2-qubit CPTP linear map $\Lambda$.

Therefore, the standard quantum process tomography of $\Lambda$ on an NMR system can be carried out as the following:

- Input the perfectly mixed state $\rho_0 = \frac{1}{4} II$ to obtain the output state $\Lambda(\rho_0) = \frac{1}{4} \Lambda(II)$;

- Measure the expectation values of all 15 nontrivial Pauli observables to obtain $p_{16}^1 = Tr(\Lambda(\rho_0)XX), p_{16}^2 = Tr(\Lambda(\rho_0)XY), \ldots, p_{16}^{16} = Tr(\Lambda(\rho_0)IZ)$;

92
• Input mixed state $\rho_1 = \frac{1}{4}(II + XX)$ to obtain the output state $\Lambda(\rho_1) = \frac{1}{4} \Lambda(II) + \frac{1}{4} \Lambda(XX);$  

• Measure the expectation values of all 15 non-trivial Pauli observables to obtain
  
  \[ p_1 = \text{Tr}(\Lambda(\rho_1)XX) - p_{16}^1, \quad p_2 = \text{Tr}(\Lambda(\rho_1)XY) - p_{16}^2, \quad \ldots, \quad p_{15} = \text{Tr}(\Lambda(\rho_1)IZ) - p_{16}^{15}; \]

• Repeat the two previous steps for $\rho_2 = \frac{1}{4}(II + XY),$ $\ldots,$ $\rho_{15} = \frac{1}{4}(II + IZ);$  

In total, 16 outputs $\Lambda(\sigma)$ are generated. For each output, 15 non-trivial Pauli observables are measured. So a total of $16 \times 15 = 240$ measurements are performed, which is equal to $4^4 - 4^2 = d^4 - d^2$ as formulated in [15, Section 8.4.2] using the $\chi$-matrix. The standard process tomography presented here using Pauli transfer matrix is equivalent to the method presented in [15, Section 8.4.2] using the $\chi$-matrix. As shown before in Chapter 2, the $\chi$-matrix is equivalent to the Choi matrix. As will be shown below, the Choi matrix $C_\Lambda$ of map $\Lambda$ has an one-to-one correspondence with the Pauli transfer matrix $\hat{P}_\Lambda$ and can be reconstructed from elements of $\hat{P}_\Lambda$.

Recall that the Choi matrix of $\Lambda$ is

\[ C_\Lambda = \sum_{i,j=1}^{4} \hat{E}_{ij} \otimes \Lambda(\hat{E}_{ij}) \]

where $(\hat{E}_{ij})_{ab} = \delta_{ia} \delta_{jb}, \ i, j, a, b = 1, 2, 3, 4.$

\[ ^{21} \text{Experimentalists working on NMR will colloquially refer to } \rho_1 \text{ as the } XX \text{ state. Obviously, this is } \text{strictly speaking incorrect since } XX \text{ is not trace one or positive semi-definite. In an actual NMR equipment, the vast majority of the molecules in the ensemble will always be in the maximally mixed state } \rho_0. \text{ For safety concerns, only weak magnetic fields are used to manipulate the ensemble. Upon adding this weak magnetic field, only a very small portion of these molecules will become polarized i.e. they will transit from the negative to the positive eigenstate of } X. \text{ This small deviation from } \rho_0 \text{ is orders of magnitude smaller than the rest of the ensemble population that are still maximally mixed. But this small deviation is the only part that will measure a non-zero expectation value of any Pauli observable. Since this deviation is very small, there is always enough molecules in the maximally mixed bulk to be added to the polarized population to form a subsystem ensemble. Then upon normalization, we describe this subsystem ensemble as the mixed state density matrix } \rho_1 = \frac{1}{4}(II + XX). \]
To reconstruct $C_\Lambda$ from $\hat{P}_\Lambda$, we first divide $\sigma = (XX, XY, ..., IZ, II)$ into four subgroups. They are $\sigma_1 = (ZZ, ZI, IZ, II)$, $\sigma_2 = (XX, XY, YX, YY)$, $\sigma_3 = (XZ, XI, YZ, YI)$ and $\sigma_4 = (ZX, ZY, IX, IY)$. It is easy to observe that, within each subgroup, all matrices have all of their non-zero elements at the same positions. All these non-zero elements have unit magnitude and only differs in sign or complex phase.

Because all non-zero elements are at the same positions, all matrices in each subgroup can be spanned by the same set of $\hat{E}_{ij}$'s. We label the spanning set of $\hat{E}_{ij}$'s for each set $\sigma_n$ as the corresponding $\varepsilon_n$. Then, $\varepsilon_1 = (\hat{E}_{11}, \hat{E}_{22}, \hat{E}_{33}, \hat{E}_{44})$, $\varepsilon_2 = (\hat{E}_{14}, \hat{E}_{23}, \hat{E}_{32}, \hat{E}_{41})$, $\varepsilon_3 = (\hat{E}_{13}, \hat{E}_{24}, \hat{E}_{31}, \hat{E}_{42})$ and $\varepsilon_4 = (\hat{E}_{12}, \hat{E}_{21}, \hat{E}_{34}, \hat{E}_{43})$.

It can be shown that the portions of the Choi matrix $C_\Lambda$ generated by the $\hat{E}_{ij}$'s within the same set $\varepsilon_n$ can also be generated by the Pauli observables within the corresponding set $\sigma_n$. We may use the pair $\varepsilon_1$ and $\sigma_1$ as an example. The process is similar for the other three pairs.

Obviously, for the four matrices in $\sigma_1$, we have

\[
ZZ = \hat{E}_{11} - \hat{E}_{22} - \hat{E}_{33} + \hat{E}_{44};
\]

\[
ZI = \hat{E}_{11} + \hat{E}_{22} - \hat{E}_{33} - \hat{E}_{44};
\]

\[
IZ = \hat{E}_{11} - \hat{E}_{22} + \hat{E}_{33} - \hat{E}_{44};
\]

\[
II = \hat{E}_{11} + \hat{E}_{22} + \hat{E}_{33} + \hat{E}_{44}.
\]

Then,

\[
ZZ \otimes \Lambda(ZZ) = \left(\hat{E}_{11} - \hat{E}_{22} - \hat{E}_{33} + \hat{E}_{44}\right) \otimes \left(\Lambda(\hat{E}_{11}) - \Lambda(\hat{E}_{22}) - \Lambda(\hat{E}_{33}) + \Lambda(\hat{E}_{44})\right);
\]
\[ ZI \otimes \Lambda(ZI) = \left( \hat{E}_{11} + \hat{E}_{22} - \hat{E}_{33} - \hat{E}_{44} \right) \otimes \left( \Lambda(\hat{E}_{11}) + \Lambda(\hat{E}_{22}) - \Lambda(\hat{E}_{33}) - \Lambda(\hat{E}_{44}) \right) ; \]
\[ IZ \otimes \Lambda(IZ) = \left( \hat{E}_{11} - \hat{E}_{22} + \hat{E}_{33} - \hat{E}_{44} \right) \otimes \left( \Lambda(\hat{E}_{11}) - \Lambda(\hat{E}_{22}) + \Lambda(\hat{E}_{33}) - \Lambda(\hat{E}_{44}) \right) ; \]
\[ II \otimes \Lambda(II) = \left( \hat{E}_{11} + \hat{E}_{22} + \hat{E}_{33} + \hat{E}_{44} \right) \otimes \left( \Lambda(\hat{E}_{11}) + \Lambda(\hat{E}_{22}) + \Lambda(\hat{E}_{33}) + \Lambda(\hat{E}_{44}) \right) . \]

By direct computation, we have

\[ \hat{E}_{11} \otimes \Lambda(\hat{E}_{11}) + \hat{E}_{22} \otimes \Lambda(\hat{E}_{22}) + \hat{E}_{33} \otimes \Lambda(\hat{E}_{33}) + \hat{E}_{44} \otimes \Lambda(\hat{E}_{44}) \]
\[ = \frac{1}{4} (Z \otimes \Lambda(Z) + Z \otimes \Lambda(Z) + I \otimes \Lambda(I) + I \otimes \Lambda(I)). \]

Cross terms like \( \hat{E}_{22} \otimes \Lambda(\hat{E}_{11}) \) all cancel out during summation due to alternating negative signs in the Pauli matrices.

Similarly, for the other three pairs of \( \varepsilon \) and \( \sigma \), we have

\[ \hat{E}_{14} \otimes \Lambda(\hat{E}_{14}) + \hat{E}_{23} \otimes \Lambda(\hat{E}_{23}) + \hat{E}_{32} \otimes \Lambda(\hat{E}_{32}) + \hat{E}_{41} \otimes \Lambda(\hat{E}_{41}) \]
\[ = \frac{1}{4} (X \otimes \Lambda(X) - X \otimes \Lambda(Y) - Y \otimes \Lambda(Y) + Y \otimes \Lambda(Y)); \]
\[ \hat{E}_{13} \otimes \Lambda(\hat{E}_{13}) + \hat{E}_{24} \otimes \Lambda(\hat{E}_{24}) + \hat{E}_{31} \otimes \Lambda(\hat{E}_{31}) + \hat{E}_{42} \otimes \Lambda(\hat{E}_{42}) \]
\[ = \frac{1}{4} (X \otimes \Lambda(X) + X \otimes \Lambda(X) - Y \otimes \Lambda(Y) - Y \otimes \Lambda(Y)); \]
\[ \hat{E}_{12} \otimes \Lambda(\hat{E}_{12}) + \hat{E}_{21} \otimes \Lambda(\hat{E}_{21}) + \hat{E}_{34} \otimes \Lambda(\hat{E}_{34}) + \hat{E}_{43} \otimes \Lambda(\hat{E}_{43}) \]
\[ = \frac{1}{4} (X \otimes \Lambda(X) - Y \otimes \Lambda(Y) + I \otimes \Lambda(I) - I \otimes \Lambda(I)). \]

Note that on the right hand side of the three equations above, all terms with input containing a single \( Y \) have a negative sign. This is due to the factor of \( i \) in Pauli matrix \( Y \). Since linear map \( \Lambda \) is transparent to such scalar factors, the overall tensor
product term will have a negative sign. For example,

\[ XY \otimes \Lambda(XY) = \left[ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \right] \otimes \Lambda(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}) \]

\[ = - \left( \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \right) \otimes \Lambda(\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}). \]

So an extra negative sign has to be added to this type of terms in order to cancel out cross terms in the same way as the pair \( \varepsilon_1 \) and \( \sigma_1 \).

Summing up the expressions for all four pairs of \( \varepsilon \) and \( \sigma \), we have exactly the Choi matrix \( C_\Lambda \) on the left hand side. On the right hand side of the equation are all 16 terms of every 2-qubit Pauli operator tensor multiplying their output from the map \( \Lambda \), e.g. \( XX \otimes \Lambda(XX) \). Once Pauli transfer matrix \( \hat{P}_\Lambda \) is determined by measurements, these terms can be expanded and computed as \( XX \otimes \Lambda(XX) = XX \otimes (p^1_1 XX + p^1_2 XY + ... + p^1_{15} IZ + p^1_{16} II) = p^1_1 XX \otimes XX + p^1_2 XX \otimes XY + ... + p^1_{15} XX \otimes IZ + p^1_{16} XX \otimes II \). Therefore, measurements of these 240 Pauli transfer matrix elements fix the Choi matrix \( C_\Lambda \) and thus completely determines CPTP linear map \( \Lambda \).
5.2 Unitary Quantum Process Tomography

If we assume that the applied CPTP map during experiment is still perfectly unitary, then we may represent it as

\[
\mathcal{V} \begin{pmatrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 & \delta_1 \\ \alpha_2 & \beta_2 & \gamma_2 & \delta_2 \\ \alpha_3 & \beta_3 & \gamma_3 & \delta_3 \\ \alpha_4 & \beta_4 & \gamma_4 & \delta_4 \end{pmatrix} \begin{pmatrix} |00\rangle \\ |01\rangle \\ |10\rangle \\ |11\rangle \end{pmatrix},
\]

where all elements of the $4 \times 4$ matrix are complex. We first prepare input state $|00\rangle$. Its output state is still pure since $\mathcal{V}$ is unitary.

\[
\mathcal{V} |00\rangle = \alpha_1 |00\rangle + \alpha_2 |01\rangle + \alpha_3 |10\rangle + \alpha_4 |11\rangle.
\]

By measuring 3 diagonal observables $IZ$, $ZI$ and $ZZ$ along with the normalization condition $|\alpha_1|^2 + |\alpha_2|^2 + |\alpha_3|^2 + |\alpha_4|^2 = 1$, magnitude values $|\alpha_1|$, $|\alpha_2|$, $|\alpha_3|$ and $|\alpha_4|$ can be determined. We temporarily choose global phase of $\mathcal{V} |00\rangle$ such that the $\alpha$ with the maximum magnitude is positive. For example, if $|\alpha_1|$ is maximum, we set $\alpha_1 = |\alpha_1|$. To determine the other three $\alpha$'s, we measure their coherence with $\alpha_1$. For example, determining the coherence between $\alpha_2$ and $\alpha_1$ is equivalent to determining the phase between $|00\rangle$ and $|01\rangle$, which requires measuring $X$ and $Y$ on the second qubit. In practice, measurement results of $IX$ and $IY$ will also include the coherence between $|10\rangle$ and $|11\rangle$. So these must be subtracted off by also measuring $ZX$ and $ZY$. But since $[IX, ZX] = [IY, ZY] = 0$, $IX$ and $ZX$ can be measured simultaneously, so can $IY$ and $ZY$. So only two measurement configurations are required to fix $\alpha_2$ relative to positive $\alpha_1$. In the same way, $\alpha_3$ and $\alpha_4$ are also fixed by two measurement
configurations. So in total, we need $3 + 2 + 2 + 2 = 9$ measurements to determine $V|00\rangle$ up to a global phase. We can determine $V|01\rangle$, $V|10\rangle$ and $V|11\rangle$ in the same way. So a total number of $4 \times 9 = 36$ measurements are required.

To determine the relative phase between $V|00\rangle$, $V|01\rangle$, $V|10\rangle$ and $V|11\rangle$, three more phase information must be obtained. So the total number of measurements required to fully determine $V$ is $36 + 2 \times 3 = 42$. To determine the relative phase between the $\alpha$–column and $\beta$–column, we input $\frac{1}{\sqrt{2}}(|00\rangle + |01\rangle)$ to obtain output

$$V \frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) = \frac{1}{\sqrt{2}}[(\alpha_1 + e^{i\theta_{\alpha\beta}}\beta_1)|00\rangle + (\alpha_2 + e^{i\theta_{\alpha\beta}}\beta_2)|01\rangle$$

$$+ (\alpha_3 + e^{i\theta_{\alpha\beta}}\beta_3)|10\rangle + (\alpha_4 + e^{i\theta_{\alpha\beta}}\beta_4)|11\rangle].$$

The phase of interest to us is $\theta_{\alpha\beta}$. To obtain it, we use the same coherence measure as before to determine the relative phase between $|00\rangle$ and $|01\rangle$. Denote this measured phase as $\theta_{exp}$, then $\theta_{\alpha\beta}$ can be obtained by solving the equation

$$\text{phase}(\alpha_1 + e^{i\theta_{\alpha\beta}}\beta_1, \alpha_2 + e^{i\theta_{\alpha\beta}}\beta_2) = \theta_{exp}.$$

### 5.3 Experiments and Results

We choose to implement our QPT protocol on the following set of five 2-qubit unitary gates:

$$H_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}, \quad H_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix};$$

98
\[
T_1 = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & e^{i\pi/4} & 0 \\
0 & 0 & 0 & e^{i\pi/4}
\end{pmatrix}, \quad T_2 = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & e^{i\pi/4} & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & e^{i\pi/4}
\end{pmatrix};
\]

\[
CNOT_{12} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0
\end{pmatrix};
\]

where

\[
H_1 = H \otimes I, \quad H_2 = I \otimes H, \quad T_1 = T \otimes I, \quad T_2 = I \otimes T
\]

and

\[
H = \frac{1}{\sqrt{2}} \begin{pmatrix}
1 & 1 \\
1 & -1
\end{pmatrix}, \quad T = \begin{pmatrix}
1 & 0 \\
0 & e^{i\pi/4}
\end{pmatrix}.
\]

This choice is due to the fact that these five unitary gates form a universal gate set from which arbitrary unitary quantum operations can be constructed [15, Section 4.5].

Following the standard QPT and adaptive unitary QPT protocols in the previous two section, two sets of data are obtained for each of the five gates. The optimal gates that fit these data are obtained by maximizing the Gaussian likelihood function.

In the standard QPT method, this is equivalent to finding a \( \hat{P} \) that minimizes the sum of the squared magnitude of all elements of \( \hat{P} - \hat{D} \), which can be expressed as \( Tr((\hat{P} - \hat{D})(\hat{P} - \hat{D})^T) \), where \( \hat{D} \) is the measured data set. This optimization must
be subject to the trace preserving constraint: \( \hat{P}(16, 16) = 1 \) and \( \hat{P}(i, 16) = 0 \) for \( i = 1, \ldots, 15 \) and the completely positive constraint \( \text{Choi}(\hat{P}) \geq 0 \).

In the adaptive unitary QPT method, this is equivalent to finding a \( \hat{U} \) that minimizes \( Tr( (\hat{U} - \hat{D})(\hat{U} - \hat{D})^\dagger ) \) subject to the constraint \( \hat{U}\hat{U}^\dagger = I \).

These optimization problems are solved using the convex optimization solver CVX [54].

To visually compare the optimized results from the two QPT protocols with the true theoretical gates used, we plot their \( \hat{P} \) matrices as shown in Figure 5.3.1.

To evaluate the closeness between the optimization results and the implement universal gate set, we use average state fidelity

\[
F = \int d\psi \left[ Tr(A(\langle \psi | \psi \rangle) \hat{U}_0 \langle \psi | \psi \rangle \hat{U}_0^\dagger) \right]^{\frac{1}{2}} \frac{1}{N} / \int d\psi,
\]

where the normalization factor \( N = \left[ Tr(A(\langle \psi | \psi \rangle)^2) Tr \left( (\hat{U}_0 \langle \psi | \psi \rangle \hat{U}_0^\dagger)^2 \right) \right]^{\frac{1}{2}} \) and \( d\psi \) integrates over the full parameter space of a pure state \( |\psi \rangle \). Note: this reduces to \( F = \int d\psi |\langle \psi | \hat{U}_0^\dagger \hat{U}_0 |\psi \rangle| / \int d\psi \) if \( A = \hat{U} \). To implement this in a computer program, we use 1000 random input states to compute the integrand value and take the average value as the fidelity. The results are shown in Table 5.1.

<table>
<thead>
<tr>
<th>Gates</th>
<th>H1</th>
<th>H2</th>
<th>T1</th>
<th>T2</th>
<th>CNOT</th>
</tr>
</thead>
<tbody>
<tr>
<td>QPT F</td>
<td>0.9953</td>
<td>0.9927</td>
<td>0.9931</td>
<td>0.9970</td>
<td>0.9933</td>
</tr>
<tr>
<td>UQPT F</td>
<td>0.9934</td>
<td>0.9961</td>
<td>0.9908</td>
<td>0.9852</td>
<td>0.9772</td>
</tr>
</tbody>
</table>

Table 5.1: Average State Fidelity: standard and unitary QPT

As shown, our adaptive unitary QPT protocol can successfully reconstruct the underlying unitary operation with 200 less measurements but only a slight decrease in fidelity.

The CNOT gates has the lowest fidelity. This is due to the special constraint
Figure 5.3.1: $\hat{P}$ Matrix of all Gates: Theory, Standard QPT and UQPT
of the NMR system. Double coherence like transitions from $|00\rangle$ to $|11\rangle$ that are generated by CNOT gate cannot be measured directly. So in real experiments, the CNOT gate has to be applied twice to transform this back to a single coherence, which will increase the error in the data.
Chapter 6

Conclusion

In this work, we first studied the special case of QPT of 1-qubit unitary gates. We demonstrated that 6 Pauli measurements are required in a non-adaptive approach to fully determine all such gates without failure. An adaptive approach to fully determine all such gates with at most 5 Pauli measurements is found. Non-adaptive experiments using only 5 Pauli measurements are performed on an NMR system. The results indicated that while non-adaptively measuring 5 Pauli operators can determine many underlying 1-qubit unitary gates with high fidelity, certain special gates will cause failure. We also showed that our 1-qubit QPT method can be nicely visualized on the Bloch sphere as spatial rotations due to homomorphism between $SU(2)$ and $SO(3)$. Next, we presented an adaptive approach to determine a $d$–dimensional pure state among all quantum states with at most $2d - 1$ measurements. Using this approach, we showed that theoretically, at most $d^2 + d - 1$ measurements are required to fully determine a $d$–dimensional unitary gate. Since a $d$–dimensional pure state and unitary gate only have $2d - 2$ and $d^2 - 1$ independent real parameters, the margin to further reduce the required number of measurements, if exists at all, is very small. This approach is then demonstrated by performing QPT on 2-qubit
unitary gates. The resulting gate fidelity is found to be well above 95% and only slightly lower than using the standard QPT method. For a 2-qubit unitary gate, the standard QPT method requires measuring 240 observables, while this new approach can be implemented on an NMR system with only 42 Pauli measurements. Our work is summarized and published in [55].
Bibliography


[54] This convex optimization solver is free to download for academic usage at http://cvxr.com/cvx/.