Facilitating the calibration of complex quantum photonic circuits with machine learning assisted gate set tomography

By

Liam Dalgarno

Department of Computer Science

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of Master of Engineering in the Faculty of Engineering.

February 2, 2021

Word count: 2392

Declaration

This dissertation is submitted to the University of Bristol in accordance with the requirements of the degree of MEng in the Faculty of Engineering. It has not been submitted for any other degree or diploma of any examining body. Except where specifically acknowledged, it is all the work of the Author.

SIGNED: ..... DATE: .....

Abstract

Fusce mauris. Vestibulum luctus nibh at lectus. Sed bibendum, nulla a faucibus semper, leo velit ultricies tellus, ac venenatis arcu wisi vel nisl. Vestibulum diam. Aliquam pellentesque, augue quis sagittis posuere, turpis lacus congue quam, in hendrerit risus eros eget felis. Maecenas eget erat in sapien mattis porttitor. Vestibulum porttitor. Nulla facilisi. Sed a turpis eu lacus commodo facilisis. Morbi fringilla, wisi in dignissim interdum, justo lectus sagittis dui, et vehicula libero dui cursus dui. Mauris tempor ligula sed lacus. Duis cursus enim ut augue. Cras ac magna. Cras nulla. Nulla egestas. Curabitur a leo. Quisque egestas wisi eget nunc. Nam feugiat lacus vel est. Curabitur consectetuer.

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

Acknowledgements

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Contents

List of Figures

List of Tables

List of Acronyms

GST Gate Set Tomography

POVM Positive Operator-Valued Measure

CPTP Completely Positive and Trace-Preserving

Todo list

Chapter 1: Notes

1.1 Tomography

Notes taken from [1], only reworded and trimmed down for my benefit.

TODO: Clean up these notes and put into a Technical Background chapter.

1.1.1 Operator-Sum Representation

Quantum operations can be represented in the *operator-sum representation*. Consider a state ρ coupled with the environment ρ_{env} which is transformed by U. The final state $\mathcal{E}(\rho)$ of the system is then

$$\mathcal{E}(\rho) = \text{tr}_{\text{env}}\left(U(\rho \otimes \rho_{\text{env}})U^\dagger\right) \tag{1.1}$$

where we trace out the environment to obtain the state of the system alone. To rewrite this in operator form, we let ρ_{env} be a pure state $|e_0\rangle\langle e_0|$ with orthonormal basis $|e_k\rangle$. Then, we have

$$\mathcal{E}(\rho) = \sum_k \langle e_k| U\Big[\rho \otimes |e_0\rangle\langle e_0|\Big]U^\dagger |e_k\rangle \tag{1.2}$$

$$= \sum_k E_k \rho E_k^\dagger \tag{1.3}$$

Where $E_k = \langle e_k|U|e_0\rangle$, and we omit the identities in $I \otimes |e_k\rangle$. See this answer for why. The operators E_k are known as *operation elements*. This form is incredibly useful for tomography, which is covered in the following sections.

TODO: Expand this more

1.1.2 State Tomography

In the classical world, characterising the dynamics of a system is trivial and known as *system identification*. The general idea is that we wish to know how the system behaves with respect to any input, thus uniquely identifying it. In the quantum world, the analogue of this is called *quantum process tomography*. To understand process tomography, we must first understand *quantum state tomography*.

State tomography is the procedure of determining an unknown quantum state. This is harder than it sounds: if we're given an unknown state ρ, we can't just measure the state and recover it immediately way since measurement will *disturb* the original state. In fact, *there is no quantum measurement which can distinguish non-orthogonal states with certainty*. However, if we have an *ensemble* of the same quantum state ρ, then it's possible to estimate ρ.

If we represent the state of the system using its density matrix ρ, we may expand ρ as

$$\rho = \frac{\text{tr}\,(\rho)\, I + \text{tr}\,(X\rho)\, X + \text{tr}\,(Y\rho)\, Y + \text{tr}\,(Z\rho)\, Z}{2} \tag{1.4}$$

Note that $\text{tr}\,(Z\rho)$ can be interpreted as the *expectation* of the observable Z. Therefore, to estimate $\text{tr}\,(Z\rho)$, we measure the observable Z m-times to obtain outcomes $z_1, \ldots, z_m$ and calculate

$$\text{tr}\,(Z\rho) \approx \frac{1}{m}\sum_i^m z_i \tag{1.5}$$

In general, this estimate is approximately a Gaussian with mean $\text{tr}\,(Z\rho)$ and standard deviation $\Delta(Z)/\sqrt{m}$, where $\Delta(Z)$ is the standard deviation of a single measurement. We can apply this same

method to estimate $\mathrm{tr}\,(X\rho)\,X$ and $\mathrm{tr}\,(Y\rho)\,Y$; with a large enough sample size we obtain a good estimate for ρ. This process can be generalised to a density matrix on n qubits to

$$\rho = \sum_{\vec{v}} \frac{\mathrm{tr}\,(\sigma_{v_1} \otimes \sigma_{v_2} \otimes \cdots \otimes \sigma_{v_n})\,\sigma_{v_1} \otimes \sigma_{v_2} \otimes \cdots \otimes \sigma_{v_n}}{2^n} \tag{1.6}$$

where $\vec{v} = (v_1, \ldots, v_n)$ with entries v_i chosen from the set $0, 1, 2, 3$, i.e. each σ_{v_i} is a particular Pauli matrix.

1.1.3 Process Tomography

To extend this notion to quantum process tomography is actually quite easy from a theoretical point of view. If the state space of the system has d dimensions ($d = 2$ for a single qubit), then we choose d^2 pure quantum states $|\psi_d\rangle, \ldots, |\psi_{d^2}\rangle$. The corresponding density matrices $|\psi_1\rangle\langle\psi_1|, \ldots, |\psi_{d^2}\rangle\langle\psi_{d^2}|$ for these states should form a *basis set* for the space of possible density matrices.

Now, for each state $|\psi_j\rangle$, we prepare the system in that state and then subject it to the process $\mathcal{E}$. Afterwards, we use state tomography to determine the output state $\mathcal{E}(|\psi_j\rangle\langle\psi_j|)$. Theoretically, this is all that we need to do, since the matrices $|\psi_j\rangle\langle\psi_j|$ form a basis set, so any other possible density matrices can be represented as a linear combination of the basis set. e.g.,

$$\mathcal{E}(|\Phi\rangle\langle\Phi| + |\Psi\rangle\langle\Psi|) = \mathcal{E}(|\Phi\rangle\langle\Phi|) + \mathcal{E}(|\Psi\rangle\langle\Psi|) \tag{1.7}$$

However, in practice, it's not that simple since operators are just a 'theoretical tool'. We only have access to measureable quantities. We can instead describe $\mathcal{E}$ using *operator-sum representation*, where our goal is to determine a set of operation elements $\{E_i\}$ for $\mathcal{E}$,

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger \tag{1.8}$$

To obtain these operators from measurements, it is convenient to use a different formulation of $\mathcal{E}$ from a *fixed* set of operators $\tilde{E}_i$. These fixed operators will form a basis set for the operators of that state space, so that

$$E_i = \sum_m e_{im} \tilde{E}_m \tag{1.9}$$

for some set of complex numbers e_{im}. We can rewrite Equation 1.8 as

$$\mathcal{E}(\rho) = \sum_{mn} \tilde{E}_m \rho \tilde{E}_n^\dagger \chi_{mn} \tag{1.10}$$

where each $\tilde{E}_i$ is a fixed operator, and χ is a complex matrix known as the *chi matrix representation*. χ is defined as

$$\chi_{mn} = \sum_i e_{im} e_{in}^* \tag{1.11}$$

Here, we have simply combined the complex numbers e_{im} into a single matrix. By definition, this is a positive Hermitian matrix. This shows that we can determine $\mathcal{E}$ entirely from χ, once the set of operators has been fixed.

Now, let $p_j, 1 \leq j \leq d^2$ be a set of matrices which form an independent basis for the the space of $d \times d$ matrices. We can determine each $\mathcal{E}(\rho_j)$ by state tomography, and then express this as a linear combination of the basis states,

$$\mathcal{E}(\rho_j) = \sum_k \lambda_{jk} \rho_k \tag{1.12}$$

Since we know $\mathcal{E}(\rho_j)$ via state tomography, we can determine λ_{jk} by linear algebra. i.e. $\mathcal{E}(\rho_j)$ is a known matrix and each ρ_k is known, so we simply need to solve standard linear equations for each λ_{jk}. Furthermore, we can construct a matrix β_{mn} from the fixed operators $\tilde{E}_m$ as follows

$$\tilde{E}_m \rho_j \tilde{E}_n^\dagger = \sum_k \beta_{jk}^{mn} \rho_k \tag{1.13}$$

where β_{jk}^{mn} are complex numbers determined by standard linear algebra algorithms given the operators $\tilde{E}_m$, $\tilde{E}_n$, and ρ_j. Note that the m and n come from the fixed operators in Equation 1.10. If we combine the previous two equations with Equation 1.10, we get

$$\sum_k \sum_{mn} \chi_{mn} \beta_{jk}^{mn} \rho_k = \sum_k \lambda_{jk} \rho_k \tag{1.14}$$

From the linear independence of ρ_k, it follows that

$$\sum_{mn} \beta_{jk}^{mn} \chi_{mn} = \lambda_{jk} \tag{1.15}$$

Equation 1.15 is a necessary condition for χ to give the correct quantum operation $\mathcal{E}$.

We can think of χ and λ as vectors, and β as a $d^4 \times d^4$ matrix. Here, β has columns indexed by mn and rows indexed by jk. Now, suppose that β has an inverse κ. χ is then completely defined by

$$\chi_{mn} = \sum_{jk} \kappa_{jk}^{mn} \lambda_{jk} \tag{1.16}$$

Which we can write more clearly in matrix form as

$$\vec{\chi} = \kappa \vec{\lambda} \tag{1.17}$$

Fortunately, once we know χ, we immediately know the operator-sum representation of $\mathcal{E}$. Suppose that some unitary matrix $U^\dagger$ diagonalises χ. Then, we can construct E_i using

$$\chi_{mn} = \sum_{xy} U_{mx} d_x \delta_{xy} U_{ny}^* \tag{1.18}$$

$$E_i = \sqrt{d_i} \sum_j U_{ji} \tilde{E}_j \tag{1.19}$$

In summary: λ is determined using state tomography, which then determines χ via $\vec{\chi} = \kappa \vec{\lambda}$. We can then use χ to completely determine each E_i.

In general χ will contain $d^4 - d^2$ independent parameters. A linear map from $d \times d$ complex matrices to $d \times d$ matrices is described by d^4 independent parameters. d^2 comes from the constraint that ρ must remain Hermitian with unit trace. For a single qubit, we'll have $2^4 - 2^2 = 12$ parameters, whilst two qubits will have $4^4 - 4^2 = 240$ parameters!

The above process looks pretty complex. Fortunately, in the case of a single qubit, we can pick specific fixed operators $\tilde{E}_i$ to simplify the process massively. We select

$$\tilde{E}_0 = I \tag{1.20}$$

$$\tilde{E}_1 = X \tag{1.21}$$

$$\tilde{E}_2 = -iY \tag{1.22}$$

$$\tilde{E}_3 = Z \tag{1.23}$$

For our basis set ρ_j, we can prepare the input states $|0\rangle, |1\rangle, |+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}, |-\rangle = (|0\rangle + i|1\rangle)/\sqrt{2}$. We determine

$$\rho_1' = \mathcal{E}(|0\rangle\langle 0|) \tag{1.24}$$

$$\rho_4' = \mathcal{E}(|1\rangle\langle 1|) \tag{1.25}$$

$$\rho_2' = \mathcal{E}(|1\rangle\langle 0|) = \mathcal{E}(|+\rangle\langle +|) - i\mathcal{E}(|-\rangle\langle -|) - (1-i)(\rho_1' + \rho_4')/2 \tag{1.26}$$

$$\rho_3' = \mathcal{E}(|0\rangle\langle 1|) = \mathcal{E}(|+\rangle\langle +|) + i\mathcal{E}(|-\rangle\langle -|) - (1+i)(\rho_1' + \rho_4')/2 \tag{1.27}$$

because $|0\rangle\langle 0|, |1\rangle\langle 1|, |1\rangle\langle 0|, |0\rangle\langle 1|$ form a basis for the set of $d \times d$ matrices. This particular choice of basis allows us to represent β as the Kronecker product $\beta = \Lambda \otimes \Lambda$ and χ in terms of block matrices:

$$\Lambda = \frac{1}{2} \begin{bmatrix} I & X \\ X & -I \end{bmatrix} \tag{1.28}$$

$$\chi = \Lambda \begin{bmatrix} p_1' & p_2' \\ p_3' & p_4' \end{bmatrix} \Lambda \tag{1.29}$$

TODO: Refine all of this down for the background chapter, because we're mostly interested in the single qubit case.

TODO: Figure out why $\beta = \Lambda \otimes \Lambda$ gives that equation for χ.

1.2 Gate Set Tomography

Notes taken from [2].
Gate Set Tomography (GST) differs from state and process tomography in that:

- It is almost entirely calibration-free. It does not depend upon a prior description of the measurements used (as in state tomography) or the states that can be prepared (as in process tomography). *These are called "reference frame" operations in the literature?*
- It estimates an entire set of logic operations, rather than a single one.

GST being calibration-free is incredibly important. Both state and process tomography are limited in that they rely on *accurate* characterisation of their "reference frame" operations. Typically, they're either unknown or misidentified.

1.2.1 Mathematical Background

A quantum system is described by a d-dimensional *Hilbert space* $\mathcal{H} = \mathbb{C}^d$, where d is the largest number of outcomes of a repeatable measurement. For a qubit, $d = 2$. GST uses the *Hilbert-Schmidt space*. The Hilbert-Schmidt space is the complex d^2-dimensional vector space of $d \times d$ matrices. We're interested in the d^2-dimensional subspace of Hermitian matrices, denoted $\mathcal{B}(\mathcal{H})$. The basis we use for $\mathcal{B}(\mathcal{H})$ is the set of normalised Pauli matrices $\{\mathbb{I}/\sqrt{2}, \sigma_x/\sqrt{2}, \sigma_y/\sqrt{2}, \sigma_z/\sqrt{2}\}$. This basis has the following properties:

- Hermicity: $B_i = B_i^\dagger$
- Orthonormality: $\mathrm{tr}\,(B_i B_j) = \delta_{ij}$
- Traceless for $i > 0$: $B_0 = \mathbb{I}/\sqrt{d}$ and $\mathrm{tr}\,(B_i) = 0 \;\; \forall i > 0$.

Elements of $\mathcal{B}(\mathcal{H})$ are represented using an 'extension' of Dirac's bra-ket notation called *super bra-ket notation*. Some element B is represented as a column vector $|B\rangle\rangle$, and an element of its dual space by a row vector $\langle\langle A|$. Everything works similarly to regular Dirac notation, the main difference is that we can represent everything as vectors in $\mathcal{B}(\mathcal{H})$ rather than the usual matrices.

Measurement of a quantum system yields an outcome from a set of k possibilities. Therefore, the ith outcome can be represented by a dual vector $\langle\langle E_i|$, so that $\Pr(i|\rho) = \langle\langle E_i|\rho\rangle\rangle = \mathrm{tr}\,(E_i\rho)$. Since they represent probabilities, we require that $E_i \geq 0$ and $\sum_i E_i = \mathbb{I}$. The E_i are called *effects*, and the set $\{E_i\}$ is called a Positive Operator-Valued Measure (POVM). Note that since both states and effects are both Hermitian, we can in fact represent them in the d^2-dimensional real subspace of $\mathcal{B}(\mathcal{H})$. Therefore, any reference to $\mathcal{B}(\mathcal{H})$ is referring to the real subspace.

1.2.2 Quantum Logic Gates

An *ideal* quantum logic gate is *reversible* and corresponds to a unitary transform of $\mathcal{H}$. Such a gate would transform ρ as $\rho \to U\rho U^\dagger$ for some unitary matrix U. This is a linear transformation from $\mathcal{B}(\mathcal{H})$ to itself; the linear transformation $\rho \to U\rho U^\dagger$ is called a *superoperator*. In reality, logic gates are not perfectly reversible. These superoperators are known as quantum processes or quantum channels. We can represent any superoperator Λ as a $d^2 \times d^2$ matrix, which acts on $|\rho\rangle\rangle \in \mathcal{B}(\mathcal{H})$ by left multiplication. This representation is called the *transfer matrix* of Λ, and is denoted by S_Λ. Thus,

$$\Lambda : |\rho\rangle\rangle \mapsto S_\Lambda\,|\rho\rangle\rangle \tag{1.30}$$

If Λ is performed on some input state ρ, then the probability of outcome E_i is therefore

$$p_i = \langle\langle E_i | S_\Lambda | \rho \rangle\rangle = \mathrm{tr}\left(E_i S_\Lambda \rho\right) \tag{1.31}$$

Not all superoperators describe physical operations. To be physically possible, they must be:

- *Trace-preserving*: $\mathrm{tr}\left(\Lambda(\rho)\right)$ must equal 1 for all ρ.
- *Completely Positive*: when Λ acts on part of a larger system, it must preserve positivity for the entire system. A superoperator is *positive* iff. $\Lambda(\rho) \geq 0$ for all ρ. A superoperator is *completely positive* iff. $\Lambda \otimes \mathbb{I}_{\mathcal{A}}$ is positive for any auxiliary state space $\mathcal{A}$.

This Completely Positive and Trace-Preserving (CPTP) constraint alone is sufficient – any CPTP superoperator can be physically implemented. The TP condition corresponds to $\langle\langle \mathbb{I}| S_\Lambda = \langle\langle \mathbb{I}|$. Since our basis is traceless for $i > 0$, then Λ is TP iff. the first row of S_Λ is $[1, 0, \ldots, 0]$. The CP condition is a lot more tricky to describe. We first rewrite S_Λ in the operator-sum representation:

$$\Lambda : \rho \mapsto \sum_{ij} \chi^\Lambda_{ij} B_i \rho B_j^\dagger \tag{1.32}$$

where $\{B_i\}$ is a basis, and χ^Λ_{ij} is a matrix of coefficients called the "Choi process matrix" which represents Λ. Similarly to the chi matrix representation from earlier, this completely describes Λ (read this answer for more about their relationship). The mapping between S_Λ and χ^Λ_{ij} is known as the Choi-Jamiołkowski isomorphism:

$$\chi^\Lambda = d(S_\Lambda \otimes \mathbb{I}) \left|\Pi_{\mathrm{EPR}}\right\rangle\!\rangle \tag{1.33}$$

where $\left|\Pi_{\mathrm{EPR}}\right\rangle\!\rangle$ is a maximally entangled state. Note that the equality above really means element-wise equality in a consistent basis. This is all quite complex. Fortunately, the process is simple once we have the Choi process matrix: Λ is CP iff. χ is positive semidefinite.

TODO: Research Choi representation vs. Chi representation

1.2.3 Gate Sets

A quantum processor's capabilities can be specified with a *gate set*. Consider a processor that can perform:

- N_G distinct gates,
- N_ρ distinct state preparations,
- N_M distinct measurements, with $N_E^{(m)}$ distinct outcomes.

We can use these to construct a gate set

$$\mathcal{G} = \left\{ \left\{ \left|\rho^{(i)}\right\rangle\!\right\rangle \right\}_{i=1}^{N_\rho} ; \{G_i\}_{i=1}^{N_G} ; \left\{ \langle\!\langle E_i^{(m)} | \right\}_{m=1,i=1}^{N_m, N_E^{(m)}} \right\} \tag{1.34}$$

where

$$\begin{aligned} G_i : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}) \quad &\text{for} \quad i = 1, \ldots, N_G \\ \left|\rho^{(i)}\right\rangle\!\right\rangle \in \mathcal{B}(\mathcal{H}) \quad &\text{for} \quad i = 1, \ldots, N_\rho \\ \langle\!\langle E_i^{(m)}| \in \mathcal{B}(\mathcal{H})^* \quad &\text{for} \quad m = 1, \ldots, N_M, i = 1, \ldots, N_E^{(m)} \end{aligned} \tag{1.35}$$

In other words, we can describe the capabilities of a quantum processor by collecting together the possible 'input states', the quantum gates, and the possible measurement outcomes for each input state.

1.2.4 Circuits

The term 'quantum circuit' can refer to different things. There are two related but distinct types of quantum circuit that we're interested in.

- **Fixed-Input, Classical-Output (FI/CO)**: a quantum circuit where each experiment is described by a quantum circuit that begins by initialising and ends by measuring all of the qubits. A FI/CO circuit describes a probability distribution over classical bit strings.
- **Quantum-Input, Quantum-Output (QI/QO)**: a quantum circuit which is an arrangement of unitary logic gates, with no explicit initialisation or measurement. This may be inserted into a large quantum circuit as a 'sub-routine'.

QI/QO circuits are defined as a sequence of layers $S = (\gamma_1, \gamma_2, \ldots, \gamma_L)$. Each layer corresponds to applying some superoperator G_{γ_i}. The entire QI/QO circuit S also corresponds to applying a superoperator. We denote the transfer matrix for S by $\tau(S)$, where

$$\tau(S) = \tau((\gamma_1, \gamma_2, \ldots, \gamma_L)) = G_{\gamma_L} \ldots G_{\gamma_2} G_{\gamma_1} \tag{1.36}$$

which is formed by composing the elements from each layer. Exponentiation of a circuit describes repetition, and since $S^2 = SS$, it follows that

$$\tau(S^n) = \tau(S)^n \tag{1.37}$$

For a set of FI/CO circuits, we can generate data sets by repeating each one N times, and recording the results. The results are summarized by observed frequencies, $f_k = n_k/N$, which approaches the corresponding probabilities

$$f_k \approx \langle\langle E_k^{(m)} | G_{\gamma_L} \ldots G_{\gamma_2} G_{\gamma_1} | \rho^{(i)} \rangle\rangle \tag{1.38}$$

and can be used to estimate them. This is not the only way to do so, but it illustrates that we can infer some of $|\rho^{(i)}\rangle\rangle$, $\langle\langle E_k^{(m)}|$, and G_j based on observed frequencies. This is called tomography. *Each type of tomography treats some of the operations as known, and uses them as the reference frame to estimate the others.*

1.2.5 Quantum State Tomography

In order to perform *any* type of tomography, we must have a *fiducial* and *informationally complete* set. "Fiducial" means that it is accepted as a fixed basis of reference, and "informationally complete" means that it will uniquely identify any target information.

In the context of state tomography, we're aiming to describe some unknown quantum state ρ, given some fiducial and informationally complete measurements. In other words, the measurements are a frame of reference and must uniquely identify ρ. This implies that the set of measurements $\{E_i^{(m)}\}_{m,i}$ must span the entire space of effects. To perform state tomography, many copies of ρ are made available, and divided into M pools. The mth fiducial measurement is applied to all the copies in the mth pool, and used to estimate the probabilities,

$$p_i^{(m)}(\rho) = \mathrm{tr}\left(\rho E_i^{(m)}\right) \tag{1.39}$$

which should uniquely identify ρ. In general, this is only true if we have infinite copies of ρ. In practice, we have limited numbers of ρ, and therefore $\hat{p}_i = f_i$ where f_i are the frequencies for each measurement. This often yields an estimate $\hat{\rho}$ which is not positive. Nevertheless, it is still useful.

Let's assume that we have the exact probabilities for each measurement outcome for the state ρ. We can in fact ignore the measurements themselves, what is important is the list of effects, so we can write the set simply as $\{E_j : j = 1, \ldots, N_{f1}\}$ where $f1$ is the total number of distinct measurement outcomes. Additionally, we can represent these as dual vectors $\langle\langle E_j|$ in $\mathcal{B}(\mathcal{H})$ just like before. To recover ρ, we can write Born's rule as an inner product,

$$\begin{aligned} p_j &= \mathrm{tr}\left(E_j \rho\right) \\ &= \langle\langle E_j | \rho \rangle\rangle \end{aligned} \tag{1.40}$$

Then, we can stack all of the effects into a single $N_{f1} \times d^2$ matrix

$$A = \begin{bmatrix} \langle\langle E_1| \\ \langle\langle E_2| \\ \vdots \\ \langle\langle E_{N_{f1}}| \end{bmatrix} \tag{1.41}$$

which gives $\vec{p} = A\,|\rho\rangle\rangle$. If A is square, then we can recover ρ with:

$$|\rho\rangle\rangle = A^{-1}\vec{p} \tag{1.42}$$

If N_{f1} is greater than d^2, making A non-square, we must solve with a pseudo-inverse:

$$\begin{aligned} (A^TA)^{-1}A^T\vec{p} &= (A^TA)^{-1}A^TA\,|\rho\rangle\rangle \\ &= (A^TA)^{-1}(A^TA)\,|\rho\rangle\rangle \\ &= I\,|\rho\rangle\rangle \\ &= |\rho\rangle\rangle \end{aligned} \tag{1.43}$$

1.2.6 Quantum Process Tomography

In the context of process tomography, we're aiming to describe some quantum process (e.g. a quantum gate), given an informationally complete set of known fiducial states. Broadly, we prepare many copies of them, pass them through the target process, and perform state tomography on the output states.

Let G be the superoperator representing the process we want to reconstruct. If state ρ_i is prepared, G is applied, and measurement is performed with possible outcomes $\{E_j\}$, then the probability of observing E_j is

$$\begin{aligned} Pj, i &= \mathrm{tr}\left(E_jG[\rho_i]\right) \\ &= \langle\langle E_j|G|\rho_i\rangle\rangle \end{aligned} \tag{1.44}$$

We can then define a $d^2 \times N_{f2}$ matrix B, similarly to A, which represents the fiducial states $|\rho_i\rangle\rangle$:

$$B = \begin{bmatrix} |\rho_1\rangle\rangle & |\rho_2\rangle\rangle & \dots & |\rho_{N_{f2}}\rangle\rangle \end{bmatrix} \tag{1.45}$$

which gives the $N_{f1} \times N_{f2}$ matrix $P = AGB$. Similarly to before, we can recover G with (pseudo)-inverses. If P is square, we can recover G with:

$$G = A^{-1}PB^{-1} \tag{1.46}$$

If P is non-square, then again we use a pseudo-inverse:

$$G = (A^TA)^{-1}A^TPB^T(BB^T)^{-1} \tag{1.47}$$

1.2.7 Calibration

The requirements of fiducial states/measurements in state and process tomography show why GST being calibration-free is important. In practice, we never have access to perfectly known states/measurements, and they're also not noiseless. In order to identify the exact fiducial measurements for state tomography, we would need perfectly known states, which would require state tomography. Similarly, process tomography relies on fiducial states and measurements which are almost always produced by applying quantum logic gates. Identifying these would again require process tomography. It's an endless loop of self-referentiality.

Bibliography

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. USA: Cambridge University Press, tenth ed., 2011.

[2] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, "Gate Set Tomography," *arXiv:2009.07301 [quant-ph]*, Sept. 2020.