# On the evaluation of quantum instruments with a consideration to measurements in trapped ion systems

by

Darian McLaren

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics (Quantum Information)

Waterloo, Ontario, Canada, 2022

## Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Statement of Contributions

1. Chapter 2 was completed under the supervision of Rajibul Islam and was done as a part of a larger project being undertaken by Rajibul Islam and the members of his research group the Laboratory for Quantum Information with Trapped Ions (QITI). The work will be included as part of a manuscript to be submitted for publication shortly after the completion of this thesis. The writing in this chapter is my own, however many members of the QITI group contributed ideas and provided feedback to the presented work.

    - Section 2.2 is my own calculation.
    - The method outlined in Section 2.3 to solve the Lindblad master equation was based on an earlier method used by Chung-You Shih. The adaptation of the method to the case of Ytterbium was performed by myself with the oversight of Nikhil Kotibhaskar.
    - Section 2.4 is an, as indicated, previously known method for creating histograms of photon counts.
    - Section 2.5 is an example developed by myself.

2. Chapter 3 was completed under the supervision of Joel Wallman and was originally completed as a draft for a manuscript authored by myself and Wallman soon to be submitted for publication. The overall structure and content of the chapter follows that of the draft of the manuscript, but reworded into my words.

    - Sections 3.2 and 3.3 are simply a review of well known results.
    - Sections 3.4 and 3.7 consist of theorems/lemmas which were suggested by Wallman. The proofs of which were completed by myself with Wallman providing minor edits/proofreading.
    - Section 3.5 consists of an analysis of errors pre- and post-measurement provided by Wallman. The idea of showing that the fidelity and trace distance of the output of the quantum instruments can reduced to classical distinguishability measures was my own.
    - The idea in section 3.6 to motivate the diamond distance by bounding the trace distance of the ideal and noisy output was Wallman's. The associated calculation was performed by myself.

- In section 3.8 the suggestion of using the unitary channel to saturate the upper bound was given by Wallman, where I performed the computation. The other examples are my own.

## Abstract

Trapped ion chains have shown promise in their application as quantum simulators. However, the close proximity of ions in the trap leads to operations such as state detection causing loss of coherence of other ions due to imperfect beam addressing and absorption of scattered photons during neighbouring ion fluorescence. In the first part of my thesis I consider a Ytterbium 171 ion trap and calculate the probability of neighbouring ion photon absorption. I additionally show how to use the Lindblad master equation to simulate the dynamics of an ion under an attenuated beam, and provide a review of determining the detection fidelity through histograms of photon detection counts.

I then return to the more abstract setting of non-destructive quantum measurements, which are instrumental for fault-tolerant quantum computers. These non-destructive quantum measurements have both a quantum and classical output and hence must be described through a quantum instrument as opposed to the typical representation of measurements through POVM's. It therefore becomes necessary to develop methods for evaluating quantum instruments. In this thesis I show that the process fidelity and diamond distance as figures of merits on unitary channels may be generalized to figures of merits on quantum instruments. I show that these figures of merit adequately capture the errors associated to non-destructive measurements, and they additionally provide upper and lower bounds for each other. Several examples are also given of computing the figures of merit on quantum instruments under various noise models.

**Acknowledgements**

I would like make a special thank you to my supervisor Dr. Kazi Rajibul Islam and all the members of QITI for welcoming me into their group and their patience and dedication aiding me in my research. I'd especially like to thank Nikhil, Sainath, Anthony, and Gilbert whose help made this thesis possible.

I thank my supervisor Dr. Joel Wallman for many fruitful discussions over the past 2 years teaching me everything I know about quantum instruments and for introducing me to the QITI group.

And finally I would like to thank Dr. Sarah Plosker for taking me under her wing during my undergrad, sparking my interest in quantum information, and mentoring me to be where I am now.

# Table of Contents

# List of Figures

# Chapter 1

# Introduction

In this thesis we will be analyzing two separate albeit somewhat related problems involving the characterization of errors in measurements on quantum computing systems. The first problem is analyzed in Chapter 2 which provides a brief look at the errors that arise when measuring the trapped ion qubits in a Ytterbium 171 ion chain. These errors arise due to the close proximity of ions in the chain. There have been many methods developed that seek to minimize these errors. One such method is ion shuttling [30, 10], in which any time an operation is needed on an ion it is "shuttled" away from neighbouring ions, the operation performed, and then the ion brought back to it's original position. Another possible solution is the use of multiple species [4, 6], where a trap may consist of a line of ions alternating between two species. One species would be designated for memory and the other designated for qubit operations. A measurement on a qubit can be performed by, for example, transferring the state of the selected memory qubit to its designated neighbour of a different species and then performing the measurement on the neighbour. This has the benefit that operations such as state detection can made to be off-resonant with the ions used for memory and hence limit their destructive effects. These methods require additional overhead on the system and so it becomes necessary to develop tools to identify when these method are needed. Chapter 2 will provide an analysis of errors in measurement for strictly *in situ* ions.

The primary focus of this thesis however is in chapter 3, which takes a more abstract approach. While there has been extensive research on evaluating unitary operations, the research on evaluating measurements is not as extensive; in many contexts, errors in the measurement process are often swept up into SPAM (state preparation and measurement) and ignored. Moreover, most methods of evaluating measurements concentrate on evaluating destructive measurements. However, with the advent of quantum non-demolition

measurements which preserve the state of the system, it becomes necessary to develop tools for evaluating these more general forms of measurement. Consequently, the representation of a measurement with a POVM, which has only classical output, is not sufficient. Measurements can more generally be represented by quantum instruments [12, 27], which have both quantum and classical output. In Chapter 3 we use the extensive literature on evaluating unitary channels to develop figures of merit that may be used to evaluate quantum instruments.

# Chapter 2

# Measurement in a Ytterbium Trap

As the qubit states in trapped ion systems are encoded in the electronic energy levels of an ion, most qubit operations are performed by exposing the ion to light of various frequencies in order to drive specific energy level transitions. Using light in this way imposes restrictions on the maximum amount of ions that can be placed in the trap while still being able to successfully perform the various qubit operations and preserving the states of the qubits.

The first limitation, referred to as the technical limit, is the simple requirement that any beam meant to perform a single qubit operation will need to be focused on a single ion. Assuming the beam being used is Gaussian it will obey an exponential decay in intensity. This sounds favorable, however ions that are a few beam widths away from the target can still have their states altered in a significant way if exposed to the beam for a sufficient amount of time.

The second limitation, referred to as the fundamental limit, is that many of the qubit operations will lead to the ion emitting a photon through its transition to a lower energy level. In some cases this is actually desirable, e.g. state detection, or can simply be an unavoidable byproduct, e.g. optical pumping. These emitted photons can be absorbed by other ions in the trap causing any information they stored to be potentially destroyed.

The natural solution to both of these problems is to simply move the ions sufficiently far apart so that the above affects are negligible. However, this still leaves open the question of exactly how large the ion separation should be: too far and you are unnecessarily reducing the number of ions in the chain, but not far enough and then your qubit operations begin to alter the states of other ions in the chain. Hence we need a way of characterizing the effects of an operation such as state detection on the other non-target ions in the system.

In order to characterize the affect on a neighbouring ion of a state detection procedure there are three relevant quantities we will look at:

1. The probability that a neighbouring ion will absorb a photon emitted during fluorescence of the target ion.
2. The state of the neighbouring ion after being exposed to an attenuated detection beam for a specified length of time.
3. The probability of a successful state detection of a target ion given a specified detection time.

The first two quantities above will be used to characterize the affect on a neighbouring ion during a detection procedure. The third quantity is necessary as there is a balancing that must be done when selecting the detection time: the detection time needs to be long enough to ensure a successful detection, while short enough to minimize the destructive effects on the neighbouring ion.

While the three mentioned quantities above are not specific to using Ytterbium, for simplicity this chapter will restrict itself to considering the case of Ytterbium 171 while noting that many of the concepts discussed can easily be generalized to different species. This chapter will proceed as follows Section 2.1 will review the use of a Ytterbium ion as a qubit. Section 2.2 will describe how to approximate the probability a neighbouring ion will absorb a photon during the fluorescence of the target ion. In section 2.3 details how to simulate the attenuated detection beam leaking to the neighbouring using the Lindblad master equation. Section 2.4 shows how to determine the probability of a successful state detection of a target ion given a specified detection time. Then finally this chapter is concluded in section 2.5 where we walk through an example using all three of these concepts.

## 2.1 Ytterbium as a qubit

This section is intended as a brief review of using Ytterbium 171 as a qubit. For a much more detailed exploration see [43]. In figure 2.1 below we see the relevant energy levels and decay paths for the Ytterbium ion. As seen in the figure, for the entirety of this article the energy levels of the ion will be numbered in ascending order according to their energy. Note that the states $|0\rangle$ and $|2\rangle$ are used as the qubit states of the ion, so the state $|2\rangle$ correspond to the 1 qubit state. Figure 2.1 is again the energy level diagram of Ytterbium, but with the relevant energy splittings indicated.

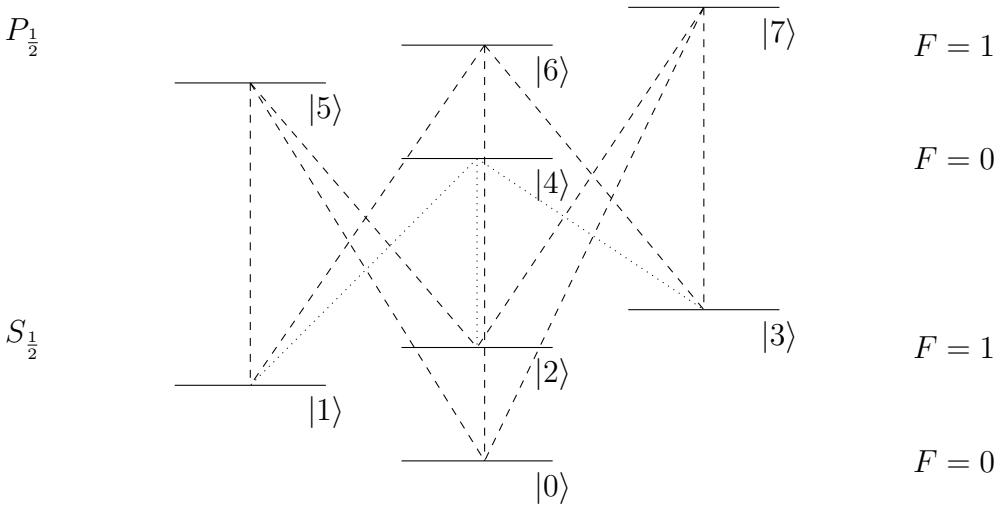


Figure 2.1: Energy level diagram of Ytterbium with each state labeled, and the possible transitions indicated. Note that several of the transitions, e.g. $|0\rangle \rightarrow |4\rangle$, are forbidden by selection rules. Each transition has a branching ratio of $\frac{1}{3}$.

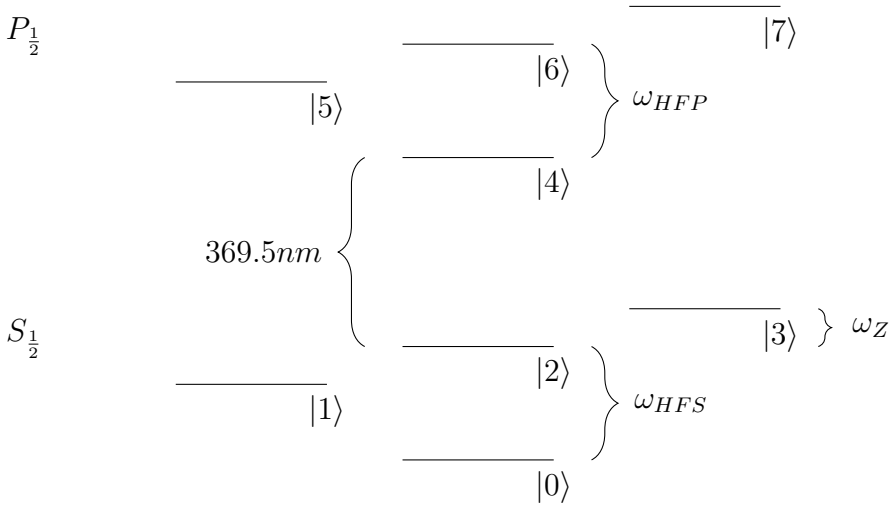


Figure 2.2: The energy level diagram for Ytterbium 171 with relevant energy gaps indicated. The frequencies for the hyperfine splittings are $\omega_{HFS} = 2\pi\, 12.6GHz$ and $\omega_{HFP} = 2\pi\, 2.1GHz$. The value of the Zeeman splitting $\omega_z$ will vary.

A simple detection procedure consists of hitting the ion with $369.5nm$ light to drive the $|2\rangle \rightarrow |4\rangle$ transition. This will cause the ion to fluoresce if it's in the $|2\rangle$ state, and remain dark if it's in the $|0\rangle$. In this article we will always assume that the detection consists of equal amounts of the three polarizations of light: $\pi$, and $\sigma^{\pm}$. Therefore as short hand we will refer to the three states $|1\rangle$, $|2\rangle$ and $|3\rangle$ as bright states, and $|0\rangle$ as the dark state.

## 2.2 Fundamental limit

Consider the case where a Ytterbium ion is undergoing a detection cycle (i.e. a beam of a wavelength $396.5nm$) for a time $\tau_d$ with a neighbouring ion at a distance $d$ and initially in the state $|2\rangle$. Assume that the distance between the ions is small enough such that the neighbouring ion will undergo decoherence through absorption of photons emitted by the fluorescing ion. Under the assumption that if the neighbouring ion absorbs a single photon that any information it contains is destroyed we may characterize the effect of this decoherence by simply computing the probability that the ion absorbs a photon.

To begin let $R_{sc}$ be the scattering rate of the fluorescing ion and $f$ being the absorption efficiency (i.e. the fraction of the emitted photons the neighbouring ion can absorb) we have the following expression for the absorption rate

$$R_{ab} = fR_{sc}. \tag{2.1}$$

For the scattering rate $R_{sc}$, we may assume the worst case scenario which corresponds to the optimal scattering rate of the fluorescing ion. This occurs when the Rabi frequency $\Omega$ of the transition driven by the $369.5nm$ light is twice the Zeeman splitting $\omega_Z$. The optimal scattering rate taking into account coherent dark states can be found in [25, 10]. Assuming no detuning of the detection beam we have

$$R_{sc} = \left(\frac{\Gamma}{6}\right) \frac{s_0}{1 + \frac{2}{3}s_0}, \tag{2.2}$$

where $\Gamma = 2\pi 19.6MHz$ is the spontaneous emission rate and $s_0 = 2\Omega^2/\Gamma^2 = 8\delta^2/\Gamma^2$.

Due to the fact the neighbouring ion is in the state $|2\rangle$, the only polarization of light that can lead to excitation into the state $|4\rangle$ is $\pi$ polarization. The absorption efficiency $f$ is then determined by two quantities: the fraction of all photons that pass through the absorption cross section of the neighbouring ion $f_{tot}$, and the fraction of those that are $\pi$ polarization $f_\pi$. And hence we have

$$f = f_{tot} f_\pi. \tag{2.3}$$

The formula for the absorption cross section can be found in [7]

$$\sigma_{ab} = \frac{\lambda^2}{2\pi}\frac{2F'+1}{2F+1}\varepsilon, \tag{2.4}$$

where $\lambda$ is the wavelength of the transition, $F'$ is the total angular momentum of the excited state, $F$ is the total angular momentum of the ground state, and $\varepsilon$ is the branching ratio of excited state to ground state transition. Assuming that the cross section is circular we let $a$ be its radius. In our case $\lambda = 369.5nm$, $F' = 0$, $F = 1$, and $\varepsilon = 1/3$ and so we have $a = 27.7nm$.

To determine $f_{tot}$, consider a sphere of radius $r$ centered at the fluorescing ion and extending so that the very edge of the absorption cross section of the neighbouring ion is touching the surface of the sphere. The projection of the absorption cross section onto the sphere creates a spherical cap of surface area $2\pi r^2(1 - \cos\alpha)$, where $\alpha$ is given by $\tan\alpha = a/d$ where $d$ is the ion separation. Assuming that the density of photons passing through the sphere is constant along its surface, $f_{tot}$ will be the ratio of the surface area of the spherical cap to the surface area of the entire sphere, which is

$$f_{tot} = \frac{1}{2}(1 - \cos\alpha) \tag{2.5}$$

or using the small angle approximation

$$f_{tot} \approx \frac{1}{4}\alpha^2 \approx \frac{1}{4}\left(\frac{a}{d}\right)^2. \tag{2.6}$$

To determine $f_\pi$ note an ion in the state $|4\rangle$ will decay equally to the three bright states, while the transition to the dark state is forbidden. While the total intensity of all light emitted is isotropic, the intensity of each individual polarization is not. Therefore we will need to find an expression for the intensity of $\pi$ polarization. We do so following section 3.8 of [7].

With the fluorescing ion set at the origin and the quantization axis set along $\hat{z}$ and perpendicular to the ion chain. Let $\hat{k}$ be the propagation direction, which is radially outward from the origin. Neglecting the existence of the neighbour ion, this system has cylindrical symmetry and hence it is sufficient to just consider emission in the $x - z$ plane. We choose $\hat{y}$ and $\hat{\theta} = \cos\theta\hat{x} - \sin\theta\hat{z}$ to be basis vectors orthogonal to the propagation direction. The polarization vectors of the light are $\hat{z}$ and $\varepsilon_\pm = \frac{1}{\sqrt{2}}(\hat{x} \pm \hat{y})$.

We have the following relationships for the intensity of each polarization of light

$$I^{\pi} \propto |\hat{y} \cdot \hat{z}|^2 + |\hat{\theta} \cdot \hat{z}|^2 = \sin^2 \theta \tag{2.7}$$

$$I^{\sigma^+} \propto |\hat{y} \cdot \varepsilon_+|^2 + |\hat{\theta} \cdot \varepsilon_+|^2 = \frac{1}{2}(1 + \cos^2 \theta) \tag{2.8}$$

$$I^{\sigma^-} \propto |\hat{y} \cdot \varepsilon_-|^2 + |\hat{\theta} \cdot \varepsilon_-|^2 = \frac{1}{2}(1 + \cos^2 \theta). \tag{2.9}$$

As previously stated, the amount of each polarization emitted by the fluorescing ion should be equal. Therefore we can compute the proportionality constant for each intensity by selecting a normalization condition and integrating all three equations separately over the sphere. However, when doing this calculation it becomes clear the proportionality constant will be the same in all three cases. Hence it is simpler and sufficient for our purposes to forgo this step and use the above equations to determine relative amounts of each polarization. In this instance we are strictly interested in the small $\theta$ case where it can be seen the amount of $\pi$ polarized photons is approximately equal to the amount of $\sigma$ polarized photons. Hence

$$f_\pi = \frac{1}{2} \tag{2.10}$$

and we now have all we need to compute $R_{ab}$.

To determine the probability $P(t)$ that the neighbouring ion will have absorbed a photon after time $t$, note that absorption rate $R_{ab}$ gives the expected number of photons absorbed per second. In other words, $1/R_{ab}$ is the average time needed in order for a single photon to be absorbed. This should be constant in time and hence the probability distribution is by definition exponential with a probability density function of $R_{ab}e^{-R_{ab}t}$. Integrating this from 0 to $\tau_d$ we find that the probability that the ion absorbs a photon after time $\tau_d$ is

$$P(\tau_d) = 1 - e^{-R_{ab}\tau_d}. \tag{2.11}$$

## 2.3 Technical limit

In order to determine the effect of beam leakage on the neighbouring ion our goal is to use the Lindblad master equation to allow us to solve directly for the state of the neighbour. First we will need the Hamiltonian of our system. As in this instance we are only considering the case of detection we will strictly consider a five state system consisting of states $|0\rangle$ through $|4\rangle$. This will of course lead to a neglect of any off-resonant

effects, however as we are considering the effect of the beam hitting the neighbouring ion, which is already experiencing a severely attenuated beam, these effects can be considered negligible. We also note here that while the following analysis considers only five states, and the operation of detection. Generalizing to consider the entire $P_{\frac{1}{2}}$ shell as well as the operation of optical pumping is straightforward.

Now, we move on to construct our Hamiltonian. The diagoanl terms will of course consist of the five energy levels of the states $|0\rangle$ through $|4\rangle$. The off-diagonal terms of the Hamiltonian will have generic form of $\frac{\Omega}{2}e^{-i\omega t}$, where $\omega$ is the angular frequency of the laser driving the transition and $\Omega$ is the Rabi frequency. The Rabi frequency for the transition can be computed, assuming the intensity of the laser is known, through the identity

$$\frac{I}{I_{sat}} = \frac{2\Omega^2}{\Gamma^2}, \tag{2.12}$$

where $I$ is the intensity of the laser, $I_{sat}$ is the saturation intensity, and $\Gamma$ is the spontaneous emission rate of the transition. As we are considering detection, the laser will be driving the excitations of the bright states to $|4\rangle$. Hence the Hamiltonian may be expressed as

$$H = \begin{pmatrix} \omega_0 & 0 & 0 & 0 & 0 \\ 0 & \omega_1 & 0 & 0 & \frac{1}{2}\Omega_d^- e^{-i\omega_d t} \\ 0 & 0 & \omega_2 & 0 & \frac{1}{2}\Omega_d^0 e^{-i\omega_d t} \\ 0 & 0 & 0 & \omega_3 & \frac{1}{2}\Omega_d^+ e^{-i\omega_d t} \\ 0 & \frac{1}{2}\bar{\Omega}_d^- e^{i\omega_d t} & \frac{1}{2}\bar{\Omega}_d^0 e^{i\omega_d t} & \frac{1}{2}\bar{\Omega}_d^+ e^{i\omega_d t} & \omega_4 \end{pmatrix}, \tag{2.13}$$

where $\omega_i$ is the energy level for the $i$-th state. Since $\omega_0$ is our lowest energy state we may set $\omega_0 = 0$. We also note that the difference in energy between the states $|1\rangle, |2\rangle, |3\rangle$ is on the order of MHz, whereas the energy of the excitation is several orders of magnitude larger, and so all three Rabi frequencies will be approximately the same and so we set

$$\Omega_d^- \approx \Omega_d^0 \approx \Omega_d^+ \approx \Omega. \tag{2.14}$$

Therefore our Hamiltonian is now

$$H = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \omega_1 & 0 & 0 & \frac{1}{2}\Omega e^{-i\omega_d t} \\ 0 & 0 & \omega_2 & 0 & \frac{1}{2}\Omega e^{-i\omega_d t} \\ 0 & 0 & 0 & \omega_3 & \frac{1}{2}\Omega e^{-i\omega_d t} \\ 0 & \frac{1}{2}\bar{\Omega} e^{i\omega_d t} & \frac{1}{2}\bar{\Omega} e^{i\omega_d t} & \frac{1}{2}\bar{\Omega} e^{i\omega_d t} & \omega_4 \end{pmatrix}. \tag{2.15}$$

Note that the Hamiltonian will of course be time dependent, however, for our purposes, it will be useful to remove the time dependence through a rotating transform. The rotating

transform is a unitary matrix $U(t)$, which we emphasize will depend on time, such that the resulting matrix $UHU^\dagger$ has no time dependence. In other words, we wish to transform to a time dependent basis where our Hamiltonian in this new basis is time independent. The Hamiltonian of our system in the rotated basis is given by

$$H_{rot} = UHU^\dagger - iU\frac{d}{dt}U^\dagger. \tag{2.16}$$

Due to the simplicity of our Hamiltonian, $U$ can be found by simple inspection or through a small amount of trial and error with the assistance of computer software. However in more general cases such having both detection and optical pumping beams on simultaneously $U$ may not be as obvious. Hence below we will describe a more algorithmic approach to determine $U$ that should be suitable in most cases.

We begin by first assuming that $U$ is diagonal, an assumption that can be justified by showing that it leads to an appropriate $U$. Hence for our basis states $\{|k\rangle\}_{k\in\mathbb{Z}_4}$ the action of $U$ on each basis state can be described by

$$U|k\rangle = u_k|k\rangle \tag{2.17}$$

where $\{u_k(t)\}_{k\in\mathbb{Z}_4}$, the diagonal entries of $U$, are some functions of time. As $U$ is intended to remove the time dependence from $H$ we generate the following three equations that need to be solved

$$\frac{1}{2}\Omega e^{-i\omega_d t}U\,|1\rangle\langle 4|\,U^\dagger = \frac{1}{2}\Omega e^{-i\omega_d t}u_1u_4^\dagger\,|1\rangle\langle 4| = \frac{1}{2}\Omega\,|1\rangle\langle 4| \tag{2.18}$$

$$\frac{1}{2}\Omega e^{-i\omega_d t}U\,|2\rangle\langle 4|\,U^\dagger = \frac{1}{2}\Omega e^{-i\omega_d t}u_2u_4^\dagger\,|2\rangle\langle 4| = \frac{1}{2}\Omega\,|2\rangle\langle 4| \tag{2.19}$$

$$\frac{1}{2}\Omega e^{-i\omega_d t}U\,|3\rangle\langle 4|\,U^\dagger = \frac{1}{2}\Omega e^{-i\omega_d t}u_3u_4^\dagger\,|3\rangle\langle 4| = \frac{1}{2}\Omega\,|3\rangle\langle 4| \tag{2.20}$$

Therefore equivalently we have

$$e^{-i\omega_d t}u_1u_4^\dagger = 1 \tag{2.21}$$

$$e^{-i\omega_d t}u_2u_4^\dagger = 1 \tag{2.22}$$

$$e^{-i\omega_d t}u_3u_4^\dagger = 1. \tag{2.23}$$

$$\tag{2.24}$$

It is obvious now that these equations may be satisfied when $u_4 = e^{-i\omega_d t}$ and all other $u_k$ are 1. Hence $U$ in this instance is given by

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & e^{-i\omega_d t} \end{pmatrix}. \tag{2.25}$$

Our rotated Hamiltonian therefore is

$$H_{rot} = UHU^\dagger - iU\frac{d}{dt}U^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \omega_1 & 0 & 0 & \frac{1}{2}\Omega \\ 0 & 0 & \omega_2 & 0 & \frac{1}{2}\Omega \\ 0 & 0 & 0 & \omega_3 & \frac{1}{2}\Omega \\ 0 & \frac{1}{2}\bar{\Omega} & \frac{1}{2}\bar{\Omega} & \frac{1}{2}\bar{\Omega} & \omega_4 - \omega_d \end{pmatrix}. \tag{2.26}$$

From this we see that we may remove the time dependence of the Hamiltonian by performing an appropriate energy shift. With the added assumption that our detection beam is on resonance we have the final form of our Hamiltonian in terms of more physically relevant variables we have

$$H_{rot} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \omega_{HF} - \omega_Z & 0 & 0 & \frac{1}{2}\Omega \\ 0 & 0 & \omega_{HF} & 0 & \frac{1}{2}\Omega \\ 0 & 0 & 0 & \omega_{HF} + \omega_Z & \frac{1}{2}\Omega \\ 0 & \frac{1}{2}\bar{\Omega} & \frac{1}{2}\bar{\Omega} & \frac{1}{2}\bar{\Omega} & \omega_{HF} \end{pmatrix}. \tag{2.27}$$

In the case of the neighbouring ion being exposed to an attenuated detection beam. The evolution of the system will be non-unitary due to spontaneous emission from the state $|4\rangle$ to the three bright states. In other words, our system is not entirely isolated as it interacts with the environment. As the evolution of our system is non-unitary we can not determine the dynamics through the Schrodinger equation. Instead we shall use the Lindblad master equation. The master equation is given by

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}[H_{rot}, \rho] + \sum_n \frac{1}{2}[2C_n\rho C_n^\dagger - \rho C_n C_n^\dagger - C_n^\dagger C_n \rho] \tag{2.28}$$

where $C_n$ are the collapse operators for our system and describe the interaction of our system to the environment. Our system has three decay paths from the state $|4\rangle$, each of

which are equally likely. Therefore we have three collapse operators $C_1, C_2, C_3$ which are given by

$$C_k = \sqrt{\frac{1}{3}\Gamma}\,|k\rangle\langle 4|\,, \quad k = 1, 2, 3. \tag{2.29}$$

where $\Gamma$ is the spontaneous emission rate from the state $|4\rangle$.

With this we now have all that is needed to simulate an attenuated detection beam hitting the neighbouring. And so, in order to determine the affect of the beam on the neighbour, we may solve the Lindblad master equation using the computer software of our choice to determine the state of the system $\rho(\tau_d)$, where $\tau_d$ is the time corresponding to the length of detection. Now that we know the state of the neighbouring ion after the detection procedure is completed we can characterize the effect on the neighbour by determining the state fidelity $\mathcal{F}(\rho(\tau_d), \rho(0))$ with the initial state of the neighbouring ion $\rho(0)$.

## 2.4 Detection Fidelity

To determine the detection fidelity (i.e. the probability of a successful measurement) we will follow a procedure outlined in [3] and apply it to our specific case of Ytterbium. This procedure can be similarly found in [10] where it is experimentally demonstrated on a Ytterbium ion.

To evaluate our detection procedure we wish to determine the probability that we successfully detect a qubit state on our target ion after running the detection procedure for a time $\tau_d$. As was previously described the detection procedure is quite simple as it merely consists of hitting the ion with light of wavelength $369.5nm$ to drive the bright state fluorescence. There are however some complications. For one any physical system can not reasonably be expected to collect all photons emitted by the target ion during its fluorescence with 100% efficiency. The second is that off resonant scattering can lead to the target ion transitioning from a bright state to the dark state and vice versa.

We begin by first defining some parameters that will be useful. The dark state leakage parameter $\alpha_0$ is defined as

$$\alpha_d := \frac{R_d}{R_{sc}} \tag{2.30}$$

where $R_{sc}$ is the scattering rate for the ion initially in the state $|2\rangle$ ignoring off resonant effects, and $R_d$ is the expected rate that an ion initially in the dark state will leak into one

of the bright states. We similarly have the bright state leakage parameter $\alpha_1$ given by

$$\alpha_b := \frac{R_b}{R_{sc}} \tag{2.31}$$

where $R_b$ is the rate that an ion in one of the bright states will leak into the dark state. The last parameter we need is the expected number of photons detected $\lambda$ given by

$$\lambda := \tau_d \eta R_{sc} \tag{2.32}$$

where $\eta$ is the collection efficiency of the detector. These parameters will allow use to compute the probability $p_d(n)$ that we detect $n$ photons for an ion initially in the dark state. And similarly the probability $p_b(n)$ that we detect $n$ photons for an ion initially in the bright state $|2\rangle$. We now recall their corresponding formulas below

$$p_d(n) = e^{\left(-\frac{\alpha_d}{\eta}\lambda\right)} \left[\delta_n + \frac{\alpha_d/\eta}{\left(1-\frac{\alpha_d}{\eta}\right)^{n+1}}\right] \mathcal{P}\left(n+1, \left(1-\frac{\alpha_d}{\eta}\right)\lambda\right) \tag{2.33}$$

$$p_b(n) = \frac{1}{n!} e^{-\left(1+\frac{\alpha_b}{\eta}\lambda\right)} \lambda^n + \frac{\alpha_b/\eta}{\left(1+\frac{\alpha_b}{\eta}\right)^{n+1}} \mathcal{P}\left(n+1, \left(1+\frac{\alpha_b}{\eta}\right)\lambda\right) \tag{2.34}$$

where $\mathcal{P}$ is the gamma incomplete function and $\delta_n$ is the Kronecker delta.

To compute $R_{sc}$ we may use the population formulas given in [3] which take into count population trapping in coherent dark states. Assuming the detuning of the detection beam is 0 we have

$$R_{sc} = \frac{\Gamma}{6} \frac{s_0}{1 + \frac{2}{3}(\frac{\Omega^4}{4\omega_z^2} + 4\omega_z^2)/\Gamma^2} \tag{2.35}$$

where $\Gamma = 2\pi 19.6 MHz$ is the spontaneous emission rate, $\omega_z$ is Zeeman splitting of the bright states, $\Delta$ is the detuning of the beam, and $s_0 = I/I_{sat}$. For the scattering rates $R_d$ and $R_b$ we no longer need to worry about coherent dark states and so there rates are determined purely through the off resonant scattering rates of their transitions. We have

$$R_d = \frac{2}{9}\frac{\Gamma}{2}\frac{s_0}{1 + s_0 + 4\frac{\Delta_d^2}{\Gamma^2}} \approx \frac{2}{9}\frac{s_0}{4}\left(\frac{\Gamma^2}{\Delta_d^2}\right) \tag{2.36}$$

$$R_b = \frac{2}{3}\frac{\Gamma}{2}\frac{s_0}{1 + s_0 + 4\frac{\Delta_b^2}{\Gamma^2}} \approx \frac{2}{3}\frac{s_0}{4}\left(\frac{\Gamma^2}{\Delta_b^2}\right) \tag{2.37}$$

where $\Delta_d = \omega_{HFS} + \omega_{HFP} = 2\pi\, 14.7GHz$ and $\Delta_b = \omega_{HFP} = 2\pi\, 2.1GHz$. In this case we may use the approximate expressions for scattering rates as the detunings $\Delta_b$ and $\Delta_d$ are significantly larger than the spontaneous emission rate. The factor of $\frac{2}{9}$ is the probability that once the ion in the bright state absorbs a photon that it will then decay to the dark state. The factor of $\frac{2}{3}$ would be the probability that the ion in the dark state will leak to the bright state once it absorbs a photon.

We now set a threshold for the number of detected photons $n_d$ where upon detection of $n_d$ or more photons we say the state of the ion was $|2\rangle$ otherwise we say it was $|0\rangle$. The above formulas for the probabilities will allow us to create histograms from the photon detection count, which when combined with our established threshold will allow us to determine the probability of a successful detection. Note however that the probability for a successful detection will differ based on whether we started in a bright state or the dark state. Hence we will characterize the success chance by considering the average $\mathcal{F}_{avg}$ of these two success rates.

## 2.5 Example

We now conclude this chapter by fixing a set of experimental parameters and then seeing how the quantities previously discussed vary as the detection time increases. We choose $s_0 = I/I_{sat} = \frac{1}{2}$ and the Zeeman splitting to be half the Rabi frequency of the detection beam, hence

$$w_Z = \frac{\Omega}{2} = \frac{1}{2}\sqrt{\Gamma^2 \frac{s_0}{2}} = \frac{1}{4}\Gamma. \tag{2.38}$$

The beam we may assume is centered on the target ion, and as it is assumed to be Gaussian it decays as $\exp(-2r^2/w_0^2)$ where $r$ is the distance from the center and $w_0$ is the beam width. If we set the beam width to be 2 microns then an ion separation of $2.5w_0$ corresponds to an attenuation of 0.000037 on the neighbouring ion. Lastly we assume that the detection efficiency is $\eta = 0.01$.

The probability of absorption is given in eq. (2.11), where here we shall instead use the probability that the neighbour does NOT absorb a photon. To determine the state of the neighbouring ion $\rho(\tau_d)$ after detection time $\tau_d$ we solve the Lindblad master equation using the mesolve function in the QuTiP software package. Since the initial state of the neighbour is $|2\rangle$ we have that the state fidelity will be the $(2,2)$ entry of $\rho(\tau_d)$. For the detection fidelity we first compute the photon count histogram as seen in figure 2.3

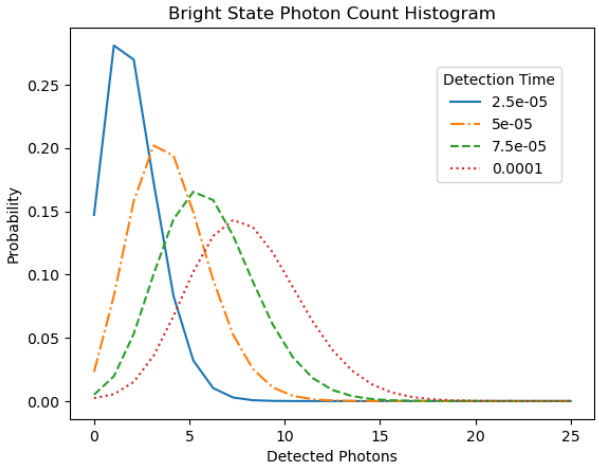


Figure 2.3: Some select histograms for the probability of detecting $n$ photons for $s_0 = 1/2$, $\eta = 0.01$, and assuming optimal scattering rate of the target ion.

We shall forgo showing the dark state histograms as they all have very low probability of emitting 1 or more photons. In figure 2.4 we show the plot of bright state detection fidelity and dark detection fidelity as a function of detection threshold for a detection time of $75\mu s$.

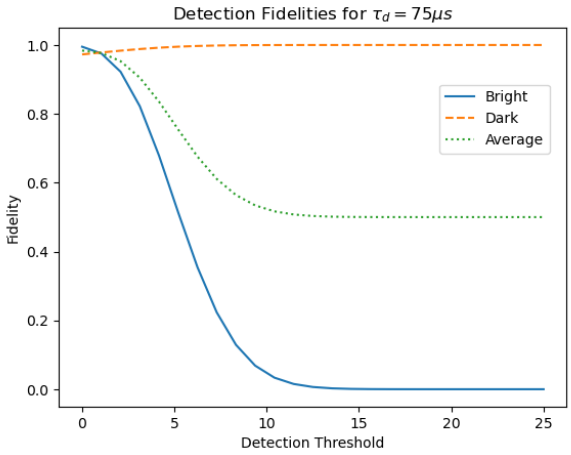


Figure 2.4: The bright state detection fidelity, and dark state detection fidelity as a function of detection threshold for the histogram corresponding to $\tau_d = 75\mu s$ seen in figure 2.3

And finally we conclude with figure 2.5 showing how all three of quantities we have

thus far discussed vary for detection times ranging from $1\mu s$ to $100\mu s$.

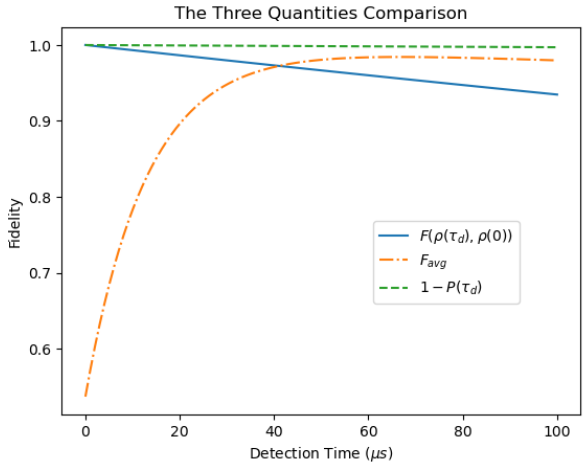


Figure 2.5: A plot showing how the probability the neighbouring ion does not absorb a photon $1 - P(\tau_d)$, the average detection fidelity $F_{avg}$, and the neighbour state fidelity $\mathcal{F}(\rho(\tau_d), \rho(0))$ vary as a function of detection time.

## 2.6 Discussion

To summarize, in this chapter we explored ways of characterizing the adverse effects of photon absorption and beam leakage on neighbouring ions, and how to determine the probability of a successful measurement. There is of course a balancing act that must be performed as attempts to improve the detection fidelity of the measurement, the simplest being an increase to detection time, can lead to increases in neighbouring ion decoherence. One of the shortest detection schemes thus far demonstrated is in [10], which attains a detection time of $11\mu s$ with a detection fidelity of over 99.9%. The example given in section 2.5 would then correspond to a slightly more modest detection time of approximately $50\mu s$ with detection fidelity 95% and chance of neighbouring ion decoherence smaller than 5%. While these numbers are likely too large/small to be of practical purpose, they were attained at a relatively small ion separation of $5\mu m$ with no attempt to improve detection fidelity or reduce ion decoherence. If we assume that the detection procedure in the previously stated example was able to be completed in the more optimistic time of $11\mu s$ this would lead to a neighbouring ion decoherence of less than 1%. These numbers can be improved in the obvious ways: slightly increased ion separation, improved detector photon

collection efficiency to reduce detection time, improved beam addressing, etc. However, there may be less obvious ways they can be improved, which we discuss below.

The fundamental limit may be improved through using the non-isotropic emission of photons by the target ion during it's fluorescence. The analysis of section 2.2 was assuming that the quantization axis was perpendicular to the ion chain as this was how the trap in [22] was constructed. This results in a neighbouring ion seeing roughly half $\pi$ polarized light and half $\sigma$. Equation 2.7 states however that the amount of $\pi$ polarized light parallel to the quantization axis will be almost negligible. Hence orienting the ion chain parallel to the quantization axis can greatly reduce decoherence as an ion in the state $|2\rangle$ can only absorb $\pi$ polarized light, excluding off-resonant effects. To implement this would require consideration when engineering the trap, and may not be necessary depending on which of the fundamental, or the technical limit, are the limiting factors of the system.

Sainath Motlakunta, a colleague, suggested applying a stark shift to the target ion during state detection. A stark shift beam is a beam of high intensity that is largely detuned from the $369.5nm$ transition and focused purely on the target ion causing a slight energy shift in the energy levels of the target. This allows the detection beam, and consequently any emitted photons, to be off-resonant to neighbouring ions thereby reducing decoherence. The downside is it increases the chances of the bright states leaking to the dark state, and vice versa, reducing the detection fidelity. Our initial calculations for Ytterbium suggest that the reduction in the detection fidelity is not worth the improved neighbour fidelity. It's possible that this trade-off may be worth it in a different species of ion.

We close this chapter by mentioning that the methods used in this chapter are by no means exhaustive and still leave open what are the best ways of characterizing the errors in state detection. As an example, the previously mentioned reference [10] analyzes the decoherence of the neighbouring ion through a T2 experiment. The method in section 2.3 can easily be adapted to simulate such an experiment if desired. However, this leaves several parameters that may be used to evaluate a state detection procedure: probability of photon absorption, T2 time, state fidelity of the neighbour, detection fidelity, among potentially others. What are the best parameters for characterizing the detection process? And what is the best way to combine the desired parameters into a single number giving the overall quality of the measurement? We now turn to this question in the next chapter.

# Chapter 3

# Figures of Merit for Quantum Instruments

## 3.1 Introduction

In the preceding chapter we examined errors in a measurement process specific to trapped ions, we now turn to the abstract setting of evaluating quantum measurements without assuming a specific implementation of a quantum computer. The performance of any quantum computer will naturally be limited by any errors in its physical implementation. While error correcting procedures of course exist, the catch is any error correcting procedure must be implemented with operations which in themselves are error prone. This naturally led to the development of fault-tolerant quantum computation [34, 31, 21, 1]. Fault-tolerant quantum computation was shown to be possible by proving that there exists a maximum error rate $\eta$ of the quantum operations of the computer below which errors can be suppressed to an arbitrarily small rate using additional quantum operations and processing time. In other words, if the error per operation in a large-scale quantum computer is sufficiently small, then quantum error correcting procedures can be implemented effectively.

Therefore, to determine whether we are close to the fault-tolerant threshold, we need figures of merit to quantify the error per operation and methods to estimate those figures of merit for a given quantum computer. For unitary operations, two common figures of merit are the diamond distance and the process infidelity between the ideal operation and its implementation, with the diamond distance being the figure of merit of choice for fault-tolerant quantum computation. The diamond distance is typically estimated by reconstructing the full process using quantum process tomography [9] or gate set to-

mography [16] and computing the diamond distance between the reconstructed and ideal processes. However, the cost associated to tomography notoriously grows exponentially with the size of the system, and there are no known methods of efficiently estimating the diamond distance for general noise processes. In contrast, the process infidelity can be efficiently estimated using techniques such as randomized benchmarking and its many variants [17], or direct fidelity estimation [11]. However, while the average gate fidelity can be used to determine an upper bound on the diamond distance and hence on $\eta$ [33], the upper bound is generally very loose as the diamond distance and the process infidelity are not directly related.

All known fault-tolerant techniques that can arbitrarily suppress errors require quantum non-demolition measurements, which must preserve the state of the quantum system [37, 29, 13]. As quantum non-demolition measurements are represented by quantum instruments it becomes necessary to develop figures of merit capable of properly evaluating them. There are then two sources of error in quantum instruments, namely, errors in the probabilities of outcomes, and errors in the post-measurement states. It is therefore important to ensure that any method used to evaluate a quantum instrument accurately represents both types of error. Several groups having been using modifications of known tomographic methods to fully characterize the quantum instrument [32, 36] or applying device-independent certification methods based on violations of Bell inequalities [38].

This chapter will proceed as follows. Section 3.2 and section 3.3 will provide a brief review of figures of merits on states and quantum channels, respectively. Section 3.4 continues with a review of quantum instruments and proves some intermediary results. In section 3.5 we analyse how the different figures of merit on quantum instruments adequately capture errors in measurements. Section 3.6 discusses a diamond distance equivalent for quantum instruments and presents an argument motivating its use. Section 3.7 generalizes bounds on the diamond distance for unitary channels to bounds on the diamond distance for quantum instruments. This chapter then concludes in section 3.8 by walking through several examples computing the different figures of merit for various error models.

## 3.2 Figures of Merit on Quantum States

We begin by first reviewing some elementary definitions and results from linear algebra. First, recall that the trace norm $\|\cdot\|_1$, otherwise known as the Schatten-1 norm, on a

matrix $A$ is given by

$$\|A\|_1 := \operatorname{Tr}\sqrt{A^\dagger A} = \operatorname{Tr}|A| = \sum_n \sigma_n(A), \tag{3.1}$$

where $\{\sigma_n\}_n$ is the set of singular values for the matrix $A$. Secondly, for a matrix $A$ a square root $B$ is defined as a matrix such that $A = B^2$. In the case where $A$ is positive semi-definite then there exists a unique positive square root, referred to as the square root. And lastly, for a $d \times d$ square matrix $A$ of rank 1, then it may be decomposed as $A = |\psi\rangle\langle\psi|$ for some vector $|\psi\rangle \in \mathbb{C}^d$

A $d$-dimensional quantum state (otherwise known as a density matrix) is a $d \times d$ positive semi-definite matrix of trace 1, with the set of all $d$-dimensional quantum states being denoted by $\mathbb{D}_d$. Hence any positive semi-definite matrix can be interpreted as an unnormalized quantum state, where unnormalized in this context refers to having potentially non-unit trace. A pure state $\rho$ is a quantum state of rank 1. A quantum state that is not pure, referred to as a mixed state, is given by a convex combination of pure states.

Generalizing the fidelity and trace distance between quantum states we define the following two functionals on the space $\mathrm{Pos}_d$ of all $d \times d$ positive semi-definite matrices (i.e. all unnormalized $d$-dimensional quantum states):

$$\mathcal{D}(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 \tag{3.2}$$

$$\mathcal{F}(\rho, \sigma) := \left(\operatorname{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2 = \|\sqrt{\rho}\sqrt{\sigma}\|_1^{2,} \tag{3.3}$$

where $\mathcal{D}$ is the trace distance, and $\mathcal{F}$ is the fidelity. In some cases it will be more convenient to work with $\sqrt{\mathcal{F}(\rho, \sigma)}$ rather than $\mathcal{F}(\rho, \sigma)$, at which point we refer to it as the square root fidelity. Additionally, it should be clear from there definitions that both functionals are symmetric in their inputs.

Note that in most contexts the above two functionals are used as a figure of merit (a measure of distinguishability) between two quantum states, and will be there primary use in this article. Presenting them in the more general context of positive semi-definite matrices is simply an expansion of their domain to include unnormalized quantum states. As we'll see later, this will be convenient as we will frequently encounter completely positive but NOT trace preserving maps. These maps input a quantum state and will, in general, output an unnormalized quantum state. This expansion of the domain will allow us to use these functionals as figures of merit, while avoiding renormalization during some intermediary steps as will be seen later.

There are two useful properties of both of these figures of merit, referred to as stability and contractivity (see [15]), which we recall in the following two lemmas.

**Lemma 1** (Stability)**.** *Let $\rho, \sigma, \tau \in \mathbb{D}_d$. Then*

$$\mathcal{F}(\rho \otimes \tau, \sigma \otimes \tau) = \mathcal{F}(\tau \otimes \rho, \tau \otimes \sigma) = \mathcal{F}(\rho, \sigma) \tag{3.4}$$
$$\mathcal{D}(\rho \otimes \tau, \sigma \otimes \tau) = \mathcal{D}(\tau \otimes \rho, \tau \otimes \sigma) = \mathcal{D}(\rho, \sigma) \tag{3.5}$$

**Lemma 2** (Contractivity)**.** *Let $\mathcal{E}$ be a trace preserving quantum operation and $\rho, \sigma \in \mathbb{D}_d$. Then*

$$\mathcal{F}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \geq \mathcal{F}(\rho, \sigma) \tag{3.6}$$
$$\mathcal{D}(\mathcal{E}(\rho), \mathcal{E}(\sigma)) \leq \mathcal{D}(\rho, \sigma) \tag{3.7}$$
$$\tag{3.8}$$

It is well known that in the case for two (emphasis on normalized) quantum states $\rho$ and $\sigma$ that if at least one of them is pure then the fidelity simplifies to $\mathcal{F}(\rho, \sigma) = \operatorname{Tr} \rho\sigma$. The following lemma shows that this is true in the more general case where one of them is assumed proportional to a pure state.

**Lemma 3.** *Let $A, B \in \mathrm{Pos}_d$. If $A$ (or equivalently $B$) is rank 1, then*

$$\mathcal{F}(A, B) = \operatorname{Tr} AB \tag{3.9}$$

*Proof.* Since $A$ has rank 1 it follows through a straightforward calculation by expressing $A$ as $A = a \left|\psi\right\rangle\!\left\langle\psi\right|$ for some constant $a > 0$ and vector $\left|\psi\right\rangle \in \mathbb{C}^d$ with $\left|\left\langle\psi|\psi\right\rangle\right|^2 = 1$. The case where $B$ is rank 1 would follow immediately by symmetry. □

While there are many figures of merit for quantum states the trace distance and fidelity are related due to a result by Fuchs and van de Graaf [14] that allows them to be used as upper and lower bounds for each other. We recall this result in the following inequality.

**Theorem 1** (Fuchs-van de Graaf Inequality)**.** *Let $\rho, \sigma \in \mathbb{D}_d$. Then,*

$$1 - \sqrt{\mathcal{F}(\rho, \sigma)} \leq \frac{1}{2} \left\|\rho - \sigma\right\|_1 \leq \sqrt{1 - \mathcal{F}(\rho, \sigma)}$$

Note that it is easy to verify that if both states are pure then the right inequality is saturated (i.e. the right inequality becomes equality). It is also known that if one of the states is pure then the left inequality can be improved to

$$1 - \mathcal{F}(\rho, \sigma) \leq \frac{1}{2} \left\|\rho - \sigma\right\|_1 \tag{3.10}$$

## 3.3 Figures of Merit on Quantum Channels

Before discussing quantum channels it will be helpful to first review the process of vectorization, or, more specifically, column stacking. Let $A \in \mathbb{C}^{m\times n}$ be some rectangular matrix and $\{j\}_{j\in\mathbb{Z}_m}$ and $\{k\}_{k\in Z_n}$ be two sets of orthonormal bases for $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively. We may express $A$ in component form as

$$A = \sum_{j\in\mathbb{Z}_m, k\in\mathbb{Z}_n} a_{jk} \left|j\right\rangle\!\left\langle k\right|. \tag{3.11}$$

From which we can see that we may vectorize $A$ through the following map

$$\left|A\right\rangle\!\right\rangle := \sum_{j\in\mathbb{Z}_m, k\in\mathbb{Z}_n} a_{jk} |k\rangle \otimes |j\rangle, \tag{3.12}$$

which is straightforward to verify that it equates to stacking the columns of $A$ to create a single vector. We could of course similarly consider vectorizing $A$ through stacking its rows. For our purposes we only need a single vectorization map, and so from here on we will strictly consider column stacking. It is worth mentioning that many of the results we will see going forward can be straight forwardly adapted to the case of row stacking.

We now recall the following useful properties of vectorization. The proofs of which can be found elsewhere but we include here for completeness.

**Theorem 2.** *Let $A$ and $B$ be $n \times n$ matrices. Then $\left|AB\right\rangle\!\right\rangle = (I \otimes A)\left|B\right\rangle\!\right\rangle = (B^T \otimes I)\left|A\right\rangle\!\right\rangle$.*

*Proof.* The process of vectorization described above takes a matrix $M$ and creates a column vector by stacking the columns of $M$, or more succintly: $\left|M\right\rangle\!\right\rangle = ([m_i]_{i=1}^n)^T$. Where $m_i$ is the $m$-th column of $M$. Therefore $\left|AB\right\rangle\!\right\rangle$ can be simply expressed as $\left|AB\right\rangle\!\right\rangle = ([Ab_i]_{i=1}^n)^T$. By noting that $I \otimes A$ is an $n \times n$ block diagonal matrix where each diagonal element is the matrix $A$. The first equality $\left|AB\right\rangle\!\right\rangle = (I \otimes A)\left|B\right\rangle\!\right\rangle$ becomes immediate.

For the second equality, note that the product $B^T \otimes I$ is a $n \times n$ block matrix where the $(i, j)$-th entry is $b_{ji}I$ with $b_{ji}$ being the $(j, i)$-th entry of B. Athough $\left|A\right\rangle\!\right\rangle$ is technically a $n^2$ dimensional vector we can instead consider it to be $n$ dimensional vector where each component is the column of $A$ (based on the discussion above). For a fixed $i$ we have a row of $B^T \otimes I$ to be $(b_{1i}I, b_{2i}I, \ldots, b_{ni}I)$ and taking the dot product of this row with the vector $\left|A\right\rangle\!\right\rangle$ gives

$$b_{1i}a_1 + b_{2i}a_2 + \cdots + b_{ni}a_n = Ab_i \tag{3.13}$$

The above equation will be the $i$-th row of $(B^T \otimes I)\left|A\right\rangle\!\right\rangle$ and so it follows that $(B^T \otimes I)\left|A\right\rangle\!\right\rangle = \left|AB\right\rangle\!\right\rangle$ from the discussion at the beginning of the proof. □

**Corollary 1.** *For $n \times n$ matrices $A$, $B$, and $C$ we have $|ABC\rangle\!\rangle = (C^T \otimes A)\,|B\rangle\!\rangle$*

*Proof.* Follows from applying theorem 2 twice:

$$|ABC\rangle\!\rangle = (I \otimes A)\,|BC\rangle\!\rangle = (I \otimes A)(C^T \otimes I)\,|B\rangle\!\rangle = (C^T \otimes A)\,|B\rangle\!\rangle \tag{3.14}$$

□

**Corollary 2.** *Let $A$ and $B$ be $n \times n$ matrices. Then $\mathrm{Tr}(A^\dagger B) = \langle\!\langle A|B\rangle\!\rangle$.*

*Proof.* We have as an identity that $\mathrm{Tr}(AB) = \langle\!\langle I|AB\rangle\!\rangle$ and so

$$\begin{aligned}
\mathrm{Tr}(A^\dagger B) &= \langle\!\langle I\big|A^\dagger B\rangle\!\rangle \\
&= \langle\!\langle I|\,(I \otimes A^\dagger)\,|B\rangle\!\rangle \\
&= ((I \otimes A)\,|I\rangle\!\rangle)^\dagger\,|B\rangle\!\rangle \\
&= |AI\rangle\!\rangle^\dagger\,|B\rangle\!\rangle \\
&= \langle\!\langle A|B\rangle\!\rangle
\end{aligned} \tag{3.15}$$

□

Now, a quantum channel is defined as any map $\mathcal{A} : \mathbb{D}_d \to \mathbb{D}_f$ between quantum states, which by necessity must be completely positive and trace preserving (CPTP). In this context, a positive map $\mathcal{C}$ (i.e. a map that sends positive semi-definite operators to positive semi-definite operators) is referred to as completely positive if the induced maps $\mathcal{I}_k \otimes \mathcal{C}$ are also positive for every $k$, where $\mathcal{I}_k$ is the identity map on $\mathbb{C}^{k\times k}$. By virtue of a quantum channel being completely positive it may decomposed in terms of a set of operators $\{K_j\}_j$ referred to as Kraus operators, where the action of $\mathcal{A}$ on a state $\rho$ may be expressed as

$$\rho = \sum_j K_j^\dagger \rho K_j. \tag{3.16}$$

This Kraus decomposition of a channel is not unique and so we define the Kraus rank of a completely positive map to be the minimum number of Kraus operators needed in the decomposition. The set of all quantum channels which maps $\mathbb{D}_d$ to $\mathbb{D}_f$ is denoted $L(\mathbb{D}_d, \mathbb{D}_f)$.

Similarly to our treatment of quantum states, whereby we discussed figures of merit that may be used to compare how "close" two quantum states are to being equal, we also

wish to compare two quantum channels. The figures of merit we are interested in are derived through the use of the Choi-Jamiolkowski isomorphism.

First we define the following maximally entangled state $\phi \in \mathbb{D}_d \otimes \mathbb{D}_d$ given by

$$\Phi := \frac{1}{d} \sum_{j,k \in \mathbb{Z}_d} |jj\rangle\langle kk| = \frac{1}{d} |I_d\rangle\rangle\langle\langle I_d| \,. \tag{3.17}$$

where $I_d$ is the $d$-dimensional identity matrix. And so we may now define the Choi-Jamiolkowski isomorphism $J : L(\mathbb{D}_d, \mathbb{D}_f) \to \mathbb{D}_d \otimes \mathbb{D}_f$ where the action of $J$ on a channel $\mathcal{A}$ is given by

$$J(\mathcal{A}) := (\mathcal{I}_d \otimes \mathcal{A})\phi, \tag{3.18}$$

where $\mathcal{I}_d$ is the $d$-dimensional identity map. The outputted state $J(\mathcal{A})$ is referred to as the Choi state for the channel $\mathcal{A}$. Note that in most contexts, and as it is introduced here, the Choi Jamiolkowski isomorphism is presented as an isomorphism between quantum channels and quantum states [8, 18]. However, analogously to how we expanded the domain of the figures of merits on quantum states, it will be convenient to expand the domain of the map $J$ to allow in its domain completely positive but not necessarily trace preserving maps. At which point for a completely positive map $\mathcal{A}$ its corresponding Choi state $J(\mathcal{A})$ will in general be unnormalized.

Assuming that the Kraus operators of a CP map are known, we have the following lemma providing a relatively easier way to compute the Choi state.

**Lemma 4.** *For a CP map $\mathcal{C}$ with Kraus operators $\{C_k\}_k$ the Choi state may be expressed as*

$$J(\mathcal{C}) = \frac{1}{d} \sum_k |C_k\rangle\rangle\langle\langle C_k| \tag{3.19}$$

*Proof.* Using the distributivity of the tensor product, as well as the theorem 2 we get

$$\begin{aligned} J(\mathcal{C}) &= \sum_k J(C_k) \\ &= \frac{1}{d} \sum_k (I_d \otimes C_k) \, |I_d\rangle\rangle\langle\langle I_d| \, (I_d \otimes C_k^\dagger) \\ &= \frac{1}{d} \sum_k |C_k\rangle\rangle\langle\langle C_k| \,. \end{aligned} \tag{3.20}$$

□

The usefulness of the Choi-Jamiolkowski isomorphism is it allows us to associate to every channel a unique quantum state. Thus any figure of merit between quantum states may be promoted to a figure of merit on quantum channels by comparing their corresponding Choi states.

We may now define the following two functionals on the space of completely positive maps:

$$\mathcal{F}_J(\mathcal{A}, \mathcal{B}) := \mathcal{F}(J(\mathcal{A}), J(\mathcal{B})) \tag{3.21}$$

$$\mathcal{D}_J(\mathcal{A}, \mathcal{B}) := \mathcal{D}(J(\mathcal{A}), J(\mathcal{B})) \tag{3.22}$$

When these two functionals are restricted to quantum channels they may be used as figures of merit. While comparing two quantum channels by comparing their Choi states may at first seem arbitrary, it has the benefit however of our figures of merit immediately inheriting the nice properites associated to the fidelity and trace distance. Our primary motivation however for choosing these figures of merit is their use in applying upper and lower bounds on the diamond distance, which we will now define.

The above approach compares two channels by comparing their Choi states. However a more natural way may be to instead compare the outputs of the two channels. As we are always going to be considering completely positive maps, or combinations of them, whose inputs are positive semi-definite matrices. We wish to define an operator norm $\|\cdot\|_1$ on superoperators of the form $\mathcal{A} : \mathrm{Pos}_d \to \mathbb{C}^{f \times f}$. Defining the operator norm in the standard way we have

$$\|\mathcal{A}\|_1 = \sup\{\|\mathcal{A}(\rho)\|_1 : \rho \in \mathrm{Pos}_d, \quad \|\rho\|_1 = 1\}. \tag{3.23}$$

From the definition it should be clear that the maximization is done over all quantum states $\rho \in \mathbb{D}_d$. As $\rho$ is a quantum state it is either a pure state or a convex combination of some set of pure states which we denote $\{\sigma_k\}$. Letting $\mathcal{A} : \mathrm{Pos}_d \to \mathbb{C}^{f \times f}$ be an arbitrary superoperator. The quantity $\sigma = \max_k\{\|\mathcal{A}(\sigma_k)\|_1\}$ must exist, where $\sigma \in \{\sigma_k\}_k$. By the triangle inequality we have

$$\|\mathcal{A}(\rho)\|_1 \leq \|\mathcal{A}(\sigma)\|_1. \tag{3.24}$$

It follows then that maximization in our operator norm can equivalently be done by only considering pure states.

From the operator norm we may now define the diamond norm

$$\|\mathcal{A}\|_\diamond = \|\mathcal{I}_d \otimes \mathcal{A}\|_1. \tag{3.25}$$

From which we can then define diamond distance $r_\diamond$ between two quantum channels $\mathcal{A}$ and $\mathcal{B}$ as

$$r_\diamond(\mathcal{A},\mathcal{B}) = \frac{1}{2}\|\mathcal{A}-\mathcal{B}\|_\diamond. \tag{3.26}$$

The addition of the identity mapping in the definition of the diamond norm serves a vital purpose in that it allows the diamond norm, and by extension the diamond distance, to have a stability property [15] analogous to that seen for the state based figure of merits. We state this stability property in the following lemma.

**Lemma 5** (Stability)**.** *Let $\mathcal{A},\mathcal{B} \in L(\mathbb{D}_d,\mathbb{D}_f)$. Then*

$$r_\diamond(\mathcal{I}_d\otimes\mathcal{A},\mathcal{I}_d\otimes\mathcal{B}) = r_\diamond(\mathcal{A},\mathcal{B}) \tag{3.27}$$

All of the figures of merit thus discussed can be related through the following inequality (see [41])

**Lemma 6.** *Let $\mathcal{A}$ be a hermiticty-preserving map between d-dimensional operators. Then we have*

$$\|J(\mathcal{A})\|_1 \leq \|\mathcal{A}\|_\diamond \leq d\|J(\mathcal{A})\|_1. \tag{3.28}$$

We note that the upper and lower bound in terms of the trace norm may be exchanged for expressions involving the fidelity by using the Fuchs-van de Graaf inequality.

## 3.4 Quantum Instruments

A quantum instrument is a special case of a quantum channel in which it takes as input a quantum state and outputs a mixed state that stores all relevant measurement statistics. More explicitly, the action of an instrument $\mathcal{A}$ on a quantum state $\rho$ is given by

$$\mathcal{A}(\rho) = \sum_k \mathcal{A}_k(\rho)\otimes|k\rangle\langle k|\,. \tag{3.29}$$

The states $\{|k\rangle\}$ are referred to as the pointer states of the measurement, and are a set of orthonormal states that represent each measurement outcome. The completely positive maps $\mathcal{A}_k$ are referred to as the pointer maps, and are defined such that the post-measurement state, assuming measurement outcome $k$ is recorded, is given by $\frac{\mathcal{A}_k(\rho)}{\mathrm{Tr}\,\mathcal{A}_k(\rho)}$. The probability of

obtaining an outcome $k$ using the instrument $\mathcal{A}$ is determined through the trace: $\operatorname{Tr} \mathcal{A}_k(\rho)$. We emphasize here that the pointer maps are by necessity not trace preserving and hence their output will be unnormalized quantum states. However their sum $\sum_k \mathcal{A}_k$ is and hence is by definition a quantum channel.

We mention here briefly that indirect measurements, which are measurements consisting of pairing a system with an ancilla state, performing a unitary operation on the combined system, and then a projective measurement on the ancilla, are special cases of quantum instruments. The pointer maps in this case would be a composition of the following maps: the unitary channel on the combined system associated to the indirect measurement, a projection of the ancilla state onto the qubit state $|0\rangle$ or $|1\rangle$, and tracing out the ancilla state.

The usefulness of the quantum instrument representation is that our measurement is now described through a quantum channel. Hence any figure of merit on quantum channels is now also a figure of merit for measurements. Secondly, the particular form of the quantum instrument means the output of the instrument will always be a summation of mutually orthogonal states due to the pointer states. An immediate consequence of this is that the square root fidelity and trace distance become additive over the pointer maps as the following lemma and subsequent theorem shows.

**Lemma 7.** *Let $\{M_k\}_{k\in K}$ be a set of hermitian matrices of the same dimension and $\{|k\rangle\}_{k\in K}$ be a set of orthonormal vectors. Then*

$$\| \sum_k M_k \otimes |k\rangle\langle k| \|_1 = \sum_k \|M_k\|_1 \tag{3.30}$$

*Proof.* From the orthonormality of $\{|k\rangle\}_{k\in K}$ we get after a brief calculation

$$\left( \sum_k |M_k| \otimes |k\rangle\langle k| \right)^2 = \left( \sum_k M_k \otimes |k\rangle\langle k| \right)^2 . \tag{3.31}$$

Taking the square root of the above and applying the definition of the trace norm we get

$$\| \sum_k M_k \otimes |k\rangle\langle k| \|_1 = \operatorname{Tr} \left| \sum_k M_k \otimes |k\rangle\langle k| \right| \tag{3.32}$$

$$= \operatorname{Tr} \sum_k |M_k| \otimes |k\rangle\langle k| \tag{3.33}$$

$$= \sum_k \|M_k\|_1 \tag{3.34}$$

□

**Theorem 3.** *Let $\mathcal{A}$ and $\mathcal{B}$ be two quantum instruments with respective pointer maps $\{\mathcal{A}_k\}_k$ and $\{\mathcal{B}\}_k$ and let $\rho$ be a quantum state of appropirate dimension. Then we have the following*

$$\sqrt{\mathcal{F}_J(\mathcal{A},\mathcal{B})} = \sum_k \sqrt{\mathcal{F}_J(\mathcal{A}_k,\mathcal{B}_k)} \tag{3.35}$$

$$\mathcal{D}_J(\mathcal{A},\mathcal{B}) = \sum_k \mathcal{D}_J(\mathcal{A}_k,\mathcal{B}_k) \tag{3.36}$$

$$\sqrt{\mathcal{F}(\mathcal{A}(\rho),\mathcal{B}(\rho))} = \sum_k \sqrt{\mathcal{F}(\mathcal{A}_k(\rho),\mathcal{B}_k(\rho))} \tag{3.37}$$

$$\mathcal{D}(\mathcal{A}(\rho),\mathcal{B}(\rho)) = \sum_k \mathcal{D}(\mathcal{A}_k(\rho),\mathcal{B}_k(\rho)) \tag{3.38}$$

$$\tag{3.39}$$

*Proof.* We first note in general for an instrument $\mathcal{A}$ we have

$$J(\mathcal{A}) = \sum_k J(\mathcal{A}_k) \otimes |k\rangle\langle k| \,. \tag{3.40}$$

The orthogonality of the pointer states gives

$$J(\mathcal{A}) - J(\mathcal{B}) = \sum_k J(\mathcal{A}_k) - J(\mathcal{B}_k) \otimes |k\rangle\langle k| \tag{3.41}$$

Since $J(\mathcal{A}_k) - J(\mathcal{B}_k)$ is of course hermitian the first equation in the theorem follows directly from lemma 7.

Recalling the definition of the fidelity using the trace norm

$$\sqrt{\mathcal{F}(J(\mathcal{A}),J(\mathcal{B}))} = \|\sqrt{J(\mathcal{A})}\sqrt{J(\mathcal{B})}\|_1. \tag{3.42}$$

Since

$$\sqrt{J(\mathcal{A})} = \sum_k \sqrt{J(\mathcal{A}_k)} \otimes |k\rangle\langle k| \tag{3.43}$$

the orthogonality of the pointer states gives

$$\sqrt{J(\mathcal{A})}\sqrt{J(\mathcal{B})} = \sum_k \sqrt{J(\mathcal{A}_k)}\sqrt{J(\mathcal{B}_k)} \otimes |k\rangle\langle k| \tag{3.44}$$

and so we may again apply lemma 7 to obtain the second equation as stated in the theorem.

Equations three and four in the theorem would following identically to the above calculations which completes the proof. □

The above theorem is highly important as it allows for simpler calculation of the fidelity and trace distance between instruments by allowing us to compare the pointer maps separately. It also illustrates that in many cases it will often be more convenient to consider the square root fidelity due to its additive nature rather than the fidelity itself.

As previously discussed, we are interested in comparing an ideal measurement to its noisy implementation. In many cases the ideal measurement would correspond to an ideal computational basis measurement whereby the pointer maps $\mathcal{A}_k$ have Kraus rank 1 and are simply rank 1 projectors. In light of this the following lemma will be useful for quickly and easily computing the fidelity $\mathcal{F}_J$.

**Lemma 8.** *For CP maps $\mathcal{A}$ and $\mathcal{B}$ with respective Kraus operators $A$ and $\{B_k\}_{k\in K}$, then we have*

$$\mathcal{F}_J(\mathcal{A},\mathcal{B}) = \frac{1}{d^2}\sum_k |\operatorname{Tr}(A^\dagger B_k)|^2$$

*Proof.* Using eq. (3.20), the Choi states for $\mathcal{A}$ and $\mathcal{B}$ are

$$J(\mathcal{A}) = \frac{1}{d}|A\rangle\!\rangle\langle\!\langle A| \tag{3.45}$$

$$J(\mathcal{B}) = \frac{1}{d}\sum_k |B_k\rangle\!\rangle\langle\!\langle B_k|\,. \tag{3.46}$$

In particular, $J(\mathcal{A})$ has rank 1. Therefore

$$\begin{aligned}
\mathcal{F}_J(\mathcal{A},\mathcal{B}) &= \mathcal{F}(J(\mathcal{A}),J(\mathcal{B})) \\
&= \operatorname{Tr}\left(J(\mathcal{A})J(\mathcal{B})\right) \\
&= \frac{1}{d^2}\operatorname{Tr}\left(|A\rangle\!\rangle\langle\!\langle A|\sum_k |B_k\rangle\!\rangle\langle\!\langle B_k|\right) \\
&= \frac{1}{d^2}\sum_k \langle\!\langle A|B_k\rangle\!\rangle\,\langle\!\langle B_k|A\rangle\!\rangle \\
&= \frac{1}{d^2}\sum_k \operatorname{Tr}(A^\dagger B_k)\operatorname{Tr}(B_k^\dagger A) \\
&= \frac{1}{d^2}\sum_k |\operatorname{Tr}(A^\dagger B_k)|^2,
\end{aligned} \tag{3.47}$$

which completes the proof. □

The intention with the above lemma is that as $\mathcal{A}_k$ has rank 1 for every $k$ the above lemma can be used to compute the fidelity between the pointer maps $\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k)$ regardless of what form the noisy implementation $\mathcal{B}$ has. Theorem 3 can then be applied to compute the fidelity or trace distance as desired.

## 3.5 Errors in Measurements

While the previous section asserts that we may use the figures of merits on quantum channels as figures of merit on quantum instruments, and hence measurements, it still leaves open the physical motivation for doing so. This section will provide a general discussion of the errors associated to measurements and how the fidelity and trace distance can be used to adequately capture these errors.

First it will be useful to recall two common distinguishability measures on classical probability distributions. Let $p : X \to [0,1]$ and $q : X \to [0,1]$ be two probability distributions over the outcome space $X$. Then the Kolmogorov distance $K$ is defined as

$$K(p,q) := \frac{1}{2}\sum_{x\in X} |p(x) - q(x)| \tag{3.48}$$

and the Bhattacharyya coefficient $B$ is defined as

$$B(p,q) := \sum_{x\in X} \sqrt{p(x)q(x)} \tag{3.49}$$

Now, suppose we have a system in the state $\rho$ and we wish to perform a measurement with outcome space $K$. Let $\mathcal{A}$ be quantum instrument representing the ideal measurement, and $\mathcal{B}$ be an instrument representing its noisy implementation. There are two possible errors that might occur with the measurement: an error in the probability of an outcome, or an error in the post-measurement state.

Letting $q = \mathrm{Tr}(\mathcal{A}_k(\rho))$ and $r = \mathrm{Tr}(\mathcal{B}_k(\rho)$. The error in the probability of outcome $k$, denoted $\Delta_p(q,r)$, may be defined simply as the difference between the two probabilities

$$\Delta_p(q,r) := |q - r|. \tag{3.50}$$

Recall that the outputs of the pointer maps $\mathcal{A}_k$ and $\mathcal{B}_k$, which dictate the post-measurement state, are unnormalized. Hence when defining the error in the post-measurement state it will be convenient to allow its domain to include unnormalized quantum states to allow us

to work with the pointer maps directly without worry of normalization. Letting $\sigma = \mathcal{A}_k(\rho)$ and $\theta = \mathcal{A}_k(\rho)$ we define the error between the post-measurement states, which we denote by $\Delta_s(\rho, \sigma)$, as

$$\Delta_s(\sigma, \theta) := \frac{1}{2} \left\| \frac{1}{\operatorname{Tr} \sigma} \sigma - \frac{1}{\operatorname{Tr} \theta} \theta \right\|_1, \tag{3.51}$$

where it can be seen that the normalization of $\theta$ and $\sigma$ has been baked into the definition.

A standard way of modeling the noise in the implementation $\mathcal{B}$ would be to assume that it consists of an ideal measurement with an additional error processes that may occur before and/or after the measurement. Under this model the pointer maps for $\mathcal{B}_k$ have the form of $\mathcal{B}_k = \mathcal{L}\mathcal{A}_k\mathcal{R}$ where $\mathcal{L}$ and $\mathcal{R}$ are quantum channels that represent errors which occur post- or pre-measurement, respectively. For simplicity this form is assuming that $\mathcal{L}$ and $\mathcal{R}$ are the same for every $k$. Since $\mathcal{L}$ is a quantum channel, and hence trace preserving, the probability of obtaining outcome $k$ using the noisy measurement $\mathcal{B}$ is $\operatorname{Tr}(\mathcal{L}\mathcal{A}_k\mathcal{R}(\rho))$ and hence will clearly be independent of $\mathcal{L}$. Consequently, any errors that happen post-measurement do not contribute to errors in probability. In general, the error in the post-measurement state will depend on both $\mathcal{L}$ and $\mathcal{R}$. However, in the commonly assumed case where the pointer maps for the ideal measurement have form $\mathcal{A}_k(\rho) = |\psi_k\rangle\langle\psi_k| \, \rho \, |\psi_k\rangle\langle\psi_k|$ for some normalized vector $|\psi_k\rangle$, then the error in post-measurement state will be independent of both $\mathcal{R}$ and the initial state $\rho$. This can immediately be seen from the fact that the post-measurement state for the outcome $k$ is proportional to $\mathcal{L}\mathcal{A}_k\mathcal{R}(\rho))$ and by assumption the output of $\mathcal{A}_k\mathcal{R}(\rho)$ will always be proportional to the state $|\psi_k\rangle\langle\psi_k|$, regardless of input. The only exception of course being if $\mathcal{R}(\rho)$ outputs a state orthogonal to $|\psi_k\rangle\langle\psi_k|$.

With the brief discussion out of the way we now wish to show that the fidelity and trace distance capture both of the errors in measurement. We begin by first fixing an initial state $\rho$ and we now wish to compare the outputs of the quantum instruments $\mathcal{A}(\rho)$ and $\mathcal{B}(\rho)$. First we shall do so by using the trace distance, however with the additional assumption that the error in the post-measurement state is 0 for every measurement outcome (i.e. $\Delta_s(\mathcal{A}_k(\rho), \mathcal{B}_k(\rho)) = 0$ for every $k$). Consequently, this is equivalent to the assumption that

$$\frac{\mathcal{A}_k(\rho)}{\operatorname{Tr}(\mathcal{A}_k(\rho))} = \frac{\mathcal{B}_k(\rho)}{\operatorname{Tr}(\mathcal{B}_k(\rho))} \tag{3.52}$$

We may now use the additivity of the trade distance and the fact that the trace norm of

a state is 1 to obtain

$$\mathcal{D}(\mathcal{A}(\rho),\mathcal{B}(\rho)) = \frac{1}{2}\sum_k \|\mathcal{A}_k(\rho) - \mathcal{B}_k(\rho)\|_1 \tag{3.53}$$

$$= \frac{1}{2}\sum_k \left\| \mathrm{Tr}(\mathcal{A}_k(\rho))\frac{\mathcal{A}_k(\rho)}{\mathrm{Tr}(\mathcal{A}_k(\rho))} - \mathrm{Tr}(\mathcal{B}_k(\rho))\frac{\mathcal{B}_k(\rho)}{\mathrm{Tr}(\mathcal{B}_k(\rho))} \right\|_1 \tag{3.54}$$

$$= \frac{1}{2}\sum_k |\,\mathrm{Tr}(\mathcal{A}_k(\rho)) - \mathrm{Tr}(\mathcal{B}_k(\rho))| \tag{3.55}$$

which is simply the Kolmogorov distance. If we had instead assumed that the error in the probability for each outcome $k$ is 0 (i.e. $\mathrm{Tr}(\mathcal{A}_k(\rho)) = \mathrm{Tr}(\mathcal{B}_k(\rho))$ for every $k$) then a similar calculation to the above gives

$$\mathcal{D}(\mathcal{A}(\rho),\mathcal{B}(\rho)) = \sum_k \mathrm{Tr}(\mathcal{A}_k(\rho))\Delta_s(\mathcal{A}_k(\rho),\mathcal{B}_k(\rho)) \tag{3.56}$$

which is simply the weighted average error in the post-measurement state of each outcome.

We now compare the outputs of both quantum instruments using the square root fidelity instead of the trace distance. By using the additivity of the trace distance and multiplicaiton by 1 to introduced factors of $\mathrm{Tr}(\mathcal{A}_k(\rho)$ and $\mathcal{B}_k(\rho)$ we obtain

$$\sqrt{\mathcal{F}(\mathcal{A}(\rho),\mathcal{B}(\rho))} = \sum_k \sqrt{\mathrm{Tr}(\mathcal{A}_k(\rho))\,\mathrm{Tr}(\mathcal{B}_k(\rho))}\sqrt{F(\tilde{\mathcal{A}}_k(\rho),\tilde{\mathcal{B}}_k(\rho))} \tag{3.57}$$

where the tilde denotes that the outputs of the pointer maps have been normalized. Following the calculations we did previously, if we assume that the error in the probabilities for each outcome are the same, then the above is simply the weighted average of the fidelities of the post measurement states. If we instead assume the error in the post-measurement states are 1 for every outcome, then the fidelity between the states is 1 and hence the above reduces to the Bhattacharyya coefficient.

## 3.6 A Diamond Distance for Measurements

The previous section was intended to provide some intuition for using the fidelity and trace distance as distinguishability measures on quantum instruments. The issue however is the analysis was conditioned on a particular input state. It would therefore be beneficial to develop an initial state agnostic figure of merit. This section is intended to provide

motivation for considering a maximization over input states and hence develop an analog of the diamond distance specific to quantum instruments. The following analysis is based on the argument in [39] for using the diamond distance as figure of merit on quantum channels.

To begin, consider a circuit consisting of a sequence of linear operations which we denote

$$\prod_{j=m\to 0} \mathfrak{O}_j := \mathfrak{O}_m \mathfrak{O}_{m-1} \dots \mathfrak{O}_0, \tag{3.58}$$

where each $\mathfrak{O}_j$ is a linear map representing either the operation of state preparation, measurement, or gate. Before preceding it will be constructive to discuss how each operation may be expressed as a linear map in this context. The operation of a gate expressed as a linear map seems obvious, while the operation of state preparation can simply be represented by the constant map which sends all inputs to a designated state. Representing the operation of measurement as a linear map is somewhat more tricky. The natural choice given the content of this article would be to represent measurements using quantum instruments. This has the benefit of being able to track measurement outcomes across the circuit through the pointer states. The downside of course being that each measurement operation will therefore expand the system by adding ancilla states. This can be corrected for by e.g. appending to each gate operation a number of identity mappings on the ancilla states equal to the number of measurement operations that precede the gate operation. This can similarly be done to the pointer maps for subsequent measurements. This initially seems cumbersome, but ultimately becomes a minor technical detail that can effectively be ignored as we'll see shortly.

Let $\psi$ represent the ideal output of the circuit and $\rho$ represent the output of the noisy

implementation, then by sub-multiplicativity and the triangle inequality we have

$$\begin{aligned}
\|\psi-\rho\|_1 &= \left\| \prod_{j=m\to 0} \mathfrak{O}_j - \prod_{j=m\to 0} \tilde{\mathfrak{O}}_j \right\|_1 \\
&= \left\| \prod_{j=m\to 0} \mathfrak{O}_j - \mathfrak{O}_m \prod_{j=m-1\to 0} \tilde{\mathfrak{O}}_j + \mathfrak{O}_m \prod_{j=m-1\to 0} \tilde{\mathfrak{O}}_j - \prod_{j=m\to 0} \tilde{\mathfrak{O}}_j \right\|_1 \\
&\leq \left\| \prod_{j=m\to 0} \mathfrak{O}_j - \mathfrak{O}_m \prod_{j=m-1\to 0} \tilde{\mathfrak{O}}_j \right\|_1 + \left\| \mathfrak{O}_m \prod_{j=m-1\to 0} \tilde{\mathfrak{O}}_j - \prod_{j=m\to 0} \tilde{\mathfrak{O}}_j \right\|_1 \\
& \qquad\qquad (\Delta \text{ inequality}) \\
&\leq \left\| \prod_{j=m-1\to 0} \mathfrak{O}_j - \prod_{j=m-1\to 0} \tilde{\mathfrak{O}}_j \right\|_1 + \left\| \mathfrak{O}_m - \tilde{\mathfrak{O}}_m \right\|_\diamond \qquad \text{(contractivity)} \\
&\vdots \\
&\leq \sum_{j=m}^{0} \|\mathfrak{O}_j - \tilde{\mathfrak{O}}_j\|_\diamond, \qquad (3.59)
\end{aligned}$$

where $\tilde{\mathfrak{O}}_j$ represents the noisy implementation of $\mathfrak{O}_j$. From this we can see that the diamond distance not only bounds the error in the circuit, but does so while allowing us to consider the error in each individual operation separately. The expression of the error in the circuit in terms of the diamond distance has the added benefit that all of the appended identity mappings that had to be added can actually be ignored due to the stability of the diamond distance.

The above is intended as physical motivation for having the diamond distance, as defined for general quantum channels, to be the figure of merit of choice when evaluating quantum instruments. There is however a subtle difference in how the diamond distance is used for quantum instruments as opposed to gates that is worth mentioning. When working with gates there is a common notion of the diamond distance from the identity, in which a sequence of noisy gates is compared to the identity channel. Comparing a quantum instrument however to the identity channel has no (obvious) physical meaning. When comparing instruments it will always be comparing a noisy implementation with its ideal counterpart in whatever form they may take in a given context. Appealing to the obvious similarities of the two uses of the diamond distance and the subtle differences it may be appropriate to refer to the diamond distance when used for quantum instruments as the diamond distance from the ideal.

## 3.7 Relations Between Figures of Merit

With several different notions for figures of merit on a quantum instrument it would be prudent to develop relations between them. The most simple of which is a quick application of the Fuchs-van de Graaf inequality applied to the Choi states of quantum instruments. This gives the following lemma

**Corollary 3.** *For any two quantum instruments $\mathcal{A}$ and $\mathcal{B}$, we have*

$$1 - \sqrt{\mathcal{F}_J(\mathcal{A},\mathcal{B})} \leq \frac{1}{2}\left\|J(\mathcal{A}) - J(\mathcal{B})\right\|_1 \leq \sqrt{1 - \mathcal{F}_J(\mathcal{A},\mathcal{B})}$$

As we saw previously in lemma 6, the diamond distance can be bounded by $\mathcal{D}_J$. Now lemma 6 will of course still trivially apply in the case of quantum instruments. However, there is a very important distinction here when bounding the diamond distance from the identity as compared to the diamond distance from the ideal. As seen in the proof of proposition 9 of [41], we may bound the diamond distance from the identity for a unitary channel $\mathcal{G}$ through the use of the improved lower bound of $1 - \mathcal{F}(\phi, J(\mathcal{G}))$, to obtain

$$1 - \mathcal{F}(\phi, J(G)) \leq \|\mathcal{G} - \mathcal{I}\|_\diamond \tag{3.60}$$

as $\phi$, the maximally entangled state used in the Choi-Jamiolkowski isomorphism, is a pure state. The problem however is the Choi state of a quantum instrument will always be a mixed state as long as there is more than one measurement outcome. The goal now is to generalize equation 3.10 to allow for mixed states. The following lemma does so for a class of states with a structure reminiscent of that of quantum instruments and hence may be used to give us the desired new lower bound.

**Lemma 9.** *Let $\rho = \sum_j \rho_j$ and $\sigma = \sum_j \sigma_j$ be decompositions of two density operators satisfying*

$$\rho_j\rho_k = \sigma_j\sigma_k = \rho_j\sigma_k = \sigma_j\rho_k = 0 \tag{3.61}$$

*for all $j \neq k$ and such that $\rho_j$ is rank 1 for all $j$. Then*

$$1 - \sum_j \frac{\mathcal{F}(\rho_j,\sigma_j)}{\sqrt{\mathcal{F}(\rho_j,\rho_j)}} \leq \tfrac{1}{2}\|\rho - \sigma\|_1. \tag{3.62}$$

*Proof.* Under the stated assumptions, $\rho$ and $\sigma$ are block-diagonal in the same basis and so from lemma 7 we have

$$\|\rho - \sigma\|_1 = \sum_j \|\rho_j - \sigma_j\|_1, \tag{3.63}$$

where $\rho_j$ and $\sigma_j$ are positive semi-definite matrices. Now for positive semi-definite matrices $A$ and $B$ we have the following identity for the trace norm

$$\frac{1}{2}\|A-B\|_1 \geq \max_{P\leq I} \operatorname{Tr}\left(P(A-B)\right), \tag{3.64}$$

where the maximization is taken over positive matrices $P \leq I$. As $\hat{\rho}_j = \rho_j / \operatorname{Tr}\rho_j$ is a projector, we have

$$\frac{1}{2}\|\rho_j-\sigma_j\|_1 \geq \operatorname{Tr}\left(\hat{\rho}_j(\rho_j-\sigma_j)\right) = \operatorname{Tr}\rho_j - \frac{\operatorname{Tr}\rho_j\sigma_j}{\operatorname{Tr}\rho_j}. \tag{3.65}$$

Summing over the previous equation while noting that

$$\sum_j \operatorname{Tr}\rho_j = 1 \tag{3.66}$$

and since each $\rho_j$ is rank 1 by assumption we may use lemma 3 to rewrite each term in terms of the fidelity. □

**Theorem 4** (Extended Fuchs-van de Graaf Inequality)**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be quantum instruments with respective pointer maps $\{\mathcal{A}_k\}_k$ and $\{\mathcal{B}_k\}_k$. If, for every $k$, the pointer maps $\mathcal{A}_k$ have Kraus rank 1 then*

$$1-\sum_k \frac{\mathcal{F}_J(\mathcal{A}_k,\mathcal{B}_k)}{\sqrt{\mathcal{F}_J(\mathcal{A}_k,\mathcal{A}_k)}} \leq \frac{1}{2}\sum_k \|J(\mathcal{A}_k)-J(\mathcal{B}_k)\|_1 \tag{3.67}$$

*Proof.* For an instrument $\mathcal{A}$ we have

$$J(\mathcal{A}) = \sum_k J(\mathcal{A}_k)\otimes|k\rangle\langle k|. \tag{3.68}$$

Since $\mathcal{A}_k$ has Kraus rank 1 the Choi states $J(\mathcal{A}_k)$ will be rank 1. Hence using lemma lemma 9 together with the stability of the fidelity obtains the desired result. □

The Choi states for the quantum instruments will by construction satisfy the requirements for theorem 4 as long as we assume that the ideal pointer maps are rank 1 projectors. We will not prove in general that the lower bound given in theorem 4 is an improvement over the lower bound of $1-\sqrt{\mathcal{F}_J(\mathcal{A},\mathcal{B})}$ given by the Fuchs-van de Graaf inequality, but we will show that it is an improvement and is able to saturate the inequality in a few select examples in the following section. We will however perform a quick sanity check that the

lower bound given in theorem 4 will in fact never be negative. Using Holder's inequality for the Schatten norms, as the Schatten 1-norm is colloquially known as the trace norm, we have

$$\begin{aligned}\sum \frac{\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k))}{\sqrt{\mathcal{F}_J(\mathcal{A}_k, \mathcal{A}_k)}} &= \sum \frac{\|\sqrt{J(\mathcal{A}_k)}\sqrt{J(\mathcal{B}_k)}\|_1^2}{\|J(\mathcal{A}_k)\|_1} \\ &\leq \sum \frac{\|\sqrt{J(\mathcal{A}_k)}\|_2^2 \, \|\sqrt{J(\mathcal{B}_k)}\|_2^2}{\operatorname{Tr} J(\mathcal{A}_k)} \\ &= \sum \operatorname{Tr}(J(\mathcal{B}_k)) \\ &= 1. \end{aligned} \tag{3.69}$$

We now generalize an argument of [41] to upper and lower bound the diamond distance from the ideal.

**Corollary 4.** *Let $\mathcal{A}$ and $\mathcal{B}$ be quantum instruments with corresponding pointer maps $\{\mathcal{A}_k\}_{k\in K}$ and $\{\mathcal{B}_k\}_{k\in K}$. If the pointer maps of $\mathcal{A}$ all have Kraus rank 1, we have*

$$1 - \sum_k \frac{\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k)}{\sqrt{\mathcal{F}_J(\mathcal{A}_k, \mathcal{A}_k)}} \leq \frac{1}{2} \|\mathcal{A} - \mathcal{B}\|_\diamond \leq d\sqrt{1 - \mathcal{F}_J(\mathcal{A}, \mathcal{B})}.$$

*Proof.* For convenience, we let $\delta = \mathcal{A} - \mathcal{B}$. To prove the lower bound, we note that

$$\|\delta\|_\diamond = \sup_{\rho \in \mathbb{D}_{d^2}} \|\mathcal{I}_d \otimes \delta(\rho)\|_1 \geq \|\mathcal{I}_d \otimes \delta(\phi)\|_1 = \|J(\mathcal{A}) - J(\mathcal{B})\|_1 \tag{3.70}$$

and use the lower bound of theorem 4. To prove the upper bound, we may write the diamond norm as [42]

$$\|\delta\|_\diamond = d \max\{\|(\sqrt{\rho_0} \otimes I_f) J(\delta) (\sqrt{\rho_1} \otimes I_f)\|_1 : \rho_0, \rho_1 \in \mathbb{D}_d\}, \tag{3.71}$$

where the factor of $d$ comes from the normalization of the Choi state. The inequality $\|ABC\|_1 \leq \|A\|_\infty \|B\|_1 \|C\|_\infty$ [5] together with $\|\sqrt{\rho}\|_\infty \leq 1$ for any quantum state $\rho$ gives

$$\|\delta\|_\diamond \leq d\|J(\delta)\|. \tag{3.72}$$

Using the upper bound from corollary 3 again completes the proof. □

## 3.8 Examples for Commnon Noise Models

With the figures of merit established it would now be constructive to look at a few examples to see how the proposed figures of merit behave under some simple noise models. The complication with the diamond distance however is that since it is a maximization over quantum states it is not always obvious or easy to compute it directly. The following examples therefore will focus on computing $\mathcal{F}_J$ and $\mathcal{D}_J$ and simply being content with using them to bound the diamond distance.

For the following examples when discussing Kraus operators for the various pointer maps we will adopt the notation $A_{i,j}$, which will denote the $j$-th Kraus operator for the pointer map $\mathcal{A}_i$ of the instrument $\mathcal{A}$. The Kraus operators for each error model we obtain from [24]. We will also frequently make use of the following lemma to more easily compute the trace distance.

**Lemma 10.** *Let* $|x\rangle, |y\rangle \in \mathbb{C}^d$. *Then*

$$\| \, |x\rangle\langle x| - |y\rangle\langle y| \, \|_1 = \sqrt{(\langle x|x\rangle + \langle y|y\rangle)^2 - 4| \, \langle y|x\rangle \, |^2}.$$

*Proof.* Let $|0\rangle$ and $|1\rangle$ be an orthonormal basis for the span of the vectors $|x\rangle$ and $|y\rangle$. Then

$$\begin{aligned} |x\rangle\langle x| &= (x_0|0\rangle + x_1|1\rangle)(\bar{x}_0\langle 0| + \bar{x}_1\langle 1|) && (3.73)\\ &= |x_0|^2 \, |0\rangle\langle 0| + x_0\bar{x}_1 \, |0\rangle\langle 1| + \bar{x}_0 x_1 \, |0\rangle\langle 1| + |x_1|^2 \, |1\rangle\langle 1| \, . && (3.74) \end{aligned}$$

and similarly for $y$. Therefore we have

$$|x\rangle\langle x| - |y\rangle\langle y| \equiv \begin{pmatrix} |x_0|^2 - |y_0|^2 & x_0\bar{x}_1 - y_0\bar{y}_1 \\ \bar{x}_0 x_1 - \bar{y}_0 y_1 & |x_1|^2 - |y_1|^2 \end{pmatrix}. \tag{3.75}$$

Without loss of generality, we choose $|1\rangle \propto v$ so that

$$|x\rangle\langle x| - |y\rangle\langle y| \equiv \begin{pmatrix} |x_0|^2 & x_0\bar{x}_1 \\ \bar{x}_0 x_1 & |x_1|^2 - \|y\|^2 \end{pmatrix}. \tag{3.76}$$

For a general $2 \times 2$ matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3.77}$$

the characteristic polynomial is

$$(a-\lambda)(d-\lambda)-bc=\lambda^2-\lambda(a+d)-bc+ad=0 \tag{3.78}$$

By the quadratic formula we have

$$\lambda=\frac{1}{2}\left((a+d)\pm\sqrt{(a+d)^2-4(ad-bc)}\right). \tag{3.79}$$

Noting that $|x_0|^2+|x_1|^2=\|x\|$ and $x_1=\langle y|x\rangle/\|x\|$, in our particular case we get the eigenvalues to be

$$\begin{aligned}
2\lambda &= \|x\|^2-\|y\|^2\pm\sqrt{(\|x\|^2-\|y\|^2)^2-4(|x_0|^2|x_1|^2-|x_0|^2\|y\|^2-|u_0|^2|u_1|^2)} && (3.80)\\
&= \|x\|^2-\|y\|^2\pm\sqrt{(\|x\|^4-2\|y\|^2(|x_0|^2+|x_1|^2)+\|y\|^4)+4|x_0|^2\|y\|^2} && (3.81)\\
&= \|x\|^2-\|y\|^2\pm\sqrt{\|x\|^4+2\|y\|^2|x_0|^2-2\|y\|^2|x_1|^2+\|y\|^4} && (3.82)\\
&= \|x\|^2-\|y\|^2\pm\sqrt{(\|x\|^4+2\|y\|^2|x_0|^2+2\|y\|^2|x_1|^2+\|y\|^4)-4\|y\|^2|x_1|^2} && (3.83)\\
&= \|x\|^2-\|y\|^2\pm\sqrt{(\|x\|^2+\|y\|^2)^2-4|\langle y|x\rangle|^2} && (3.84)
\end{aligned}$$

If we assume that $\|y\|<\|x\|$ then from the above we see that

$$0\leq\|x\|^2-\|y\|^2\leq\sqrt{(\|x\|^2+\|y\|^2)^2-4|\langle y|x\rangle|^2} \tag{3.85}$$

It follows that

$$\sum_i s_i(|x\rangle\langle x|-|y\rangle\langle y|)=\sqrt{(\|x\|^2+\|y\|^2)^2-4|\langle y|x\rangle|^2} \tag{3.86}$$

Note that the above analysis was contingent on the span of $|x\rangle$ and $|y\rangle$ forming a two dimensional space. This would not be true if $|y\rangle=\alpha|x\rangle$ for some constant $\alpha\in\mathbb{C}$. However in that case it is straightforward to compute the trace norm

$$\|\,|x\rangle\langle x|-|y\rangle\langle y|\,\|_1=\|\,|x\rangle\langle x|-|\alpha|^2\,|x\rangle\langle x|\,\|_1=|1-|\alpha|^2|\,\langle x|x\rangle \tag{3.87}$$

which is actually the same result we would get if we had instead used the formula stated in the lemma. □

**Bit Flip**

Lets consider the case of where the instrument $\mathcal{A}$ corresponds to an ideal computational basis measurement for a single qubit system. It will therefore have two pointer maps $\mathcal{A}_0$ and $\mathcal{A}_1$, which each, respectively, have a singular Kraus operator

$$A_{0,0} = |0\rangle\langle 0| \tag{3.88}$$
$$A_{1,0} = |1\rangle\langle 1|\,. \tag{3.89}$$

Now, consider an instrument $\mathcal{B}$ representing a non-ideal measurement consisting of a bit flip error with probability $p$ followed by an ideal computational basis measurement. The Kraus operators for a bit flip error are

$$\{\sqrt{p}(|0\rangle\langle 0| + |1\rangle\langle 1|\,, \quad \sqrt{1-p}(|0\rangle\langle 1| + |1\rangle\langle 0|)\}. \tag{3.90}$$

The instrument $\mathcal{B}$ will have two pointer maps $B_0$ and $B_1$ each with two Kraus operators, which we list below

$$B_{0,0} = \sqrt{1-p}\,|0\rangle\langle 0|\,, \quad B_{0,1} = \sqrt{p}\,|0\rangle\langle 1| \tag{3.91}$$
$$B_{1,0} = \sqrt{1-p}\,|1\rangle\langle 1|\,, \quad B_{1,1} = \sqrt{p}\,|1\rangle\langle 0| \tag{3.92}$$

Since $\mathcal{A}_0$ has a single Kraus operator we may use lemma 8 to compute the fidelity between $\mathcal{A}_0$ and $\mathcal{B}_0$. We find

$$\sqrt{\mathcal{F}(\mathcal{A}_j, \mathcal{B}_j)} = \frac{1}{2}\sqrt{\sum_{k\in\mathbb{Z}_4} |\operatorname{Tr}(A_j B_{j,k})|^2} \tag{3.93}$$
$$= \frac{1}{2}\sqrt{1-p}. \tag{3.94}$$

Therefore

$$\mathcal{F}(\mathcal{A}, \mathcal{B}) = 1 - p. \tag{3.95}$$

and

$$\sum \frac{\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k))}{\sqrt{\mathcal{F}_J(\mathcal{A}_k, \mathcal{A}_k)}} = 1 - p. \tag{3.96}$$

Hence the extended Fuchs-van de Graaf inequality would give a lower bound of $p$, which is tighter compared to the lower bound as given by the Fuchs-van de Graaf inequality of $1 - \sqrt{1-p}$.

The trace norm between $J(\mathcal{A}_0)$ and $J(\mathcal{B}_0)$ can then be computed using lemma 10 as follows:

$$\|J(\mathcal{A}_0) - J(\mathcal{B}_0)\|_1 = \frac{1}{2}\| \, |A_0\rangle\!\rangle\langle\!\langle A_0| - |B_{0,0}\rangle\!\rangle\langle\!\langle B_{0,0}| - |B_{0,1}\rangle\!\rangle\langle\!\langle B_{0,1}| \, \|_1 \tag{3.97}$$

$$= \frac{1}{2}\| \, |00\rangle\langle 00| - (1-p) \, |00\rangle\langle 00| - p \, |10\rangle\langle 10| \, \|_1 \tag{3.98}$$

$$= \frac{p}{2}\| \, |00\rangle\langle 00| - |10\rangle\langle 10| \, \|_1 \tag{3.99}$$

$$= p. \tag{3.100}$$

A similar calculation will show that

$$\|J(\mathcal{A}_1) - J(\mathcal{B}_1)\|_1 = p. \tag{3.101}$$

Hence by additivity

$$\mathcal{D}(\mathcal{A}, \mathcal{B}) = p. \tag{3.102}$$

Giving saturation in the lower bound of the extended Fuchs-van de Graaf inequality.

### Phase Flip

Again considering $\mathcal{A}$ to be an ideal computational basis measurement of a qubit system. Let the instrument $\mathcal{B}$ represent a non-ideal measurement consisting of a phase flip error with probability $p$ followed by an ideal computational basis measurement. The Kraus operators for phase flip channel are given below

$$\{\sqrt{1-p} \, |0\rangle\langle 0| + \sqrt{1-p} \, |1\rangle\langle 1| \, , \quad \sqrt{p} \, |0\rangle\langle 0| - \sqrt{p} \, |1\rangle\langle 1|\}. \tag{3.103}$$

The two pointer maps $B_0$ and $B_1$ for the instrument $\mathcal{B}$ will each have two Kraus operators, as given

$$B_{0,0} = \sqrt{1-p} \, |0\rangle\langle 0| \, , \quad B_{0,1} = \sqrt{p} \, |0\rangle\langle 0| \tag{3.104}$$

$$B_{1,0} = \sqrt{1-p} \, |1\rangle\langle 1| \, , \quad B_{1,1} = -\sqrt{p} \, |1\rangle\langle 1| \, . \tag{3.105}$$

Since $\mathcal{A}_0$ has a single Kraus operator we may use lemma 8 to compute the fidelity between $\mathcal{A}_0$ and $\mathcal{B}_0$. We find

$$\sqrt{\mathcal{F}(\mathcal{A}_j, \mathcal{B}_j)} = \frac{1}{2}\sqrt{\sum_{k \in \mathbb{Z}_4} |\operatorname{Tr}(A_{0,0}^\dagger B_{0,k})|^2} = \frac{1}{2}. \tag{3.106}$$

Therefore

$$\mathcal{F}(\mathcal{A}, \mathcal{B}) = 1. \tag{3.107}$$

The trace norm between $J(\mathcal{A}_0)$ and $J(\mathcal{B}_0)$ can then be computed using lemma 10 as follows:

$$\begin{aligned} \|J(\mathcal{A}_0) - J(\mathcal{B}_0)\|_1 &= \frac{1}{2}\| \, |A_0\rangle\!\rangle\langle\!\langle A_0| - |B_{0,0}\rangle\!\rangle\langle\!\langle B_{0,0}| - |B_{0,1}\rangle\!\rangle\langle\!\langle B_{0,1}| \, \|_1 && (3.108) \\ &= \frac{1}{2}\| \, |00\rangle\langle 00| - (1-p) \, |00\rangle\langle 00| - p \, |00\rangle\langle 00| \, \|_1 && (3.109) \\ &= 0. && (3.110) \end{aligned}$$

A similar calculation will show that

$$\|J(\mathcal{A}_1) - J(\mathcal{B}_1)\|_1 = 0. \tag{3.111}$$

Hence by additivity

$$\mathcal{D}(\mathcal{A}, \mathcal{B}) = 0. \tag{3.112}$$

The fidelity being 1 and trace distance 0 are reflective of the fact that phase flip will induce no error in probabilities or error in the post-measurement state due to the projective nature of $\mathcal{A}$.

### Amplitude Damping

We will now consider amplitude damping, which corresponds to the decay of the qubit system from its excited state $|1\rangle\langle 1|$ to its ground state $|0\rangle\langle 0|$. It has the following two Kraus operators

$$\{|0\rangle\langle 0| + \sqrt{1-p} \, |1\rangle\langle 1| \, , \quad \sqrt{p} \, |0\rangle\langle 1|\}, \tag{3.113}$$

where the parameter $p$ corresponds to the probability of state decay.

Now, similarly to other examples we will let $\mathcal{A}$ correspond the an ideal computational basis measurement for a qubit system. The instrument $\mathcal{B}$ will correspond to a noisy measurement consisting of amplitude damping followed by an ideal computational basis measurement. A quick calculation for the Kraus operators of the two pointer maps $\mathcal{B}_0$ and $\mathcal{B}_1$ using equation 3.113 gives

$$B_{0,0} = |0\rangle\langle 0|\,, \quad B_{0,1} = \sqrt{p}\,|0\rangle\langle 1| \tag{3.114}$$

$$B_{1,0} = \sqrt{1-p}\,|1\rangle\langle 1|\,. \tag{3.115}$$

Since the pointer maps for $\mathcal{A}$ have a single Kraus operator we may first use lemma 8 to obtain

$$\sqrt{\mathcal{F}(\mathcal{A}_0, \mathcal{B}_0)} = \frac{1}{2}\sqrt{|\operatorname{Tr} A_{0,0}^\dagger B_{0,0}|^2 + |\operatorname{Tr} A_{0,0}^\dagger B_{0,1}|^2} = \frac{1}{2} \tag{3.116}$$

and

$$\sqrt{\mathcal{F}(\mathcal{A}_1, \mathcal{B}_1)} = \frac{1}{2}\sqrt{|\operatorname{Tr} A_{0,0}^\dagger B_{0,0}|^2 + |\operatorname{Tr} A_{0,0}^\dagger B_{0,1}|^2} = \frac{1}{2}\sqrt{1-p} \tag{3.117}$$

Therefore using additivity we get

$$\sqrt{\mathcal{F}(\mathcal{A}, \mathcal{B})} = \frac{1}{2}(1 + \sqrt{1-p}) \tag{3.118}$$

and also

$$\sum \frac{\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k))}{\sqrt{\mathcal{F}_J(\mathcal{A}_k, \mathcal{A}_k)}} = 1 - \frac{p}{2}. \tag{3.119}$$

In this instance we may compute the trace distance between the pointer maps directly to find

$$\|J(\mathcal{A}_0) - J(\mathcal{B}_0)\|_1 = \frac{1}{2}\| \,|A_{0,0}\rangle\!\rangle\langle\!\langle A_{0,0}| - |B_{0,0}\rangle\!\rangle\langle\!\langle B_{0,0}| - |B_{0,1}\rangle\!\rangle\langle\!\langle B_{0,1}| \,\|_1 \tag{3.120}$$

$$= \frac{1}{2}\| \,|00\rangle\langle 00| - |00\rangle\langle 00| - p\,|10\rangle\langle 10| \,\|_1 \tag{3.121}$$

$$= \frac{p}{2} \tag{3.122}$$

and

$$\|J(\mathcal{A}_1) - J(\mathcal{B}_1)\|_1 = \frac{1}{2}\| \,|A_{1,0}\rangle\!\rangle\langle\!\langle A_{1,0}| - |B_{1,0}\rangle\!\rangle\langle\!\langle B_{1,0}| - |B_{0,1}\rangle\!\rangle\langle\!\langle B_{0,1}| \,\|_1 \tag{3.123}$$

$$= \frac{1}{2}\| \,|11\rangle\langle 11| - (1-p)\,|11\rangle\langle 11| \,\|_1 \tag{3.124}$$

$$= \frac{p}{2}. \tag{3.125}$$

Therefore using additivity we find the trace distance between $\mathcal{A}$ and $\mathcal{B}$ to be

$$\mathcal{D}(\mathcal{A},\mathcal{B}) = \frac{p}{2}, \tag{3.126}$$

which again will give saturation in the lower bound of the extended Fuchs-van de Graaf inequality.

### Depolarization

As in the previous examples we let $\mathcal{A}$ correspond to a computational basis measurement on a qubit system. We will consider the polarization channel, which corresponds to a decay of the state of the qubit into a maximally mixed state. The polarizing channel has the following Kraus operators

$$\{\sqrt{1-\frac{3p}{4}}I, \quad \sqrt{\frac{p}{4}}X, \quad \sqrt{\frac{p}{4}}Y, \quad \sqrt{\frac{p}{4}}Z\}. \tag{3.127}$$

Therefore if we let $\mathcal{B}$ a noisy implementation of $\mathcal{A}$ consisting of the depolarizing channel followed by a computational basis measurement we attain the following Kraus operators for the pointer maps of $\mathcal{B}$

$$B_{0,0} = \sqrt{1-\frac{3p}{4}}\,|0\rangle\langle 0|\,,\ B_{0,1} = \sqrt{\frac{p}{4}}\,|0\rangle\langle 1|\,,\ B_{0,2} = -i\sqrt{\frac{p}{4}}\,|0\rangle\langle 1|\,,\ B_{0,3} = \sqrt{\frac{p}{4}}\,|0\rangle\langle 0| \tag{3.128}$$

$$B_{1,0} = \sqrt{1-\frac{3p}{4}}\,|1\rangle\langle 1|\,,\ B_{1,1} = \sqrt{\frac{p}{4}}\,|1\rangle\langle 0|\,,\ B_{1,2} = i\sqrt{\frac{p}{4}}\,|1\rangle\langle 0|\,,\ B_{1,3} = -\sqrt{\frac{p}{4}}\,|1\rangle\langle 1|\,. \tag{3.129}$$

Continuing as we have before, we first start by computing the fidelity of the pointer maps to get

$$\sqrt{\mathcal{F}(\mathcal{A}_0,\mathcal{B}_0)} = \frac{1}{2}\sqrt{\sum_{k\in\mathbb{Z}_4} |\operatorname{Tr}(A_{0,0}^\dagger B_{0,k})|^2} \tag{3.130}$$

$$= \frac{1}{2}\sqrt{(1-\frac{3p}{4})+\frac{p}{4}} \tag{3.131}$$

$$= \frac{1}{2}\sqrt{1-\frac{p}{2}} \tag{3.132}$$

and similarly

$$\sqrt{\mathcal{F}(\mathcal{A}_1, \mathcal{B}_1)} = \frac{1}{2}\sqrt{1 - \frac{p}{2}}. \tag{3.133}$$

Hence by additivity

$$\sqrt{\mathcal{F}(\mathcal{A}, \mathcal{B})} = \sqrt{1 - \frac{p}{2}}, \tag{3.134}$$

and additionally

$$\sum \frac{\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k))}{\sqrt{\mathcal{F}_J(\mathcal{A}_k, \mathcal{A}_k)}} = 1 - \frac{p}{2}. \tag{3.135}$$

Continuing with the trace distance we get

$$\|J(\mathcal{A}_0) - J(\mathcal{B}_0)\|_1 = \frac{1}{2}\| \, |A_{0,0}\rangle\!\rangle\langle\!\langle A_{0,0}| - \sum_k |B_{0,k}\rangle\!\rangle\langle\!\langle B_{0,k}| \, \|_1 \| \tag{3.136}$$

$$= \frac{1}{2}\| \, |00\rangle\langle 00| - (1 - \frac{3p}{4}) \, |0\rangle\langle 0| - \frac{p}{2} \, |10\rangle\langle 10| - \frac{p}{4} \, |00\rangle\langle 00| \, \|_1 \tag{3.137}$$

$$= \frac{p}{4}\| \, |0\rangle\langle 0| - |10\rangle\langle 10| \, \|_1 \tag{3.138}$$

$$= \frac{p}{2}, \tag{3.139}$$

and similarly

$$\|J(\mathcal{A}_1) - J(\mathcal{B}_1)\|_1 = \frac{p}{2}. \tag{3.140}$$

Therefore

$$\mathcal{D}(\mathcal{A}, \mathcal{B}) = \frac{p}{2} \tag{3.141}$$

and we obtain another example of saturation in the lower bound of the extended Fuchs-van de Graaf inequality.

### Unitary Channel

All of the previous examples had saturation in the lower bound of the extended Fuchs-van de Graaf inequality. We now construct an example that gives a saturated upper

bound for the Fuchs-van de Graaf inequality. Again, let $\mathcal{A}$ be an ideal computational basis measurement on a qubit system. For any unitary map $\mathcal{U}$, consider the non-ideal instrument $\mathcal{B}$ which has pointer maps $\mathcal{B}_k = \mathcal{U}(\mathcal{A}_k)$. The Kraus operators in this case will be

$$B_0 = U\left|0\right\rangle\!\left\langle 0\right|, \quad B_1 = U\left|1\right\rangle\!\left\langle 1\right|. \tag{3.142}$$

Letting $k \in \{0, 1\}$ and using lemma 10 we find that

$$\begin{aligned}
\|J(\mathcal{A}_k) - J(\mathcal{B}_k)\|_1 &= \frac{1}{2}\| \left|A_k\right\rangle\!\!\left\rangle\!\left\langle\!\left\langle A_k\right| - \left|UA_k\right\rangle\!\!\left\rangle\!\left\langle\!\left\langle UA_k\right| \|_1 && (3.143)\\
&= \frac{1}{2}\sqrt{(\langle\!\langle A_k|A_k\rangle\!\rangle + \langle\!\langle UA_k|UA_k\rangle\!\rangle)^2 - 4|\,\langle\!\langle UA_k|A_k\rangle\!\rangle\,|^2} && (3.144)\\
&= \frac{1}{2}\sqrt{4\,\mathrm{Tr}(A_k^\dagger A_k) - 4|\,\mathrm{Tr}(U^\dagger A_k^\dagger A_k)|^2} && (3.145)\\
&= \sqrt{1 - |\langle k|U|k\rangle|^2}. && (3.146)
\end{aligned}$$

Therefore if we choose $U = \cos\theta I - i\sin\theta X$, then

$$\frac{1}{2}\sum_k \|J(\mathcal{A}_k) - J(\mathcal{B}_k)\|_1 = |\sin\theta|. \tag{3.147}$$

Computing the fidelity we obtain

$$\begin{aligned}
\mathcal{F}_J(\mathcal{A}, \mathcal{B}) = \frac{1}{4}\left(|\,\mathrm{Tr}(A_0^\dagger U A_0)| + |\,\mathrm{Tr}(A_1^\dagger U A_1)|\right)^2 = \cos^2\theta && (3.148)\\
&&(3.149)
\end{aligned}$$

and hence the upper bound of the Fuchs-van de Graaf inequality is.

$$\sqrt{1 - \mathcal{F}_J(\mathcal{A}, \mathcal{B})} = |\sin\theta|. \tag{3.150}$$

Additionally we have

$$\sum \frac{\mathcal{F}_J(\mathcal{A}_k, \mathcal{B}_k))}{\sqrt{\mathcal{F}_J(\mathcal{A}_k, \mathcal{A}_k)}} = \cos^2\theta. \tag{3.151}$$

Using the earlier expression for the trace distance we have

$$\|J(\mathcal{A}_k) - J(\mathcal{B}_k)\|_1 = |\sin\theta|. \tag{3.152}$$

And so in this example we see saturation in the upper bound of the Fuchs-van de Graaf inequality.

## 3.9 Discussion

In this chapter we explored how the figures of merit on unitary channels may be adapted to figures of merit on quantum instruments. In particular, how figures of merit based on comparing the outputs of quantum instruments capture errors in measurements in intuitive ways. This leads to the diamond distance being seen as a favorable figure of merit for measurements due to its nice properties and physical motivation. However, it's worth emphasizing here that this work is merely meant to introduce these figures of merit. It is still necessary to develop concrete applications, and ultimately develop methods for experimentally estimating them. As we previously mentioned, the diamond distance for unitary channels is typically determined by doing a full process tomography on the system allowing one to directly compute the diamond distance. This is of course inefficient, but there are no known efficient methods for computing the diamond distance directly. As the figures of merit for quantum instruments are based so strongly on the figures of merit for unitary channels we suspect that being able to directly and efficiently determine the diamond distance for a physical measurement to be unlikely. Hence the emphasis in section 3.7 on relating the diamond distance to other figures of merit which may be used to estimate the diamond distance, rather than focus on developing methods of computing it directly.

# Chapter 4

# Conclusion

To summarize, we saw in chapter 2 the errors that may arise in a Ytterbium 171 ion traps and some approaches to characterize them. The obvious limitation of these approaches is that they are rooted in the structure of trapped ions and hence don't necessarily apply to general quantum computing systems. This leads to Chapter 3, which looks at the more abstract setting of quantum instruments and figures of merit that may be used to evaluate them. The hope would be that these figures of merits could provide more robust ways of characterizing the sub system measurements necessary in a trapped ion system. Moreover, as they figures of merit are system agnostic they would facilitate the comparison of measurement procedures across multiple types of implementations of a quantum computer.

# References


[1] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing*, 38(4):1207–1282, 2008.

[2] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. page 20–30, 1998.

[3] D. J. Berkeland and M. G. Boshier. Destabilization of dark states and optical spectroscopy in zeeman-degenerate atomic systems. *Phys. Rev. A*, 65:033413, Feb 2002.

[4] I. I. Beterov and M. Saffman. Rydberg blockade, förster resonances, and quantum state measurements with different atomic species. *Phys. Rev. A*, 92:042710, Oct 2015.

[5] Robin Blume-Kohout, John King Gamble, Erik Nielsen, Jonathan Mizrahi, Jonathan D. Sterk, and Peter Maunz. Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit. 2013.

[6] C. D. Bruzewicz, R. McConnell, J. Stuart, J. M. Sage, and J. Chiaverini. Dual-species, multi-qubit logic primitives for ca+/sr+ trapped-ion crystals. *npj Quantum Information*, 5(1):102, Nov 2019.

[7] Dmitry Budker, Derek Kimball, Derek F Kimball, and David P DeMille. *Atomic physics: an exploration through problems and solutions*. Oxford University Press, USA, 2004.

[8] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[9] Isaac L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

[10] Stephen Crain, Clinton Cahall, Geert Vrijsen, Emma E. Wollman, Matthew D. Shaw, Varun B. Verma, Sae Woo Nam, and Jungsang Kim. High-speed low-crosstalk detection of a 171yb+ qubit using superconducting nanowire single photon detectors. *Communications Physics*, 2(1):97, Aug 2019.

[11] Marcus P. da Silva, Olivier Landon-Cardinal, and David Poulin. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.*, 107:210404, Nov 2011.

[12] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Communications in Mathematical Physics*, 17(3):239–260, Sep 1970.

[13] Omar Fawzi, Antoine Grospellier, and Anthony Leverrier. Efficient decoding of random errors for quantum expander codes. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, page 521–534, New York, NY, USA, 2018. Association for Computing Machinery.

[14] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, may 1999.

[15] Alexei Gilchrist, Nathan Langford, and Michael A. Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, jun 2005.

[16] Daniel Greenbaum. Introduction to quantum gate set tomography. 2015.

[17] J. Helsen, I. Roth, E. Onorati, A.H. Werner, and J. Eisert. General framework for randomized benchmarking. *PRX Quantum*, 3:020357, Jun 2022.

[18] A. Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3(4):275–278, 1972.

[19] Kaveh Khodjasteh, Daniel A. Lidar, and Lorenza Viola. Arbitrarily accurate dynamical control in open quantum systems. *Phys. Rev. Lett.*, 104:090501, Mar 2010.

[20] A Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, dec 1997.

[21] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation. *Science*, 279(5349):342–345, 1998.

[22] Kotibhaskar, Nikhil. Design and construction of an ion trapping apparatus for quantum simulation experiments. Master's thesis, 2019.

[23] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Physics Letters A*, 303(4):249–252, 2002.

[24] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[25] Rachel Noek, Geert Vrijsen, Daniel Gaultney, Emily Mount, Taehyun Kim, Peter Maunz, and Jungsang Kim. High speed, high fidelity detection of an atomic hyperfine qubit. *Opt. Lett.*, 38(22):4735–4738, Nov 2013.

[26] S. Olmschenk, D. N. Matsukevich, P. Maunz, D. Hayes, L.-M. Duan, and C. Monroe. Quantum teleportation between distant matter qubits. *Science*, 323(5913):486–489, 2009.

[27] Masanao Ozawa. Quantum measuring processes of continuous observables. *Journal of Mathematical Physics*, 25(1):79–87, 1984.

[28] G. A. Paz-Silva, G. K. Brennen, and J. Twamley. Fault tolerant quantum information processing with holographic control. 2010.

[29] Gerardo A. Paz-Silva, Gavin K. Brennen, and Jason Twamley. Fault tolerance with noisy and slow measurements and preparation. *Phys. Rev. Lett.*, 105:100501, Aug 2010.

[30] J. M. Pino, J. M. Dreiling, C. Figgatt, J. P. Gaebler, S. A. Moses, M. S. Allman, C. H. Baldwin, M. Foss-Feig, D. Hayes, K. Mayer, C. Ryan-Anderson, and B. Neyenhuis. Demonstration of the trapped-ion quantum ccd computer architecture. *Nature*, 592(7853):209–213, Apr 2021.

[31] John Preskill. Fault-tolerant quantum computation. 1997.

[32] Kenneth Rudinger, Guilhem J. Ribeill, Luke C.G. Govia, Matthew Ware, Erik Nielsen, Kevin Young, Thomas A. Ohki, Robin Blume-Kohout, and Timothy Proctor. Characterizing midcircuit measurements on a superconducting qubit using gate set tomography. *Phys. Rev. Applied*, 17:014014, Jan 2022.

[33] Yuval R Sanders, Joel J Wallman, and Barry C Sanders. Bounding quantum gate error rate based on reported average fidelity. *New Journal of Physics*, 18(1):012002, dec 2015.

[34] Peter W. Shor. Fault-tolerant quantum computation. 1996.

[35] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Review*, 41(2):303–332, 1999.

[36] Roman Stricker, Davide Vodola, Alexander Erhard, Lukas Postler, Michael Meth, Martin Ringbauer, Philipp Schindler, Rainer Blatt, Markus Müller, and Thomas Monz. Characterizing quantum instruments: From nondemolition measurements to quantum error correction. *PRX Quantum*, 3:030318, Aug 2022.

[37] Roman Stricker, Davide Vodola, Alexander Erhard, Lukas Postler, Michael Meth, Martin Ringbauer, Philipp Schindler, Thomas Monz, Markus Müller, and Rainer Blatt. Experimental deterministic correction of qubit loss. *Nature*, 585(7824):207–210, Sep 2020.

[38] Sebastian Wagner, Jean-Daniel Bancal, Nicolas Sangouard, and Pavel Sekatski. Device-independent characterization of quantum instruments. *Quantum*, 4:243, March 2020.

[39] Joel J. Wallman. Bounding experimental quantum error rates relative to fault-tolerant thresholds. 2015.

[40] Joel J. Wallman and Joseph Emerson. Noise tailoring for scalable quantum computation via randomized compiling. *Phys. Rev. A*, 94:052325, Nov 2016.

[41] Joel J Wallman and Steven T Flammia. Randomized benchmarking with confidence. *New Journal of Physics*, 16(10):103032, oct 2014.

[42] John Watrous. Simpler semidefinite programs for completely bounded norms. *arXiv preprint arXiv:1207.5726*, 2012.

[43] Kenneth Earl Wright II. *Manipulation of the quantum motion of trapped atomic ions via stimulated Raman transitions*. PhD thesis, University of Maryland, College Park, 2017.