# Measurement-Based non-Gaussian Quantum State Engineering

Miller Thomas Eaton
Colfax, IL, USA

B.S. Southern Illinois University, 2013

*A Dissertation presented to the Graduate Faculty of the University of Virginia in Candidacy for the degree of Doctor of Philosophy*

Department of Physics

UNIVERSITY OF VIRGINIA

May, 2022

# *Abstract*

Advances in quantum technologies promise to deliver a great deal of progress in information science, including applications in sensing, secure communication, and quantum computation. If successfully implemented, an error-corrected universal quantum computer would deliver an exponential boost in efficiency when solving certain problems. However, this requires a way to generate and entangle many low-noise quantum resources in a scalable way. Fortunately, quantum computing with the continuous-variable nature of light in a quantum-optics setting has recently emerged as a potential contender due to the experimental availability of massively scalable entangled states.

After connecting continuous-variable and qubit-based quantum computing, this dissertation introduces several novel proposals to engineer specific non-Gaussian quantum resources using projective measurement. These resources include Schrödinger cat, Gottesman-Kitaev-Preskill, binomial code, and coherent photon-subtracted states that would be used to achieve quantum advantages in quantum sensing and quantum computing. Each protocol is designed with current technologies in mind, and a full computational model with realistic imperfections is included to assess the experimental feasibility.

The second part of this dissertation transitions to experimental implementations with a specific focus on the transition-edge sensor, which is a highly quantum efficient photon-number-resolving detector. I detail new methods of processing the output signal and demonstrate the ability to improve resolving capabilities to up to 30 photons per detection channel. This has important implications for quantum information applications, including quantum random number generation and realization of the state engineering protocols discussed in the first portion of the dissertation. Additionally, I present variations on recently developed methods to perform full quantum state tomography that would lead to a quadratic reduction in the number of required measurement settings to completely characterize an unknown state.

Finally, I discuss experimental progress on the generation and characterization of exotic states using the protocol of photon catalysis. This process utilizes a single-photon resource and projection with photon-number-resolving detection to engineer quantum states with non-Gaussian Wigner functions. Current results indicate that this is a viable method to generate exotic states if stability and control of experimental conditions can be improved.

# *Acknowledgements*

Attaining a graduate degree is a multi-faceted challenge that one cannot hope to complete alone. There have been a myriad of people from mentors and teachers to friends and family that have helped me get to this point — too many to name individually, but I am truly grateful to them all. Although my few paltry words cannot convey my real gratitude, I will at least make a short attempt for those who have helped me with this work.

First, I have Olivier Pfister to thank for being a truly fantastic advisor and mentor, and a genuinely good human being. His nearly insufferable optimism and energy has often been just the bit of encouragement I needed to make a final attempt to solve a problem. I truly couldn't imagine having a better mentor under which to complete a PhD.

Although I have very much enjoyed working with all of the people I have collaborated with over the last few years, I would specifically like to thank my lab mates Rajveer Nehra, Chun-Hung Chang, Amr Houssameldin, Xuan Zhu, and Carlos González-Arciniegas. It has been wonderful to work with all of these people and I am grateful for the unique insight and expertise each bring. Working on an experiment by oneself can be challenging, and occasionally demoralizing, so having Raj and Amr to commiserate with made things both considerably easier and much more fun. I am certainly glad for the continued camaraderie.

Beyond the work environment, I am grateful for the continued support of family and friends, both old and new, over the last several years. Most importantly, I could not have made it this far without the loving support of my wife, Madelyn. Putting up with me while being a graduate student herself can't have been easy, and I will always be so glad that we made it through these challenges together.

# *Preface*

The work in this dissertation is part of a series of attempts to bridge the gap between what is available experimentally and what is theoretically necessary to attain a quantum advantage in quantum information. As such, it brings together several aspects of theoretical quantum optics, quantum information theory both with discrete qubits and continuous-variable qumodes (the analog of qubits), and experimental techniques involving nonlinear optics, detection, electronics, and signal processing. This document assumes familiarity with quantum mechanics, electromagnetism, and principles of optics, but the majority of necessary concepts that are related to quantum information and quantum optics are presented for those unfamiliar.

Chapter 1 (Quantum Optics and Quantum Computation) begins with a brief introduction to the quantization of the electromagnetic field and presents formalism necessary for treating quantum optical fields, including common state representations and operators. The beginning attempts to foster a more intuitive understanding of quantum phase space through the use of Wigner functions and how they transform under common operations. Concepts including non-unitary evolution, such as the decoherence of a quantum state as it interacts with the environment, are also discussed. Once the basic elements of quantum optics are in place, this chapter transitions to discussion on quantum computation. Starting from qubit formalism, measurement-based quantum computing with cluster states is introduced, which is a natural segue to continuous variables. From here, I introduce certain types of states having Wigner function negativity that would be highly desirable for universal quantum computation and error correction.

After the background is in place, chapters 2 (Photon Catalysis), 3 (Embedding Non-Gaussianity in a Cluster State), and 4 (Rotation-Symmetric Code States) focus on different proposed methods to generate some of the useful quantum states mentioned in Ch. 1. Each of these methods makes use of the ability to perform projections in the photon-number basis on a portion of an entangled state through the use of photon-number-resolving detectors. The 'quantumness' of these projective measurements allows for imprinting exotic, non-classical features on the unmeasured portion of the state. Chapter 5 (Quantum Sensing with Non-Gaussian States) then describes a variation to the method of Ch. 4, which would allow for the preparation of quantum states useful for improving interferometric sensitivity.

Chapter 6 (Photon-Number-Resolving Detection) switches to an experimental point of view by focusing on how to perform the projective measurements so casually used in the previous chapters. This chapter details optimal ways of using transition-edge sensors to achieve maximum photon-number resolution. This is the culmination of the work of many individuals and several years of detector operation experience, but only recently has the detector been capable of resolving the 30 photons it can now routinely see. Chapter 7 (Tomography with PNR Detectors) steps halfway back to theory by describing methods to characterize unknown quantum states through the process of tomography. However, it still keeps one foot firmly rooted in the experimental world, as everything is geared towards ensuring the task is experimentally feasible with the photon-counting abilities of our detectors.

Chapter 8 (Experimental Design) groups together experimental results with ideas originating from all previous chapters. The chapter begins by focusing on important experimental tools and concepts, such as using a pulsed laser system and accessing single photons produced from cavity-enhanced spontaneous parametric down-conversion in an optical parametric oscillator. Once the tools and the specific experimental setup are described, I present results for quantum state tomography from Ch. 7 and quantum state engineering from Ch. 2. The current results for the quantum state engineering experiment indicate some aspects of the experimental design may need to be reconsidered, but even this preliminary data hints that the protocol has merit. Finally, Ch. 9 concludes with a summary of the work in context with all current theoretical and experimental progress.

In addition to the usual tangentially related derivations (Appendices A- D), the appendices also contain Python code segments. Appx. E presents a guided tutorial meant to introduce the reader to quantum simulation with continuous variables. If used in conjunction with the first chapter, this tutorial will aid in the development of phase-space understanding and quantum state evolution. It will also give the reader the necessary tools to reproduce various simulations obtained in Chs. 2-5, if desired. Documented codes for the characterization and processing of transition-edge sensor data is included in Appx. F and G, and codes to perform overlap tomography are given in Appx. H. These codes are additionally hosted online with associated data on GitHub[1].

[1] https://github.com/me3nq/Thesis_docs

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Quantum Optics and Quantum Computation

## 1.1 Overview

Quantum optics is an exciting field of research where one has the opportunity to develop theoretical predictions with deep physical meaning while also being able to experimentally test these predictions with a relative ease not afforded to larger physics-based collaborations. These experiments are able to access fundamental quantum-mechanical behavior such as wave-particle duality, superposition, and entanglement, all while having practical applications. Fortunately, because of these practical applications, quantum optics is enmeshed with the current quantum-technological boom in quantum information.

As this document aims to develop a variety of quantum optical techniques for engineering certain types of exotic quantum states, I have devoted this chapter to a broad overview of quantum optics principles and quantum computation. Some sections have rigorous derivation, but generally this is only the case where it will become useful later. Several sources have been used to compile the material here, but Refs. [1–5] are specific review articles and books that I might recommend for a more complete introduction.

## 1.2 Quantizing the Field

The first step of quantum optics is to arrive at the 'quantum' part. This can be achieved by starting from Maxwell's equations for classical electrodynamics and quantizing the

electromagnetic field. This derivation is something that everyone who wishes to understand quantum optics should go through at least once. That said, it is covered in more detail and elegant fashion elsewhere [4, 6, 7] so I will only hit the main points here. The general concept is that we assume the classical Hamiltonian has the same form as the quantum Hamiltonian due to the correspondence principle. For classical fields, we can write

$$H = \frac{1}{2}\int dV \left[\epsilon_0 \boldsymbol{E}^2(\boldsymbol{r},t) + \tfrac{1}{\mu_0}\boldsymbol{B}^2(\boldsymbol{r},t)\right], \tag{1.1}$$

where $\boldsymbol{E}$ and $\boldsymbol{B}$ are the electric and magnetic field vectors.

Now, imagine we wish to write the solution to Maxwell's equations in a region of space. The travelling wave is a solution of course, but let's instead imagine we are looking inside a perfectly conducting box with fixed dimensions, and then take the limit of infinite volume later. In this case, lacking the presence of charges or currents, Maxwell's equations become

$$\nabla \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t}, \quad \nabla \times \boldsymbol{B} = \mu_0\epsilon_0\frac{\partial \boldsymbol{E}}{\partial t}, \quad \nabla \cdot \boldsymbol{E} = 0, \quad \nabla \cdot \boldsymbol{B} = 0, \tag{1.2}$$

where $\mu_0$ and $\epsilon_0$ in the above are the respective permeability and permittivity of vacuum. These equations along with the boundary conditions of the conducting box have a one-dimensional solution along the $z$-direction given by

$$E_x(z,t) = \left(\frac{2\omega^2}{V\epsilon_0}\right)^{1/2} q(t)\sin(kz), \tag{1.3}$$

$$B_y(z,t) = \frac{\mu_0\epsilon_0}{k}\left(\frac{2\omega^2}{V\epsilon_0}\right)^{1/2} \dot{q}(t)\cos(kz). \tag{1.4}$$

In the above, $V$ is the volume of the box we introduced, $\omega$ is the frequency of the wave[1], and $q(t)$ is a time-dependent factor with dimensions of length that will be thought of as canonical position. If we use canonical momentum and define $p(t) = \dot{q}(t)$, then the Hamiltonian of the single-mode field can be reduced to

$$H = \frac{1}{2}(p^2(t) + \omega^2 q^2(t)). \tag{1.5}$$

Everything so far is classical, but we can promote position and momentum in the Hamiltonian to Hermitian operators, $\hat{q}$ and $\hat{p}$, such that the canonical commutation

[1] The boundary conditions allow waves of frequency $\omega_m = \frac{m\pi c}{L}$ for integer $m > 0$ and box of length $L$, but for now consider only one frequency.

relation of $[\hat{q}, \hat{p}] = i\hbar\mathbb{1}$ is observed.

This is now recognizable as the quantum harmonic oscillator, so we can perform exactly the same treatment for quantized light! One can then define creation ($\hat{a}^\dagger$) and annihilation ($\hat{a}$) operators as usual such that

$$\hat{a} = \frac{\omega\hat{q} + i\hat{p}}{\sqrt{2\hbar\omega}} \tag{1.6}$$

$$\hat{a}^\dagger = \frac{\omega\hat{q} - i\hat{p}}{\sqrt{2\hbar\omega}} \tag{1.7}$$

with the Hamiltonian taking the familiar form of

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \tfrac{1}{2}). \tag{1.8}$$

We can take one additional step and define dimensionless operators such that

$$\hat{Q} = \frac{\hat{a}^\dagger + \hat{a}}{\sqrt{2}} \tag{1.9}$$

$$\hat{P} = \frac{\hat{a} - \hat{a}^\dagger}{i\sqrt{2}} \tag{1.10}$$

where the canonical commutator is given by

$$[\hat{Q}, \hat{P}] = i\mathbb{1}. \tag{1.11}$$

These canonical operators of $\hat{Q}$ and $\hat{P}$ are known respectively as the amplitude and phase quadrature operators of the field. Since the quantized field is that of a harmonic oscillator, $\hat{Q}$ and $\hat{P}$ can be thought of as dimensionless analogs of position and momentum for a particle in a quantum harmonic oscillator potential.

Using the Hamiltonian and the commutation relations, we can arrive at the single-mode electric field which takes the form of an operator given by

$$\hat{\boldsymbol{E}}(\boldsymbol{r}, t) = i\left(\frac{\hbar\omega}{2\epsilon_0 V}\right)^{1/2} \boldsymbol{e}(\hat{a}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}), \tag{1.12}$$

where $\boldsymbol{e}$ is the polarization unit vector and $\boldsymbol{k}$ is the wavevector. The multi-mode field operator is simply a sum of all individual $\hat{\boldsymbol{E}}_{\omega,\boldsymbol{k},\boldsymbol{e}_j}$ field operators defined by frequency, wave-vector direction, and polarization direction. Each of these fields exists in its own

Hilbert space and does not interact with other modes under the evolution of the 'natural' Hamiltonian given by Eq. 1.8 (light does not interact with itself without a medium or other influencing particles). Mathematically, this can also be expressed in terms of commutation — all operators of one mode commute with operators of another mode,

$$[\hat{a}_j, \hat{a}_k^\dagger] = \delta_{j,k}\mathbb{1}. \tag{1.13}$$

Taking the limit of $V \to \infty$ allows for a continuum of frequency modes as solutions to the wave equations.

Similar to the electric field, a single-mode $\hat{\boldsymbol{B}}$ field operator is

$$\hat{\boldsymbol{B}}(\boldsymbol{r}, t) = \frac{i}{c}\left(\frac{\boldsymbol{k}\times\boldsymbol{e}}{|\boldsymbol{k}|}\right)\left(\frac{\hbar\omega}{2\epsilon_0 V}\right)^{1/2}(\hat{a}e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}). \tag{1.14}$$

An important point to note is that the creation and annihilation operators contained within Eqs. 1.12 & 1.14 have implicit time dependence. Written this way, these operators can be thought of as operators in the Heisenberg picture that are evolved in time by the Hamiltonian. This will be discussed more in the next section, but the result corresponds with the evolution of the classical field as is given by

$$\hat{a}(t) = \hat{a}(0)e^{i\omega t}; \quad \hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{-i\omega t}. \tag{1.15}$$

Since each mode has its own associated Hilbert space, I will drop all subscripts of $\omega, \boldsymbol{k}, \boldsymbol{e}$ when describing single mode states and operators with the implicit assumption that all operators pertain to the same mode. When describing multi-mode states and operators, I will simply label each mode with a single alphanumeric index to differentiate it from others.

## 1.3 State Representations

Once one has quantized the field and has a Hamiltonian, it is appropriate to start finding ways to represent a quantum state. As a quantum harmonic oscillator, this prescription is fairly straightforward. A first way to consider representing states is to use the energy eigenstates of the Hamiltonian, $|n\rangle$, such that the creation and annihilation

operators act on the state as

$$\hat{a}\,|n\rangle = \sqrt{n}\,|n-1\rangle \tag{1.16}$$

$$\hat{a}^\dagger\,|n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \tag{1.17}$$

with $\hat{a}\,|0\rangle = 0$. These eigenstates determine the number of energy quanta, $n$, in the state; thus, $n$ is exactly the number of photons contained in the light mode. We can then define the photon number operator,

$$\hat{n} \equiv \hat{a}^\dagger \hat{a} \tag{1.18}$$

such that $\hat{n}\,|n\rangle = n\,|n\rangle$. Note that just as the ground state of any harmonic oscillator has nonzero energy, the ground state $|0\rangle$, otherwise known as vacuum, has *nonzero* energy while also containing *zero* photons. Despite there being zero photons, the vacuum still possesses non-zero fluctuations in the field quadratures. When measured, the recorded quadrature values average to zero, but they do possess a non-zero variance due to the Heisenberg relation that $\Delta Q \Delta P \geq \frac{1}{2}$. As an additional oddity, the continuum of modes means that each has a vacuum state with non-zero energy leading to infinite total vacuum energy. This infinity is regularized by some unknown mechanism, leaving the fundamental nature of the vacuum up for debate [4]. Fortunately, these issues don't crop up anywhere in this work as we are only concerned with quadrature distributions and the number of photons in a particular mode, which are both quantities unaffected by an arbitrary energy offset.

These energy eigenstates are known as Fock states and form an orthonormal basis for our Hilbert space. We have both

$$\langle m|\,n\rangle = \delta_{m,n} \tag{1.19}$$

$$\sum_{n=0}^{\infty} |n\rangle\,\langle n| = \mathbb{1}. \tag{1.20}$$

Thus, any quantum state can be represented as

$$|\psi\rangle = \sum_{n=0}^{\infty} c_n\,|n\rangle\,, \qquad \langle\psi|\,\psi\rangle = \sum c_n^* c_n = 1. \tag{1.21}$$

This representation allows for a discrete, but infinite, Hilbert space to describe each

mode. Alternatively, we could use a continuous basis based on eigenstates of the amplitude and phase operators. These quadrature eigenstates satisfy

$$\hat{Q}\,|s\rangle_q = s\,|s\rangle_q \tag{1.22}$$

$$\hat{P}\,|t\rangle_p = t\,|t\rangle_p\,. \tag{1.23}$$

In the above, $s$, $t \in \mathbb{R}$ and the subscript indicates the basis of the eigenstate. These bases are also orthogonal so that

$${}_q\langle s|s'\rangle_q = \delta(s-s') \tag{1.24}$$

$${}_p\langle t|t'\rangle_p = \delta(t-t') \tag{1.25}$$

$$\int_{-\infty}^{\infty} ds\,(|s\rangle\,\langle s|)_q = \int_{-\infty}^{\infty} dt\,(|t\rangle\,\langle t|)_p = \mathbb{1}. \tag{1.26}$$

Additionally, the eigenstates of these conjugate bases are related by the Fourier transform:

$$|t\rangle_p = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ds\, e^{ist}|s\rangle_q \tag{1.27}$$

$$|s\rangle_q = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dt\, e^{-ist}|t\rangle_p. \tag{1.28}$$

Quadrature eigenstates are quite interesting in that they, just like definite position and momentum states of a particle in a potential well, are not normalizeable. This means that only approximate quadrature eigenstates are actually physical, as a true eigenstate of $\hat{Q}$ would require an infinite series of Fock-basis components to represent and would thus require infinite energy. Nevertheless, the quadrature basis is useful to describe the wavefunction of a normalized state,

$$\psi_q(x) = {}_q\langle x|\psi\rangle \tag{1.29}$$

$$\psi_p(x) = {}_p\langle x|\psi\rangle, \tag{1.30}$$

so that the quantum state can be described by

$$|\psi\rangle = \int dx\,\psi_q(x)|x\rangle_q = \int dx\,\psi_p(x)|x\rangle_p. \tag{1.31}$$

Occasionally, other texts use tilde notation to denote the Fourier transform of a function, such as writing $\psi_q(x) \equiv \psi(x)$ and $\psi_p(x) \equiv \tilde{\psi}(x)$. In some cases, however, it is useful

to consider a $q$ or $p$-quadrature wavefunction at the same time as a wavefunction of a rotated basis in a linear combination of both $q$ and $p$. Thus, reserving the subscript for basis (and mode index) generally keeps things more clear.

The Fock state basis can be related to the quadrature basis by using the wavefunction of the energy eigenstates of the harmonic oscillator, which end up involving Hermite polynomials [8]. As energy eigenstates, the Fock states have no defined phase and are invariant under the Fourier transform: $\psi_q(x) = \psi_p(x) \equiv \psi_n(x)$. Written out, the Fock states are given by

$$\psi_n(x) = \frac{\pi^{-1/4}}{\sqrt{n!2^n}} e^{-x^2/2} H_n(x), \tag{1.32}$$

and $H_n(x)$ are the physicist's Hermite polynomials.

When manipulating states, it can be extremely useful to represent any quantum state as a function of a quadrature operator applied to the zero eigenstate of the conjugate quadrature. This arises through Fourier transform and the freedom to alternate between writing the eigenvalue of an operator applied to an eigenmode and that operator acting on the same eigenmode. For example, we can write a generic ket as

$$\begin{aligned} |\psi\rangle &= \int dx\, \psi_q(x) \, |x\rangle_q \\ &= \int dx\, \psi_q(\hat{Q}) \, |x\rangle_q \qquad (1.33) \\ &= \sqrt{2\pi} \psi_q(\hat{Q}) \, |0\rangle_p \,. \qquad (1.34) \end{aligned}$$

In the second line, promoting the wavefunction $\psi_q(x)$ to a function of an operator, $\psi_q(\hat{Q})$, allows us to pull it out of the integral and use Eq. 1.27 to write the state as a zero-momentum eigenstate.

### 1.3.1 Density Operators

This section introduced the Fock and quadrature bases as a way to represent pure states, but this can be extended to statistical mixtures through the use of the density matrix. In the Fock basis, the most general density matrix can be written as

$$\rho = \sum_{m,n} \rho_{mn} \, |m\rangle \, \langle n| \,, \tag{1.35}$$

where $\rho_{mn} \in \mathbb{C}$ and we have the constraints that $\rho$ is a normalized, positive semi-definite matrix with

$$\mathrm{Tr}[\rho] = 1; \quad \rho = \rho^{\dagger}. \tag{1.36}$$

For pure states, we have that $\rho = |\psi\rangle \langle\psi|$ which implies the additional identity that $\rho^2 = \rho$, thus $\mathrm{Tr}[\rho^2] = 1$. Impure states are statistical mixtures that can be obtained by performing a partial trace over an entangled state. These mixtures have

$$\rho = \sum_i p_i \, |\psi_i\rangle \langle\psi_i| \, , \tag{1.37}$$

where $p_i \in [0, 1]$ are the probabilities of having the pure state $|\psi_i\rangle$ with $\sum_i p_i = 1$.

In addition to this matrix representation, $\rho$ can be represented in a continuous basis as a density operator. In the $q$-basis, this is written as

$$\rho = \iint_{-\infty}^{\infty} ds dt \, \rho_{st} \left( |s\rangle \langle t| \right)_q \tag{1.38}$$

Regardless of representation, the expectation value of an operator $\hat{A}$ is calculated as

$$\langle \hat{A} \rangle = \mathrm{Tr}[\rho \hat{A}]. \tag{1.39}$$

### 1.3.2 Wigner function

The density matrix provides a means to completely describe any quantum state whether it be pure or mixed. However, there are alternative representations using quasi-probability distributions that also provide complete state information. Throughout much of this work, I will make use of the Wigner quasi-probability distribution, or Wigner function, which is a valuable tool for visual representation of quantum states as well as providing simple indicators of non-classicality. The Wigner function of a quantum state is defined from the density matrix according to

$$W(q, p) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{2ipy} {}_q\langle q - y|\rho|q + y\rangle_q dy. \tag{1.40}$$

Here, $q$ and $p$ are the eigenvalues of the respective amplitude and phase quadratures defined earlier. The Wigner function is known as a 'quasi' probability distribution due to the fact that it doesn't yield probabilities directly, but integrating over one quadrature

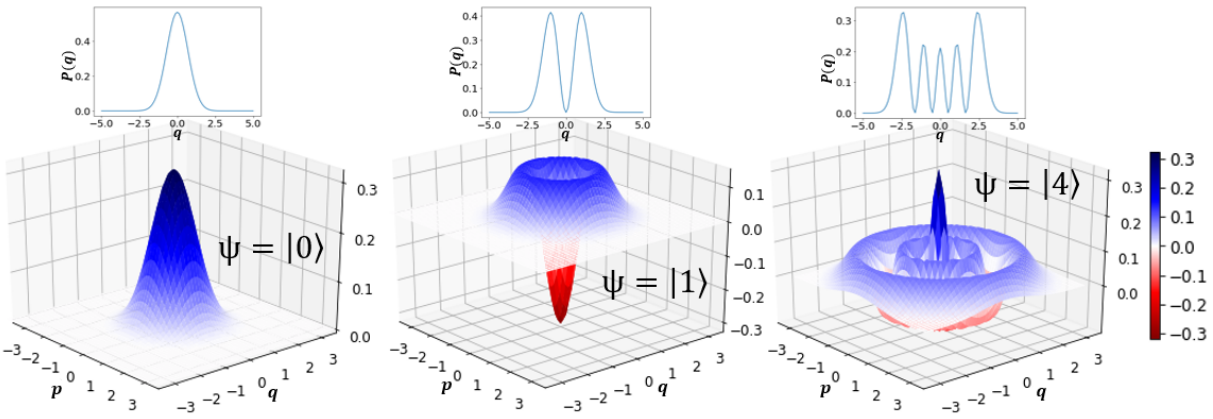



FIGURE 1.1: Wigner functions are plotted for the Fock states $|0\rangle$ , $|1\rangle$ , and $|4\rangle$.

will result in the probability distribution for the other. Mathematically, we have

$$\int_{-\infty}^{\infty} dp\, W(q,p) = |\psi_q(q)|^2 \tag{1.41}$$

$$\int_{-\infty}^{\infty} dq\, W(q,p) = |\psi_p(p)|^2, \tag{1.42}$$

which directly implies normalization of

$$\iint_{-\infty}^{\infty} dpdq\, W(q,p) = 1. \tag{1.43}$$

Additionally, as a quasi-probability distribution, the Wigner function has the potential to be non-positive[2]. This is extremely powerful, as negativity of the Wigner function is a *necessary* condition to achieve a quantum advantage over classical systems [9]. This is a direct result of the fact that continuous-variable states with Gaussian distributions are efficiently simulable with classical computation [10], and only states with non-Gaussian distribution can have Wigner functions with any negativity [11]. As an example, Fig. 1.1 shows the Wigner functions for several Fock states with negative regions shown in red. The vacuum state is the only Fock state with a Gaussian distribution — thus a solely positive Wigner function. The graphic also shows the associated probability distribution, $P(q) = |\psi_q(q)|^2$, which was obtained by integrating the quasi-probability distribution over the $p$-quadrature using Eq. 1.41 An important point to note is that while negativity of the Wigner function is necessary for a quantum

[2]The Wigner function is bounded by the interval $[-\frac{1}{\pi}, \frac{1}{\pi}]$.

speed-up, it is not sufficient, even for pure states [12].

## 1.4 Operators and State Evolution

Quantum-mechanical quantities, such as states, follow evolutionary processes that are coherent and reversible. When a coherent process is well-separated from the environment such that the state in question has no interaction with a reservoir so is not influenced by dissipation (no partial trace needed), and the state is not subject to quantum measurement, then both the purity and norm of a state are maintained. Under this form of evolution, we can operate in the Schrödinger picture, where states evolve with respect to some parameter (e.g., time) and operators remain fixed. Alternatively, we can use the Heisenberg picture which allows the operators to change while the state remains fixed.

Intuitively, the distinction between these two pictures can be understood in relatively simple terms. For any quantum mechanical system, we must eventually be concerned with a measurement associated with an observable. Having an observable in mind from the outset of any quantum experiment means we can manipulate our state by applying different operators (Schrödinger picture) and see how the outcomes of measurements of the same observable change as the state is evolved. Alternatively, instead of allowing the state to evolve, we can leave the state alone and perform *different* measurements that correspond to observables that have been evolved by the same physical process (Heisenberg picture). In both cases the measurement results will be the same, so effectively both pictures are equivalent. The duality of these pictures comes about by the physicist asking the questions: Is it easier to modify the state so I can measure it how I'd like? Or, is it easier to leave the state alone and change the way I measure so that the result would be the same?

There is actually a third picture, known as the interaction picture, which is just a compromise where both states and operators evolve. Here, the Hamiltonian is broken into 'free' and 'coupled' evolution terms, where state evolution is easier under the 'free' terms. The operators are then evolved under the 'coupled' terms where the states would otherwise prove difficult.

### 1.4.1 Schrodinger vs. Heisenberg

Begin by considering a quantum state, $|\psi\rangle$, subjected to a coherent physical process. Because the state is pure and a superposition of basis states, the evolution must follow a linear differential equation. Additionally, because we require that $|\psi\rangle$ contains complete information about the state before evolution, the equation must be of only first order in the parameter $t$; a higher order equation would mean that derivatives of $|\psi\rangle$ are *also* necessary to define initial conditions [4]. In the Shrödinger picture, we have

$$i\hbar\frac{\mathrm{d}\,|\psi\rangle}{\mathrm{d}t} = \hat{H}\,|\psi\rangle\,, \tag{1.44}$$

where the Hamiltonian, $\hat{H}$, is a linear operator that is related to energy when $t$ is time. By requiring that the norm of $|\psi\rangle$ must be preserved, we find that $\hat{H}^\dagger = \hat{H}$, so $\hat{H}$ is Hermitian. The Schrödinger evolution of the state can be written as

$$|\psi(t)\rangle = \hat{U}\,|\psi(t_0)\rangle\,, \tag{1.45}$$

and if $\hat{H}$ is time-independent[3], then we have that the unitary operator $\hat{U}$ is given by

$$\hat{U} = e^{-\frac{i}{\hbar}\hat{H}(t-t_0)} \tag{1.46}$$

By examining how expectation values of operators evolve, we can discover the Heisenberg picture, where only operators evolve. This is found by differentiating the expectation value of operator $\hat{A}$ and using the Hermiticity of $\hat{H}$ along with the Schrödinger equation to arrive at the Heisenberg equation:

$$i\hbar\frac{\mathrm{d}\hat{A}_H}{\mathrm{d}t} = [\hat{A}_H, \hat{H}_H] + i\hbar\frac{\partial\hat{A}_s}{\partial t}. \tag{1.47}$$

The $H$ and $S$ subscripts in the above equation denote in which picture the operators are represented. If the operator $\hat{A}$ has no explicit time dependence in the Schrödinger picture, then the equation reduces to

$$i\hbar\frac{\mathrm{d}\hat{A}}{\mathrm{d}t} = [\hat{A}, \hat{H}] \tag{1.48}$$

[3]If $\hat{H}$ depends on time, then the argument of the exponential in $\hat{U}$ is an integral of $\hat{H}$ over $t_0$ to $t$, which must be time-ordered if $\hat{H}$ does not commute with itself at different times. None of this will apply for this work.

with the $H$ subscripts dropped. Now, operators can be evolved in the Heisenberg picture according to

$$\hat{A} = \hat{U}^\dagger \hat{A} \hat{U}, \tag{1.49}$$

where $\hat{U}$ is the same unitary operator as in the Schrödinger picture. Evolution in this way ensures that expectation values agree regardless of which picture is used.

For a density matrix in the Schrödinger picture, we have that

$$\rho(t) = \hat{U}\rho(t_0)\hat{U}^\dagger \tag{1.50}$$

and after evolution, the expectation value of $\hat{A}$ is given by

$$\langle \hat{A} \rangle = \mathrm{Tr}[\rho(t)\hat{A}_S] = \mathrm{Tr}[\hat{U}\rho(t_0)\hat{U}^\dagger \hat{A}_S]. \tag{1.51}$$

In the Heisenberg picture, we have

$$\langle \hat{A} \rangle = \mathrm{Tr}[\rho(t_0)\hat{A}_H(t)] = \mathrm{Tr}[\rho(t_0)\hat{U}^\dagger \hat{A}_S \hat{U}], \tag{1.52}$$

which is the same result due to the cyclic property of the trace.

### 1.4.2 Baker-Campbell-Hausdorff formula

When evolving operators, one can solve the differential equation given by the Heisenberg picture. However, one can also make use of the fact that since unitary operators satisfy $\hat{U}^\dagger \hat{U} = \mathbb{1}$, it can be easier to simply make use of commutation relations. This allows one to write

$$\hat{U}^\dagger \hat{O} \hat{U} = \hat{U}^\dagger \hat{U} \hat{O}' = \hat{O}', \tag{1.53}$$

where $\hat{O}'$ can be found from Eq. 1.47 or the commutation

$$\hat{O}' = \hat{O} + \hat{U}^\dagger [\hat{O}, \hat{U}]. \tag{1.54}$$

This commutator is not always easy to solve, but the Baker-Campbell-Hausdorff (BCH) formula can be particularly useful. Suppose we have two operators, $\hat{X}$ and $\hat{Y}$. If $[\hat{X}, \hat{Y}]$

is *central* (i.e. $[\hat{X},[\hat{X},\hat{Y}]] = [\hat{Y},[\hat{X},\hat{Y}]] = 0$), then we have the formulas

$$e^{s\hat{X}}\hat{Y}e^{-s\hat{X}} = \hat{Y} + s[\hat{X},\hat{Y}], \tag{1.55}$$

$$e^{t(\hat{X}+\hat{Y})} = e^{t\hat{X}}\, e^{t\hat{Y}}\, e^{-\frac{t^2}{2}[\hat{X},\hat{Y}]}. \tag{1.56}$$

The BCH formula extends to non-central $[\hat{X},\hat{Y}]$, but this can become quite cumbersome as often when the commutator of commutators does not vanish, the extended commutators with $\hat{X}$ and $\hat{Y}$ must be computed ad infinitum. Occasionally one might be fortunate and the infinite series will converge to something recognizable, but this is not always the case. The extended version of BCH is given by the identity

$$e^{\hat{X}}e^{\hat{Y}} = e^{(\hat{Y}+[\hat{X},\hat{Y}]+\frac{1}{2!}[\hat{X},[\hat{X},\hat{Y}]]+\frac{1}{3!}[\hat{X},[\hat{X},[\hat{X},\hat{Y}]]]+\cdots)}e^{\hat{X}}. \tag{1.57}$$

### 1.4.3 Rotations in phase space: optical phase shifts

From quantizing the field, we found the form of the evolution Hamiltonian in Eq. 1.8. When writing this as a unitary evolution acting over time $t$, we have that

$$\hat{U}(t) = e^{-i\omega t\left(\hat{a}^\dagger\hat{a}+\frac{1}{2}\right)}. \tag{1.58}$$

By parameterizing such that $\omega t = \theta$ and ignoring the global phase of $e^{-i\frac{\omega t}{2}}$ since it will not change expectation values, we can redefine this operator as[4]

$$\hat{R}(\theta) = e^{-i\theta a^\dagger\hat{a}} = e^{-\frac{i\theta}{2}(\hat{Q}^2+\hat{P}^2+1)}. \tag{1.59}$$

For reasons that will soon become apparent, this is known as the rotation operator, and it can be implemented by performing an optical phase shift on the particular mode.

To see how this operator behaves, let's see how it transforms the creation and annihilation operators if we use the Heisenberg picture. Using Eq. 1.47 and using natural units such that $\hbar = 1$, we have that

$$i\frac{\mathrm{d}\hat{a}}{\mathrm{d}t} = [\hat{a},\hat{a}^\dagger\hat{a}], \tag{1.60}$$

[4]Sometimes defined with $\theta \to -\theta$.

which leads to

$$\dot{\hat{a}} = -i\hat{a}$$
$$\hat{a}(\theta) = e^{-i\theta}\hat{a}(0). \tag{1.61}$$

Thus, we know that the operator transforms $\hat{a}$ and $\hat{a}^\dagger$ as

$$\hat{R}^\dagger(\theta)\hat{a}\hat{R}(\theta) \to e^{-i\theta}\hat{a} \tag{1.62}$$
$$\hat{R}^\dagger(\theta)\hat{a}^\dagger\hat{R}(\theta) \to e^{i\theta}\hat{a}^\dagger. \tag{1.63}$$

Using the definition given by Eq. 1.9, we see that the $\hat{Q}$ operator transforms as

$$\hat{R}^\dagger(\theta)\hat{Q}\hat{R}(\theta) \to \hat{A}(\theta) \equiv \hat{Q}\cos(\theta) + \hat{P}\sin(\theta), \tag{1.64}$$

where $\hat{A}(\theta)$ is a general rotated quadrature. This is why the operator described so far is known as the rotation operator: it acts to rotate the state in phase space. Since this is what happens as the field evolves in time under the free Hamiltonian, this operation can be physically enacted simply by changing the path length of a traveling field, or by applying an optical phase shift. Note that an *optical phase* shift is very different from a shift, i.e., a translation or displacement, of the *phase quadrature*, which will be discussed in the next section. The unfortunate naming conventions can be a major source of confusion if one is not yet accustomed.

### 1.4.4 Translations in phase space: field displacements

Besides rotations, translations are also a fundamental quantum mechanical operation. For a given operator, the translation operator is generated by the gradient of that operator. This means that for a pair of conjugate operators, such as position and momentum, the translation operator for one is generated by the other. Thus, a translation by $s$ in the $q$-quadrature is enacted by

$$\hat{X}(s) \equiv e^{-is\hat{P}} \tag{1.65}$$
$$\hat{X}(s)\,|q\rangle_q = |q+s\rangle_q\,. \tag{1.66}$$

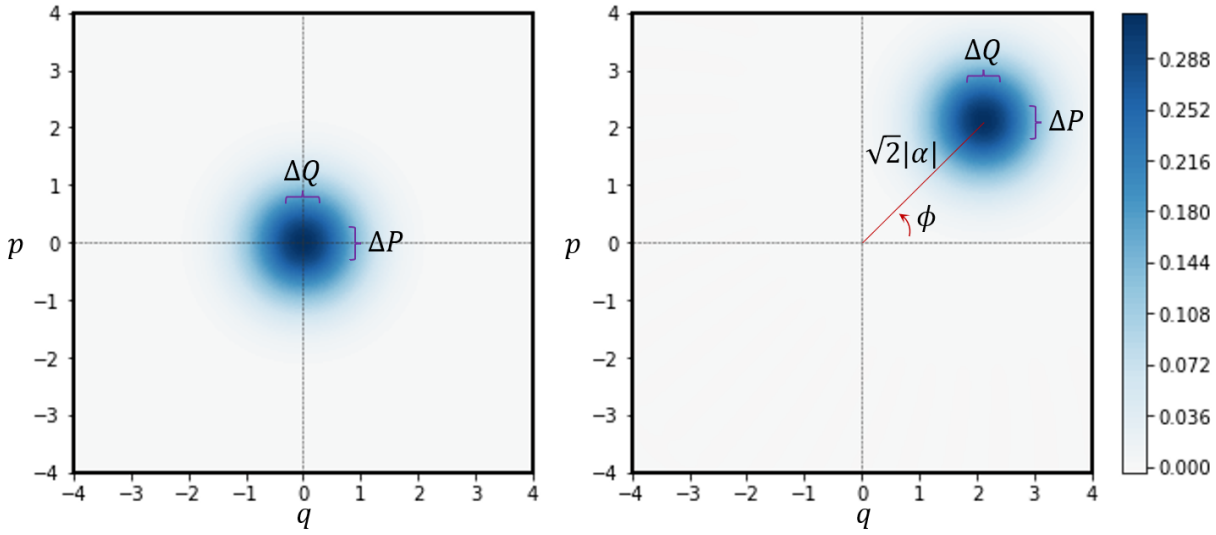



FIGURE 1.2: Wigner functions are plotted for the vacuum and displaced vacuum.

Translations in the conjugate phase quadrature by an amount $t$ is given by

$$\hat{Z}(t) \equiv e^{-it\hat{Q}} \tag{1.67}$$

$$\hat{Z}(t)\,|p\rangle_p = |p+t\rangle_p\,. \tag{1.68}$$

These operators are named as $\hat{X}$ and $\hat{Z}$ due to their uses in CV quantum computation as we will see later. $\hat{X}$ and $\hat{Z}$ provide a means to shift quantum states along the quadrature axis, but often one makes use of a combined displacement operator given by

$$\hat{D}(\alpha) \equiv e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}}. \tag{1.69}$$

Using the BCH formula and Eq. 1.9 reveals that the displacement operator can also be written as

$$\hat{D}(\alpha) = e^{-i\mathrm{Re}[\alpha]\mathrm{Im}[\alpha]}\hat{Z}\left(\sqrt{2}\mathrm{Im}[\alpha]\right)\hat{X}\left(\sqrt{2}\mathrm{Re}[\alpha]\right). \tag{1.70}$$

This displacement operator acts to push states around in phase space. In terms of the Wigner function, the distribution remains the same but is merely translated in both $q$ and $p$.

### Coherent states

When applied to vacuum, the displacement operator creates coherent states, which are eigenstates of the annihilation operator. These states are given by

$$\hat{D}(\alpha)\,|0\rangle \equiv |\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle\,. \tag{1.71}$$

Coherent states have the property that $\hat{a}\,|\alpha\rangle = \alpha\,|\alpha\rangle$, which essentially means that coherent states remain coherent states after photon loss[5]. These states are described by coherent emission of photons with Poissonian statistics and have a defined phase and amplitude but are the most 'classical' of the quantum states. These states are easily accessible by experiment as they are produced by the output of a phase and intensity-stabilized laser.

Viewing coherent states as displaced vacuum makes it evident that beyond being shifted in the quadratures, the uncertainty of $\hat{Q}$ and $\hat{P}$ are symmetric and the same as $|0\rangle$. This can be easily calculated as

$$\begin{aligned}(\Delta Q)^2 &= \langle\alpha|\,\hat{Q}^2\,|\alpha\rangle - \langle\alpha|\,\hat{Q}\,|\alpha\rangle^2 \\ &= \tfrac{1}{2}\,\langle\alpha|\,(\hat{a}^2 + \hat{a}^{\dagger 2} + \hat{a}\hat{a}^\dagger + \hat{a}^\dagger\hat{a})\,|\alpha\rangle - \tfrac{1}{2}\,\langle\alpha|\,(\hat{a} + \hat{a}^\dagger)\,|\alpha\rangle^2 \\ &= \tfrac{1}{2}(\alpha^2 + \alpha^{*2} + 2|\alpha|^2 + 1) - \tfrac{1}{2}(\alpha + \alpha^*)^2 = \tfrac{1}{2},\end{aligned} \tag{1.72}$$

where in the last line of the above I have used $\hat{a}\hat{a}^\dagger = [\hat{a}, \hat{a}^\dagger] + \hat{a}^\dagger\hat{a}$. Because the vacuum is symmetric in phase space, the uncertainty in the displaced vacuum will also be symmetric (can be quickly checked) which leads to $\Delta Q = \Delta P$. Thus, vacuum and all coherent states are minimum uncertainty states as they saturate the Heisenberg uncertainty relation:

$$\Delta Q \Delta P = \tfrac{1}{2} \tag{1.73}$$

Figure 1.2 plots the Wigner functions for a coherent state and vacuum state. The phase of the coherent state is determined by the phase of the complex number $\alpha = |\alpha|e^{i\phi}$, and the distance from the origin is related to the average number of photons present in the state,

$$\langle\hat{a}^\dagger\hat{a}\rangle = \tfrac{1}{2}\langle\hat{Q}^2 + \hat{P}^2 - 1\rangle = |\alpha|^2. \tag{1.74}$$

[5] Albeit with a different amplitude. This is discussed more after introducing Kraus operators for loss.

### 1.4.5 Phase-space symmetry

Now that I have introduced the basics of phase-space and state representations with the Wigner function, it's worth building a little more intuition into phase-space by looking at rotational symmetries. Here I give a simple proof demonstrating that quantum state symmetry in phase space is directly related to its Fock basis support: if a pure state has $N$-fold symmetry in phase-space, then the Fock components comprising it are all separated by multiples of $N$.

To better quantify phase-space symmetry, we say that a pure state $|\psi\rangle$ has $N$-fold symmetry if there are exactly $N$ values of $\theta \in [0, 2\pi)$ such that the Wigner function remains unchanged after rotating through the angle $\theta = \frac{2\pi}{N}k$, $k \in \mathbb{Z}$ about the origin of phase space. For example, a coherent state has symmetry of $N = 1$ as it must undergo a full $2\pi$ rotation before returning to the initial state, a cat state has symmetry of $N = 2$, and the example states in Fig. 1.3 have symmetry of $N = 4$. Each Fock state is an eigenstate of the rotation operator, and thus has infinite symmetry. Defined this way, if $|\psi\rangle$ is $N$-symmetric, then $|\psi\rangle$ must be an eigenstate of the rotation operator with a phase of $\frac{2\pi}{N}$ and satisfies

$$\hat{R}\left(\frac{2\pi}{N}\right)|\psi\rangle = \lambda|\psi\rangle, \tag{1.75}$$

where the rotation operator is defined by Eq. 1.59. Starting with a generic pure quantum state defined as

$$|\phi\rangle = \sum_{n=0}^{\infty} \phi_n |n\rangle, \tag{1.76}$$

applying the rotation operator leads to

$$\hat{R}\left(\frac{2\pi}{N}\right)|\phi\rangle = \sum_{n=0}^{\infty} \phi_n e^{-i\frac{2\pi}{N}a^\dagger a}|n\rangle \tag{1.77}$$

$$= \sum_{n=0}^{\infty} \phi_n e^{-i\frac{2\pi}{N}n}|n\rangle. \tag{1.78}$$

Suppose we consider the class of states such that that all $\phi_n = 0$ unless $n = M + kN$ for $k \in \mathbb{Z}$. Then we have that

$$\hat{R}\left(\frac{2\pi}{N}\right)|\phi\rangle = \sum_{k=0}^{\infty} \phi_n e^{-i\frac{2\pi(M+kN)}{N}a^\dagger a}|M + kN\rangle = e^{-i\frac{2\pi M}{N}}|\phi\rangle. \tag{1.79}$$

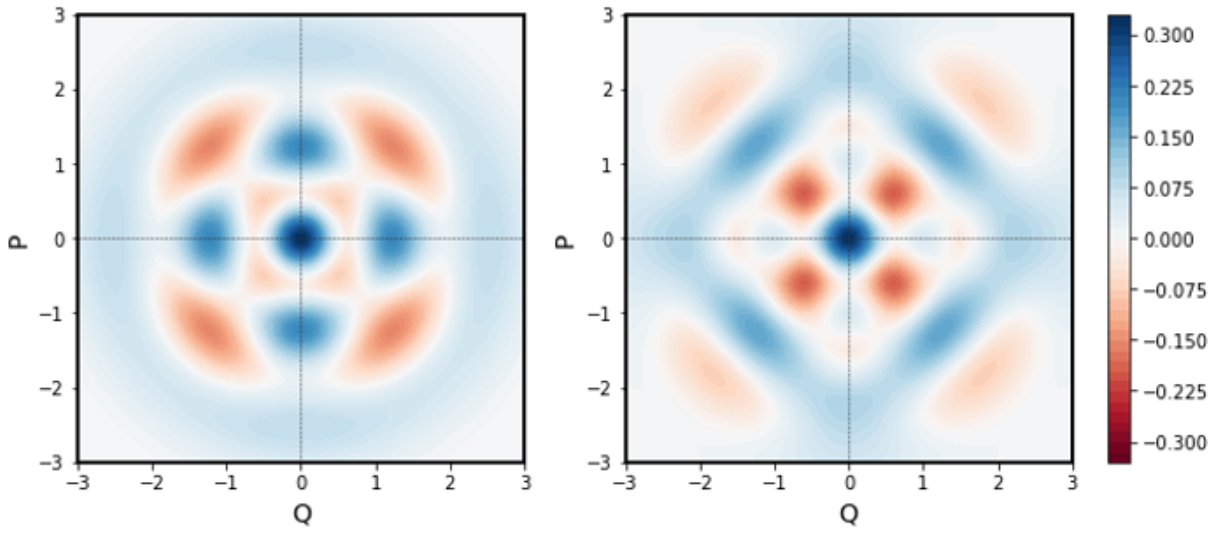



FIGURE 1.3: The Wigner quasi-probability distribution of binomial states $\frac{1}{2}\left(|0\rangle + \sqrt{3}|4\rangle\right)$ (left) and $\frac{1}{2}\left(\sqrt{3}|2\rangle + |6\rangle\right)$ (right) have $\frac{\pi}{2}$-symmetry in phase space.

Therefore, $|\phi\rangle$ has $N$-fold symmetry in phase space if it is of the form

$$|\phi\rangle = \phi_0|M\rangle + \phi_1|M+N\rangle + ... + \phi_k|M+kN\rangle \tag{1.80}$$

for integers $M$, $N$, and $k$.

An example of this type of symmetry is shown in Fig. 1.3, where the states each have photon number components separated by four, leading to $\frac{\pi}{2}$ rotation symmetry. From this argument, one can simply look at a Wigner function and be able to determine not only signatures of 'quantumness' from Wigner negativity but also information about photon number through simple symmetry arguments. In fact, if one has a Wigner function that has rotational symmetry about an axis other than out-of-plane through the origin, then it is merely a phase-space translation (displacement) away from returning to a state centered at the origin having support on only a subset of the Fock basis. As we will see in the next section, the squeezed vacuum has only even photon number. This could be guessed simply by realizing that while vacuum has unbroken rotation symmetry (thus a single Fock component), squeezing the vacuum in phase-space acts to break this symmetry and reduce it to $\pi$ symmetry only, which means it must have only every-other photon number in the Fock-basis superposition.

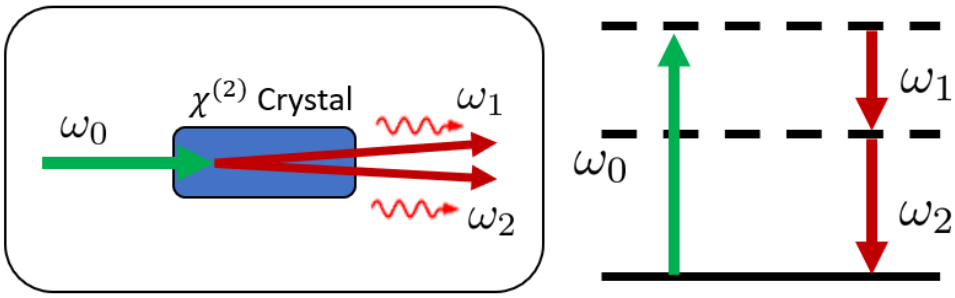



FIGURE 1.4: Parametric down-conversion. Pump photons excite virtual energy levels that decay into photon pairs. This process is reversible and is known as second-harmonic generation when $\omega_1 = \omega_2 = \omega$ are used to generate light at $2\omega$.

### 1.4.6 Squeezing

While the vacuum and coherent states are both minimum uncertainty states, their uncertainty is symmetric across the generalized quadrature $\hat{A}(\theta)$ (defined in Eq. 1.64) in that $\Delta\hat{A}(\theta) = 2^{-1/2}$ for all values of $\theta$. Can we instead make a state that saturates the Heisenberg inequality but is asymmetric in the uncertainty? The answer is yes, as this is precisely what the squeezing operator does.

The squeezing operator arises from the nonlinear optical process of down-conversion depicted in Fig. 1.4 where there is a probability for a pump photon at frequency $\omega_0$ to down-convert into two photons such that energy conservation is satisfied:

$$\omega_1 + \omega_2 = \omega_0. \tag{1.81}$$

This process must also be phase-matched so that momentum is conserved:

$$\Delta k = k_0 - k_1 - k_2 \approx 0. \tag{1.82}$$

The phase-matching requirement is written as being approximately zero since the nonlinear process can still occur for nonzero $\Delta k$, where the intensity of the down-converted light will scale as $I \propto \text{sinc}^2(\Delta k \frac{L}{2})$ for a crystal of length $L$. This condition allows for clever engineering of nonlinear crystal optics that are only quasi-phase matched, which allows for a compromise to make experimental implementation simpler while also preserving a large enough effective nonlinearity [13].

For now, suppose that we only focus on the degenerate process such that $\omega_1 = \omega_2 = \omega_0/2$. In this case and in the regime where a small portion of the pump beam

is down-converted, known as the undepleted pump approximation, the pump intensity combines with phase matching and intrinsic crystal properties to give an effective squeezing Hamiltonian of

$$\hat{H}_{sq} = \frac{i\hbar}{2}\chi^{(2)}(\hat{a}^{\dagger 2} - \hat{a}^2). \tag{1.83}$$

Here, the pump is undepleted and we are considering the degenerate frequency case, this is a single-mode Hamiltonian for the down-converted field at $\omega_0/2$. Parameterizing as $\imath = \chi^{(2)}t$ leads to the single-mode squeezing operator given by

$$\hat{S}(\imath) = e^{\frac{\imath}{2}(\hat{a}^{\dagger 2} - \hat{a}^2)} = e^{-\frac{i\imath}{2}(\hat{Q}\hat{P} + \hat{P}\hat{Q})}. \tag{1.84}$$

As the name suggests, the squeezing operator acts to scale the quadratures by compressing one axis. To maintain the uncertainty relation, the conjugate axis must be stretched. This scaling can be understood by determining the effects of $\hat{S}$ on the quadrature operators. Using Eq. 1.47, we find that

$$\hat{S}^\dagger(\imath)\hat{Q}\hat{S}(\imath) \to e^{\imath}\hat{Q} \tag{1.85}$$

$$\hat{S}^\dagger(\imath)\hat{P}\hat{S}(\imath) \to e^{-\imath}\hat{P}. \tag{1.86}$$

The application of the squeezing operator to vacuum is shown in Fig. 1.5, where the vacuum is squeezed by 6 dB in the central plot and 12 dB in the plot at right. The amount of squeezing is often measured in decibels below the vacuum noise leading to a conversion of

$$s_{dB} = 20\log_{10}\left(\frac{\Delta P_0}{\Delta P_\imath}\right) = 20\log_{10}(e^{\imath}). \tag{1.87}$$

Occasionally in other quantum optics texts, the squeezing operator will be defined with a complex $\imath$ so that an arbitrary quadrature can be squeezed. However, here I will simply leave $\imath$ as a real parameter where $\imath > 0$ means the $p$-quadrature will be squeezed, $\imath < 0$ means the $q$-quadrature will be squeezed, and squeezing of an arbitrary quadrature operator $\hat{A}(\theta + \pi/2)$ can be achieved by applying the operator $\hat{R}^\dagger(\theta)\hat{S}\hat{R}(\theta)$.

Experimentally, squeezing the vacuum has been the most successful route to produce the largest squeezing, with levels reaching 15 dB [14]. In the photon-number basis, we can write squeezed vacuum as

$$\hat{S}(\imath)\,|0\rangle = \frac{1}{\sqrt{\cosh\imath}}\sum_{n=0}^{\infty}\tanh^n\imath\,\frac{\sqrt{(2n)!}}{2^n n!}\,|2n\rangle\,. \tag{1.88}$$

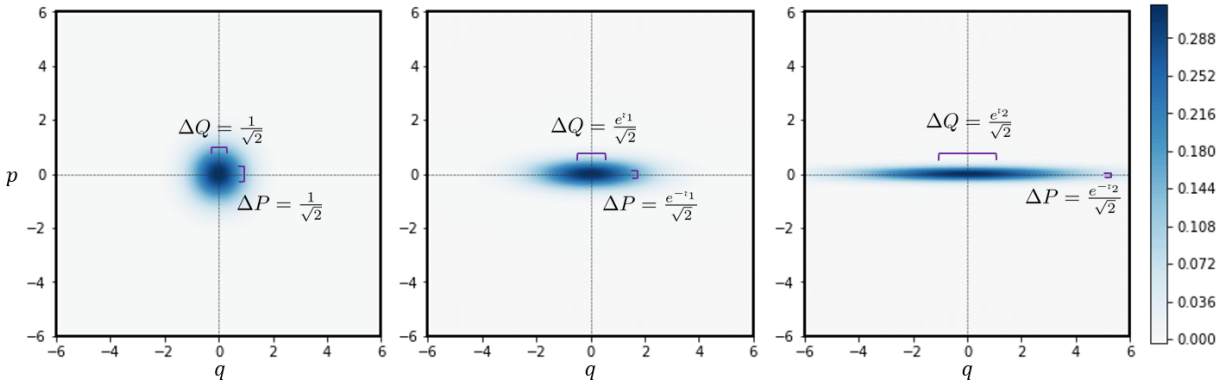



FIGURE 1.5: Theoretical Wigner functions are plotted for the vacuum and squeezed vacuum. The middle plot is squeezed by 6 dB while the right plot is squeezed by 12 dB.

This formula is really quite interesting as it shows that regardless of how strongly the vacuum is squeezed, the state will only ever have even photon numbers! In hindsight, thinking about the process in terms of down-conversion and examining the Hamiltonian in Eq. 1.83 makes this obvious; in the absence of loss and other imperfections, each pump photon that is down-converted will contribute exactly two photons the field. For stronger squeezing, many pump photons can be converted simultaneously leading to pairs of photon pairs, and so on, leading to even photon numbers.

If we instead express squeezed vacuum in the quadrature basis, we can achieve a deeper understanding in a different way. Here,

$$\begin{aligned}\hat{S}(\mathfrak{r})\,|0\rangle &= \hat{S}(\mathfrak{r})\pi^{-1/4}\int dt\, e^{-t^2/2}\,|t\rangle_p \\ &= \sqrt{2}\pi^{1/4}\hat{S}(\mathfrak{r})e^{-\hat{P}^2/2}\hat{S}^\dagger(\mathfrak{r})\hat{S}(\mathfrak{r})\,|0\rangle_q \\ &= s^{1/2}\pi^{-1/4}\int dt\, e^{-\frac{s^2t^2}{2}}\,|t\rangle_p\,,\end{aligned} \tag{1.89}$$

where we let $s = e^{\mathfrak{r}}$. After the first line, we used Eq. 1.34 and then that zero-quadrature eigenstates are also eigenstates of the squeezing operator:

$$\hat{S}(\mathfrak{r})\,|0\rangle_q = e^{-\frac{i\mathfrak{r}}{2}(2\hat{P}\hat{Q}+i)}\,|0\rangle_q = e^{\frac{\mathfrak{r}}{2}}\,|0\rangle_q\,. \tag{1.90}$$

If we take the large squeezing limit where $\mathfrak{r} \to \infty$, then we have that Eq. 1.89 becomes a normalized version of $|0\rangle_p$. Thus, large squeezing is the closest physical means of

approximating the unphysical quadrature eigenstates. Later, we will see how to convert quadrature eigenstates into finitely-squeezed states using non-Hermitian operators that arise from some physical processes.

### 1.4.7 Two-mode beamsplitter

Up until now, we have only considered operators that act on a single mode. Quantum information can only take advantage of scaling when we include multi-mode operators. One important passive two mode operator is the beamsplitter.

Given annihilation operators $\hat{a}$ and $\hat{b}$ for two modes in separate Hilbert spaces, the beamsplitter operator is given by the unitary operator

$$\hat{B}_{ab} = e^{\theta(\hat{a}^\dagger \hat{b} e^{i\phi} - \hat{a}\hat{b}^\dagger e^{-i\phi})}, \tag{1.91}$$

where $\theta$ is the beamsplitter angle that determines mode coupling, and $\phi$ is the beamsplitter phase. This can also be written in terms of the quadrature operators as

$$\hat{B}_{ab} = exp\left[\theta \sinh\phi(\hat{Q}_a\hat{Q}_b + \hat{P}_a\hat{P}_b) + i\theta\cosh\phi(\hat{Q}_a\hat{P}_b - \hat{P}_a\hat{Q}_b)\right]. \tag{1.92}$$

Using the Heisenberg picture, operators $\hat{a}$ and $\hat{b}$ are transformed according to

$$\dot{\hat{a}} = \hat{b}e^{i\phi}, \quad \dot{\hat{b}} = -\hat{a}e^{-i\phi}, \tag{1.93}$$

which lead to

$$\ddot{\hat{a}} = -\hat{a}, \quad \ddot{\hat{b}} = -\hat{b}. \tag{1.94}$$

Solving the differential equations with the conditions that

$$\hat{a}(\theta = 0) = \hat{a}, \quad \hat{b}(\theta = 0) = \hat{b} \tag{1.95}$$

leads to the transformation that

$$\hat{B}^\dagger_{ab}\begin{pmatrix}\hat{a}\\ \hat{b}\end{pmatrix}\hat{B}_{ab} = \begin{pmatrix}\hat{b}e^{i\phi}\sin\theta + \hat{a}\cos\theta\\ -\hat{a}e^{-i\phi}\sin\theta + \hat{b}\cos\theta,\end{pmatrix} \tag{1.96}$$

where the creation operators transform similarly. Setting the phase $\phi = \pi$ and letting $r = \sin\theta$, $t = \cos\theta$ leads us to symplectic matrix representation of the beamsplitter

transformation given by

$$\begin{pmatrix} t & -r \\ r & t \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ta - rb \\ ra + tb \end{pmatrix}. \tag{1.97}$$

Symplectic transformations are the most general linear operations that preserve the bosonic commutator, $[\hat{a}, \hat{a}^\dagger] = \mathbb{1}$.

It is important to note that the phase $\phi$ can be changed by applying single-mode rotations to the states before and after the beamsplitter. For this reason, the beamsplitter phase can be chosen for convenience with the assumption that an optical phase can be applied for free as needed. The beamsplitter convention used throughout this text is not always fixed, but changes are noted as they occur by defining the precise beamsplitter operator used.

When writing circuit diagrams, it will be useful to represent the beamsplitter interaction between two qumodes as an arrow,

$$\theta \;\downarrow \;\begin{matrix} j \\ k \end{matrix} = \hat{B}_{jk}(\theta) = \hat{B}_{kj}(-\theta) := e^{\theta(a_j a_k^\dagger - a_j^\dagger a_k)}. \tag{1.98}$$

This will almost always be used to represent the balanced beamsplitter where $\theta = \frac{\pi}{4}$, but the value of $\theta$ will be specified in the text if this is not the case.

### 1.4.8 Two-mode squeezing

The process in Fig. 1.4 is not restricted to a single output mode as the pump photon can produce photon pairs at different frequencies, thus requiring a two-mode output state. Alternatively, the process might be degenerate in frequency, but the output photons may be crossed-polarized, such that they are still each in a separate mode. For these cases, the single-mode squeezing operator from Sec. 1.4.6 can be extended to multiple modes.

For output modes $\hat{a}$ and $\hat{b}$, the two-mode squeezing operator is given by

$$\hat{S}_{ab}(\imath) = e^{\imath(\hat{a}^\dagger \hat{b}^\dagger - \hat{a}\hat{b})} = e^{-2i\imath(\hat{Q}_a \hat{P}_b + \hat{P}_a \hat{Q}_b)}. \tag{1.99}$$

The photons are emitted in pairs as before, but now they are emitted as one in each mode. In the Fock basis, applying the two-mode squeezing operator to vacuum leads

to

$$\hat{S}_{ab}(\imath)\left|0\right\rangle_a\left|0\right\rangle_b = \frac{1}{\cosh\imath}\sum_{n=0}^{\infty}\tanh^n\imath\left|n\right\rangle_a\left|n\right\rangle_b. \tag{1.100}$$

Alternative to jointly squeezing vacuum modes directly, $\hat{S}_{ab}$ acting on vacuum can also be decomposed into single-mode squeezing and a passive beamsplitter operation. This arises from the more general formula [15]

$$\begin{aligned}&\hat{B}_{ab}(\tfrac{\pi}{4},\phi)\hat{S}_a(\imath_a)\hat{S}_b(\imath_b)\left|0\right\rangle_a\left|0\right\rangle_b\\ &=\hat{S}_a\left(\tfrac{1}{2}(\imath_a+\imath_b e^{2i\phi})\right)\hat{S}_b\left(\tfrac{1}{2}(\imath_a e^{-2i\phi}+\imath_b)\right)\hat{S}_{ab}\left(\tfrac{1}{2}(\imath_a e^{i\phi}-\imath_b e^{-i\phi})\right)\left|0\right\rangle_a\left|0\right\rangle_b,\end{aligned} \tag{1.101}$$

which for $\imath_a=\imath_b=\imath$ and $\phi=\frac{\pi}{2}$ implies that

$$\hat{B}_{ab}(\tfrac{\pi}{4},\tfrac{\pi}{2})\hat{S}_a(\imath)\hat{S}_b(\imath)\left|0\right\rangle_a\left|0\right\rangle_b = \hat{S}_{ab}(\imath)\left|0\right\rangle_a\left|0\right\rangle_b. \tag{1.102}$$

Note that the squeezing here is preserved before and after the beamsplitter. Before, each of the single modes was squeezed by $\imath$ such that the two squeezed quadratures are $\hat{S}_a^\dagger\hat{P}_a\hat{S}_a = \hat{P}_a e^{-\imath}$ and $\hat{S}_b^\dagger\hat{P}_b\hat{S}_b = \hat{P}_b e^{-\imath}$. After the beamsplitter, there are still two squeezed and two anti-squeezed quadratures, but these are now linear combinations of the individual field quadratures. After the beamsplitter, we have that

$$\hat{S}_{ab}^\dagger(\imath)(\hat{Q}_a-\hat{Q}_b)\hat{S}_{ab}(\imath) = (\hat{Q}_a-\hat{Q}_b)e^{-\imath} \tag{1.103}$$

$$\hat{S}_{ab}^\dagger(\imath)(\hat{P}_a+\hat{P}_b)\hat{S}_{ab}(\imath) = (\hat{P}_a+\hat{P}_b)e^{-\imath}. \tag{1.104}$$

The two quadratures that are now squeezed are the phase-sum and amplitude-difference quadratures, $\hat{P}_+=(\hat{P}_a+\hat{P}_b)$ and $\hat{Q}_-=(\hat{Q}_a-\hat{Q}_b)$. Similarly, the conjugate quadratures of $\hat{P}_-$ and $\hat{Q}_+$ are anti-squeezed. Because of this, squeezing can be seen to be an irreducible resource that passive transformations such as rotations and beamsplitters just redistribute [16]. All signs in Eq. 1.103 can be flipped through choice of beamsplitter phase, which amounts to a rotation of the squeezed quadrature.

It is also helpful to have the explicit forms for how each quadrature rotates individually[6]. The individual quadratures transform according to a hyperbolic rotation:

$$\begin{pmatrix} \cosh(r) & 0 & \sinh(r) & 0 \\ 0 & \cosh(r) & 0 & -\sinh(r) \\ \sinh(r) & 0 & \cosh(r) & 0 \\ 0 & -\sinh(r) & 0 & \cosh(r) \end{pmatrix} \begin{pmatrix} \hat{Q}_a \\ \hat{P}_a \\ \hat{Q}_b \\ \hat{P}_b \end{pmatrix} = \begin{pmatrix} \hat{Q}_a \cosh(r) + \hat{Q}_b \sinh(r) \\ \hat{P}_a \cosh(r) - \hat{P}_b \sinh(r) \\ \hat{Q}_b \cosh(r) + \hat{Q}_a \sinh(r) \\ \hat{P}_b \cosh(r) - \hat{P}_a \sinh(r) \end{pmatrix}. \tag{1.105}$$

### 1.4.9 Controlled gates

The two-mode squeezing operation can be related to other mathematical operations that are very useful in quantum computation. When considering qubits, one often defines controlled gates that act to perform an operation on one qubit depending on the state of the other. This will be discussed more in Sec. 1.6, but it is useful to consider CV gates that act to perform quadrature shifts, or displacements, on one mode depending on the state of the other. The naming convention will become more clear in future sections, but for now, let's define the CV gates

$$\hat{C}_X(g) \equiv e^{-ig\hat{Q}_1\hat{P}_2} \tag{1.106}$$

$$\hat{C}_Z(g) \equiv e^{ig\hat{Q}_1\hat{Q}_2}. \tag{1.107}$$

When applied to states, the second qumode is shifted by the scaled eigenvalue of the first according to

$$\hat{C}_X(g)|s\rangle_{q_1}|t\rangle_{q_2} = e^{-igs\hat{P}_2}|s\rangle_{q_1}|t\rangle_{q_2} = |s\rangle_{q_1}|t+gs\rangle_{q_2} \tag{1.108}$$

$$\hat{C}_Z(g)|s\rangle_{q_1}|t\rangle_{p_2} = e^{igs\hat{Q}_2}|s\rangle_{q_1}|t\rangle_{p_2} = |s\rangle_{q_1}|t+gs\rangle_{p_2}. \tag{1.109}$$

[6]These equations come about as an intermediary step to deriving Eq. 1.103 using the Heisenberg picture.

Note that the second qumode in the second line is a momentum ket. We can also use the BCH formula to see how the quadrature operators transform, where we have

$$\hat{C}_X^\dagger(g)\hat{P}_1\hat{C}_X(g) = \hat{P}_1 - g\hat{P}_2, \tag{1.110}$$
$$\hat{C}_X^\dagger(g)\hat{Q}_2\hat{C}_X(g) = \hat{Q}_2 + g\hat{Q}_1, \tag{1.111}$$
$$\hat{C}_Z^\dagger(g)\hat{P}_1\hat{C}_Z(g) = \hat{P}_1 + g\hat{Q}_2, \tag{1.112}$$
$$\hat{C}_Z^\dagger(g)\hat{P}_2\hat{C}_Z(g) = \hat{P}_2 + g\hat{Q}_1. \tag{1.113}$$

The operators $\hat{Q}_1$ and $\hat{P}_2$ both commute with $\hat{C}_X$ and are thus unaffected by the action of the gate. Similarly, $\hat{C}_Z$ has no effect on $\hat{Q}_1$ and $\hat{Q}_2$. For both gates, the value of the coefficient will generally be taken to be $g = 1$ as it can be tuned through a single-mode squeezing rotation on either or both modes. When the gate is written denoted without the argument simply as $\hat{C}_X$ or $\hat{C}_Z$, it can be assumed that $g = 1$.

### 1.4.10 EPR states

The canonical EPR state is a two-mode state such that one particle, or mode, is perfectly correlated in position or momentum with the other [17]. Defined in the $q$-basis, we have

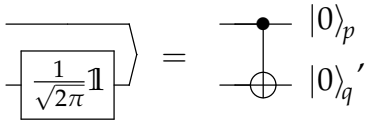
$$|\text{EPR}\rangle_{12} = \frac{1}{\sqrt{2\pi}}\int dt\, |t\rangle_{q_1}|t\rangle_{q_2} = \frac{1}{\sqrt{2\pi}}\int dt\, |t\rangle_{p_1}|-t\rangle_{p_2}, \tag{1.114}$$

which is perfectly correlated in position and perfectly anti-correlated in momentum. Supposed one starts with two zero-quadrature eigenstates, one $|0\rangle_p$ and one $|0\rangle_q$. These can be transformed into an EPR state by applying the $\hat{C}_X$ gate defined in the previous section. We have

$$\begin{aligned}\hat{C}_X|0\rangle_{p_1}|0\rangle_{q_2} &= \tfrac{e^{-i\hat{Q}_1\hat{P}_2}}{\sqrt{2\pi}}\int dt\, |t\rangle_{q_1}|0\rangle_{q_2} \\ &= \tfrac{1}{\sqrt{2\pi}}\int dt\, |t\rangle_{q_1}|t\rangle_{q_2}.\end{aligned} \tag{1.115}$$

This can be written as the circuit

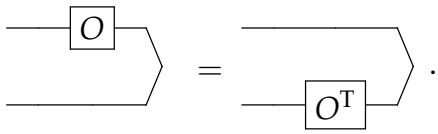
$|\text{EPR}\rangle$ :=




, (1.116)

where the connection with an open circle represents the $\hat{C}_X$ gate, the boxes represent single-mode operators, and the ended circuit represents an EPR pair. Here and everywhere else I show circuit diagrams[7], I adopt the convention that the circuit proceeds from right to left. This is so that when the circuit is written out mathematically with operators, it can be written down and read left to right.

One particular use of a diagrammatic approach is that operators can be easily transferred, or 'bounced', from one mode to the other across an EPR pair [18]. This bouncing can be represented as




(1.117)

In the above, the closed off portion on the right of each diagram represents an incoming EPR pair. The circuit on the left-hand side represents applying operator $\hat{O}$ to mode one of the EPR state, and the right-hand circuit, which is equivalent, represents taking an EPR state and applying the transpose of the operator, $\hat{O}^T$, to mode two. It is important to be aware that the transposed operator is not the same as the daggered operator. Here, the transpose is taken in the Q-basis, so that $\hat{Q}^T = \hat{Q}$ while $\hat{P}^T = -\hat{P}$.

To see how these relations come about, first consider a single-mode operator that is a function of $\hat{Q}$ only, $O(\hat{Q})$, applied to an EPR state. We have

$$\begin{aligned} O(\hat{Q}_1)\,|EPR\rangle_{12} &= \frac{1}{\sqrt{2\pi}}\int dt\, O(t)\,|t\rangle_{q_1}\,|t\rangle_{q_2} \\ &= \frac{1}{\sqrt{2\pi}}\int dt\,|t\rangle_{q_1}\,O(\hat{Q}_2)\,|t\rangle_{q_2} \\ &= O(\hat{Q}_2)\,|EPR\rangle_{12}\,. \end{aligned} \tag{1.118}$$

[7] Other than one or two exceptions, namely Fig. 1.8.

Similarly, for an operator that is a function of $\hat{P}$ only, we have that

$$\begin{aligned} O(\hat{P}_1)\,|EPR\rangle_{12} &= \frac{1}{\sqrt{2\pi}}\int dt\, O(t)\,|t\rangle_{p_1}\,|-t\rangle_{p_2} \\ &= \frac{1}{\sqrt{2\pi}}\int dt\,|t\rangle_{q_1}\,O(-\hat{P}_2)\,|-t\rangle_{q_2} \\ &= O(-\hat{P}_2)\,|EPR\rangle_{12}\,. \end{aligned} \tag{1.119}$$

We have already discussed how quadrature eigenstates are unphysical. However in the large squeezing limit, two-mode squeezed vacuum approaches the EPR state. This arises since

$$\lim_{\varepsilon\to\infty}\hat{S}_{12}(\varepsilon)|0\rangle_{N_1}|0\rangle_{N_2} \to \sum_{n=0}^{\infty}|n\rangle_{N_1}|n\rangle_{N_2} = \int dt\,|t\rangle_{q_1}|t\rangle_{q_2}. \tag{1.120}$$

## 1.5 Kraus Operator Formalism and the Lindblad equation

Up to this point, we have only considered evolving states and operators using the methods in Sec. 1.4; however, a more general formalism is required.

**An example: photon loss**

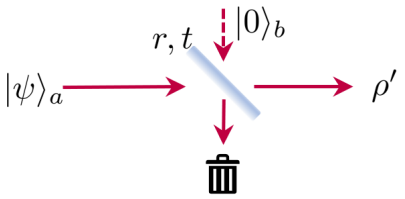



FIGURE 1.6: Photon loss can be modeled as a quantum state interacting with the vacuum at a beamsplitter, where a partial trace is performed over the reflected mode. The transmitted mode is now a mixture and must necessarily be represented as a density operator.

As an example, consider the process of photon loss. This can be modeled as shown in Fig. 1.6, where a quantum state interferes with the vacuum at an entangling beamsplitter and the second of the two modes is 'thrown away'. This would be achieved by performing a partial trace over subsystem $b$, leaving us with a statistical mixture. In

the specific case where mode $a$ is the pure state $|\psi\rangle = \sum_n \psi_n |n\rangle$ and mode $b$ is vacuum, $|0\rangle$, our loss-degraded density operator would be [19]

$$\begin{aligned}
\rho' &= \mathrm{Tr}_b \left[ \hat{B} |\psi\rangle_a \langle\psi| \otimes |0\rangle_b \langle 0| \, \hat{B}^\dagger \right] \\
&= \mathrm{Tr}_b \left[ \sum_{n,m} \tfrac{\psi_n \psi_m^*}{\sqrt{n!m!}} (t\hat{a}^\dagger - r\hat{b}^\dagger)^n |00\rangle_{ab} \langle 00| \, (t\hat{a} - r\hat{b})^m \right] \\
&= \mathrm{Tr}_b \left[ \sum_{n,m}^{\infty} t^{n+m} \psi_n \psi_m^* \sum_{k,k'}^{n,m} \binom{n}{k}^{\frac{1}{2}} \binom{m}{k'}^{\frac{1}{2}} \left( \frac{-r}{t} \right)^{k+k'} |n-k\rangle_a \langle m-k'| \otimes |k\rangle_b \langle k'| \right] \\
&= \sum_{k}^{\infty} \sum_{\substack{n=k \\ m=k}}^{\infty} t^{n+m} \psi_n \psi_m^* \binom{n}{k}^{\frac{1}{2}} \binom{m}{k}^{\frac{1}{2}} \left( \frac{r}{t} \right)^{2k} |n-k\rangle_a \langle m-k| \, . \qquad (1.121)
\end{aligned}$$

We have managed to write an expression for our state, albeit a cumbersome one, but it's there and it makes sense considering what we might expect of loss. If there is no loss ($r = 0$) then the state remains pure and $\rho' = \rho = |\psi\rangle \langle\psi|$. As $r$ increases, there is a higher weighting on the components of the density matrix where the photon number inside the ket has been decreased, which corresponds to lost photons.

Instead of having to calculate the state evolution through the beamsplitter, involve the ancillary vacuum mode, and perform the partial trace, it would be nice if we had an alternative means of applying loss directly as an operator, such as in the form

$$\rho' = \sum_{k=0}^{\infty} \hat{K}_k \rho \hat{K}_k^\dagger . \qquad (1.122)$$

Here, the operators $\hat{K}_k$ are known as Kraus operators. I will circle back around to the specific problem of loss, but the desire to write the loss process in the form of Eq. 1.122 brings us to a wider problem: How do we evolve states that potentially interact with external systems that may then involve non-unitary evolution such as detection or a partial trace? This is where we can rely on the Kraus operator formalism and the Lindblad master equation.

### 1.5.1 Lindblad master equation

The Lindblad equation describes the general evolution of a density operator that is subjected to open and potentially non-unitary evolution with the environment. Phrased

differently, it describes the evolution of a subsystem of a larger, more complicated system. An excellent introduction and pedagogical derivation of the Lindblad equation, density matrix evolution, and Kraus operators can be found in Ref. [20]. The question of evolving an arbitrary subsystem $\rho$ in Hilbert space $\mathcal{H}$ of dimension $N$ through general physical processes amounts to determining maps $\mathcal{V}$ that transform $\rho$ into other physical density matrices existing in Hilbert space $\mathcal{H}$. More formally, if $\rho(\mathcal{H})$ is the space of all density matrices within Hilbert space $\mathcal{H}$, we wish to find maps of the form $\mathcal{V} : \rho(\mathcal{H}) \to \rho(\mathcal{H})$. These operators necessarily must:

- Preserve the trace; $\mathrm{Tr}[\mathcal{V}A] = \mathrm{Tr}[A]$ for all valid operators $\hat{A}$ acting in $\mathcal{H}$
- Be a completely positive map.

With these conditions, $\mathcal{V}$ is generally known as a completely-positive trace-preserving (CPT) map. The first condition is fairly self-explanatory, as all valid density matrices must have trace 1. The second is a bit more subtle. We need that $\mathcal{V}$ acting in $\mathcal{H}$ must map positive, semi-definite matrices to other positive semi-definite matrices, but also that $\mathcal{V} \otimes \mathbb{1}^n$ is a positive map for all integers $n$, where $\mathbb{1}^n$ represents the tensor product of $n$ identity operators. This last bit essentially means that if $\mathcal{V}$ is acting on subsystem $\rho_a$ of large system $\rho_a \otimes \sigma_b$, then the result after taking a partial trace must still be positive semi-definite: $\mathrm{Tr}_b[(\mathcal{V}_a \otimes \mathbb{1}_b)(\rho_a \otimes \sigma_b)] \geq 0$.

By the Choi-Kraus theorem, the map $\mathcal{V}$ is completely positive and trace-preserving if and only if it can be expressed as

$$\mathcal{V}\rho = \sum_i V_i \rho V_i^\dagger, \tag{1.123}$$

where $V_i$ satisfy the condition:

$$\sum_i V_i^\dagger V_i = \mathbb{1}. \tag{1.124}$$

The operators $V_i \equiv K_i$ are known as *Kraus operators*. Note that these operators need not be orthonormal, and furthermore, need not be unique[8]. Although we will only need the Kraus operator formalism and not the full Lindblad equation, it is worth providing the latter for completeness. In the spirit of the Schrödinger equation, the goal is to find a differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = \mathcal{L}\rho(t), \tag{1.125}$$

[8]It is possible for a physical process to be accurately described by more than one set of Kraus operators. The set used is determined by the basis of the trace-out mode.

such that we can write our time-dependent map as $\mathcal{V}(t) = e^{\mathcal{L}t}$. Choosing an orthonormal operator basis, $\{F_i\}$, spanning $\mathcal{H}$ such that $\mathrm{Tr}[F_i^\dagger F_j] = \delta_{i,j}$ and letting the last one be proportional to the identity, $F_{N^2} = \frac{1}{\sqrt{N}}\mathbb{1}$, allows the Lindblad master equation to be written as

$$\dot{\rho} = -\tfrac{i}{\hbar}[\hat{H}, \rho] + \sum_{j,k}^{N^2-1} g_{j,k}\left(F_j \rho F_k^\dagger - \tfrac{1}{2}(F_k^\dagger F_j \rho + \rho F_k^\dagger F_j)\right). \tag{1.126}$$

In the above, the Hamiltonian $\hat{H}$ and the coefficient matrix $g$ act together to define the evolution of the system. The coefficient matrix $g$ can be diagonalized by a unitary transformation, which allows one to rewrite the Lindblad equation in a diagonal form of

$$\dot{\rho} = -\tfrac{i}{\hbar}[\hat{H}, \rho] + \sum_{k}^{N^2-1} \gamma_k\left(L_k \rho L_k^\dagger - \tfrac{1}{2}(L_k^\dagger L_k \rho + \rho L_k^\dagger L_k)\right). \tag{1.127}$$

In this form, $\gamma_k$ are the eigenvalues of $g$ and $L_k$ are known as the Lindblad jump operators for the system,

$$L_k = \sum_{j=1}^{N^2-1} u_{jk} F_j, \tag{1.128}$$

where $u_{jk}$ are the elements of the unitary matrix that diagonalizes $g$.

One example for the use of the Lindblad equation is cavity dynamics. Imagine a cavity field with the only interaction being to couple to a zero-temperature bath, meaning that occasionally the cavity field may leak photons into the environment of vacuum modes, but no additional photons will get back in. For a decay rate of $\kappa$ and in the rotating frame of the cavity (such that $[\hat{H}, \rho] = 0$), the Lindblad equation for this system is [21]

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho = \kappa\left(\hat{a}\rho\hat{a}^\dagger - \tfrac{1}{2}(\hat{a}^\dagger\hat{a}\rho + \rho\hat{a}^\dagger\hat{a})\right). \tag{1.129}$$

The term in the inner parentheses on the right is the unitary portion that describes the evolution of $\rho$ by the free-field Hamiltonian. The first term after the equal sign, however, describes the non-unitary interaction with the environment where there is a potential for photons to be lost, as $\hat{a}\rho\hat{a}^\dagger$ reduces the photon number of $\rho$ by one. This equation is exactly related to the effects of loss that will be described next.

### 1.5.2 Kraus operator for loss

Let us return to the problem of loss discussed at the beginning of this section, and see if the Kraus operator formalism can make things a bit easier. The goal is to take the density operator we ended up with in the last line in Eq. 1.121 and turn in into some sum of operators acting on the initial density matrix. First, we can make use of the identity

$$\hat{a}^k |n\rangle = \sqrt{\tfrac{n!}{(n-k)!}} |n-k\rangle \tag{1.130}$$

to write that the loss-degraded density operator is given by

$$\rho' = \sum_{k=0}^{\infty} \sum_{\substack{n=k \\ m=k}}^{\infty} \frac{t^{n+m}}{k!} \psi_n \psi_m^* \left(\frac{r}{t}\right)^{2k} \hat{a}^k |n\rangle\langle m| \hat{a}^{\dagger k}. \tag{1.131}$$

Next, we can reset the lower limits of the second sum to $n = 0$ and $m = 0$ since $\hat{a}^k |n\rangle = 0$ when $k > n$. Additionally, we can use an eigenvalue trick to promote the coefficient $t^n$ to an operator,

$$t^n |n\rangle = t^{\hat{n}} |n\rangle = t^{\hat{a}^\dagger \hat{a}} |n\rangle , \tag{1.132}$$

which allows us to pull this term outside of the sum, but only *after* the factor of $\hat{a}^k$ since $\hat{n}$ and $\hat{a}$ do not commute. This gives

$$\rho' = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{r}{t}\right)^{2k} \hat{a}^k t^{\hat{a}^\dagger \hat{a}} \left[ \sum_{\substack{n=0 \\ m=0}}^{\infty} \psi_n \psi_m^* |n\rangle \langle m| \right] t^{\hat{a}^\dagger \hat{a}} \hat{a}^{\dagger k}, \tag{1.133}$$

which is exactly the form of Eq. 1.122 we wished to find where $\hat{K}_k$ are the Kraus operators associated with loss given by

$$\hat{K}_k = \sqrt{\frac{(1-\eta)^k}{k! \eta^k}} \hat{a}^k e^{\frac{1}{2} \hat{a}^\dagger \hat{a} \ln \eta}, \tag{1.134}$$

and I have let the overall loss be $\eta = t^2$ and rewritten the $t^{\hat{a}^\dagger \hat{a}}$ term as an exponential. This derivation started with a pure state, but we could have just as easily begun with an arbitrary density matrix and arrived at the same result. Thus, these Kraus operators are general and Eqs. 1.133 & 1.134 completely describe single-mode loss. Multi-mode loss is a simple extension of this theory, where one sums over tensor products of the Kraus operators in Eq. 1.134 applied to each mode of the state. Next, I will discuss properties

and uses of a specific operator that showed up here, $e^{-\beta\hat{a}^\dagger\hat{a}}$, where $\beta = -\frac{1}{2}\ln\eta$, which is otherwise known as the damping operator.

### 1.5.3 The damping operator

The damping operator is a non-unitary operator that acts to symmetrically de-amplify quantum states and can be defined as

$$\hat{N}(\beta) := e^{-\beta\hat{a}^\dagger\hat{a}} = e^{-\frac{\beta}{2}(\hat{Q}^2+\hat{P}^2-1)}. \tag{1.135}$$

When applied to a quadrature eigenstate, the damping operator brings the unphysical state to a displaced finitely squeezed vacuum state. This can be seen by first applying $\hat{N}$ to the zero-valued quadrature eigenstate $|0\rangle_p$. Writing $|0\rangle_p$ in the Fock-basis in the infinite squeezing limit, we have

$$|0\rangle_p = \lim_{\imath\to\infty} \frac{1}{\sqrt{\cosh\imath}} \sum_{n=0}^{\infty} (\tanh\imath)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle_N \tag{1.136}$$

Applying the damping operator to the above state yields

$$\hat{N}(\beta)|0\rangle_p = \lim_{\imath\to\infty} \frac{1}{\sqrt{\cosh\imath}} \sum_{n=0}^{\infty} (e^{-2\beta}\tanh\imath)^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle_N. \tag{1.137}$$

Because $\hat{N}(\beta)$ is non-unitary, this state must now be normalized. Additionally, by replacing $e^{-2\beta}\tanh\imath$ with $\tanh\imath'$, we have that

$$\hat{N}(\beta)|0\rangle_p = n \sum_{n=0}^{\infty} (\tanh\imath')^n \frac{\sqrt{(2n)!}}{2^n n!} |2n\rangle_N, \tag{1.138}$$

which is just a finitely squeezed state of squeezing parameter $\imath' = \tanh^{-1}\left[e^{-2\beta}\right]$ in the limit $\imath \to \infty$ with normalization given by

$$n = \sqrt{\frac{1}{\cosh\imath'}}. \tag{1.139}$$

Using these results, we can also examine the effects of applying the damping operator to a non-zero quadrature eigenstate, $|m\rangle_p$, in which case we have

$$\begin{aligned}\hat{N}(\beta)|m\rangle_p &= \hat{N}(\beta)\hat{Z}(m)\hat{N}(-\beta)\hat{N}(\beta)|0\rangle_p \\ &= \hat{N}(\beta)\hat{Z}(m)\hat{N}(-\beta)\hat{S}(\imath')|0\rangle_N,\end{aligned} \tag{1.140}$$

since the inverse of $\hat{N}(\beta)$ is $\hat{N}(-\beta)$. Using commutation relations, we can derive that

$$\hat{N}(\beta)\hat{a}\hat{N}(-\beta) = \hat{a}e^{\beta} \tag{1.141}$$

$$\hat{N}(\beta)\hat{a}^\dagger\hat{N}(-\beta) = \hat{a}^\dagger e^{-\beta} \tag{1.142}$$

which leads to

$$\hat{N}(\beta)\hat{Q}\hat{N}(-\beta) = \hat{Q}\cosh\beta + i\hat{P}\sinh\beta. \tag{1.143}$$

The above equation is useful to determine the transformation of the momentum-shift operator, which will be used later in Ch. 3, to be

$$\hat{N}(\beta)\hat{Z}(m)\hat{N}(-\beta) = e^{im(\hat{Q}\cosh\beta + i\hat{P}\sinh\beta)} \tag{1.144}$$

$$= e^{\frac{m^2}{2}\sinh\beta\cosh\beta}e^{-m(\hat{P}\sinh\beta)}\hat{Z}(m\cosh\beta). \tag{1.145}$$

Putting everything together, we now have that

$$\hat{N}(\beta)|m\rangle_p = \mathcal{A}_\beta e^{-m(\hat{P}\sinh\beta)}\hat{Z}(m\cosh\beta)\hat{S}(\imath')|0\rangle_N, \tag{1.146}$$

as announced, where

$$\mathcal{A}_\beta = e^{\frac{m^2}{2}\sinh\beta\cosh\beta}\eta^{-1}. \tag{1.147}$$

### 1.5.4 Projective measurement

In Eq. 1.133, we have a nice example where we were able to explicitly derive the operators that comprise our CPT map from a channel having known loss. But one might ask, what happens if we don't perform a partial trace, but instead, place a detector in that channel to perform a projection in the Fock basis?

This process is otherwise commonly known as photon subtraction, which is when a beamsplitter is used to direct a small portion of light to a detector where photons are 'subtracted' from the main beam. It turns out that the Kraus operator describing the

process will look much like the loss operators in Eq. 1.134, but the new density matrix will no longer be a sum of these operators applied to the initial state, but rather just a single one of them. This will in actuality not preserve the trace, as the trace of the new density matrix will instead reflect the probability for the specified detection event to occur. This is due to the fact that we are indeed post-selecting on a specific event, thus all detections other than the specified one are discarded.

### Photon subtraction

Direct application of the annihilation operator in the optical domain remains a challenge, but we have the capability to probabilistically apply $\hat{a}^n$ followed by unavoidable damping of both quadratures due to the physical implementation involving a beamsplitter. This method is shown in Fig. 1.7 and is executed by a beamsplitter and a photon-number resolved detector (PNRD), where a small amount of one mode is coupled to vacuum and sent to the PNRD. Detecting $n$ photons leads to the approximate application of $\hat{a}^n$ when beamsplitter reflectivity is low, but we will derive the Kraus operator representation of this system in general.

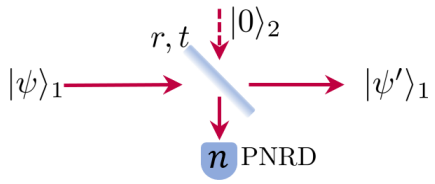



FIGURE 1.7: Experimental photon subtraction, where a weakly reflective beamsplitter is used to direct a portion of the light to a photon counting detector. The transmitted beam is then the initial state with the number of detected photons having been removed, or 'subtracted'.

For a beamsplitter operator given by [19]

$$\hat{B} = e^{\theta(\hat{a}_1^\dagger \hat{a}_2 - \hat{a}_1 \hat{a}_2^\dagger)}, \tag{1.148}$$

we can use Sec. 1.4.7 to see that the annihilation and creation operators are transformed to

$$\hat{B}^\dagger \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} \hat{B} = \begin{pmatrix} t & r \\ -r & t \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \end{pmatrix} = \begin{pmatrix} t\hat{a}_1 + r\hat{a}_2 \\ t\hat{a}_2 - r\hat{a}_1 \end{pmatrix}, \tag{1.149}$$

where $r = \sin\theta$ and $t = \cos\theta$. If we write $|\psi\rangle = \sum_m \psi_m |m\rangle$, then coupling $|\psi\rangle$ to vacuum with $\hat{B}$ and detecting $n$ photons in one output leads to the final output

$$\begin{aligned}
{}_2\langle n|\psi'\rangle_{12} &= {}_2\langle n| \sum_{m=0} \frac{\psi_m}{\sqrt{m!}} \hat{B}\hat{a}_1^{\dagger m}|0\rangle_1|0\rangle_2 \\
&= {}_2\langle n| \sum_{m=0} \frac{\psi_m}{\sqrt{m!}} \sum_{k=0}^{m} \binom{m}{k} (t\hat{a}_1^\dagger)^{m-k}(-r\hat{a}_2^\dagger)^k|0\rangle_1|0\rangle_2 \\
&= \frac{1}{\sqrt{n!}}\left(\frac{-r}{t}\right)^n \sum_{m=0} t^m \psi_m \hat{a}_1^n |m\rangle_1
\end{aligned} \tag{1.150}$$

Recognizing that $t^m |m\rangle$ can be rewritten as an exponentiated number operator acting on a Fock state, we can replace $t^m$ with a damping operator having argument $(-\ln t)$ and arrive at

$${}_2\langle n|\psi'\rangle_{12} = \frac{1}{\sqrt{n!}}\left(\frac{-r}{t}\right)^n \hat{a}^n \hat{N}(\beta)\,|\psi\rangle\,, \tag{1.151}$$

where $\beta = -\ln t$. Using the commutation relation from Eq. 1.141 allows us to write the Kraus operator representing photon-subtraction for any beamsplitter reflectivity as

$$\hat{\mathcal{S}}_n = \frac{(-1)^n e^{-n\beta/2}}{\sqrt{n!}} (2\sinh\beta)^{n/2}\, \hat{N}(\beta)\hat{a}^n. \tag{1.152}$$

The probability of successfully subtracting $n$ photons from an input density matrix, $\rho$, is given by

$$P(n) = \mathrm{Tr}\left[\hat{\mathcal{S}}_n^\dagger \hat{\mathcal{S}}_n \rho\right], \tag{1.153}$$

and the new subtracted density matrix becomes

$$\rho' = \frac{\hat{\mathcal{S}}_n \rho \hat{\mathcal{S}}_n^\dagger}{P(n)}. \tag{1.154}$$

As a quick check, we can verify the results of the limiting cases. When $t \to 1$, $\beta \to 0$, so the damping term becomes identity and we would have perfect application of the annihilation operator. Of course, however, this limit also has $P(n) \to 0\,\forall n \neq 0$, so no photons will ever reach the PNRD to be subtracted. In the other limit where $t \to 0$, we have that $\beta \to \infty$, so the output is 'infinitely damped', which means the output mode is only vacuum.

**Photon addition**

A similar process can be used to effectively enact a photon addition. Rather than leaving the second beamsplitter port empty, one can instead input a single photon, $|1\rangle$, and post-select on an event where the detector didn't record any photons which amount to a projection on the vacuum. The derivation follows similarly to the above and results in the photon-addition Kraus operator of

$$\hat{\mathcal{A}} = \sqrt{\frac{\sinh \beta}{2}} e^{\beta/2} e^{-\beta \hat{a}^\dagger \hat{a}} \hat{a}^\dagger, \tag{1.155}$$

where again $t = e^{-\beta} = \cos\theta$ is the beamsplitter coefficient.

## 1.6 Quantum Computation

Quantum computation is often considered the 'Holy Grail' of quantum information sciences and has the potential to outclass classical computing on certain problems. First proposed forty years ago by Feynman as a possible method of simulating other quantum systems [22], the field has made substantial progress in a relatively short period of time. Once Peter Shor discovered a quantum algorithm that could factor numbers exponentially faster than a classical computer and destroy RSA security encryption [23], the field understandably gained attention far beyond the initial scientific curiosity. Today, quantum computers have demonstrated an ability to first compete with [24] and now outclass the world's largest supercomputers at a few select problems with both superconducting transmon qubits and photons [25–27]. These quantum supremacy tasks are generally designed favorably for the quantum computer and involved sampling a random quantum process that is difficult for a classical computer to simulate. It is now nearly a certainty that real quantum computers will eventually find a niche in which they either assist or outperform classical computers for certain tasks, but precisely which technology will prevail, to what extent quantum will overtake classical, and exactly when this will happen is all up in the air [28, 29].

Scott Aaronson, who maintains an informative — and often entertaining — blog on quantum computing, recently attempted to put into words why explaining the origin of a quantum advantage in simple terms is challenging [30]. Without mathematics, the best that words can generally do is to say that quantum mechanics allows access to

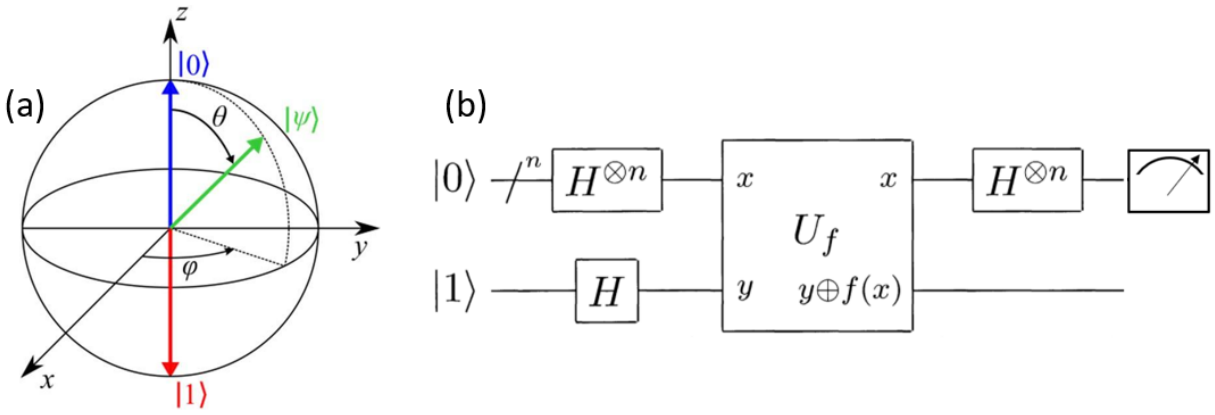



FIGURE 1.8: (a) The Bloch sphere, where the state of a qubit is represented by a vector lying on the surface. (b) Deutsch-Jozsa algorithm circuit taken from Ref. [5] where f(x) is determined to be either constant or balanced. The '/$^n$' in the top wire of the circuit represents a register of $n$ qubits.

a larger (better scaling than classical) vector space (the Hilbert space) that can be used to perform calculations through the use of extended superposition across constituent particles (entanglement) and that the nature of quantum interference of probability amplitudes from controllable interactions allows the cancellation of many possibilities, leading to a high-likelihood to measure the result in a state that answers a problem at hand. Ignoring the experimental challenges, theoretical difficulties involve attempting to coerce the 'wrong' probability amplitudes to destructively interfere without already knowing the result in advance. Additionally, even proving that a particular problem is hard for a classical machine is difficult, and the best one can often do is to demonstrate that a particular quantum algorithm has a speed-up over the best known classical algorithm. The Quantum Algorithm Zoo[9] is an online repository that attempts to catalog all currently known quantum algorithms with potential speed-ups, including the wider known algorithms such as factoring, sorting, and optimization.

### 1.6.1 Qubits

In most cases, introductions to quantum computing begin with the Bloch sphere, depicted in Fig. 1.8. Here, we start with the basic building block of the device, the qubit, which when measured takes on a binary value of 0 or 1. One could equally well use a *qudit* that could take on several discrete values of 0, 1, 2, ... etc., but partly due to the

[9] https://quantumalgorithmzoo.org/

simplicity and partly due to the propensity for classical computing to use binary bits, the majority of quantum implementations focus on *qubits*. As a quantum object, the state of a single qubit can be represented by a superposition of basis vectors $|0\rangle$ and $|1\rangle$. An arbitrary superposition for a pure state of this form can be described by

$$|\psi\rangle = a\,|0\rangle + b\,|1\rangle\,, \tag{1.156}$$

where $a$ and $b$ are complex numbers. By enforcing normalization such that $|a|^2 + |b|^2 = 1$, which is equivalent to saying that when measured $|\psi\rangle$ must collapse into either $|0\rangle$ with probability $|a|^2$ or $|1\rangle$ with probability $|b|^2$, and by ignoring an overall global phase, the two complex parameters can be reduced to two continuous real parameters. These real parameters, $\theta$ and $\varphi$, give the location of a vector on the surface of the Bloch sphere (for a pure state). The most general representation of a qubit then, up to a global phase, is given by

$$|\psi\rangle = \cos(\theta/2)\,|0\rangle + \sin(\theta/2)e^{i\varphi}\,|1\rangle\,, \tag{1.157}$$

where $0 \leq \theta \leq \pi$ and $0 \leq \varphi < 2\pi$ are the two continuous parameters that describe the state. This state is normalized since $\cos^2(\theta/2) + \sin^2(\theta/2) = 1$, and the probability amplitudes are allowed to take on complex values. When working with discrete quantum information in particular, it is often helpful to use canonical representation where

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{1.158}$$

This allows for an intuitive understanding of quantum gates as unitary matrices that map one state vector to a next so that an arbitrary single-qubit unitary gate can be represented as

$$U = \begin{pmatrix} u_{00} & u_{01} \\ u_{10} & u_{11} \end{pmatrix} \tag{1.159}$$

Some of the most fundamental quantum gates are the Pauli gates, $(X, Y, Z)$, which give rise to general rotations around the respective $\hat{x}$, $\hat{y}$, and $\hat{z}$ axes of the Bloch sphere

when exponentiated, $e^{-\frac{i\theta}{2}\sigma_k}$ where $\sigma_k \in \{X, Y, Z\}$. Explicitly, these matrices are:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{1.160}$$

where their commutation relation leads to $\sigma_j\sigma_k = \delta_{jk}\mathbb{1} + i\epsilon_{jkl}\sigma_l$. This states that a product of any two different Pauli matrices yields the third up to a phase, and a product of two identical Pauli matrices is just the identity. From this, one can collapse any arbitrarily long series of Pauli gates into a single Pauli gate multiplied by a phase.

Yet another oft-discussed gate is the Hadamard gate,

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \tag{1.161}$$

which allows one to convert to the plus/minus basis with $H|0\rangle = |+\rangle$ and $H|1\rangle = |-\rangle$, where

$$\begin{aligned} |+\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \\ |-\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \end{aligned} \tag{1.162}$$

The Hadamard gate can be visualized in terms of the Bloch sphere as a $\pi/2$ rotation about the $\hat{y}$-axis followed by a reflection through the $\hat{x} - \hat{y}$ plane [5].

The Pauli and Hadamard gates are very useful and allow rotations of $\pi/2$, but arbitrary rotations are needed to rotate a state vector from one location on the Bloch sphere to any other. Interestingly, an arbitrary single-qubit unitary operation can be decomposed into just three operations and only four real parameters [5],

$$U = e^{i\alpha} \begin{pmatrix} e^{-i\beta/2} & 0 \\ 0 & e^{i\beta/2} \end{pmatrix} \begin{pmatrix} \cos(\gamma/2) & -\sin(\gamma/2) \\ \sin(\gamma/2) & \cos(\gamma/2) \end{pmatrix} \begin{pmatrix} e^{-i\delta/2} & 0 \\ 0 & e^{i\delta/2} \end{pmatrix}, \tag{1.163}$$

where $\alpha$, $\beta$, $\gamma$, and $\delta$ are real-valued. It is not always practical to experimentally enact arbitrary rotation gates, so yet another way to achieve an arbitrary single-qubit $U$ is by combining the Pauli and Hadamard gates with the ability to apply the $T$ gate,

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}, \tag{1.164}$$

which is a $\frac{\pi}{4}$ rotation about $\hat{z}$. Thus far we have really only been talking about a single qubit, but the real magic of quantum computation comes from considering multiple qubits. Unfortunately, while the Bloch sphere is useful for visualizing a single qubit with two basis vectors, multiple qubit system exist in a higher-dimensional space and can realistically only be described in the language of linear algebra. For an $n$ qubit system, the multi-dimensional Hilbert space is spanned by $2^n$ mutually orthogonal vectors, and the most general pure state can be written as

$$|\psi\rangle = \sum_{x=1}^{2^n} \psi_x |x\rangle , \tag{1.165}$$

where $\psi_x \in \mathbb{C}$ and each of the $2^n$ orthogonal component vectors is

$$|x\rangle = |x_1\rangle \otimes |x_2\rangle \otimes ... \otimes |x_n\rangle , \quad x_i \in \{0, 1\}. \tag{1.166}$$

Considering normalization and ignoring an overall global phase, $|\psi\rangle$ has $2^{n+1} - 2$ independent real parameters. Note that it is generally more convenient to suppress the tensor product notation when writing states comprising several qubits, so for example, one may write $|0\rangle \otimes |1\rangle \equiv |0\rangle_1 |1\rangle_2 \equiv |01\rangle_{12}$, where in this case the subscript is used to label which qubit is which.

With a multiple qubit system, we now have the option of applying gates that act on more than one qubit at a time. This is necessary to build entanglement between qubits and gain access to the entire Hilbert space. Without entangling interactions, the best we could do is have a tensor product of qubits of the form $|\psi_1\rangle \otimes |\psi_2\rangle \otimes ... |\psi_n\rangle$, and while each *individual* qubit is in a superposition, the whole tensor product only has a linear $2n$ parameters. This provides an argument why entanglement is *necessary* for an exponential quantum speed-up, but it is in no way *sufficient*. Only entanglement can give rise to states having the full $2^{n+1} - 2$ parameters as opposed to the $2n$ parameters of product states.

A particularly common class of two-qubit gates are the controlled gates which are the equivalent of 'if-then' logic statements in traditional computer science. These gates work by taking a pair of qubits and applying a single-qubit gate to the second qubit contingent on the state of the first. A general controlled-unitary gate can be written in block diagonal form as

$$C_{U_{12}} = \begin{pmatrix} \mathbb{1} & \mathbb{0} \\ \mathbb{0} & U \end{pmatrix}, \tag{1.167}$$

and alternatively as

$$C_{U_{12}} = |0\rangle_1\langle 0| \otimes \mathbb{1}_2 + |1\rangle_1\langle 1| \otimes \mathcal{U}_2. \tag{1.168}$$

Under the application of this gate, the state is left unchanged if qubit 1 is in state $|0\rangle$ and the unitary $\mathcal{U}_2$ is applied to qubit 2 if qubit 1 is in state $|1\rangle$. Things become really interesting, however, when the first qubit is in a superposition. Suppose we were to start with a tensor product of the two qubits, $|\psi\rangle_1 |\phi\rangle_2$. Applying the controlled unitary leads to, using Eq. 1.157,

$$C_{U_{12}} |\psi\rangle_1 |\phi\rangle_2 = \cos(\theta_1/2) |0\rangle_1 |\phi\rangle_2 + \sin(\theta_1/2) e^{i\varphi_1} |1\rangle_1 U_2 |\phi\rangle_2 . \tag{1.169}$$

The above state *can no longer* be written as a product state[10]! This is the mathematical definition of entanglement. Thus, through the use of multi-qubit gates, one can 'distribute' a superposition localized on a single qubit to a larger set of qubits.

The problem of universal quantum computation can be simply stated as requiring one to find a set of operations that can take any multi-qubit state and rotate it into any other multi-qubit state (of the same dimension). That is, if we have that

$$U |\psi\rangle = |\phi\rangle \tag{1.170}$$

for arbitrary $|\psi\rangle$ and $|\phi\rangle$, our task becomes one of determining a set of gates that will allow us to enact $U$. Of course, we could always imagine some physical process acting on the entire system that can be perfectly described by a Hamiltonian to enact $U$ under Schrödinger evolution, this is highly unlikely to be found for any given $U$. Thus we try to find a set of simple gates that only act on a subset of the physical qubits at any given time which can be combined to create $U$. Although it is no trivial statement, we can in fact generate $U$ simply by combining single and two-qubit gates without the need for three-or-more qubit gates. The problem of determining a universal gate set does not have a unique solution, and there is much more to be said on the topic than what is included here [5]. For now, it is sufficient to know that the $H$, $C_X$, and $T$ gates are enough to achieve universality. Interestingly, if only has access to $H$, $C_X$, and $T^2$, then this amounts to what is known as the Clifford gate set. The Clifford gates are sufficient for simulating classical computation and for enacting error-correction on qubits, but a

[10] Slight caveat here: the state can be reduced to a product state under the restricted cases where $\theta_1 = 0$, $\theta_1 = \pi$, or when $|\phi\rangle_2$ is an eigenstate of $U_2$.

quantum computer with only the Clifford gate set will never achieve a quantum speed-up. Thus much emphasis is placed on the $T$ gate, which we will later find has an analog in non-Gaussianity for continuous-variables.

### 1.6.2 Qubit cluster states

Quantum computation is generally formulated in terms of a circuit model such as the circuit depicted in Fig. 1.8(b). Here, the user begins with a register of qubits on the left-hand side as inputs, and then a series of gates[11] are applied to them followed by measurements to read out the answer on the right. This is the model of computation currently employed by machines such as used by Google [24] and IBM [31]. However, there exists an alternative, but equivalent, scheme of quantum computation where all entanglement is generated up front in the initial state. Once this state is created, no more gates need be directly applied, but instead, all qubits are individually measured in particular bases to enact the computation and read out the result. This method is known as 'one-way'[12] quantum computing [32, 33] and the initial resource state is known as a 'cluster state' [34].

An example of this is shown in Fig. 1.9. The cluster state is initialized by beginning with qubits in the $|+\rangle$ state at every node of the square-grid lattice and then interacting each of them with their four nearest neighbors using a $\hat{C}_Z$ interaction. From here, a universal gate set can be enacted simply by measuring each qubit in an appropriate basis. The whole $N$-qubit cluster state can thus be written as

$$\prod_{j|j\in n_k} \hat{C}_{Z_{jk}} \bigotimes_{k}^{N} |+\rangle_j |+\rangle_k, \tag{1.171}$$

where the set $n_k$ contains all indices for nearest-neighbor qubits to qubit $k$, excluding those qubits already connected to qubit $k$. Because the $\hat{C}_Z$ gates all commute, they can be applied in any order, or even simultaneously.

[11] Again, any multi-mode unitary can be decomposed into a series of single and two-qubit gates.

[12] The term 'one-way' can also mean 'single use' since the cluster state is used up via measurement to perform gates on the unmeasured portions.

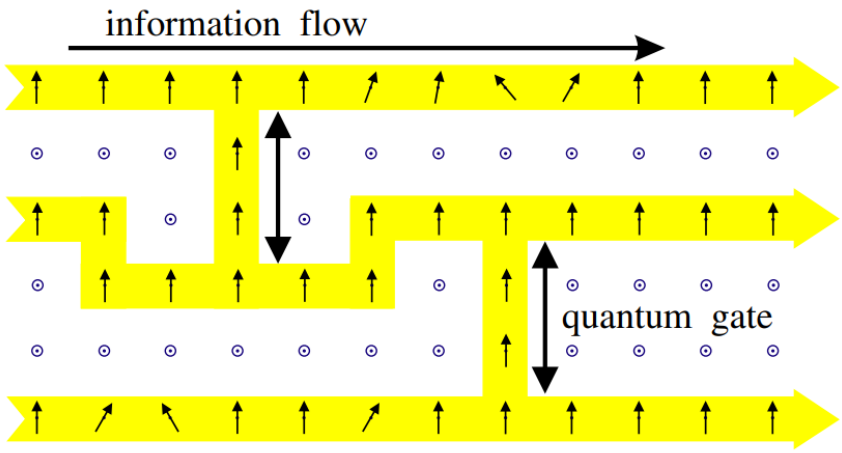



FIGURE 1.9: An example of a cluster state formed from two-state qubits taken from Ref. [33], where qubits have been initially entangled in a square-grid pattern. Then, single-qubit measurements are applied everywhere to apply gates and teleport quantum information. Circled dots indicate a $Z$-basis measurement, vertical arrows an $X$-basis measurement, and tilted arrows a measurement in the $x-y$ plane of the Bloch sphere.

The simplest example is the two-qubit cluster state:

$$\begin{aligned} |C_2\rangle_{12} &= \hat{C}_{Z_{12}} |+\rangle_1 \otimes |+\rangle_2 \\ &= \tfrac{1}{2} \left( |0\rangle_1 |0\rangle_2 + |1\rangle_0 |0\rangle_2 + |0\rangle_1 |1\rangle_2 - |1\rangle_1 |1\rangle_2 \right). \end{aligned} \tag{1.172}$$

The two and three-qubit cluster states are maximally entangled and equivalent the Bell states [35] and GHZ states [36], respectively, when local operators and classical communication (LOCC) are used to relate them. However, all cluster states of higher dimension are not LOCC-equivalent to highly entangled states such as the GHZ state family as local measurements do not have long-range correlations with all qubits in the cluster. This essentially allows qubits to be 'trimmed' without disturbing others, and for effective single-qubits gates to be enacted through measurement.

The fundamental primitive for cluster state QC is based on teleportation, where measuring one qubit teleports the information to a neighbor while also applying a local unitary operator that is determined solely by the measurement basis. Continuing the two-qubit example, suppose the first qubit was in state $|\psi\rangle = a|0\rangle + b|1\rangle$ instead of

$|+\rangle$. In this case, we have

$$\begin{aligned} |\mathcal{C}_2^{(\psi)}\rangle &= \hat{C}_{Z_{12}}|\psi\rangle_1 \otimes |+\rangle_2 \\ &= \tfrac{1}{\sqrt{2}}\left(a|0\rangle_1|0\rangle_2 + b|1\rangle_0|0\rangle_2 + a|0\rangle_1|1\rangle_2 - b|1\rangle_1|1\rangle_2\right). \end{aligned} \tag{1.173}$$

To perform a teleportation of $\psi$ from qubit one to qubit two, we first measure qubit one in the $\pm$ basis. Depending on the measurement result, this gives

$${}_1\langle m|\mathcal{C}_2^{(\psi)}\rangle_{12} = \tfrac{1}{\sqrt{2}}\left(a|+\rangle_2 + (-1)^m b|-\rangle_2\right), \tag{1.174}$$

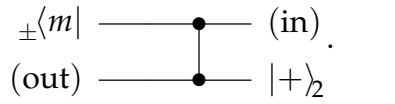
where $m$ is the measurement result in the $\pm$ basis. Thus, the initial state, $\psi$, has been teleported to mode two and both a Hadamard gate and a Z gate dependent on the measurement outcome have been applied (and can therefore be subsequently corrected for):

$${}_1\langle m|\mathcal{C}_2^{(\psi)}\rangle_{12} = \hat{H}\hat{Z}^{\frac{m-1}{2}}|\psi\rangle_2. \tag{1.175}$$

Since the first state has been teleported to qubit two, the above equation can also be written by relabeling the ket index without harm, making qubit two become qubit one (which has been 'measured out'), as

$${}_1\langle m|\mathcal{C}_2^{(\psi)}\rangle_{12} = \hat{H}\hat{Z}^{\frac{m-1}{2}}|\psi\rangle_1. \tag{1.176}$$

This process can also be written out in circuit form, which we write as proceeding from right to left:




(1.177)

In the above circuit, the input state to wire two is the qubit $|+\rangle_2$, the line connecting the wires represents a $\hat{C}_Z$ gate between the two qubits, and the top wire is measured in the $\pm$ basis. An important point to note is that this circuit specifies neither the input state nor the output state as it just leaves open spots blank for and input and output. This is to leave the form of the input completely general; there is no need to restrict the input to even a single qubit state as it could be an entangled state with many qubits, such as the remainder of the cluster state. Because this circuit only connects the two wires and measures the first, it acts locally on whichever qubit enters, regardless of any other entanglement present at the input, or even at the output that happens later. Circuits

like these are known as 'quantum gadgets' and can in fact be simplified into a single effective Kraus operator applied to the input.

The effect of this circuit has already been derived in Eq. 1.176. Thus applying this circuit to an input state can equivalently be written as applying the Kraus operator

$$\hat{K}_m = \hat{H}\hat{Z}^{\frac{m-1}{2}} \tag{1.178}$$

to the input state. By simply applying this same circuit with a different measurement basis repeatedly on the cluster state, one can perform a universal set of gates up to measurement-dependent $\hat{Z}$ and $\hat{X}$ gates. These can either be fixed by applying gates to undo the effects, or one can simply plan accordingly with the measurement basis and by shifting results. Thus the power of cluster state computation is that one only needs to perform single-qubit measurements.

While this method seems nice, the unfortunate bit is that applying gates through measurement consumes qubits. This means that cluster-state computing requires an overhead in the number of qubits to perform gates; yet this is only at worst a polynomial overhead [33], thus preserving a quantum advantage for algorithms with an exponential speed-up. Nevertheless, as producing many stable qubits is a huge problem in quantum computation, cluster state computing with qubits is not likely to be a realistic option. However, this is where continuous-variable systems come in, as large-scale entanglement is much more feasible [37].

## 1.7 Continuous-Variable Quantum Computation

So far we have only talked about quantum computing with quantum systems having a discrete number of basis vectors, but what happens if we were to try and use a continuous-variable (CV) system, such as described using the phase and amplitude quadratures of a quantum-optical field[13]? Suppose we were able to achieve the infinite squeezing limit mentioned earlier and had access to $p$-quadrature eigenstates, including the ability to resolve a quadrature measurement to arbitrary precision. In principle, provided everything were perfect, one could implement *any* computation using a single quantum-optical mode, or qumode.

[13] Or, perhaps position and momentum of a particle? Really, any CV conjugate variables will do, but we'll stick to field quadratures for this discussion.

The method would proceed as follows: take the qumode and initialize it to have some predetermined wavefunction,

$$|\psi\rangle = \int dx \psi(x) \, |x\rangle_p \, . \tag{1.179}$$

Next, program some unitary gate, $U$, that maps $|\psi\rangle$ to a single quadrature eigenstate and measure the result. Because the CV state has an infinite number of parameters (i.e., a dense set of eigenvalues), there would be enough freedom to map any computational input into the initialized $\psi(x)$. Similarly, if $U$ rotated $|\psi\rangle$ into a single quadrature eigenstate and one had infinite precision on the measurement result, then the continuum of measurement results would again give an infinite number of parameters as an answer. Of course, this is infeasible on many practical levels. One must first devise a $U$ and an encoding scheme for the desired problem in addition to the fact that generating each specific $|\psi\rangle$ and $U$ would likely prove impractical if not impossible. Furthermore, infinite measurement precision will never be a reality. Nonetheless, this simple thought experiment leads one to believe that the continuous nature of qumodes may be beneficial in quantum information. As a practical implementation, let's look deeper into a cluster state implementation.

### 1.7.1 Qumode cluster states

First, as a note of bookkeeping, I will redefine $\hat{R}(\theta) = e^{i\theta a^\dagger a}$ for the remainder of this section, such that

$$\hat{R}^\dagger(\theta)\hat{Q}\hat{R}(\theta) = \hat{Q}\cos\theta - \hat{P}\sin\theta \tag{1.180}$$

$$\hat{R}^\dagger(\theta)\hat{P}\hat{R}(\theta) = \hat{P}\cos\theta + \hat{Q}\sin\theta. \tag{1.181}$$

This is to keep in line with numerous other sources on the topic of CVQC with cluster states, such as Refs. [18, 38–40], to name a few.

To begin with, assume that we have access to infinitely squeezed states in the form of $|0\rangle_p$ and $|0\rangle_q$ states. After developing the ideas of QC with ideal quadrature eigenstates, we will relax the infinite squeezing condition and see that QC is still possible with finite squeezing provided we have the ability to perform error correction. Analogous to qubit cluster states, the canonical CV qumode cluster state is formed by placing qumodes in a lattice and entangling each qumode with its nearest neighbors.

Each qumode is initialized to the zero-momentum eigenstate $|0\rangle_p$ and then $\hat{C}_Z$ gates of the form of Eq. 1.106 of weight $g = 1$ are applied to all neighboring pairs. This results in the cluster state

$$\prod_{j|j\in n_k} e^{i\hat{Q}_j\hat{Q}_k} \bigotimes_k^N |0\rangle_{p_j}|0\rangle_{p_k}, \tag{1.182}$$

which is analogous to the qubit cluster. Additionally, note that

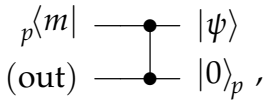
$$\hat{C}_Z|0\rangle_{p_1}|0\rangle_{p_2} = \hat{R}_2(\tfrac{\pi}{2})\hat{C}_X|0\rangle_{p_1}|0\rangle_{q_2}, \tag{1.183}$$

so essentially each neighboring pair in the cluster is an EPR pair up to a single-mode rotation. The exact form of the qumode cluster state, or any graph state, can be determined through the nullifier formalism [41, 42], which is equivalent to stabilizers in qubit graph states [43]. These are also valuable tools to verify an experimentally generated cluster state [37], but here I will focus on what to do with a cluster state once it is already made.

Although a 2-D cluster state is the minimum substrate required to enact universal QC, we will start by understanding the 1-D chain, which is enough to demonstrate all single-qumode operations. From there, extending to 2-D becomes straightforward. First, let's start by deriving the analogous circuit to the qubit teleportation circuit in Eq. 1.177.

### Teleportation

As before, consider the simplest two-mode cluster state, which consists of taking arbitrary state $|\psi\rangle$ and entangling it to state $|0\rangle_p$ with a $\hat{C}_Z$ interaction. Next, to complete the teleportation, we will measure qumode one with a homodyne measurement, which is a simple form of Gaussian measurement that projects the state into a quadrature basis. To start, consider performing the homodyne measurement in the $p$-quadrature basis and obtaining result $m$. This circuit will be drawn as:



(1.184)

which can be written out as

$$\begin{aligned} |\phi\rangle_{out} &= {}_{p_1}\langle m|\hat{C}_{Z_{12}}|\psi\rangle_1|0\rangle_{p_2} \\ &= {}_{p_1}\langle m|e^{i\hat{Q}_1\hat{Q}_2}\int dx\,\psi_q(x)|x\rangle_{q_1}|0\rangle_{p_2}, \end{aligned} \tag{1.185}$$

where $\psi_q$ is the amplitude (or position) wavefunction of the input state. Next, we can move the operator and projector through the integral to apply them to the states to arrive at

$$\begin{aligned} |\phi\rangle_{out} &= \int dx\,\psi_q(x)\,{}_{p_1}\langle m|e^{ix\hat{Q}_2}|x\rangle_{q_1}|0\rangle_{p_2} \\ &= \tfrac{1}{\sqrt{2\pi}}\int dx\,\psi_q(x)e^{-ixm}|x\rangle_{p_2}, \end{aligned} \tag{1.186}$$

where we used

$${}_p\langle s|t\rangle_q = \tfrac{1}{\sqrt{2\pi}}e^{-ist}. \tag{1.187}$$

Next, rotate back to the $q$-basis by inserting the identity in the form of $\hat{R}(\frac{\pi}{2})\hat{R}^\dagger(\frac{\pi}{2})$, and then promote the exponential to an operator so it can be pulled out of the integral:

$$\begin{aligned} |\phi\rangle_{out} &= \tfrac{1}{\sqrt{2\pi}}\hat{R}_2(\tfrac{\pi}{2})\int dx\,\psi_q(x)e^{-imx}|x\rangle_{q_2} \\ &= \tfrac{1}{\sqrt{2\pi}}\hat{R}_2(\tfrac{\pi}{2})e^{-im\hat{Q}_2}\int dx\,\psi_q(x)|x\rangle_{q_2} \\ &= \tfrac{1}{\sqrt{2\pi}}\hat{R}_2(\tfrac{\pi}{2})\hat{Z}_2^\dagger(m)|\psi\rangle_2. \end{aligned} \tag{1.188}$$

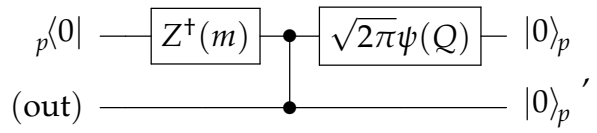
Thus, we see that the input state $|\psi\rangle$ in mode one has been perfectly teleported to mode two, with an additional measurement-dependent quadrature shift followed by a $\pi/2$ rotation. Furthermore, we could now just relabel the modes as mode one has been projected out, so we could write the whole process as a single Kraus operator applied to the input,

$$|\phi\rangle_{out} = \hat{K}_m\,|\psi\rangle\,; \qquad \hat{K}_m = \tfrac{1}{\sqrt{2\pi}}\hat{R}(\tfrac{\pi}{2})\hat{Z}^\dagger(m). \tag{1.189}$$

It is also useful to note that the rotation of $\frac{\pi}{2}$ is equivalent to a single-mode quantum Fourier transform as it rotates the $q$-basis to the $p$-basis. This is similar to the Hadamard gate transforming a qubit from the computational basis to the $\pm$ conjugate basis — an analogy that will be further developed later.

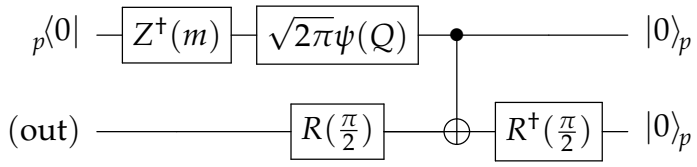
Instead of performing the derivation as above, it is instructive to develop the derivation graphically starting from the circuit in Eq. 1.184. Using the trick from Eq. 1.34

to write $|\psi\rangle$ as a function of an operator acting on a momentum eigenstate and pulling out the measurement result from the detection, we can write



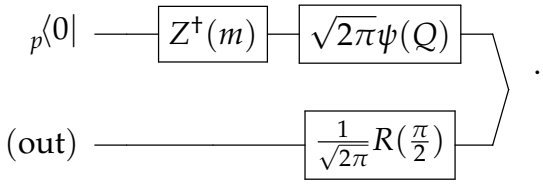

(1.190)

where I will neglect hats on operators in circuit diagrams with the assumption that a capitalized $Q$ or $P$ indicates an operator while lowercase indicates an eigenvalue. Because the $\hat{C}_Z$ gate commutes with any operators that are functions of $\hat{Q}$ only, we can move $\psi(\hat{Q})$ through the gate. After moving this through, we can transform the $\hat{C}_Z$ gate to a $\hat{C}_X$ by using that $\hat{C}_Z = \hat{R}_2(\frac{\pi}{2})\hat{C}_X\hat{R}_2^\dagger(\frac{\pi}{2})$. This yields the circuit



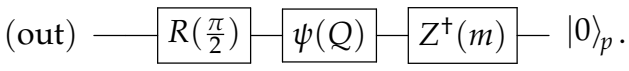

(1.191)

The right half of the circuit, from the $\hat{C}_X$ gate on can be recognized as the circuit for the EPR state from Eq. 1.116, so we can close off the end and write




(1.192)

Finally, we can 'bounce' the operators from the top wire to the bottom through the EPR pair using Eq.1.117 with $\hat{Q}^T = \hat{Q}$, and pull the circuit taut, so to speak, to arrive at the Kraus operator for teleportation:




(1.193)

Writing down these operators applied to the zero-momentum eigenstate recovers Eq. 1.189.

### Disconnecting nodes

We just saw how teleportation works with $p$-quadrature measurements, but let's instead change the measurement basis and measure in the $q$-quadrature instead. However, when the measurement is a $q$-eigenstate, the entangling $\hat{C}_Z$ gate can equally well

act to the left instead of the right in the circuit drawings, and in this way the entangling gate only acts to add a single-mode displacement to the unmeasured mode — the entanglement is broken and thus there is no teleportation. Explicitly, we have that

$$\begin{aligned} &{}_{q_1}\langle m|e^{i\hat{Q}_1\hat{Q}_2}|\psi\rangle_1|0\rangle_{p_2} \\ =&e^{im\hat{Q}_2}{}_{q_1}\langle m||\psi\rangle_1|0\rangle_{p_2} \\ \propto&\hat{Z}_2(m)|0\rangle_p. \end{aligned} \tag{1.194}$$

Regardless of the unmeasured state, whether it is a momentum eigenstate or not, performing a $q$-basis homodyne measurement on a canonical cluster state node will disconnect that node and apply local quadrature shifts on the neighboring nodes. Note that if one had defined the cluster state differently, say with $\hat{C}_X$ gate acting on alternating zero $q$ and zero $p$ eigenstates, then the disconnecting and teleporting homodyne measurement bases would be reversed.

### General homodyne measurement

Next, instead of a $q$ or $p$ homodyne measurement, considering measuring in a general rotated quadrature. This we will label as $p_\theta$ where

$$\hat{P}_\theta = \hat{R}^\dagger(\theta)\hat{P}\hat{R}(\theta) = \hat{P}\cos\theta - \hat{Q}\sin\theta, \tag{1.195}$$

which is the same setting as measuring the scaled operator

$$\frac{\hat{P}_\theta}{\cos\theta} = \hat{P} - \hat{Q}\tan\theta. \tag{1.196}$$

This happens to be the same effect as applying the *shear* operator[14],

$$\hat{\mathcal{P}}(t) = e^{i\frac{t}{2}\hat{Q}^2} \tag{1.197}$$

which acts on $\hat{P}$ as

$$\hat{\mathcal{P}}^\dagger(t)\hat{P}\hat{\mathcal{P}}(t) = \hat{P} + t\hat{Q}. \tag{1.198}$$

[14]Which is, unfortunately, referred to as the phase gate/operator in some sources, which further complicates the terminology.

The above equation is derivable from the BCH formula, but also can be found quite easily from the Heisenberg equation in Eq. 1.47. If we measure in the rotated basis with $\theta = \arctan(-t)$ and scale the outcome, then a measurement of $m$ in the $p_\theta$ basis is the same as measuring

$$|m\rangle_{p_\theta} = e^{-\frac{i\tan\theta}{2}\hat{Q}^2}|\tfrac{m}{\cos\theta}\rangle_p \tag{1.199}$$

Applying this as the measurement in the teleportation circuit earlier, we find that the Kraus operator associated with a rotated measurement is

$$\hat{K}_{p_\theta} = \hat{R}(\tfrac{\pi}{2})\hat{Z}^\dagger(\tfrac{m}{\cos\theta})\hat{\mathcal{P}}(\tan\theta). \tag{1.200}$$

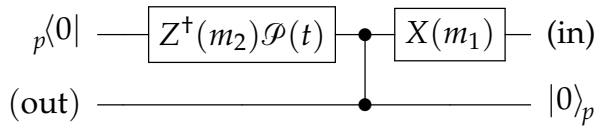
The reader might note that so far, all of these measurements apply unitary operators. This might seem strange, as we are in fact performing a measurement to project a portion of the quantum state, which one associates with non-unitary evolution. However, in this case, the presence of the ancillary $|0\rangle_p$ state in the second mode and the entangling operator $\hat{C}_Z$ ensure that measurement in the $p$-basis of the mode after entanglement provides no information at about the input state $|\psi\rangle$. The measurement result is completely and uniformly random without collapsing the wavefunction of the input state, leading to a perfect teleportation. This is a result of the infinite squeezing limit, and thus the ideal unitary gate evolution will only approximately be true with finite squeezing, as will be discussed later.

**Feed-forward displacements**

It has been mentioned that the effects of the quadrature shifts can be accounted for by feed-forward. Let's examine exactly how this can be done by tracking what happens to the measurement-dependent shift from one measurement to the next. Suppose we have just performed a teleportation by performing a homodyne measurement on a linear cluster state, and we would like to make a subsequent homodyne measurement on the next mode. The Kraus operator applied for a general homodyne measurement in Eq. 1.200 can be rewritten with the shift operator at the back to read

$$\hat{K}_{p_\theta} = \hat{X}(\tfrac{m}{\cos\theta})\hat{R}(\tfrac{\pi}{2})\hat{\mathcal{P}}(\tan\theta). \tag{1.201}$$

If we push everything except the shift into the state, then we can write the next teleportation gadget as

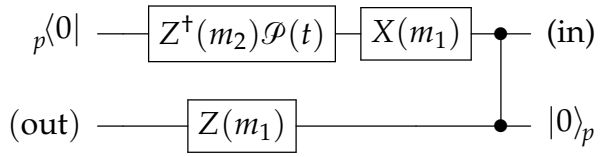



(1.202)

where the effects of the physical measurement result and the scaling on the eigenvalue from the rotated basis have been combined into general measurement results $m_1$ and $m_2$ for simplicity. The $\hat{X}$ shift can now be commuted with the $\hat{C}_Z$ gate using

$$\hat{C}_Z \hat{X}(m_1) = e^{-im_1(\hat{P}_1 - \hat{Q}_2)} \hat{C}_Z \tag{1.203}$$

to yield the circuit




(1.204)

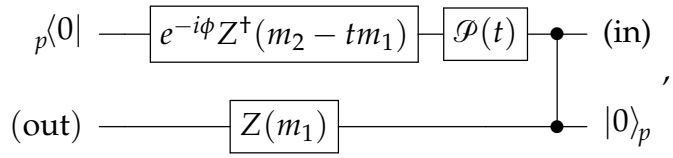
Using the BCH identities, we can commute the shear gate with the quadrature shift to find

$$\hat{\mathcal{P}}(t)\hat{X}(s) = e^{i\frac{st}{2}} \hat{X}(s)\hat{Z}(st)\hat{\mathcal{P}}(t), \tag{1.205}$$

and commuting two orthogonal quadrature shifts gives

$$\hat{Z}^\dagger(m)\hat{X}(s) = e^{-ism} \hat{X}(s)\hat{Z}^\dagger. \tag{1.206}$$

Using the additional fact that $\hat{X}(s)|0\rangle_p = |0\rangle_p$, we arrive at the circuit




(1.207)

where the picked up global phase is $\phi = m_1(m_2 - \frac{t}{2})$. To proceed, we can simply replace $m_2 - tm_1$ with $m_2'$[15], and everything is the same as before up to an additional shift on the output. Thus, each prior shift can be accounted for at the current measurement by using the result of the previous measurement value to inform how the user should modify the current measurement result. This can proceed as measurements are performed all the way across the cluster state. An important caveat to mention is that this methodology

[15]If the experimenter measured a homodyne result of $m_2$, he would just subtract $tm_1$ and use the new value in place of $m_2$ everywhere.

is less simple when one no longer performs a quadrature measurement.

### General measurement

Instead of simply performing a homodyne measurement in the $p_\theta$-basis, consider projecting onto an arbitrary pure state. This can written as projecting onto a bra, $\langle m|$, which can written in the $q$-basis as

$$\langle m| = \int dx\, m(x)_q\langle x| = \sqrt{2\pi}_p\langle 0|m(\hat{Q}). \tag{1.208}$$

This will act to replace ${}_p\langle m|$ in the circuit in Eq. 1.184, which after following the same derivation, will yield the circuit

$$\text{(out)} \quad \boxed{R(\tfrac{\pi}{2})} \boxed{\psi(Q)} \boxed{\sqrt{2\pi}m(Q)} \; |0\rangle_p\,. \tag{1.209}$$

Since we wrote this measurement operator in the $q$-basis initially, we can commute it with the wavefunction operator of the initial state and arrive at the general Kraus operator for the process of arbitrary pure state measurement as

$$\hat{K}_m = \hat{R}(\tfrac{\pi}{2})m(\hat{Q}), \tag{1.210}$$

with the measurement result occurring with probability

$$P_m = \mathrm{Tr}[\hat{K}_m\,|\psi\rangle\,\langle\psi|\,\hat{K}_m^\dagger]. \tag{1.211}$$

Suppose we were to use a more general projector that accounted for mixed states. In this case, the new output state, up to a normalization, would be given by

$$\rho_{out} = \sum_j \hat{K}_{m_j}\,|\psi\rangle\,\langle\psi|\,\hat{K}_{m_j}^\dagger, \tag{1.212}$$

where $\hat{K}_{m_j}$ are Kraus operators associated with the measurement operator diagonalized in the $q$-basis.

Regardless of what measurement operator is applied, there is always a rotation operator applied from the teleportation. Supposing that this rotation is unwanted and one can only apply operations through measurement, this can be removed by commuting the rotation to the back and successively teleporting the state three more times to

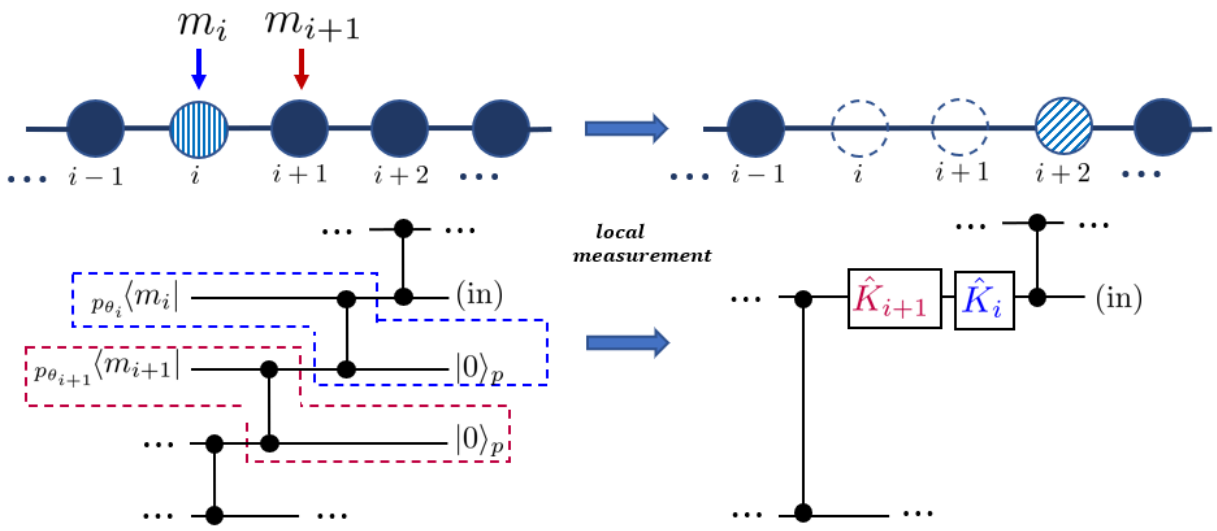



FIGURE 1.10: Measurements on a cluster state teleport quantum information and apply gates. Each step can be look at as applying a Kraus operator with measurement-dependent parameters to the input. Each graph can be represented by the circuit shown below. More details in the main text.

erase the effects since $\hat{R}^4(\frac{\pi}{2}) = \mathbb{1}$.

## 1.8 Universal Gate Set

We just saw how performing a measurement on a two-qumode cluster state can be used to teleport the state and apply an operator. This idea can be extended to an arbitrarily long 1-D cluster state, where measurements simply act to teleport the state along and apply gates. This is depicted in Fig. 1.10, where a state at node $i$ is teleported to node $i+2$ after two measurements. Recall that a universal gate set requires one to implement any arbitrary unitary gate on the quantum system. We will start by showing how one can implement the entire Gaussian gate set, analogous to the Clifford gate set with qubits, simply by performing homodyne detection on the cluster.

### 1.8.1 Gaussian gate set

The single-mode general Gaussian unitary can be written as a transformation of phase and amplitude quadrature operators in the Heisenberg picture,

$$\begin{pmatrix} \hat{Q}' \\ \hat{P}' \end{pmatrix} = \hat{U}_G^\dagger \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix} \hat{U}_G = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}, \tag{1.213}$$

where $ad - bc = 1$. The affine portion containing $e$ and $f$ describe the effects of displacements, which can be accounted for by shifting homodyne measurement results as desired [38, 44], so we will only consider the linear portion. We now need a sequence of gates that can form an arbitrary symplectic transformation to allow us to choose $a$, $b$, $c$ and $d$. The teleportation circuit with general homodyne measurements allows us to perform the shear gate, which in symplectic form is given as

$$\begin{pmatrix} \hat{Q}' \\ \hat{P}' \end{pmatrix} = \hat{\mathcal{P}}^\dagger(\kappa) \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix} \hat{\mathcal{P}}(\kappa) = \begin{pmatrix} 1 & 0 \\ \kappa & 1 \end{pmatrix} \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix}, \tag{1.214}$$

where $\kappa$ is a real-valued parameter.

We can also perform on-demand Fourier transforms of $\hat{F} = \hat{R}(\frac{\pi}{2})$ by measuring in the $p$-basis, which are described by

$$\begin{pmatrix} \hat{Q}' \\ \hat{P}' \end{pmatrix} = \hat{F}^\dagger \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix} \hat{F} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \hat{Q} \\ \hat{P} \end{pmatrix}. \tag{1.215}$$

Overall, the linear matrix transformation resulting from a single homodyne measurement on a cluster state is given by

$$M(\kappa) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \kappa & 1 \end{pmatrix} = \begin{pmatrix} -\kappa & -1 \\ 1 & 0 \end{pmatrix}. \tag{1.216}$$

By combining matrices of this form with suitable choices for $\kappa$, it turns out that one can implement an arbitrary symplectic transformation on a single mode with just four quadrature measurements [39]. This lets us apply

$$\prod_{i=1}^{4} M(\kappa_i) = \begin{pmatrix} \kappa_4\kappa_3\kappa_2\kappa_1 - \kappa_4\kappa_3 - \kappa_2\kappa_1 - \kappa_4\kappa_1 + 1 & \kappa_4\kappa_3\kappa_2 - \kappa_4 - \kappa_2 \\ -\kappa_3\kappa_2\kappa_1 & -\kappa_3\kappa_2 + 1 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{1.217}$$

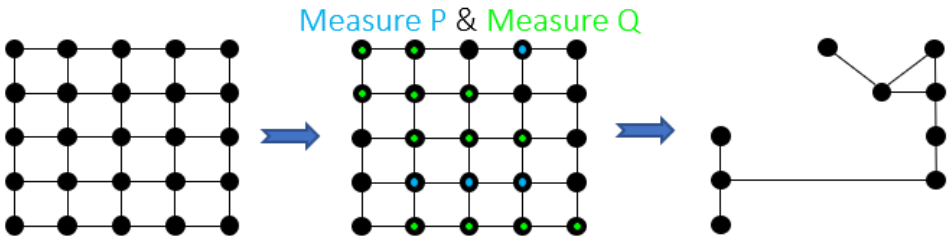



FIGURE 1.11: The cluster state resource can be shaped into a different graph state by performing homodyne measurements.

where the decomposition is given by

$$\kappa_2 = \frac{1-d}{\kappa_3}, \qquad \kappa_3 = c - d\kappa_1, \qquad \kappa_4 = \frac{1-a+b\kappa_1}{\kappa 3}, \tag{1.218}$$

and $\kappa_1$ is a free parameter. Ref. [39] is careful to note that many single-mode symplectic operations can in principle be realized with only three measurements on the cluster; however, one cannot in general choose an arbitrarily large value for $\kappa_i$ due to finite squeezing. Using all four operations gives one some freedom to avoid finite-squeezing noise if done properly.

As an example, suppose we wished to apply a single-mode squeeze gate, $\hat{S}(z)$ which can be decomposed into three operations. This can be achieved by applying

$$M_3(s)M_2(s^{-1})M_1(s) = \begin{pmatrix} -s & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -s^{-1} & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} -s & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} s & 0 \\ 0 & s^{-1} \end{pmatrix}, \tag{1.219}$$

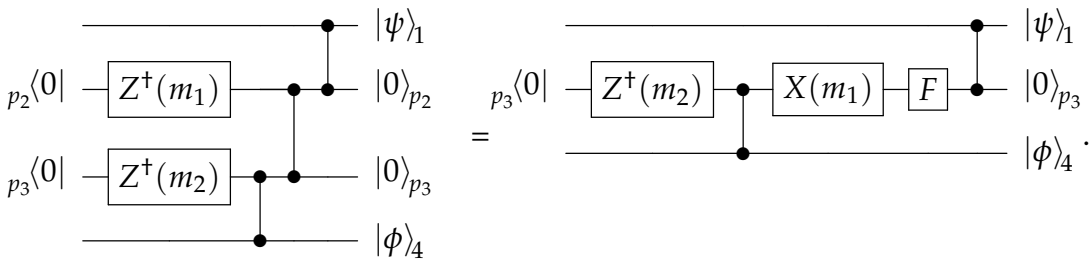
where $s = e^z$.

### Two-mode gates

Combining an entangling two-mode operation, such as the native $\hat{C}_Z$ gate, with the ability to perform universal single-mode gates is sufficient to generate a universal gate set over any number of qumodes[16]. The cluster state is initialized with $\hat{C}_Z$ gates connected to each of the neighboring nodes, but a $\hat{C}_Z$ gate can be enacted between non-neighboring nodes by forming a 'bridge' with measurements. This is done by performing $p$-basis homodyne measurements on all nodes along a path connecting them. Similarly, any undesired nodes can be disconnected, as depicted in Fig. 1.11.

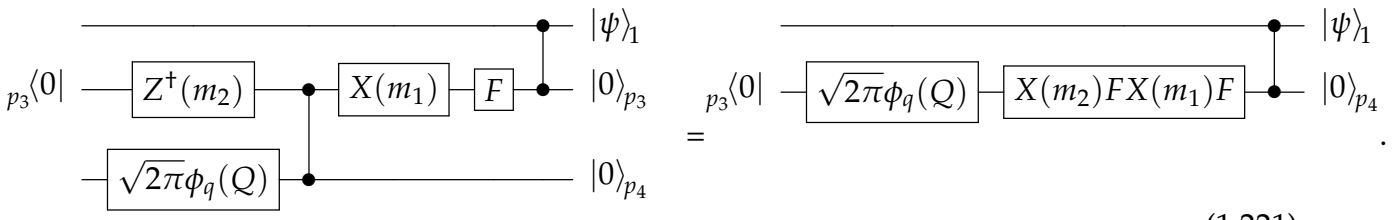
[16] We still need non-Gaussianity, of course, which will be discussed in the next section.

To see how moving around the $\hat{C}_Z$ gates works in practice, consider states $|\psi\rangle$ and $|\phi\rangle$ sitting at two nodes on a cluster state separated by two intermediate nodes. If we measured out the middle nodes, these two states will now become connected. This is shown by first taking the initial circuit with four qumodes, where nodes two and three are the intervening nodes, and measuring node 2:



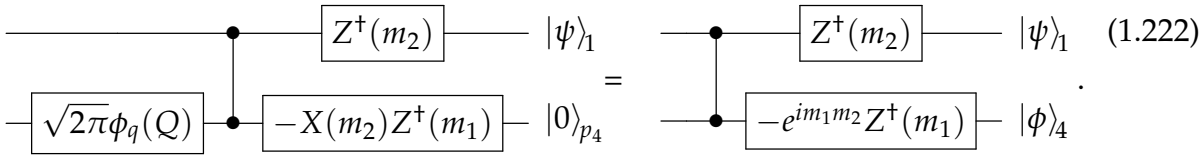

(1.220)

Before measuring the remaining node, we rewrite $|\phi\rangle$ as $|\phi\rangle = \sqrt{2\pi}\phi_q(\hat{Q})|0\rangle_p$ and commute the operator in $\hat{Q}$ with the $\hat{C}_Z$ gate on the bottom wire. After this, we measure the remaining node to arrive at




(1.221)

We now commute the Fourier transforms to use that $\hat{F}^2 = -\mathbb{1}$, and then commute the remaining shift operators through the $\hat{C}_Z$ gate using $\hat{X}_2\hat{C}_Z = \hat{C}_Z\hat{Z}_1^\dagger\hat{X}_2$. From here, we can apply the $\hat{X}_2$ shift on the momentum eigenstate which just results in a global phase. This lets us write




(1.222)

Looking at the above, we see that we have implemented a $\hat{C}_Z$ gate between our two qumodes, up to a global phase and single-mode displacements that can be accounted for with feed-forward as discussed in Sec. 1.7.1. It is important to measure two 'empty' nodes in between the two nodes with the quantum information in order to have the teleportation-induced Fourier transforms cancel out. Otherwise, the quantum information in one mode will be Fourier transformed, and the effects of the measurement

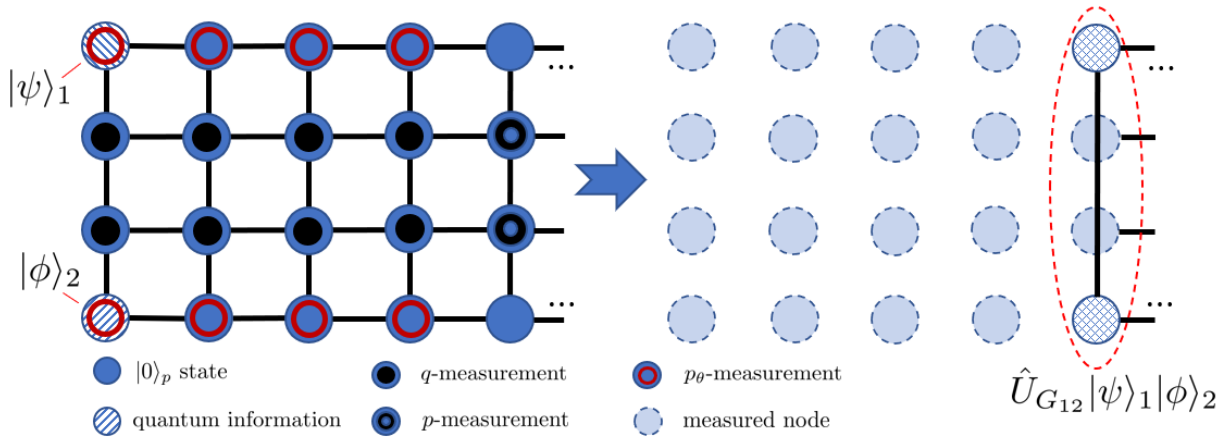



FIGURE 1.12: Any two-mode Gaussian unitary, up to displacements, can be implemented by performing up to 10 homodyne measurements. Additional $q$-basis measurements are shown to disconnect nodes that are not used for implementing the gate, but these measurement results will still matter in terms of the overall measurement-dependent displacements.

result will no longer be as easily undone.

Combining this mechanism for enacting entanglement between two nodes and the universal single-mode Gaussian gate means that one can enact an arbitrary two-mode Gaussian unitary operation by strategically performing at most 10 homodyne measurements. This is depicted in Fig. 1.12. Note that more than 10 measurements are used in this figure, as additional $q$-measurements are used to disconnect the unneeded cluster state nodes. If one had control over the initial cluster state construction, the measured-out nodes could have been left out entirely.

### 1.8.2 Non-Gaussian gates

So far, this section has shown how one can perform arbitrary single and two-mode Gaussian unitary gates using the cluster state. This allows one to build up arbitrary multi-mode Gaussian unitaries, but it will never lead to a quantum speed-up as per the Gottesman-Knill theorem [45]. In order to achieve something non-trivial, it is necessary to enact a gate or use a state that is non-Gaussian. It is important to note that this is a necessary criterion and not sufficient, as implementing certain gates sets on classes of non-Gaussian states can still be simulated efficiently [12, 46].

As all symplectic matrices translate to a Gaussian transformation, this non-Gaussian

gate will be beyond the symplectic formalism. One way to achieve universality is to have the capability to enact gates of higher-order polynomials in the quadrature operators, such as the cubic phase gate

$$\hat{V}_\gamma = e^{i\gamma\hat{Q}^3}. \tag{1.223}$$

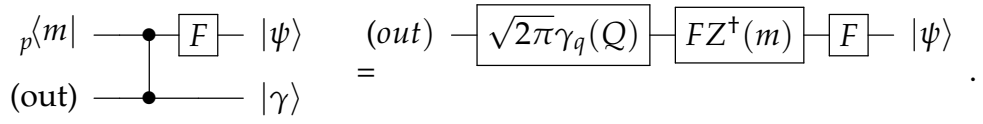
Applying Gaussian gates that are at most quadratic in the field operators along with the cubic phase gate allows one to build up unitary gates of arbitrary polynomials. This is due to the fact that commutators of the form $[\hat{Q}^k, \hat{P}^2]$ where $k \geq 3$ generate polynomials in the field of degree greater than $k$. In this way, one can approximate any Hamiltonian to an arbitrary degree of precision using a polynomial expansion, and enact the unitary $e^{-i\hat{H}t}$ from combinations of lower-order gates. Although this is feasible in principle with just cubic phase gates and Gaussian operations, the true composition may be quite cumbersome. Arbitrary gate decompositions can be calculated using Ref. [47].

### Gate teleportation

Directly implementing higher-order Hamiltonians optically is an experimental challenge due to weak higher-order nonlinearities. Instead, an approach that can be taken is to generate offline resource states and then teleport the wavefunction onto our quantum information as a way to apply the desired gate. For example, suppose our goal is to apply a cubic phase gate. To do so, we can instead generate a cubic phase state and teleport our quantum information through it to apply the cubic phase gate. Taking the cubic phase state defined as [48]

$$|\gamma\rangle = \tfrac{1}{\sqrt{2\pi}} \int dx e^{i\gamma x^3} |x\rangle_q, \tag{1.224}$$

we can attach this state to our cluster state and implement the circuit




. (1.225)

Remember that the Fourier transform in the initial circuit can be implemented by performing a normal teleportation through a $|0\rangle_p$ state. The final result of this circuit, up to a shift of $\hat{X}(m)$ at the end and global phase, is to apply the operator

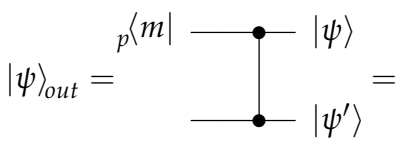
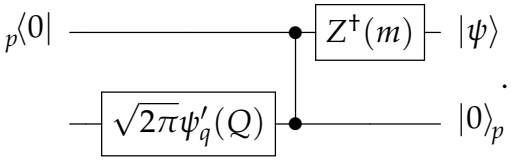
$$\hat{U} = \sqrt{2\pi}\gamma_q(\hat{Q} + m) = e^{i\gamma(\hat{Q}+m)^3}. \tag{1.226}$$

However, this doesn't precisely apply the cubic phase gate we wanted due to commuting the measurement-dependent shift through the cubic phase state wavefunction. We can recover the cubic phase gate of Eq. 1.223 by applying the gate

$$\hat{U}_G = e^{i\left[\gamma\hat{Q}^3 - \gamma(\hat{Q}+m)^3\right]}, \tag{1.227}$$

which is only quadratic in the field. In general, the measurement-dependent shift from gate teleportation will necessitate a correction operation that is a polynomial of one degree lower than the gate applied by teleportation. In this way, cubic gates require a quadratic correction, and as we saw in Sec. 1.7.1, applying the natural quadratic gate on the cluster (shear and rotation) necessitates a linear correction that can be mitigated with feed-forward corrections to future measurement bases.

The idea of gate teleportation can be generalized to more general types of operations beyond unitary gates dependent on the form of the ancillary state. Teleporting through an ancillary state acts to filter the $p$-basis wavefunction of the input with the $q$-basis wavefunction of ancilla shifted by the homodyne result. This is seen generically by replacing the momentum eigenstate from the teleportation circuit in Eq. 1.184 with an arbitrary quantum state, $|\psi'\rangle$. This can be rewritten as an operator in $\hat{Q}$ applied to a zero momentum eigenstate and commuted through the $\hat{C}_Z$ gate to have the general circuit

$|\psi\rangle_{out} =$  $=$  . (1.228)


Pulling the circuit taught and commuting the resultant Fourier transform to the end, we have that

$$|\psi\rangle_{out} = \sqrt{2\pi}\hat{F}\psi'_q(-\hat{P})\hat{Z}^\dagger(m)\,|\psi\rangle\,, \tag{1.229}$$

where the subscript on $\psi'_q(-\hat{P})$ indicates that it is the $q$-basis wavefunction of the starting ancillary state, even though it is a function of the $\hat{P}$ operator. Finally, commuting the displacement through to the left reveals the result of

$$|\psi\rangle_{out} = \sqrt{2\pi}\hat{X}(m)\hat{F}\psi'_q(-\hat{P}+m)\,|\psi\rangle\,. \tag{1.230}$$

This shows that up to an overall displacement and rotation, the initial $P$-basis wavefunction for $|\psi\rangle$ transforms as $\psi_p(x) \to \psi_p(x)\psi'_q(m-x)$. Thus, in the $P$-basis, $|\psi\rangle$ is

filtered by the $Q$-wavefunction of $|\psi'\rangle$.

If we had first applied a Fourier transform to the input state as we did for the cubic phase gate teleportation circuit in Eq. 1.225, the operator applied would be

$$\hat{O}_{\psi'} = -\sqrt{2\pi}\hat{X}(m)\psi'_q(\hat{Q}+m) \tag{1.231}$$

If one has the ability to engineer non-Gaussian states offline and then attach them to the cluster, gate teleportation gives the ability to teleport the non-Gaussianity onto the quantum information in a given calculation. However, this will in general not be a unitary operation unless $\psi'_q(x)$ is of the form $e^{if(x)}$. Because of the non-unitarity, the measurement-dependent shift will not be correctable with unitary operations. For example, suppose that $\psi'$ was the wavefunction of a Fock state. In this case, gate teleportation would apply the operation (up to normalization and shift)

$$\hat{O}_{\psi'} = e^{-\frac{1}{2}(\hat{Q}+m)^2}H_n(\hat{Q}+m), \tag{1.232}$$

where $H_n(x)$ is a Hermite polynomial. This motivates the possibility of applying polynomial gates, which I will explore later in Ch. 3.

## 1.9 Realistic Cluster States

These cluster states discussed have so far been the canonical cluster, based on the $\hat{C}_Z$ entangling interaction, and have been ideal with the infinite squeezing limit. First, let's maintain the infinite squeezing limit for now and briefly discuss experimental cluster states.

### 1.9.1 Macronode cluster states

Although the entangling $\hat{C}_X$ gate has been implemented optically [49], performing many $\hat{C}_X$ or $\hat{C}_Z$ is challenging as they are not passive gates — the energy of the system is not conserved. As such, all-optical cluster states to date have been based on using the beamsplitter as the entangling operator [50–57]

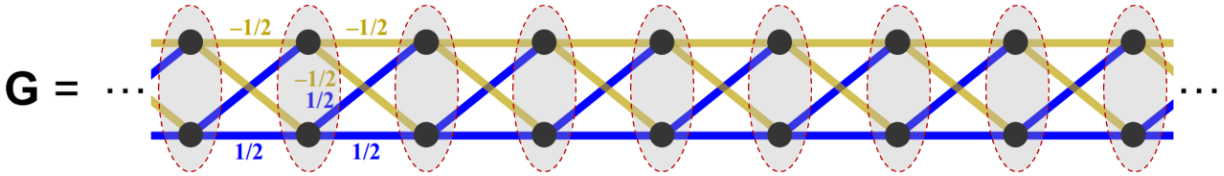


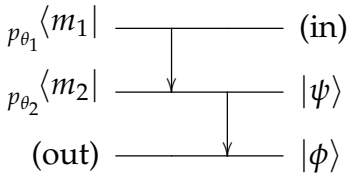

FIGURE 1.13: The graph state of a dual-rail quantum wire, which is the macronode equivalent of a 1-D canonical cluster state. Weighted connections represent entanglement between modes as outlined in the graphical calculus [42]. Quantum information is distributed between pairs of qumodes (circled), and teleportations are enacted by performing homodyne measurements on both qumodes in a macronode. Figure adapted from Ref [54].

These types of cluster states are formed from *macronodes*, which are a collection of two or more physical modes containing non-locally distributed information encoding a single logical state [40]. The macronode cluster states are created by beginning with many entangled resources in the form of two-mode squeezed states, which can be generated either directly or by interfering two single-mode squeezed states on a beamsplitter, and then entangled pairs are linked up through more beamsplitter interactions. An example of the dual-rail quantum wire is shown in Fig. 1.13, which is the equivalent of a single 1-D chain canonical cluster state. A single node from the 1D canonical cluster state now maps to quantum information distributed between a pair of qumodes, and quantum information processing can proceed analogously to the canonical case, where now all physical modes within each macronode are subject to homodyne measurements.

The simplest macronode cluster state, the quantum wire, contains two physical modes per macronode and has been used to experimentally implement Gaussian operations [58]. Weaving together quantum wires can produce higher-dimensional cluster states [59, 60]; the 2D case of this now has four qumodes per macronode and has recently been used to implement a set of universal Gaussian gates [61]. This is a straightforward, if cumbersome, extension of the 1D macronode wire.

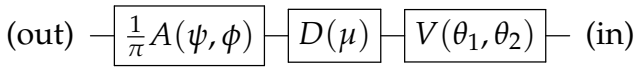
We briefly review formalism recently developed to treat general teleportation on a macronode wire, where arbitrary and potentially non-Gaussian ancillary inputs can be used to impart non-Gaussianity on the cluster through teleportation [18]. Further details and examples can be found elsewhere [62–66], including a dictionary protocol that can be used to map operations from macronodes to canonical clusters [40]. The general teleportation operation can be described by a local teleportation gadget which

has the circuit form:




(1.233)

where the down arrow represents the balanced beamsplitter interaction as in Eq. 1.98. Note that an up arrow would be the daggered operation, which is the same as the beamsplitter with the mode indices swapped. For the choice of beamsplitter phase in Eq. 1.98, the $\uparrow$ and $\downarrow$ operations are not equivalent. The overall effect of this teleportation circuit is given by [18, 40]




(1.234)

which is the same as applying the Kraus operator,

$$\hat{K}_{m_1,m_2} = \frac{1}{\pi}\hat{A}(\psi,\phi)\hat{D}(\mu)\hat{V}(\theta_1,\theta_2), \tag{1.235}$$

to the input state. Here, $\hat{D}(\mu)$ is a displacement based on the homodyne measurement with

$$\mu = \frac{m_1 e^{i\theta_2} + m_2 e^{i\theta_1}}{\sin(\theta_2 - \theta_1)}. \tag{1.236}$$

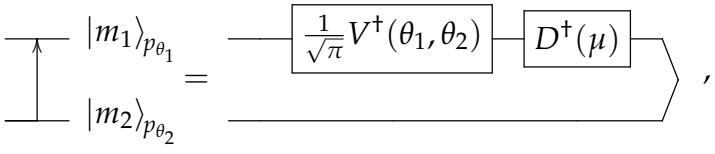
and $\hat{V}$ is a Gaussian operation that depends only on the measurement basis given by

$$\hat{V}(\theta_1,\theta_2) = \hat{R}(\theta_+ - \tfrac{\pi}{2})\hat{S}\left(\ln(\tan\theta_-)\right)\hat{R}(\theta_+). \tag{1.237}$$

We have additionally defined

$$\theta_\pm = \frac{1}{2}(\theta_1 \pm \theta_2). \tag{1.238}$$

This was derived in part using the useful identity [18],




(1.239)

where the ended circuit represents an EPR pair.

The potentially non-Gaussian operation, $\hat{A}(\psi,\phi)$, comes from non-Gaussianity within the input states $|\psi\rangle$ and $|\phi\rangle$, and can be determined by

$$\hat{A}(\psi,\phi) = \iint d^2\alpha \widetilde{\psi}(\alpha_I)\phi(\alpha_R)\hat{D}(\alpha) \tag{1.240}$$

with complex parameter $\alpha = \alpha_R + i\alpha_I$. The respective position and momentum wavefunctions of the ancillary states are given by $\widetilde{\psi}(t) = {}_p\langle t\,|\psi\rangle$ and $\phi(s) = {}_q\langle s\,|\phi\rangle$. In the case where the input states are Gaussian only, then $\hat{A}(\psi,\phi)$ is also Gaussian. In the specific case where the inputs are position and momentum eigenstates, $|\psi\rangle = |0\rangle_p$ and $|\phi\rangle = |0\rangle_q$, then the operator $\hat{A}(\psi,\phi)$ reduces to the identity, and we have the ability to apply a complete Gaussian gate set as in the canonical case. These methods and the dictionary protocol will be used more in Ch. 3.

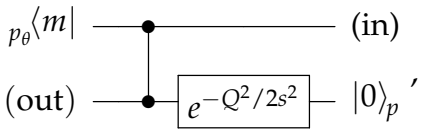
### 1.9.2 Finite squeezing

Now it is time to at least partially address the realistic nature of finite squeezing. To show the effects of this alone, I will simply use the canonical cluster state, since everything can be converted to macronodes by using the proper dictionary protocol. The use of finite squeezing means that every regular node of a cluster state that was a $|0\rangle_p$ must be replaced with squeezed vacuum given by Eq. 1.89, which can also be represented as

$$\hat{S}(\iota)|0\rangle_N = \frac{\sqrt{2}\pi^{1/4}}{\sqrt{s}} e^{-\frac{\hat{Q}^2}{2s^2}} |0\rangle_p, \tag{1.241}$$

where $s = e^r$.

With this, the general teleportation circuit is described by



(1.242)

where I have neglected the normalization for now. Commuting the $\hat{Q}$ operators through the $\hat{C}_Z$ and using the ideal teleportation leads to the circuit applying the measurement-dependent Kraus operator

$$\hat{K}_{p_\theta,s} = \frac{\sqrt{2}\pi^{1/4}}{\sqrt{s}} \hat{X}(m') e^{-\frac{(\hat{Q}+m')^2}{2s^2}} \hat{F}\hat{\mathcal{P}}(\tan\theta), \tag{1.243}$$

where $m' = m\sec(\theta)$ for general homodyne measurement at angle $\theta$. Looking at the form of this Kraus operator, we see that the Fourier transform and shear operator are applied exactly the same as in the ideal case, and the measurement-dependent shift still appears at the back to be undone with feed-forward later. However, we now have a Gaussian envelope that filters the state in the $q$-basis centered about the inverse of the measurement result, $-m'$. For large enough squeezing, $s$ becomes large and this envelope has a small impact on the state. For moderate squeezing, however, this can severely damage the state, and the measurement result will now play an important role. Before, the measurement result had no impact other than as a parameter to keep track of to modify future measurement bases. Fortunately, the measurement results will on average tend to be such that the Gaussian envelope is centered in phase space about the input quantum state, which we can determine by calculating the expectation value of the homodyne measurement.

Let's start with input state $|\psi\rangle$, and since the shearing operation only acts to scale the measurement result, we can commute the shearing through to let $|\psi'\rangle = \mathcal{P}|\psi\rangle$. Now, given that the measurement is in the $p$-basis on mode one after the entangling gate, we can use the Heisenberg picture to see that the measured operator before the $\hat{C}_Z$ gate is in fact

$$\hat{P}_1' = \hat{C}_Z^\dagger \hat{P}_1 \hat{C}_Z = \hat{P}_1 + \hat{Q}_2. \tag{1.244}$$

With input state of $|\psi'\rangle_1$ and $\hat{S}_2(\iota)|0\rangle_{N_2}$, we find that the expectation value is given by

$$\begin{aligned}\langle \hat{P}_1' \rangle &= {}_1\langle\psi'|\hat{P}_1|\psi'\rangle_1 + {}_{N_2}\langle 0|\hat{S}_2^\dagger(\iota)\hat{Q}_2\hat{S}_2(\iota)|0\rangle_{N_2} \\ &= {}_1\langle\psi'|\hat{P}_1|\psi'\rangle_1.\end{aligned} \tag{1.245}$$

Thus the distribution of measurement results will be centered about the input state. One might notice that the expectation value here was in the $p$-basis, but the Gaussian filter is on the $q$-quadrature. However, the filter will still be centered since the state is also rotated with a Fourier transform[17].

[17] Or, alternatively, commuting $\hat{F}$ to the back of the filter shows that the filter is applied to the $p$-quadrature, and then the filtered state is rotated.

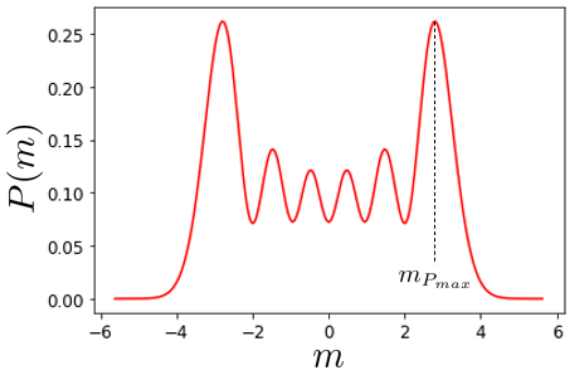



FIGURE 1.14: Potential measurement result distribution obtained by teleporting a five-photon Fock state on a finitely squeezed cluster state. This figure is obtained when the ancilla qumode is a state with 9 dB of quadrature squeezing.

It is also useful to look at the variance of the measurement result, which is given by

$$\begin{aligned}(\Delta P_1')^2 &= \langle(\hat{P}_1 + \hat{Q}_2)^2\rangle - \langle\hat{P}_1 + \hat{Q}_2\rangle^2 \\ &= \langle\hat{P}_1^2\rangle + \langle\hat{Q}_2^2\rangle + 2\langle\hat{P}_1\hat{Q}_2\rangle - \langle\hat{P}_1 + \hat{Q}_2\rangle^2. \end{aligned} \tag{1.246}$$

Since $\langle\hat{Q}_2\rangle = 0$, this reduces to

$$(\Delta P_1')^2 = (\Delta P_1)^2 + \tfrac{1}{2}e^{2r}, \tag{1.247}$$

which is simply the sum of the $p$-variance of the state $|\psi'\rangle$ and the $q$-variance of squeezed vacuum.

Because of this finite-squeezing induced envelope, all gates applied can only approximately be unitary, and the build-up of Gaussian noise after several teleportation steps can cause errors in the calculation. This is where the need for error correction comes into play. Unfortunately, Gaussian resources cannot be used to correct for Gaussian noise — any attempts to do so will actually make things worse [67]. Thus, we must find a way to perform error correction using non-Gaussian resources, which is the subject of Sec. 1.10.1.

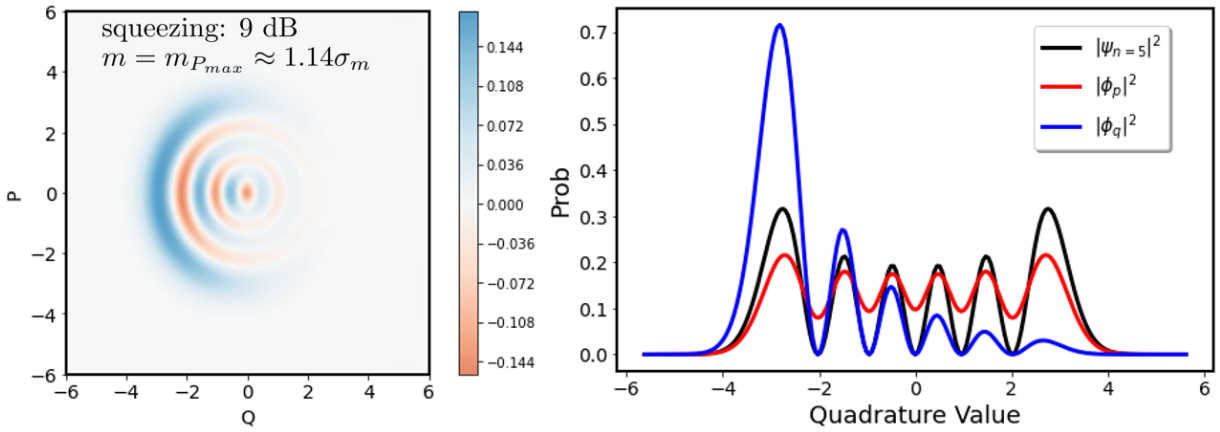



FIGURE 1.15: Wigner function and quadrature probabilities, $|\psi_q|^2$ and $|\psi_p|^2$, for a five-photon Fock state after a cluster state teleportation. The ancilla state was squeezed by 9 dB, and the measurement result was chosen the result with the maximum probability, $m_{P_{max}}$ in Fig. 1.14.

**Fock state example**

To see how finite squeezing effects the teleportation of quantum states, let's consider an example with a Fock state. Suppose we take an input state that is a Fock state, $|n\rangle_N$, and send it through the teleportation circuit shown in Eq. 1.242 with a $p$-basis measurement ($\theta = 0$). Since the Fock state is invariant under rotations, the output state after applying the operator in Eq. 1.243, disregarding the $\hat{X}(m)$ shift, will be

$$|\psi\rangle_{out} \propto \int dx\, e^{-\frac{1}{2s^2}(x+m)^2} e^{-\frac{x^2}{2}} H_n(x)|x\rangle_q. \tag{1.248}$$

To get a better feel for what happens, let's take the five-photon Fock state as the input $n = 5$, and suppose the squeezing in the cluster state is 9 dB at each node, so $s \approx 2.82$. In this case, the likely measurement outcomes are given by the probability distribution shown in Fig. 1.14. This distribution has a distinct resemblance to the Fock state distribution by itself, but it has be broadened and smoothed by the contribution from the $q$-quadrature distribution of the squeezed ancilla.

This distribution has a standard deviation of $\sigma_m \approx 2.5$, but let's let $m$ be the most likely measurement of $m_{P_{max}} \approx 2.815$, which is just a bit larger than $1\sigma$. If this is the case, then we can plot the Wigner function and quadrature probability distributions for the output state, which are shown in Fig. 1.15. As we can see, the Wigner function and $q$-distribution fade out on the right side of the plots, which is not surprising

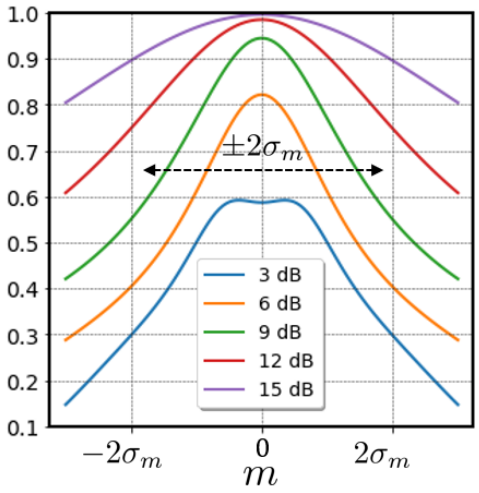



FIGURE 1.16: Fidelity of the teleported Fock state with the ideal $|5\rangle$ state as squeezing in the ancilla mode and detection results vary. Because the detection results are dependent on the squeezing level in the cluster, the x-axis in the figure is scaled linearly with the standard deviation of detection results, $m$, for each curve at a given squeezing.

considering that the Gaussian filter is centered about $-m$. The other quadrature, however, is not filtered, but broadened, and remains symmetric. This broadening is due to taking a Fourier transform of Eq. 1.248, which essentially means that the $p$-quadrature distribution of the original Fock state (black curve in the figure) is convolved with the finite-squeezing envelope.

In the ideal case, this state should be exactly the same as our input state, but it is clearly not. To characterize the difference, we plot the fidelity of the input five-photon Fock state with the filtered output in Fig. 1.16 over a range of detection outcomes and at several squeezing levels. Because the probability distribution for the measurement outcomes is dependent on the squeezing, the x-axis for the plot is scaled in terms of the standard deviation of the measurement outcome probabilities, $\sigma_m$. For a $m = 0$ measurement outcome, the teleportation fidelity is quite good for even moderate levels of squeezing. However, the only case where the fidelity remains above 0.9 over the entire $\pm 2\sigma_m$ range is the case of 15 dB squeezing.

## 1.10 Error Correction and State Encoding

Due to finite squeezing and noise in any physical system, continuous-variable QC, like all forms of realistically viable QC, will need a way to implement error correction. Error correction with qubits comes in several forms, but the general idea is to encode quantum information into a subspace of the entire Hilbert space available. This way, the remainder of the Hilbert space can be used as a resource to track down errors. If we find that our quantum state has deviated from the subspace used for calculations, we know that an error has occurred. Furthermore, if we are clever about how we subdivide the Hilbert space, measurement can be used to both determine that an error has occurred and specifically, which error has occurred without perturbing the quantum information we care about.

An example of this is the bit-flip protection algorithm that uses redundancy encoding with each logical qubit being composed of several physical qubits

$$|0\rangle \rightarrow |0_L\rangle \equiv |000\rangle \tag{1.249}$$

$$|1\rangle \rightarrow |1_L\rangle \equiv |111\rangle\,, \tag{1.250}$$

so the general logical qubit would be

$$|\psi_L\rangle = \cos(\theta/2)\,|000\rangle + \sin(\theta/2)e^{i\varphi}\,|111\rangle\,. \tag{1.251}$$

If we ever find that the physical qubits do not all agree, then we know that an error has occurred. Unlike classical computing, we *cannot* check the state to see the values of the individual qubits without destroying the encoded information, but we *can* measure whether the physical qubits are all the same without finding out what they are, and furthermore, we can measure which qubits agree. Supposing that they didn't agree upon measuring and assuming a low enough error rate, it is likely that only one qubit flipped so we can flip the disagreeing qubit back. There has been a great deal of work done on error correction with qubits, and an introduction to the field can be found in Ref. [5].

Continuous-variables quantum states are substantially different from qubits due to the continuous nature of the Hilbert space. However, we can leverage similar ideas to qubit-based schemes and attempt to encode information in a subspace of the entire Hilbert space. The prevailing method of doing this is to section the continuous space

into discrete sections which effectively amounts to having a logical computation space of qubits or qudits embedded within the larger continuous Hilbert space. The first proposal of doing so was aptly called "Encoding a qubit in an oscillator" and works by separating the Hilbert space into two quadrature 'combs' that act as orthogonal qubits known as Gottesman-Kitaev-Preskill (GKP) states [48]. An alternative method, which has recently been extended, is to build logical states with a particular symmetry in phase space [68].

These specific encodings are designed to optimize error correction against certain types of errors - GKP states are adept at correcting displacement errors while rotation-symmetric codes are optimized to correct against photon loss and optical dephasing. With any type of qudit encoding, one can take the discrete QC universal gate set and map it onto the physical system used. It is still true that a universal gate set will necessitate non-Gaussianity, but specifically how this comes in can be more easily determined. Interestingly, in the case of GKP qubit encoding, the universal gate set can be achieved with additional resources being all Gaussian, as the GKP states themselves have the requisite non-Gaussianity [69]. As a side point, Gaussian operations together with the cubic phase gate will remain sufficient for universality in any encoding[18]. However, this may not be the easiest or best route, as attempting to combine cubic phase gates with the GKP encoding would spell disaster for error-correctability [70].

Before proceeding to describe these states, it is worth making two additional points. The first is that these encodings are more than simply mapping CVQC onto discrete QC. These logical encodings may utilize discrete quantum information, but they are still embedded into a continuous physical space. This imparts some benefits, especially against continuous errors. The second and perhaps more important point is that all encodings are compatible with the CV cluster state where the majority of nodes are still just finitely squeezed states. This is because the formalism introduced earlier using teleportation to apply gates is agnostic with respect to the input state, regardless of its encoding.

### 1.10.1 GKP states

The GKP states are a family of bosonic code states that attempt to discretize phase space using Dirac combs. These states were first proposed in Ref. [48], but there has

[18] This is because these ingredients allow one to enact any arbitrary CV Hamiltonian.

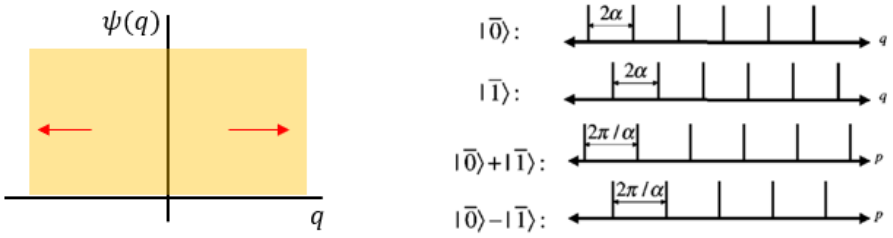



FIGURE 1.17: Phase-space can be discretized by carving out specific logical states in the form of combs from the continuous-variable space. An example of this is shown at right for a qubit encoding, taken from Ref. [48].

been a considerable amount of effort extending these codes in recent years. Refs. [71–73] present good overviews and representations of recent progress.

The ideal states are depicted in Fig. 1.17, where the $q$-basis wavefunction is a series of delta functions with spacing $2\alpha$. The point of these codes is for each logical state to be an eigenstate of translations in both quadratures simultaneously, since the Fourier transform of a comb is yet another comb. The ideal choice for error correction is to choose $\alpha = \sqrt{\pi}$, which ensures that the combs are symmetric and are eigenstates of $\hat{X}(2\sqrt{\pi})$ and $\hat{Z}(2\sqrt{\pi})$. One can define a logical encoding where $|0_L\rangle$ is the qubit with the central delta function corresponding to the origin and $|1_L\rangle$ is a comb shifted by half the comb spacing,

$$|\mu_L\rangle = (2\sqrt{\pi})^{1/2} \sum_{m=-\infty}^{\infty} |(2m+\mu)\sqrt{\pi}\rangle_q \tag{1.252}$$

$$= (\sqrt{\pi})^{1/2} \sum_{m=-\infty}^{\infty} e^{ij\pi m} |m\sqrt{\pi}\rangle_p, \tag{1.253}$$

where $\mu \in \{0,1\}$. With this, one can adopt qubit-based formalism, like the $\pm$ basis with $|\pm_L\rangle = \frac{1}{\sqrt{2}}(|0_L\rangle \pm |1_L\rangle)$. Additionally, we have that

$$\hat{X}(\sqrt{\pi})\,|0_L\rangle = |1_L\rangle \rightarrow \hat{X}(\sqrt{\pi}) \equiv \hat{X}_L, \tag{1.254}$$

$$\hat{Z}(\sqrt{\pi})\,|+_L\rangle = |-_L\rangle \rightarrow \hat{Z}(\sqrt{\pi}) \equiv \hat{Z}_L, \tag{1.255}$$

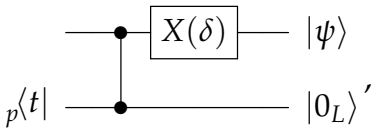
$$\hat{R}(\tfrac{\pi}{2})\,|0_L\rangle = |+_L\rangle \rightarrow \hat{R}(\tfrac{\pi}{2}) \equiv \hat{H}_L. \tag{1.256}$$

These, along with the unit shear (logical phase gate), $\hat{S}_L \equiv e^{i\frac{1}{2}\hat{Q}^2}$, give the correspondence of a logical gate set for single-qubit Pauli and Clifford operations, $\{\hat{X}_L, \hat{Z}_L, \hat{H}_L, \hat{S}_L\}$.

As one can immediately see, any superposition of these logical states only ever has support in phase space on integer multiples of $\sqrt{\pi}$, and each state is periodic with period $2\sqrt{\pi}$. Because of this, any small shift in either position or momentum of less than $\frac{\pi}{2}$ can be corrected. If one performs a modulo $\sqrt{\pi}$ quadrature measurement and finds that the remainder is anything other than zero, then one can be sure that an error has occurred. If this error is less than $\frac{\pi}{2}$, then one can also know the sign of the quadrature error — otherwise one doesn't know which direction the state should be shifted to perform a correction.

**Error-correction protocol**

Suppose that we use the GKP states for our encoding — this means that the two logical states will comprise our computational basis, and any $|\psi\rangle$ will only have support on quadratures values that are multiples of $\sqrt{\pi}$ in the absence of errors. Suppose this state obtains some $q$-quadrature shift error, $\delta$, which we can model by applying $\hat{X}(\delta)$. We need a way to determine if this $\delta$ shift happened, without performing a measurement that would collapse our quantum state. To do this, we can entangle $|\psi\rangle$ to an ancillary GKP state and perform a homodyne detection on the ancilla. Because the ancilla is a GKP state with phase-space support only on integer multiples of $\sqrt{\pi}$, we know that any remainder from the homodyne measurement modulo $\sqrt{\pi}$ must originate from the shift error. We can then use this information to apply the opposite shift and fix our state. This error-correction circuit is described by [48]




(1.257)

where the error, $\hat{X}(\delta)$, is included in the circuit.

This circuit is worth deriving in detail. First, consider a pure input state before the error in the form of the GKP encoding

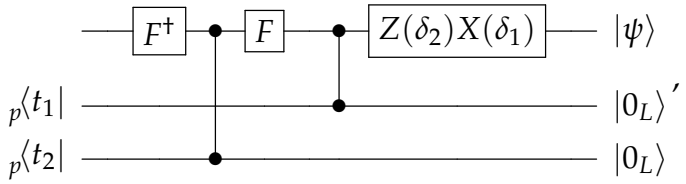
$$|\psi\rangle = a\,|0_L\rangle + b\,|1_L\rangle\,. \tag{1.258}$$

We can determine the action of this circuit on a single component of state and then extend the result to the general superposition, $|\psi\rangle$. For logical qubit $|\mu_L\rangle$, the circuit

acts as:

$$
\begin{aligned}
&{}_{p_2}\langle t|e^{i\hat{Q}_1\hat{Q}_2}\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}|(2m+\mu)\sqrt{\pi}+\delta\rangle_{q_1}|n\sqrt{\pi}\rangle_{p_2}\\
&=\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}|(2m+\mu)\sqrt{\pi}+\delta\rangle_{q_1}{}_{p_2}\langle t|(n+2m+\mu)\sqrt{\pi}+\delta\rangle_{p_2}. \qquad (1.259)
\end{aligned}
$$

Since $n$, $m$, and $\mu$ are all integers, $t = k\sqrt{\pi}+\delta$. Thus, we can do one of two things. We can either apply a shift of $\hat{X}^\dagger(\mathrm{mod}(t,\sqrt{\pi}))$ to undo just the error portion, $\delta$, or we can apply the shift $\hat{X}^\dagger(t)$, which still fixes the error but may also require the application of an additional logical gate $\hat{X}(\pm\sqrt{\pi})$ [69]. Since the circuit in Eq. 1.257 acts the same way on both GKP states $|\mu_L\rangle$, it can be used to correct an arbitrary encoded state, $|\psi\rangle$. This error correction scheme is ideally suited to cluster states, as the $\hat{C}_Z$ gate and homodyne measurements are already native operations. In this way, one can embed GKP states in the cluster state and perform error-correction [74]. Furthermore, by modifying the above circuit slightly, one can instead teleport through GKP resources, which allows for enacting Gaussian gates and quantum error-correction all in one step, which is compatible with realistic macro-node clusters [18]. The error-correcting circuit in Eq. 1.257 only corrects small shift errors in $q$, so a second circuit must be used to correct errors in $p$. The full error-correction can be done with the circuit



(1.260)

where the correction that still need be applied takes the form of $\hat{Z}^\dagger(t_2)\hat{X}^\dagger(t_1)$, up to a logical operation. Remember, these corrections need not actually be applied with physical displacements, as they can be taken into account with feed-forward changes to future homodyne measurement bases. Although the GKP encoding is designed to protect against errors in the form of quadrature shifts, all sufficiently small errors can be expanded in terms of shift errors, since the Pauli operators form a complete basis [48].

### Finite squeezing

The states in Eq. 1.252 extend infinitely in phase space and are unphysical. We can, however, acquire physically relevant approximations to the ideal GKP states by using

a series of narrow-width Gaussian peaks in place of delta functions. There are several different equivalent ways to arrive at approximate GKP state [75], but one can be obtained from the ideal case simply by applying the damping operator,

$$|\tilde{\mu}_L\rangle \propto e^{-\Delta^2\hat{a}^\dagger\hat{a}}\,|\mu_L\rangle\,. \tag{1.261}$$

Defined this way, the approximate code states are $+1$ eigenstates of the nonunitary operators

$$\hat{X}^\Delta = e^{-\Delta^2\hat{a}^\dagger\hat{a}}\hat{X}_L^2 e^{\Delta^2\hat{a}^\dagger\hat{a}}, \tag{1.262}$$

$$\hat{Z}^\Delta = e^{-\Delta^2\hat{a}^\dagger\hat{a}}\hat{Z}_L^2 e^{\Delta^2\hat{a}^\dagger\hat{a}}, \tag{1.263}$$

where logical operators in the ideal codespace can be transformed to $\hat{O}_L^\Delta = e^{-\Delta^2\hat{a}^\dagger\hat{a}}\hat{O}_L e^{\Delta^2\hat{a}^\dagger\hat{a}}$ where $\hat{O}_L = \hat{X}_L, \hat{Z}_L$ [76]. The approximate states $|\mu_L\rangle$ can now be seen as a superposition of displaced squeezed states with an additional Gaussian envelope. An example is depicted in Fig. 1.18, where the top panel shows the Wigner functions for the two logical states, $|\tilde{0}_L\rangle$ and $|\tilde{1}_L\rangle$ with peak width $\Delta = 0.3$. The $|\tilde{\pm}\rangle$ states can be obtained from this simply by rotating the Wigner function through a $\frac{\pi}{2}$ phase. The bottom panel shows the quadrature probability distributions, $|\psi(q_\mu)|^2$ for the two states. Note that the $p$-quadrature distribution for the two states only differs by a complex phase in Eq. 1.252, so the probabilities displayed are the same for both logical state.

Because finite squeezing gives an envelope to the ideal states, the two logical states are no longer orthogonal. The inset in Fig. 1.18 shows the fidelity $|\langle\tilde{0}_L|\,\tilde{1}_L\rangle|^2$ between the two logical states as a function of $\Delta$. For the value of $\Delta = 0.3$ shown in the rest of the figure, the state overlap is already zero with an error less than $10^{-6}$. If one thinks of realistic GKP states as a superposition of displaced squeezed states, one can relate the peak width $\Delta$ to the squeezing in dB through the formula

$$\mathcal{S}_{dB} = -10\log_{10}(\Delta^2). \tag{1.264}$$

Even in the case of finite squeezing, an error-correction circuit of the form of Eq. 1.260 can still be applied; there will just be some remaining residual noise. If the squeezing is large enough such that the residual error can be made arbitrarily small through repeated application of error-correction with noisy states, we say that the error-correcting

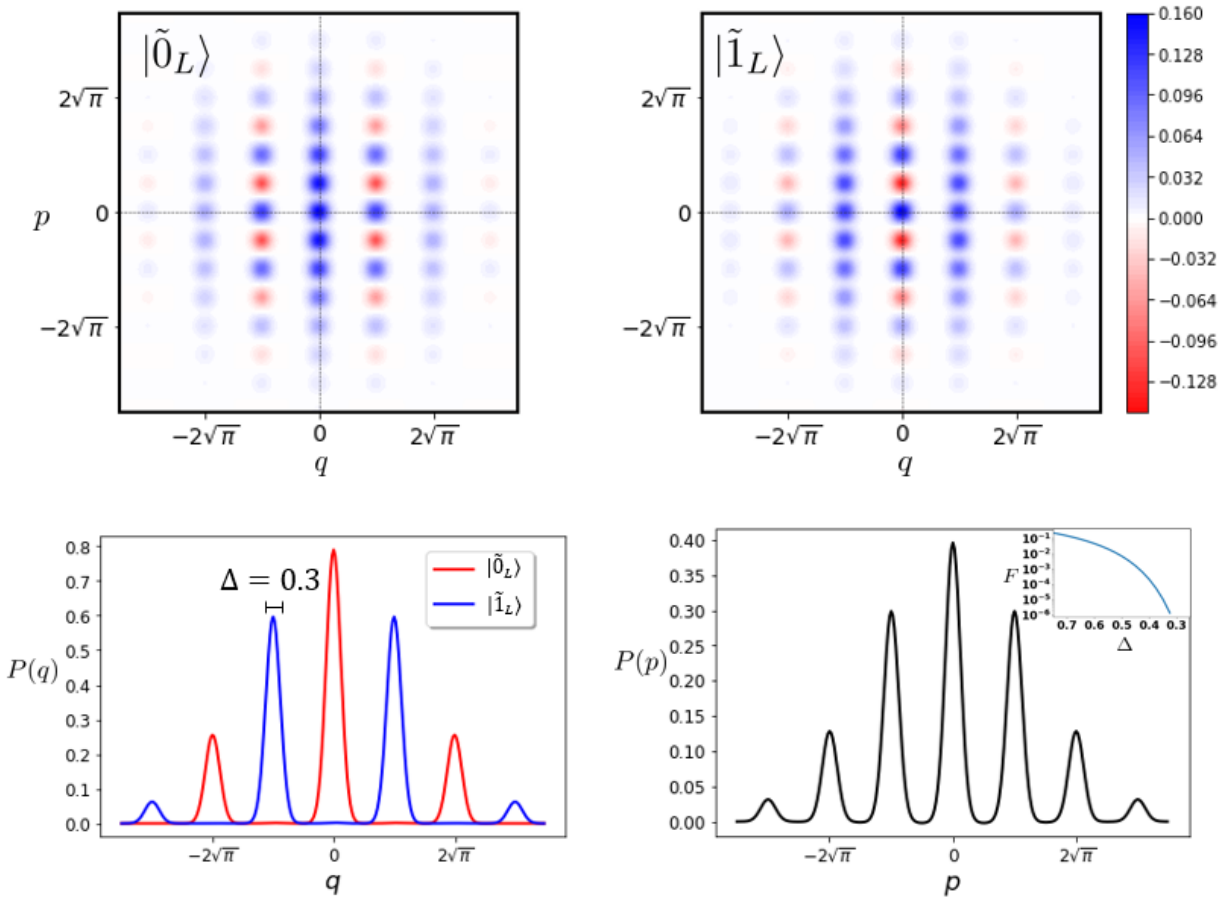



FIGURE 1.18: Top: wigner function for logical GKP states with peak width $\Delta = 0.3$. Bottom: probability distribution for the above states, where the $p$-quadrature distribution is identical for both. The inset shows the fidelity between the logical states, which tends to zero as $\Delta$ decreases.

code is fault-tolerant. Fault-tolerant thresholds depend on the specific types of error-correcting code used since the CV error-correction schemes will also need to be combined, or concatenated, with other qubit-based error-correction to correct for logical errors. It turns out that cluster-state QC with GKP states can be made fault-tolerant with finite squeezing [74], and these required squeezing levels have been reached [14]. Furthermore, this threshold has been reduced for certain encoding schemes [64, 66, 77], and in the presence of anti-squeezing in excess of the squeezing [78]. The current fault-tolerant threshold lies below 10 dB of squeezing [79, 80].

An important point to emphasize is that these fault-tolerant thresholds are all for a qubit encoding, so the quantum information being manipulated must be encoded in GKP logical states. Efforts have also been made to develop GKP error correction for general CV states, where the GKP ancilla qubits are used to correct errors in a state within the full CV Hilbert space [81]. Unfortunately, fault-tolerance in this context is ill-defined, as even small errors may be difficult to suppress for complex non-Gaussian states.

Yet another subtle but important point to note is the relationship between fault-tolerant thresholds of GKP squeezing and the levels of experimentally achieved vacuum squeezing. Although the threshold for fault-tolerance of 10 dB has certainly been well-surpassed in vacuum squeezing [14], and quality of vacuum-squeezed modes will very much matter in terms of the amount of Gaussian noise added at a given teleportation step when applying gates in the cluster state formalism, 10 dB of vacuum squeezing does not equate to 10 dB of GKP squeezing. Methods to synthesize GKP states will be discussed later in Chs. 2 and 3, but the majority of optical methods require some squeezed resources, whether it be to generate the supply of precursor cat states which can be bred into GKP states [82–84][19], or as the Gaussian resource input to a Gaussian boson sampling (GBS) device [71, 85, 86]. When considering cat state breeding, the input cat states must have sufficient squeezing *and* amplitude to eventually produce the target GKP. Although the process of photon subtraction from squeezed vacuum produces squeezed cat states with roughly the same squeezing, producing a large enough cat state requires either iterated photon subtraction or the subtraction of a large number of photons — both of which will degrade the squeezing below the initial squeezed

[19] Photon catalysis, as presented in Ch. 2, is an exception as it requires no squeezed resources yet still generates squeezed cat states. However, this does require a supply of single photons, and the output squeezing is weak which indicates further in-line squeezing may be needed to reach fault-tolerant GKP states.

vacuum resource. In the case of GBS devices, the best current parameters for GKP state synthesis could produce a 10 dB-squeezed GKP resource with $> 99\%$ fidelity, but this requires a 12 dB-squeezed vacuum resource and assumes lossless detection [71]. When including as little as 0.02 dB of loss ($\approx 0.5\%$), the fidelity is reduced to just above 80%, which will likely not suffice for error correction. A recent take on iterating layers of GBS devices claims to be able to synthesize fault-tolerant GKP states [86], but without the use of quantum memories (which would further degrade a produced state), the overall success probabilities are on the order of $10^{-29}$. Additionally, the parameters for the Gaussian resources that optimize this GBS, including squeezing, are not provided. All this together indicates that the jury is still out on the best way to produce optical GKP states. Furthermore, one should use caution when relating experimentally achieved vacuum squeezing to any fault-tolerant threshold based on GKP state squeezing.

**Experimental progress**

GKP states have been demonstrated experimentally with circuit QED and trapped ion systems [87–89], but an all-optical approach with traveling fields has remained elusive. Many approaches have been proposed for creating and enlarging cat states and GKP states optically [71, 82, 84, 85, 90–95], with the majority of these protocols utilizing photon subtraction to generate cats while GKP states are formed by using higher-order nonlinearities, non-deterministic boson-sampling devices, or by breeding pairs of cat states. In spite of the difficulty to demonstrate GKP states optically, several approaches have successfully generated low-amplitude cat states [96–102]. Unfortunately, all approaches for cat state generation currently rely on nondeterministic, probabilistic methods with success rates below 10% [93] which demand "repeat-until-success" approaches [103, 104].

Additionally, it is desirable to have these non-Gaussian states somehow integrated within the Gaussian cluster state to utilize the fully measurement-based advantages of one-way QC. Current proposals plan to do by generating non-Gaussian states offline and later coupling the ancillary modes to the cluster [65, 104, 105]. This method could potentially introduce further errors with the need to perform entangling operations between the offline GKP resources and the cluster, and additionally non-deterministic state generation may complicate matters with requirements for quantum memories if many copies of GKP states are needed. To help address these difficulties,

Ch. 3 proposes a method to generate Schrödinger cat and GKP states embedded within a cluster state.

### 1.10.2 Rotation symmetric codes

The GKP encoding is only one of several types of bosonic encodings [68, 106–108]. Instead of discretizing phase-space through periodic repetition of $q$ or $p$, one can instead use number-phase symmetry, such as introduced in Sec. 1.4.5. There are actually many classes of codes that can be developed in this way, and unlike the GKP states, ideal rotation-symmetric code states are physically realizable. However, the universal gate set is generally more complicated. Rotation-symmetric codes remain a viable option and would outclass GKP code error-correction against photon loss under certain conditions [106].

In order for a code to successfully correct errors, the code states must satisfy the Knill-Laflamme conditions [109]. For a given set of operators describing the interaction with the environment, $\mathcal{E} = \{\hat{E}_1, \hat{E}_2, ..., \hat{E}_k\}$ such that $\sum_k \hat{E}^\dagger \hat{E} = \mathbb{1}$, a necessary and sufficient criterion for a faithful error correction is defined as:

$$\langle 0_L | \hat{E}_j^\dagger \hat{E}_k | 1_L \rangle = 0 \tag{1.265}$$

$$\langle 0_L | \hat{E}_j^\dagger \hat{E}_k | 0_L \rangle = \langle 1_L | \hat{E}_j^\dagger \hat{E}_k | 1_L \rangle . \tag{1.266}$$

These conditions essentially state that if an error occurs on both or either logical state, then they must remain orthogonal. Additionally, if an error occurs, it should modify the coefficients of the logical states in the same way. This way, when we determine that an error has occurred through syndrome measurement, the recovery operation will correct the state without altering the superposition.

#### Binomial code states

One class of states that I will highlight here are the binomial code states, defined as [21]

$$|\mu\rangle = \frac{1}{\sqrt{2^M}} \sum_{p\,even/odd}^{[0,M+1]} \binom{M+1}{p}^{1/2} |p(L+G+1)\rangle_N . \tag{1.267}$$

Depending on the particular parameters, these codes can correct up to $L$ photon loss, $G$ photon gain, and $D$ dephasing errors, where $M = \max\{L, G, 2D\}$. These code states satisfy the Knill-Laflamme conditions for errors in the set

$$\mathcal{E} = \{\hat{\mathbb{1}}, \hat{a}, \hat{a}^2, ..., \hat{a}^L, \hat{a}^\dagger, ..., \hat{a}^{\dagger G}, \hat{a}^\dagger \hat{a}, ..., (\hat{a}^\dagger \hat{a})^D\}. \tag{1.268}$$

An example of these states for $L = 1$, $G = 0$, and $D = 1$ is shown in Fig. 1.3. A more detailed discussion on these states and potential ways to generate them will be found in Ch. 4.

# Chapter 2

# Photon Catalysis

From Ch. 1, we saw that obtaining a universal gate set for CV quantum information processing requires non-Gaussian elements with signatures of negativity in the Wigner function. Wigner function negativity is a required resource for achieving an exponential speedup [10], for entanglement distillation [110], Bell inequality violation [35], and for quantum error correction [67]. Because of this, it is very important to develop techniques to engineering quantum states and measurements with the requisite non-Gaussianity. Although non-Gaussianity is necessary, it is not sufficient, which means that finding a prescribed way of introducing the vital 'quantum-ness' may vary widely between use cases for different encoding schemes. In addition to the limited availability of non-Gaussian experimental resources, this makes developing universal resources difficult, but nonetheless, quite important.

This chapter will develop methods of engineering certain non-Gaussian resources using experimental techniques that are realistic for near-term applications. The main resource of non-Gaussianity will be photon-number-resolving (PNR) measurements, which I will show is quite feasible experimentally in Ch. 6. Additionally, for the photon catalysis scheme, I will make use of other limited non-Gaussian resources such as single-photons in addition to Gaussian resources in the form of displacements and rotations. This chapter will be based on published work on engineering cat and GKP states with photon catalysis [84].

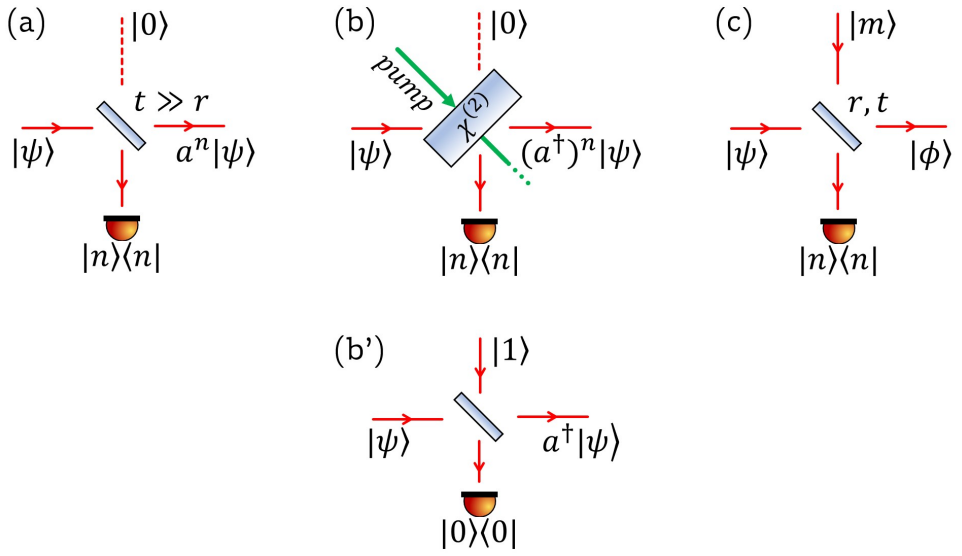



FIGURE 2.1: Various feasible techniques for generating quantum states with non-Gaussian Wigner functions. Diagram (a) depicts $n$ photon subtraction, (b) and (b') denote variants of $n$ and single-photon additions, respectively, and (c) generalizes the previous cases for an $m$ fock state with arbitrary beampslitter parameters and $n$ photon detection.

## 2.1 Overview

Single-photon states, which have been generated and characterized using heralding detection of downconverted photon pairs [111–114], are excellent non-Gaussian ingredients. Sophisticated techniques such as photon subtraction [90, 91] as discussed in Ch. 1 and photon addition [115] have led to very promising advances. In photon subtraction, Fig. 2.1(a), a nonclassical state of light impinges onto a highly unbalanced, say transmissive, beamsplitter and a photon-number-resolving (PNR) detector, ideally, or at least a single-photon sensitive one, detects the reflected light. Conditioned on the detection of $n$ photons, one can show that the transmitted light is, to a good approximation, the state $a^n |\psi\rangle$, where $\psi$ denotes the initial state of light and $a$ the photon annihilation operator [91, 116]. This was recently generalized to multimode light [117]. Photon subtraction cannot work on coherent states, of course, but photon addition, Figures(b,b'), does [115]. In Fig. 2.1(c), a more general process is presented, called photon catalysis [118]. The Kraus operators for photon addition and subtraction were given in Sec. 1.5.4

The more general technique of photon catalysis is derived from the "quantum

scissors" scheme [119] and consists in interfering a quantum input (not necessarily pure) with a Fock state and performing PNR detection of one beamsplitter output. The beamsplitter is no longer necessarily balanced, and its coefficients are parameters of the process as well. Bartley et al. explored this process and showed its potential for the creation of non-Gaussian coherent state superpositions and squeezed states [120]. By varying the parameters of the beamsplitter, Bartley et. al experimentally performed single-photon catalysis on a coherent state input to demonstrate the creation of states exhibiting nonclassical photon statistics. Photon catalysis has also been explored in utilizing multi photon detection [121], with general single-mode Gaussian inputs [122], and for use as an entanglement enhancer [123]. Note that near-unity efficiency PNR is now an experimentally available resource, thanks to superconducting transition-edge sensors [124].

This chapter will explore the use of single-step and sequential photon catalysis to generate non-Gaussian states of interest to CVQC, in particular exact displaced single-photon states and phase-symmetric superpositions of squeezed vacuum (SSV) states, of which squeezed Schrödinger cat states are a subset. By including a PNR-based breeding protocol, we demonstrate the potential to enlarge SSV states and generate GKP states. In this method, no squeezed states are required, the needed resources being coherent states and linear optics, single-photon states, and PNR detection. Note that photon catalysis can be related to a generalized Hong-Ou-Mandel effect, the result of which has recently been calculated for arbitrary $m$ if the input state, $|\psi\rangle$ in Fig. 2.1, is either a coherent state or a thermal state [125].

## 2.2 Derivation

For the remainder of this work, we take the restriction of photon catalysis in Fig. 2.1(c) to $m$=1, i.e., to single-photon input resource states, but we keep the option of variable $n$ for PNR detection. For an arbitrary input mode $|\psi\rangle_a|1\rangle_b = \sum_{m=0}^{\infty}\psi_m|m\rangle_a|1\rangle_b$, the output state is [19, 126]

$$|\text{out}\rangle_{ab} = \sum_{m=0}^{\infty}\psi_m\sum_{k=0}^{m}\binom{m}{k}^{\frac{1}{2}}r^{m-k}t^k\left[t\sqrt{m-k+1}|m-k+1\rangle_a|k\rangle_b - r\sqrt{k+1}|m-k\rangle_a|k+1\rangle_b\right], \tag{2.1}$$

where the beamsplitter operator is defined by $U_{ab} = \exp[\theta(ab^\dagger - a^\dagger b)]$ with the reflection and transmission coefficients being $r = \cos\theta$ and $t = \sin\theta$. If, say, output $a$ is sent to a PNR detector which measures $n$ photons, then output $b$ is projected into the state

$$\begin{aligned} |\phi\rangle_b &= {}_a\langle n|\, out\,\rangle_{ab} \\ &= \sum_{\ell=0}^{\infty} \frac{\psi_{\ell+n-1}}{\sqrt{n}} \binom{\ell+n-1}{\ell}^{\frac{1}{2}} r^{n-1} t^{\ell-1} \left(nt^2 - \ell r^2\right) |\ell\rangle_b \,. \end{aligned} \tag{2.2}$$

If the beamsplitter is designed so that destructive quantum interference $nt^2 = \ell r^2$ occurs, then the corresponding Fock state $|nt^2/r^2\rangle$ is absent from the output. An application of this situation is given in the next section on exact state displacement by such Fock-state filtering. Note that the state amplitudes are shifted by $n-1$ in the process. When postselecting on $n = 1$, this shift disappears and, by setting $r^2 = 1/(q+1)$ for the beamsplitter, one can remove the $q$-photon amplitude:

$$|\phi_{\bar{q}}\rangle_b = \sum_{\ell=0}^{\infty} \psi_\ell \left(\frac{q}{q+1}\right)^{\frac{\ell+1}{2}} \left(1 - \frac{\ell}{q}\right) |\ell\rangle_b \,. \tag{2.3}$$

If $|\psi\rangle$ has a maximum number $\ell_{\max}$ of amplitudes and if $q \gg \ell_{\max}$, then the Fock-state filtering is almost perfect: only $\psi_q$ is removed and the other amplitudes are practically unchanged.

However, the number of free parameters allows us to Fock-filter a state in several different ways. Rather than postselecting $n = 1$ and tuning the beamsplitter, it would be much more advantageous, from an experimental point of view, to postselect as little as possible and use PNR detection to the fullest. Let us also fix $r = t = \frac{1}{\sqrt{2}}$, then

$$|\phi\rangle_b \propto \sum_{\ell=0}^{\infty} \psi_{\ell+n-1} \binom{\ell+n-1}{\ell}^{\frac{1}{2}} 2^{-\frac{\ell}{2}} (n-\ell) |\ell\rangle_b \,, \tag{2.4}$$

thereby removing the $n$-photon Fock state from the $b$ output if $n$ photons are detected in port $a$. Again, one should keep in mind that the state amplitudes $\psi_k$ are shifted by $n-1$ in the process, hence this operation of photon catalysis is more complex than just Fock-state filtering. However, as we now show, photon catalysis can nontrivially generate *exact* displaced single-photon states, as well as arbitrarily good approximations of non-Gaussian states that are of interest in quantum information processing, such as

Schrödinger-cat states and GKP quantum error code resource states. In all the rest of the paper, we will consider either a coherent-state input, $|\psi\rangle = |\alpha\rangle$, or inputs derived from previous photon catalysis steps, $|\psi'\rangle = |\phi\rangle$.

To practically apply photon catalysis to arbitrary quantum states, we developed a numerical procedure to take an input density matrix, possibly impure, and transform it by photon catalysis. These calculation details are described in AppendixA.1.

## 2.3 Exact Displaced Single-photon States

The use of phase space displacements is very important for implementing a variety of operations in continuous-variable quantum information, since the Weyl-Heisenberg group of field quadrature shifts is the CV analog of the Pauli group for qubits [10]. These displacements can be experimentally realized by combining the state to be displaced with a coherent state at a highly unbalanced beamsplitter [127]. However, this method does require a partial trace over the other output port, yielding a statistical mixture that only approaches the exact displaced state as the reflectivity approaches unity, which in turn limits the amount of displacement. Moreover, as discussed in [128], the displacement operation becomes less accurate as the energy constraint for the input states increases. By contrast, photon catalysis is not subject to this limitation and can generate an exact displaced single-photon state.

### 2.3.1 Lossless case

We recall the derivation to approximately displace a single-photon where we use $|\alpha\rangle\,|1\rangle$ as the input [127]. The photon catalysis output is, before PNR detection,

$$|\text{out}\rangle_{ab} = D_a(r\alpha)D_b(t\alpha)(ta^\dagger - rb^\dagger)\,|0\rangle_a\,|0\rangle_b\,. \tag{2.5}$$

Performing a partial trace over mode $a$ results in a mixture of displaced single-photon and weakened coherent given by

$$\rho_{out} = t^2\,|t\alpha\rangle\,\langle t\alpha| + r^2 D(t\alpha)\,|1\rangle\,\langle 1|\,D^\dagger(t\alpha). \tag{2.6}$$

In the limiting case of $\alpha \to \infty$ while $t \to 0$, we recover a pure single-photon state displaced by $t\alpha$. In all experimental cases where $\alpha$ is finite and $t > 0$, we see that there is unavoidable mixing with a coherent state of the same amplitude as the desired displacement [127]. Defining the fidelity between $\rho_{out}$ and a target state, $\rho_T$, by

$$F = \left| Tr\sqrt{\sqrt{\rho_{out}}\rho_T\sqrt{\rho_{out}}} \right|^2, \tag{2.7}$$

we see that the fidelity of the approximate displaced photon given with the ideal pure state, $\rho_T = D(t\alpha)\,|1\rangle\,\langle 1|\,D^\dagger(t\alpha)$, is simply $F = r^2$.

However, if we don't perform the partial trace and instead send mode $a$ of Eq. 2.5 to a PNR detector and condition on the detection of $n$ photons, we get

$$|\phi\rangle_b \propto {}_a\langle\, n \,|\, \text{out}\,\rangle_{ab} \tag{2.8}$$

$$\propto (nt - \alpha r^2 b^\dagger)\,|t\alpha\rangle_b\,. \tag{2.9}$$

We can then evaluate the overlap of $|\phi\rangle_b$ with a single-photon state displaced by $\beta$:

$$\langle 1|\,D^\dagger(\beta)\,|\phi\rangle \propto nt(t\alpha - \beta) - \alpha r^2[1 + \beta^*(t\alpha - \beta)]. \tag{2.10}$$

Normalizing Eq. (2.9) and taking $\alpha, \beta \in \mathbb{R}$ yields the fidelity

$$F = e^{-|\beta - t\alpha|^2}\frac{\left|nt(\beta - t\alpha) + r^2\alpha(1 + t\alpha\beta - |\beta|^2)\right|^2}{r^4|\alpha|^2 + t^2(n - r^2|\alpha|^2)^2}. \tag{2.11}$$

Examining this result, we see that if $r^2|\alpha|^2$ is specifically chosen to be an integer so that $n = r^2|\alpha|^2$ photons are detected, then $F = 1$ for $\beta = \sqrt{|\alpha|^2 - n} = \alpha t$. While this result holds for any value of displacement, $\beta$, the specific case of taking an integer value for $|\beta|^2$ allows us to view the method of photon catalysis to enact displacements as a process that removes a single Fock component as per Eq. 2.2. To see this, consider the

displaced single-photon state,

$$D(\beta)\,|1\rangle = D(\beta)a^\dagger D^\dagger(\beta)\,|\beta\rangle \tag{2.12}$$

$$= a^\dagger\,|\beta\rangle - \beta^*\,|\beta\rangle \tag{2.13}$$

$$\propto \sum_{m=1}^{\infty} \frac{\beta^m}{\sqrt{m!}}(m - |\beta|^2)\,|m\rangle - |\beta|^2\,|0\rangle\,. \tag{2.14}$$

Clearly, the $m = |\beta|^2$ Fock-component is absent, which is precisely the component that was removed from the initial coherent state, $\alpha$, by the photon catalysis process, i.e., the component $nt^2/r^2 = |\alpha|^2 t^2 = |\beta|^2$.

Our method gives the same result as the limit case given by Ref. [127]; however, here, the beamsplitter retains $t > 0$ and the limit case is not needed to reach exact displacement. Therefore, we see that by tuning the reflectivity of a beamsplitter and post-selecting on the desired $n$ detection, it is possible to use photon catalysis with a coherent state to prepare an exact displaced single-photon state of displacement amplitude $\beta = t\alpha$.

### 2.3.2 Lossy case

If we now consider an imperfect detector of quantum efficiency $\eta < 1$, the final state is no longer pure and the fidelity with the target displaced Fock-state is no longer unity. However, the fidelity can be improved slightly by modifying the displacement of the target state, $\rho_T = D(\beta)\,|1\rangle\,\langle 1|\,D(\beta)^\dagger$. In this case, we postulate, and numerically verify, that the maximum fidelity is achieved when

$$\alpha = \sqrt{\frac{n}{\eta r^2}} \tag{2.15}$$

$$\beta = \sqrt{\alpha^2 - \frac{n}{\eta}} \tag{2.16}$$

and this value approaches unity as $t \to 0$ ($r \to 1$) as illustrated in Fig.2.2(c), where the Wigner functions for various parameters are evaluated numerically using the open-source Python modules QuTip [129] and Strawberry Fields [130]. In order to verify that the codes perform as expected, we computed the above photon catalysis step using QuTip and confirmed that the output state is identically a displaced Fock state for correctly chosen parameters.

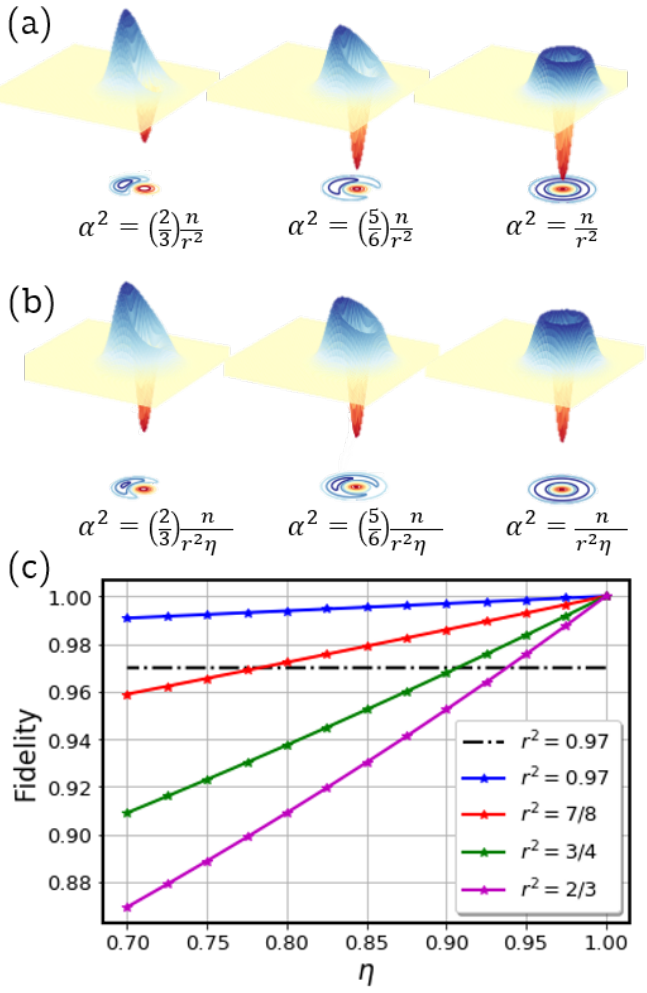



FIGURE 2.2: (a): Wigner functions for lossless photon catalysis per Eq. (2.9). As the amplitude of the initial coherent state approaches the optimal ratio, the output Wigner function becomes that of a displaced single-photon Fock State. When $\alpha^2 = n/r^2$, the fidelity between the two states is unity. (b): Wigner functions for photon catalysis with detector efficiency $\eta = 0.9$. (Amplitudes have the same ratio of optimal values as in (a)) (c): Maximum fidelity achievable with a displaced Fock-state of displacement $\beta = \sqrt{\alpha^2 - n/\eta}$ for detector efficiency $\eta$ when the beamsplitter parameter $r$ is varied. The black dot-dashed line indicates the fidelity of an ideal displacement with that obtained by the usual technique [127] for $r^2 = 0.97$, where the PNR detector is replaced by a partial trace.

Fig.2.2 shows the resulting Wigner functions after photon catalysis as the coherent state amplitude is tuned. If the PNR detector is perfect (Fig.2.2(a)), then the negativity of the Wigner function increases and becomes maximal as $|\alpha|^2$ approaches the optimal value to achieve an ideal displacement. Fig.2.2(b) demonstrates that when $|\alpha|^2$ is the same fraction of the optimal value as in Fig.2.2(a), but the detector is no longer perfect ($\eta = 0.9$), the Wigner function has the same qualitative shape in each case but exhibits an overall decreased negativity due to the effective loss at the detector. When comparing to a typical experimental displacement such as using the beamsplitter parameter $r^2 = 0.97$ [114](black dot-dashed line in Fig.2.2(c)), the addition of even imperfect PNR detectors improves the attainable fidelity with the ideal displaced single-photon state.

## 2.4 Schrödinger Cat States

We now turn to Schrödinger-cat coherent superpositions (SCSs), which are of the type

$$|SCS_{\pm}(\zeta)\rangle = N(|\zeta\rangle \pm |-\zeta\rangle), \tag{2.17}$$

where $N$ is the normalization constant. These non-Gaussian states have been proposed for quantum computing [131] and small-amplitude optical SCSs have been created by several methods [91, 96–99, 132], but there has yet to be a reliable approach to generate larger photon-number SCSs. Methods have been proposed to "breed" SCSs using two smaller SCSs, a beamsplitter, and conditional homodyne detection to create a larger SCS of amplitude $\sqrt{2}\zeta$ [133, 134], including an approach of particular interest involving an iterative process similar to photon catalysis where the PNR detectors are replaced with homodyne measurements [135]. These approaches will also require quantum memories. To our best knowledge, the largest created optical SCS to-date made use of two squeezed vacuum resources and a breeding step to create a squeezed SCS with $\zeta = 2.15$ and a fidelity of 0.74 (according to the fidelity defined by Eq. 2.7) [102].

As exemplified by the work of Sychev et al. [102], most SCS preparation methods result in a squeezed cat state, $S(r)\,|SCS_{\pm}(\zeta)\rangle$, where $S(r) = \exp[r(a^{\dagger 2} - a^2)/2]$ is the single-mode squeezing operator. Without loss of generality, we take the squeezing parameter $r \in \mathbb{R}$ for the numerical simulations in this work.

Using a procedure of iterated photon catalysis steps with only a coherent state and single photons as inputs, we show that one can create states that approach exact

squeezed SCS, where the final amplitude and squeezing of the state increases with the number of photon catalysis steps. We numerically demonstrate that for low photon numbers, the final SCS amplitude after photon catalysis increases more rapidly with each additional step than with the number of steps required by the homodyne detection based breeding protocol [133].

The general idea to generate squeezed SCS states from multiple photon catalysis steps is somewhat counterintuitive: consider, for example, the "odd" SCS,

$$|SCS_{-}(\zeta)\rangle = \frac{1}{\sqrt{\sinh|\zeta|^2}} \sum_{n=0}^{\infty} \frac{\zeta^{2n+1}}{\sqrt{(2n+1)!}} |2n+1\rangle\,, \tag{2.18}$$

which only contains odd photon numbers. Naïvely, it would be tempting to consider using photon catalysis to filter the even Fock components from a coherent state in an attempt to approximate an odd SCS. However, this approach would require cascaded stages which would "undo" one another in general because of the shift of the probability amplitudes by $n-1$, at each stage, when $n > 1$. Thus, the previously filtered Fock amplitudes in a cascaded scheme do reappear, in general, after the next stage. Nonetheless, and remarkably, cascaded photon catalysis *can* be used to generate excellent approximations to *squeezed* SCSs — which are useful to generate the GKP resource states crucial to CV quantum error correction, as we will see in the next section.

### 2.4.1 Cascaded photon catalysis

We turn to the squeezed SCS generation protocol, which consists of photon catalysis iterated over $N$ steps as shown in Fig.2.3. If we consider allowing the beamsplitter parameters and PNR measurements to vary at each step, then each iteration gives us two additional degrees of freedom: the deterministically controllable beamsplitter reflectivity, $r_i$, and the probabilistic detection, $n_i$. We limit the scope of this paper to measurements that are subsequently performed, which means that after being post-selected for an outcome $n_i$ at the $i^{th}$ step, the output state will be the input to the $(i+1)^{th}$ step along with the single-photon as displayed in Fig.2.3. The possibility of using measurement results as feed-forward parameters, so as to optimize protocol efficiency, will be investigated in future work.

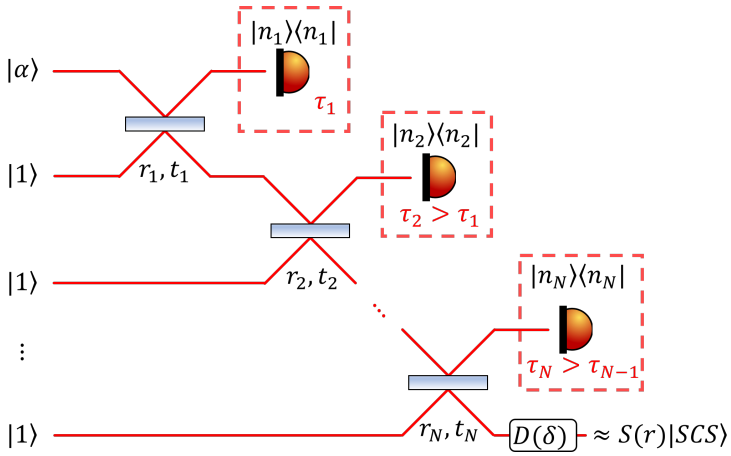



FIGURE 2.3: Protocol for generating an approximate squeezed SCS using photon catalysis. The initial coherent state of amplitude $\alpha$ is successively interfered with single-photon Fock states on a tree of variable beamsplitters, each of which is followed by a PNR detector. These steps are sequential in that each prior PNR detection occurs before the intermediate state is interfered at the next beamsplitter in the tree. By tuning the beamsplitter parameters and initial coherent state amplitude, the output conditioned by $N$ PNR detection steps is very near a squeezed SCS, up to a final displacement.

Using the general output given by Eq. A.10 for a coherent state, $|\alpha\rangle$, we arrive at the $N$-catalyzed state given by

$$|\phi_N\rangle \propto \sum_{m=0}^{\infty} \frac{\alpha^m}{\sqrt{m!}} \prod_{k=0}^{N-1} t_{k+1}^m \left\{ n_{N-k} t_{N-k}^2 - r_{N-k}^2 \left[ m + \sum_{j=0}^{k-1} (n_{N-j} - 1) \right] \right\} |m\rangle \,. \tag{2.19}$$

Following the intuition that filtering Fock components may lead to an SCS, we wish to compare the fidelity of the state given by Eq. 2.19 with that of a squeezed SCS. However, in order to simplify the resulting expression and decouple the squeezing parameter from the SCS amplitude, we instead choose to compare the photon-catalyzed state with a superposition of squeezed vacuum (SSV) given by

$$|SSV_{\pm}(\beta)\rangle \propto [D(\beta) \pm D(-\beta)]\, S(r)\, |0\rangle \,. \tag{2.20}$$

Using the property [136]

$$D(\beta)S(r) = S(r)D(\zeta) \tag{2.21}$$

$$\zeta = \beta \cosh r + \beta^* \sinh r, \tag{2.22}$$

it is easy to see that the squeezed SCS and SSV states are equivalent when

$$\zeta = e^{-r}\beta \tag{2.23}$$

where $\beta$ is chosen to be real without loss of generality. By allowing for a potential displacement following the multiple photon catalysis procedure, we can derive an expression for the fidelity with the $N$ step catalyzed state and the target state,

$$\rho_T = D(\delta) \left|SSV\right\rangle \left\langle SSV\right| D^\dagger(\delta). \tag{2.24}$$

Here, $\delta$ is the amplitude by which the final catalyzed state must be displaced back to the origin of phase-space in order to recover the desired $|SSV\rangle$ state defined in Eq. 2.20. The analytical expression for the fidelity with the target state is derived in Appendix A.3 and given by

$$F = \frac{C_{SSV\pm}^2 C_{|\phi_N\rangle}^2}{\cosh r} \left| \sum_{\ell,m=0}^{\infty} \left\{ (\tanh r)^\ell \frac{(2\ell-1)!!\alpha^m}{\sqrt{(2\ell)!m!}} [A(2\ell, m, -\beta-\delta) - A(2\ell, m, \beta-\delta))] \right.\right.$$
$$\left.\left. \times \prod_{k=0}^{N-1} \left[ t_{k+1}^m \left( n_{N-k} t_{N-k}^2 - r_{N-k}^2 (m + \sum_{j=0}^{k-1} (n_{N-j} - 1)) \right) \right] \right\} \right|^2, \tag{2.25}$$

where $C_{SSV\pm}$ normalizes Eq. 2.20, $C_{|\phi_N\rangle}$ normalizes Eq. 2.19, and we define $A(n, m, \gamma) \equiv \langle n| D(\gamma) |m\rangle$. This expression can be optimized with respect to all available parameters. In order to limit the computation time and adhere to realistic PNR detection capabilities, we constrain each $n_i < 10$. For each combination of possible detections subject to the constraints, we implemented the Nelder-Mead algorithm to perform a multivariate numerical optimization in Mathematica. Additionally, although all optimizations were performed with respect to the SSV amplitude, $\beta$, we will give the results in terms of an ideal squeezed SCS amplitude for easier comparison with other work.

We performed the fidelity optimization for a given coherent state passing through two, three, and four-step cascaded photon catalysis process. We noticed that in each

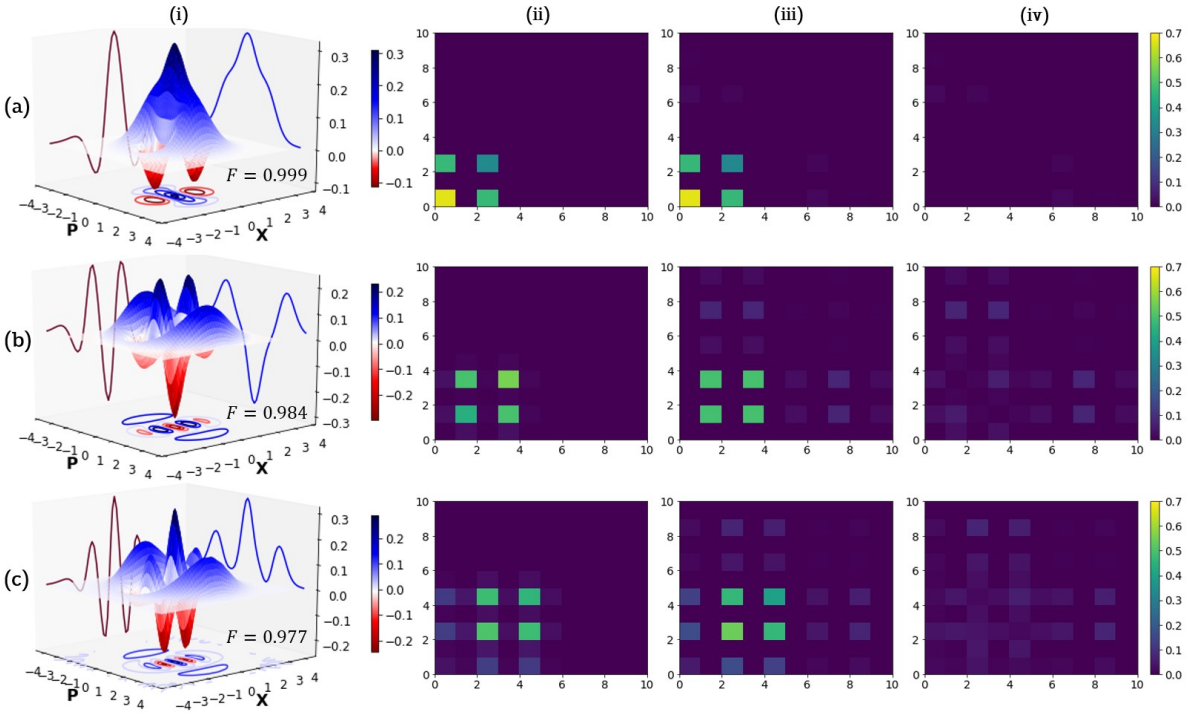



FIGURE 2.4: Calculated Wigner functions and density matrix elements for the approximate squeezed cat states resulting from varying the number of cascaded photon catalysis iterations with an initial coherent state input. Row (a) shows the results after two iterations, (b) after three iterations, and (c) after four. In each separate case (a-c), all beamsplitter parameters, PNR detections, and the input coherent state amplitude are tuned independently to optimize the fidelity with the nearest squeezed cat state. Column (i) shows the plotted Wigner functions for the photon catalyzed states, and (ii) gives the density matrix elements. Column (iii) gives the density matrix elements of the ideal associated squeezed cat, and (iv) shows the magnitude of the difference between (ii) and (iii).

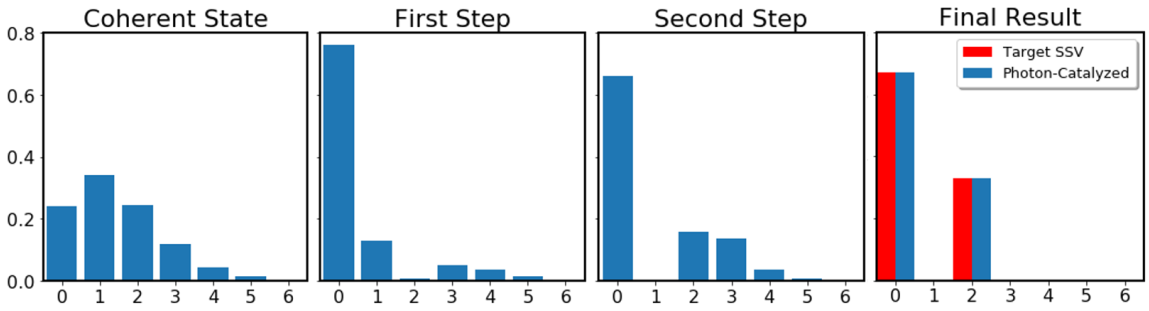



FIGURE 2.5: Photon number probability amplitudes for various states. Top: distributions for coherent state input and the quantum state after the first detection step. Bottom: distributions for the quantum state after the second PNR detection and for the final catalyzed state after a displacement(blue) compared to the target SSV state(red).

case, there are multiple sets of parameters that generate a high-fidelity SSV state. In hindsight, this may not be a surprise, as the number of free parameters likely underconstrains the system for low amplitude target states. We selected the results that have the highest probability of success, which are depicted both in terms of the plotted Wigner function and density matrix elements in Fig.2.4. We find that we can produce a nearly ideal SCS with $\zeta = 1.50$ squeezed by 1.91 dB after a two step process ($F > 0.999$). Optimizing a three-step process yields an SCS with $\zeta = 2.18$ squeezed by 4.17 dB ($F = 0.984$), and using $N = 4$ results in $\zeta = 2.67$ with squeezing of 4.52 dB ($F = 0.977$). The full list of optimized parameters for each case is provided in Appendix A.4. We note that scaling from $N = 2$ to $N = 3$ in this procedure already increases the cat amplitude by a factor of $\sqrt{2}$. This approach only includes a single additional interference and detection stage whereas the other breeding procedure discussed earlier requires three additional steps provided that the two smaller cats to be bred came from an $N = 2$ photon catalysis procedure. This indicates that an iterative photon-catalysis process might scale better for larger SCS creation than homodyne-based breeding.

It is important to keep in mind that each $N$-step photon catalysis process is optimized to produce an approximate SSV state as a final result; simply adding an additional catalysis step to any of these final approximate cats will not increase the cat amplitude or squeezing under the current approach. As an example, consider Fig.2.5, which shows the photon number distributions before and after each catalysis step for $N = 2$. The filtering of Fock components at each step is not perfect due to the numerical procedure, but it is easy to see that the first projective PNR measurement, $|1\rangle\langle 1|$, acts

to nearly eliminate the $|2\rangle$ component from the initial coherent state, while the second step effectively shifts and re-scales the distribution. A final displacement of the Wigner function back to the origin of phase-space results in a very near approximation to the target state as shown by the comparison of probability amplitudes in the rightmost panel of Fig.2.5. Although the first catalysis step acts to de-Gaussify the coherent state, applying the displacement at this point would result in poor overlap with any class of SSV.

### 2.4.2 Loss tolerance

Experimental imperfections and associated optical losses will contribute to non-ideal single-photon resources, in which case we must consider a mixture of single-photon and vacuum as the realistic input at each interference beamsplitter [111, 114]. This can be modeled by sending each pure single photon through a lossy channel with transmission $\gamma$, so that each input is now

$$\rho_{mix} = \gamma \left|1\right\rangle \left\langle 1\right| + (1-\gamma) \left|0\right\rangle \left\langle 0\right|. \tag{2.26}$$

By using the optimized parameters for the ideal procedure and replacing the pure single-photons with $\rho_{mix}$ at each step, we see in Fig.2.6 how including a vacuum contribution adversely affects the fidelity with the target states in each case. The decreased purity more greatly harms the procedures with larger $N$, unsurprisingly.

However, we note that just as in Ref. [135], our protocol in of itself adds little imperfection to the result, which can be seen directly in Fig.2.6 by comparing the drop in fidelity between neighboring values of $N$. For purities above $\gamma = 0.8$, the decrease in the fidelity of our output state with the target between $N = 2$ and $N = 3$ step processes is nearly equivalent to the fidelity difference between $N = 3$ and $N = 4$ step procedures, which indicates the change in fidelity is nearly linear with the number of impure single-photon inputs.

Fig.2.6 indicates the importance of pure single photons for our protocol; fortunately, there has been much work in the area of on-demand, highly efficient single-photon emitters using solid state devices [137–139]. On-demand quantum dot emitters have already demonstrated single-photon purities above 99% [140], and a recent proposal making use of quantum feedback control suggests that on-chip sources with

greater than 99% on-demand extraction efficiency are possible with current technology [141]. As in Sec. 2.3, we consider the effects of imperfect detectors in each step,

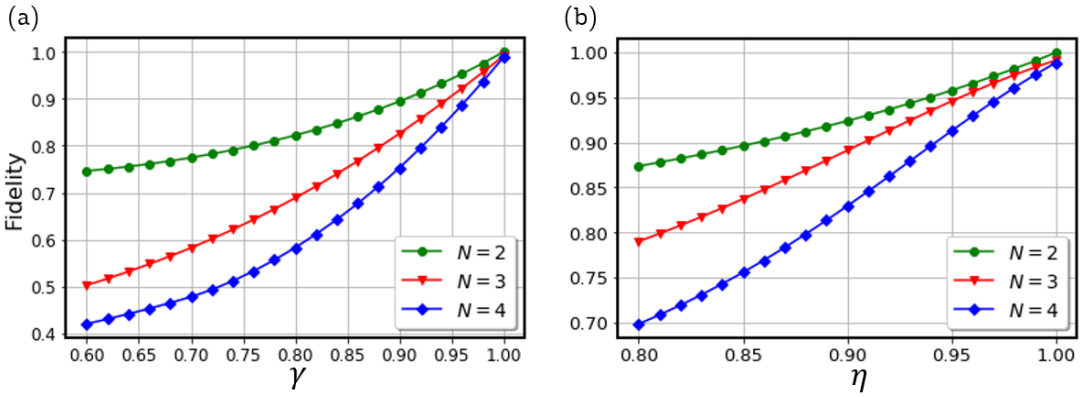



FIGURE 2.6: The effects of vacuum mixing due to (a) impure single-photon input and (b) imperfect detectors are shown separately for $N$-step photon catalysis with $N = 2, 3$ and 4. The fidelity is calculated with respect to the same target SSV states as for the optimized procedure shown in Fig.2.4.

where we assume identical detector efficiencies of $\eta \leq 1$. Similarly to having impure single-photons, each additional imperfect PNR measurement acts to further decrease the fidelity with the target state in a way that appears linear with increasing photon catalysis steps. We see from Fig. 2.6 that generating high-fidelity cat states would require the use of highly efficient PNR detectors, as $\eta \leq 0.85$ prohibits even $N = 2$ step procedures from yielding states above a fidelity of 0.9. Thanks to recent advances, high efficiency PNR is achievable with current transition-edge sensors [124] and superconducting nanowire single-photon detector arrays [142], which have been recently reported to have PNR capabilities [143].

### 2.4.3 Success probability

For methods of non-Gaussian state engineering that use homodyne detection as in Ref. [135], one can select a confidence window of $\Delta x$ on acceptable post-selection measurement results that allow for an increase in procedure success probability at the cost of decreasing fidelity. In our case, however, the projective operator of PNR detection lies in a discrete subspace, so we cannot allow for continuous deviations from the ideal detection scheme. Nonetheless, we can still examine the fidelity and success probability in cases where the initial coherent state amplitude and $N$ deterministic beamsplitter

parameters are held fixed at values optimized for SSV generation, but some $n_i$ deviate, necessarily in integer steps, from the ideal scenario.

For each combination of detection events that deviates, we can re-optimize the fidelity from Eq. 2.25, where now only the target state parameters are allowed to vary (SSV amplitude, squeezing, and the displacement amplitude). In this way, it is possible to generate a finite list of PNR detection combinations that differ from the combination that best approximates an ideal SSV state, yet still yields a final state above a given threshold fidelity, $F_{thr}$, with a target state. This suggests that approaches involving feed-forward based on measurement results might also be fruitful. If one only discards the photon-catalyzed states where the detected combination $\{n_1, n_2, \ldots, n_N\}$ is outside of the acceptable list, then the photon-catalyzed state is guaranteed to have a fidelity $F \geq F_{thr}$ with an appropriate SSV state. This target state may vary slightly in amplitude and squeezing from the SSV target for the optimal detection scheme (removing a different total photon number changes the total energy of the state), but the target is still required to be an SSV state, and we show later how two similar SSV states can be enlarged by utilizing PNR detection. By decreasing $F_{thr}$, more combinations of possible PNR detections are included in the "accept" list, thereby increasing the probability of successfully generating an SSV state. The effects of lowering $F_{thr}$ for the two, three, and four-step photon catalysis procedures considered in the previous section are depicted in Fig.2.7. For the $N = 2$ and $N = 3$ cases, we see that setting $F_{thr} = 0.9$ yields nearly an order of magnitude increase in success probability from the perfect detection scenario, while the $N = 4$ case shows an increase of nearly two orders of magnitude for the same fidelity threshold. Paired with state-of-the-art technology in deterministic single-photon sources [140, 144] with emission rates on the order of $10^6$ counts/sec [137], an $N = 3$ step photon catalysis process would yield nearly perfect SSV states at rates in the kilohertz regime.

### 2.4.4 Higher-symmetry SSV states

The versatility of cascaded photon catalysis allows us to examine classes of SSV states that exhibit beyond two-fold phase space rotational symmetry. These states can be more generally expressed as superpositions in the form of

$$|SSV^{(M)}\rangle \propto \sum_{n=0}^{M-1} D(\beta e^{\frac{2n}{M}i\pi}) S(re^{\frac{4n}{M}i\pi}) |0\rangle , \tag{2.27}$$

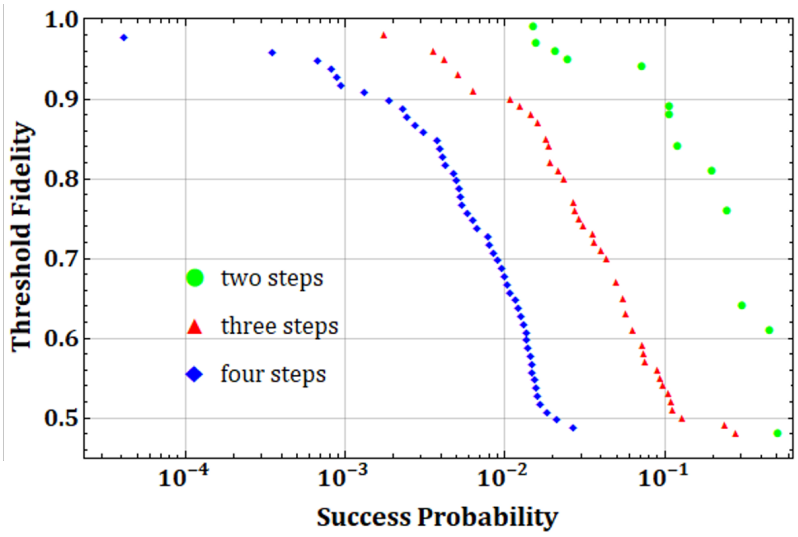



FIGURE 2.7: The success probability to generate an approximate SSV state with at least a given threshold fidelity is shown for a cascaded photon catalysis procedure of two(green), three(red), and four(blue) steps. Allowing the PNR detection results to deviate from the optimized values yields an increased probability of success at a cost to fidelity.

where $M$ is the number of components in the superposition and degree of rotational symmetry. Each component consists of squeezed vacuum that has been displaced with an integer multiple of phase increment of $2\pi/M$, where the squeezer anti-squeezes the vacuum perpendicularly to the axis of the subsequent displacement. Although there is no restriction on $r$ for this class of states, we consider $r < 0$, as this corresponds to the states we generated with cascaded photon catalysis. While we only explicitly consider protocols to generate states with $M = 2$ and $M = 3$ in this work, we expect that it is possible to further generalize the generation of $SSV^{(M>3)}$ by optimizing photon-catalysis.

For the next highest symmetry state (three-fold symmetry), we can again use the protocol depicted in Fig.2.3 with three photon catalysis steps, but with different input coherent state amplitude and beamsplitter parameters, to create the state

$$\left(D(\beta)S(r) + D(\beta e^{\frac{2\pi i}{3}})S(re^{\frac{4\pi i}{3}}) + D(\beta e^{\frac{4\pi i}{3}})S(re^{\frac{2\pi i}{3}})\right)|0\rangle\,, \tag{2.28}$$

where $r = -0.24$, $\beta = 1.255$, and the fidelity is greater that 0.99. This is achieved with detection events of $\{n_1, n_2, n_3\} = \{6, 4, 2\}$, which were found serendipitously. The

comparison of the ideal and procedurally generated states is depicted in Fig.2.8(a), and Fig.2.8(b) demonstrates the 120° rotational symmetry of the Wigner function. As this state is created with a cascaded photon catalysis procedure, the success probability and dependence on purity of single-photon inputs and detector efficiency is very similar to the other $N = 3$ case considered earlier. The utility of the $M = 3$ SSV state is developed further in the next section.

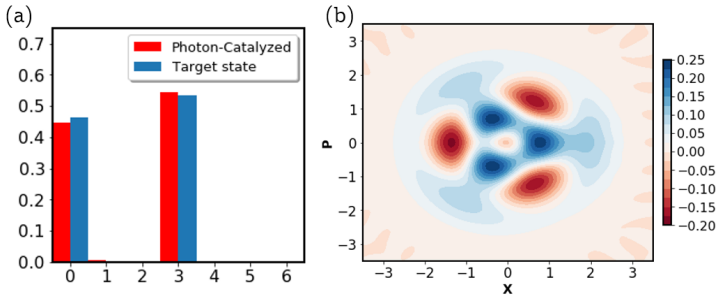



FIGURE 2.8: Photon-number distribution, (a), and Wigner function, (b), of the approximate $SSV^{(3)}$ state generated by a three-step photon catalysis protocol, where the squeezing parameter is $r = -0.24$ and amplitude is $\beta = 1.255$. The Wigner function nicely displays 120° phase-space rotational symmetry.

## 2.5 Breeding Protocols

Methods to combine SCS with breeding protocols have been widely explored to generate larger amplitude SCS states [99, 102, 133, 134, 145] and other classes of comb states [82, 135]. A recent work shows how feed-forward displacements can allow for deterministic creation of optical GKP states from a supply of squeezed SCS [83]. Here, we motivate the extension of these breeding protocols to PNR detection as opposed to traditional homodyne measurements. Additionally, we make use of the phase-insensitivity of direct photon-counting measurements to demonstrate how PNR detectors can be used to breed higher symmetry states.

### 2.5.1 SSV states

First, we begin by examining a procedure with the aim to take two lower amplitude $M^{\text{th}}$-order SSV states and produce a larger SSV state of the same order of symmetry.

Taking two identical states given by Eq. 2.27 and interfering them on a balanced beamsplitter, we have the state

$$|\phi\rangle \propto U_{ab} \sum_{n=0}^{M-1} D_a(\beta_n) S_a(\xi_n) |0\rangle_a \otimes \sum_{n=0}^{M-1} D_b(\beta_n) S_b(\xi_n) |0\rangle_b , \tag{2.29}$$

where $\beta_n = |\beta| e^{\frac{2n}{M} i\pi}$ and $\xi_n = r e^{\frac{4n}{M} i\pi}$, with the beamsplitter operation given by $U_{ab} = e^{\frac{\pi}{4}(ab^\dagger - a^\dagger b)}$ . Developed further in Appendix A.5 and numerically verified for the first few SSV states, it can be seen that for finite squeezing and relatively larger values of $|\beta|$, a PNR measurement of zero photon, $|0\rangle\langle 0|$, on output mode $b$ leads to

$${}_b\langle 0 |\phi\rangle_{ab} \approx \sum_{n=0}^{M-1} D_a(\sqrt{2}\beta_n) S_a(\xi_n) |0\rangle_a . \tag{2.30}$$

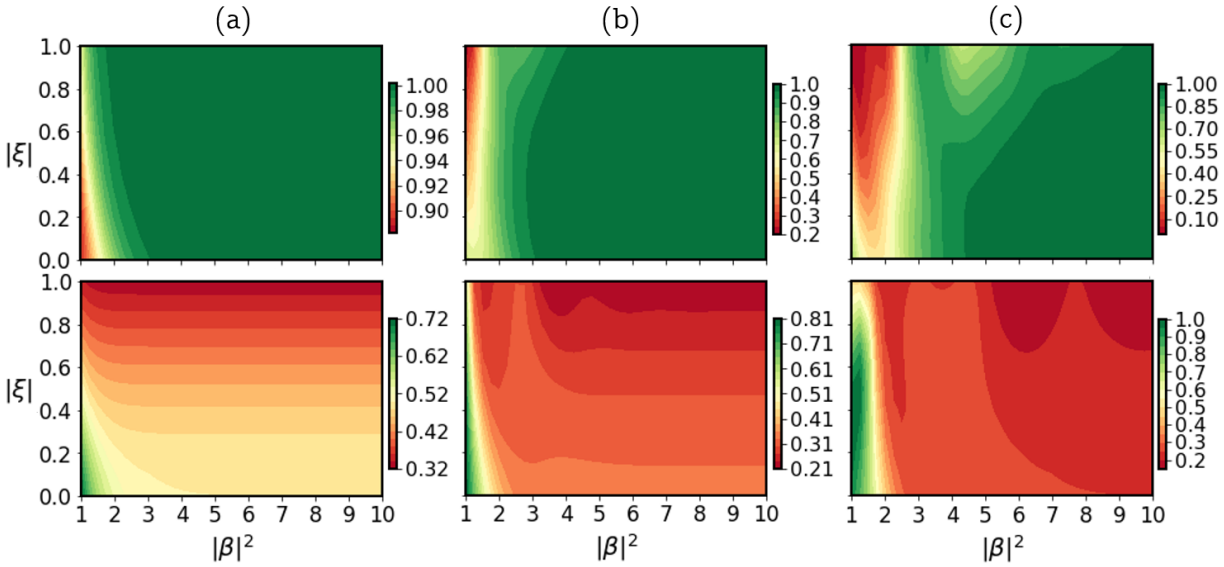



FIGURE 2.9: Fidelity and success probability of the PNR detection-based breeding protocol for various initial amplitudes, $\beta$, and squeezing, $\xi$. The colors bars indicate fidelity (top panel) and success probability (bottom panel) for the respective plots, where dark green is unity. $M = 2$ (a), $M = 3$ (b), and $M = 4$ (c) symmetric SSV states are bred into larger states in the same class with amplitude $\sqrt{2}\beta$ and the same squeezing parameter.

This means that using PNR detectors, SSV states can be bred into states with the same squeezing, but with increased amplitude $|\beta| \to \sqrt{2}|\beta|$. The factor of $\sqrt{2}$ increase is the same result as achieved for SSV states with $M = 2$ using homodyne detection-based breeding protocols, but the nature of performing a homodyne measurment along

a particular axis in phase-space means that higher symmetry features are not preserved, thus indicating that PNR-based protocols such as this will be necessary to breed general SSV-type states. The fidelity of the bred states with the enlarged target states as well as success probabilities are shown in Fig.2.9. In all cases examined, the fidelity quickly approaches unity for values of $|\beta|^2 \gtrapprox M$. Additionally, the success probabilities are greater than $10^{-1}$ in all regions where fidelities are numerically indistinguishable from one, and the probability for success in the $M = 2$ case holds steady at 0.5 for squeezing $|\xi| \lessapprox 0.3$ and $|\beta|^2 > 4$.

### 2.5.2 Square-lattice GKP states

Here, we adapt the experimental proposal of Vasconcelos et al. [82] to our cascaded photon catalysis method of producing SCSs followed by PNR detector-based breeding. The essence of their original idea is as follows: by entangling two squeezed SCS on a balanced beamsplitter and performing a homodyne measurement of $p = 0$ on one output, the output of the other port is projected into a state exhibiting a series of evenly spaced Gaussian peaks along the $x$-quadrature axis where the peak amplitudes follow a binomial distribution about a central peak at $x = 0$, and the peak width is determined by the SCS squeezing. Repeating an identical protocol with two of these outputs serves to increase the number of peaks by elevating the order of the binomial distribution. Recently, Weigand et. al have shown that including feed-forward displacements based on the homodyne measurement results makes this process deterministic [83]. After many iterations and in the limit of large initial SCS squeezing, this method accurately approximates an ideal square-lattice GKP state, which can be explicitly written as [48]

$$|\mu_{\text{sqr}}\rangle \propto \sum_{n_1,n_2=-\infty}^{\infty} e^{-i\hat{p}\sqrt{\frac{2\pi}{d}}(dn_1+\mu)} e^{i\hat{q}\sqrt{\frac{2\pi}{d}}n_2} |0\rangle, \tag{2.31}$$

where $d$ is the dimensionality of the code space and $\mu = 0, \ldots, d-1$ labels the specific GKP state in the code space. To make the state physically realizable in the case of finite energy, the state can be modulated by a Gaussian envelope [146]. The finite energy GKP state is thus given by

$$|\mu^{\Delta}_{sqr}\rangle \propto e^{-\Delta^2\hat{N}} |\mu_{sqr}\rangle \tag{2.32}$$

where $\hat{N} = a^\dagger a$ is the photon-number operator. These finite-energy GKP states

compose a periodic square-lattice grid in phase-space, where each peak has width $\Delta$. The result of following this homodyne-detection based procedure with two states generated by our four-step photon catalysis procedure is shown in the top left plot of Fig.2.10.

If we perform PNR detection on one mode following the entangling beamsplitter instead of a homodyne measurement, we can obtain different final states according to the selected measurement result that each exhibit the periodic comb-like structure of a GKP state as demonstrated in Fig.2.10. Using a balanced beamsplitter, $U_{ab}$, and phase-shift operator $R(\Theta) = e^{i\Theta a^\dagger a}$, the overall result can be concisely written as

$$|\phi\rangle \propto {}_b\langle n| U_{ab} R_a(\Theta) |SSV^{(2)}\rangle_a |SSV^{(2)}\rangle_b . \tag{2.33}$$

This is exactly the process that we used earlier to breed SSV states, but with two important changes. Because PNR detection is phase-insensitive, precisely controlling the relative phases between the states impinging on the entangling beamsplitter becomes vital, and here we set the phase difference to $\Theta = \frac{\pi}{2}$ to achieve the square-grid comb-like periodicity. Additionally we allow for PNR detection other than vacuum.

As a concrete example, if we entangle two SSV states generated from photon catalysis (Fig.2.4(c)) and detect $n = 4$ photons in one mode, we achieve a state that has a fidelity of $F = 0.996$ with a finite-energy square GKP state of width $\Delta = 0.545$. This single event occurs with a success probability of 0.09, but the probability to achieve at least one of the comb states shown in Fig.2.10 exceeds 0.60. It has recently been shown that in the absence of squeezing, the processes of breeding a pair of cat states as per Eq. 2.33 deterministically produces one of a family of four-component cat states, otherwise known as compass states [147]. Under certain conditions, low-amplitude compass states are very close to approximate GKP states, which is precisely what is shown in Fig. 2.10.

Since finite-energy GKP code states are no longer orthogonal within the $d$-dimensional code space, error correction protocols become probabilistic due to the non-zero overlap between the logically encoded states. In the original work, the probability to successfully correct an error hinges solely on the ability to accurately distinguish code states, leading to a 99% success rate for $\Delta \sim 0.5$ [48], such as is the case for our states. However, a more realistic error-correcting protocol considers potential shift errors on the ancilla qubits as well as the encoded qubit to be corrected, leading to a more stringent

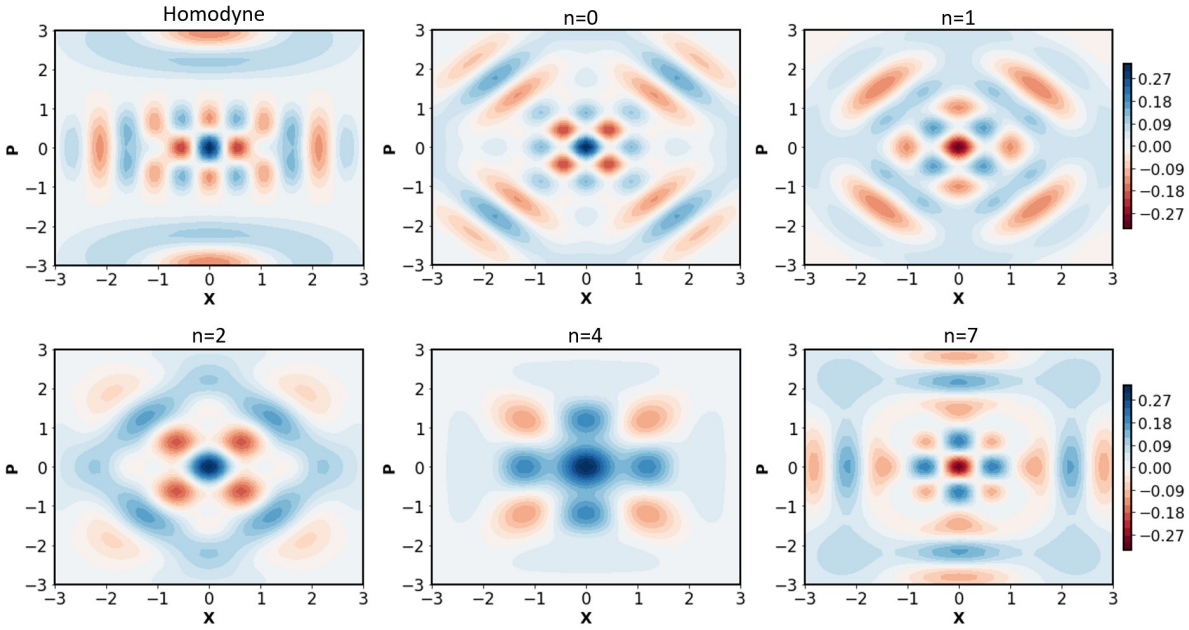



FIGURE 2.10: Wigner functions of bred two-fold symmetry SSV states (squeezed SCS states) achieved by the photon catalysis protocol after various detector measurements. The homodyne plot demonstrates the grid state of the result after a $p = 0$ quadrature measurement, while the other plots show the resultant state after a PNR detection of $n$ photons. The $n = 4$ case has a fidelity $F = 0.996$ with a square-grid GKP state with finite-energy peak-width $\Delta = 0.545$.

requirement of $\Delta < 0.15$ to achieve the same success rate [148]. Utilizing more photon catalysis iterations to increase the initial SCS squeezing and performing multiple breeding operations as in [83] will result in suitable states for realistic quantum error correction.

### 2.5.3 Hexagonal GKP states

One principle advantage exhibited by the GKP encoding scheme is an inherent resistance to shift errors in the conjugate position and momentum quadratures $x$ and $p$. The hex GKP state, formulated in the original work [48] and recently expanded upon in greater detail in Refs. [146] and [106], improves upon the initial square lattice state by being able to correct larger displacement errors. As a hexagonal lattice in phase space, the hex GKP state can now account for shift errors with a new upper-bound phase

space radius of $r \leq \sqrt{\frac{\pi}{\sqrt{3}d}}$, as opposed to $r \leq \sqrt{\frac{\pi}{2d}}$ intrinsic to the original, where $d$ is the dimensionality of the code space. This hex GKP state can be explicitly written as

$$|\mu_{\text{hex}}\rangle \propto \sum_{n_1,n_2=-\infty}^{\infty} e^{-\frac{i}{2}(\hat{q}+\sqrt{3}\hat{p})\sqrt{\frac{4\pi}{d\sqrt{3}}}(dn_1+\mu)} e^{i\hat{q}\sqrt{\frac{4\pi}{d\sqrt{3}}}n_2}|0\rangle, \tag{2.34}$$

where as in the square-lattice case, applying the non-unitary operator $e^{-\Delta^2\hat{N}}$ makes the state physically realizable [146]. Fig.2.11 shows the Wigner function for a linear combination of hex GKP states of dimension $d = 2$ plotted with different peak-widths.

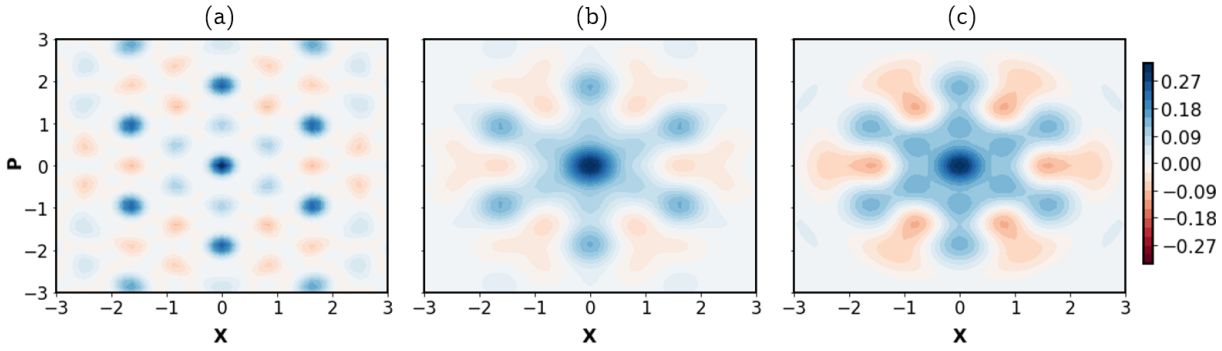



FIGURE 2.11: Comparison of finite energy hex GKP states to that of two 120° rotationally symmetric SSV states after breeding. (a) and (b) show the Wigner function of a mixture of hex GKP states of code dimension 2 with peak widths $\Delta = 0.2$ and $\Delta = 0.46$, respectively. (c) shows the Wigner function of bred SSV states each developed by an $N = 3$ photon catalysis protocol.

Noting that the SSV$^{(2)}$ states (squeezed SCS) can be bred into square-lattice GKP states, we are motivated to examine what happens when breeding SSV$^{(3)}$ states, such as those generated in Sec. 2.42.4.4. Because the states have $2\pi/3$ rotational symmetry, we fix the phase difference of the two initial states to be $\Theta = \pi/3$ at the entangling beamsplitter, and send one output mode to a PNR detector. Just as with the SSV$^{(2)}$ breeding protocol, we find that several possible measurement results lead to a final state exhibiting the desired hexagonal symmetry. Fig.2.11(c) shows the result of breeding two of the states depicted in Fig.2.8 after post-selecting on a PNR measurement of $n = 0$ with a probability of 0.31. While only having a fidelity of 0.8 with the mixed hex GKP state in Fig.2.11(b), the visual similarity and relatively high success probability indicate that there may be a suitable SSV$^{(3)}$ for more realistic hex GKP synthesis.

## 2.6 Summary and Goals

This chapter further explored the photon catalysis process [118] where interfering a single-photon with an input state on a beamsplitter can result in specifically engineered non-Gaussian quantum states contingent upon beamsplitter parameters and conditioned by PNR detection. By using a coherent state as an input, we demonstrated that one photon catalysis step can implement arbitrary displacements on the single photon Fock state and preserve state purity. Further, by cascading the photon catalysis process, we are able to filter out multiple photon number components and accurately approximate relatively large superpositions of squeezed vacuum, including what is to our knowledge the largest proposed optical SCS without the need for large-number Fock states or breeding. Furthermore, we adapted breeding protocols based on homodyne detection [102, 133–135, 145] to a PNR-based scheme that allows for breeding $M$-fold symmetry SSV states. By changing the phase between input modes of the PNR-based breeding, we showed that it is possible to reconstruct square and hexagonal lattice comb-states that exhibit the structure of finite-energy GKP-class code states. Future avenues of exploration may include utilizing PNR measurement results to implement feed-forward modifications to subsequent photon catalysis steps in the hopes to increase protocol success rates.

As mentioned previously in Ch. 1, it was shown that combining a supply of GKP states with Gaussian QC alone is sufficient for fault-tolerant universal QC; therefore, developing a technique to synthesize GKP states is vital [69]. Machine learning optimization algorithms have demonstrated the potential to create high-fidelity GKP states [149], but these methods require in excess of 100 gate operations. The method of photon catalysis and PNR-based breeding we describe here provides an experimentally feasible approach to generate various nonclassical states, including GKP, important for quantum information and quantum computing.

Future work may focus on optimizing and extending the protocol discussed, such as perhaps compacting the design or including machine learning methods along with tunable parameters from step to step of the catalysis process. For example, consider Fig. 2.12. Provided one has a supply of single photons, one may use an optical cavity with tunable beamsplitters to apply many steps of photon catalysis until a desired result is achieved. At each step, $i$, the coherent state $|\alpha_i\rangle$ can be tuned to apply a displacement of $t\alpha_i$, which then interferes with the input photon at a beamsplitter with tunable parameters. The result of the PNR detection, $n_i$, along with the displacement value can

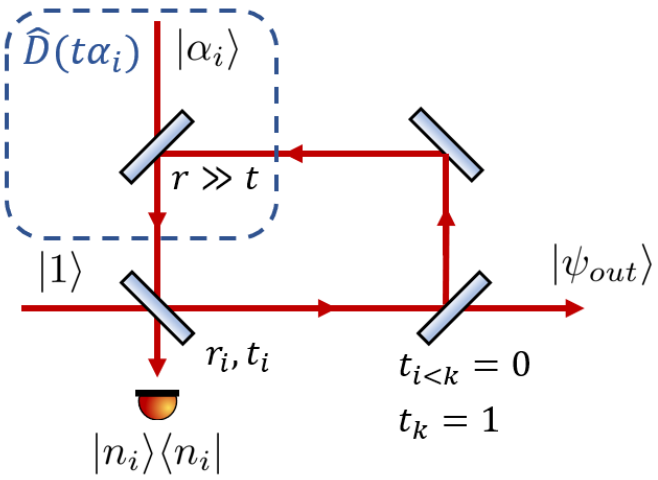



FIGURE 2.12: Multi-stage photon catalysis may be 'folded' into a loop, where the round-trip time of the loop corresponds to the single-photon pulse separation time. On each step, the single photon interferes with the displaced outcome of a previous catalysis step at a beamsplitter. The displacement and beamsplitter parameters could be controlled based on prior round detection outcomes.

be used as feed-forward parameters that determine the future displacement $\alpha_{i+1}$ and values of the beamsplitter parameters, $r_{i+1}, t_{i+1}$. This is an example of a Markov process as the output of a photon catalysis step depends only on the previous output state and the current modifiers,

$$|\phi_i\rangle = {}_2\langle n_i|\hat{B}_{12}(r_i, t_i)\hat{D}(\alpha_i)|\phi_{i-1}\rangle_1|1\rangle_2. \tag{2.35}$$

Two things can happen:

- The output state can converge to a desired target state, at which point the state can be released from the loop on the $k$-th step and used as desired.
- The output state of the $k$-th step is determined to be unable to converge to a desired result regardless of future steps and must be thrown away.

This is still an ongoing area of research, but preliminary results suggest that machine learning models used to determine feed-forward beamsplitter and displacement parameters may drastically increase state generation success probability and allow for convergence to a desired resource, such as a cat state, in a reasonable amount of time.

# Chapter 3

# Embedding Non-Gaussianity in a Cluster State

The previous chapter focused on a means to design some gadget that would generate a useful non-Gaussian resource. Once generated, the user would have the freedom to use the resource state as desired. However, the probabilistic state generation and challenge of integrating the resource state with another quantum information platform could pose difficulties. Additionally, single-photons are themselves a non-Gaussian resource, and while photon emitters have seen tremendous advances, they still pose challenges to produce on-demand [137]. With this in mind, this chapter proposes a new method to generate cat states embedded directly within a cluster state [150].

## 3.1 Overview

We present a *near-deterministic* approach to generating non-Gaussian states, such as cat and GKP states, directly *embedded* in a CV cluster state, by fully leveraging the cluster state's measurement-based QC capability. This approach, which we dub the Photon-counting-assisted Node-Teleportation Method (PhANTM), allows us to build polynomial gates and use them to generate cat states. An essential feature of this method is that it also stabilizes cat states against Gaussian noise and prevents amplitude decay. As such, this protocol preserves non-Gaussianity in the CV cluster, which was previously thought only possible through the use of GKP error-correction [44].

Practically, all PhANTM requires is an initial CV cluster state and its standard QC tools that are balanced homodyne field measurements and field displacements, to

which we adjoin photon-number-resolving (PNR) measurements in the small number regime by way of photon subtraction.

This chapter is organized as follows: section 3.2 describes the PhANTM algorithm and how it can be applied by performing measurements on a cluster state to generate cat states. The process is first motivated with ideal cluster states and subsequently extended to realistic cases with finite squeezing and amplitude damping introduced by experimental photon subtraction. Section 3.3 presents numerical results demonstrating how a 1D cluster state can be reduced to a large amplitude cat state with high probability and additionally shows that the PhANTM algorithm can be used to preserve cat states already present within a cluster state. Section 3.4 applies previous results on breeding cat states with beamsplitters and homodyne detection [82, 83, 133, 135, 151] to breeding cat states within the cluster state for generating enlarged cats and grid states, the later of which can be produced without post-selection [83] and used as a resource for universal QC [69]. We also make the connection to phase-estimation protocols and show how they can be implemented with cat states in the cluster. Section 3.5 motivates an extension to macronode-based cluster states, which are the predominant form of current CV cluster-state architectures [53, 54, 56, 57], and in Sec. 3.6 we conclude.

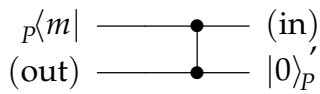
It should be emphasized that this approach utilizes feed-forward Gaussian quantum information processing steps on the cluster state interspersed between PNR detection events, and as such is beyond the scope of Gaussian Boson Sampling (GBS) machines previously proposed for non-Gaussian state generation [85], in which PNR detection acts terminally. Note also the recent negative result concerning the impossibility of generating general entangled cat states with a GBS-type machine where displacements are neglected[152].

## 3.2 PhANTM

### 3.2.1 Ideal teleportation

As mentioned in Ch. 1, measurement-based CVQC with cluster states is fundamentally based on CV teleportation, where Gaussian measurements in the form of homodyne detection teleport quantum information between neighboring nodes of a graph. The

freedom to choose the measurement basis of the the homodyne detection by controlling a classical local oscillator phase additionally allows for performing Gaussian operations on the teleported state [38, 44]. The ideal teleportation circuit in canonical form is described by




(3.1)

where $P$-subscripts indicate the quadrature basis. In this circuit, quantum information enters from the right and is coupled to a zero momentum eigenstate, $|0\rangle_P$, with the control-Z gate defined as $\hat{C}_Z = e^{i\hat{Q}_1\hat{Q}_2}$. A homodyne measurement is then performed in the $P$ basis on the top wire yielding result $m$. We adopt the convention that circuits proceed from right-to-left, so as to coincide with the order of operator action when writing out the mathematics. The result of this circuit is simply to teleport the input quantum information in the top wire to the bottom wire with an additional Fourier transform (rotation of $\pi/2$) and a displacement that depends on the measurement result; that is to say, the operator $\hat{X}(m)\hat{R}(\frac{\pi}{2})$ is applied to the input. Here, we have defined the rotation and quadrature shift operators as

$$\hat{Z}(y) = e^{i\hat{Q}y} \tag{3.2}$$

$$\hat{X}(x) = e^{-i\hat{P}x} \tag{3.3}$$

$$\hat{R}(\theta) = e^{i\theta\hat{a}^\dagger\hat{a}}, \tag{3.4}$$

where a general displacement is expressed as

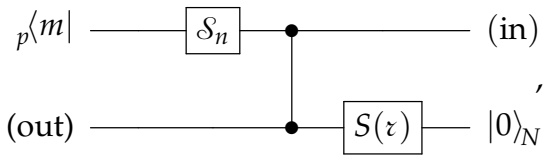
$$\hat{D}(\alpha) = e^{\alpha\hat{a}^\dagger - \alpha^*\hat{a}} = \hat{Z}\left(\sqrt{2}\mathrm{Im}[\alpha]\right)\hat{X}\left(\sqrt{2}\mathrm{Re}[\alpha]\right)e^{-i\mathrm{Re}[\alpha]\mathrm{Im}[\alpha]}. \tag{3.5}$$

By varying the measurement phase, all Gaussian operations on cluster states can be realized with homodyne measurement alone[38] up to a displacement.

However, in order to achieve a quantum speed-up, non-Gaussianity must be introduced; and furthermore, this non-Gaussianity must not be efficiently simulable classically [12]. This can be done by introducing the ability to perform photon addition or subtraction, which introduces negativity into the multi-mode Wigner function [153, 154]. In our approach, we examine the effects of applying PhANTM, which utilizes PNR detection to perform photon-subtraction within the teleportation circuit.

Suppose we wish to perform a photon subtraction of $n$ photons immediately prior to the homodyne detection step of the teleportation circuit. This circuit, including finite

squeezing on the ancillary mode, is represented by




(3.6)

where realistic photon subtraction of $n$ photons is implemented by applying the Kraus operator

$$\hat{\mathcal{S}}_n = \frac{(-1)^n e^{n\beta/2}}{\sqrt{n!}} (2\sinh\beta)^{n/2}\, \hat{a}^n \hat{N}(\beta), \tag{3.7}$$

where

$$\hat{N}(\beta) := e^{-\beta \hat{a}^\dagger \hat{a}} = e^{-\frac{\beta}{2}(\hat{Q}^2+\hat{P}^2-1)} \tag{3.8}$$

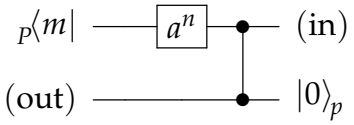
is the damping operator with $\beta \in \mathbb{R}$. Here, a squeezed momentum state is achieved by applying the squeezing operator

$$S(\xi) = e^{\frac{\xi}{2}(\hat{a}^{\dagger 2}-\hat{a}^2)} \tag{3.9}$$

to the vacuum state, denoted as a zero-valued ket with subscript $N$ denoting the photon-number eigenbasis. Eq. 3.7 can be implemented by coupling the top wire to a vacuum mode with a beamsplitter of transmissivity $t = e^{-\beta}$ and then performing PNR detection on the reflected mode as derived in Sec. 1.5.4. Additionally, Sec. 1.5.3 gives further information about the non-unitary damping operator.

Before analyzing this circuit in depth, we wish to simplify it to a mathematically idealized case where its usefulness is more readily apparent. First, we take the high squeezing limit such that we can approximate the squeezed vacuum as a zero momentum eigenstate, $\hat{S}(\xi)|0\rangle_N \approx |0\rangle_p$. Next, consider taking the limit $\beta \to 0$ in Eq. 3.7, which is physically equivalent to tuning the subtraction beamsplitter to have vanishing reflectivity. Measuring $n$ photons in the weakly reflected mode will result in the application of $\mathcal{S}_n \approx \hat{a}^n$, up to a normalization. In this limit, the probability to measure photon numbers with $n > 0$ vanishes; however, when properly combined with the high squeezing limit such that the beamsplitter reflectivity becomes small as the energy in the input state becomes very large, the probability to subtract zero photons can also vanish. Later, we will relax both constraints on the squeezing and beamsplitter reflectivity (value of $\beta$) to demonstrate that our method can perform successfully in a realistic scenario.

For now, consider the idealized non-Gaussian subtraction and teleportation circuit given by

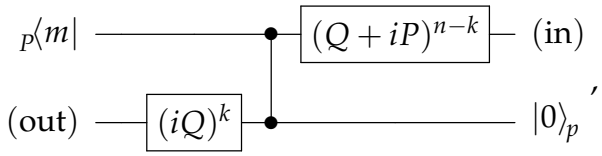



(3.10)

Writing the annihilation operator as $\hat{a} = \frac{1}{\sqrt{2}}(\hat{Q} + i\hat{P})$ and commuting this with the $\hat{C}_Z$ gate, we have that the operator applied to the output state is given by

$$\begin{aligned} &{}_{1P}\langle m|\frac{\hat{C}_Z}{\sqrt{2^n}}\left(\hat{Q}_1 + i\hat{P}_1 + i\hat{Q}_2\right)^n|0\rangle_{2P} \\ =&{}_{1P}\langle m|\frac{\hat{C}_Z}{\sqrt{2^n}}\sum_{k=0}\binom{n}{k}\left(i\hat{Q}_2\right)^k\left(\hat{Q}_1 + i\hat{P}_1\right)^{n-k}|0\rangle_{2P}. \end{aligned} \tag{3.11}$$

Since the $\hat{C}_Z$ gate commutes with $\hat{Q}$, we can graphically represent each term in the sum as the circuit




(3.12)

which is simply the teleportation circuit applied to a modified input followed by the application of $\hat{Q}$ to a power. Representing the action of this circuit as a Kraus operator applied to the input, we see that the resultant operator would be

$$\begin{aligned} \hat{K}_{m,n} &= \frac{1}{\sqrt{2^n}}\sum_{k=0}^{n}\binom{n}{k}\left(i\hat{Q}\right)^k\hat{R}(\tfrac{\pi}{2})\hat{Z}^\dagger(m)(\hat{Q} + i\hat{P})^{n-k} \\ &= \hat{X}(m)\hat{R}(\tfrac{\pi}{2})\frac{1}{\sqrt{2^n}}\sum_{k=0}^{n}\binom{n}{k}\left(-i\hat{P} + im\right)^k(\hat{Q} + i\hat{P})^{n-k}. \end{aligned} \tag{3.13}$$

The position shift of $\hat{X}(m)$ at the end can be effectively ignored, as its effect can be removed with feed-forward displacements or accounted for by shifting the result of subsequent homodyne detections [38]. Surprisingly, the $\hat{P}$ contribution in the sum vanishes and the operator can be written as $\hat{K}_{m,n} = \hat{X}\hat{R}f_n(\hat{Q})$, where $f_n(\hat{Q})$ takes the form of a polynomial with generally complex coefficients. This can be seen by examining the commutator of $\hat{K}_n$ with $\hat{Q}$. Furthermore, in the specific case where $m = 0$, Eq. 3.13 can be written in terms of a Hermite polynomial in $\hat{Q}$ as

$$\hat{K}_{0,n} = \hat{R}(\tfrac{\pi}{2})H_n(\tfrac{\hat{Q}}{\sqrt{2}}), \tag{3.14}$$

where $H_n(x)$ is the $n$-th order physicist's Hermite polynomial. This is derived in Appx. B.2. By iterating this procedure $M$ times successively, actively undoing the displacements between steps, and applying a regular teleportation at each intervening step to enact a Fourier transform and keep the overall operator a function of $\hat{Q}$ only, we can build up powers of $\hat{Q}$ to form the overall operator

$$\hat{K}_M = \hat{R}(M\pi) \prod_k f_k \left( (-1)^k \hat{Q} \right), \tag{3.15}$$

which is a polynomial in $\hat{Q}$ follow by a rotation of $M\pi$. The degree of this polynomial is equal to the total number of photons subtracted over all steps. It is worth pointing out that while we have presented this analysis from the standpoint of the idealized circuit in Eq. 3.10, including realistic effects will still result in the application of a polynomial in $\hat{Q}$ of the same order, but with additional Gaussian noise. Before continuing on to the full treatment, we wish to further motivate the use of applying polynomials of this form. Developing polynomial gates can be used to approximately implement non-Gaussian unitaries and the cubic phase gate [155, 156], but when specifically applied to squeezed vacuum, they can give rise to cat states [96], which is our main interest here.

**Polynomial operator applications**

As a segue to a fully realistic implementation, consider applying this idealized overall $\hat{K}_M$ operator to a cluster state node with finite squeezing, and consider the case where all homodyne measurements are null-valued (a constraint we will relax later), in which case $f_n(\hat{Q})$ is a Hermite polynomial. For moderate to large squeezing on the state to which the polynomial operator is applied, the leading power of $\hat{Q}$ in each Hermite polynomial will dominate as the anti-squeezing gives support over a large range of position in phase-space, thus eliminating the contribution of lower order terms upon normalizing the state after the operator is applied. This is demonstrated in Fig. 3.1, where the fidelity of $H_n(\hat{Q}/\sqrt{2})S(\imath)|0\rangle_N$ and $\hat{Q}^n S(\imath)|0\rangle_N$ is plotted as a function of squeezing. We use the fidelity between two density matrices $\rho$ and $\sigma$ defined as

$$F = \mathrm{Tr}\left[ \sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}} \right]^2. \tag{3.16}$$

For $n = 1$, the cases are trivially the same, and for larger $n$, the fidelity approaches one as squeezing increases. Thus, for the immediate sake of illustration, we can examine

the effects of applying only $\hat{Q}^M$ to a squeezed state, where $M$ is now the total number of photons subtracted over all iterations of the process.

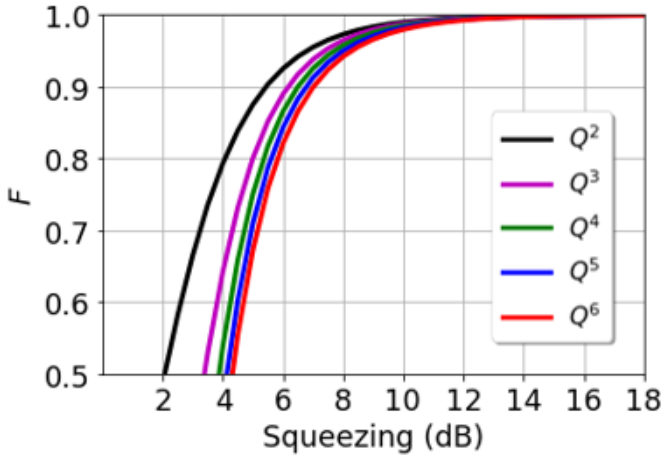



FIGURE 3.1: Application of Hermite polynomials in $\hat{Q}$ compared to leading order powers of $\hat{Q}$. Fidelity approaches unity as squeezing increases.

The Wigner functions of the resulting states after applying $\hat{Q}^M$ to squeezed vacuum for several values of $M$ are plotted in Fig. 3.2. As one can see the results resemble anti-squeezed Schrödinger Cat states, where increasing the power of $Q$ applied to the squeezed state increases the amplitude of the approximated cat. The closeness to cat states can be seen by looking at the form of the wavefunction. In the $Q-$basis, the state wavefunction is given by

$$\psi_Q(v) = {}_q\langle v|Q^M S(\iota)|0\rangle_N = \mathcal{N} e^{-(\frac{v^2}{2s^2})} v^M, \tag{3.17}$$

where $s = e^{\iota}$ and the normalization is given by

$$\mathcal{N} = s^{-M}\sqrt{\frac{1}{s(M-\frac{1}{2})!}}. \tag{3.18}$$

Transforming to the $P-$basis, we have

$$\psi_P(t) = \mathcal{N}' e^{(-\frac{s^2t^2}{2})} H_M(\tfrac{st}{\sqrt{2}}), \tag{3.19}$$

where $\mathcal{N}' = (-2)^{-\frac{M}{2}} s^{M+\frac{1}{2}} \mathcal{N}$ and $H_M(x)$ are Hermite polynomials. This wavefunction resembles that of a squeezed $M-$photon Fock state, but with the subtle difference that

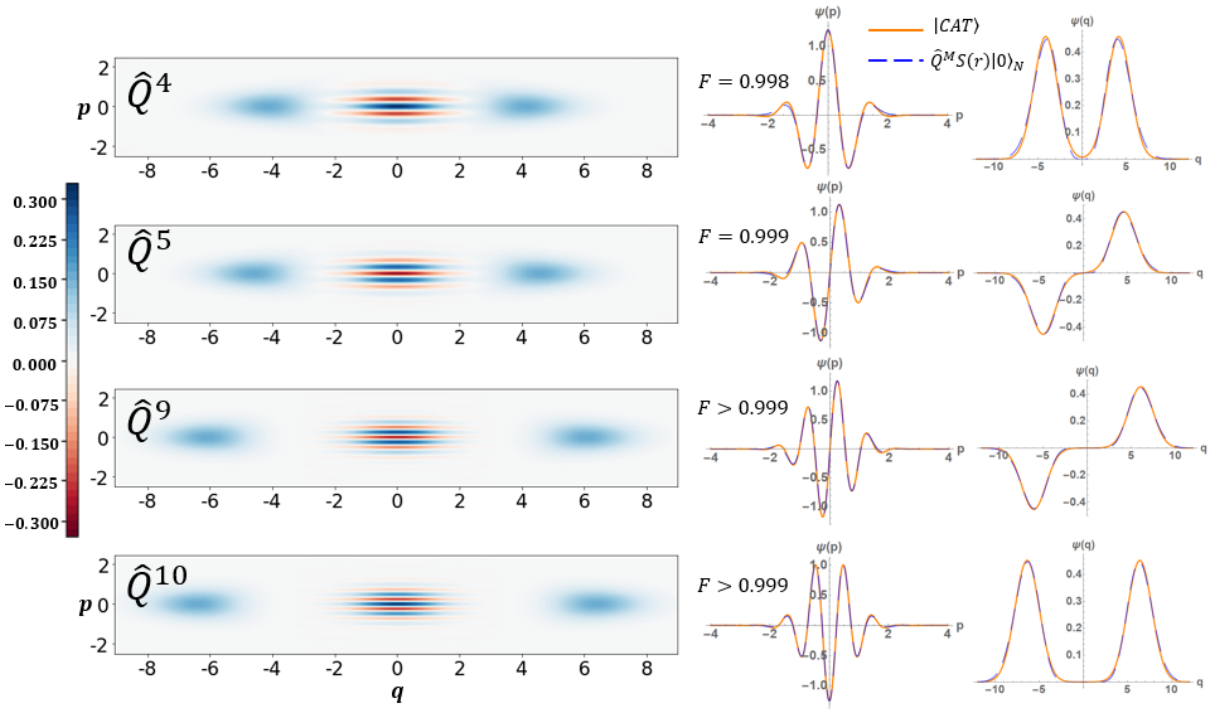



FIGURE 3.2: Idealized PhANTM steps with null homodyne measurement results iteratively apply powers of $\hat{Q}$ to an input state based on the total number of subtracted photons over all rounds of teleportation. Here, the input is a squeezed vacuum state with squeezing of 6 dB ($r = 0.7$). The final state is a nearly perfect match to an even (odd) squeezed cat state when an even (odd) number of photons was subtracted.

the argument of the Hermite polynomial is scaled by a factor of $2^{-1/2}$. No additional squeezing operation will transform it back to a Fock state as scaling the argument of the exponent will also scale the argument of the Hermite polynomial. Furthermore, this scaling is independent of the power of $Q$ applied. Examining this wavefunction shows that the number of ripples is determined by the order of the Hermite polynomial, which is based on the total photon subtraction counts. As shown in Fig. 3.2, a nearly identical anti-squeezed cat state can be found for each of these, where fidelities with cats having numerically optimized parameters are above 0.999. Additionally, it has previously been shown that the wavefunction in Eq. 3.17 asymptotically approaches a cat state as $M$ becomes large [96].

As a final step before proceeding to the full treatment, we relax the restrictions imposed on homodyne measurement and allow for arbitrary quadrature detection results. This leads us to the slightly more general operator given by a product of Hermite polynomials in $\hat{Q}$ with arguments shifted by the measurement result, which arises

from a limiting case of the general derivation that will be discussed later (Eq. B.38 in Appx. B.1). However, random homodyne measurement results do not pose a significant obstacle for cat state generation. This effect leads to producing cat states with a general phase between the 'classical' displacement components of the form:

$$|cat\rangle = \mathcal{N}\left(D(\alpha) + e^{i\phi}D(-\alpha)\right) S(\zeta)|0\rangle_N. \tag{3.20}$$

As shown in Fig. 3.3, protocols with stochastic homodyne detection generate cat states with a phase-space interference fringe that has been scanned in a direction perpendicular to the displacements. The Wigner functions plotted in Fig. 3.3 depict typical results of a multistage process where finite squeezing of 6 dB has been used for all ancillary inputs, and a single photon was subtracted in each teleportation iteration before a stochastic homodyne measurement.

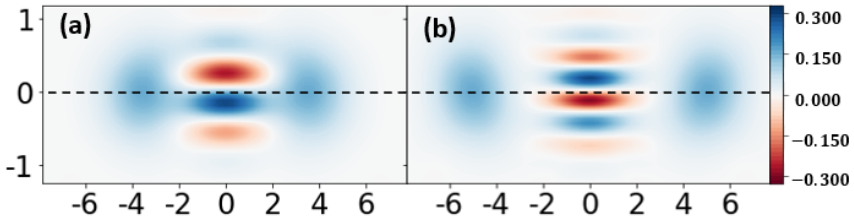
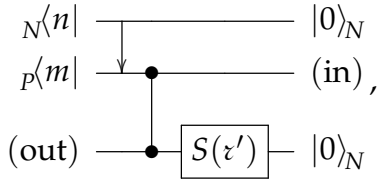



FIGURE 3.3: Typical results from iterated subtracted teleportation applied to squeezed vacuum with one photon subtracted each instance and stochastic homodyne results. The dotted black line is added as a guide to show the phase of the resultant cat state. Note that the figures have been rescaled to better visualize the fringes.

### 3.2.2 Realistic teleportation

Until now, we have only described an idealized case where teleportation proceeded with infinite squeezing and photon-number subtraction was modeled as a perfect application of the annihilation operator. The true circuit that represents realistic photon-subtraction during a cluster-state teleportation will instead appear as



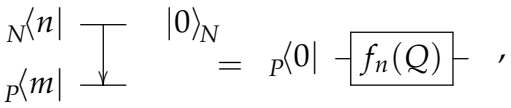

(3.21)

where the application of the annihilation operator is instead replaced by an additional circuit component to couple the input state in the middle wire to vacuum in the top wire with a beamsplitter,

$$\hat{B} = e^{\theta(\hat{a}_1\hat{a}_2^\dagger - \hat{a}_1^\dagger\hat{a}_2)}, \tag{3.22}$$

which is represented by the down arrow in the diagram. We take the beamsplitter to have real reflectivity and transmissivity such that $r = \sin\theta$ and $t = \cos\theta$. The momentum eigenstate in the third wire has also been replaced with a finitely squeezed vacuum state. In the limit of weak beamsplitter reflectivity, projecting the top wire into an $n$-photon Fock state after the beamsplitter effectively acts to apply $\hat{a}^n$, although the probability to successfully measure $n$ photons vanishes as the reflectivity drops to zero. Away from this limit, this portion of the circuit acts to apply non-unitary damping to the state in addition to a factor of $\hat{a}^n$, as discussed in Sec. 1.5.4.

The top portion of the circuit in Eq. 3.21 can be simplified to

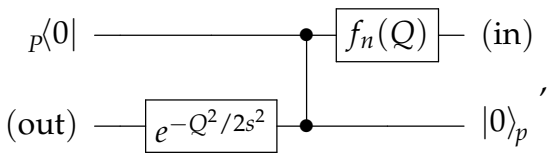



(3.23)

where $f$ is a function in $\hat{Q}$ only. On the bottom wire, we can write the input momentum-squeezed state as

$$\begin{aligned}\hat{S}(r')|0\rangle_N &= \pi^{-1/4}\int dt e^{-t^2/2s^2}|t\rangle_q \\ &= \pi^{-1/4}e^{-\hat{Q}^2/2s^2}\int dt|t\rangle_q \\ &= \pi^{1/4}\sqrt{\frac{2}{s}}e^{-\hat{Q}^2/2s^2}|0\rangle_p,\end{aligned} \tag{3.24}$$

where $s = e^{r'}$. This allows us to commute all operations in $\hat{Q}$ to either the back or front of the circuit to arrive at




(3.25)

which is recognizable as the same form as Eq. 3.1 with additional operators before and

after the teleportation. We can now directly write down the Kraus operator representation of the action of the circuit on the input state as

$$\hat{K}_n = \pi^{1/4}\sqrt{\frac{2}{s}}e^{-\hat{Q}^2/2s^2}\hat{R}(\tfrac{\pi}{2})f_n(\hat{Q}). \tag{3.26}$$

The function $f_n(\hat{Q})$ contains a displacement from the measurement result that may be useful to keep track of separately, so $f$ can be decomposed into a displacement term, a quadrature damping term, and a polynomial in $\hat{Q}$. Commuting the displacement term through the Fourier transform allows us to arrive at

$$\begin{aligned}\hat{K}_n =& \tfrac{\pi^{1/4}\sqrt{2}}{\sqrt{s}}e^{\frac{m^2}{s^2}\sigma^2}\hat{X}(m\sigma)\hat{R}(\tfrac{\pi}{2}) \\ &\times e^{-\frac{1}{2s^2}(\hat{P}-m\sigma)^2}e^{-\hat{Q}^2\frac{\sigma t(1-t^2)}{4}}f_n'(\hat{Q}),\end{aligned} \tag{3.27}$$

where we define $\sigma = \frac{2t}{1+t^2}$. This is derived in detail in Appx. B.1, where we find that

$$f_n'(x) = \tfrac{i^n}{\sqrt{2^{n-1}n!(1+t^2)^{n+1}}}\sum_{k=0}^{n}\sum_{j=0}^{\lfloor\frac{n-k}{2}\rfloor}\binom{n}{k}\binom{n-k}{2j}\tfrac{(2j)!r^{2j}}{2^j j!}\left(\tfrac{1+t^2}{2}\right)^{k/2}(-i)^{k+2j}H_k(\tfrac{-x\sqrt{2}rt^2}{1+t^2})H_{n-k-2j}(\tfrac{-mr}{t\sqrt{1+t^2}}) \tag{3.28}$$

is a polynomial of degree $n$. In the limit of weak beamsplitter reflectivity and large squeezing so that $r \to 0$, $t \to 1$ and $s \to \infty$, the Kraus operator reduces to

$$\hat{K}_n \propto \hat{X}(m)\hat{R}(\tfrac{\pi}{2})H_n\left(\frac{i\hat{Q}-m}{\sqrt{2}}\right). \tag{3.29}$$

Subtracting photons during the teleportation process is stochastic in nature, where the probability of any particular $n$-photon subtraction and $m$ homodyne measurement occurring is given by

$$P_n = \mathrm{Tr}[\hat{K}_n\rho_{in}\hat{K}_n^\dagger], \tag{3.30}$$

where $\rho_{in}$ is the general quantum state being sent through the PhANTM gadget, and the final state is evolved to

$$\rho_{out} = \frac{\hat{K}_n\rho_{in}\hat{K}_n^\dagger}{P_n}. \tag{3.31}$$

As in the idealized case, we assume that one has the ability to undo the accumulated displacement operations. If the photon subtraction was successful, then any Gaussian input has been de-Gaussified. If, however, no photons were subtracted, the

state will accrue additional Gaussian noise due to the finite squeezing in the cluster state and the non-vanishing beamsplitter reflectivity used to perform the subtraction. For large enough squeezing and low beamsplitter reflectivity, this additional Gaussian noise can be made small.

Whether subtraction was successful or not, teleporting the resultant state through the next node of the quantum wire without attempting to subtract will enact another $\pi/2$ rotation and re-align the squeezed-axis of the state, at which point another PhANTM step can be performed. Provided there is sufficient squeezing in the cluster state, this process of subtracting photons through teleportation can build up non-Gaussianity faster than the Gaussian noise at each step can degrade the state and result in large, high-fidelity cat states.

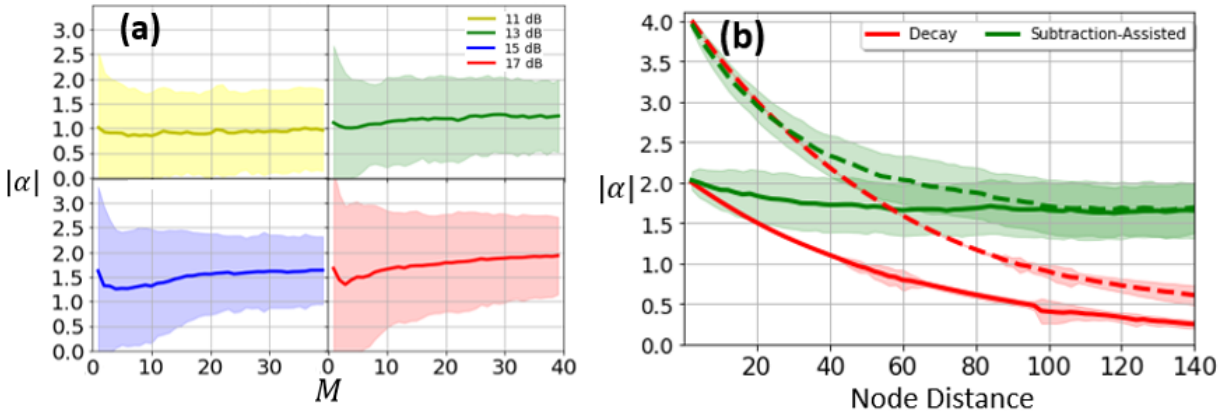



FIGURE 3.4: (a) Solid lines show the average cat amplitude obtained from numerical simulations of performing $M$ iterations of the subtraction protocol on a 1D cluster state where the homodyne and PNRD measurements were randomly sampled from the state distribution at each step. The shaded regions shows $\pm 2\sigma$ from the sampled data and squeezing values indicate the squeezing level of each node in the initially prepared cluster state. (b) Cat states with respective amplitudes of $|\alpha| = 2$ and $|\alpha| = 4$ introduced to a cluster state can be protected as they are teleported through the cluster if subtraction is attempted every other teleportation step to prevent state decay. Shading shows $\pm 1\sigma$.

## 3.3 Results

### 3.3.1 Cat states

In this section, we present the results of numerical calculations demonstrating that iteratively performing the PhANTM algorithm on cluster states can generate and stabilize large cat states. Even when considering finite squeezing effects and the probabilistic nature of the Kraus operator given by Eq. 3.27, repeatedly applying operators of this form gradually shapes the teleported quantum information into cat states and preserves cat states already present in the cluster.

For these calculations, the python packages of QuTip [129, 157] and Strawberry Fields [130, 158] were used to perform simulations of quantum state evolution using the Rivanna high-performance computing system at the University of Virginia. A single step of the PhANTM along a 1D cluster state was first simulated by sampling the probability distribution for detection outcomes from the circuit shown in Eq. 3.21, and then the input state was updated accordingly by applying the corresponding Kraus operator from Eq. 3.26. This protocol was applied in succession, with a regular teleportation to reapply a $\pi/2$ rotation occurring between PhANTM steps to simulate the algorithm along several nodes of a cluster state. Cluster states with different squeezing parameters were examined, and statistics were collected by simulating the entire procedure repeatedly.

In each case, the value of beamsplitter reflectivity for the photon subtraction step was chosen so that the added Gaussian noise from subtraction effectively adds the same amount of noise as finite squeezing in the cluster. The condition is given by

$$\cos^2\theta = \tanh|z'| \tag{3.32}$$

where $z'$ is the squeezing parameter for the cluster state nodes. This condition arises by realizing that applying the damping operator, $\hat{N}(\beta)$, to a zero-momentum eigenstate is equivalent to finitely squeezing the vacuum state, as has been derived previously [18] and discussed further in Appx. 1.5.3. Because squeezing can instead be thought of as applying the damping operator to a momentum eigenstate and the damping operator is also what separates the realistic photon subtraction operator $\hat{\mathcal{S}}_n$ from ideally applying annihilation operators, implementing the above condition essentially makes the real circuit in Eq. 3.6 different from the idealized circuit in Eq. 3.10 only by the inclusion of

damping operators applied both to the ancilla and before homodyne detection.

For each trial, we started with a fresh 1-D cluster state and applied many steps of the PhANTM along the cluster while allowing stochastic measurement results. After each step in the simulation, the evolved state was fitted to the nearest cat state of the form of Eq. 3.20 by optimizing the fidelity. Fidelity and fit values for all trials with a given set of initial cluster state squeezing parameters were averaged at each step, $M$, and compared to trials with different squeezing parameters.

The results are depicted in Fig. 3.4(a) for cluster states with $11 - 17$ dB of squeezing at each node. The average fitted value of the nearest cat state amplitude (solid lines), $|\alpha|$, is plotted against the number of times that a PhANTM had occurred, where each step corresponds to consuming two physical nodes of the cluster due to the necessary intervening teleportation step between PhANTMs to apply a Fourier transform. After $M = 10$ steps, the average fidelities with the associated cats are greater than 0.98 in all cases examined. The shaded regions indicate two standard deviations in the spread of obtained values from 100 trials.

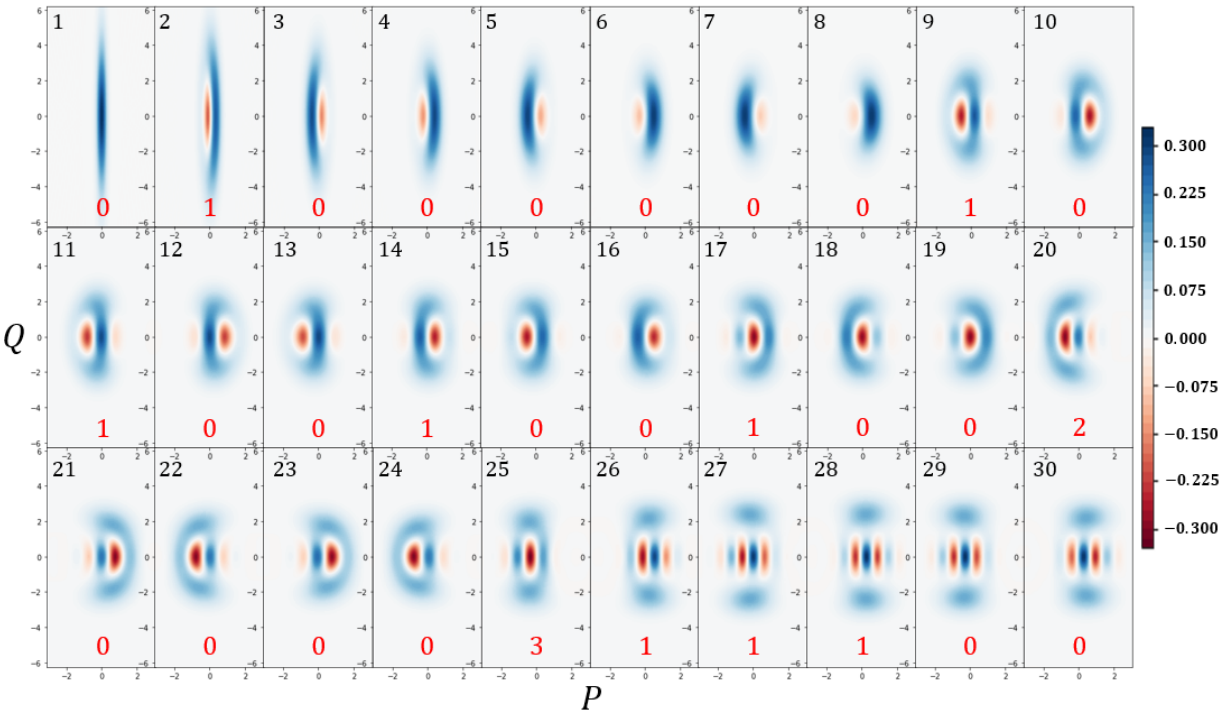



FIGURE 3.5: Typical results from iterated PhANTM where measurement results are stochastic. Each Wigner function plotted shows what state is now $\hat{C}_Z$'ed to the cluster after the $M$-th step (black numbering) where the number of subtraction photons from the corresponding step is shown in red.

An example evolution of a single trial as quantum information is teleported along the canonical cluster state is shown in Fig. 3.5. Steps without attempting photon subtraction are used for teleporting the state and applying a $\pi/2$ rotation to reorient the state. The Wigner function for the state at each step where subtraction was attempted is shown, with the number of actual photons subtracted listed in red. Each teleportation takes into account finite squeezing of 15 dB in the cluster state, randomly sampled homodyne and subtraction measurement outcomes, and a subtraction beamsplitter with reflectively chosen to have the same order damping effect as finite squeezing.

After just a few steps, it is possible to generate a strongly anti-squeezed cat state of weak amplitude, but this is highly probabilistic and dependent on the precise subtraction measurements. However, if there is sufficient squeezing in the cluster state, each PhANTM on average succeeds in adding more non-Gaussianity than is washed away by Gaussian noise, and thus there is a high probability of obtaining a cat state with reasonable amplitude after several measurement steps. For 17dB of cluster state squeezing, 30 steps of PhANTM are sufficient to succeed in generating a high-fidelity cat state with amplitude $|\alpha| > 1$ with a 95% success rate. It should be noted that although anti-squeezed cats are obtained in the idealized case of Fig. 3.2, the effects of finite squeezing in the cluster and overall stochastic nature of the photon subtraction result in cat states very low anti-squeezing ($< 3$ dB) beyond the first few steps, which is the region of interest.

#### Stabilization of cat states

In addition to generating cat states from a Gaussian resource using only local measurements, the PhANTM protocol can also be used to protect cat states already present within the cluster state. Suppose one has a cat state embedded in a cluster state, either generated through the PhANTM as detailed above or offline through some other means, and entangles it to a cluster state with a $\hat{C}_Z$ interaction for later use. If this state needs to be moved through the cluster to a suitable location, it will suffer amplitude decay from Gaussian noise at each teleportation step, as shown by the red lines in Fig. 3.4(b). If one applies the PhANTM algorithm in place of regular teleportation at every-other node, then the cat state can instead be preserved (green curves) to a level dependent on the squeezing present in the cluster.

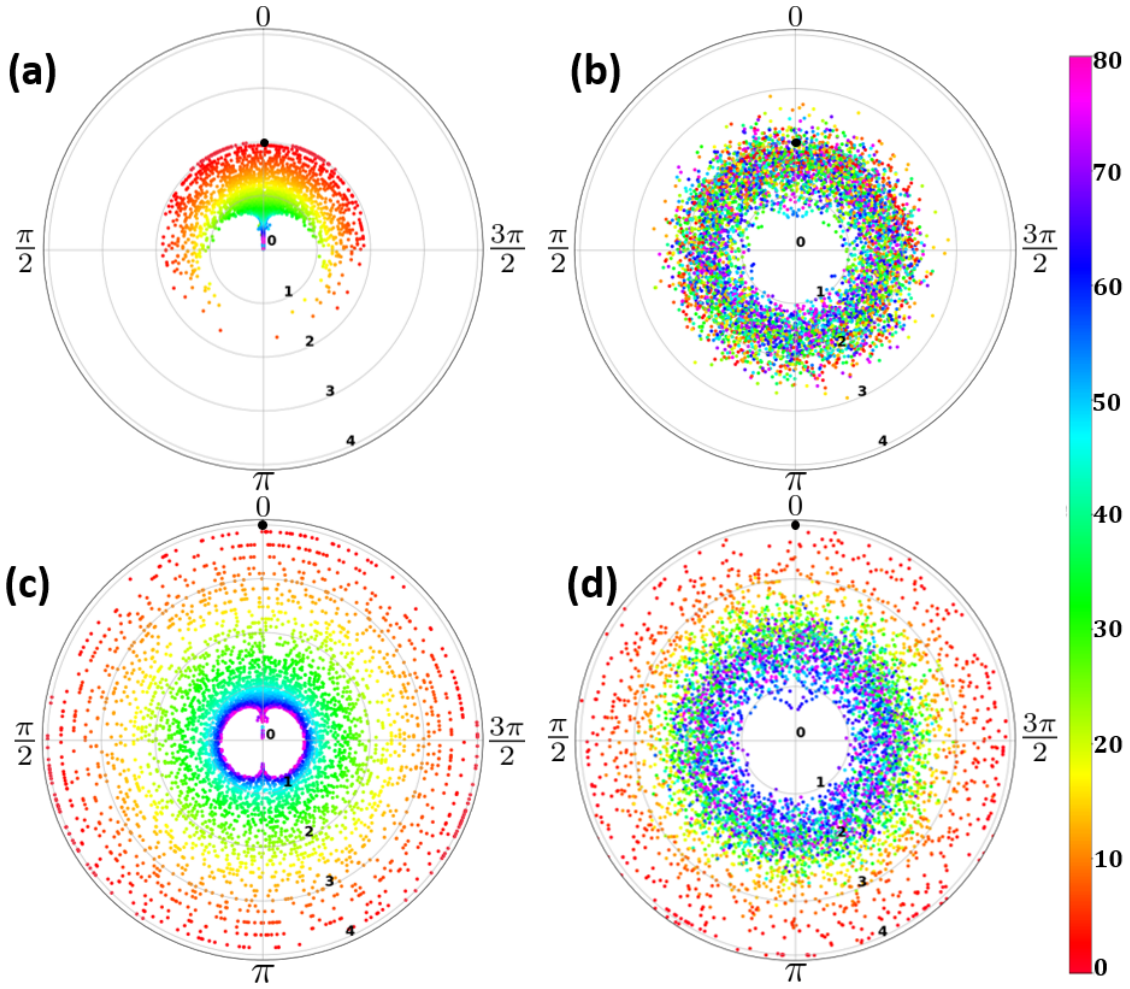



FIGURE 3.6: Alternate representation of the data in Fig. 3.4b for 80 trajectories at each initial condition, where the radial direction indicates cat state amplitude, $|\alpha|$, the polar coordinate gives cat state phase, $\phi$, and the color shows the number of node-teleportations from the the initial conditions (black dots). (a) and (c) depict regular teleportation through a cluster state where cat states have initial amplitude $\alpha = 2$ and $\alpha = 4$, respectively. (b) and (d) have the same initial conditions as at left, but now PhANTM is performed as opposed to regular teleportation. A random (but known) phase kick is applied at each step by homodyne measurement in all cases, but one can see that PhANTM preserves cat state amplitude up to a level dependent on cluster-state squeezing (here 15dB).

Under this paradigm, including the PhANTM with a Gaussian cluster state can be seen as a machine that *perpetuates* non-Gaussianity even in the presence of Gaussian noise. There exists a stable equilibrium at which some amount of non-Gaussianity will remain, which was previously only formulated for CV cluster states with fault-tolerant

GKP error correction [74]. This protocol can thus work as a pseudo quantum memory for cat states when paired with a time-domain cluster state, as an input cat state can be teleported through many time-separated cluster state nodes and retrieved later. The retrieved state will have randomized but known parameters dependent on specific measurement results at each step, and the retrieved state will remain within the cat-state manifold.

This is demonstrated more clearly in Fig. 3.6, which depicts the same data as shown in Fig. 3.4(b). Fig. 3.6 shows the evolution of the phase, $\phi$, as the polar coordinate and amplitude, $|\alpha|$, as the radial value according to the fit with a cat state of the form of Eq. 3.20 for cat states as they are teleported across a cluster state. In each plot, the large black dot shows the parameters of the initial cat state introduced to the cluster which was taken to be an even cat ($\phi = 0$) with no squeezing and amplitude of $\alpha = 2$ (a and b) or $\alpha = 4$ (c and d). In Figs. 3.6(a) and (c), the cat states were simply teleported normally using $P-$basis homodyne measurements. Figs. 3.6(b) and (d) begin with the same initial cat state but depict the effects of including PhANTM during teleportation across the cluster. As the states were teleported further along the cluster state, as indicated by color to represent node-distance, the amplitude of the cat states decayed as seen by the final magenta dots being clustered about the origin in Figs. 3.6(a) and (c). When the PhANTM was included, the cat state amplitudes at no point spiraled to the center, but instead became clustered about an orbit at radius $|\alpha| \approx 1.5$, thus preserving the cat state and the associated non-Gaussianity.

At each teleportation, finite-squeezing effects introduce Gaussian noise that dampens the amplitude of the cat state and introduces measurement-dependent randomization to the phase, which is also dependent on the amplitude of the cat state. This effect can be understood by realizing that as the cat state amplitude increases, the interference fringes between the classical coherent state portions that determine the phase, $\phi$, oscillate at a higher frequency. Due to finite squeezing, the teleportation step applies a Gaussian envelope not centered about the origin in phase space, but instead centered about the measurement outcome. Shifts to this Gaussian envelope cause a slight 'scanning' effect of the interference fringe, which changes the phase at each step more drastically for larger cat states. Because this phase-randomization is known from previous homodyne measurement results, a feed-forward displacement can be applied to shift the cat and realign the fringes to reset the phase.

Examining Fig. 3.6(a) and (c), we can also see that the amplitude decays to zero

through the poles of the plot when $\phi = 0$ or $\phi = \pi$. This can be understood by realizing that as the cat shrinks, the interference fringes become dampened and very wide, so that a phase cannot be easily determined when the oscillations begin to vanish. These are the two extremes of shrinking cat states: either the even ($\phi = 0$) cat, which must tend toward the vacuum state as $|\alpha| << 1$, or the odd ($\phi = \pi$) cat, which tends toward a squeezed single photon.

Additionally, it is important to note that finite squeezing dictates a threshold at which the cat state amplitude saturates. If one wishes to achieve non-Gaussianity beyond this threshold in a given cluster, e.g., larger amplitude cats, it will be necessary to make use of distillation-type protocols. Many such protocols exist to take cat states and breed them to generate more non-Gaussianity [82, 83, 99, 102, 133, 135, 145, 159], and we will show later how analogous protocols can be developed to breed cat states within the cluster state architecture.

**Higher-dimensional clusters**

As demonstrated above, the PhANTM can be used to transform a 1-D quantum wire of sufficient length into a cat state. A simple extension of this idea is to apply the algorithm on any 1-D topological wire connected to a higher dimensional cluster state to produce a cat state embedded within the remainder of the cluster. With this method, the algorithm can be performed simultaneously on many quantum wires to engineer a non-Gaussian cluster state with cat states embedded at strategic locations. Each of these cat states will have an average amplitude dependent on the initial cluster state squeezing and the number of nodes in each 1-D wire consumed by the PhANTM algorithm. Gaussian operations can then be performed as per the usual cluster state formalism with homodyne measurements to manipulate the exotic quantum states as desired, such as perhaps breeding the cat states as we will discuss further in the next section.

The process of embedding cat states in a higher-dimensional cluster can be visualized by Fig. 3.7, where a 2D cluster state is reduced to a 1-D cluster peppered with cat states. The PhANTM can be applied to all chains simultaneously, but measurements within each chain must proceed in the correct order — top down in this example.

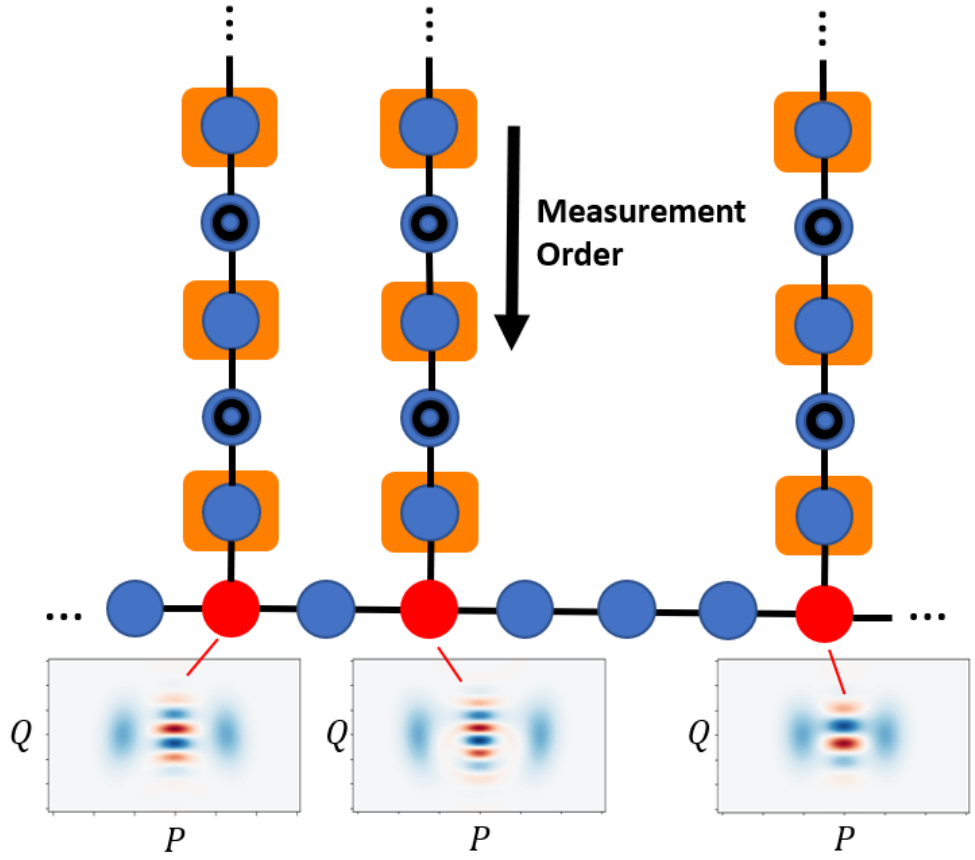



FIGURE 3.7: Measurements on 1-D quantum wires can be used to generate cat states embedded within a higher-dimensional cluster state. Orange blocks indicate a step of the PhANTM algorithm (attempted photon subtraction, homodyne detection, and feed-forward displacement) and black rings indicate $P$-basis homodyne measurements. The red nodes represent fully de-Gaussified cat states.

## 3.4 Breeding Cat States

It has long been known that cat states can be probabilistically bred into larger amplitude cat states [99, 102, 133, 135, 145, 159], and recently it was demonstrated that a supply of squeezed cat states can be deterministically bred into a class of grid states [83]. These procedures were formulated in terms of beamsplitter interactions, but each method can proceed analogously by performing homodyne measurements on cat states embedded in a cluster state. Instead of teleporting through a momentum eigenstate, each cat state is teleported through another cat state to produce the desired effect. Just as teleporting through a finite squeezed state leads to filtering of the teleported state with a Gaussian envelope [40, 44], teleporting through an ancillary squeezed cat state filters the $P$-basis

wavefunction of the input with the $Q$-basis wavefunction of ancilla shifted by the homodyne result. This is seen generically by replacing the momentum eigenstate from the teleportation circuit in Eq. 3.1 with an arbitrary quantum state, $|\psi'\rangle$. This can be rewritten as an operator in $\hat{Q}$ applied to a zero momentum eigenstate and commuted through the $\hat{C}_Z$ gate to have the general circuit

$$|\psi\rangle_{out} = \quad {}_P\langle 0| \;\text{—}\bullet\text{—}\; \boxed{Z^\dagger(m)} \;\text{—}\; |\psi\rangle \quad \boxed{\sqrt{2\pi}\psi'_Q(Q)} \;\text{—}\bullet\text{—}\; |0\rangle_P \quad . \tag{3.33}$$

Pulling the circuit taught and commuting the resultant Fourier transform to the end, we have that

$$|\psi\rangle_{out} = \sqrt{2\pi}\hat{R}(\tfrac{\pi}{2})\psi'_Q(-\hat{P})\hat{Z}^\dagger(m)\,|\psi\rangle\,, \tag{3.34}$$

where the subscript on $\psi'_Q(-\hat{P})$ indicates that it is the $Q$-basis wavefunction of the starting ancillary state, even though it is a function of the $\hat{P}$ operator. Finally, commuting the displacement through to the left reveals the result of

$$|\psi\rangle_{out} = \sqrt{2\pi}\hat{X}(m)\hat{R}(\tfrac{\pi}{2})\psi'_Q(-\hat{P}+m)\,|\psi\rangle\,. \tag{3.35}$$

This shows that up to an overall displacement and rotation, the initial $P$-basis wavefunction for $|\psi\rangle$ transforms as $\psi_P(x) \to \psi_P(x)\psi'_Q(m-x)$. Thus, in the $P$-basis, $|\psi\rangle$ is filtered by the $Q$-wavefunction of $|\psi'\rangle$.

For the sake of illustration, supposed we input two equivalent squeezed cat states into the circuit but with one Fourier transformed with respect to the other, so that their wavefunctions are each proportional to the sum of two Gaussian peaks,

$$\psi_P(x) = \psi'_Q(x) = \mathcal{N}\left(e^{-\frac{1}{2s}(x-\alpha)^2} + e^{-\frac{1}{2s}(x+\alpha)^2}\right), \tag{3.36}$$

where

$$\mathcal{N} = \frac{\pi^{-1/4}\sqrt{s}}{\sqrt{2+2e^{-2|\alpha|^2}}}. \tag{3.37}$$

Supposing the measurement result is $m = 0$, we have that up to an overall rotation, the new wavefunction, $\phi$, is

$$\begin{aligned}\phi_P(x) &= \psi_P(x)\psi'_Q(x)\\ &\propto e^{-\frac{1}{s}(x-\alpha)^2} + e^{-\frac{1}{s}(x+\alpha)^2} + 2e^{-\frac{1}{s}(x^2+\alpha^2)}.\end{aligned} \tag{3.38}$$

For peak separations of $\alpha \gtrsim 1$, the third term becomes negligible, and the wavefunction is the same as the initial wavefunction $\psi$, but with narrower peaks. If we were to apply a squeezing operation to bring the peak widths back to the starting width, the peak separation would increase, and it is evident that this is just a larger cat state than what we started with. The above process is an example of breeding to enlarge cat states.

Now, suppose the inputs were the same cat states as above, but first each mode was Fourier transformed so that

$$\psi_P(x) = \psi'_Q(x) = \mathcal{N} e^{-\frac{1}{2}s^2x^2} (2\cos\alpha x). \tag{3.39}$$

Again taking a measurement outcome of $m = 0$ for illustration, the output wavefunction, $\phi$, from this process before the rotation would be

$$\phi_P(x) \propto e^{-s^2x^2} \cos\alpha x^2, \tag{3.40}$$

which in the $Q$-basis is

$$\phi_Q(x) \propto \left( e^{-\frac{1}{s}(x-\alpha)^2} + e^{-\frac{1}{s}(x+\alpha)^2} + 2e^{-\frac{1}{s}x^2} \right). \tag{3.41}$$

The above equation shows that with this case, the final state is now a superposition of three peaks with a binomial distribution as opposed to the two-peak superposition of the starting cat states. One can see how repeating this process can eventually give rise to a state with many peaks in the superposition which can approximate a grid state. These two cases are depicted in Fig. 3.8, where the process at left breeds to enlarge cats and the process at right breeds to create grids.

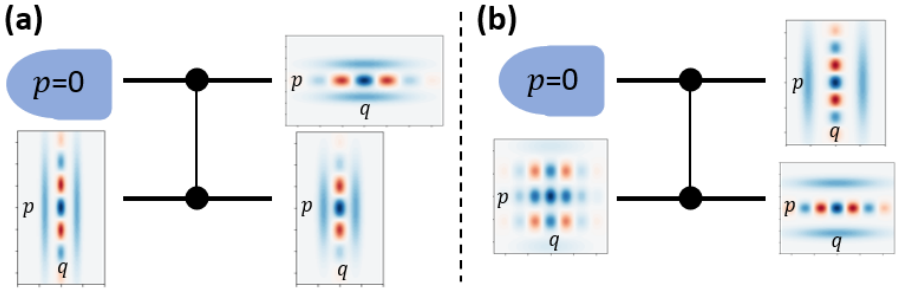



FIGURE 3.8: Cat states can be bred by teleporting one through the other to produce larger amplitude cat states as shown in (a) or grid states as shown in (b). The type of bred state is dependent on input cat state orientation.

As one can see, the output state depends on the phase of the input cat states, which is known from the iterated preparation method and can be controlled by performing regular teleportations to enact $\hat{R}(\pi/2)$ gates. Cat state breeding has been described in detail elsewhere [82, 83, 99, 102, 133, 135, 145, 159], but the point of interest here is that it can be done *within* the cluster state using only Gaussian measurements once all initial cats have be embedded within the cluster. Previous methods have examined breeding with beamsplitters and homodyne measurements in great detail, and we show in Appx. B.3 how the beamsplitter breeding maps directly to breeding with $\hat{C}_Z$ gates, where the only difference is that one input must be first Fourier transformed and the output has an additional Gaussian operation that can be undone with feed-forward displacement and Gaussian information processing on the cluster state. This mapping implies that all previous results based on breeding cat states with beamsplitters holds within the canonical cluster state.

One may ask why including a probabilistic scheme to enlarge cats is necessary when we have already demonstrated a scheme to generate cat states with near-unity success after enough iterations. However, it is important to remember that the mean amplitude of cat state produced by the PhANTM protocol is limited in part by the squeezing present in the cluster state. Breeding, however, allows for one to have a chance at sacrificing pairs of smaller cat states to enlarge them. By performing PhANTM in paralell on several 1-D chains, one could foreseeable have enough weak cats that it would be advantageous to attempt breeding them into larger resource states.

### 3.4.1 GKP states

Instead of taking a pair of cat states and breeding to generate a larger cat state, one can breed with a different outcome in mind, such as the generation of grid states. For this purpose, we more closely examine the case at right in Fig. 3.8, which when repeated, can succeed without post-selection as described previously [83]. Next, we will provide the Kraus operator and resultant state after an arbitrary breed step.

#### Breeding on the cluster state

Grid state breeding with beamsplitters has been discussed in detail in Ref. [83]. Here we provide the analogous derivation for cluster states. Suppose we have a superposition of evenly spaced Gaussian peaks, which can generally be described by a product of

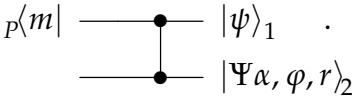
cat-like operators acting on squeezed vacuum as

$$|\Psi_{\alpha,\varphi,r}\rangle \propto \hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi})\hat{S}(r)\,|0\rangle\,. \tag{3.42}$$

where we define

$$\hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi}) := \prod_{k}^{N}\left(\hat{Z}(\alpha) + e^{i\varphi_k}\hat{Z}^{\dagger}(\alpha)\right). \tag{3.43}$$

If one teleports through this state as opposed to the normal momentum squeezed state, then the circuit will appear as




. (3.44)

Writing the squeezed vacuum of mode two as an operator in $\hat{Q}$ applied to a momentum eigenstate as in Eq. 3.24 and commuting all remaining operators on mode two through the $\hat{\mathrm{C}}_Z$ reveals that the effect of this circuit is to apply the operator

$$\hat{\mathcal{G}} = \hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi})e^{-\hat{Q}^2/2s^2}\hat{R}(\tfrac{\pi}{2})\hat{Z}^{\dagger}(m), \tag{3.45}$$

where $s = e^r$. Suppose the input to this circuit was of the same form as Eq. 3.42, but first having undergone a Fourier transform. The transformed state would then be

$$\begin{aligned}|\phi\rangle &\propto \hat{\mathcal{G}}\hat{R}(\tfrac{\pi}{2})\,|\Psi_{\alpha',\varphi',r'}\rangle \\ &\propto \hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi})e^{-\hat{Q}^2/2s^2}\hat{X}(m)\hat{R}(\pi)\hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi'})\hat{S}(r')\,|0\rangle\,.\end{aligned} \tag{3.46}$$

Making use of the relations that

$$\hat{Z}(\alpha_k)\hat{X}(\alpha_j) = e^{i\alpha_k\alpha_j}\hat{X}(\alpha_j)\hat{Z}(\alpha_k), \tag{3.47}$$

$$e^{-\hat{Q}^2/2s^2}\hat{X}(m) = e^{m^2/2s^2}\hat{X}(m)e^{-\frac{1}{s^2}\left(m\hat{Q}+\hat{Q}^2\right)}, \tag{3.48}$$

we can write the evolved state as

$$|\phi\rangle \propto \hat{X}(m)\hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi}-2m\alpha)\hat{\mathcal{D}}(-\alpha,\boldsymbol{\varphi'})e^{-\frac{1}{s^2}\left(m\hat{Q}+\hat{Q}^2\right)}\hat{S}(r')\,|0\rangle\,, \tag{3.49}$$

where $\boldsymbol{\varphi}-2m\alpha$ indicates that each phase in the expanded product has been shifted by the measurement result and the value of $\alpha$ to be $\varphi_k \to \varphi_k - 2m\alpha$. With the help of Eq. 3.24, the exponentiated quadratic in $\hat{Q}$ applied to squeezed vacuum can be rewritten

as

$$\begin{aligned} e^{-\frac{1}{s^2}\left(m\hat{Q}+\hat{Q}^2\right)}\hat{S}(r')\left|0\right\rangle &= e^{-\frac{1}{s^2}\left(m\hat{Q}+\hat{Q}^2\right)-\frac{1}{s'^2}\hat{Q}^2}\left|0\right\rangle_p \\ &= \hat{X}(B)e^{-A\hat{Q}^2}\hat{X}^\dagger(B)\left|0\right\rangle_p \\ &= \hat{X}(B)e^{-A\hat{Q}^2}\left|0\right\rangle_p \end{aligned} \tag{3.50}$$

where $s' = e^{r'}$ and

$$A = \frac{1}{2}\left(\frac{1}{s^2}+\frac{1}{s'^2}\right) \tag{3.51}$$

$$B = \frac{ms^2s'}{s^2+s'^2}. \tag{3.52}$$

Commuting the obtained position-quadrature shift to the left and converting the exponentiated $\hat{Q}^2$ back to a squeezer applied to vacuum, we have

$$\begin{aligned} \left|\phi\right\rangle \propto& \hat{X}(m+B)\hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi}-2(m+B)\alpha) \\ &\times\hat{\mathcal{D}}(-\alpha,\boldsymbol{\varphi}'+2B\alpha)\hat{S}(r'')\left|0\right\rangle \\ \propto& \hat{X}(m+B)\hat{\mathcal{D}}(\alpha,\boldsymbol{\varphi}'')\hat{S}(r'')\left|0\right\rangle, \end{aligned} \tag{3.53}$$

which is exactly a state of the form of Eq. 3.42 with an additional displacement and modified squeezing

$$r'' = r+r'-\frac{1}{2}\ln\left(e^{2r}+e^{2r'}\right) \tag{3.54}$$

The process of performing alternating PhANTM steps as detailed in the previous section acts to produce cat states on the cluster state, but breeding useful grid states will require squeezed cats. This is no obstacle, however, as any single-mode Gaussian operation can be implemented with a series of four homodyne measurements on the cluster [39]. For the specific case of squeezing, it suffices to apply three successive shear gates of $e^{\frac{i\gamma_k}{2}\hat{Q}^2}$ with properly selected $\gamma_k$ to effectively apply the gate $\hat{S}(\imath)$. For squeezed parameter $\imath$, the values of $\gamma_k$ are $\gamma_1 = \gamma_3 = \imath$ and $\gamma_2 = \imath^{-1}$ Note that the maximum squeezing operation that can be applied in this way is limited by the squeezing level of the cluster.

Suppose we begin with a cluster state consisting of several quantum wires and perform many steps of the PhANTM protocol on each quantum wire until we have successfully generated a cat state, which is now entangled to the remaining portion of the cluster state. Each of these can be squeezed through a series of three homodyne

measurements, and then pairs can be bred to make grid states without post-selection. An example of the resultant grid state is shown in Fig. 3.9(a), where we have used states generated from average results after $M = 35$ steps of the PhANTM protocol in a cluster of 17 dB squeezing, squeezed each one with three shear operations, and bred the result. Gaussian noise due to finite squeezing is included in all parts of the calculation to simulate realistic application of information processing on the cluster state.

To have a benchmark with some consistency, the target state was chosen to be the GKP qunaught, which is the GKP state with equal grid spacing in both quadratures. In order to make this state, the spacing of the initial cats must be correctly chosen so that the final spacing of the grid will be symmetric. This was achieved by tuning the shearing parameters used for the applied squeeze operation.

The asymmetries in the Wigner function shown arise from the stochastic homodyne measurements used when applying the squeezing with cluster state processing. The final homodyne measurement for the breed step was post-selected on zero to achieve consistent states within the approximate GKP family to be used for direct comparison of the quality of the state. Varying the measurement result will not change finial grid spacing or width in the $Q$-quadrature, but the introduced phase will alter the symmetry between the quadratures. Nevertheless, a grid state will always be created under asymptotic breeding regardless of measurement results [83]. We repeated this procedure for each cat state created with the PhANTM algorithm using the same data as shown in Fig. 3.4(a). Each cat state was squeezed with quantum information processing on the cluster state, and then bred with a copy to produce a similar grid state to the one shown. The resultant state was then fitted to a GKP qunaught with varying peak widths, $\Delta$. The resulting histogram is shown in Fig. 3.9(b), where the fidelity with the matched grid state averaged above 0.95, but was above 0.98 for all states with $\Delta < 0.6$.

With sufficiently large cat states and high squeezing in the cluster state, breeding within the cluster allows for the generation of a supply of GKP states that are embedded within the computational resource. These can then be teleported throughout the cluster where they can then be used for fault-tolerant error correction [74, 160] and universal QC [69], or alternatively, states can be teleported through the embedded GKP state to enact error correction directly [161]. Thus, while GKP states may be essential for error correction on cluster states, cluster states may also be valuable ingredients for the synthesis of GKP states.

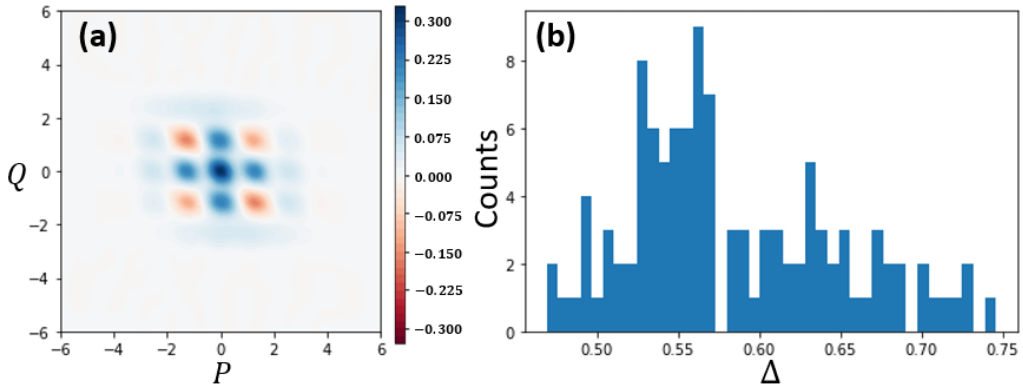



FIGURE 3.9: (a) Average GKP qunaught obtained by performing Gaussian operations and measurements within the cluster state on a pair of states generated from the PhANTM algorithm. (b) Histogram of $\Delta$ fit parameters for the GKP qunaughts bred from all states obtained by the PhANTM algorithm.

### 3.4.2 Phase estimation

Quantum phase estimation for unitary operator $\hat{U}$ is a procedure to measure the eigenvalue of the operator, $e^{i\theta}$, and project the input state into the corresponding eigenstate, $\hat{U}\left|\psi_\theta\right\rangle = e^{i\theta}\left|\psi_\theta\right\rangle$. Phase estimation, which can proceed through several methods [162], was proposed to be implemented with ancilla qubits to generate GKP states [163]. This has been successfully demonstrated with microwave cavity fields and transmon qubits [164]. A similar phase-estimation scheme without qubits was developed to create GKP states from cat states [83], which is the method we connect to above for breeding cat states. We now complete the connection to phase-estimation on a cluster state for unitary operators that are phase-space displacements.

For a general unitary operation $\hat{U} = e^{i\hat{\mathcal{T}}}$, a single round of phase estimation can be implemented using an ancillary qubit described by the circuit

$$\begin{array}{lcl} \text{(out)} & \boxed{Ue^{i\varphi}} & \text{(in)} \\ {}_{qubit}\langle\pm| & \bullet & |+\rangle_{qubit} \end{array} \tag{3.55}$$

Here, a controlled-$Ue^{i\varphi}$ is applied to the input state state in mode one where a qubit initialized to the $|+\rangle$ state is used as a control. This controlled unitary can be applied

by implementing the operator

$$\begin{aligned} C_{U_{12}} &= e^{\frac{i}{2}(\hat{\mathcal{P}}_1+\varphi)} e^{-\frac{i}{2}\left((\hat{\mathcal{P}}_1+\varphi I_1)\otimes\hat{\sigma}_{z_2}\right)} \\ &= e^{\frac{i}{2}(\hat{\mathcal{P}}_1+\varphi)} \left(\cos(\tfrac{\hat{\mathcal{P}}_1+\varphi}{2})\otimes I_2 - i\sin(\tfrac{\hat{\mathcal{P}}_1+\varphi}{2})\otimes\hat{\sigma}_{z_2}\right). \end{aligned} \tag{3.56}$$

The ancilla qubit is then measured in the $\pm$-basis. Consider the case when the applied unitary is a displacement. Without losing the generality of arbitrary displacements, we can take

$$\hat{U} = e^{i\alpha\hat{P}}, \tag{3.57}$$

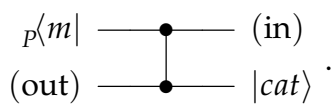
and note that rotating the input state before and after applying $\hat{U}$ will give the desired freedom for arbitrary displacements. In this case, the effective measurement-based operator applied to the input state to the circuit in Eq. 3.55 is given by

$$m_+ = e^{\frac{i}{2}(\alpha\hat{P}_1+\varphi)}\cos(\tfrac{\alpha\hat{P}+\varphi}{2}) \tag{3.58}$$

$$m_- = ie^{\frac{i}{2}(\alpha\hat{P}_1+\varphi)}\sin(\tfrac{\alpha\hat{P}+\varphi}{2}). \tag{3.59}$$

Thus up to a global phase and a displacement, a single round of phase estimation applies a sinusoid in $\hat{P}$ with an additional measurement-dependent phase shift of either 0 or $\frac{\pi}{2}$.

We can now compare the result of this circuit to teleporting an input through a squeezed cat state on a cluster state using the circuit




(3.60)

If the cat state is oriented such that the fringes are along the $P$-quadrature, that is if

$$|cat\rangle \propto \int dx e^{-x^2/2s^2}\cos(\alpha x)|x\rangle_q \tag{3.61}$$

where $s = e^{\iota}$ is the squeezing, then using the results of the general teleportation circuit given by Eq. 3.35, shows that this circuit applies the operator

$$m_m = \sqrt{2\pi}\hat{X}(m)\hat{R}(\tfrac{\pi}{2})e^{-\hat{P}^2/2s^2}\cos(\alpha\hat{P} - m). \tag{3.62}$$

After undoing the measurement-induced displacement and rotation with further cluster state processing, and for large enough squeezing, this operator reduces to

$$m_m \approx \sqrt{2\pi}\cos(\alpha\hat{P} - m), \tag{3.63}$$

which is identical to the qubit-based phase estimation up to an overall displacement, where the applied sinusoid operation now has a phase shift that is continuous based on the measurement result. By performing Fourier transforms before and after this circuit (through teleportation on the cluster state), the operator can instead be transformed to a sinusoid in $\hat{Q}$.

Inserting another cat state for the input state in Eq. 3.60 is exactly the method of GKP systhesis as described in the previous section and completes the connection to Ref. [83] where repeated phase-estimation is used to make grid states. More generally, we can borrow ideas from other bosonic phase-estimation schemes and enact them using cluster states and PhANTM generated cat states.

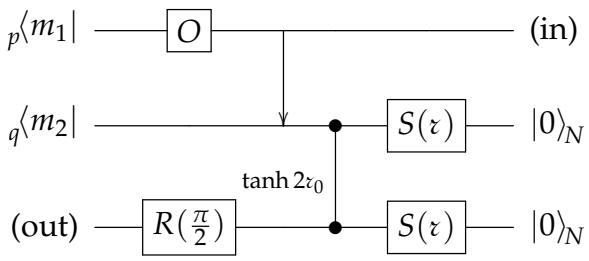
## 3.5 Macronode Extension

Many experimental implementations of cluster states have not generated the canonical cluster based on $\hat{C}_Z$ entangling gates, which would require inline squeezing, but instead rely on linear optics with all squeezing generated up-front [53, 54, 56, 57]. These types of cluster states are formed from *macronodes*, which we introduced in Sec. 1.9.1. Here, we develop a dictionary protocol that maps photon subtraction on a macronode cluster onto the canonical cluster that we have been discussing. In this way, the details of the operators change, but the same general result allowing the embedding of cat states into the cluster remains.

### 3.5.1 Dictionary protocol

Consider a macronode teleporation circuit where we have applied some operator $\hat{O}$ before the first homodyne measurement, and the measurement angles are chosen such that we perform respective $P$ and $Q-$basis measurements ($\theta_1 = 0, \theta_2 = \frac{\pi}{2}$). With the

inclusion of finite squeezing, this circuit will appear as

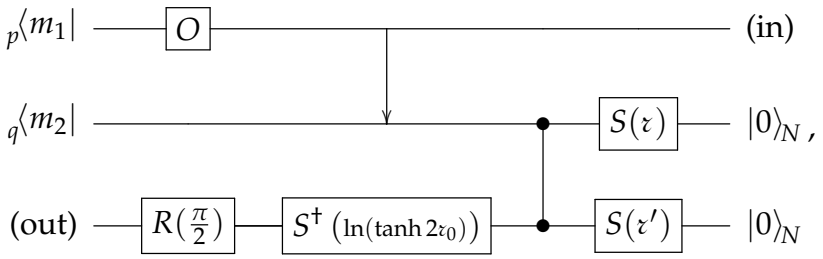



(3.64)

Note that for the formation of the finite-energy EPR pair in the bottom two modes we have used a weighted $\hat{C}_Z$ gate of $e^{i \tanh 2r_0 \hat{Q}_2 \hat{Q}_3}$, which is the same as the finite-squeezing version of the beamsplitter case in Eq. 1.233, up to a rotation on the output and the second input squeezed state, and a re-scaling of the initial squeezing. If the single-mode squeezed states before the entangling beamsplitter have squeezing $r_0$, then the effective squeezing before the weighted $\hat{C}_Z$ gate will be $r = \ln \sqrt{\text{sech} 2r_0}$. The weighting on the $\hat{C}_Z$ gate can be returned to one by noting that

$$e^{i \tanh 2r_0 \hat{Q}_2 \hat{Q}_3} = \hat{S}_3^\dagger(\ln[\tanh 2r_0]) e^{i \hat{Q}_2 \hat{Q}_3} \hat{S}_3(\ln[\tanh 2r_0]). \tag{3.65}$$

With this, the circuit becomes




(3.66)

where $r' = \ln\left(\tanh 2r_0 \sqrt{\text{sech} 2r_0}\right)$. We now derive the specific dictionary protocol that will determine the effect of applying $\hat{O}$ as if it were applied to a canonical cluster. We start by commuting the beamsplitter with the $\hat{C}_Z$ gate using the identity

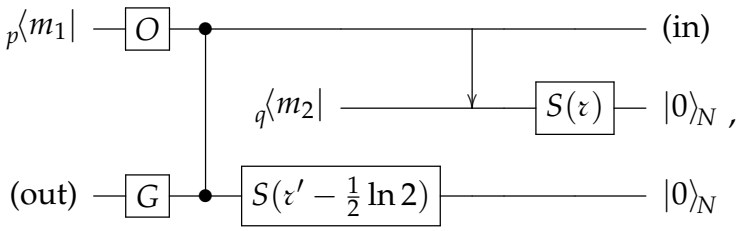
$$\hat{B}_{12}(\tfrac{\pi}{4}) \hat{C}_{Z_{23}} \hat{B}_{12}^\dagger(\tfrac{\pi}{4}) = e^{\frac{i}{\sqrt{2}}(\hat{Q}_2 - \hat{Q}_1)\hat{Q}_3}, \tag{3.67}$$

which is just two weighted $\hat{C}_Z$ gates. We can act this to the left on the measurement on mode two and fix the weight of the remaining $\hat{C}_Z$ by rotating the operator in the

Heisenberg picture with squeezing operators. This is written as

$$
\begin{aligned}
{}_{2Q}\langle m_2|e^{\frac{i}{\sqrt{2}}(\hat{Q}_2-\hat{Q}_1)\hat{Q}_3} &= {}_{2Q}\langle m_2|\hat{Z}_3(\tfrac{m_2}{\sqrt{2}})e^{-\frac{i}{\sqrt{2}}\hat{Q}_1\hat{Q}_3} \\
={}_{2Q}\langle m_2|\hat{Z}_3(\tfrac{m_2}{\sqrt{2}})\hat{R}_3(\pi)\hat{S}_3(\tfrac{1}{2}\ln 2)\hat{C}_{Z_{13}}\hat{S}_3^\dagger(\tfrac{1}{2}\ln 2)\hat{R}_3^\dagger(\pi). &
\end{aligned}
\tag{3.68}
$$

The rotations can both be pushed to the outer edges of the circuit, where a rotation of $\hat{R}_3^\dagger(\pi)$ to the right has no effect, and the rotation of $\hat{R}_3(\pi)$ on the left changes the sign of the $P$-quadrature displacement and combines with the end rotation. This results in the circuit




(3.69)

where the Gaussian operation $\hat{G}$ is

$$
\hat{G} = \hat{R}^\dagger(\tfrac{\pi}{2})\hat{Z}^\dagger(\tfrac{m_2}{\sqrt{2}}\tanh 2r_0)\hat{S}^\dagger\left(\ln(\tfrac{1}{\sqrt{2}}\tanh 2r_0)\right). \tag{3.70}
$$

From the above circuit, it is easy to see that if we can reduce the beamsplitter and measurement on the second wire to a single Kraus operator acting on the top wire, then we have a macronode teleportation circuit with some additional operator acting on the input and a Gaussian operator $\hat{G}$ on the output. To do this, we need to find the operator for

$$
{}_{2Q}\langle m_2|\hat{B}_{12}(\tfrac{\pi}{4})\hat{S}_2(r)|0\rangle_{2N}. \tag{3.71}
$$

Using Eq. 3.24, the fact that $\hat{B}_{12}(\theta) = \hat{B}_{12}^\dagger(-\theta) = \hat{B}_{21}(-\theta)$, and the beamsplitter decomposition given in Ref. [18] of

$$
\hat{B}_{j,k}(\tfrac{\pi}{4}) = e^{i\hat{P}_j\hat{Q}_k}\hat{S}_j(\tfrac{1}{2}\ln 2)\hat{S}_k^\dagger(\tfrac{1}{2}\ln 2)e^{-i\hat{Q}_j\hat{P}_k}, \tag{3.72}
$$

we can rewrite Eq. 3.71 as

$$
\begin{aligned}
&\pi^{1/4}\sqrt{\tfrac{2}{s}}\hat{X}_1^\dagger(m_2)\hat{S}_1(\tfrac{1}{2}\ln 2)\times \\
&\qquad {}_{2Q}\langle m_2|\hat{S}_2^\dagger(\tfrac{1}{2}\ln 2)e^{-i\hat{Q}_1\hat{P}_2}e^{-\hat{Q}_2^2/2s^2}|0\rangle_{2P},
\end{aligned}
\tag{3.73}
$$

where again $s = e^{r}$. The operators between the bra and ket in Eq. 3.73 can be commuted to have all operators in $\hat{Q}_2$ to the left and all operators in $\hat{P}_2$ to the right, which gives

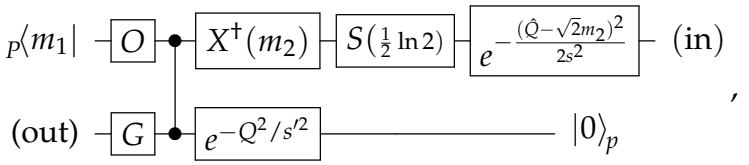
$$e^{-\hat{Q}_2^2/s^2} e^{\sqrt{2}\hat{Q}_1\hat{Q}_2/s^2} \hat{S}_2^\dagger(\tfrac{1}{2}\ln 2) e^{-\hat{Q}_1^2/s^2} e^{-i\hat{Q}_1\hat{P}_2}. \tag{3.74}$$

When acted on by the $Q$-basis bra and $P$-basis ket, along with noting that squeezing a quadrature eigenstate of zero has no effect, this reduces to an operator acting on wire one only; Eq. 3.71 becomes

$$\pi^{1/4}\sqrt{\tfrac{2}{s}}\hat{X}_1^\dagger(m_2)\hat{S}_1(\tfrac{1}{2}\ln 2) e^{-\frac{(\hat{Q}_1-\sqrt{2}m_2)^2}{2s^2}}. \tag{3.75}$$

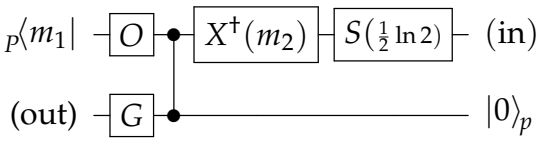
This expression is the same result as Eq.(23) in Ref. [60], where the wavefunction of the input state used here from mode two is that of a squeezed vacuum state.

Taking Eq. 3.75 and using Eq. 3.24 on the third wire in the circuit to commute the effect of finite squeezing with the $\hat{C}_Z$ gate, we have that our circuit becomes



, (3.76)

with $s' = e^{r'}$. For high initial squeezing, this circuit reduces to



. (3.77)

Thus, up to a Gaussian operation on the output, we can see that the result of applying any operator $\hat{O}$ before the first homodyne measurement in the macronode teleportation circuit is equivalent to applying $\hat{O}$ before the homodyne measurement in the canonical cluster-state teleportation, but first squeezing and displacing the input state. For finite squeezing, the input state also undergoes $Q$-quadrature damping.

Suppose $\hat{O}$ is the operator for photon subtraction of $n$ photons, $\hat{\mathcal{S}}_n$, given previously by Eq. 3.7. This is exactly the process described earlier in terms of the canonical cluster, and when followed by homodyne detection, can be written as an operator in $\hat{Q}$.

Using Eq. 3.23, photon subtraction and homodyne detection becomes

$$ {}_p\langle m|\hat{S}_n = {}_p\langle 0|f_n(\hat{Q}), \tag{3.78} $$

where $f_n(\hat{x})$ is the same function derived previously, and is given by Eq. B.25 in the Appendix. It is now clear that the macronode equivalent of the canonical case Kraus operator is just a slight modification of Eq. 3.27, and is given by

$$ \hat{K}_n^{mac} = \sqrt{\tfrac{2\pi^{1/2}}{s}}\hat{G}\hat{R}(\tfrac{\pi}{2})\hat{K}_n\hat{X}_1^\dagger(m_2)\hat{S}(\tfrac{1}{2}\ln 2)e^{-\frac{(\hat{Q}-\sqrt{2}m_2)^2}{2s^2}}. \tag{3.79} $$

For large squeezing and weak subtraction beamsplitter reflectivity, we can use Eq. 3.29 and this becomes

$$ \begin{aligned} &\hat{K}_n^{mac} \propto \\ &\hat{R}(\tfrac{\pi}{2})\hat{S}^\dagger\left(\ln(\tfrac{1}{\sqrt{2}}\tanh 2r_0)\right)\hat{Z}^\dagger(\tfrac{m_1}{\sqrt{2}})H_n(i\hat{Q}-\tfrac{m_1+im_2}{\sqrt{2}}). \end{aligned} \tag{3.80} $$

Just as before, subtracting photons before the homodyne measurement will apply a polynomial in $\hat{Q}$ to the teleported quantum information. Here, however, the measurement dependent shift has both real and imaginary components. Additionally, there is a residual squeezing term due to the use of beamsplitters as the entangling gates in the macronode implementation from Eq. 1.233. Failing to subtract any photons will only act to damp the state slightly and teleport it further along the cluster state, just as in the canonical case.

An important point to note is that the subtraction must only be attempted on one wire of the macronode for the above derivation to hold. We have derived the results for applying the subtraction to the top wire before detection, but the derivation remains the same if the subtraction was instead performed on the middle wire, up to a rotation on the input. Because of this, there is no need to waste an intermediate step in applying an 'empty' teleportation to enact a Fourier transform and reorient the state, as simply alternating which wire the subtraction is performed on will effectively apply the necessary rotation and allow for operators of the same quadrature to build up through several repetitions of the modified teleportation gadget.

### 3.5.2 Effects of $m_2$

We have shown how the PhANTM protocol can be translated from the canonical cluster state to a macronode implementation using the dictionary protocol, but we will leave an in-depth analysis on cat state breeding and generation to future work. However, here we motivate the main differences created by the measurement-induced $Q$- quadrature shift by $m_2$. Sec. 3.2 demonstrated that the overall effect of $m_1$ is to shift the phase between the components of the resulting cat state. Here, consider the case where $m_1 = 0$, so that the measurement results of $m_2$ can be dealt with separately. Instead of making a cat state that is a balanced superposition of coherent states as per Eq. 3.20, $H_n(i\hat{Q} - \frac{im_2}{\sqrt{2}})$ applied to squeezed vacuum will result in the weighted superposition

$$|cat\rangle_{A,B} \propto \left(A\hat{X}(\alpha) \pm B\hat{X}^{\dagger}(\alpha)\right) S(\mathfrak{r})|0\rangle_N, \tag{3.81}$$

where $A^2 + B^2 = 1$ are the coefficients weighting the different components in the superposition. When $m_2 = 0$, we have the case illustrated previously in Figs. 3.1 and 3.2, which corresponds to weighting coefficients of $A^2 = B^2 = \frac{1}{2}$. For nonzero $m_2$, the coefficients become unbalanced. This is shown in Fig. 3.10(a) for the case of applying $H_4(i\sqrt{2}\hat{Q} - im_2)$ to vacuum squeezed by 6 dB. The fourth-order Hermite polynomial was chosen since $\hat{Q}^4$ and higher polynomials applied to squeezed vacuum become nearly indistinguishable from cat states in the idealized case, as illustrated by Fig. 3.2. For each value of $m_2$, a weighted cat state of the form of Eq. 3.81 was selected and numerically optimized to fit the resultant state, and we plot the fitted value of $B^2$ against $m_2$. Each optimized weighted cat had fidelity greater than 0.99 with the resultant state, and it can be seen from the figure that the value of $B$ monotonically increases as $m_2$ increases.

Once the state becomes unbalanced, one can ask what happens when it is sent into future rounds of the PhANTM algorithm. To explore this, we consider input states to the circuit in Eq. 3.64 that are small weighted cat states of the form of Eq. 3.81, and calculate the probability $P(m_2)$ of performing a homodyne measurement in mode two and obtaining outcome $m_2$. The probability of the $Q$-quadrature homodyne measurement $m_2$ in the macronode circuit from Eq. 3.64 is given by

$$P(m_2) \propto \int_{-\infty}^{\infty} dy \left| {}_{q_1}\langle y|_{q_2}\langle m_2|\hat{B}_{12}e^{-\hat{Q}_2^2/2s_2^2}|\psi\rangle_1|0\rangle_{p_2}\right|^2, \tag{3.82}$$

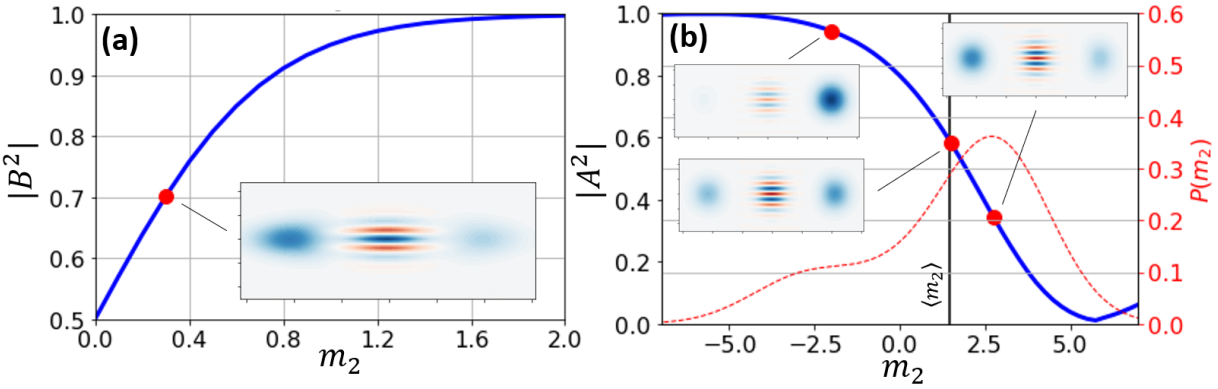



FIGURE 3.10: (a) The effects of the $Q$-basis homodyne measurement result in the macronode teleportation scheme with photon subtraction occurring before the $P$-basis measurement. As $m_2$ increases, the negative displacement in the resulting cat-state superposition becomes more heavily weighted. (b) Sending a weighted cat state into one round of PhANTM tends to reduce the more strongly weighted component. Here the initial state is set with $A^2 = \frac{4}{5}$, $\alpha = 4$, and $s_1 = 0$ while the cluster state is squeezed by 15 dB.

where here $\hat{B}_{12}$ is a balanced beamsplitter, $s_2 = e^{r_2}$ is the squeezing in the cluster state, and $|\psi\rangle_1$ is the input quantum state. This can be seen more easily by looking at the reduced circuit in Eq. 3.69, where the $m_2$ homodyne measurement can be seen as occurring immediately after the first beamsplitter. Suppose the input state has the form of a weighted cat state

$$|\psi\rangle_1 \propto \left(A\hat{X}(\alpha) + B\hat{X}^\dagger(\alpha)\right)\hat{S}(r_1)\,|0\rangle\,, \tag{3.83}$$

where we take $A^2 + B^2 = 1$ and let $\alpha \geq 0$ without loss of generality. Using this form of $\psi$, we have that

$$\hat{B}_{12}e^{-\hat{Q}_2^2/2s_2^2}|\psi\rangle_1|0\rangle_{p_2} = e^{-\frac{(\hat{Q}_2-\hat{Q}_1)^2}{4s_2^2}}\psi_Q\left(\frac{\hat{Q}_1+\hat{Q}_2}{\sqrt{2}}\right)|0\rangle_{p_1}|0\rangle_{p_2} \tag{3.84}$$

and so the integrand is, in the $Q$-basis,

$$\iint dx dx' e^{-\frac{(x-x')^2}{4s_2^2}}\psi_Q\left(\frac{x+x'}{\sqrt{2}}\right)|x\rangle_{q_1}\langle m_2|x'\rangle_{q_2}. \tag{3.85}$$

Thus, up to a normalization,

$$P(m_2) = \int_{-\infty}^{\infty} dx \left| e^{-\frac{(x-m_2)^2}{4s_2^2}}\left(Ae^{-\frac{x+m_2-\alpha\sqrt{2}}{4s_1^2}} + Be^{-\frac{x+m_2+\alpha\sqrt{2}}{4s_1^2}}\right)\right|^2. \tag{3.86}$$

Performing the integral leads to

$$P(m_2) = C\Big(A^2 e^{-\frac{(\alpha-\sqrt{2}m_2)^2}{s_1^2+s_2^2}} + B^2 e^{-\frac{(\alpha+\sqrt{2}m_2)^2}{s_1^2+s_2^2}} + 2AB e^{-\frac{\alpha^2}{s_1^2}-\frac{2m_2^2}{s_1^2+s_2^2}}\Big), \tag{3.87}$$

where $s_1 = e^{r_1}$ is the squeezing in the input weighted cat state, $s_2 = e^{r_2}$ is the squeezing in each physical mode of the macronode cluster state, and the coefficient is

$$C = \left(\sqrt{\tfrac{\pi}{2}}\sqrt{s_1^2+s_2^2}\left(1 + AB e^{-\frac{\alpha^2}{s_1^2}}\right)\right)^{-1}. \tag{3.88}$$

This distribution is just two Gaussians originating from the input weighted cat state with broadening dependent on both the squeezing of the input state and the squeezing in the macronode cluster. The expectation value for a $Q$-basis homodyne measurement of mode two is given by

$$\begin{aligned}\langle \hat{Q}_2 \rangle &= \int_{-\infty}^{\infty} dm_2 P(m_2) m_2 \\ &= \frac{\alpha(2A^2-1)}{\sqrt{2}(1+AB e^{-\alpha^2/s_1^2})},\end{aligned} \tag{3.89}$$

which for $\alpha > 0$, is always positive when $A^2 > \frac{1}{2}$. This indicates that for an input cat state weighted more strongly toward the positive displacement ($\alpha > 0$, $A > 0$), the measurement result $m_2$ is likely to be positive, and thus the current round of the PhANTM algorithm will tend to increase $B$ and re-balance the weighting in the coherent state superposition.

The re-balancing can be thought of as a 'restoring force' tending to prevent either coherent state term from dominating the other, and is shown for a particular example in Fig. 3.10(b), where an unbalanced cat state with amplitude $\alpha = 4$ and $A^2 = \frac{4}{5}$ is sent through a macronode teleportation circuit with a single photon subtraction occurring. The measurement result $m_2$ is varied while $m_1$ is fixed to be zero to highlight the effects of changing $m_2$, and the output state is then numerically optimized to a fidelity above 0.99 with a weighted cat of the form of Eq. 3.81. The blue curve plots the final value of $A^2$ against the measurement result where the probability distribution is shown as a red-dashed curve with the expectation value given by the solid vertical line. The

figure insets display the Wigner functions for the output state and demonstrate that as $m_2$ increases, the weighting shifts from favoring the positive displacement ($A > B$) to favoring the negative displacement ($A < B$). As shown by the figure, the value of $A$ is likely to decrease. If the weighting remains unbalanced after the teleportation, then the process can be repeated until the output state has displacement components with weightings that are within a predetermined acceptable range.

A salient point of this process is that runaway affects will not occur; as shown by Eq. 3.89, the measurement result $m_2$ will be such that unbalancing is, on average, not exacerbated in the same direction. Overshooting is possible, but a strongly weighted coefficient will not progressively become larger until only one displacement remains in the superposition.

## 3.6 Summary

Starting with a Gaussian cluster state, we have presented PhANTM, a method for using PNR detection and feed-forward Gaussian operations to locally de-Gaussify the cluster state to systematically generate Schrödinger cat states. Although each individual photon-subtraction event is probabilistic, teleporting quantum information along the cluster state and repeatedly applying steps of PhANTM leads to the production of cat states with high probability and mean amplitude dependent on the squeezing present in the cluster state. This process can be thought of as an adiabatic 'cooling' toward a cat state basis, as additionally seen by the ability for PhANTM to preserve embedded cat states as they are teleported throughout the cluster, preventing the build-up of excess Gaussian noise. The phase of the teleported cat will be randomized by homodyne measurement, but this randomization can be tracked by recording the measurement results at each step and fixed by performing a feed-forward displacement to align the cat fringe. This therefore allows a cluster state with PNR measurement capabilities to be used as a way to perpetuate a particular class of non-Gaussianity (cat states), without the need for GKP error-correction.

Portions of 1-D quantum wires can each be converted into a cat state by repeatedly applying PhANTM to embed multiple cat states within a large cluster state, making this process compatible with state-of-the-art massively scalable 2D cluster states [56, 57]. Additionally, the photon-subtraction style measurement only requires low photon-number resolution, which is well within the current experimental capabilities of both

TES systems [124, 165] and number-resolving silicon nanowire detectors [166, 167]. Note that semiconductor detector technology, e.g. avalanche photodiodes, can be used in multiplexed detection schemes to achieve PNR [168–170], though this requires large numbers of multiplexed detectors since the individual detectors aren't PNR. While error-correction and universal QC with cat states alone is possible [131, 171], we show how breeding protocols are compatible with cat states embedded within the cluster state solely by performing Gaussian measurements and feed-forward displacements locally on the cluster. This eliminates the need for offline resource-state generation as in current CV one-way error-corrected QC proposals [65, 104], and allows for cat state enlargement and GKP state synthesis. Taken together with the PhANTM cat-state generation protocol, we have provided a means to take Gaussian cluster states and transform them into universal quantum-computational resources by performing local homodyne detections, PNR measurements, and feed-forward displacements.

# Chapter 4

# Rotation-Symmetric Code States

As mentioned in the first chapter, the GKP encoding is only one way in which one can build qubits into a CV Hilbert space. Another is the family of rotation-symmetric encodings. In this chapter, I will discuss work that aims to use squeezed resources and PNR measurements to engineering specific rotation-symmetric states [172]. It should be noted that like photon catalysis from Ch. 2, the scheme proposed here can be generalized as a type of Gaussian boson sampling, where performing PNRD on several modes of a multimode Gaussian state allows for the projection of the remaining modes into non-Gaussian resources [85]. Machine-learning methods have been proposed to optimize fidelity and success probability with desired target states [85, 173, 174], but the nature of Gaussian to non-Gaussian conversion may not be unique, therefore a global minimum to such optimization problems may not exist and it may be difficult to optimize for many realistic scenarios including resource efficiency. Thus, devising specific protocols may still prove more useful that an arbitrary brute-force boson sampling device.

## 4.1 Overview

The rotation-symmetric codes are a promising avenue to exploit non-Gaussianity and include include quantum states that are rotationally symmetric in phase space in analog to translation phase space symmetry in GKP states. Error correcting codes designed with such states take advantage of the fact that states with $K$-fold rotational symmetry have decompositions over the Fock-basis with $K$-periodic spacing, making it easier to ascertain when photon loss and dephasing errors occur. Such codes include binomial codes and cat codes, and Pegg-Barnett codes [21, 68, 175]. While binomial and cat codes

have been recently demonstrated in superconducting quantum circuits [176, 177], their implementations in the optical domain have remained elusive.

In this work, we introduce an all-optical method to generate rotationally symmetric states with 2-fold and 4-fold-symmetry, in particular, binomial code states and truncated cat code states. Our method requires resources as minimal as a two-mode squeezed vacuum (TMSV) state, easily accessible linear optics, and PNRD at low photon numbers. Our proposal requires squeezing levels below what has already been demonstrated [14], and low PNR detection which is now feasible due to advances in highly efficient transition-edge sensors [124] and number-resolving superconducting nanowire single-photon detectors detectors [143, 166, 167], and time-multiplexed detection schemes [169].

We show that by tuning the initial available resource squeezing and post-selecting on desired the PNR outcomes, our method can generate a variety of states in addition to exact binomial states and approximated cat states with the generation rates in the KHz-MHz depending on the detection scheme using state-of-the-art PNR detectors with high count rates. Moreover, these rotationally symmetric states can be exactly produced as opposed to GKP states, which are only approximated due to available finite squeezing limit.

Our paper is structured as follows. In Section 4.2, we provide an overview of bosonic quantum error correction with binomial codes. Section 4.3 discusses the general framework of our method to generate a various rotation-symmetric non-Gaussian states. In Section 4.5, we focus on generating binomial and cat-like codes in the ideal case as well as in the presence of experimental imperfections. In Section 4.6, we derive the conditions for desired detection efficiency for faithful error correction. Section 4.7 concludes and offers a brief outlook.

## 4.2 Rotation-symmetric Bosonic Codes: Recap

A detailed discussion on rotationally symmetric quantum error-correcting codes can be found in Refs. [21, 68, 178, 179]; here we briefly summerize the main concepts. The rotation-symmetric bosonic codes include the bosonic encodings that remain invariant under discrete rotations in phase space in the same manner as GKP encodings are invariant under phase space translations. Some representative examples of rotation-symmetric bosonic codes are binomial codes and cat codes, which have been recently

demonstrated in circuit quantum electrodynamics (cQED) architectures [21, 106, 177]. The logical code words are stabilized by the photon-number super parity operator defined as

$$\hat{\Pi}_K := e^{\frac{i2\pi\hat{n}}{K}}, \tag{4.1}$$

where $\hat{n} = \hat{a}^\dagger\hat{a}$ is the photon-number operator. Therefore, in order to satisfy $\hat{\Pi}_K|\psi\rangle = |\psi\rangle$, the state $|\psi\rangle$ must have support on every K-th Fock basis.

For a given error set $\mathcal{E} = \{\hat{E}_1, \hat{E}_2, ..., \hat{E}_l\}$, a necessary and sufficient criterion for a faithful error correction is known as the Knill-Laflamme condition, mathematically defined as:

$$\hat{P}_\mathcal{C}\hat{E}_l^\dagger\hat{E}_m\hat{P}_\mathcal{C} = \alpha_{l,m}\hat{P}_\mathcal{C}; \quad \forall\ \hat{E}_l, \hat{E}_m \in \mathcal{E}, \tag{4.2}$$

where $\hat{P}_\mathcal{C}$ is the projector defined over the code space $\mathcal{C}$ and $\alpha_{l,m}$ are entries of a Hermitian matrix [109]. For example, let's consider the (4-fold symmetric) binomial code words

$$|0_L\rangle = \frac{1}{2}|0\rangle + \frac{\sqrt{3}}{2}|4\rangle, \quad |1_L\rangle = \frac{\sqrt{3}}{2}|2\rangle + \frac{1}{2}|6\rangle, \tag{4.3}$$

whose amplitudes are the square roots of binomial coefficients. One can easily see that these code words are able to detect and faithfully correct the errors given by error set $\mathcal{E} = \{\mathbb{I}, \hat{a}, \hat{a}^\dagger\hat{a}\}$ as per Knill-Laflamme condition.

Physically, the Knill-Laflamme conditions ensure that the probability for a single-photon loss or a dephasing error to occur are the same for each code word, making it impossible for the environment to distinguish between the logical basis states $|0_L\rangle$ and $|1_L\rangle$. This preserves the encoded quantum information (the state's amplitudes in the logical basis), as it is not deformed as a result of the error and can thus be recovered by means of unitary operations. To understand this further, let us consider a single-photon loss error for quantum state $|\psi\rangle$ encoded using the code words defined in 4.3. The logical state $|\psi\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$ is an even photon-number parity state which transforms to odd parity state $|\psi'\rangle = \sqrt{3}\alpha|3\rangle + \sqrt{3/2}\beta(|1\rangle + |5\rangle)$ under a single-photon loss error. Thus we get:

$$\hat{a}|\psi\rangle \rightarrow |\psi'\rangle = \alpha|\tilde{0}_L\rangle + \beta|\tilde{1}_L\rangle, \tag{4.4}$$

where $|\tilde{0}_L\rangle = |3\rangle$ and $|\tilde{1}_L\rangle = 1/\sqrt{2}(|1\rangle + |5\rangle)$ are the error words, which are orthogonal to the original code words. We note that the quantum information encoded in the complex amplitudes $\alpha$ and $\beta$ is preserved and can be faithfully recovered by mapping the error words $\{|\tilde{0}_L\rangle, |\tilde{1}_L\rangle\}$ to code words $\{|0_L\rangle, |1_L\rangle\}$ by means of unitary operations [21]. Since the single-photon loss changes the photon-number parity of the

original state from even to odd, the parity measurements can be used as an error syndrome measurement. In optics, error syndrome measurements can be performed using highly efficient photon-number-resolving transition-edge sensors [114, 124, 165].

In general, the photon loss process can be viewed more precisely by subjecting a quantum state to a completely-positive and trace-preserving (CPTP) bosonic channel with non-unity transmission. In this case, the effect of the channel is given by the Kraus operator-sum representation formulated as

$$\rho' = \mathcal{L}(\rho_0) = \sum_{k=0}^{\infty} \hat{E}_k \rho_0 \hat{E}_k{}^{\dagger}, \tag{4.5}$$

where $\rho'$ and $\rho_0$ are the output and input states, respectively, and to ensure the CPTP channel, one needs $\sum_{k=0}^{\infty} E_k^{\dagger} E_k = \mathbb{I}$. In this work, we treat loss as an amplitude-damping Gaussian channel, which can be modeled as an optical mode traveling a distance $L$ through a medium with loss coefficient $\kappa$ dB/cm. The Kraus operator associated with losing $k$ photons is

$$\hat{E}_k = \sqrt{\frac{(1-e^{-\gamma})^k}{k!}} e^{-\frac{1}{2}\gamma \hat{a}^{\dagger}\hat{a}} \hat{a}^k, \tag{4.6}$$

where we define $\gamma = \kappa L$. A $k$-photon loss to $\rho$ occurs with the probability of

$$P_k = \langle \hat{E}_k{}^{\dagger} \hat{E}_k \rangle = \frac{\gamma^k}{k!} \mathrm{Tr}[\hat{a}^k \rho \, \hat{a}^{\dagger k}] + O[\gamma^{k+1}]. \tag{4.7}$$

With this in mind, QEC codes with the ability to correct up to $k$-photon losses can be reframed as codes that correct operators $E_k$ up to $k$-th order in $\gamma$. As a result, the code words we propose to create in Eq. 4.3 have 4-fold-symmetry in phase space with the ability to correct single photon losses, making it effective up to first order in $\gamma$ and corrects for $E_0$ and $E_1$ corresponding to 'no-jump' and 'single-jump' errors, respectively.

Along with the error-correcting capabilities, a universal gate set for quantum computation with rotation-symmetric codes has recently been described that is agnostic with respect to the type of code used, and depends only on the degree of rotational symmetry [68]. The set of Clifford gates can be implemented with a controlled-rotation gate, which can be tuned to enact a logical control-Z gate, and a single-mode phase gate. These respective gates can be enacted optically through the use of cross-Kerr and self-Kerr interactions. The universal gate set is then completed through gate teleportation and logical code-basis measurements. Fault-tolerance with these gates is possible

by concatenating number-phase codes with the Bacon-Shor subsystem codes [68, 180–182]. As the gate set and error-correct schemes have been sufficiently described previously, we focus here on methods for state generation.

## 4.3 Proposed Method: Analytical Model

In this section, we detail the proposed method displayed in Fig. 4.1. Our method starts by preparing by a two-mode squeezed vacuum (TMSV) state by interfering two orthogonal single-mode squeezed vacuum (SMSV) states produced by optical parametric amplifiers (OPAs) at the first balanced beamsplitter labeled as BS1 in Fig. 4.1. This is followed by two highly unbalanced beamsplitters (BS2) used for photon subtractions from each mode of the TMSV state. Next, a balanced beamsplitter (BS1) is placed to interfere the subtracted photons in order to erase the information about from which mode the subtracted photons came from. As a result, this combination of photon subtractions and the balanced interference allows one to coherently subtract photons from the TMSV state. Finally, photon-number-resolving (PNR) measurements are performed on three output modes prepares the desired state $|\psi\rangle$ in the fourth mode for a certain combination of PNR measurement outcomes.

Interfering two SMSV states at a balanced beamsplitter is equivalent to preparing a TMSV state by action of the two-mode squeezed operator $S(z)_{ab} = e^{[z\hat{a}^\dagger\hat{b}^\dagger - z^*\hat{a}\hat{b}]}$ on two-mode vacuum, where $z = Re^{i\phi}$ with real squeezing parameter $R$ and phase $\phi$. In the photon-number basis, the TMSV state prepared in modes 'a' and 'b' after the first BS1 is given as [136]

$$|0,R\rangle_{ab} = \frac{1}{\cosh R}\sum_{n=0}^{\infty} e^{in\phi}\tanh^n R|nn\rangle_{ab}. \tag{4.8}$$

After encountering the highly unbalanced beamsplitters (BS2), the state interferes with vacuum modes 'c' and 'd', which leads to the four-mode state

$$|\psi\rangle_{\text{abcd}} = \hat{U}_{ac}(\theta)\hat{U}_{bd}(\theta)|0,R\rangle_{ab}|00\rangle_{cd}, \tag{4.9}$$

where $\hat{U}_{ac}(\theta), \hat{U}_{bd}(\theta)$ are beamsplitter unitary operators where $(t,r) = (\cos\theta, \sin\theta)$ are the transmission and reflections coefficients, respectively. We first consider the simplest case where the probability of subtracting more than one photon is negligible, i.e., at

most one photon is subtracted from each mode. For single photon subtraction, we set both beamsplitters such that $t \to 1$ ($\theta << 1$) which allows one to approximate the unitary operators as

$$\begin{aligned}\hat{U}_{ac}(\theta) &= \exp[\theta(\hat{a}^\dagger\hat{c} - \hat{a}\hat{c}^\dagger)] \approx \mathbb{I} + \theta[\hat{a}^\dagger\hat{c} - \hat{a}\hat{c}^\dagger] + O(\theta^2),\\ \hat{U}_{bd}(\theta) &= \exp[\theta(\hat{b}^\dagger\hat{d} - \hat{b}\hat{d}^\dagger)] \approx \mathbb{I} + \theta[\hat{b}^\dagger\hat{d} - \hat{b}\hat{d}^\dagger] + O(\theta^2),\end{aligned} \tag{4.10}$$

where we consider $\mathcal{O}(\theta^2)$ terms negligible to ensure that mostly single photons are subtracted from each mode. From Eq. 4.9 and Eq. 4.10, we get the unnormalized four-mode state

$$\hat{U}_{ac}(\theta)\hat{U}_{bd}(\theta)|0,R\rangle_{ab}|00\rangle_{cd} \propto \sum_n^\infty e^{in\phi}\tanh^n R[\mathbb{I} + \theta(\hat{a}^\dagger\hat{c} - \hat{a}\hat{c}^\dagger + \hat{b}^\dagger\hat{d} - \hat{b}\hat{d}^\dagger)]|nn\rangle_{ab}|00\rangle_{cd}, \tag{4.11}$$

Finally, the action of last balanced beamsplitter, $\hat{U}_{bd}(\frac{\pi}{4})$, which is the second BS1 in the figure, acts to transform the operators $\hat{c}^\dagger \to \frac{1}{\sqrt{2}}(\hat{c}^\dagger + \hat{d}^\dagger)$ and $\hat{d}^\dagger \to \frac{1}{\sqrt{2}}(\hat{d}^\dagger - \hat{c}^\dagger)$. This, along with the observation that any terms with $\hat{c}$ or $\hat{d}$ vanish when acting on vacuum, leads to the resulting state

$$|\psi\rangle_{\text{abcd}} \propto \sum_n^\infty e^{in\phi}\tanh^n R\left[|00\rangle_{cd} - \frac{\theta}{\sqrt{2}}\left[(\hat{a}-\hat{b})|10\rangle_{cd} + (\hat{a}+\hat{b})|01\rangle_{cd}\right]\right]|nn\rangle_{ab}. \tag{4.12}$$

We now perform the PNR measurements on three output modes as depicted in Fig. 4.1. Looking at the state in Eq. 4.12, it can be immediately seen that for $n_1 = n_2 = 0$, the initial TMSV state is unchanged as one expects. We now consider two cases of $n_1 = 1, n_2 = 0$ and $n_1 = 0, n_2 = 1$, i.e., only one of the two detectors detects a single photon. The two orthogonal output states in these two cases are

$$|\psi\rangle_{ab}^{\pm} \propto \sum_n^\infty \sqrt{n}e^{in\phi}\tanh^n R[|n-1,n\rangle \pm |n,n-1\rangle], \tag{4.13}$$

where the positive (negative) sign corresponds to $[n_1, n_2] = [1, 0]$ ($[n_1, n_2] = [0, 1]$).

It is worth mentioning that states like $|\psi\rangle^{\pm}$ have been shown to achieve Heisenberg-limit of phase measurements in quantum interferometry [183, 184]. We now consider a third PNR detector placed in the path of mode 'b' to detect $n_3$ photons. The third PNR measurement projects the $|\psi\rangle^{\pm}$ to the pure state given by

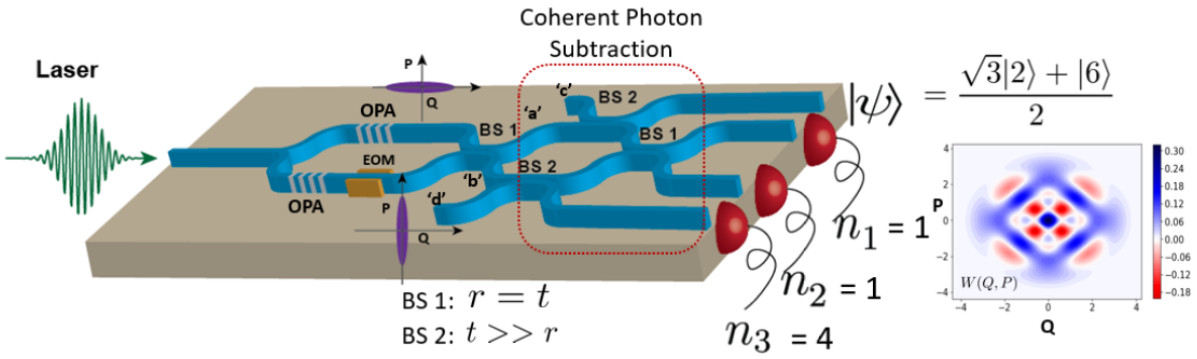



FIGURE 4.1: The proposed scheme for generating rotation-symmetric error-correcting codes. A two-mode squeezed vacuum (TMSV) state is generated by combining two orthogonal single-mode squeezed vacuum (SMSV) states at the first balanced beamsplitter (BS1). Photons are then subtracted from each modes of the TMSV state using highly unbalanced beamsplitters (BS2). The subtracted photons are interfered at the second balanced beamsplitter followed by PNR measurements on three output modes, which prepares the fourth mode in the desired code for a certain measurement combination of $n_1, n_2$, and $n_3$. The dotted red box represents the coherent photon subtraction process. A particular case shown here considers $[n_1, n_2, n_3] = [1, 1, 4]$, resulting in $|\psi\rangle \propto |2\rangle + |6\rangle$ with 4-fold rotational symmetry as evident in the displayed Wigner function. OPA: Optical parametric amplifier. BS1: Balanced beamsplitter. BS2: Unbalanced beamsplitter for photon subtraction. EOM: Electro-optic modulator.

$$|\psi\rangle_a^{\pm} = \mathcal{N}[\sqrt{n_3 c_{n_3,n_3}}|n_3 - 1\rangle \pm \sqrt{(n_3+1)c_{n_3+1,n_3+1}}|n_3+1\rangle], \tag{4.14}$$

where

$$\mathcal{N} = \sqrt{n_3 c_{n_3,n_3} + (n_3+1)c_{n_3+1,n_3+1}} \tag{4.15}$$

$$c_{n,n'} = e^{i\phi(n+n')} \tanh^{n+n'} R / \cosh^2 R \tag{4.16}$$

From Eq. 4.14, one can see that for $n_3$ = odd (even), the final state has two consecutive even (odd) Fock components corresponding to $n_3 - 1$ and $n_3 + 1$. We note that this particular task of coherent single-photon subtraction with highly unbalanced beamsplitters can be performed with click detectors by ensuring that the probability of

reflecting more than one photon is vanishing. However, the third detector needs to be PNR detector. In Section 4.5, we show how one can generate rotation-symmetric states with 2-fold and 4-fold phase space symmetry in the Wigner function.

We now generalize the proposed method for arbitrary detection of $n_1$ and $n_2$ photons. In this case, to the leading order approximation, only the terms involving up to $\theta^{n_1+n_2}$ will contribute in the expansion of unitary operators of the photon subtracting beamsplitters (BS2 in Fig. 4.1) in Eq. 4.11. As a result, the Taylor expansion can be approximated to

$$\hat{U}_{ac}(\theta)\hat{U}_{bd}(\theta) \approx \sum_{k=0}^{N} \frac{\theta^N}{k!(N-k)!}(\hat{a}^\dagger\hat{c} - \hat{a}\hat{c}^\dagger)^{N-k}(\hat{b}^\dagger\hat{d} - \hat{b}\hat{d}^\dagger)^k, \tag{4.17}$$

where $N = n_1 + n_2$ is the total number of photons detected after the photon subtraction step. Since we have vacuum inputs to modes $c$ and $d$, and as a consequence of detecting $N$ photons in total, Eq. 4.17 reduces to

$$\hat{U}_{ac}(\theta)\hat{U}_{bd}(\theta) \approx \sum_{k=0}^{N} \frac{(-\theta)^N}{k!(N-k)!}(\hat{a}\hat{c}^\dagger)^{N-k}(\hat{b}\hat{d}^\dagger)^k. \tag{4.18}$$

This approximation can now be transformed by the second balanced beamsplitter, $\hat{U}_{cd}(\theta)$, into

$$\hat{U}_{cd}\hat{U}_{ac}(\theta)\hat{U}_{bd}(\theta)\hat{U}_{cd}^\dagger \approx \left(\frac{-\theta}{\sqrt{2}}\right)^N \sum_{k=0}^{N} \frac{\hat{a}^{N-k}\hat{b}^k}{k!(N-k)!}(\hat{c}^\dagger + \hat{d}^\dagger)^{N-k}(-\hat{c}^\dagger + \hat{d}^\dagger)^k. \tag{4.19}$$

The action of these beamsplitters to the TMS resource state and vacuum modes $c$ and $d$ followed by detecting $n_2$ and $n_1$ photons in the transformed modes $c$ and $d$, respectively, leads to the final state

$$|\psi\rangle_{ab} = \hat{U}_{cd}\hat{U}_{ac}(\theta)\hat{U}_{bd}(\theta)\hat{U}_{cd}^\dagger|0,R\rangle_{a,b} \otimes |0,0\rangle_{c,d} \propto \sum_{k=0}^{N} A_{n_1,n_2,k}\hat{a}^{N-k}\hat{b}^k|0,R\rangle_{ab}, \tag{4.20}$$

where the coefficient $A_{n_1,n_2,k}$ is

$$A_{n_1,n_2,k} = \sum_{i_{min}}^{i_{max}} \frac{(-1)^{i+k-n_>}\sqrt{n_1!n_2!}}{i!(i+k-n_>)!(n_>-i)!(N-i-k)!}, \tag{4.21}$$

and the summation goes from $i_{min} = \text{Max}(0, n_> - k)$ to $i_{max} = \text{Min}(n_>, N-k)$, where

$n_> = \mathrm{Max}(n_1, n_2)$. The overall success probability of this combination of PNR measurements is

$$P([n_1, n_2]) = \left(\frac{\theta}{\sqrt{2}}\right)^{2N} \mathrm{Tr}\Big[\sum_{k=0}^{N}\sum_{j=0}^{N} A_{n_1,n_2,k} A_{n_1,n_2,j} \hat{a}^{N-k}\hat{b}^{k}|0,R\rangle_{ab}\langle 0,R|\hat{a}^{\dagger N-j}\hat{b}^{\dagger j}\Big]. \quad (4.22)$$

One salient point to note about the coherent photon subtraction is that the approximation of weakly reflective subtraction beamsplitters used in Eq. 4.18 is not strictly necessary. In fact, the full treatment of perfect coherent photon subtraction from a TMSV state with arbitrary beamsplitter parameter $\theta$ reduces to the same two-mode state in Eq. 4.20, but with a different effective squeezing parameter. If the initial experimental squeezing parameter is $R$, then the nonzero beamsplitter reflectivity acts to replace $R$ in Eq. 4.20 with an effective squeezing parameter of $R' = \tanh^{-1}(t^2 \tanh R)$, where $t^2 = \cos^2\theta$ is the transmission of each subtraction beamsplitter.

### 4.3.1 Perfect coherent photon subtraction

In a laboratory setting where beamsplitters with non-vanishing reflectivity are used, we can recover the ideal subtraction case with a modification to the value of initial squeezing. In the case where subtraction beamsplitters are given by the operators $U_{ac} = U_{bd} = e^{-\theta(a^\dagger b - a b^\dagger)}$ and the third beamsplitter to make the subtraction coherent is balanced,$U_{cd} = e^{-\frac{\pi}{4}(a^\dagger b - a b^\dagger)}$, applying these operators to a TMS vacuum yields

$$U_{cd}U_{ac}U_{bd}|0,R\rangle_{ab} = \sum_{n=0}^{\infty}\frac{(e^{i\phi}\tanh R)^n}{n!}\left(ta^\dagger + \frac{r}{\sqrt{2}}(c^\dagger + d^\dagger)\right)^n\left(tb^\dagger + \frac{r}{\sqrt{2}}(-c^\dagger + d^\dagger)\right)^n|0\rangle_{abcd}, \quad (4.23)$$

where $r = \sin\theta$ and $t = \cos\theta$ are the real reflectivity and transmissivity coefficients of the subtraction beamsplitters. Detecting $n_1$ photons in mode $c$ and $n_2$ photons in mode $d$ leads to the unnormalized output state of

$$\begin{aligned}|\phi\rangle &= \sum_{n=0}^{\infty}\sum_{k=0}^{n_1+n_2} B\frac{(t^2\tanh R)^n}{k!(n_1+n_2-k)!}a^{n_1+n_2-k}b^k|nn\rangle_{ab} \\ &= \sum_{k=0}^{N} B a^{N-k}b^k|0,R'\rangle_{ab}, \end{aligned} \quad (4.24)$$

where $N \equiv n_1 + n_2$ and $B = A_{n_1,n_2,k}$ from Eq. 4.21. The above equation is then exactly the same as the perfect coherent subtraction case of Eq. 4.20 with an effective decreased squeezing parameter of

$$R' = \tanh^{-1}(t^2 \tanh R). \tag{4.25}$$

This shows that the idealized states discuss ed are in fact equivalent to realistic states with modified squeezing. Furthermore, this fact can be used to increase the overall success probability of a protocol if one has the ability to generate an initial TMS vacuum state with squeezing in excess of what is required to produce a specific target. By beginning with a large squeezing parameter, $R$, one can increase the reflectivity of the subtraction beamsplitters to increase the likelihood of successfully subtracting $n_1$ and $n_2$ photons until the effective squeezing has been reduced to the $R'$ value required to produce the target.

Later we show how one can utilize this feature of coherent photon subtraction to increase the success rates of preparing the desired states.

We now investigate the third PNR measurement on the coherently photon-subtracted state represented by Eq. 4.20. A measurement outcome of $n_3$ photons projects the two-mode state $|\psi\rangle_{a,b}$ to the following parity state

$$|\psi\rangle_a = \mathcal{N} \sum_{k=k_{min}}^{N+n_3} A'_{n_1,n_2,k} |2k - N - n_3\rangle_a, \tag{4.26}$$

where we have

$$A'_{n_1,n_2,k} = A_{n_1,n_2,(k-n_3)} \frac{(e^{i\phi} \tanh R)^k k!}{\sqrt{(2k - N - n_3)!}}, \tag{4.27}$$

and $k_{min} = \text{Max}(n_3, \lceil \frac{1}{2}N + \frac{1}{2}n_3 \rceil)$ with $\lceil . \rceil$ being the ceiling function. The normalization coefficient is given by

$$\mathcal{N} = \frac{\left(\frac{1}{2}\theta^2\right)^N}{n_3!(\cosh R)^2 \sqrt{P_{\text{succ}}}}, \tag{4.28}$$

where $P_{\text{succ}}$ is the overall success probability of the measurement outcome configuration $[n_1, n_2, n_3]$ given as

$$P_{\text{succ}} = \frac{\left(\frac{1}{2}\theta^2\right)^N}{n_3!(\cosh R)^2} \sum_{k=k_{min}}^{N+n_3} (A'_{n_1,n_2,k})^2. \tag{4.29}$$

This completes the generalized derivation for coherent photon subtraction scheme for generating rotation-symmetric error correcting codes.

Examining Eq. 4.26, we see that in addition to the state having definite photon-number parity, the number of Fock components in the superposition is determined by the number of subtracted photons to reach a maximum of $N+1$ when $n_1 \neq n_2$ and $n_3 > N$. When $n_3 < N$, the maximum number of components drops to $\lfloor \frac{1}{2}N + \frac{1}{2}n_3 \rfloor + 1$. A particularly interesting case occurs when $n_1 = n_2$. In this case the coefficient $A_{n_1,n_2,k}$ vanishes for $k =$ odd as seen below:

$$A_{n_1,n_2,k_{odd}} = \sum_{i=0}^{n_2-\frac{k-1}{2}} (-1)^{n_2+k-i}\big[(2n_2-k-i)!(k+i-n_2)!i!(n_2-i)!)\big]^{-1} + (-1)^{-n_2-i}\big[i!(n_2-i)!(2n_2-k-i)!(k+i-n_2)!)\big]^{-1} = 0. \quad (4.30)$$

Therefore, only the terms where $k =$ even contribute to the state in Eq. 4.26 meaning that the final state has half as many Fock-basis components as the $n_1 \neq n_2$ case, and each of these components is separated by integer multiples of 4. Indeed, this is a consequence of extended Hong-Ou-Mandel (HOM) type interference between the subtracted photons [185]. Since we are considering only the case where $n_1 = n_2$, the subtracted photons from each mode of the TMSV can only occur in even numbers. The simplest case of such scenario is $n_1 = n_2 = 1$, which means both the subtracted photons came from either mode 'a' or 'b', as one expects in a typical HOM experiment. As we show later, this extended HOM interference property can be used to prepare rotationally-symmetric states which can be used in bosonic quantum error correction schemes [21, 68].

## 4.4 Numerical Experiments

In general for arbitrary $n_1, n_2$, and $n_3$, we show in Fig. 4.2 how the probability to successfully generate any non-Fock state with either 2-or 4-fold phase space symmetry can be optimized by varying the subtraction beamsplitter (BS2 in Fig. 4.1) reflectivies based on the initial resource squeezing level. Since changing the subtraction beamsplitter reflectivity has no effect on the form of the output state other than to reduce the original TMS state squeezing parameter to a smaller effective squeezing, this feature

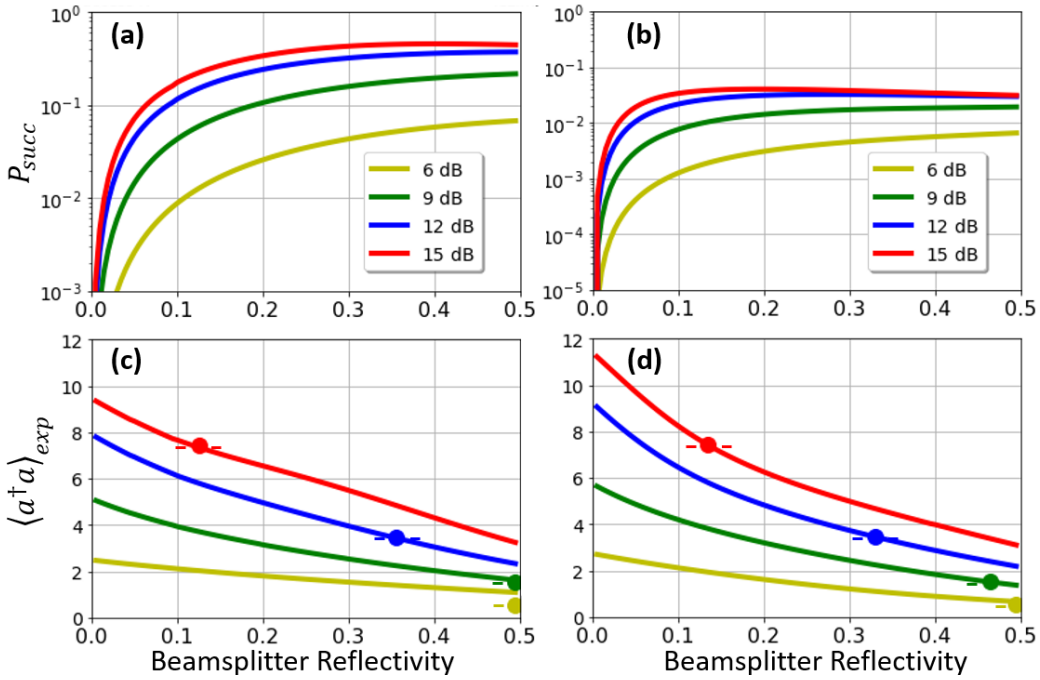



FIGURE 4.2: (a) The probability to successfully generate any non-Fock parity state (2-fold or 4-fold-symmetry) as the subtraction beamsplitter reflectivity is changed for many values of initial two-mode squeezing. (b) Probability to successfully generate only 4-fold-symmetry states. (c),(d) The average of the mean-photon number of the successfully generated states. The large dots indicate the mean photon number of a single mode of input state at each respective squeezing. Regions on the curves to the left of the dots show it is likely to increase the average energy of the mode when the state is successfully generated.

can be used as an external experimental knob to tune the mean output state energy and boost the generation rates. The 'non-Fock' qualifier means we exclude generated states that are composed of only a single Fock-basis component. Thus, the probabilities plotted simply sum the probabilites to obtain any interesting superposition state of the desired symmetry for a given subtraction beamsplitter reflectivity at each initial squeezing. Additionally, 4-fold rotation symmetric states also possess 2-fold symmetry, and thus Fig. 4.2(a) shows the probability to generate either type of states. In Fig. 4.2(b), we exclusively show the case of $n_1 = n_2$ where 4-fold symmetric states are generated.

The plots demonstrate that there is a threshold value for the beamsplitter reflectivity beneath which the generation probability drops precipitously, but increasing the reflectivity beyond about 10% ($r^2 = 0.1$) gives limited returns. Furthermore, in Fig. 4.2(c) and 4.2d we show the expected mean photon number, $\langle a^\dagger a \rangle_{\text{exp}}$ of the output

state given the specified initial conditions. These values are obtained by averaging over the mean photon number of each output state weighted by the probability of obtaining that state while excluding the non-superposition states. As one can see, the value of $\langle a^\dagger a \rangle_{exp}$ for the generated states only decreases as reflectivity increases, which intuitively follows considering more of the original TMS state energy is diverted toward the subtraction detectors. In all cases however, it is interesting to see that for all values of squeezing considered, the energy of the generated state can be increased beyond what was present in a single mode of the initial TMS state.

The large dot corresponding to each squeezing curve in the plots shows the mean photon-number of each single mode of the initial resource state and thus the regions to the left of this point indicate that on average, successfully generating the desired symmetry state with the indicated squeezing and beamsplitter reflectivities will increase the mean photon-number of the output state. This effect is attributable to subtracting photons from two-mode squeezed vacuum, and has been noted previously in the context of quantum state interferometry [186].

## 4.5 State Engineering for Bosonic Error Correcting Codes

In this section, we show how the proposed scheme can be tuned to generate exact binomial- and truncated cat-like codes by choosing certain combinations of PNR measurements. We provide a detailed analysis on required resource squeezing, success probabilities, and how lossy PNR detection affects the quality of the generated bosonic codes.

### 4.5.1 General two-component states

As mentioned earlier in Sec. 4.3, simply by tuning the initial squeezing parameter and post-selecting for either $n_1 = 1, n_2 = 0$ or $n_1 = 0, n_2 = 1$ produces any 2-fold symmetric two-component state of the form

$$|\psi_\pi\rangle \propto |m-1\rangle + \xi\,|m+1\rangle\,, \tag{4.31}$$

where the only requirement is that $|\xi| < \sqrt{\frac{m+1}{m}}$, and we let $\psi_\pi$ indicate a state whose Wigner function possess $\pi$ rotation symmetry (2-fold symmetric). Similarly, detecting

one photon in each of modes $c$ and $d$ ($n_1 = n_2 = 1$) allows the production of arbitrary 4-fold symmetric two-component states of the form

$$|\psi_{\pi/2}\rangle \propto |m-2\rangle + \zeta\,|m+2\rangle\,, \tag{4.32}$$

where $|\zeta| < \sqrt{\frac{(m+1)(m+2)}{m(m-1)}}$ and $\psi_{\pi/2}$ indicates a state with $\frac{\pi}{2}$ rotation symmetry (4-fold symmetric). In the above formulas, $\xi$ and $\zeta$ are complex numbers determined by the amplitude and phase of the squeezing parameter for the initial resource TMSV state. If the initial squeezing parameter is written in complex form as $z = Re^{i\phi}$, then beginning with a state that satisfies

$$e^{i\phi}\tanh R = -\xi\left(\frac{m}{m+1}\right)^{\frac{1}{2}} \tag{4.33}$$

and performing a PNR detection of $n_3 = m$ after the coherent subtraction leads to the creation of the 2-fold-symmetry state with parameter $\xi$ in Eq.4.31. In order to generate the 4-fold symmetric state in Eq. 4.32, the condition on the initial squeezing is

$$e^{i\phi}\tanh R = \sqrt{-\zeta}\left(\frac{m(m-1)}{(m+1)(m+2)}\right)^{\frac{1}{4}}, \tag{4.34}$$

where again a PNR detection of $n_3 = m$ must follow the coherent photon subtraction of the $n_1 = n_2 = 1$ case.

### 4.5.2 Binomial code state generation

We now consider two particular cases of 4-fold symmetric two-component states, in particular the binomial code words of Eq. 4.3. These binomial code words $|0_L\rangle$ and $|1_L\rangle$ can be exactly created by coherently subtracting $n_1 = n_2 = 1$ photons and performing respective PNR measurement of $n_3 = 2$ and $n_3 = 4$, provided the effective squeezing is chosen to satisfy the relation in Eq. 4.34 for $\zeta_{|0_L\rangle} = \sqrt{3}$ and $\zeta_{|1_L\rangle} = 1/\sqrt{3}$. In order to satisfy these requirements, the experimental resources must achieve squeezing thresholds of 10.63 dB to perfectly create the $|0_L\rangle$ state and 6.08 dB to create the $|1_L\rangle$ state. Note that for experimental squeezing above these thresholds, the beamsplitter reflectivity must be tuned to decrease the effective squeezing back to these quantities. However, squeezing below these thresholds does not mean that QEC with rotationally-symmetric code words is no longer possible. Lower squeezing means that the lower Fock-components of the code states will contribute more to the superposition, but as

long as the coefficients for the two code words are chosen such that they have the equal photon number moments, error correction is still possible as in accordance with Knill-Laflamme condition. In fact, the coefficients deviating from a binomial distribution may prove beneficial in specific circumstances if they are optimized to reduce specific error rates [21].

For the present discussion, we limit our focus to the binomial code states of Eq. 4.3. We show in Fig. 4.3(a) how the probability to successfully generated the binomial code states varies with initial resource squeezing. When the squeezing is exactly at the thresholds (10.63 dB for $|0_L\rangle$ and 6.08 dB for $|1_L\rangle$) to generate the desired state, the subtraction beamsplitters must have vanishing reflectivities, resulting in small success probabilities as shown on the left most parts of the plots in Fig. 4.3. For larger available initial squeezing, however, increasing the subtraction reflectivity until the effective squeezing is at the required thresholds boosts the success probability by three orders of magnitude for both code states. Note that these larger squeezing values for optimized success probability are within the realm of realistically achievable squeezing [14].

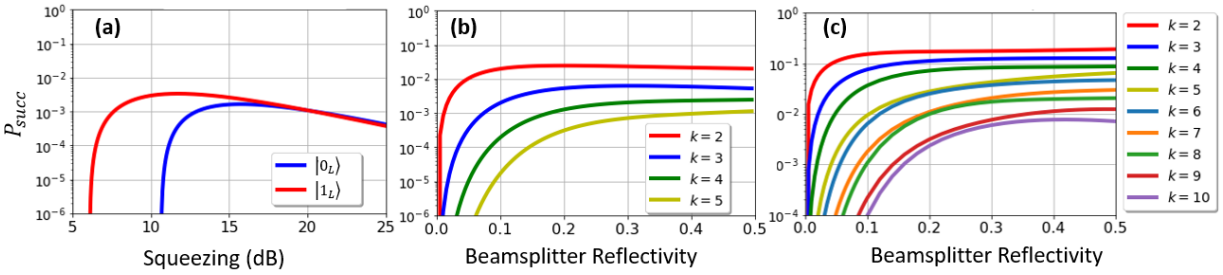



FIGURE 4.3: (a) Probability to successfully generate the code states $|0_L\rangle$ and $|1_L\rangle$ as initial squeezing is increased. The subtraction beamsplitters are tuned to satisfy Eq. 4.34. (b) Probability to successfully generate any 4-fold symmetric state and (c) any 2-fold symmetric state with $k$ Fock-basis components in each superposition as subtraction beamsplitter reflectivity is increased given an initial squeezing of 12 dB.

In Fig. 4.3(b), we display the success probability to generate multi-component 4-fold symmetric states with $k$ Fock-basis components in the superposition given an initial resource squeezing value of 12 dB as the subtraction beamsplitter reflectivity varies. Generating states with $k$ components in the superposition is a higher order process in the beamsplitter reflectivity as it requires $n_1 = n_2 \geq k - 1$, so the two component state described by Eq. 4.32 is the most common result. It is worth mentioning that the growing number of components in the superposition is a consequence of extended HOM interference. The post-selection is done for the particular case of $n_1 = n_2$

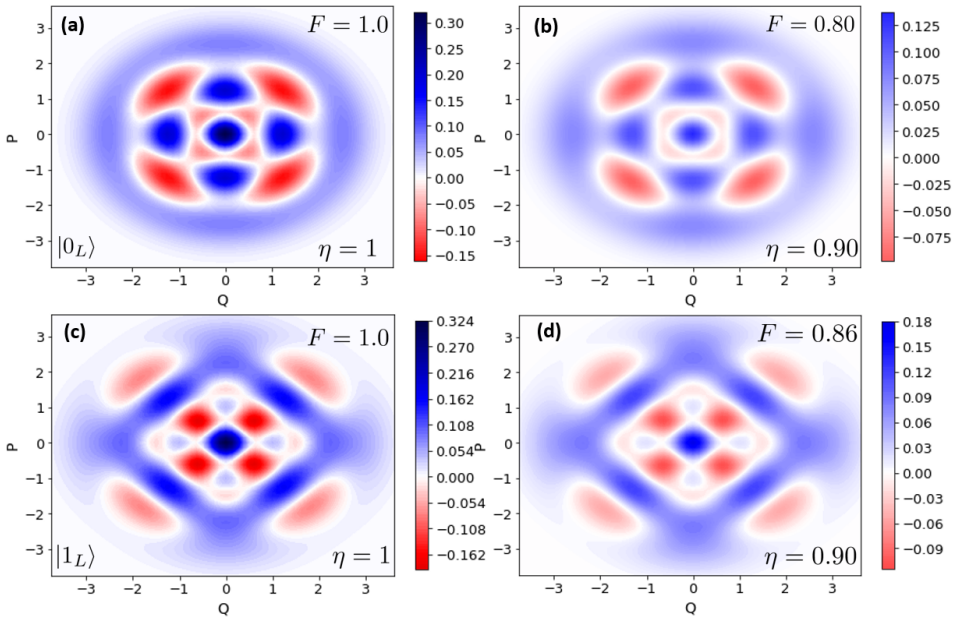



FIGURE 4.4: Wigner function visualization of binomial codewords. (a) and (c) are the ideal codewords generated for $|0_L\rangle$ and $|1_L\rangle$ when no losses are considered in our protocol. (b) and (d) are in the presence of 10% losses in the PNR measurement of $n_3$ photons. We see that the features such as phase space interference and the negativity of the Wigner functions are preserved with 10% losses.

after the HOM interference, but this can happen in many ways as long as we have $n_1^{(-)} = N - 2l, n_2^{(-)} = 2l; l \in [0, N/2]$, where $n_1^{(-)}$ and $n_2^{(-)}$ denotes the number of photons subtracted from modes '$a$' and '$b$', respectively before the HOM inferences at the second balanced beamsplitter, $BS$ 1. Likewise, Fig. 4.3(c) shows the success probability to generate $k$-component 2-fold-symmetry states with an initial 12 dB of squeezing.

In Fig. 4.4, we plot the Wigner functions of the generated binomial code words with the optimal effective squeezing of 10.63 dB and 6.08 dB for $|0_L\rangle$ and $|1_L\rangle$, respectively. On the left, the resultant Wigner functions (a) and (c) are for the case of perfect detection ($\eta = 1$) for all three PNR detectors. As expected, the Wigner functions display 4-fold symmetry owing to the photon-number support on every 4th Fock state. In an experimental realization of such state engineering protocol, the photon losses due to propagation though lossy optical components and inefficient PNR detections are inevitable and lead to decoherence of the desired state.

Next, we investigate the effects of imperfect detection ($\eta = 0.90$) for $n_3$ measurement for generating both the binomial code words. As evident from the Wigner

functions in Fig. 4.4 (b) and (d), the negativity of the Wigner functions is reduced while the phase space symmetry is still preserved. Note that we have considered the ideal detection for $n_1$ and $n_2$ PNR measurements since the deleterious effects of their imperfect detection can be reduced by lowering the reflectivity of the subtracting beamsplitters to ensure that the probability of subtracting more than one photon is negligible. While this prevents the decoherence of the generated code words caused by imperfect detection of $n_1$ and $n_2$, it lowers the overall success rate. Therefore, the PNR detectors with high detection efficiency are desired in such state engineering protocols.

We then evaluate the state Fidelity, Wigner Negativity, and Wigner Log Negativity (WLN) against the imperfect detection efficiency, $\eta$ for $n_3$ detection. The losses are modeled by setting up a fictitious beamsplittier of transmissivity $\eta$ in the path of $n_3$ photons.

Numerical simulations are displayed in Fig. 4.5. The generated state fidelity, defined as $F(\rho,\sigma) = \left(\mathrm{Tr}[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}]\right)^2$, with the ideal binomial code words in Eq. 4.3 is plotted against the overall detection efficiency in Fig. 4.5(a). We see that for a given detection efficiency, $|1_L\rangle$ performs better than $|0_L\rangle$, as evident from the larger fidelity difference at lower overall detection efficiency. In CV resource theories, the Wigner

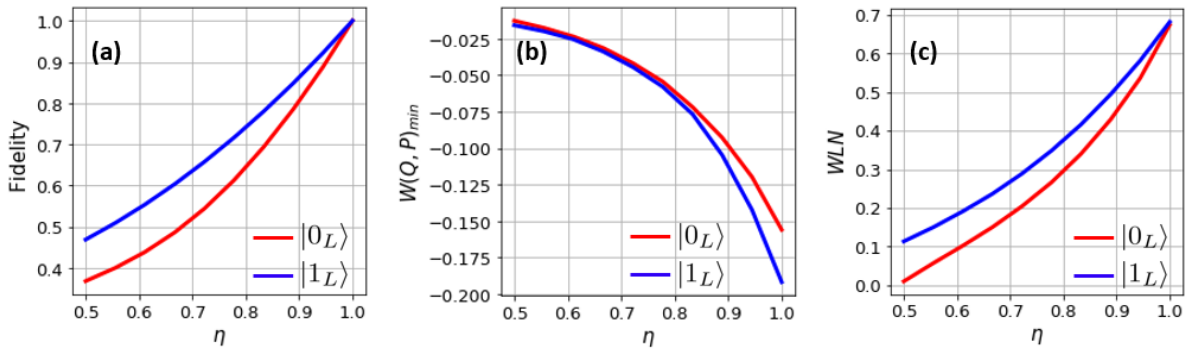



FIGURE 4.5: Effects of lossy PNR detection for $n_3$ photons. (a) State Fidelity with the ideal codewords. The minimum Wigner function amplitude and Wigner Log Negativity (WLN), (b) and (c), respectively.

Negativity (WN), i.e., the minimum value of the negativity of Wigner function, and Wigner Log Negativity (WLN) are the key measures to quantify the non-Gaussian nature of quantum resources which are essential for quantum computational advantages

and fault-tolerance [187]. The WLN is defined as

$$\text{WLN} := \ln\left[\int |W(q,p)|dqdp\right], \tag{4.35}$$

and it immediately follows that a non-negative Wigner function has zero WLN. The WN and WLN are plotted in Fig. 4.5(b) and Fig. 4.5(c), respectively.

We found that the non-Gaussian nature of the code words is preserved even for the detection efficiency as low as $\eta = 0.5$, as seen from the negative minimum amplitude of the Wigner function, $W(Q,P)_{\text{min}}$ and non-zero value of WLN. It is worth pointing out that the PNR detection efficiency of 98% is a experimentally feasible with state-of-the-art transition-edge sensors [188], and methods to obtain $> 99\%$ are possible [189]. This concludes the realistic implementations of our state engineering scheme for binomial codes.

### 4.5.3 Truncated cat code state generation

In this section, we investigate the feasibility of generating code words having support on higher photon-number. We show that such codes satisfy the Knill-Laflamme condition for faithful error correction against single-photon loss. Consider two events where we perform the detection combinations of $[n_1, n_2, n_3] = [0, 1, m]$ and $[n_1, n_2, n_3] = [0, 2, m]$ with $m \geq 2$, which we denote as $|\psi_{0,1,m}\rangle$ and $|\psi_{0,2,m}\rangle$, respectively. Using the results from Eqs. 4.26 and 4.27, the former case yields a two component 2-fold symmetric state of the form

$$|\psi_{0,1,m}\rangle \propto e^{-i\frac{\phi}{2}}(\tanh R)^{-\frac{1}{2}}\sqrt{m}\,|m-1\rangle - e^{i\frac{\phi}{2}}(\tanh R)^{\frac{1}{2}}\sqrt{m+1}\,|m+1\rangle\,, \tag{4.36}$$

while the latter case gives us

$$|\psi_{0,2,m}\rangle \propto e^{-i\phi}(\tanh R)^{-1}\frac{\sqrt{m}}{\sqrt{m+1}}\,|m-2\rangle + 2\frac{\sqrt{m+1}}{\sqrt{m-1}}\,|m\rangle + e^{i\phi}\tanh R\frac{\sqrt{m+2}}{\sqrt{m-1}}\,|m+2\rangle\,. \tag{4.37}$$

By inspection, it is easy to see that these states are mutually orthogonal where $|\psi_{0,2,m}\rangle$ has parity of $m$ and $|\psi_{0,1,m}\rangle$ has the opposite parity. As seen from the Wigner function plots of the example cases of $m = 5$ and $m = 7$ in Fig. 4.6, these types of states resemble truncated cat-like states with the associated fast interference fringes near the origin in phase-space. If we can arrange for the states to satisfy the Knill-Laflamme condition

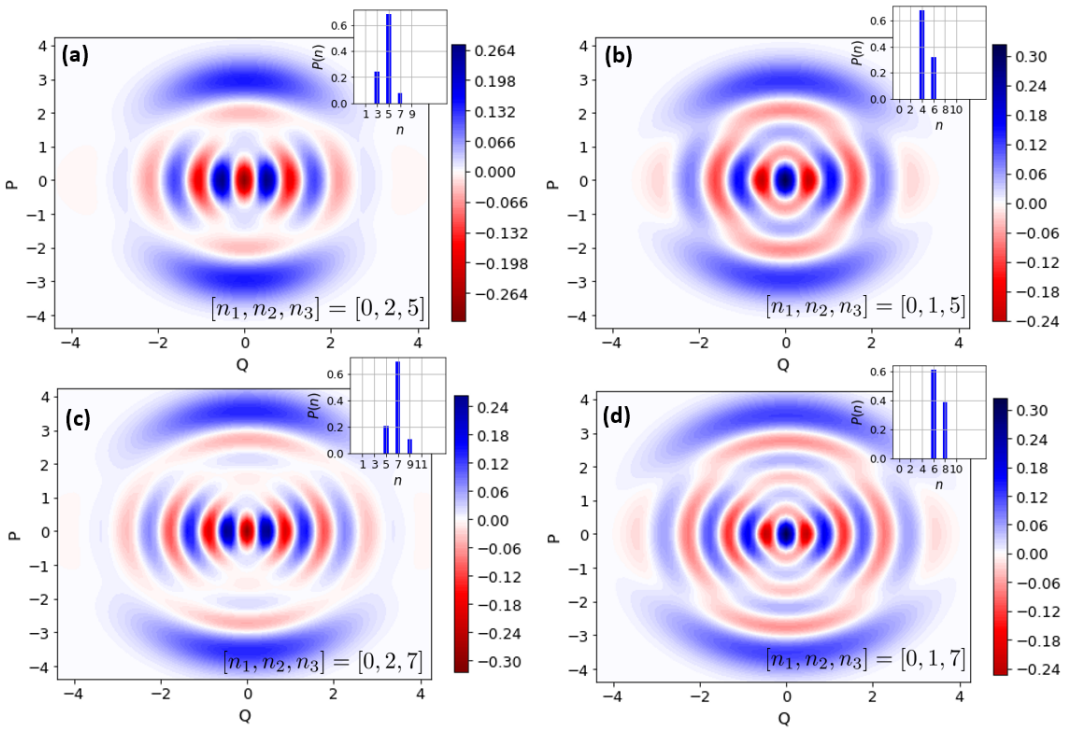



FIGURE 4.6: Wigner functions and photon-number distributions of codewords with 2-fold phase space symmetry. (a) and (b) are two orthogonal codewords with mean photon-number $\bar{n} \approx 4.6$ for the measurement outcomes of $[n_1, n_2, n_3] = [0, 2, 5]$ and $[n_1, n_2, n_3] = [0, 1, 5]$, respectively. Likewise, (c) and (d) are for $[n_1, n_2, n_3] = [0, 2, 7]$ and $[n_1, n_2, n_3] = [0, 1, 7]$ with mean photon number $\bar{n} \approx 6.7$. Insets show the photon-number distributions having support on odd and even photon numbers. The effective squeezing for (a) and (b) is 6.42 dB and 8.30 dB for (c) and (d).

defined in Eq. 4.2

$$\langle \psi_{0,1,m} | \hat{a}^\dagger \hat{a} | \psi_{0,1,m} \rangle = \langle \psi_{0,2,m} | \hat{a}^\dagger \hat{a} | \psi_{0,2,m} \rangle , \tag{4.38}$$

then the two states can be used as code words to protect against single dephasing errors.

Using the level of squeezing as a tuning experimental knob, the Knill-Laflamme condition can be exactly satisfied. In Fig. 4.7, we show how each pair of states can be made with the same experimental scheme without changing the subtraction beamsplitter reflectivities, and that for the correct squeezing (indicated by the solid dot where the solid and dotted lines cross in Fig. 4.7(a) ), the two orthogonal states each have the same mean photon number. This means that the scheme can be set with some initial squeezing and beamsplitter parameters and is capable of generating either of the pair of orthogonal states. For higher detected $m$, the two orthogonal states have nearly the

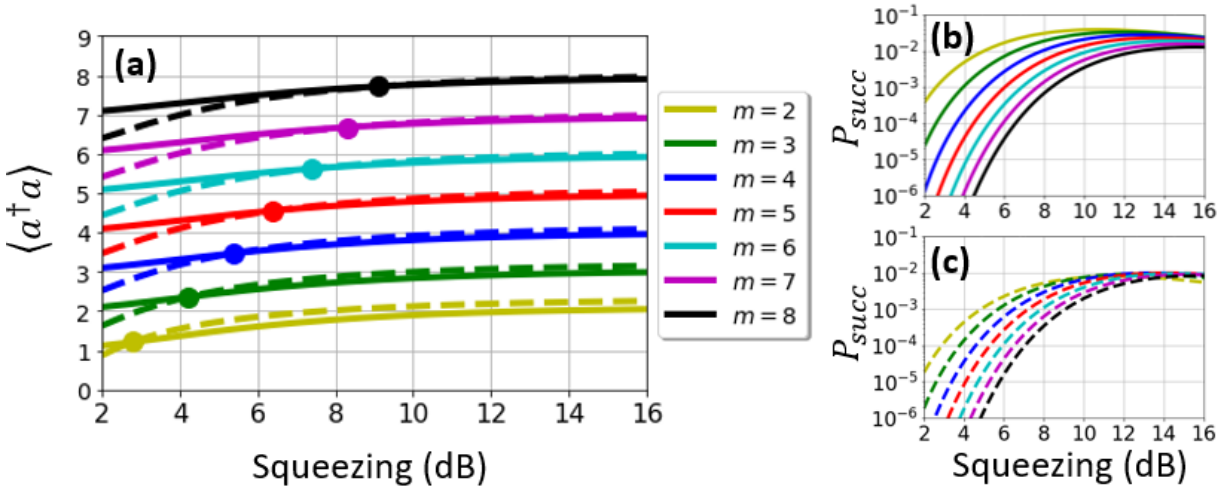



FIGURE 4.7: (a) Mean photon number of the orthogonal states generated by detecting $[0, 1, m]$ (dashed lines) and $[0, 2, m]$ (solid lines) photons as initial squeezing increases for a fixed subtraction beamsplitter reflectivity of 10%. For each $m$, there is a squeezing value that yields orthogonal states with the same mean photon number (indicated by solid dots) which is necessary to satisfy the Knill-Laflamme criteria. (b) The probability to generate the $[0, 1, m]$ or $[1, 0, m]$ states, which differ only by an optical phase, and (c) the probability to generate the $[0, 2, m]$ or $[2, 0, m]$ states.

same mean photon number for values of squeezing past the exact point as well, leading to slightly larger success probabilities for both $|\psi_{0,1,m}\rangle$ and $|\psi_{0,2,m}\rangle$ as shown in the respective plots in Fig. 4.7(b) and 4.7(c).

The relevant data associated with the intersection points in Fig. 4.7 is shown in Table 4.1. For each end detection of $n_3 = m$, we give the squeezing of the input state, the overall probability to successfully generated the states $|\psi_{0,1,m}\rangle$ and $|\psi_{0,2,m}\rangle$ at that input squeezing, and the mean photon number, $\langle \hat{a}^\dagger \hat{a} \rangle$, which is the same for both output states.

The general form of Eq. 4.26 can be used to find other pairs of orthogonal truncated cat-like states with more terms in the superposition beyond the simple case of measuring only zero and one or zero and two photons in the subtraction step. However, states created by subtracting larger numbers of photon generally arise less frequently due to decreasing success probabilities with unbalanced beamsplitters, and for arbitrary subtractions of $n_1$ and $n_2$, cat-like state production may not happen for every end PNR detection of $n_3$. In Sec. 4.5.5, we will consider how to enlarge the success probability by using multiplexing techniques.

| $m$ | $\langle a^\dagger a \rangle$ | Squeezing (dB) | Prob. [0,1,m] | Prob. [0,2,m] |
|---|---|---|---|---|
| 2 | 1.21 | 2.78 | $1.39 \times 10^{-3}$ | $8.46 \times 10^{-5}$ |
| 3 | 2.36 | 4.22 | $1.41 \times 10^{-3}$ | $1.82 \times 10^{-4}$ |
| 4 | 3.47 | 5.40 | $1.51 \times 10^{-3}$ | $3.08 \times 10^{-4}$ |
| 5 | 4.56 | 6.44 | $1.53 \times 10^{-3}$ | $4.32 \times 10^{-4}$ |
| 6 | 5.62 | 7.40 | $1.72 \times 10^{-3}$ | $6.28 \times 10^{-4}$ |
| 7 | 6.68 | 8.30 | $1.86 \times 10^{-3}$ | $8.34 \times 10^{-4}$ |
| 8 | 7.72 | 9.12 | $1.94 \times 10^{-3}$ | $1.03 \times 10^{-3}$ |

TABLE 4.1: Numerical data for the experimental parameters when the mean photon-number is equal for both the code words in Fig. 4.7

### 4.5.4 Arbitrary superposition states

True quantum information process will require the need to encode not only the logical states in the given code space, but also arbitrary superpositions of the two. This becomes challenging if generating the two logical states requires different experimental settings, such as when the initial squeezing is different for each of the states. When considering the class of binomial code states, however, there are several codes to choose from. Rather than the code words from Eq. 4.3, we could instead use code words

$$|\tilde{0}_L\rangle = |2\rangle\,, \quad |\tilde{1}_L\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |4\rangle\right), \tag{4.39}$$

which can faithfully correct against single-photon loss errors [21]. Our proposed scheme can be modified slightly in order to generate the arbitrary superposition $|\psi\rangle = u\,|\tilde{0}_L\rangle + v\,|\tilde{1}_L\rangle$ using the design shown in Fig. 4.8(a), where we make use of an additional ancillary state, $A\,|0\rangle + B\,|2\rangle$, where $A$ and $B$ are complex numbers. This ancillary state can be generated with a secondary device as described in the previous section, where a single photon was coherently subtracted from the TMS state and a single photon was detected and the end step for $m = 1$. As shown by Eqs. 4.31 and 4.33, the values of $A$ and $B$ for this ancilla can be tuned by changing the input squeezing to this secondary device.

Considering Fig. 4.8(a), let $|\psi\rangle_{ab}$ be the two-mode state after coherent-photon subtraction has occurred with $n_1 = n_2 = 1$. The output state after interfering with the

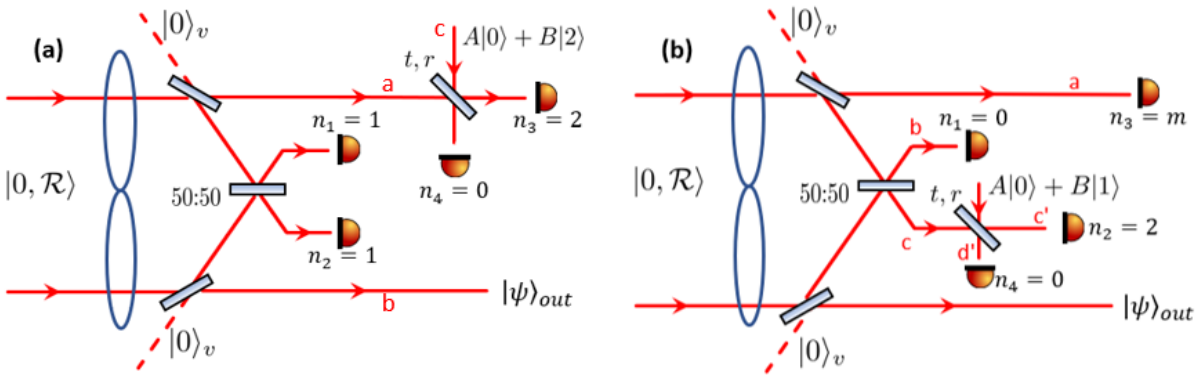



FIGURE 4.8: (a) Schematic to generate an arbitrary superposition of logical binomial code states. After performing a coherent-photon subtraction of $n_1 = n_2 = 1$, an ancillary state generated by a secondary device can be coupled to the system with an additional beamsplitter. The ancilla state coefficients and beamsplitter parameters can be tuned to control the weights of the output superposition obtained by successfully measuring $n_3 = 2$ and $n_4 = 0$. (b) Similar scheme for generative superpositions of orthogonal cat-like states $|\psi_{0,1,m}\rangle$ and $|\psi_{0,2,m}\rangle$. Here the ancilla state is injected prior to one of the photon-subtraction detections in order to generate the desired superposition.

ancilla state and subsequent detections is then given by

$$\begin{aligned} |\psi\rangle_{out} &= {}_a\langle 2|\,{}_c\langle 0|\,\hat{U}_{ab}(\theta)\,|\psi\rangle_{ab} \otimes (A\,|0\rangle + B\,|2\rangle)_c \\ &= {}_a\langle 0|\,{}_c\langle 0|\,\tfrac{1}{\sqrt{2}}(t\hat{a} + r\hat{c})^2\,|\psi\rangle_{ab} \otimes (A\,|0\rangle + B\,|2\rangle)_c \\ &= At^2\,{}_a\langle 2|\,\psi\rangle_{ab} + Br^2\,{}_a\langle 0|\,\psi\rangle_{ab}. \end{aligned} \tag{4.40}$$

Using Eqs. 4.32 and 4.34, we have that

$${}_a\langle 0|\,\psi\rangle_{ab} = |\tilde{0}_L\rangle \tag{4.41}$$

$${}_a\langle 2|\,\psi\rangle_{ab} = |\tilde{1}_L\rangle \tag{4.42}$$

when the initial squeezing parameter is $R = 0.756$ which corresponds to 6.51 dB of squeezing. Thus, by tuning the coefficients of the ancillary state in concert with the final beamsplitter parameters, it is possible to generate any arbitrary superposition of the logical states of the form $|\psi\rangle = u\,|\tilde{0}_L\rangle + v\,|\tilde{1}_L\rangle$ where $u = Br^2$ and $v = At^2$.

A similar procedure can be applied to the truncated cat code scenario. As discussion in Sec. 4.5.3, it is possible to generate two orthogonal cat-like code states through a single device with the same input squeezing. For the same post-selection on $n_3 = m$,

different values recorded on the coherent photon-subtraction detectors can lead to orthogonal parity states. However, if we make the addition shown in Fig. 4.8(b) by including the ancillary state $A|0\rangle + B|1\rangle$, we can make arbitrary superpositions between the two possible orthogonal states for any given $m$. This type of superposition can be approximated with a very weak coherent state or generated by other means [118, 190]. The result can be seen by examining the projective measurement that is made by the detectors. Using the mode indexing shown in the figure, the measurement is a projection on the Fock states $|m\rangle_a |0\rangle_b |2\rangle_{c'} |0\rangle_{d'}$. Note this mode indexing differs from previous mode indexing, but is used here for the sake of visual continuity. This measurement can be evolved backwards to one step before the the ancillary input and final beamsplitter, to give the effective measured state of

$$\begin{aligned} |\phi\rangle &= (A|0\rangle + B|1\rangle)_d \hat{U}^\dagger_{c'd'} |m\rangle_a |0\rangle_b |2\rangle_{c'} |0\rangle_{d'} \\ &= |m\rangle_a |0\rangle_b \tfrac{1}{2}(tc^\dagger - rd^\dagger)^2) |0\rangle_c |0\rangle_d \\ &= |m\rangle_a |0\rangle_b \left(At^2 |2\rangle - 2Brt |1\rangle\right)_c . \end{aligned} \tag{4.43}$$

From earlier, we found that taking the TMS state, sending in through the subtraction beamsplitters and projecting on $|n_3 = m\rangle_a |n1 = 0\rangle_b |n_2 = 1\rangle_c$ led to the state $|\psi_{0,1,m}\rangle$, while projecting on $|n_3 = m\rangle_a |n1 = 0\rangle_b |n_2 = 2\rangle_c$ led to the orthogonal state $|\psi_{0,2,m}\rangle$. With the modified scheme that effectively projects onto $|\phi\rangle$, we now have an experimental scheme to generate classes of states given by

$$|\psi\rangle = At^2 |\psi_{0,2,m}\rangle - Brt |\psi_{0,1,m}\rangle . \tag{4.44}$$

Tuning the new beamsplitter's phase and reflectivity allows for the creation of states with arbitrary coefficients in the superposition.

### 4.5.5 Success probability enhancement by resource multiplexing

Thus far, we have considered the generation of bosonic codes using only one setup presented in Fig. 4.1. To boost the success probability of generating a particular code one can employ multiplexing schemes, which have been used to enlarge the heralding probability of single-photons [191–193]. For a given success probability $P_{\text{succ}}$ per

state generation setup, the probability of generating the desired code with $N_{\text{MUX}}$ multiplexed resources is

$$P_{\text{MUX}} = 1 - (1 - P_{\text{succ}})^{N_{\text{MUX}}}. \tag{4.45}$$

For a given success probability tolerance $\delta = 1 - P_{\text{MUX}}$, one can estimate the number of multiplexed sources as

$$N_{\text{MUX}} = \left\lceil \frac{\log_{10}(\delta)}{\log_{10}(1 - P_{\text{succ}})} \right\rceil, \tag{4.46}$$

where $\lceil . \rceil$ is the ceiling function. In Fig.4.9(a), we plot the success probability against the number of multiplexed sources for states with both 2-fold (solid line) and 4-fold (dotted line) symmetry at assumed experimental squeezing values of 15 dB (red) and 6 dB (blue). The probabilities are calculated using beamsplitter parameters that maximize the single-scheme success probabilities for each case according to Fig. 4.2. Fig. 4.9(a) also shows the the multiplexed probability to generate the $|0_L\rangle$ (dashed green) and $|1_L\rangle$ (dashed black) states where again the other parameters are chosen to maximize individual scheme success probabilities. The figure inset zooms in to show where the multiplexed probabilities exceed 0.90. For respective probability tolerances of 0.1 and 0.01, only $N_{MUX} = 4$ and 8 schemes are needed to successfully generation a 2-fold symmetric superposition when initial squeezing is 15 dB.

The second panel of the figure, Fig. 4.9(b) performs the same multiplexing analysis on the truncated cat-like states whose Wigner function are shown in Fig. 4.6. The solid lines show the $P_{MUX}$ for the cases where the initial squeezing is chosen such that the mean photon number of orthogonal pairs ($|\psi_{0,1,m}\rangle$ and $|\psi_{0,2,m}\rangle$) is the same. However, the success probability can be significantly increased when larger initial squeezing is available, and orthogonal states still have approximately the same mean number of photons as seen in Fig. 4.7. The dotted lines in Fig. 4.9(b) show $P_{\text{MUX}}$ when the input squeezing is 12 dB.

## 4.6 Imperfect Binomial Code Words

In order to engineer the states $|0_L\rangle$ and $|1_L\rangle$, the final PNR detector must register two and four photons, respectively. However, when the detector is imperfect and has efficiency $\eta < 1$, this leads to mixtures instead of ideal code words. We wish to evaluate

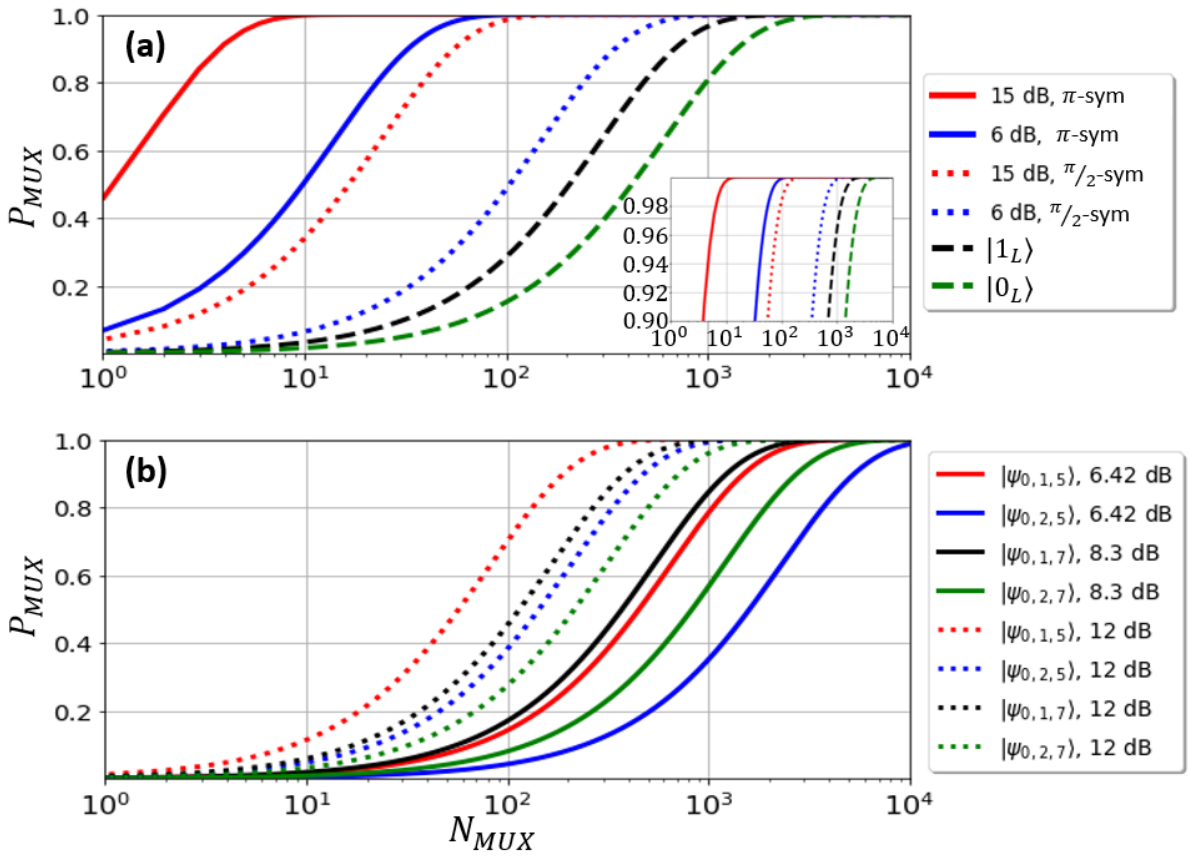



FIGURE 4.9: Including $N_{MUX}$ multiplexed protocols increases the overall success probability, $P_{MUX}$. (a) Curves are shown for the probability to generate any superposition 2-fold symmetric state using 15 dB and 6 dB initial squeezing (solid red and blue, respectively), any 4-fold symmetric at 15 and 6 dB (dotted red and blue), and the binomial code words $|1_L\rangle$ and $|0_L\rangle$ (dashed black and green) when initial squeezing maximizes the single-shot success probability. (b) The $P_{MUX}$ values are plotted for the truncated cat-like states of the Wigner functions shown in Fig. 4.6. Solid lines correspond to states generated from initial squeezing values that ensure the mean photon number is the same for each orthogonal state in the pair of $|\psi_{0,1,m}\rangle$ and $|\psi_{0,2,m}\rangle$. Dotted lines assumed a flat 12 dB of initial squeezed, and the mean photon numbers of the pairs of states are now only approximately equal.

how well error correction can proceed and try to find a bound on the detector inefficiency beneath which QEC is still possible. In our analysis, we ignore the dark counts of the detector which is a reasonable assumption for superconducting single-photon detectors. Let's begin with the state after coherent photon subtraction where a single photon was measured in each of the subtraction detectors, i.e., $n_1 = n_2 = 1$. The state after photon subtraction is

$$|\psi\rangle \propto (\hat{a}^2 - \hat{b}^2)\,|0, R\rangle_{ab}\,, \tag{4.47}$$

Note that this process can be assumed to be ideal, as slight inefficiencies in the subtraction detectors are less important than the end PNR efficiency since the probability to detect values of $n_1$ and $n_2$ larger than one is an order of magnitude smaller than $n_1 = n_2 = 1$ due to highly unbalanced beamsplitters. It is, however, desirable to have near-unity efficiency in order to subtract photons at higher rates.

For the $|0_L\rangle$ code word, performing lossy PNR detection of two photons on mode '$a$' of Eq. 4.47 yields the mixtures

$$\begin{aligned}\rho_{0_L} &= \frac{\mathrm{Tr}_a\left[\Pi_2\,|\psi\rangle\,\langle\psi|\right]}{\mathrm{Tr}\left[\Pi_2\,|\psi\rangle\,\langle\psi|\right]} \\ &\propto \mathrm{Tr}_a\Bigg\{\sum_{n,n'=0}^{\infty} e^{i\phi(n-n')}(\tanh R_{0_L})^{n+n'}\Big(\langle 2|_a\,(a^2-b^2)\,|n,n\rangle\,\langle n',n'|\,(a^{\dagger 2}-b^{\dagger 2})\,|2\rangle_a \\ &\qquad + 3(1-\eta)\,\langle 3|_a\,(a^2-b^2)\,|n,n\rangle\,\langle n',n'|\,(a^{\dagger 2}-b^{\dagger 2})\,|3\rangle_a + O[(1-\eta)^2]\Big)\Bigg\},\end{aligned} \tag{4.48}$$

where $\Pi_n = \sum_{m=n} p(n|m)|m\rangle\langle m|$ is the POVM for 'm' photons detection from a PNR detector with efficiency $\eta$ and $p(n|m) = \binom{m}{n}\eta^n(1-\eta)^{m-n}$ is the conditional probability of detecting 'n' photons out of 'm' photons incidented to the detector. For high efficiency PNR detectors with near unit efficiency, we can neglect orders of $(1-\eta)^2$ and higher. If we assume that the effective squeezing was chosen properly to for the desired state, then we can write the final output as the mixture

$$\rho_{0_L} = (1-\delta_{0_L})|0_L\rangle\,\langle 0_L| + \delta_{0_L}\rho_{E_{0_L}}, \tag{4.49}$$

where the erroneous component is

$$\rho_{E_{0_L}} = \frac{1}{8}\left(\sqrt{3}\,|1\rangle + \sqrt{5}\,|5\rangle\right)\left(\sqrt{3}\,\langle 1| + \sqrt{5}\,\langle 5|\right) \tag{4.50}$$

and the ratio

$$\frac{\delta_{0_L}}{1-\delta_{0_L}} = 3\sqrt{2}(1-\eta). \tag{4.51}$$

Similarly, one can go through the same procedure for the other code word with a lossly PNR detection of four photons. In this case we have

$$\begin{aligned}
\rho_{1_L} &= (1-\delta_{1_L})|1_L\rangle\langle 1_L| + \delta_{1_L}\rho_{E_{1_L}}, \\
\rho_{E_{1_L}} &= \frac{1}{32}\left(5|3\rangle + \sqrt{7}|7\rangle\right)\left(5\langle 3| + \sqrt{7}\langle 7|\right), \\
\frac{\delta_{1_L}}{1-\delta_{1_L}} &= \frac{8\sqrt{2}}{\sqrt{15}}(1-\eta).
\end{aligned} \tag{4.52}$$

The Wigner functions plotted in Fig. 4.4(b) and 4.4(d) show the effects the detector inefficiency when $\eta = 0.90$. Additionally, Fig. 4.5 plots the fidelity and Wigner log negativity of the imperfect code states as efficiency deviates from unity.

Now, we wish to examine how the error and recovery process is affected by the imperfections introduced by the non-ideal detector. Because the chosen code only protects against single photon loss, we need only consider effects up to first order in $\gamma$ from the Kraus operators describing the loss. We will use the approximate code word of $\rho_{0_L}$ as an example and treat the effects of no-jump errors ($E_0$) and single-jump errors ($E_1$) separately.

**No-jump errors**

Expanding in factors of $\gamma$, the Kraus operator for a no-jump event is

$$E_0 = 1 - \frac{1}{2}\gamma\hat{a}^\dagger\hat{a} + O[(\gamma)^2]. \tag{4.53}$$

To leading order in $\gamma$, this error acts on $\rho_{0_L}$ to give

$$\begin{aligned}
E_0\rho_{0_L}E_0^\dagger =& (1-\delta_{0_L})E_0|0_L\rangle\langle 0_L|E_0^\dagger + \delta_{0_L}\rho_{E_{0_L}} \\
&- \tfrac{1}{2}\delta_{0_L}\gamma\left(\hat{a}^\dagger\hat{a}\rho_{E_{0_L}} + \rho_{E_{0_L}}\hat{a}^\dagger\hat{a}\right).
\end{aligned} \tag{4.54}$$

The first term in the above expression is the ideal code word transformed by $E_0$

and is proven to be correctable with unitary operations [21]. The middle term contains the error from state generation and is orthogonal to the ideal code word. Due to the trace preserving nature of unitary matrices and the orthogonality of the first two components in the mixture, applying the unitary recovery operations will preserve the probability ratio of the error component with the desired corrected component. This means that after the recovery, the error from this term will not grow in magnitude but have the same contribution as the original state preparation error. Thus, if we can neglect the last term in Eq. 4.54, then error in the final code word after error and recovery will be the same as that of state generation and not grow in magnitude. In order to neglect this final term, we need $\delta_{0_L}$ to be of the same order as $\gamma$. From Eq. 4.51, we can determine a condition on the PNR detector efficiency in order for the error correction to be feasible up to $O(\gamma)$. This condition is

$$1 - \eta \sim \gamma. \tag{4.55}$$

**Single-jump errors**

The expanded single-jump Kraus operator is

$$E_1 = \left(\gamma - \tfrac{1}{2}(\gamma)^2\right)^{\frac{1}{2}} a + O[(\gamma)^{\frac{3}{2}}], \tag{4.56}$$

and acts on $\rho_{0_L}$ to yield

$$E_1 \rho_{0_L} E_1^\dagger = (\gamma - \tfrac{1}{2}(\gamma)^2)\Big[(1-\delta_{0_L})\hat{a}|0_L\rangle\langle 0_L|\hat{a}^\dagger + \delta_{0_L}\hat{a}\rho_{E_{0_L}}\hat{a}^\dagger\Big]. \tag{4.57}$$

The photon loss error is detected by measuring the photon number mod two for our chosen code space, which in this case, is a parity measurement. The error portion of the mixture, $\rho_{E_{0_L}}$, is orthogonal to the ideal state $|0_L\rangle\langle 0_L|$, so a measurement after a photon loss might incorrectly yield an even parity result since $\hat{a}\rho_{E_{0_L}}\hat{a}^\dagger$ has the same parity as $|0_L\rangle\langle 0_L|$. The probability to incorrectly diagnose the loss is thus given by the ratio

$$P_{inc} = \frac{\mathrm{Tr}[\delta_{0_L}\hat{a}\rho_{E_{0_L}}\hat{a}^\dagger]}{\mathrm{Tr}[(1-\delta_{0_L})\hat{a}|0_L\rangle\langle 0_L|\hat{a}^\dagger]} \approx \frac{9\delta_{0_L}}{10(1-\delta_{0_L})}. \tag{4.58}$$

Again, since the accuracy of our error correcting code is only up to first order in

$\gamma$, if the probability to incorrectly diagnose the error is negligible compared to $\gamma$ then our code still has efficacy. Using Eq. 4.51, this leads to the slightly stronger condition on the detector efficiency of

$$\eta \gg (1-\gamma). \tag{4.59}$$

This condition states that in order for the QEC to be effective, the losses from detector inefficiency must be considerably smaller than the expected losses the code words will protect against.

## 4.7 Summary and Outlook

This chapter explored the coherent photon subtraction from a two-mode squeezed vacuum to generate non-Gaussian quantum states with 2-fold and 4-fold phase space symmetry in the Wigner function, including exact binomial code states and truncated cat-like states. Such states have the desired Wigner function negativity required for quantum advantages in CV quantum information processing. While the proposed method is not deterministic, desired codes can be generated with rates of KHz-MHz by pairing up with state-of-the-art PNR detectors with high count rates. Furthermore, the success probabilities can be enlarged by starting with higher initial resource squeezing and multiplexing schemes, which could be achieved by monolithic integration on the same nanophotonic chip.

Moreover, this protocol can be extended to increase the rotation symmetry by utilizing PNR based breeding methods [68]. Higher rotation symmetry leads to enlarged spacing in the photon number basis, offering to correct for larger photon loss, gain, and dephasing errors. The truncated finite dimensional Hilbert space makes these codes hardware efficient and amenable for high fidelity unitary operations, thereby opening promising avenues for their applications in some other areas, in addition to fault-tolerant quantum computation, such as quantum metrology and quantum communication in optical quantum information processing. An important direction of future research is to further investigate such state engineering protocols to generate higher order codes allowing for correction of higher photon loss, photon gain, and dephasing errors.

# Chapter 5

# Quantum Sensing with Non-Gaussian States

Other than quantum computation and communication, quantum mechanics offers several routes to improve our abilities to perform measurements through quantum metrology. In terms of a phase-space perspective, the best classical physics can do is to reach a symmetric Heisenberg limit, such that uncertainty is equally small in two canonical observables, such as position and momentum. The majority of the field of classical metrology is concerned with optimizing technology to consistently achieve this symmetrized limit. In terms of measurement with light-based interferometers, the 'shot noise' arises from quantum mechanical vacuum fluctuations and provides the limit for classical physics.

However, with quantum mechanics and in particular squeezed light, we have the ability to asymmetrize the uncertainty such that, while the Heisenberg limit is maintained, one observable is more accurately measured while the conjugate acquires the excess noise. This is particularly of use when one is only interested in on of the observables, such as the phase of an optical beam, for instance. While squeezed light has been used in interferometry to reduce measurement noise below classical levels, this method will never attain a truly Heisenberg-limited sensitivity. To get to the limit, one must employ non-Gaussian states in some form.

This Chapter motivates a potential design of using photon-number-resolved detection to engineer a particular class of non-Gaussian states that would achieve Heisenberg-limited sensitivity [184]. The protocol design is similar to the coherent photon subtraction from the previous chapter, but the scheme is modified slightly to use only a single PNR detector.

## 5.1 Overview of Quantum-Enhanced Interferometry

A general interferometer, typified by the Mach-Zehnder interferometer (MZI) of Fig.5.1, measures the phase difference between two propagation paths by probing them with mutually coherent waves. From a purely undulatory standpoint, a sure way of ensuring such mutual coherence is to split an initial wave into two waves, for example by use of a beamsplitter. However, the unitarity of quantum evolution mandates that any two-wave-output unitary have a two-mode input as well — rather than a classical, single-mode input. Thus, the quantum description of a "classical" interferometer must feature an "idle" vacuum field in addition to the initial wave, and the quantum

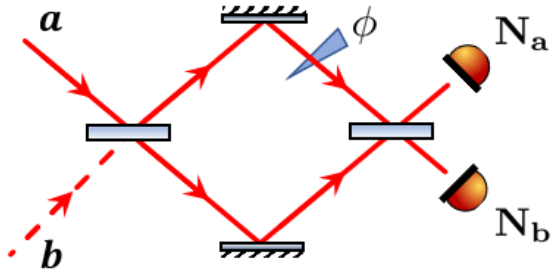



FIGURE 5.1: A Mach-Zehnder interferometer with phase difference $\phi$ between two optical paths. Both beamsplitters are balanced. Quantum splitting of input field $a$ implies interference with the vacuum field $b$.

fundamental limit of interferometric measurements is then dictated by the corpuscular statistics of the interference between the two inputs of the beamsplitter (Fig.5.1). In a classical interferometer, the vacuum fluctuations at the idle input port limit the phase difference sensitivity between the two interferometer arms to the quantum limit of classical interferometry [126], the input beamsplitter's shot-noise limit (SNL)[1]

$$\Delta\phi_{SN} \sim \langle N \rangle^{-\frac{1}{2}}, \tag{5.1}$$

where $\phi$ is the phase difference to be measured and $N = N_a + N_b$ is the total photon number operator. This limit is that of phase noise *inside the interferometer* and has nothing to do with, say, the single-mode properties of a coherent state (e.g., laser) input $|\alpha\rangle$ of photon-number deviation $\Delta N = |\alpha| = \langle N \rangle^{1/2}$ and phase deviation [2] $\Delta\theta \sim \langle N \rangle^{-1/2}$

[1]The SNL is often called "standard quantum limit." However, the latter was initially defined with a different meaning, in order to address the optimum error of quantum measurements in the presence of back action, such as radiation pressure on interferometer mirrors. [194, 195].

[2]From the number-phase Heisenberg inequality $\Delta N \Delta\phi \geqslant 1/2$, easily derived from the energy-time inequality [196].

*before the interferometer*. In fact, Caves showed that a Fock-state input $|n\rangle$, for which $\Delta N = 0$ and hence $\Delta\theta \to \infty$, still yields the SNL of Eq. (5.1) [126].

When both input modes of the interferometer are properly "quantum engineered," one can, in principle, reach the ultimate limit, called the Heisenberg limit (HL),

$$\Delta\phi_H \sim \langle N\rangle^{-1}, \tag{5.2}$$

which can clearly be many orders of magnitude lower than the SNL when $\langle N\rangle \gg 1$. A recent comprehensive review of quantum interferometry can be found in Ref. [197]. The first quantum engineering proposal to break through the SNL was Caves' idea to replace the vacuum state input with a squeezed vacuum [198], which has since been shown to optimize the quantum Cramér-Rao bound when the input field is a coherent state [199]. This was demonstrated experimentally [200, 201] and is now the approach adopted for high-frequency signals (above the standard quantum limit) in gravitational-wave detectors [202, 203]. Many other approaches have been investigated [204, 205], such as twin beams [206–211], N00N states [212–216], or two-mode squeezed (TMS) states. These different schemes were recently compared in terms of the quantum Cramér-Rao bound on their phase sensitivity [217].

It is important to recall here the essential result of Escher, de Matos Filho, and Davidovich: operating a realistic, i.e., lossy, interferometer at the HL requires losses to be no greater than $\langle N\rangle^{-1}$ [218], i.e., the grand total of the loss can never exceed one photon, on average. This result had been obtained earlier by Pooser and Pfister in the particular case of Holland-Burnett interferometry [219]: using Monte Carlo simulations for up to $n = 10\,000$ photons, it was shown that the phase error of a non-ideal Holland-Burnett interferometer scales with the HL if the losses are of the order of $n^{-1}$, and that larger losses degrade the scaling to a limit proportional to the SNL of $N^{-1/2}$, staying sub-SNL as long as photon correlations are present in the twin Fock input. This is consistent with the general result of Escher, de Matos Filho, and Davidovich for phase estimation [218]. Additionally, in the case of loss, optimal input states have been numerically calculated whose form generally depends on the interferometric efficiency and average photon-number input [220].

A direct consequence is that, if the total photon number is too large, ultimate-sensitivity interferometry cannot be Heisenberg-limited in the current state of technology: the most sensitive classical interferometer to date, the Laser Interferometer Gravitational-wave Observatory (LIGO) before the introduction of squeezed light boasted

$\Delta\phi_{SN} \sim 10^{-11}$ rad and is shot-noise-limited in some spectral regions, therefore featuring $\langle N \rangle \sim 10^{22}$ photons. While a Heisenberg-limited version of LIGO would only require $\langle N \rangle \sim 10^{11}$ photons to reach the same sensitivity, it would also require an unrealistic loss level of $10^{-11}$, the optical coatings on LIGO's mirrors "only" reaching already remarkable sub-ppm loss levels [221].

However, the maximally efficient use of photons by Heisenberg-limited interferometry can still be interesting provided we take into account this constraint of an ultimate loss level of $10^{-6}$. At this level, a 1064 nm interferometer with (arbitrarily chosen) 10 ms measurements would be allowed to reach the $10^6$-photon HL of 1 $\mu$rad with only 200 pW, whereas a classical interferometer would need $10^{12}$ photons, i.e., 200 $\mu$W, to have its SNL at 1 $\mu$rad. This can be of interest in situations where low light levels are beneficial, such as phase imaging of living biological tissue.

### 5.1.1 SU(2) interferometry

In order to motivate the approach of this paper, we review and compare and contrast some different quantum-enhanced sensing proposals in Table 5.1. The key points we examine are:
*(i)*, whether the input state enables performance at the HL;
*(ii)*, whether a direct interference fringe is observable;
*(iii)*, whether the $\langle N \rangle \gg 1$ regime is experimentally accessible. As will be shown, the new input state we propose in this paper is the only one that fulfills all three criteria.

The first two cases are interferometry with the vacuum field in one beamsplitter input port, leading to no quantum enhancement.

The third case makes use of Caves' squeezed input [198] into the previously unused port of the beamsplitter. This benefits from mature, high-level laser and quantum optics technology, with large average photon numbers from well stabilized lasers [222]. Case 3 benefits from the recent 15 dB squeezing record [14], but it does require that the phase difference between the squeezed state and the coherent state be controlled [223]. The gravitational-wave observatories of Advanced LIGO, Advanced VIRGO, and GEO600 all currently utilize squeezed light to improve sensitivity [224–226], and Advanced LIGO will soon implement frequency-dependent squeezing to improve sensitivity over a larger bandwidth [227].

TABLE 5.1: Characteristics and performance of different input states — except the N00N state* which is a state specified *inside* the interferometer. Also the fringe signal for the N00N state requires $n$-photon detection$^{(\ddagger)}$. The phase error $\Delta\phi$ is the quantum Cramér-Rao bound [217]. The state whose use we propose in this paper is the last one.

| | Input* state | Ref. | *(i)* $\Delta\phi$ | | *(ii)*fringe $\langle N_a - N_b \rangle$ | *(iii)* $\langle N \rangle \gg 1$? |
|---|---|---|---|---|---|---|
| 1. | $\|n\rangle_a \|0\rangle_b$ | [126] | $\frac{1}{\sqrt{n}}$ | SNL | $n\cos\phi$ | yes |
| 2. | $\|\alpha\rangle_a \|0\rangle_b$ | [126] | $\frac{1}{\sqrt{\langle N\rangle}} = \frac{1}{\|\alpha\|}$ | SNL | $\|\alpha\|^2\cos\phi$ | yes |
| 3. | $\|\alpha\rangle_a \|0,r\rangle_b$ | [198] | $\simeq \frac{e^{-r}}{\|\alpha\|}$ | sub-SNL | $\|\alpha\|^2\cos\phi$ | yes |
| 4. | $\|n\rangle_a \|n\rangle_b$ | [206] | $\frac{1}{\sqrt{2n(n+1)}}$ | HL | 0 | yes [210] |
| 5. | $\|n\rangle_a \|n-1\rangle_b$ | [217] | $\frac{1}{\sqrt{2n^2-1}}$ | HL | $\frac{1}{2}\cos\phi$ | possible |
| 6.* | $\frac{1}{\sqrt{2}}(\|n\rangle_a \|0\rangle_b + \|0\rangle_a \|n\rangle_b)$ | [212, 213] | $\frac{1}{n}$ | HL | $\sim \cos(n\phi)^{(\ddagger)}$ | unknown |
| 7. | $\frac{1}{\sqrt{2}}(\|n\rangle_a \|n\rangle_b + \|n+1\rangle_a \|n-1\rangle_b)$ | [204] | $\frac{1}{\sqrt{2n(n+1)-1}}$ | HL | $\frac{\cos\phi}{2} - \frac{\sin\phi}{4}\sqrt{n(n+2)}$ | unknown |
| 8. | $\sum_{k=0}^{Z} c_k \|n-Z+k\rangle_a \|n-k\rangle_b$ | | $\simeq \frac{\sqrt{Z}}{n}$ | HL | $\sim \frac{n}{2}\sin\phi$ | possible |

Case 4 in Table 5.1 is the twin Fock state input first proposed by Holland and Burnett [206], and which is implementable, to a good approximation, with large photon numbers by using an optical parametric oscillator above threshold [210, 211, 228–230]. The input density operator is then of the form, in the absence of losses,

$$\rho = \sum_{n,n'} \rho_{n,n'} \left|nn\right\rangle \left\langle n'n'\right|, \tag{5.3}$$

which can be a pure state ($\rho_{n,n'} \mapsto \rho_n \rho^*_{n'}$), e.g. the TMS state emitted by a lossless optical parametric oscillator (OPO) below threshold, or can be a general statistical mixture as

emitted by a lossless OPO above threshold [3]. It thus also benefits from the same mature OPO-based squeezing technology, with a record 9.7 dB reduction on the intensity-difference noise [230]. Moreover, the phase difference between the twin beams is irrelevant (being actually very noisy from being anti-squeezed) and thus need not be controlled before the interferometer. The generalized [231] Hong-Ou-Mandel [185] quantum interference responsible for twin beams breaking the SNL was demonstrated experimentally in an ultrastable phase-difference-locked OPO above threshold [210, 232, 233], with several mW of CW power.

An inconvenient feature of the Holland-Burnett scheme, however, is that the direct interference fringe disappears ($\langle N_a - N_b \rangle = 0$ in Table 5.1, a property also shared by the classical input $|\alpha\rangle_a |\alpha\rangle_b$) in contrast to all previous cases for which the fringe signal is proportional to the total photon number. This inconvenience can be circumvented by the use of Bayesian reconstruction of the probability distribution [206, 208, 219]. Although recently demonstrated for low photon numbers [234], this requires photon-number-resolved detection, which has yet to be experimentally accessible for large photon numbers. Another workaround is to use the variance of the photon-number difference, which is sensitive to $\phi$ [207] but whose signal-to-noise ratio is bounded by $\sqrt{2}$ [208]. Another idea is to use a heterodyne signal, which presents high visibility but is restricted to phase shifts ever closer to zero as the squeezing increases [235]. This was demonstrated experimentally as heterodyne polarimetry 4.8 dB below the SNL [211].

Case 5 in Table 5.1 is a variant of the twin Fock state, the "fraternal" twin Fock state [217], which does provide a direct fringe signal which being Heisenberg-limited, but the fringe signal is still extremely small as it does not scale with the input photon number. Recently, other variants of this state in the family $|n\rangle_a |m\rangle_b$ have been shown better performance than Holland-Burnett and N00N states in the presence of loss [236].

Case 6 stands out for several reasons. The N00N state refers not to an interferometer input state but to a state inside the Mach-Zehnder interferometer [212, 213]. While it yields performance at the HL, its experimental generation is experimentally inaccessible for $n \gg 1$; previous experimental realizations have used post-selected outcomes for $n = 3$ [214] and 4 [215], a method which fails to scale up to large photon numbers, though a more scalable method using coherent state displacement was also demonstrated [216]. Furthermore, N00N states are extremely vulnerable to loss [237],

[3] This is the most general mixture for which $\langle N_a - N_b \rangle = 0$ and $\Delta(N_a - N_b) = 0$.

although experimental efforts have attempted to side-step the issue by using more loss-tolerant techniques [101]. Last but not least, the use of a N00N state with $n$ photons requires $n$-photon detection, which is currently inaccessible in optics experiments with $n \gg 1$ (but may be easier to reach in atomic spectroscopy [212]).

Case 7 is the theoretical proposal of Yurke, McCall, and Klauder (YMCK) [204]. It features both performance at the HL and a strong fringe signal, but an experimental realization has yet to be determined.

Case 8 features the input proposed in this paper; it is the only one of the table that features HL performance, a clear interference fringe signal, and is experimentally feasible with demonstrated technology for large photon numbers. The state can be generated by using bright twin beams from which one or multiple photons have been coherently subtracted.

In addition to the above cases, photon subtraction has been suggested to be used in other interferometry schemes, such as interfering a coherent state and photon-subtracted squeezed vacuum [238] or subtracting a photon from each mode of a TMS state [186] (*distinguishable* subtraction, unlike what we propose here), but these schemes would require either a parity measurement or large photon-number resolving (PNR) measurements for end-detection, both of which are currently unfeasible for large photon numbers and will be discussed further in Sec. 5.7. As will be demonstrated, our method can make use of bright twin beams, thus scaling to large numbers of photons, and provides a directly measurable fringe with conventional photo-diodes. As mentioned, the presence of a strong fringe mitigates the need for numerically intensive reconstruction of the measured distribution.

### 5.1.2 A note on SU(1,1) interferometry

We have to our knowledge covered many of the basic approaches to quantum-enhanced optical interferometry with SU(2) based interferometers, but it is also important to mention SU(1,1) interferometry where the traditional passive beamsplitters are replaced by nonlinear two-mode squeezers. These devices do not work with intensity difference measurements [239], but must instead rely on other techniques, such as parity detection, which requires a PNR detector and is highly susceptible to loss [240]. This weakness to loss can be mitigated by seeding the input with a coherent state [241] and the PNR detector can be replaced with a click detector [242], but in neither case will the

detection sensitivity achieve the HL scaling. Alternatively, both the full SU(1,1) interferometer and a truncated version [243] can reach the HL with homodyne detection. Sensitivity below the SNL has been experimentally demonstrated [244] and this idea has been used to propose measuring angular displacement [245]. However, the use of homodyne detection on each mode means that the local oscillator phases on two auxiliary beams must be carefully controlled. Even if they are well phase-locked, an overall additive phase inside the interferometer may be disguised as the differential phase we wish to measure with the traditional SU(2) interferometer. In fact, the use of dual homodyne detection in SU(1,1) interferometric devices actually lends them the ability to more generally detect displacements, which prompted Caves to propose renaming them SU(1,1) displacement sensors [246]. In this work, we suggest using a squeezer in Sec. 5.5 as part of the state generation process which is reminiscent of the non-linearity present in SU(1,1) interferometry. However, this is just for convenience, as any twin-correlated source will work as stated in Sec. 5.6.C. Squeezing has also been used post-interferometer as a means to amplify the quantum signal to a classical one before detection in order to mitigate problems with inefficient detection [247]. From this point on, we restrict our study to that of SU(2) interferometry.

This paper is organized as follows. In section 5.2, we introduce the Schwinger-spin representation and calculate the Cramér-Rao bound by way of the quantum Fisher information for our proposed state. Section 5.3 demonstrates that the $Z$-photon coherently-subtracted twin-beam state has phase-sensitivity scaling with the HL when implemented with a traditional MZI. We then propose an experimental scheme to generate the desired state in section 5.4 and derive the result from a twin-beam input. Section 5.5 shows the results of numerical calculations when all approximations are disregarded, and section 5.6 discusses practical considerations of loss, detection imperfections, and the use of click detectors in place of PNR detectors during state generation. Section 5.7 compares cases with distinguishable subtraction in place of coherent subtraction, and then we conclude.

## 5.2 Quantum Fisher Information

The Fisher information is a well-known parameter that provides a means to quantify the amount of information contained by a parameterized random variable, and the Cramér-Rao inequality formulates an upper bound on the precision of an estimator in

terms of the Fisher information [248, 249]. This inequality has been extended to the quantum case [250] which is very useful in determining the bounds on the sensitivity for quantum interferometry given a specified input state [199], and is independent of the nature of the estimator. For a standard MZI, the estimator of interest is the phase difference between the two arms of the interferometer, for which the quantum Fisher information of a general pure input state is given by [199, 251]

$$\mathcal{F} = \langle \psi_{in} | \, U_{BS}^{\dagger} N_d^2 U_{BS} \, | \psi_{in} \rangle - \langle \psi_{in} | \, U_{BS}^{\dagger} N_d U_{BS} \, | \psi_{in} \rangle^2 , \tag{5.4}$$

where $U_{BS} = exp(\frac{i\pi}{4}(a^{\dagger}b + ab^{\dagger}))$ is a balanced beamsplitter and $N_d = a^{\dagger}a - b^{\dagger}b$ is the photon number difference operator. For the sake of convenience, we adopt for our calculation the Schwinger-spin SU(2) representation [252] initially proposed by Yurke et al. for quantum interferometers [204]. A fictitious spin $\boldsymbol{J}$ is defined from a pair of bosonic modes $(a, b)$ as

$$J_z = \frac{1}{2}(a^{\dagger}a - b^{\dagger}b) \tag{5.5}$$

$$J_x = \frac{1}{2}(a^{\dagger}b + ab^{\dagger}) \tag{5.6}$$

$$J_y = -\frac{i}{2}(a^{\dagger}b - ab^{\dagger}), \tag{5.7}$$

where $a$ and $b$ are the annihilation operators of each mode. These operators satisfy the canonical angular momentum commutation relations of the SU(2) algebra

$$[J_i, J_j] = i\epsilon_{ijk} J_k. \tag{5.8}$$

The operator $J_z$ represents the photon number difference operator between the two modes whereas $J_{x,y}$ are interference terms. The ease of working with Schwinger operators is further simplified by noting that the common eigenstates of $J_z$ and $J^2$, $|j, m\rangle$, take the form of two-mode Fock states:

$$|j, m\rangle = |n_a\rangle_a \, |n_b\rangle_b , \tag{5.9}$$

where we have

$$j = \frac{1}{2}(n_a + n_b); \tag{5.10}$$

$$m = \frac{1}{2}(n_a - n_b). \tag{5.11}$$

We can then see that the Fisher information can be expressed as

$$\begin{aligned}\mathcal{F} =& 4 \langle \psi_{in}| e^{(-i\frac{\pi}{2}J_x)} J_z^2 e^{(i\frac{\pi}{2}J_x)} |\psi_{in}\rangle \\ & - 4 \langle \psi_{in}| e^{(-i\frac{\pi}{2}J_x)} J_z e^{(i\frac{\pi}{2}J_x)} |\psi_{in}\rangle^2 \qquad (5.12)\\ =& 4 \langle \psi_{in}| J_y^2 |\psi_{in}\rangle - 4 \langle \psi_{in}| J_y |\psi_{in}\rangle^2 \qquad (5.13)\\ =& -\langle (a^\dagger b - a b^\dagger)^2 \rangle + \langle a^\dagger b - a b^\dagger \rangle^2. \qquad (5.14)\end{aligned}$$

This quantity is related to the phase estimation bound by

$$(\Delta\phi_d)^2 \geq \frac{1}{\mathcal{F}}. \tag{5.15}$$

It is important to note that $\Delta\phi_d$ is the general phase difference measurement whereas the quantum Fisher-limited phase error, i.e., when the inequality achieves equality, can be achieved for the correct estimator when $\phi$ deviates from an initially specified optimal value.

For an interferometry scheme using solely a coherent state input, the Fisher information can be easily calculated to yield exactly the expectation value of the input photon number, $\mathcal{F} = \langle N \rangle$. This leads to a maximum phase sensitivity of $\Delta\phi_d = \langle N \rangle^{-1/2}$, which is the well-known limit to interferometric measurement due to quantum noise with a classical input shown in Table 5.1, Case 2. In order to beat this limit, it becomes necessary to include something other than coherent states and vacuum as input, such as squeezed light addressed by Case 3 of Table 5.1. Although this will beat the $\langle N \rangle^{-1/2}$ scaling, it still cannot achieve the HL and has a maximum scaling of $\sim \langle N \rangle^{-2/3}$, which is only achievable with large squeezing ($\sim |\alpha|^{2/3}$ photons in the squeezed field for $|\alpha|^2$ photons in the coherent state) [126]. Because of this, reaching the HL requires the use of other types of quantum states. As a start, consider the simplest case of the final state in Table 5.1, (Z=1 case), the state given by

$$|\psi^{(i)}\rangle = A |n-1\rangle_a |n\rangle_b + B |n\rangle_a |n-1\rangle_b, \tag{5.16}$$

where $|A|^2 + |B|^2 = 1$. Calculating the quantum Fisher information for $\psi^{(i)}$ in Eq. 5.16, gives

$$\mathcal{F}^{(i)} = 2n^2 - 4n^2 \mathrm{Im}[A^* B] - 1, \tag{5.17}$$

which leads to the inequality

$$(\Delta\phi_d^{(i)})^2 \geq \frac{1}{2n^2 - 1} \tag{5.18}$$

when $A = B = \frac{1}{\sqrt{2}}$. In general, it is possible to achieve the Cramér-Rao bound by judiciously choosing the ideal measurement scheme [253]. If the ideal measurement is physically realizable, such as photon-number difference or single-mode parity, Eq. 5.18 demonstrates that the maximum sensitivity of $\Delta\phi_d$ can achieve a Heisenberg-limited scaling of $\langle N \rangle^{-1}$.

We now consider the most general state with arbitrary $Z$:

$$\begin{aligned} |\psi\rangle =& c_0 |n - Z\rangle_a |n\rangle_b + c_1 |n - Z + 1\rangle_a |n - 1\rangle_b \\ &+ \ldots + c_Z |n\rangle_a |n - Z\rangle_b \\ =& \sum_{k=0}^{Z} c_k |n - Z + k\rangle_a |n - k\rangle_b , \end{aligned} \tag{5.19}$$

where some initial $|n\rangle_a |n\rangle_b$ has been subjected to a coherent $Z$ photon subtraction, where the $Z$ photons could have been removed from either mode in any given combination. Derived in Appx. C.4, the Fisher information when all $c_k$ are the same and the mean photon number of the state is much larger than the number of photons subtracted is determined to be

$$\mathcal{F} \overset{(i)}{=} \frac{N^2}{Z}, \tag{5.20}$$

where $(i)$ represents the use of the condition that $N = 2n - Z \gg Z$. This leads to a bound on the phase sensitivity to be

$$(\Delta\phi_d)^2 \geq \frac{Z}{N^2}, \tag{5.21}$$

which shows the potential to achieve scaling that is proportional to HL. Although the bound given by the Fisher information is general, true infererometric performance is dependent upon the measurement scheme. We will now demonstrate that the class of states given by Eq. 5.19 follows the $N^{-1}$ scaling in a realistic MZI implementation when subtracting the resultant photo-detection currents, and will also yield a measurable interference fringe scaling with input photon number. From there, we will provide

a feasible experimental scheme based on correlated twin-beams and a single multi-photon subtraction event.

## 5.3 MZI Implementation Sensitivity

The above discussion demonstrated that the generalized photon-subtracted states are potentially viable for Heisenberg-limited interferometry as verified by the quantum Fisher information, but now it is important to ensure that the same $\langle N \rangle^{-1}$ scaling can be reached with a realistic implementation, such as subtracting photocurrents, and that there is a measurable fringe *also* scaling with $\langle N \rangle$ to ensure a sufficient signal above any electronics noise floor. The standard MZI is shown for reference in Fig. 5.1, where $\Delta\phi$ is the phase difference between the arms of the interferometer and the end detectors measured the respective mean photocurrents, $N_a$ and $N_b$. The input state, $|\psi\rangle_{ab}$, is transformed using the Schrödinger Picture by the interferometer to

$$|\psi'\rangle_{ab} = U_{BS}^{\dagger} P_{\Delta\phi} U_{BS} |\psi\rangle_{ab}, \tag{5.22}$$

where $P_{\Delta\phi}$ applies the relative phase shift $\Delta\phi$. This is equivalent, in the Schwinger representation, to applying the rotations

$$|\psi'\rangle_{ab} = e^{-i\frac{\pi}{2}J_x} e^{i\phi J_z} e^{i\frac{\pi}{2}J_x} |\psi\rangle_{ab}. \tag{5.23}$$

However, in order to determine if there is a measurable fringe, it is necessary to find $\langle N_a - N_b \rangle$, which is much easier to find when working in the Heisenberg Picture, as it simply corresponds to the transformed rotation operator, $\langle J_z' \rangle$, where any operator, $O$, is transformed by the MZI to become

$$O' = e^{-i\frac{\pi}{2}J_x} e^{-i\phi J_z} e^{i\frac{\pi}{2}J_x} O e^{-i\frac{\pi}{2}J_x} e^{i\phi J_z} e^{i\frac{\pi}{2}J_x}. \tag{5.24}$$

Using the transformation, we find

$$J_z' = \cos\phi J_z - \sin\phi J_x. \tag{5.25}$$

Making use of Eq. 5.9-5.11, we can rewrite the state of interest from Eq. 5.19 in the Schwinger representation to be

$$|\psi\rangle = \sum_{m=-s}^{s} c_m |jm\rangle , \tag{5.26}$$

where $s = \frac{1}{2}Z$, $j = n - s$, and from this point forward we consider all coefficients real and symmetric such that

$$c_m^* = c_m = c_{-m}, \tag{5.27}$$

which is the case for the physical state as derived in Appx. C.1. Supposing condition $(i)$ that the number of photons removed from the state by the subtraction process is considerably smaller than the number of photons remaining, i.e. $j \gg s$, the measurable fringe for the state in Eq. 5.26 is given by

$$\frac{j}{s}|\sin\phi| \overset{(i,ii)}{\leq} 2|\langle J_z'\rangle| \overset{(i)}{\leq} 2j|\sin\phi|, \tag{5.28}$$

with $(ii)$ denoting the use of results from Appx. C.3 where we bound the closeness of neighboring coefficients $c_m, c_{m+1}$ obtained from the experimental design. Eq. 5.28 shows that there is a direct measurable fringe scaling with $j$. Now, in order to determine the phase sensitivity when using $J_z'$ as a phase estimator, it is necessary to calculate the quantity

$$(\Delta\phi)^2 = \left[\Delta J_z' \left(\frac{\partial d}{\partial d\phi}\langle J_z'\rangle\right)^{-1}\right]^2 . \tag{5.29}$$

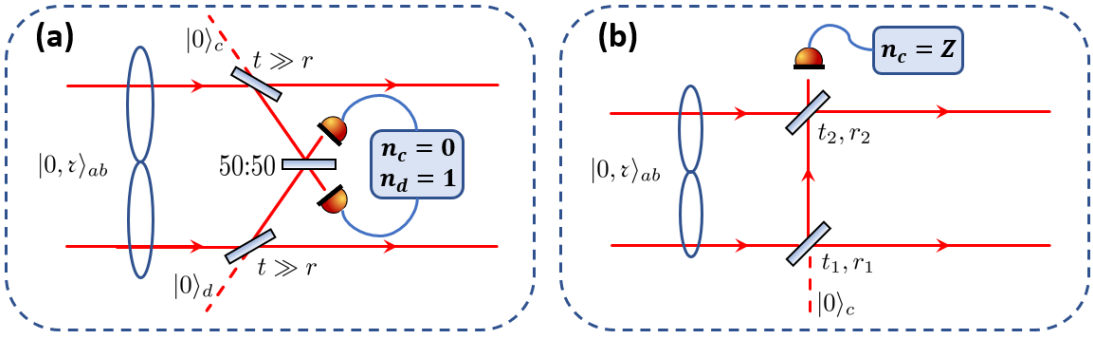



FIGURE 5.2: (a) Scheme using two click detectors to indistinguishably subtract a single photon from two squeezed modes. (b) Simplified version using a single detector that includes the case of coherent single-photon subtraction, but also extends the method to include coherent $Z$-photon subtraction when a PNR detector is used.

Derived in Appx. C.5, $\Delta\phi$ takes on a minimum value about the angle $\phi = 0$, yielding an upper bound on the phase-sensitivity of

$$\frac{\sqrt{s}}{j} \overset{(i)}{\leq} \Delta\phi_{min} \overset{(i,ii)}{\leq} \frac{2s^2}{j}. \tag{5.30}$$

We note again that $s$ is a small number based on the number of photons subtracted, so is simply a constant factor that does not affect scaling. Together, Eqs. 5.28 and 5.30 show that the general $Z$-photon subtracted state has sensitivity proportional to the HL and has a measurable fringe that scales with photon number, meaning that this class of states is potentially useful for quantum-enhanced interferometry, provided a physical realization can be found. Such a realization is provided in the next section, along with a detailed analysis of the possible performance.

## 5.4 State Generation

The specific case of $Z = 1$ has been previously examined in Ref. [183], where the proposed experimental design is shown here in Fig. 5.2(a). By sending each mode of a TMS state to a highly unbalanced beamsplitter, a single photon is subtracted from one of one of the modes. However, before the detection occurs, a third balanced beamsplitter is cleverly placed to erase the identifying path information about from which mode the photon came. By detecting exactly one photon on the combined two detection modes, the scheme implements a superposition of performing the subtraction on each mode to create a superposition of the type of states given in Eq. 5.16, which has the form of

$$\sum_n c_n(|n\rangle_a |n-1\rangle_b + |n-1\rangle_a |n\rangle_b), \tag{5.31}$$

where the coefficients $c_n$ depend on the strength of the initial TMS state.

Here, we demonstrate that this state, along with all classes of state given by Eq. 5.19 can be generated using only a single detector as per the scheme shown in Fig. 5.2(b). In this implementation, instead of mixing both modes with vacuum to perform the subtraction, an asymmetric scheme is used where only one mode mixes with vacuum while the other is mixed with the siphoned-off portion of the first beam. By tuning the beamsplitter coefficients, it is possible to make a variety of interesting entangle states beyond the scope of Eq. 5.19, but we will show that utilizing beamsplitters

with high transmissivity will preserve higher mean photon numbers of the outputs for any given detected number of photons, $Z$.

The first step of Fig. 5.2(b) starts with a TMS input state given by

$$|\phi_{in}\rangle = |0,\zeta\rangle_{ab}\,|0\rangle_c = \frac{1}{\cosh\zeta}\sum_{n=0}^{\infty}(\tanh\zeta)^n\,|n\rangle_a\,|n\rangle_b\,|0\rangle_c\,, \tag{5.32}$$

where $\zeta$ the squeezing parameter. The state is sent to the beamsplitters each with real reflectivity and transmissivity coefficients such that $t_i^2 + r_i^2 = 1$. Next,the mode $c$ is sent to a PNR detector to measure $Z$ photons and project the remaining modes into the desired state. The output state for the general case is derived in Appx. C.1 and may be useful for engineering interesting two-mode quantum states, but the most desirable case for use in sensitive interferometric measurements occurs when both beamsplitters are highly transmissive such that $t \gg r$. This leads to the approximate output state, conditioned on a detection of $Z$ photons, to be

$$\begin{aligned}|\phi_{out}\rangle \propto \sum_{n=0}^{\infty}(t^2\tanh\zeta)^n \sum_{k_{min}}^{Z}\sqrt{\binom{n}{k}\binom{n}{Z-k}\binom{Z}{k}} \\ \times e^{ik\varphi}\,|n+k-Z\rangle_a\,|n-k\rangle_b\,,\end{aligned} \tag{5.33}$$

where $\varphi$ is the phase difference between the two input modes to the second beamsplitter $(t_2, r_2)$ and $k_{min} = \mathrm{Max}(0, Z-n)$. This state is simply a superposition of the general class of states given in Eq. 5.19, where for each $|nn\rangle$ term, the $Z$ photons have been coherently subtracted from both modes in every possible configuration. When $\varphi$ is set to zero, the coefficients exactly follow the properties specified above in Eq. 5.27. An important point to note is that the value of $n$ in each input $|nn\rangle$ term is determine by the initial squeezing, but modified by the beamsplitter transmissivity to yield a new decreased effective squeezing, $\tanh\zeta \rightarrow t^2\tanh\zeta$. Regardless of the initial squeezing value, this reduction in squeezing means that only values of $t \approx 1$ will allow for useful interferometric measurements as large $n$ is desirable. Cases with $t$ deviating from unity may still work for interferometry applications if a large value of $Z$ is measured; however, this is prohibited by challenges in performing such a large PNR measurement.

## 5.5 Squeezed Input

Numerical simulations were performed using the Python package QuTip [129] to determine the minimum resolvable fringe from an MZI according to Eq. 5.29, where the input state was a TMS state that had undergone photon-subtraction as per Fig. 5.2b.

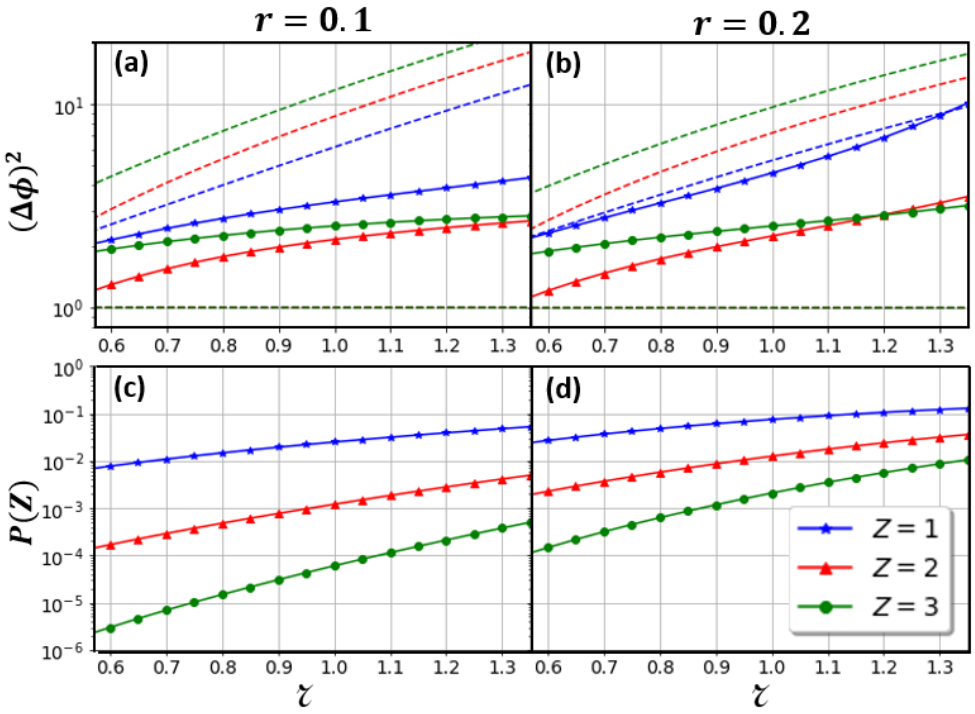



FIGURE 5.3: Minimum phase sensitivity for a photon-subtracted state where a PNR detector has detected $Z$ photons during state generation where beamsplitter reflectivities are $r_1 = r_2 = 0.1$ (a) and $r_1 = r_2 = 0.2$ (b). States are normalized to their respective HL (lower dotted line) with upper dotted lines showing the respective shot-noise levels. Probability to successfully detect $Z$ photons vs input state squeezing for the beamsplitter coefficients $r_1 = r_2 = 0.1$ (c) and $r_1 = r_2 = 0.2$ (d) are also shown.

The results are shown in the top panel of Fig. 5.3, where $Z$ varies from one to three photons as the initial squeezing is increased, and both beamsplitters are taken to be equivalent with reflectivity $r_1 = r_2 \equiv r$. All cases beat the shot-noise limit, but it is clear that decreasing the reflectivity leads to a larger quantum-enhancement, and larger values of $Z$ increase the overall phase-sensitivity. This behavior may seem odd, as looking at Eq. 5.30 shows that the overall scaling of the sensitivity is slightly worse as $Z$ increases; however, detecting a larger number of photons acts to increases the mean photon-number of the resultant state before the interferometer, and higher-order

error terms from the use of beamsplitters with non-vanishing reflectivity drop off more quickly with increasing $Z$. For input states with lower mean-photon number, post-selecting on a larger $Z$ can actually beat the HL set by the states heralded at smaller $Z$. However, this comes at the cost of diminishing success rates.

By comparing Fig. 5.3a and b, it is clear that for the relatively large value of $r = 0.2$ the weak beamsplitter approximation employed earlier is no long valid for smaller $Z$, as the state with $Z = 1$ is only marginally better than the classical case. However, decreasing the reflectivity to $r = 0.1$ suffices to give a quantum advantage that scales when increasing the input energy for all detected values of $Z$. This highlights the importance of choosing an optimal beamsplitter coefficient that pairs with the desired post-selection value, $Z$, to yield increased sensitivity while balancing success rates for a given input state.

## 5.6 Practical Considerations

### 5.6.1 PNR detector loss

In a realistic implementation, one might think that the state generation is highly sensitive to PNR detector efficiency. However, provided any imperfections in the detector are not too drastic, the present scheme is tolerant to this inefficiency. Instead of projecting mode $c$ of the intermediate state following the beamsplitters onto an ideal measurement of $Z$ photons, consider a detector positive-operator valued measure (POVM) given by

$$D(Z) = \sum_{l=Z}^{\infty} \binom{l}{Z} \eta^Z (1-\eta)^{l-Z} \left|l\right\rangle \left\langle l\right|, \tag{5.34}$$

where the detector has efficiency $\eta$ and the device registers a detection of $Z$ photons. The output density matrix will be given by

$$\rho \propto \mathrm{Tr}_c \left[D_c(Z) \left(\left|\psi\right\rangle \left\langle\psi\right|\right)_{abc}\right], \tag{5.35}$$

where $\left|\psi\right\rangle$ is the state following both beamsplitters given by Eq. C.4. This leads to

$$\rho \propto \sum_{l=Z}^{\infty} \binom{l}{Z} \eta^Z (1-\eta)^{l-Z} \rho_l, \tag{5.36}$$

where $\rho_l$ is the density matrix of the pure state for an $l$-photon detection given in Eq. C.5. However, taking $r_1 = r_2 \equiv r$ and $r \ll 1$ yields

$$\rho \approx \sum_{l=Z}^{\infty} \binom{l}{Z} \eta^Z (1-\eta)^{l-Z} \left(\frac{r}{t}\right)^{2l} \rho_l' \tag{5.37}$$

$$\propto \sum_{l=Z}^{\infty} \binom{l}{Z} \left(\frac{r^2(1-\eta)}{t^2}\right)^l \rho_l', \tag{5.38}$$

where $\rho_l' = |\psi_l'\rangle \langle\psi_l'|$ is an unnormalized pure state with

$$|\psi_l'\rangle = \sum_{n=0}^{\infty} \frac{(t^2 \tanh \mathfrak{r})^n}{\cosh \mathfrak{r}} \sum_{k_{min}}^{l} \sqrt{\binom{n}{k}\binom{n}{l-k}\binom{l}{k}} \times e^{ik\varphi} |n+k-l\rangle_a |n-k\rangle_b . \tag{5.39}$$

Eq. 5.38 reveals that even for values of $\eta$ deviating significantly from unity, we have a final state that is approximately pure, i.e., $\rho \approx \rho_Z'$ when $r^2(1-\eta) \ll 1$, since all terms with $l > Z$ in the sum can be neglected. This regime can always be reached by decreasing $r$ until the approximation holds. However, it is important to note that an imperfect detector reduces the success probability of the scheme by a factor of $\eta^Z$.

### 5.6.2 Click/no-click detector

In a similar manner to the above discussion, the PNR detector can be safely replaced with a click/no-click detector without ill-effect. Since we showed above that the effects of losses during state generation can be considered negligible if the beamsplitter reflectivity is tuned correctly, here we consider an ideal click detector with perfect efficiency. In the case where the detector reads no signal, then the POVM is simply the vacuum as before. In the case of a registered click, the POVM is given by

$$D_{clk} = \sum_{l=1}^{\infty} |l\rangle \langle l| . \tag{5.40}$$

Now, the output density matrix is simply a mixture of all possible detected states with $Z \geq 1$,

$$\rho \propto \sum_{l=1}^{\infty} P(l) \rho_l, \tag{5.41}$$

where each of the pure state in the mixture is weighted by the probability of the number of photons that actually went to the detector. Since the quantities of interest are $\langle J_z' \rangle$ and $\Delta J_z'$, we can use the linearity of Eq. 5.41 along with the results of Appx. C.5 to see that

$$\langle J_z' \rangle = -\sin\phi \sum_{l=1}^{\infty} \mathrm{Tr}\left[J_x \rho\right] \tag{5.42}$$

$$\propto \frac{-\sin\phi}{n} \tag{5.43}$$

for each component state in the mixture, $\rho_l$. We also know that the maximum phase sensitivity is achieved about the angle of $\phi = 0$, so the only necessary term from $\Delta J_z'$ is $\Delta J_z$. Thus, we can find that

$$(\Delta J_z)^2 = \sum_{l=1}^{\infty} \mathrm{Tr}[J_z^2 \rho] \tag{5.44}$$

$$\overset{(i)}{\leq} \sum_{l=1}^{\infty} \frac{l^2}{4} \left(\frac{r}{t}\right)^{2(l-1)}, \tag{5.45}$$

which leads to a phase sensitivity about $\phi = 0$ of

$$\Delta\phi \propto \frac{1}{n}. \tag{5.46}$$

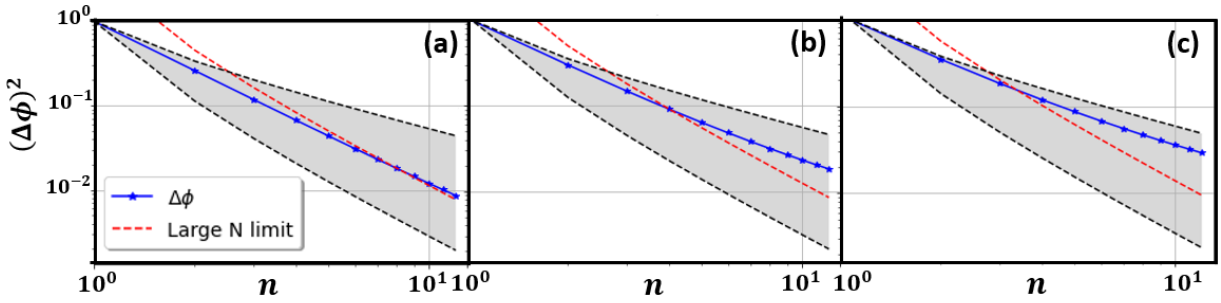



FIGURE 5.4: Minimum phase sensitivity (blue) for a photon-subtracted state with beamsplitter reflectivities $r_1 = r_2 = 0.1$ and input state $|n, n\rangle$, where the PNR detector has been replaced by an ideal click detector. Efficiency $\eta$ determines the combined transmission of the interferometer and end detector efficiencies and is $\eta = 1$ (a), $\eta = 0.95$ (b), $\eta = 0.9$ (c). The shot-noise and HL bound the shaded region, and the red dotted line shows the Cramér-Rao bound for the large $N$ limit with $Z = 1$ given by Eq. 5.18.

These results can be easily understood by realizing that each of the possible pure

state components in the mixture that results from the click detection has phase sensitivity scaling with the HL around *the same* reference phase of $\phi = 0$. Additionally, the high unbalancing of the beamsplitters makes the components in the mixture diminish with increasing $l$.

The results from this section are experimentally significant in that not only does the detector not need the ability to resolve photon-numbers, but the imperfect efficiency negligibly degrades the purity of the resultant state, provided that there is precise control in ensuring the beamsplitter reflectivity is small.

### 5.6.3 General OPO output

The general output of an arbitrary twin-beam source, such as obtained from an above-threshold OPO, can be described as a mixture of the form

$$\rho = \sum_{n,n'} \rho_{n,n'} \left|nn\right\rangle \left\langle n'n'\right| . \tag{5.47}$$

Despite being a mixture, we show in Appx. C.6 that the indistinguishable multi-photon subtraction protocol also works for this input and leads to a Heisenberg-limited output with phase sensitivity scaling as $\Delta\phi \sim \langle N\rangle^{-1}$. The intuition behind this result follows similar reasons for why a click/no-click detector also fails to ruin Heisenberg-limited sensitivity; although the input is a mixture, each $|n_i, n_i\rangle$ component from the initial twin-beam is formed into a superposition with phase sensitivity of $\sim n_i^{-1}$ about the reference angle of $\phi = 0$. Since all of the components of the mixture are Heisenberg-limited at the same reference phase, then so is the entirety of the mixture. As such, when considering imperfections such as loss, simulations with the input state $|n, n\rangle$ can be reliably used to gauge to effectiveness of the process for an arbitrary twin-beam source with mean photon number $n$.

Fig. 5.4 compares the phase sensitivity of a photon-subtracted state with an $|n, n\rangle$ input for the realistic scenario of a heralding PNR detector replaced by a click detector, and when the interferometer has losses. As shown in the previous section, inefficiencies in the click detector channel can be accounted for by choosing a small enough beamsplitter reflectivity, so the click detector here is considered ideal for simplicity. Additionally, if we assume that losses on both interferometer arms are identical, then the losses can be commuted and combined with detection losses to result in a single

overall efficiency of $\eta$ in each arm [254].

Each plot shows the phase sensitivity along with the large $n$ limit given by the Cramér-Rao bound for a single photon-subtraction given by Eq. 5.18 (red dotted line). When the beamsplitter reflectivities are small enough for the approximation to apply, as in Fig. 5.4a, it can be seen that the phase-sensitivity readily follows the Cramér-Rao bound, which scales with the HL up to a constant factor. This constant factor comes from the fact that since each mode in the twin beam has $n$ input photons, the total number of photons in the system is $2n$ leading to a HL of $(2n)^{-1}$; however, the ideal scaling with the single-photon subtracted state goes as $n^{-1}$. The second and third panels demonstrate that realistic interferometric losses do not drastically reduce the phase sensitivity, such as occurs with states less resilient to loss, such as $N00N$ states. As shown in Fig. 5.4b, the slope of the phase sensitivity still scales better than $\langle n\rangle^{-1/2}$ for $\eta = 0.95$, and it is possible to achieve considerably larger overall $\eta$ in practical experiments with advanced low-loss optical coatings and highly efficient detectors [221]. Fig. 5.4c, with $\eta = 0.9$, shows how larger losses on the order of $\langle N\rangle^{-1}$ bring the resultant phase-sensitivity away from the HL and back to scaling with the shot-noise, in agreement with the general result of Escher, de Matos Filho, and Davidovich [218].

## 5.7 Distinguishable Subtraction without Coherence

Thus far we have focused solely on the state we propose, the $Z$-photon coherently-subtracted twin-beam state. However, it is illustrative to compare with the case where $Z$ photons have been distinguishably subtracted from a twin-beam state *without* preserving coherence. One example of this is the case explored previously by Carranza and Gerry [186], where a single photon had been subtracted from each mode of a TMS state. As discussed in Sec.5.6, the coherent photon subtraction presented in this article will yield a state with HL sensitivity for any twin-beam correlated photon input, but for the comparison here, we consider the input to be a TMS state and the subtraction process to be ideal with beamsplitter reflectivity r$\rightarrow$ 0. In this case, the beamsplitter operator can be approximated to first order be

$$U_{BS} = e^{\theta(a^\dagger b - a b^\dagger)} \approx 1 + \theta(a^\dagger b - a b^\dagger), \tag{5.48}$$

where $r = \sin\theta << 1$. A distinguishable subtraction of $z_1$ photons in mode $a$ and $z_2$ photons in mode $b$ transforms the TMS input to

$$\begin{aligned} &|\phi_{out}\rangle \propto \\ &{}_c\langle z_1|\,{}_d\langle z_2|\left(1+\theta(a^\dagger c - a c^\dagger)+\theta(b^\dagger d - b d^\dagger)\right)|0,\zeta\rangle_{ab}\,|00\rangle_{cd} \\ &\qquad \propto \sum_{n=0}^{\infty}\frac{n!(\tanh\zeta)^n}{\sqrt{(n-z_1)!(n-z_2)!}}\,|n-z_1\rangle_a\,|n-z_2\rangle_b\,. \end{aligned} \tag{5.49}$$

Although subtracting photons from the TMS state has the interesting effect of increasing the average photon number of the remaining distribution as noted previously [186, 255], each term of the superposition has only one Fock-basis component in each mode. Each term takes the form of $c_{nm}\,|n\rangle\,|m\rangle$, which does not yield an intensity difference fringe scaling with the photon number unless $n >> m$, which is experimentally unfeasible with photon-subtraction due to the low probability of subtracting large numbers of photons. Because of this, states of the form of Eq. 5.49 will require a different measurement scheme to reach the HL scaling. Introduced by Bollinger *et al.* [212] for trapped ions and later for optical interferometry by Gerry [256], measuring the parity of one output mode, $\Pi_b = (-1)^{b^\dagger b}$, allows for reaching HL sensitivity with these types of input states. With the parity detection scheme, the phase uncertainty becomes

$$(\Delta\phi_\Pi)^2 = \frac{1-\langle\Pi_b\rangle^2}{|\partial\langle\Pi_b\rangle/\partial\phi|^2}. \tag{5.50}$$

Optical parity detection is currently an experimental challenge and requires large PNR measurements unless the measured state is of Gaussian nature [257]. Even without this obstacle, the nature of parity predisposes the measurement to be highly susceptible to loss and external noise as the presence or absence of just one photon flips the value of the measurement. This effect is shown in the top panel of Fig. 5.5, where the phase sensitivity is normalized to the Heisenberg limit (black dot-dashed line) and plotted for several states as the overall interferometric efficiency decreases from $\eta = 1$ to $\eta = 0.90$. Solid curves show the phase sensitivity for the state in Eq. 5.49 with squeezing parameter $\zeta = 0.9$ where no photons were subtracted (blue), a single photon was subtracted from one of the two modes (red), and the case where one photon was subtracted from each mode (green). The SNL for each curve is shown as a dashed line in the corresponding color. Additionally, Fig. 5.5 shows the phase sensitivity of the coherently subtracted TMS states with $\zeta = 0.9$, Eq. 5.33, in fine dotted lines with colors corresponding to the

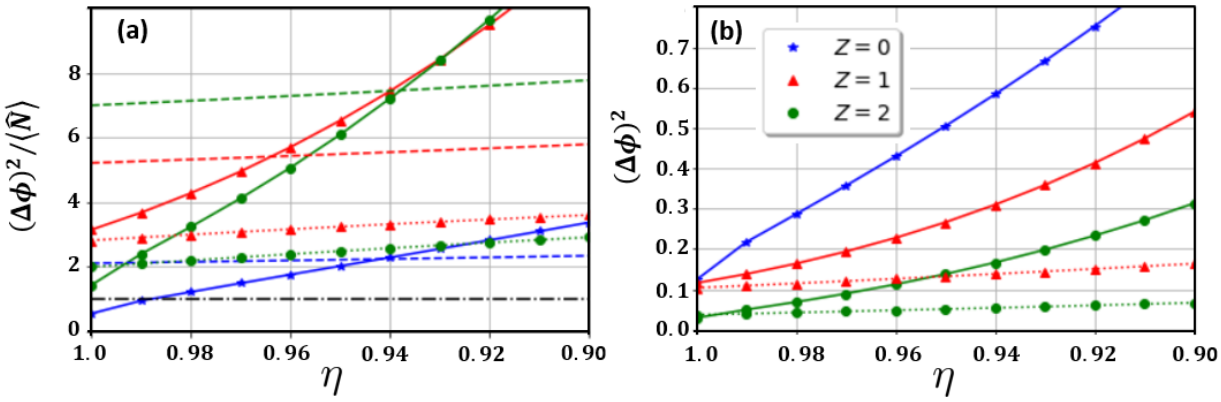



FIGURE 5.5: Minimum phase uncertainties for TMS states of squeezing parameter $\tau = 0.9$ with coherent and distinguishable photon subtraction where $Z$ total photons have been subtracted. Solid lines indicate distinguishable subtraction with $Z = 1$ (red triangles) indicating a photon was subtracted from just one of the two modes and $Z = 2$ (green circles) showing the case where one photon was subtracted from each mode. Dotted lines with markers indicate coherent subtraction between the two modes. Phase sensitivity was calculated using Eq. 5.50 (parity detection) for cases with distinguishable subtraction and with Eq. 5.29 (intensity different detection) for coherent subtraction. Top: Sensitivities normalized to HL (black dot-dashed line) where the dashed unmarked lines show the corresponding SNL for each corresponding color. Bottom: Raw phase sensitivity comparison.

total number of subtracted photons (red for $Z = 1$, green for $Z = 2$). For these states, the phase sensitivity is evaluated with intensity difference detection as per Eq. 5.29. Note that the state for $Z = 0$ is unchanged from the input TMS state to leading order in $r$. When there is no loss ($\eta = 1$), it can be seen that all of the states considered have phase sensitivity nearing the HL, and in fact, parity detection with a pure TMS state can actually beat the HL as noted by Anisimov *et al.* [258]. However, even for small losses well below $\langle N\rangle^{-1}$, the states with parity based detection veer from the HL and quickly become worse than the classical sensitivity bound, unlike the class of states proposed in this article. This behavior is well in line with the results of Ref. [240]. The bottom panel of Fig. 5.5 shows the phase sensitivity for each state on a true scale without normalization and demonstrates three key points:

*(i)*, increasing the number of subtracted photons acts to increase the overall mean photon number of the remaining distribution and thus increase phase sensitivity;

*(ii)*, in the absence of loss, both types of states achieve HL sensitivity if the correct measurement scheme is used;

*(iii)*, finally, for nonzero losses the phase sensitivity of schemes relying on parity detection is quickly degraded whereas the states utilizing an intensity difference measurement retain much of their sensitivity.

## 5.8 Summary

This chapter proposed and studied a nontrivial modification of the twin-beam input for Heisenberg-limited quantum interferometry, which features coherently indistinguishable multi-photon subtraction that leads to a superposition of photon-subtractions. This modification brings about a strong fringe signal — absent from the unmodified twin-beam input — while preserving Heisenberg-limited operation. The loss behavior is consistent with what is now well known about Heisenberg-limited interferometry. The experimental implementation should be feasible with state-of-the-art technology, for example using a stable OPO above threshold [210, 232, 233] and photodetectors with single-photon sensitivity. Detectors with PNR capability may improve the phase-sensitivity of the resultant photon-subtracted state by detecting $Z > 1$, but this is not a requirement, and reasonable experimental losses still result in phase-sensitivity beating the standard quantum limit. We believe it is possible to operate at no more than $10^6$ photons per detection time bin, so as to be compatible with the lowest achievable optical losses and splitting ratios.

# Chapter 6

# Photon-Number-Resolving Detection

In order for a quantum optics-based implementation to achieve a quantum speed-up in QC, reach the Heisenberg limit in quantum sensing, or provided error correction against Gaussian noise (such as a lossy channel), it is necessary to have access to non-Gaussian states and operations. Although directly engineering a higher-order Hamiltonian would allow for deterministic non-Gaussian state generation, the relative weakness of available third-order and higher non-linearities make alternative options more appealing. One option is to pursue single-photon sources, and there has been a great deal of experimental progress in this area [137, 259, 260]. These sources are seen partly as an 'engineering problem' and are likely to continue to improve. However, end-to-end efficiencies of on-demand pure single photons have yet to eclipse 60%, and building up arbitrary non-Gaussianity from single photons remains difficult.

Alternative to directly producing non-Gaussian states, one of the most powerful tools in an experimentalist's toolbox is detection, which gives access to large projection-based nonlinearities. Any detection result is, of course, probabilistic, but we saw in each chapter thus far that projecting part of an entangled state can give rise to a variety of potentially interesting non-Gaussianity. Due to the repeated focus on the use of photon-number-resolving (PNR) detection, this chapter will focus on a more detailed description of PNR detectors and experimental implementations. Low PNR detection is feasible with several schemes including time-multiplexed detection [169] and number-resolving superconducting nanowire single-photon detectors detectors [143, 166, 167], but the highly efficient transition-edge sensor (TES) remains the detector with the ability to resolve the most photons at the highest quantum efficiency [124, 165], although with detection speed considerably slower than the 100 MHz repetition rates made possible by SNSPDs [261].

In this chapter, I will first outline the basic theory of positive-operator-valued measures for PNR detectors, which is a way to completely characterize the detector response. From here, I will delve into the specifics of PNR detection with a TES system. These detectors have been chosen for this work due to their high quantum efficiency, extremely low dark count, and ability to resolve high photon numbers. I will discuss experimental methods of optimizing detection signal and show how applying similar methods as outlined in Ref. [165] can lead to even high PNR capability, with up to 30 photons resolvable per detection channel. These results usher in a new mesoscopic era of quantum detection, where goals such as the 50 photon-number detection required by the well-known cubic-phase generation protocol are nearly within reach [48].

## 6.1 Detector POVMs

Quantum detectors can be completely described by their positive-operator-valued measure (POVM), which is a way to write all possible measurement outcomes as eigenstates of some Hermitian observable [262]. Each measurement outcome, $k$, is associated with the Hermitian operator that is the POVM element $\hat{\Pi}_k$. These operators are complete, $\sum_k \hat{\Pi}_k = \hat{\mathbb{1}}$, and for a detector insensitive to phase (such as ones that measures photon-number), can be written as

$$\hat{\Pi}_k = \sum_{n=0}^{\infty} P(k|n) \left|n\right\rangle \left\langle n\right| . \tag{6.1}$$

Here, $P(k|n)$ is the conditional probability that a detector will register $k$ events for an input of $n$ photons. For an ideal detector, such as one that registers $k$ counts if and only if there was a true $k$ photon measurement of the quantum state, we have that the ideal POVM elements are

$$\hat{\Pi}_k = \sum_{n=0}^{\infty} \delta_{n,k} \left|n\right\rangle \left\langle n\right| = \left|k\right\rangle \left\langle k\right| . \tag{6.2}$$

Unfortunately, few detectors can be approximated as ideal, even though it it possible to measure large numbers of photons with high quantum efficiency using TES devices [124, 165]. One main imperfection to consider is detector inefficiency. This information can be directly incorporated into the POVM by examining the conditional probability, $P(k|n)$. For a detector of efficiency $\eta \leq 1$, this conditional probability is given by

$$P(k|n) = \binom{n}{k} \eta^k (1-\eta)^{n-k}. \tag{6.3}$$

The above formula can be simply understood as follows: suppose we send $n$ photons to a detector, but it only registers $k < n$ events. In what way could this happen? Well, each of the $n$ photons has a probability of $\eta$ to successfully reach the detector, and a probability of $(1-\eta)$ to be lost. Thus, if we measured $k$, we know $k$ photons made it to the detector and $n-k$ photons were lost, and hence the $\eta^k(1-\eta)^{n-k}$ factor. The binomial coefficient comes because we have no way of knowing *which* of the $n$ photons were the $k$ that actually made it to the detector, so we must consider all possible combinations. Thus, the POVM for a detector with loss is

$$\hat{\Pi}_k^{(\eta)} = \sum_{n=0}^{\infty} \binom{n}{k} \eta^k (1-\eta)^{n-k} \left|n\right\rangle \left\langle n\right| . \tag{6.4}$$

Detector inefficiency can be equivalently modeled as subjecting the quantum state to be measured to a beamsplitter and tracing out over the reflected mode before sending what remains to an ideal detector, but this has the potential to be more computationally complex, as one must calculate the evolution of the quantum state through the beamsplitter including interference with an ancillary vacuum mode and a partial trace (or use the Kraus operator formalism for loss). The scaling of this problem fairs worse as more modes are considered. Depending on the experimental situation or calculation being performed, having both the beamsplitter and POVM methods at one's disposal to model reality can be quite useful.

Now, consider the additional inclusion of external noise, such as dark counts. These effects can be modeled by imagining that in addition to the signal we attempt to measure, there is some other aspect of the environment also contributing to the measured signal. Suppose the noise is described by a photon-number distribution with probabilities given as $P_{env}(m)$. For each registered event of $k$ photons, some number may have come from the $n$ photons in the true distribution and some may come from the $m$ photons in the environment. Because additional noise photons may be registered, $k$ is no longer necessarily smaller than $n$. The conditional probability with noise is then

$$\begin{aligned} P_{env}(k|n) &= P(k|n)P_{env}(0) + P(k-1|n)P_{env}(1) + ... + P(0|n)P_{env}(k) \\ &= \sum_{j=0}^{k} P(k-j|n)P_{env}(j), \end{aligned} \tag{6.5}$$

where $P(k-j|n)$ is just the conditional probability from Eq. 6.3. In fact, the above equation is simply a discrete convolution between the noiseless conditional probability

distribution and the distribution of the environmental noise alone. Suppose the environmental noise is given by a thermal distribution,

$$P_{env}(m) = \left(\frac{1}{\bar{m}+1}\right)\left(\frac{\bar{m}}{\bar{m}+1}\right)^{m}, \tag{6.6}$$

where $\bar{m}$ is the average photon number registered by the detector of the environment. In this case, the conditional probability becomes

$$P_{env}(k|n) = \frac{1}{\bar{m}+1}\sum_{j=0}^{k}\left(\frac{\bar{m}}{\bar{m}+1}\right)^{j}\binom{n}{k-j}\eta^{k-j}(1-\eta)^{n-k+j}. \tag{6.7}$$

The detector POVM element including loss and environmental noise is thus given by

$$\hat{\Pi}_{k}^{env} = \sum_{n=0}^{\infty}\sum_{j=0}^{k} P(k-j|n)P_{env}(j)\,|n\rangle\langle n|\,. \tag{6.8}$$

The above are all conditional probabilities based on an $n$ photon input, but we can now calculate the probability to register $k$ photons for a given signal density matrix, $\rho_s$, by

$$P(k) = \mathrm{Tr}\left[\hat{\Pi}_{k}^{env}\rho_s\right]. \tag{6.9}$$

Alternative to the above, one can instead view environmental noise on the detector as measuring the photon number of two independent density matrices, and then only recording the sum of the two measured outcomes. This is exactly how one would treat sending separate optical modes to a bucket detector, where the detector cannot distinguish between the pair of modes [117]. In this case, the probability to detect $k$ overall photons can be written as

$$P(k) = \sum_{j=0}^{k}\mathrm{Tr}\left[(\hat{\Pi}_{j}^{(\eta)}\otimes\hat{\Pi}_{k-j}^{(\eta)})\rho\right], \tag{6.10}$$

where $\rho$ is the two-mode density matrix comprising the signal and the environment. When the signal is in a product state with the environmental noise, $\rho = \rho_s\otimes\rho_{env}$,

Eq. 6.10 becomes

$$\begin{aligned} P(k) &= \sum_{j=0}^{k} \mathrm{Tr}\left[\hat{\Pi}_j^{(\eta)}\rho_s \otimes \hat{\Pi}_{k-j}^{(\eta)}\rho_{env}\right] \\ &= \sum_{j=0}^{k} \mathrm{Tr}\left[\hat{\Pi}_j^{(\eta)}\rho_s\right] \mathrm{Tr}\left[\hat{\Pi}_{k-j}^{(\eta)}\rho_{env}\right] \\ &= \sum_{j=0}^{k} P_s(j) P_{env}(k-j), \end{aligned} \tag{6.11}$$

which is again just a convolution of the two independent probability distributions.

Fortunately, TES systems have dark counts of less than 1 Hz, so these effects can be neglected for our PNR detectors. Thus the POVM for a detector with loss given by Eq. 6.4 is sufficient when modeling the TES.

## 6.2 Transition-Edge Sensor

A TES is a PNR detector that makes use of the extremely temperature-dependent resistance of a superconductor near the phase transition. Our detection system is comprised of superconducting tungsten wafers that operate with a critical temperature near 100 mK. When light is incident on the chip, the thermal energy of an absorbed photon acts to locally break the superconducting state and induce a spot of non-zero resistance. Near the critical temperature, the resistivity of the superconductor is nearly linear with temperature, as shown in Fig. 6.1(a). These detectors are optimized to be highly absorptive at the desired wavelength, and while our detectors achieve above 90% quantum efficiency at the target wavelength of 1064 nm, TES systems can achieve higher efficiencies of $\eta = 0.95$ [124] or even $\eta = 0.98$ [188].

Each TES chip is connected to an induction loop in parallel with a reference resistance and fed a biasing current as shown in Fig. 6.1(b). Once a photon is absorbed by the detector, the resistivity of the chip increases linearly and causes the current flowing through the inductor to drop. This is then picked up and amplified by the extremely sensitive Superconducting Quantum Interferometric Device (SQUID) magnetometers to be read out as a classical signal. These SQUIDs themselves must be properly biased to achieve maximum sensitivity. When testing the sensitivity with an external input, the SQUID fringe should be clearly visible and reach maximum amplitude as shown

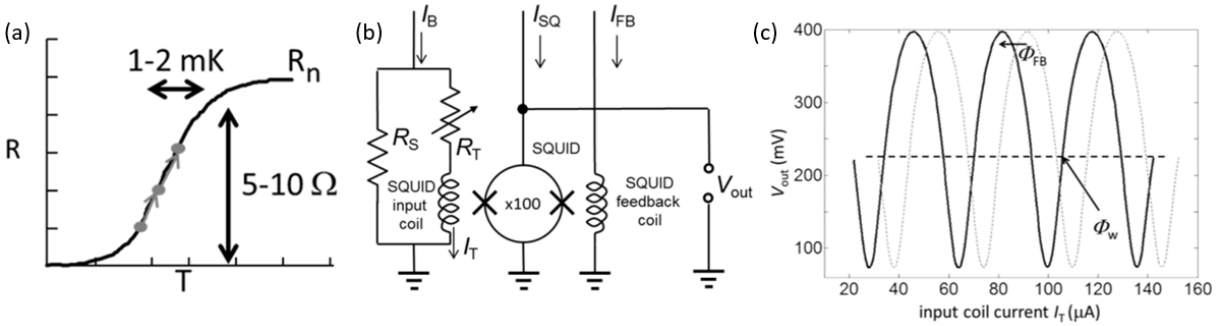



FIGURE 6.1: Figure taken from Ref. [189]. (a) Resistivity of TES as temperature increases near the superconducting phase transition. (b) Readout circuit with SQUID magnetometer for detecting changes in resistance. (c) SQUID fringe tuned such that all SQUID signals are in-phase and constructively interfere.

in Fig. 6.1(c), which is achieved by tuning the SQUID bias setting on the cryostat. Tuning the 'flux bias' (FB) allows one to set the baseline location along the SQUID fringe, which should be set to the steepest part of the slope to achieve maximum responsivity. A full biasing procedure can be found in the externally hosted PowerPoint tutorial[1]. Extensive details on the TES detectors, their design, and several selected applications can be found in Ref. [189]. The detectors used for this work were obtained from the Sae Woo Nam group at the National Institute of Standards and Technology (NIST).

### 6.2.1 Basic TES response

Once the TES has been cooled to base temperature and has been properly biased, it is ready to receive signal and detect photons. It should be emphasized that each TES is a different detector, so the bias settings and even critical temperature vary slightly for the devices. The parameters can all be tweaked to elicit various responses, but I have found that the optimal base temperature to achieve maximum sensitivity for several TES channels at once is around 95-100 mK. The TES channels will begin to respond to external signal around 80-85 mK, but it becomes more difficult to resolve differences between vacuum and single-photon events. Similarly, on the high end, some detectors still work all the way up to 115-120 mK. However, at this temperature, very few photons can be resolved before the detectors saturate. Thus the operating temperature

[1] https://github.com/me3nq/Thesis_docs

should remain just below or around 100 mK, but some experimentation is encouraged when using a new or less frequently used detection channel.

An important point to note is that the TES cannot easily respond to continuous-wave (CW) laser input. In addition to having a very slow cool-down time (1-20 $\mu$s, depending on biasing), there is also no 'reset time' in that the TES can always pick up an additional signal. When CW light is incident on the detector, there is always some probability to detect an event at any given time, and if more than one event occurs within a given time window, this can lead to pile-ups that complicate photon-number discrimination [114, 263]. In spite of these difficulties, low-flux CW light is still compatible with the TES, especially in the case of heralding experiments where pile-ups are quite rare [114, 264]. To avoid the pile-up problem, we implemented a controllable pulsed scheme (see Ch. 8) which allows for tuning the time between pulses to allow the TES to entirely cool to base temperature before the next laser pulse arrives.

When a pulsed signal is sent to the detector, the TES samples the probability distribution in the Fock basis; that is, it implements a phase insensitive projector given by Eq. 6.4. An example of this signal is shown in the top panel of Fig. 6.2, where an external 14-bit digitizer (AlazarTech ATS9440) is used to acquire the output signal. After collecting many pulses, the height obtained by each peak can be used to generate a histogram, as shown in the bottom panel of the figure. Once the histogram is created, it is clear that there are distinct values of heights that the peaks attain and these height values fit nicely into bins. Because we are working at a single wavelength, each discrete bin corresponds to the number of quanta of energy absorbed by the TES during a pulse window, which is exactly the number of photons in that pulse. In this case, for a coherent state input signal, binning the peak heights and constructing a probability distribution from the histogram yields the Poissonian distribution shown in the inset of Fig. 6.2.

Although the PNR capability with peak heights alone is impressive, the top photon resolvability saturates around 8 or 9 photons, depending on the TES channel used. However, it was recently demonstrated that by post-processing the data in a different way to calculate peak areas, up to 16 photons could be resolved per detection channel [165]. We further develop this method in collaboration with Jefferson Lab and have been able to demonstrate PNR capabilities beyond 30 photons. This requires specialized external electronics and hardware to properly sample the output signal.

The motivation for switching to area measurements is clearly seen by looking

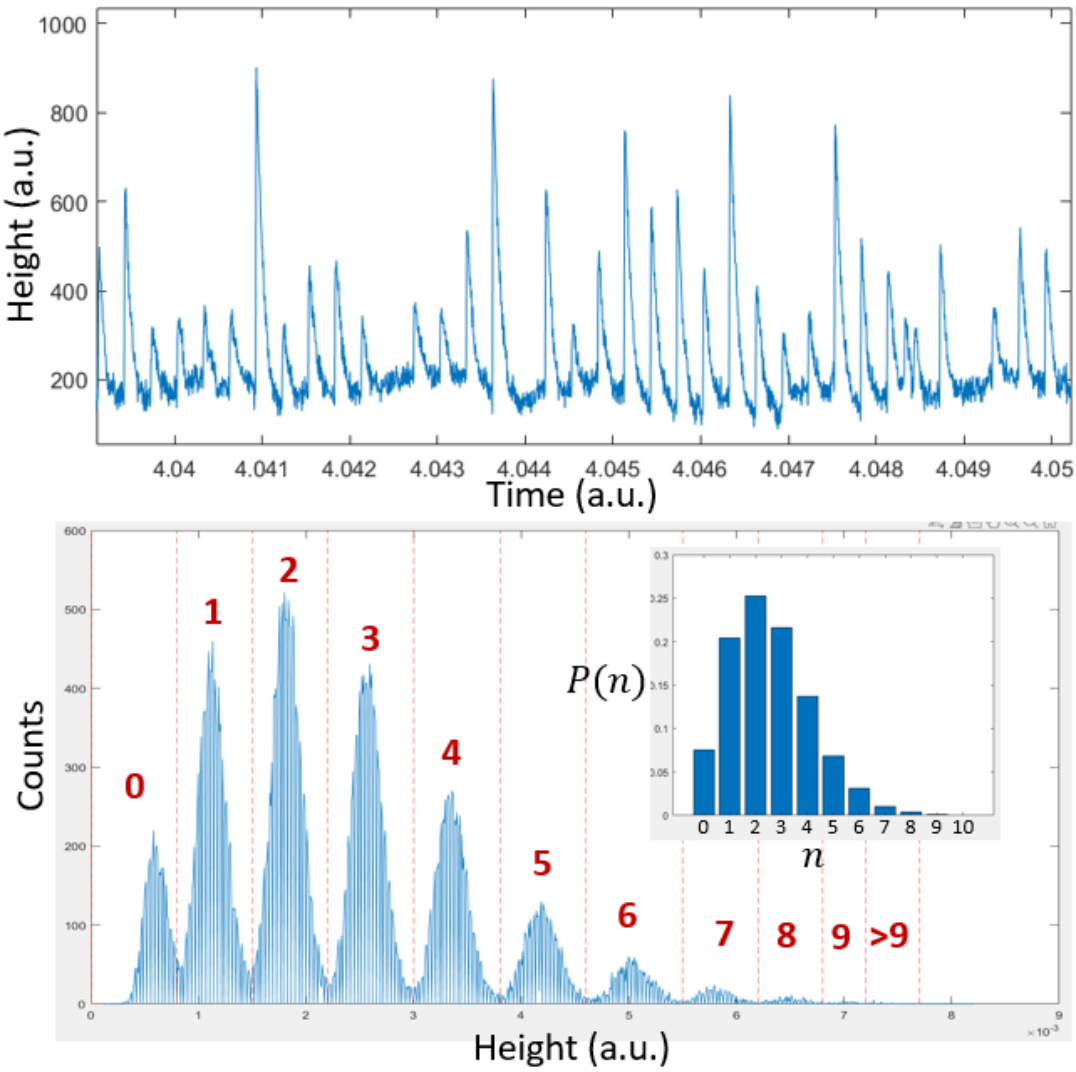



FIGURE 6.2: Raw TES signal obtained from a pulsed coherent state. The signal heights can be converted to a histogram, which can then be binned to determine photon number. The inset depicts the Poissonian probability distribution for the input coherent state.

at the persistence measurement in Fig. 6.3. This figure shows an overlay of many laser pulses (red) as they occur after a trigger signal (blue). Although the distinction between the height of events washes out around 8 photons, many more discrete signal 'tracks' are clearly visible. This image beautifully captures the quantized nature of light — events clearly follow well-defined paths and no signal falls between event groupings!

## 6.3 TES Data Acquisition

In order to calculate the area of each peak and improve the PNR ability of each TES, it is necessary to quickly sample the output signal to achieve high resolution. From there, the area could be computed in post-processing, but this must be done for potentially hundreds of thousands or millions of separate events, and is thus computationally intensive. To avoid this, Jefferson Lab provided us with a fast, Ethernet-based flash analog-to-digital converter (EFADC) that would acquire data and calculate pulse parameters on the fly.

### 6.3.1 Hardware and firmware

The EFADC device was custom-built to collect and process TES signals for up to 8 channels. The device is based on a field-programmable gate array (FPGA) which samples a signal with 12-bit resolution at a rate of 250 MHz (once every 4 ns). The internal memory and processing speed allows the device to collect up to 32 $\mu$s worth of signal points, perform rudimentary calculations on the data to determine key parameters, and transfer the calculated parameters to a hard disk all before the next signal pulse arrives. In this way, the EFADC can collect and save up to $10^7$ pulses at a time. The limitation is in fact not on the hardware, but on a file size limit within Windows and the C-code used to operate the hardware.

The interface between the TES, EFADC, and lab computer is shown in Fig. 6.4. For each pulse, the EFADC begins to look for data when a trigger signal between $2-4$ V is received. This trigger should be the same signal used to pulse the laser, and in this implementation is the signal used to quickly switch the acousto-optic modulator on and off. The EFADC has a 50 Ohm input impedance and digitizes an analog signal in the range of 0-500 mV. Note that if the input signal exceeds 500 mV, the EFADC will simply report the highest value while the input remains above 500 mV. This can be

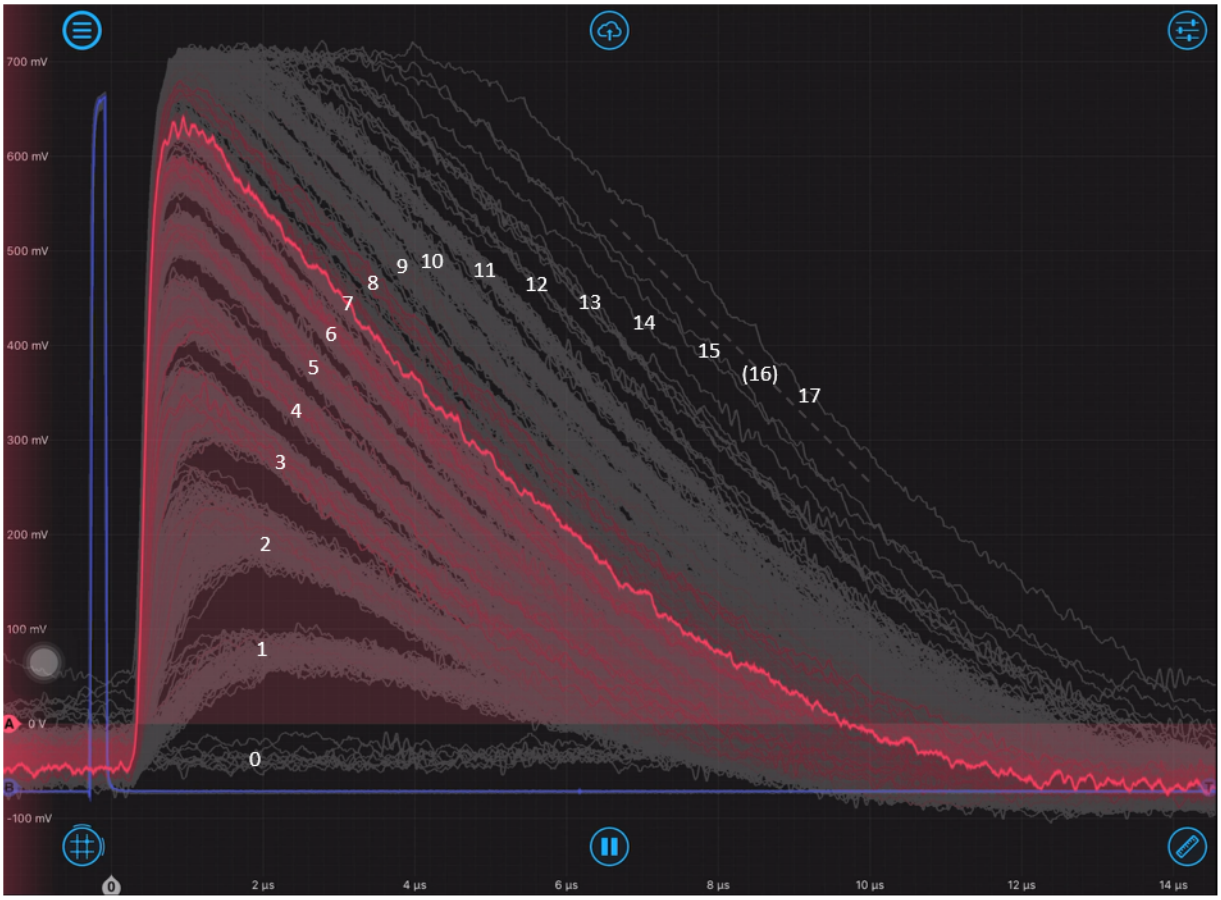



FIGURE 6.3: A coherent state laser is pulsed and sent to the TES. The pulse is triggered by the signal in blue, and the TES response after external amplification is given in red, where the signal is allowed to build up over time. The photon number differentiation based on height alone becomes poor beyond about 8 photons, but higher number photon incidents clearly follow well-defined signal traces. This allows for PNR of many more photons if peak area is used. Here, 0 to 17 photons are clearly resolvable (a 16 photon event did not occur).

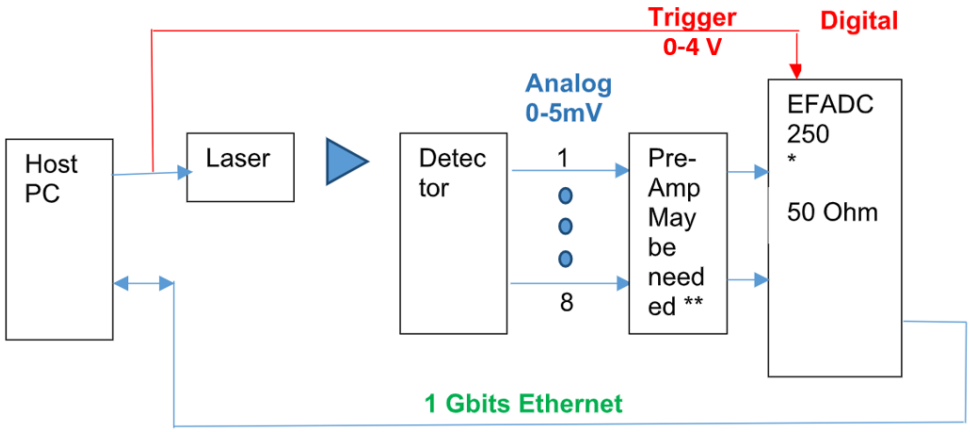



FIGURE 6.4: Diagram of data collection. An external trigger is synced to pulse the laser and prepare the EFADC to look for TES output pulse data. Each pulse is temporarily stored while calculations are performed on the hardware, then key parameters are permanently stored.

quite useful when characterizing PNR values, as the tip of the peak has a small impact on the area, so in actuality, tuning the input signal such that it may occasionally exceed the saturation threshold is not a bad thing. This point will be covered in more detail in the next section. Additionally, the EFADC has a user-programmable DAC offset value. This must be set with some trial and error for each channel to ensure that the baseline noise fluctuates around the zero point on the digitized data. Furthermore, the DAC inverts the analog signal before digitization, thus the true input signal should consist of peaks spiking negative in the range of 0 to -500 mV. The programmable DAC value allows for compensating for small offset voltages on the input signal.

Once the EFADC receives a signal, the internal firmware is programmed to calculate and output several important parameters, including the peak height, area, trigger timestamp, timestamp of peak with respect to the trigger, and several other parameters. The calculations performed are also dependent on several user input parameters. Full operational details of the EFADC data acquisition system include format of output data can be found in external documentation on the associated firmware[2], but the main operation can be summarized by examining Fig. 6.5

The user first programs trigger thresholds (TET) such that the device will only integrate the signal if it first crosses the upper threshold, TETHI, and it will continue to calculate the pulse parameters until either the signal drops below the lower threshold, TETLO, or the signal continues beyond the maximum duration specified by the user.

[2]https://github.com/me3nq/Thesis_docs

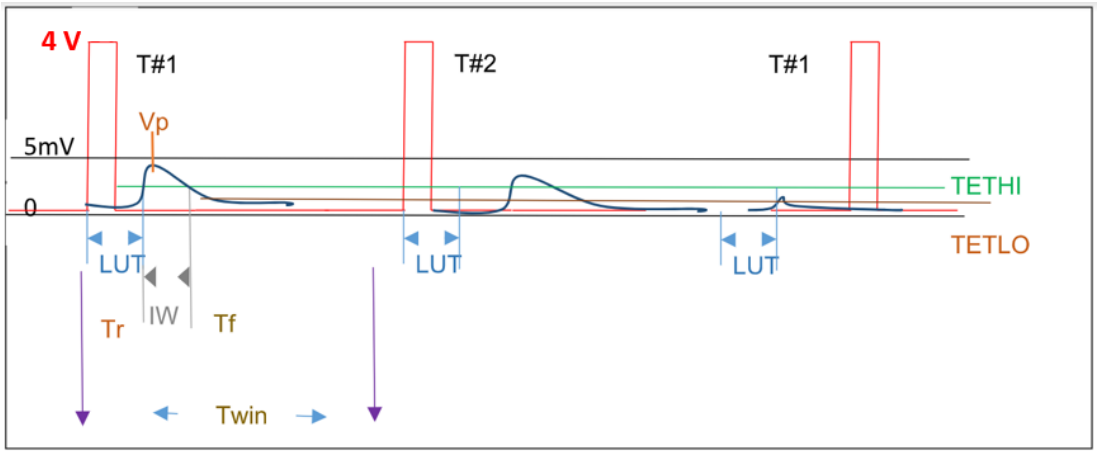



FIGURE 6.5: A basic depiction of the response of the EFADC firmware. For an incoming trigger signal (red), the firmware decides whether to acquire the signal (blue) based on user-defined controls. Further details can be found in the main text.

After the device receives a trigger, if the signal rises above TETHI for a user-specified number of ADC samples (NSAT), and if this occurs within a certain number of samples from the trigger define by the lookup time (LUT), the device will behave as follows:

- A Y user-programmable number of samples (PedNumOfAdcSamp) starting at the Trigger time will be averaged to be reported as a Pedestal for this trigger. Y is restricted to be 1,2,4, or 8. If any of these ADC samples is above TETHI, the Pedestal quality bit will set to TRUE. If an ADC sample within PedNumOfAdcSamp is greater than the user-programmable value (MAXPED), MaxPedDetect bit is set to TRUE.
- If a pulse started within PedNumOfAdcSamp, PulseInPed is set to TRUE
- The samples (including the Pedestal samples) in that channel will be summed until a sample drops below TETLO (the final sample below TETLO). In other words, the samples within the integration window (IW) will be summed. If an ADC sample is either above or below the limits set by the input signal, then either an Overflow (Ov) or underflow (Uv) bit will be set to TRUE and recorded in the output file. The sum is terminated either when the EFADC receives a new trigger or if the value for Twin is reached. Twin is a user-programmable quantity that determines how long, in ADC samples, the sum should continue as measured from the start of the trigger.

- Twin must be less than the minimum time between pulses. If Twin is greater than the time between pulses, the second pulse will not be recognized, i.e., if more than one pulse occurs within Twin, only the parameters of the first pulse will be processed.
- The time that the 1st sample is above TETHI (Tr) and the time that the 1st sample falls below TETLO (Tf) will be reported. This time is relative to the rising edge of the trigger signal.
- The peak ADC sample (Vp) and the time (Tp) will be reported.
- If after a Trigger signal, no NSAT number of samples is larger than TETHI within the specified LUT, only the time of the trigger is stored in the output file.
- The difference between TETHI and TETLO is essentially hysteresis. As such the user should not set TETHI and TETLO to the same value. The difference should be set to be equal to or greater than the baseline noise of the signals.

Of the three triggers shown in Fig. 6.5, only T#1 has data to report. T#2 does not have data to report because samples cross TETHI after LUT time. Similarly, T#3 does not have data to report because no or not enough sample(s) cross TETHI. It should be noted that all quantities, whether it be a measure of voltage or time, are stored in terms of ADC samples. For example, if one wishes to specify a time of 800 ns for LUT, the user would set LUT $= 200$ since the device performs one sample every 4 ns.

### 6.3.2 External amplification

The magnitude of the analog output signal from the TES is dependent on several factors including bias and temperature in addition to the number of photons actually absorbed. Even for large PNR signals, this voltage remains on the order of 0-25 mV. Because the EFADC digitizes a 0-500 mV range with 12-bit accuracy, the best peak differentiation will occur if the input signal is designed to occupy the majority of the available voltage input range. In order to do this, I designed an amplifier circuit based on the ultra-low-noise OPA818 op-amp.

In order to design the circuit to properly amplify the photon peaks and reduce noise, it is useful to first examine the frequency-domain spectrum of the signal. This is shown in Fig. 6.6. From this plot, we see that the signal drops off quickly beyond

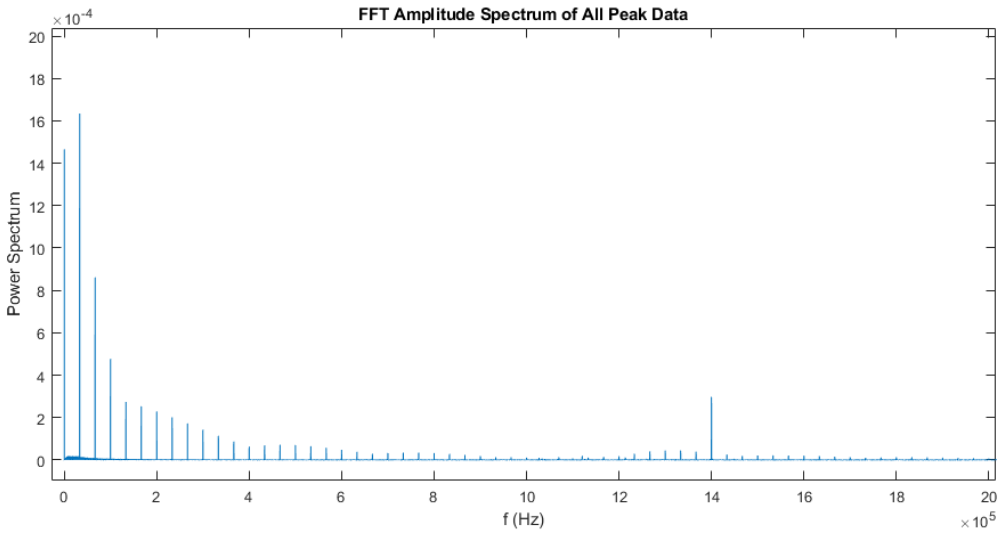



FIGURE 6.6: FFT data of TES signal.

1.4 MHz, but there is a substantial contribution at low frequency, even down to 1 kHz. This contribution can be attributed to the slow cooling tail that trails after each event. This portion of the signal will be important to the area calculation so should not be excluded. However, there is a DC offset that one would like to suppress, so some high-pass filtering would also be helpful.

To test possible filtering frequencies, low and high-pass filtering was applied to the data in post-processing at several test frequencies as shown in Fig. 6.7. Both panels of this figure show the original data (blue) compared to the data filtered by the given frequency (red). It can be seen that the test high-pass filter at 40 kHz in Fig. 6.7(a) performs poorly as several features are qualitatively quite different; this shows that the low-frequency peaks in the power spectrum are components of the signal, and important to be kept.

The high-pass filter in Fig. 6.7(b), however, seems to perform quite well. The filtered data gives similar quality peaks but smooths the fast noise. This indicates that the 1.4 MHz peak on the power spectrum may be noise unrelated to the signal. However, it was eventually decided that this 1.4 MHz component may also relate to the initial peak height, as the rise time can be quite fast depending on the bias, and 1 MHz Fourier components are expected for a signal with a rise time of 1 $\mu$s.

The final circuit design is shown in Fig. 6.8, where the top panel shows the simplified circuit diagram and the picture shows the modified demo board for the op-amp

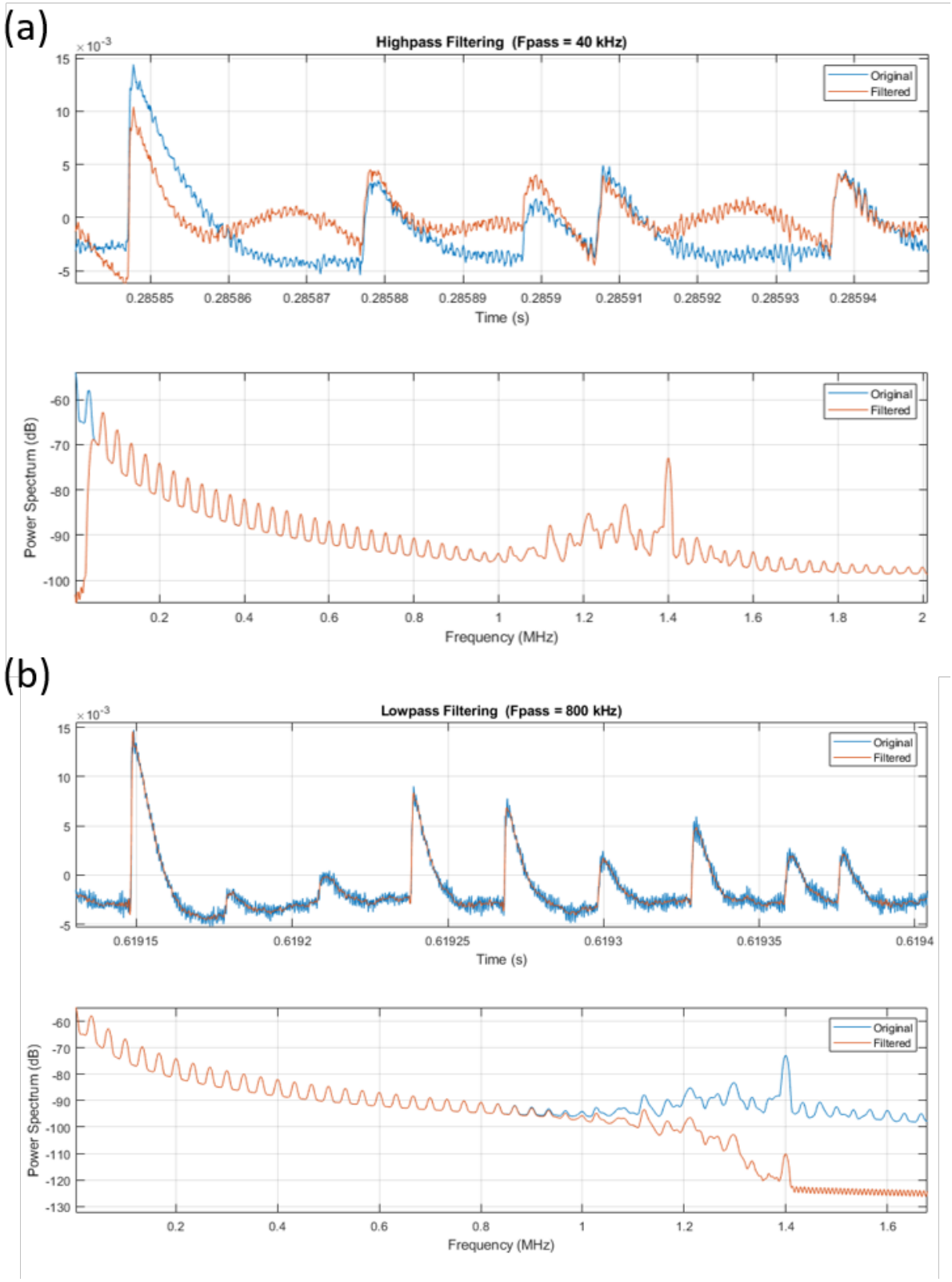



FIGURE 6.7: TES signal (blue) compared to a filtered signal (red) for a high-pass filter applied at 40 kHz (a) and a low-pass filter applied at 800 kHz (b). The top panel of each shows the comparison of the signal in time while the bottom compares the power spectrum in the frequency domain.

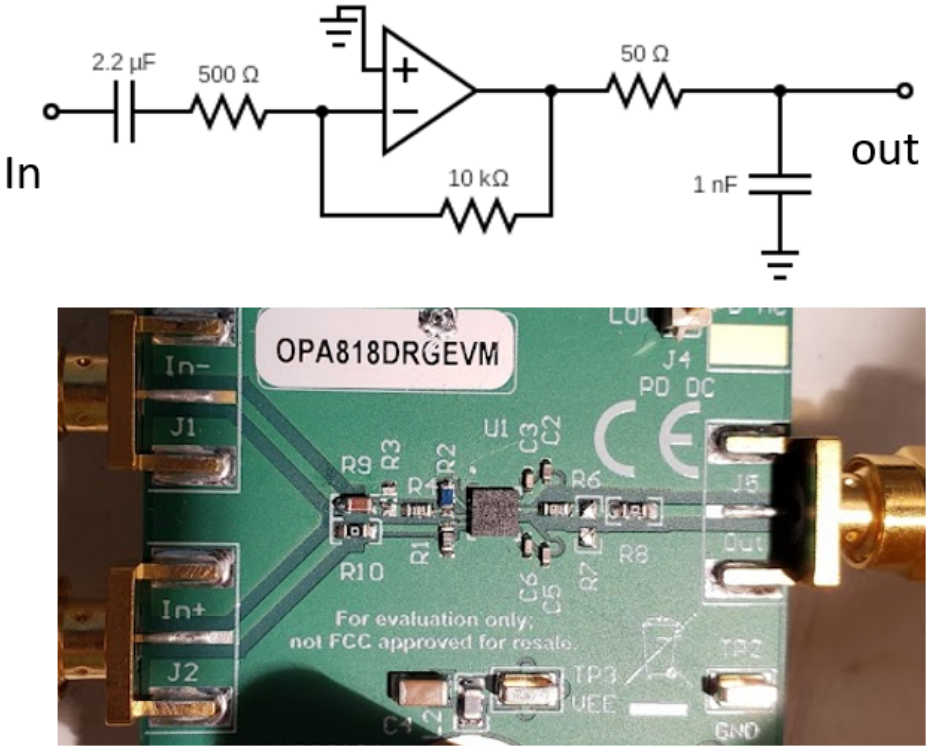



FIGURE 6.8: Circuit used to amplify the TES output signal into the 0-500 mV range. The circuit is an active high-pass filter with corner of 145 Hz, gain of 20, and low-pass filter at 3.2 MHz. The modified OPA818 demo board is also shown.

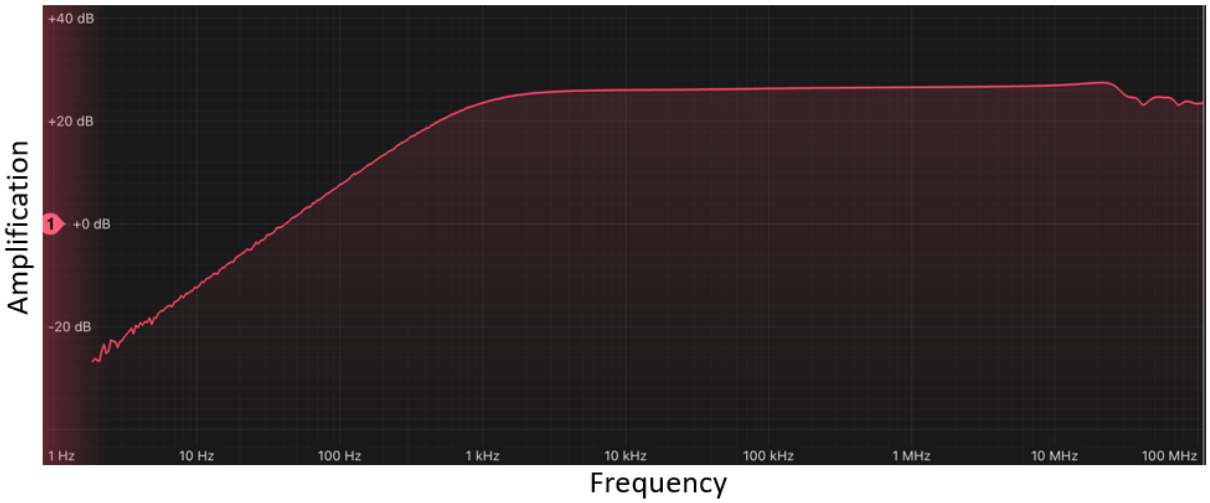



FIGURE 6.9: Frequency response of the OPA818 amplifier circuit.

use. First, an active high-pass filter with a cutoff frequency of 145 Hz is used to remove the DC component of the signal and provide a gain of 20. At the end of the circuit, a low-pass filter with cutoff frequency of 3.2 MHz is used to clean any residual noise from the op-amp or other high-frequency noise from the TES signal.

The demo board can be turned into the amplifier described by replacing R2 with a 10 k$\Omega$ resistor, removing R3 entirely, replacing R4 with a 500 $\Omega$ resistor, adding a 1 nF capacitor at R7 (not shown in picture), and replacing R9 with a 2.2 $\mu$F capacitor. After these modifications, the frequency response was measured as shown in Fig. 6.9

## 6.4 TES Efficiency Calibration

Transition-edge sensors have managed to reach up to 98% quantum efficiency [188], but it is important to characterize the precise response of our detection system. Although our detectors were optimized for our system wavelength at 1064 nm, the quantum efficiency (QE) is still imperfect, although it remains at or above 90% for the detectors measured.

Before our system was optimized to count up to 30 photons, each TES channel

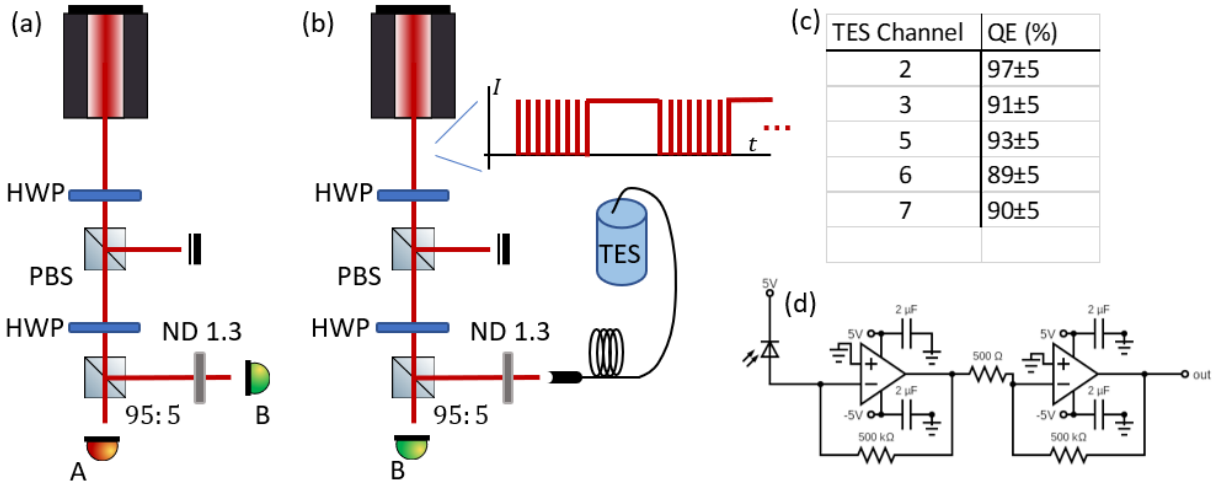
could handle no more than several photons at a time before saturating. At our wavelength (1064 nm), each photon has energy

$$\hbar\omega = \frac{2\pi\hbar c}{\lambda} = 1.868 \times 10^{-19} J. \tag{6.12}$$

If each pulse is 100 ns wide and I wish to send on average 1 photon per pulse to the TES, then the power in a pulse is one the order of 2 pW. This power is considerably smaller than what can be resolved by commercial power meters, and in the absence of highly accurate fiber attenuators as used by NIST [124], a more complicated calibration scheme is necessary. One could, of course, attempt to accurately calibrate several free-space attenuators and use these in succession, but this method will lead to large error in calculated efficiency, as each attenuated placed will ever-so-slightly alter the beam path, and thus change the alignment on the fiber coupled to the TES. Another option would be to use several accurately calibrated beamsplitters, but each beamsplitter can only reduce beam power by about three orders of magnitude. In order to bring the power down from the several hundred $\mu W$ necessary to achieve an accurate power reading on a commercial sensor to the TES level, at least three or four beamsplitters with associated waveplates would be required, which becomes cumbersome.

Instead, an alternate scheme was used as depicted in Fig. 6.10. First, I constructed a high-amplification photodetection circuit which consisted of a photodiode connected to two back-to-back inverting amplifiers. Although these op-amps were used to amplify a DC signal, utilizing two amplifiers in series with slightly smaller gains as opposed to a single, high-gain amplifier allows for a larger amplification bandwidth. The overall gain for the circuit was on the order of $10^6$, which gave it a low-power sensitivity threshold at approximately 200 pW.

The scheme in Fig. 6.10(a) was used to accurately calibrate the the home-built detector, depicted as detector 'B', with respect to a Scientech pyroelectric calorimeter as detector 'A'. In this scheme, the closest beamsplitter to the detector is set to a static 5 : 95 splitting ratio, and a fixed neutral density (ND) filter of optical density (OD) 1.3 is placed in front of the detector under test. Next, to avoid any movement of detectors or other optics that may slightly cause beam path deviation and additional error, a single half-wave plate (HWP) is rotated further back in the beam path before a second beamsplitter to change the power of the whole system. Power measurements at detector 'A' and voltage output levels at detector 'B' are measured for a variety of input powers. It is expected that far from either the minimum or maximum detectable value of the




| TES Channel | QE (%) |
|---|---|
| 2 | 97±5 |
| 3 | 91±5 |
| 5 | 93±5 |
| 6 | 89±5 |
| 7 | 90±5 |

FIGURE 6.10: Protocol to measure TES efficiency. (a) Set-up used to to calibrate the home-built low-light detector, where A is the commercial power meter and B is the detector to be calibrated. (b) The newly calibrated detector B is now used in place of the commercial power meter, and the TES takes the place of the detector to be calibrated. The power of light can be lowered by several orders of magnitude to a level the TES can handle. (c) Measured efficiency of several TES channels. (d) Circuit for the home-built low-light photodiode.

photodiode circuit, the response of detector 'B' will be linear with increasing power. For a voltage $V_m$, measured by the photodiode circuit, the power can be written as

$$P_B = \gamma \left( V_m - V_{offset} \right), \tag{6.13}$$

where $V_{offset}$ is the offset due to dark noise or amplified residual current. By taking several power measurements and fitting the data to a line, values for $\gamma$ and $V_{offset}$ can be determined. In this instance, for $V_m$ measured in Volts, $\gamma = 63 \pm 1$ $\mu$A and $V_{offset} = 0.5 \pm 0.2$ mV. It is important to note, however, that these values will be useless if used again. Both the offset and $\gamma$ are highly sensitive to environmental conditions and exact lens placement to focus the light on the photodiode. These values are listed here only for order-of-magnitude reference and should not be used for future calibrations.

Once the home-built detector was calibrated, the set-up in Fig. 6.10(b) was used to measure the TES efficiency. This set-up is nearly identical to the first using the newly built detector as a reference, but with the difference that the beam was alternated between being pulsed and held CW. Additionally, the ND filter was recalibrated with the

fiber coupling efficiency included. Thus the ND and fiber were considered as a single system that together just acted to attenuate the beam. During the CW phase, the home-built detector could register a large enough power ($\sim$ nW) within the regime it was calibrated, and then the TES would acquire signal during the pulse phase, which consisted of 100 ns pulsed separated by 30 $\mu$s. Although the power per pulse should be $\frac{1}{300}$ of the CW power, the response of the AOM/pulse generator may mean that the pulse width is not precisely 100 ns (see Ch. 8 for details). To check, one can simply measure the average integrated power of a pulse train with a commercial detector at higher power to determine what fraction of the CW power is in each pulse.

Now, to determine the detector efficiency, the mean photon number over the course of $10^5$ pulses was collected over a small range of powers, with mean photon number varying from $0-4$, but ensuring that the detection probability of photon number above 8 was low. This mean photon number can be related back to the average power per pulse, and thus compared with the power measured by the calibrated home-built detector. Once this is done, the set-up can be used for several detectors sequentially, and the QE can be extracted as shown in Fig. 6.10(d). Only five detectors were measured, as detectors 1, 4, and 8 were offline at the time of calibration.

The main source of error is the 3% absolute calibration error on the commercial detector used to begin the process. The remainder arises from small propagating errors in the fit of $\gamma$ and $V_{offset}$ for the photodiode circuit and ND filter calibration. Together, this amounts to about 5% error on the TES QE. However, each detector measured achieved a QE above 90% to within measurement error.

## 6.5 Counting 30 Photons

Once the TES signal is properly amplified, filtered, and digitized by the EFADC, the results still need a series of post-processing algorithms before they are ready to be turned into photon number distributions. An extended explanation of these algorithms is given in Appx. F, but I will motivate some of the key components here.

First, it is insufficient simply to use the raw heights or areas as acquired by the EFADC. This is due to the fact that the reference point for all signals is the zero-point on the DAC instead of the signal baseline. If the actual signal has a small offset such that the signal zero is above the DAC zero, then each measured area will have an additional,

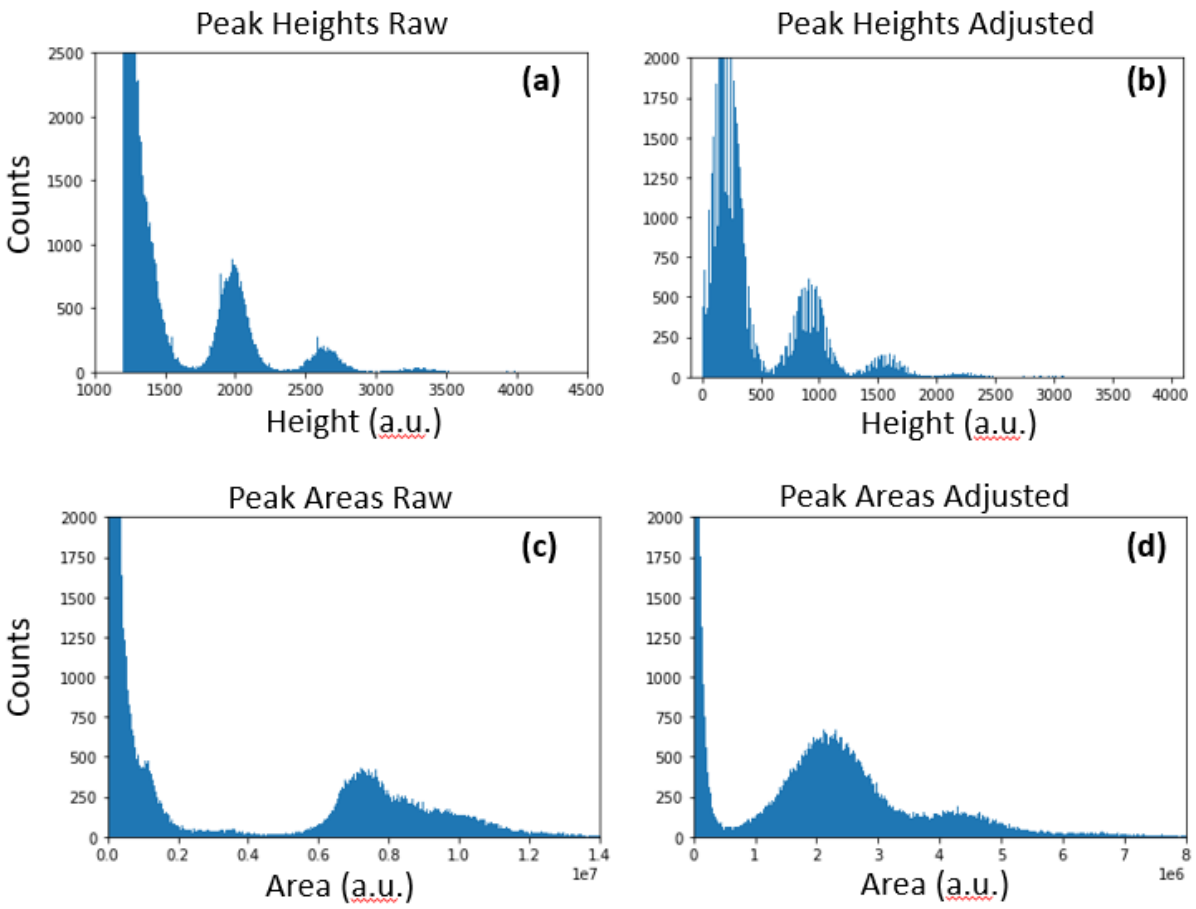



FIGURE 6.11: Using raw heights and areas to generate histograms may produce errors if DAC settings are imperfect. Calculating the adjusted heights and areas by subtracting off the baseline in post-processing improves the photon-number discrimination.

unwanted rectangular area portion added to it. Because the baseline occasionally fluctuates at low frequencies (the He compressor imparts 60 Hz environmental noise), this additional area can negatively impact the peak separation in the histogram.

To fix this problem, the pedestal value (baseline at the trigger) can be used to set a new baseline, and the area of the rectangle formed from Tr to Tf on the time axis and zero to TETHI on the voltage axis can be subtracted off. The new area is referred to as the 'relative area', as it is the measured area relative to the signal baseline instead of the DAC baseline. A similar process can be used on the signal heights, but this has a smaller effect. The results of applying this correction can be seen in Fig. 6.11, where the peak heights give clear PNR distinction. The areas for the exact same data set in Fig. 6.11(c), however, show very poor signal distinction. Applying the relative area

corrections offers some improvement as seen in Fig. 6.11(d). It should be noted that even with the correction, the peak distinction with areas for this data set is not optimal. This likely means that the TETHI or TETLO setting was chosen sub-optimally, such that some portions of the correct area were prematurely cut off.

As yet another reason to use relative areas as a metric instead of raw areas, I have found that the signal offset from zero is also dependent on the average intensity of the state being measured by the TES. This may at first seem highly problematic, as the detector response should be independent of the input state. However, on further thought, this is not too surprising that the detector behaves this way. Because there is already some DC offset, I included a high-pass filter to remove low frequencies below $\sim$150 Hz. Recall, however, that the TES signal saturates beyond about 8 photons, meaning the signal stays at maximum, and the remainder of the PNR capability comes from the area, which has its origins in how long the signal takes to drop back down. If the input signal is large enough such that photon numbers larger than 8 occur frequently, then the TES will spend a majority of the time saturated, and thus the maximum signal starts to have a DC-like quality in the spectral decomposition of the signal and will be suppressed by the filter. One way to correct this problem is to increase the time between pulses. This was the method used for this work whenever high energy inputs were used; even though the cool-down time remained below 30 $\mu$s, a temporal pulse spacing of 80 $\mu$s was used. Even so, there remains a small effect of slight signal offsets that vary with average input power which is avoidable using the relative area calculations.

Fig. 6.11 also illustrates another important quality of the histogramming process: peak height can be a better differentiator of photon-number binning at low photon counts than area. This is not always the case, but heights are more robust to imperfections in DAC settings as TETLO does not come into play, and offset values are less important. Thus, one can use the peak heights at low counts as a way to verify that the area measurements are correctly calibrated. This way, when larger photon number measurements are made that cannot rely on height alone, we can rest assured that the user settings optimize the peak differentiation given by areas.

Once all settings are verified and the relative areas are calculated, one can produce a histogram of the areas as shown in Fig. 6.12(a). This figure plots the log-scale histogrammed areas for $10^7$ pulsed events where the coherent state amplitude was tuned to be as large as possible without saturating in order to test the limits of the TES. This histogram can now be sectioned based on the grouping of events, and each section can

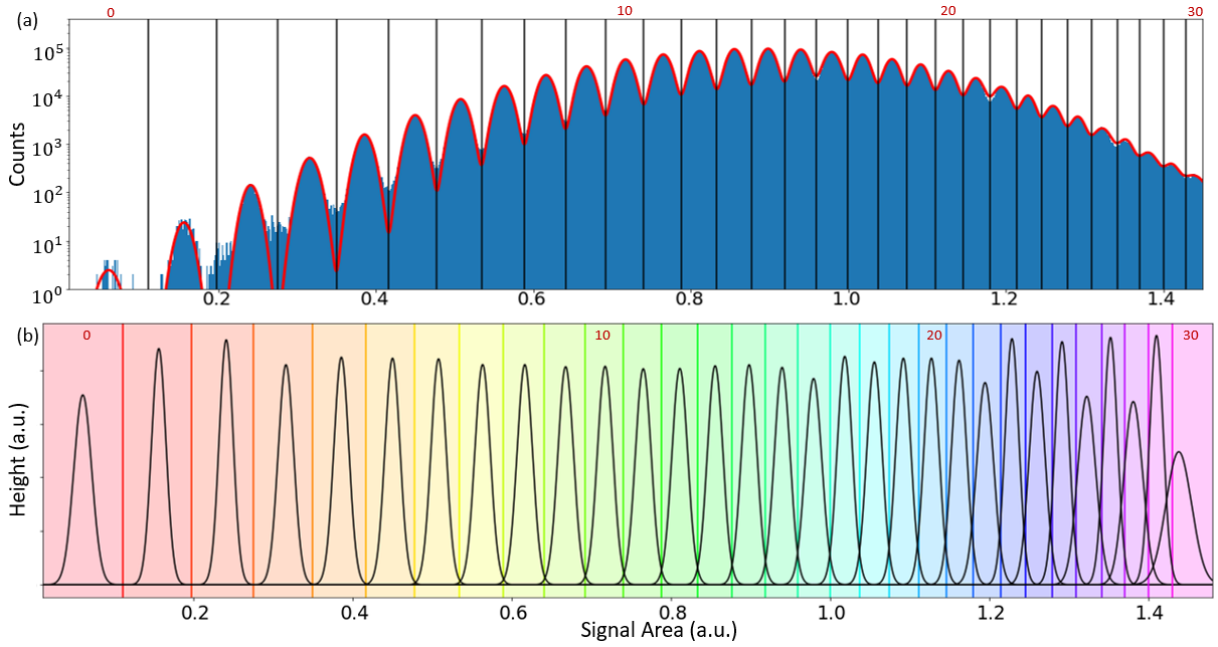



FIGURE 6.12: (a) Histogram of integrated peak areas for $10^7$ measured events. The distribution is fit to a sum of Gaussian peaks, and peak intersections are used to determine photon-number binning. (b) Individual normalized Gaussian fits. The overlap of neighboring Gaussians lead to miscounting errors.

be assigned a photon number.

The spread in areas in each bin is caused by Gaussian noise and thermal fluctuations, so each of these can be fitted to a Gaussian. Thus, the whole histogram is fitted to a sum of Gaussian curves, and the intersection point of neighboring Gaussians defines the bins used to determine photon number [165]. It is important to ensure that the histogram has enough smaller area bins initially so that the distribution approaches a continuous curve that can be well-fitted to the Gaussian sum. When attempting to characterize the TES for large PNR, fitting the entire histogram to many Gaussians all at once can be difficult depending on the fitting algorithm used (least sum of squares, for instance) due to the several orders-of-magnitude difference in counts for different area occurrences. To avoid this, one can section the histogram and fit each individual piece to several Gaussian curves. This process is described and implemented in Appx. F.

The process of fitting the histogram to a series of Gaussian peaks and finding the intersections is the most accurate method of determining the photon number bins, but it is computationally intensive for large data sets. Additionally, this characterization must

be redone for each channel and for every data set as the area calibrations will vary with TES temperature, bias settings, and EFADC settings. Even if all parameters are kept unchanged, re-cooling the TES will invariably introduce minor changes that shift the calibration. Because of the time-consuming nature of this method, a 'manual' method can occasionally be useful, where a user simply zooms in on the region between two peaks and chooses the minimum histogram value between the peaks. Surprisingly, this method is nearly indistinguishable from a numeric fit as far as accuracy is concerned, and is generally faster than waiting for the Gaussian fit to find a valid solution.

As a good rule of thumb, this TES calibration should always be performed with a pulsed coherent state, so the resultant probability distribution can be checked with what is expected. Once the calibration is completed and the measured distribution is acquired, one can determine the mean photon number of the measured data, which is equivalent to $|\alpha|^2$ for the input coherent state $|\alpha\rangle$. One can then calculate the sum-squared error (SSE) as

$$\sum_n \left( P_{meas}(n) - \frac{e^{-|\alpha|^2}|\alpha|^{2n}}{n!} \right)^2, \tag{6.14}$$

where $P_{meas}(n)$ is the measured probability. The SSE for this data set is on the order of $\sim 10^{-6}$ so the measured distribution agrees quite well with the Poissonian we would expect based on the mean. The SEE can be reduced slightly more if we consider amplitude fluctuations as well, where we imagine we have actually measured a mixture of coherent states,

$$\rho = \frac{1}{2\pi\sigma} \int d\beta e^{-\frac{(\beta-\alpha)^2}{2\sigma}} |\beta\rangle \langle\beta|, \tag{6.15}$$

where $|\alpha|^2$ is the mean photon number of the mixture and $\sigma$ is a small parameter describing the fluctuations. Over the course of $10^7$ measurements, these fluctuations are small but non-zero, and we find that $\sigma \approx 0.04$ for the data set shown in Fig. 6.12. This is also considered further in Appx. F.

### 6.5.1 Bin overlap error

Due to the fact that there is a distribution of peak areas for each photon number bin, there is some error in the discrimination of photon counting. This manifests in the fact that each Gaussian overlaps with its neighbor, so after calibrating the detector, any given photon that is assigned a bin, $n$, has some probability of actually being either an

| n | Error (%) | n | Error (%) |
|---|---|---|---|
| 0 | 0.00075% | 16 | 5.23630% |
| 1 | 0.00028% | 17 | 4.43508% |
| 2 | 0.00491% | 18 | 5.29261% |
| 3 | 0.02702% | 19 | 5.75297% |
| 4 | 0.06191% | 20 | 6.43787% |
| 5 | 0.19571% | 21 | 7.47639% |
| 6 | 0.35943% | 22 | 11.60825% |
| 7 | 0.57692% | 23 | 7.35573% |
| 8 | 0.77740% | 24 | 10.94260% |
| 9 | 1.03799% | 25 | 9.35130% |
| 10 | 1.33954% | 26 | 16.75042% |
| 11 | 1.79497% | 27 | 12.46927% |
| 12 | 2.19522% | 28 | 22.38296% |
| 13 | 2.54734% | 29 | 13.94831% |
| 14 | 3.03144% | 30 | 37.69627% |
| 15 | 3.99945% | | |

FIGURE 6.13: Probability that an event assigned to be an $n$ photon peak is actually either an $n-1$ or $n+1$ peak for the calibration data in Fig. 6.12

$n-1$ or $n+1$ photon event. Because the Gaussians are already fairly well separated, errors beyond nearest-neighbor overlap are negligible.

Errors are shown in the Table depicted in Fig. 6.13, where the error shown is the probability that an event was incorrectly assigned to an $n$ photon event. Because the calibration itself is different for various external parameters, these binning errors are again only valid for the specific data set in Fig. 6.12 and should merely act as a guide. These values are calculated in detail with the documented code in Appx. G.

The versatility of the bias controls actually allows for the user to optimize certain detection regimes. For instance, if the FB and TES biases are set such that the cooling time is slow and the events peaks are as large and wide as possible, then the discrimination at higher photon number is better, and one can count all the way out to 30 photons. However, this also means that lower and intermediate regimes are subject to underdamped oscillations and may have slightly larger binning errors at low photon counts. If one instead tweaks the TES bias to shorten the peak width, then not only will the TES be ready to measure the subsequent pulse more quickly, but the distinction between lower photon-number events can be improved. However, this does come at the cost of reducing the maximum number of resolvable photons.

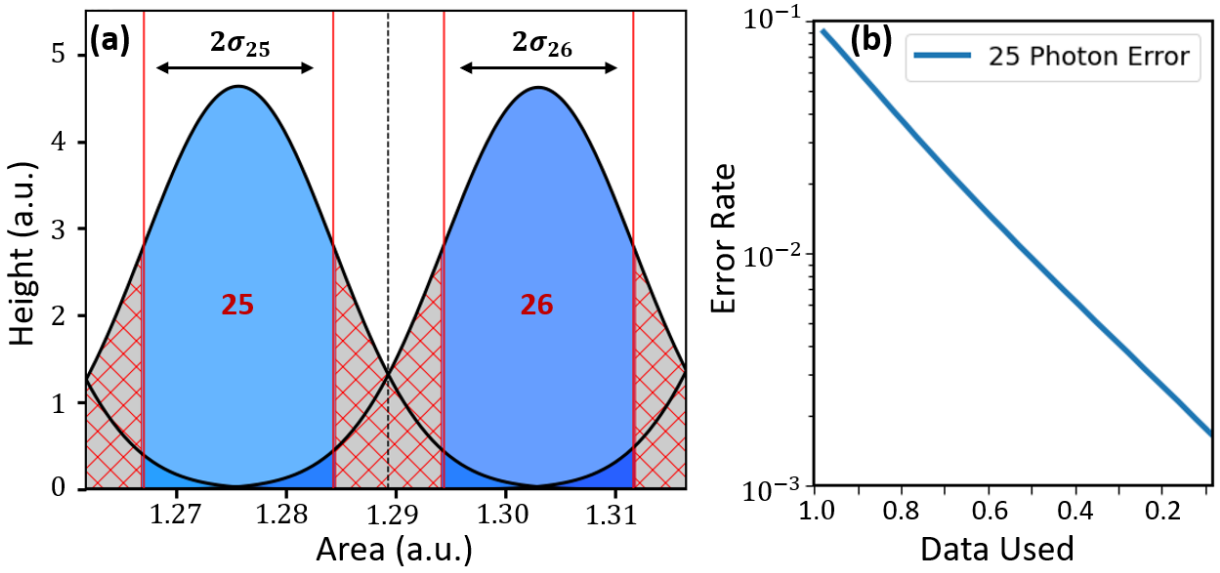



FIGURE 6.14: (a) By defining regions of uncertainty near the bin edges with respect to each $n$-th Gaussian width, $\sigma_n$, and ignoring measured areas in this region, the photon-resolving accuracy can be improved. (b) Example error rate of incorrectly counting a true 25 photon event as the proportion of data used is decreased. Using half of the data set reduces error by an order of magnitude.

**Bin error mitigation**

These errors in bin calibration can be almost entirely eliminated if one also implements post-selected detection to ignore events with areas that are 'too close' to the bin boundary. Events with areas falling nearer to the center of each bin can be considered as having a high chance to actually be within the bin they were found. Rather than choosing the intersection points of the Gaussians to fix the bin edges, one could instead use a specific distance from the bin center in terms of the Gaussian width, $\sigma_n$, for each peak and ignore all events with areas landing outside of any defined bin. Fortunately, because the areas of the histogram follow Gaussian distributions in each bin, systematically ignoring events will not alter the measured probability distribution provided that each newly defined bin is the same size as all of the others in terms of fractions of the associated $\sigma_n$.

An example of this method is shown in Fig. 6.14(a), where a region of uncertainty is defined near the bin intersection and we zoom in on the 25 photon bin. In Fig. 6.14(b), we suppose that we have a perfect 25 photon event and plot the probability to miscount this 25 photon event as a function of the fraction of data used. We see that cutting out

half of the measured data reduces the error rate by an order of magnitude to $\sim$ 1%. Provided one can take enough statistics, the error rate can be drastically reduced on even large photon-number counts.

## 6.6 Random Number Generation with the TES

One practical use of the high PNR capability of the TES is to generate truly random numbers. Random numbers are an essential requirement in many applications including simulation and cryptography [265]. Utilizing the inherent randomness in quantum mechanics enables us to generate truly random numbers as opposed to algorithmically generated pseudo-random numbers. We were able to experimentally verify a recent proposal for quantum random-number generation (QRNG) based on measuring the photon-number parity of coherent light [266]. The proposed method does not require post-processing and is immune to phase noise. Furthermore, we can prove that this method is resilient against external noise and we provide arguments for its robustness against small amplitude fluctuations, detector inefficiency, and eavesdropping. We also argue that it is tamper-proof provided we use a frequency filter and monitor the output distribution periodically. With a photon-counting resolution of up to 30 photons, we construct the distribution of a coherent state with an average photon number greater than 14. This enables us to measure an expectation value of our parity operator close to zero ($10^{-4}$), eliminating bias. Detector inefficiency is no obstacle as coherent states are eigenstates of the annihilation operator and therefore globally invariant under losses (which transform a coherent state into a lower-amplitude coherent state). In addition to the near-zero bias, the high-resolution ability allows us to double our generation rates (bit rates) by using modulo-4 photon-number binning instead of the modulo-2, parity binning, which also yields a uniform distribution. Finally, we prove that QRNG with parity detection only requires a state with zero-valued parity expectation value. The main idea is that the expectation value of the parity operator in a coherent state exponentially approaches zero as the average photon number increases. This means that we have an equal probability of getting an even or odd number of photons every time we detect a pulse on our photon number resolving detector. Assigning a 0 to even numbers and 1 to odd, we can then construct binary strings of random numbers.

We begin by looking at the expectation value of the parity operator. Intuitively, the expectation value being near zero can be seen by remembering that the value of the

Wigner function at the origin is the same as the expectation value of the parity [267] (discussed more in the next chapter). Because the coherent state Wigner function is Gaussian, the value of the Wigner function exponentially trails to zero away from the center of the displacement in phase space. Thus for large displacements, the value of the Wigner function at the origin is very near zero.

Mathematically, this is easily obtained by looking at the form of a coherent state: $|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$. The parity operator can be written as $\hat{\Pi} = (-1)^{\hat{n}} = e^{i\pi\hat{n}}$, where $\hat{n} = \hat{a}^\dagger \hat{a}$ is the photon-number operator. Thus we can derive

$$\langle\alpha| \hat{\Pi} |\alpha\rangle = e^{-2\bar{n}}, \tag{6.16}$$

where $\bar{n} = \langle\alpha| \hat{n} |\alpha\rangle$. Next, we consider the case where we do not have a pure coherent state, but a statistical mixture of coherent states with the same amplitude and totally random phase,

$$\rho_{coh} = \frac{1}{2\pi} \int_0^{2\pi} d\phi \, |re^{i\phi}\rangle \langle re^{i\phi}| , \tag{6.17}$$

where $r = |\alpha| = \sqrt{\bar{n}}$. This yields

$$\langle\hat{\Pi}\rangle = \mathrm{Tr}[\rho_{coh}\hat{\Pi}] = e^{-2\bar{n}}, \tag{6.18}$$

which shows that phase noise does not affect the parity expectation value.

Changes in the amplitude of the coherent state amount to changes in the mean photon number $\bar{n}$. For a change $\delta$ in the mean photon number, the parity expectation value becomes $e^{-2(\bar{n}\pm\delta)}$ which is approximately $e^{-2\bar{n}}$ for small $\delta$. As for detector inefficiency, we can always model an inefficient detector as a perfect one with a beamsplitter before it of transmittivity $\eta$. Doing so, the detector always gets a coherent state from the beamsplitter output, albeit of a different $\bar{n}$. And as long as the final $\bar{n}$ is large enough, all the previous reasoning holds. We stay with the beamsplitter to explore another idea — that of eavesdropping. Suppose an eavesdropper uses a beamsplitter to sample the light used to generate random numbers. Due to the nature of coherent light, the two beamsplitter outputs are in a product state, hence not correlated and no information about the results at one output port can be used to determine the results at the other.

Consider now the addition of environmental noise, which can be understood as follows: What is the expectation value of the parity operator on the whole system where $\rho = \rho_{coh} \otimes \rho_{env}$? Deriving the expectation value of the new parity operator, $e^{i\pi\sum\hat{n}_i}$,

where subscript $i$ denotes the different subsystems, we obtain

$$\langle e^{i\pi\sum\hat{n}_i}\rangle = Tr[e^{i\pi\hat{n}_1}\rho_{coh}\otimes e^{i\pi\hat{n}_2}\rho_{env}] = e^{-2\bar{n}}\langle\hat{\Pi}\rangle_{env}, \tag{6.19}$$

where $\langle\hat{\Pi}\rangle_{env}$ is bounded between 1 and -1. For large enough $\bar{n}$, the whole expectation value goes to zero regardless of the form of $\rho_{env}$. Finally, we explore the general requirements needed to satisfy our method of quantum random number generation. The necessary condition imposed is that a state with $\langle\hat{\Pi}\rangle = 0$ impinges upon an ideal PNR detector. Such a state will still be immune to mixing with the environment and phase shifts since PNR detection is phase-insensitive, but it will no longer necessarily be immune to eavesdropping, amplitude fluctuations, and imperfect detectors. These attributes arise because the coherent state is an eigenstate of the annihilation operator, and a coherent state exits a beamsplitter as yet another coherent state.

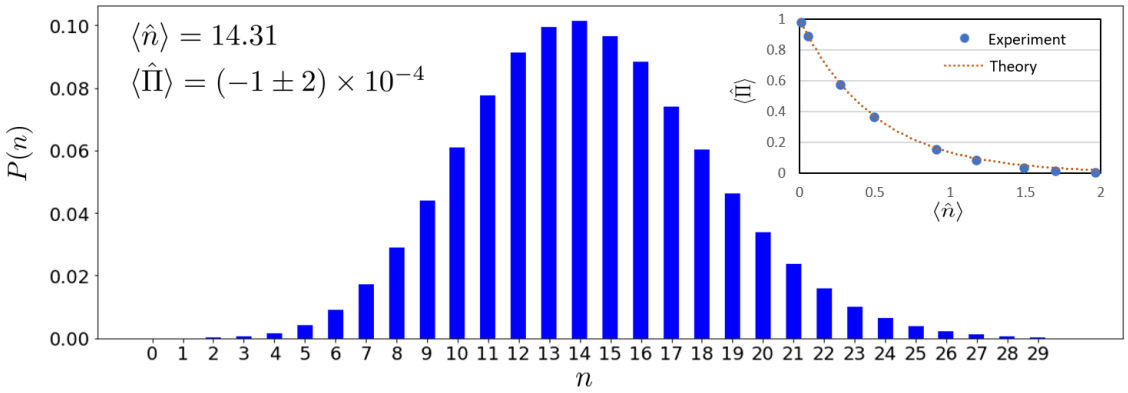



FIGURE 6.15: Measured photon number distribution of coherent laser output from $1.07 \times 10^8$ samples. The inset plot shows the measured parity exponentially decreases with increasing mean photon number, as predicted by Eq. 6.16. Error due to finite sampling falls within the size of the data markers.

We implemented QRNG by coupling a strongly attenuated pulsed coherent light source at 1064 nm into a TES device cooled to 100 mK and analyzed the response on-the-fly with an Ethernet-based flash analog-to-digital converter (EFADC). The EFADC computes the signal pulse areas, which can be used to accurately resolve photon number in real time [165]. Our system achieved PNR capability up to 30 photons. By only keeping the parity of the measured event, we can ensure a nearly bias-free coin flip when sampling a coherent state with large enough mean photon number. We collected $10^8$ samples using a coherent state with $|\alpha^2| = 14.31$ and experimentally measured the expectation of parity to be zero within finite-sampling error. The results are presented

in Fig. 6.15.

Furthermore, we demonstrated the robustness of our system to external noise and possible tampering. We note that the parity of the measured distribution still tends to zero even when the coherent state is phase-randomized and thermal environmental noise is introduced. Additionally binning the distribution by modulo 4 successfully doubled the bit generation rate while remaining bias-free.

# Chapter 7

# Tomography with PNR Detectors

Applications in quantum information often require a process where one can completely characterize a quantum state in what is known as quantum state tomography [268]. This is useful not only when one would like to verify some types of quantum state engineering experiments are successful, but also if one is sent a completely unknown state. It is, of course, impossible to determine the state of a single unknown quantum system as any measurement will disturb the state, but it is possible to perform quantum state tomography on an ensemble of identically prepared states. Even still, this is not a straightforward task in general. Suppose Alice prepares many copies of the pure quantum state, $|\psi\rangle$, and sends them to Bob. When Bob measures each state, he projects it into a specific eigenstate in the eigenbasis of the measurement operator. By performing this measurement many times, he can construct a histogram of the probabilities, which according to the Born rule, allows the determination of $|\psi(m)|^2$ where $m$ is an eigenstate of the measurement operator $\hat{M}$. Clearly, this does not provide complete information about the full wavefunction, $\psi(m)$, so more information is required. If Bob now measures the same ensemble in the conjugate basis using the measurement operator $\hat{M}'$ to construct $|\tilde{\psi}(m')|^2$, then he now has enough information to reconstruct the quantum state due to the relationship of the wavefunction $\psi$ to $\tilde{\psi}$ through the Fourier transform. The determination of a quantum state in this way is an example of phase retrieval [269].

Quantum state tomography is not restricted to pure states, as mixtures will always be the predominant form of any experimentally generated state. Thus one must entirely reconstruct an unknown density operator, $\rho$, where measuring in a single basis only provides access to the main diagonal. For an $N$ mode (or particle) system where each mode can be described by a $d$-dimensional Hilbert space, the density operator of

the whole system can be written as a $d^N \times d^N$ matrix. Since $\mathrm{Tr}[\rho] = 1$ and $\rho$ is Hermitian, the density matrix has $d^{2N} - 1$ real parameters that must be found. The problem of full state tomography is thus inefficient as it requires a number of measurement settings that is exponential in the number of modes in general. Certain situations may reduce the number of required measurements if one has *a priori* information about the system, or if $\rho$ is sparse [270].

Alternative to the density operator, the Wigner function [271, 272] can equivalently be used as a full state descriptor:

$$W(q,p) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{2ipy}\langle q-y|\rho|q+y\rangle dy, \tag{7.1}$$

where $q$ and $p$ are the eigenvalues of the respective amplitude and phase quadratures, which are defined in terms of the bosonic creation and annihilation operators as discussed in Chapter 1. The experimental determination of the Wigner function was first proposed and realized utilizing homodyne quadrature measurements, as process in which a bright local oscillator is used as a phase reference and interfered with the state to be measured [273]. Homodyne tomography is a mature technology that is commonly used to tomograph CV quantum optical states; however, many difficulties still prevent this technique from being the final word on tomographic methodology [274]. In addition to the necessity for computation-intensive reconstruction algorithms, maximum-likelihood (MaxLik) and maximum-entropy (MaxEnt) methods may fail to converge or give unphysical results, especially in the presence of noise during the measurement process.

In this chapter, I will discuss optical methods of tomography alternative to homodyne detection that involve PNR detection schemes. The majority of this chapter is based on the recently published overlap tomography method [264], but I will also include a variety of extensions that are not found elsewhere.

Before moving on to other methods of tomography, it is interesting to fundamentally examine *why* homodyne measurement gives rise to difficulties in state reconstruction even though one has access to fast and reliable measurements in linear combinations of two non-commuting operators ($\hat{Q}$ and $\hat{P}$). Suppose one has a single-mode density operator that can be written out in the Fock basis with a Hilbert space truncation of $d$ (no photon-number components greater than $d$), and thus has $d^2 - 1$ real parameters. As a continuous variable measurement, performing an ideal homodyne

detection would perfectly project $\rho$ onto an eigenstate of $\hat{Q}$ or $\hat{P}$, which are both unphysical infinite energy eigenstates. Although the density operator has a finite number of elements in the Fock basis, the quadrature basis requires a continuous representation of the density operator, robbing the experimenter of the ability to perform only $d^2 - 1$ measurement settings. Thus for homodyne tomography to work properly, one must measure using *many* settings, where each setting corresponds to a different angle of the phase reference and the measurement projects upon the generalized quadrature

$$\hat{A}(\theta) = \hat{Q}\cos(\theta) + \hat{P}\sin(\theta). \tag{7.2}$$

Furthermore, many measurements must be made at each measurement setting to accurately construct the probability distribution at that setting. Even in the absence of experimental noise, a perfect reconstruction requires measurements over a *continuum* of $\theta$ for measurement settings. This is where methods such as MaxLik and MaxEnt make their debut — to compensate for lack of information at all continuous $\theta$ and aid the reconstruction algorithm when information is lacking. In spite of the difficulties mentioned here, homodyne tomography is useful in many applications.

## 7.1 Point-by-point Tomography

Rather than using a quadrature basis to perform measurements, one could make use of the photon-number basis to measure quantum states. Fortunately, the value of the Wigner function at the origin of phase space corresponds to the expectation value of the parity operator [267]. The parity operator can be expressed as

$$\hat{\Pi} = (-1)^{\hat{n}} = e^{i\pi\hat{n}}. \tag{7.3}$$

From here, one can measure the photon number distribution of the unknown state and directly obtain the value of the unknown Wigner function of state $\rho$ at the origin according to

$$W_\rho(0,0) = \frac{1}{\pi}\sum_{n=0}^{\infty}(-1)^n P(n). \tag{7.4}$$

Wallentowitz-Vogel [275] and Banaszek-Wodkiewicz [276] (WVBW) proposed a method by which this fact can be used to perform full state tomography. In this method,

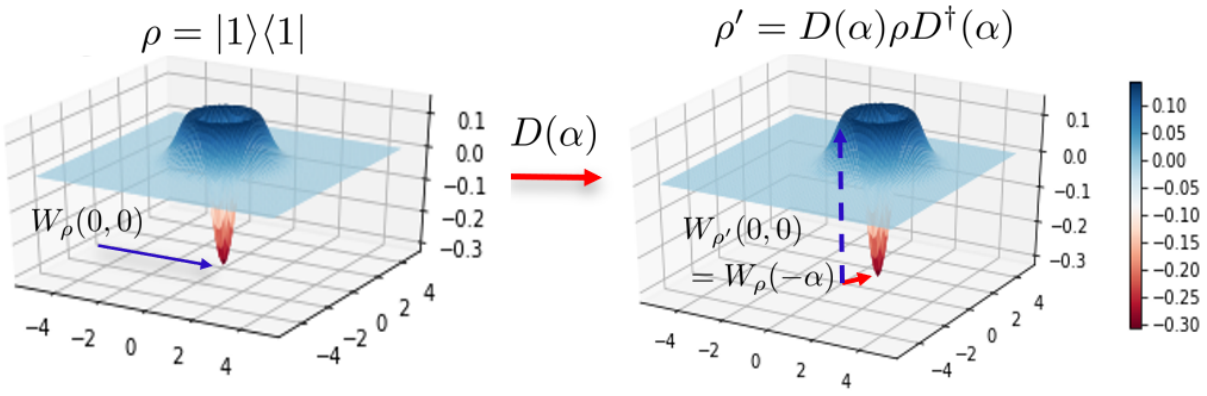



FIGURE 7.1: Point-by-point tomography method described in the main text.

a simple phase-space translation, i.e., displacement, of the quantum state to be characterized, followed by parity measurements, allows easy determination of the Wigner function. An example of this is shown in Fig. 7.1, where the Wigner function of an unknown state (in this case a single-photon) is first measured at the origin, and then translated before another measurement is performed. This methodology was first experimentally implemented on the phononic field vibration of a single trapped ion [277] and on microwave cavity fields [278, 279]. More recently, optical PNR detection such as with TES devices has allowed the use of the full WVBW method on traveling optical fields with no prior knowledge of the measured quantum state [114, 263]. While the WVBW method presents clear advantages in terms of the numerical demands on reconstruction, it requires a phase space raster scan involving a large number of optical displacements, and the required measurement-point density of the raster scan is determined by the specific features of the unknown Wigner function to be resolved. Because the method maps out the Wigner function instead of the density operator, one can never determine the smooth function exactly, but only interpolate a continuous function between a finitely-grained sample grid. This will often be sufficient, but there are times where the interpolation will fail to give accurate results unless many measurement points are obtained, as we will discuss later. Moreover, the best experimental implementation of phase space displacements is intrinsically imperfect [127].

## 7.2 Overlap Tomography

In Ref. [264], we presented a generalization of the WVBW approach which uses a Wigner function overlap measurement to reconstruct the density operator, rather than the Wigner function, using computationally efficient semidefinite programming. This general method requires considerably less data acquisition and ensures physical results which are robust against measurement noise. Since this method reconstructs the density operator directly, it can be done with an efficient number of measurements settings.

### 7.2.1 Theory of wigner function overlap

To understand overlap tomography, consider the situation depicted in Fig. 7.2(a): a

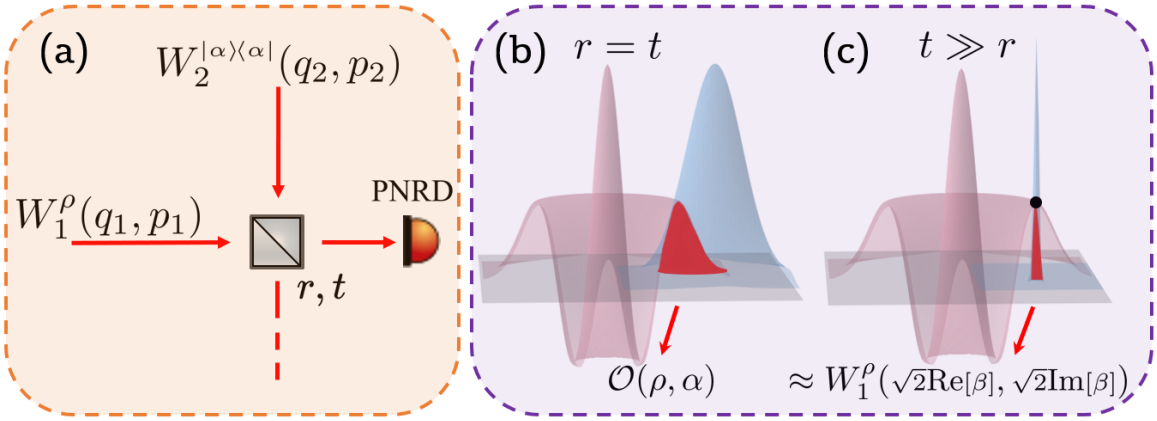



FIGURE 7.2: (a), Schematic of the experiment: the field to be measured, of density operator $\rho$, interferes with a calibrated field in coherent state $\alpha$ at a beamsplitter of field reflectance $r \in \mathbb{R}$ and transmittance $t = (1 - r^2)^{1/2}$. PNRD: photon-number-resolving detector. (b), Principle of generalized overlap tomography exemplified with a two-photon Fock state. (c), Limit case of (b), where a highly unbalanced beamsplitter merely implements a displacement of $\rho$ by $-\beta$.

field with unknown density operator $\rho$ and Wigner function $W_1(q_1, p_1)$ interferes at a beamsplitter with a reference coherent state $|\alpha\rangle\langle\alpha|$ of Wigner function $W_2(q_2, p_2)$. Measuring the Wigner function at the origin of phase space of just one output mode directly yields the Wigner function overlap between $\rho$ and a probing function. To see this, we adopt the Heisenberg picture and determine the evolved output quadratures under the beamsplitter interaction to be $q_1' = tq_1 - rq_2$ and $p_1' = tp_1 - rp_2$. Likewise, $q_2' = rq_1 + tq_2$ and $p_2' = rp_1 + tp_2$.

The two-mode input state is written in the Wigner function representation as

$$W_{1,2}(\mathbf{x}) = W_1(q_1, p_1)W_2(q_2, p_2). \tag{7.5}$$

Next, by using the evolved quadratures, one can write the Wigner function of the beamsplitter output as

$$W'_{1,2}(\mathbf{x}') = W_1(tq'_1 + rq'_2, tp'_1 + rp'_2)W_2(-rq'_1 + tq'_2, -rp'_1 + tp'_2), \tag{7.6}$$

where $\mathbf{x}$ and $\mathbf{x}'$ are column vectors consisting of quadratures corresponding to the input and output modes, respectively. The value of the Wigner function of output mode 1 at the origin can be obtained by setting $q'_1 = p'_1 = 0$ and tracing out over mode 2 to yield

$$\iint W'_{1,2}(\mathbf{x}')dq'_2dp'_2|_{q'_1,p'_1=0} = \iint W_1(rq'_2, rp'_2)W_2(tq'_2, tp'_2)dq'_2dp'_2 \tag{7.7}$$

$$= \frac{1}{r^2}\iint W_1(q, p)W_2(\tfrac{t}{r}q, \tfrac{t}{r}p)dqdp. \tag{7.8}$$

Setting $r = t = \frac{1}{\sqrt{2}}$ yields the Wigner function overlap between $\rho$ and $|\alpha_j\rangle\langle\alpha_j|$, i.e., the overlap $\mathcal{O}$ of the unknown $\rho$ with $|\alpha\rangle\langle\alpha|$:

$$\mathcal{O} = \mathrm{Tr}[\rho|\alpha\rangle\langle\alpha|] = \pi W'_1(0, 0; \tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}) = \sum_n^N (-1)^n \rho'_{nn}, \tag{7.9}$$

which can be obtained by a parity measurement on mode 1 ($\rho'$) after the beamsplitter. Note that $\mathcal{O}$ is proportional to the Husimi Q function, $Q(\alpha)$, which we sample sparsely [280].

In the case of an unbalanced beamsplitter where the input probe is still a coherent state, $\alpha_j = \frac{1}{\sqrt{2}}(q_{\alpha_j} + ip_{\alpha_j})$, we now have

$$W_2(\tfrac{t}{r}q, \tfrac{t}{r}p) = W_{|\alpha_j\rangle\langle\alpha_j|}(\tfrac{t}{r}q, \tfrac{t}{r}p) = \frac{1}{\pi}\exp\left[-\left(\tfrac{t}{r}q + q_{\alpha_j}\right)^2 - \left(\tfrac{t}{r}p + p_{\alpha_j}\right)\right)^2\Big] \tag{7.10}$$

$$= \frac{1}{\pi}\exp\left\{-\frac{1}{\sigma^2}\left[\left(q + q_{\beta_j}\right)^2 + \left(p + p_{\beta_j}\right)^2\right]\right\}, \tag{7.11}$$

where $\sigma = \frac{r}{t}$ and $\beta_j = \sigma\alpha_j$. The integral overlap from Eq. 7.8 is then

$$\frac{1}{t^2}\iint W_1(q, p)W_\sigma(q + q_{\beta_j}, p + p_{\beta_j})dqdp, \tag{7.12}$$

where $W_\sigma(q+q_{\beta_j}, p+p_{\beta_j}) = \frac{1}{\pi\sigma^2}\exp\left\{-\frac{1}{\sigma^2}\left[\left(q+q_{\beta_j}\right)^2+\left(p+p_{\beta_j}\right)^2\right]\right\}$. When $\sigma > 1$, Eq. (7.12) gives a scaled overlap between $\rho_{in}$ and a thermal state displaced by $\beta_j$. When $\sigma < 1$, however, the overlap is between $\rho_{in}$ and a mathematical object that approaches a displaced delta function in the limit where $\sigma \to 0$ and $|\alpha| \to \infty$, which exactly probes the Wigner function of $\rho_{in}$ point-by-point as in the unbalanced homodyne technique of Refs. [275] and [276]. One can find more detailed information on calculating the Wigner function overlap between $\rho$ and the $\sigma$-width Gaussian function in the supplementary material of Ref. [264].

One might also wonder about the outcome of a similar measurement performed on the other port of the beamsplitter. If we go back to Eq. (7.6) and now determine the value of the Wigner function of output mode 2 at the origin while tracing out over mode 1, we have

$$\iint W'_{1,2}(\mathbf{x}')dq'_1dp'_1|_{q'_2,p'_2=0} = \iint W_1(tq'_1, tp'_1)W_2(-rq'_1, -rp'_1)dq'_1dp'_1 \tag{7.13}$$

$$= \frac{1}{t^2}\iint W_1(q,p)W_2(-\tfrac{r}{t}q, -\tfrac{r}{t}p)dqdp. \tag{7.14}$$

This time when we set $r = t$, the measured Wigner function overlap is between $\rho_{in}$ and $|-\alpha_j\rangle\langle-\alpha_j|$, i.e., the very same coherent state probe with a phase factor of $e^{i\pi}$. From this, we can conclude that measuring the Wigner function at the origin of both beamsplitter outputs would yield the overlap between the signal and coherent-state probes at opposite phases. When performing the tomographic reconstruction, it is possible to utilize both outputs to collect overlap measurements and only externally vary the probe phases by half of the desired range; however, ensuring that both detection channels following the beamsplitter are identical in losses, detector efficiency, etc., is experimentally challenging, and this also imposes additional requirements on PNR detection capabilities. Therefore, it is simpler to utilize a single output mode to perform the tomographic reconstruction and correct for known losses as detailed below and in the main text.

### 7.2.2 Tomographic reconstruction with semidefinite programming

We now use the formalism discussed above to perform the complete state tomography of an arbitrary quantum state. For a given single-mode quantum state, one can write

the density operator in the photon-number basis as

$$\rho = \sum_{n,n'=0}^{\infty} \rho_{nn'} |n\rangle\langle n'|. \tag{7.15}$$

Complete characterization of $\rho$ requires determining $\rho_{nn'}$for all $n, m$. To do that, we choose a set of distinct coherent states, $|\alpha_j\rangle$. Using Eq. (7.8), we obtain the fidelity between $|\alpha_j\rangle$ and $\rho$, formulated as

$$\mathcal{O}^j = \mathrm{Tr}\left[|\alpha_j\rangle\langle\alpha_j|\rho\right] = \langle\alpha_j|\rho|\alpha_j\rangle. \tag{7.16}$$

Using Eq. (7.16) and the coherent state represented in the photon-number basis, $|\alpha_j\rangle = \sum_{n=0}^{\infty} c_{jn}|n\rangle$, we get

$$\mathcal{O}^j = \sum_{m'=0}^{\infty} c^*_{jm'}\langle m'| \sum_{n,n'=0}^{\infty} \rho_{nn'}|n\rangle\langle n'| \sum_{n'=0}^{\infty} c_{jm}|m\rangle. \tag{7.17}$$

Further simplification leads to

$$\mathcal{O}^j = \sum_{n,m=0}^{\infty} c_{jm}c^*_{jn}\rho_{nm}. \tag{7.18}$$

Ideally, the sum over $n, m$ goes to infinity but for practical purposes one needs to truncate it at, say $N$, such that any terms $n, m > N$ do not significantly contribute to the sum. As a result, we have

$$\mathcal{O}^j = \sum_{n,m=0}^{N} c_{jm}c^*_{jn}\rho_{nm}, \tag{7.19}$$

where $c_{jm}c^*_{jn} = e^{-|\alpha_j|^2}\frac{(\alpha_j^*)^n(\alpha_j)^m}{\sqrt{n!m!}}$.

Furthermore, by using $N_p = (N+1)^2$ coherent states, Eq. (7.19) can be written in the matrix form as

$$\begin{pmatrix} \mathcal{O}^{(0)} \\ \mathcal{O}^{(1)} \\ \vdots \\ \mathcal{O}^{N_p} \end{pmatrix} = \begin{pmatrix} c_0^0 c_0^{0*} & c_0^0 c_1^{0*} & \dots & c_N^0 c_N^{0*} \\ c_0^1 c_0^{1*} & c_0^1 c_1^{1*} & \dots & c_N^1 c_N^{1*} \\ \vdots & \vdots & \ddots & \vdots \\ c_0^{N_p} c_0^{N_p*} & c_0^{N_p} c_1^{N_p*} & \dots & c_N^{N_p} c_N^{N_p*} \end{pmatrix} \begin{pmatrix} \rho_{0,0} \\ \rho_{0,1} \\ \vdots \\ \rho_{N,N} \end{pmatrix}. \tag{7.20}$$

We can rewrite the above matrix equation in compact form as

$$\mathbf{O} = \mathbf{CP}, \tag{7.21}$$

where $\mathbf{O} \in \mathbb{R}^{(N+1)^2}$, $\mathbf{P} \in \mathbb{C}^{(N+1)^2}$ and $\mathbf{C} \in \mathbb{C}^{(N+1)^2 \times (N+1)^2}$. Next, we can invert Eq. (7.21) to reconstruct $\mathbf{P}$. This can be achieved by solving the following semidefinite program:

$$\begin{aligned} \underset{\mathbf{P}}{\text{Minimize}} \quad & ||\mathbf{O} - \mathbf{CP}||_2 \\ \text{Subject to} \quad & \rho \geq 0 \quad \text{and} \quad \text{Tr}[\rho] = 1, \end{aligned} \tag{7.22}$$

where $||.||_2$ is the $l_2$ norm defined as $||V||_2 = \sqrt{\sum_i |v_i|^2}$. Note that this kind of quadratic convex techniques have been extensively discussed in the the context of quantum detector tomography [281–283]. The optimization problem is convex and can be efficiently solved for a guaranteed unique $\mathbf{P}$, and hence for the unknown state, $\rho$, using open source Python module CVXPY [284].

It is important to emphasize that the coefficient matrix in Eq. 7.20 is a square matrix with a number of rows equal to the number of overlap coherent states, i.e., the number of measurements settings used. Thus, in order to have enough measurement settings to reconstruct the density matrix and solve the matrix inversion, we need one measurement setting for each element in $\rho$, so $N^2$ measurement settings. This is only one more than the lower limit of $N^2 - 1$, which is set by the number of independent parameters in the density matrix. Thus, this method is considerably more efficient in terms of required measurement settings than either homodyne tomography or the WVBW method, which I will elaborate on in the next section.

Although the above method holds for general quantum states, we restrict our simulations to real-valued density matrices for numerical ease. However, we do show the reconstruction for complex-valued density matrices in Ref. [264]. All numerical simulations were performed in QuTip [285] where the Hilbert space for each optical mode was constructed in the Fock basis with a high enough dimensionality to ensure state probability amplitudes decayed to less than $10^{-7}$ before truncation. Under these parameters, the numerically efficient SDP algorithms converged on the order of $10^{-2}$ seconds on a 3GHz Intel i5 quad-core processor with 16Gb RAM.

Our method is demonstrated in Fig. 7.3 for the example cases of the small amplitude cat state, $|\psi\rangle \propto |\alpha\rangle + |-\alpha\rangle$ where $\alpha = \sqrt{3}$, and a Gottesman-Kitaev-Preskill (GKP) state of mean photon number 5. These states were reconstructed using 400 different probing coherent states of 20 amplitude increments from $\beta = 0$ to $\beta = \sqrt{6}$ and 20 phase increments from 0 to $2\pi$, to achieve fidelities with the target states greater than 0.999 for the cat state, and a fidelity of 0.985 for the GKP state. Because the observer is assumed to have no prior knowledge of the state to be characterized, it is important to scan the entirety of phase space in question with different coherent states so as to have sufficient overlap between all portions of the state under test. If some prior knowledge of the state is obtained, then the probing coherent states can be restricted to a localized region of phase space near the unknown quantum state.

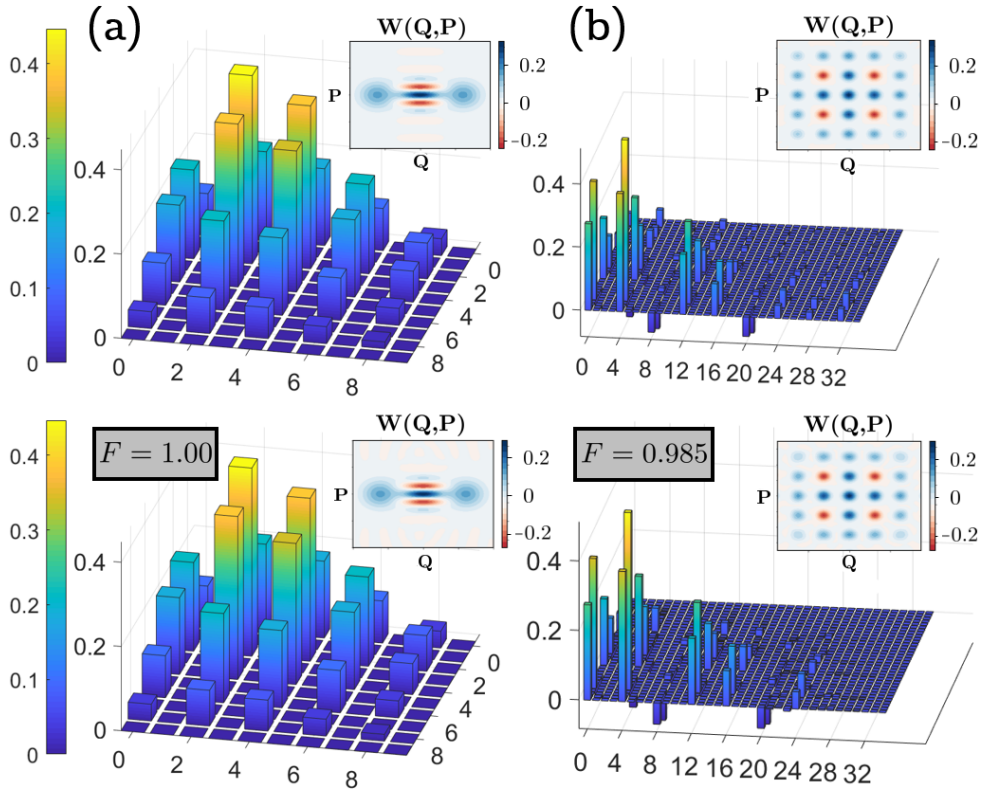



FIGURE 7.3: Tomographic reconstruction using 400 coherent state probes for **(a)** a cat state of amplitude $\sqrt{3}$, and **(b)** a GKP state with a mean photon number of 5. The top row displays the density matrix for the ideal theoretical state, and the bottom shows the reconstructions. Insets display the plotted Wigner function for each state.

This method also exhibits some resilience to noise in that acquiring more measurements adds some redundancy to the SDP optimization problem. This redundancy can help average out some measurement-induced noise, such as fluctuations in the phase or amplitude of the coherent state used to probe the unknown density matrix. This is discussed in more detail in the supplementary material in Ref. [264] and in R. Nehra's thesis [286], so I with omit further discussion on this point. This point will also be evident in the experimental success of the procedure, which I will show later in this chapter.

### 7.2.3 Phase-Space resolution

In this section, I compare the number of measurement settings required to accurately reconstruct a Wigner function using the WVBW method vs. the overlap tomography method. Using WVBW, the ability to perform a parity measurement allows one to map out the entire Wigner function in a point-by-point manner without the need for any numerical reconstructions. However, the Wigner function can, and in general, will acquire complicated structures with sub-Planck length features [287]. For a quantum state localized to a region of phase-space with area $A$, this structure can appear on scales with area $a \approx 1/A$ and require fine-grained sampling in order to resolve when performing WVBW tomography. Additionally, if one has little knowledge of the quantum state in question (excluding possible information about total energy or even the exact photon-number distribution), the location of the quickly-varying Wigner function structure may appear anywhere in phase space, meaning that the WVBW method will need to sample a minimum of one point per phase-space area $a$ over the entire region. This leads to the requirement that the Wigner function be probed at $\approx A^2$ points. However, using the overlap tomography method, we only need $d^2$ distinct measurements to completely reconstruct the unknown density matrix, where $d$ is the dimension of the Hilbert space in the Fock basis. Interestingly enough, the number of measurements between each method scales similarly, but simply measuring the Wigner function at one point per phase-space area of $a$ may not be sufficient if one is unlucky, as I will demonstrate. But first, one might ask - given that I know where to truncate my state, how far should I measure in phase-space? Or, if I want to measure the Wigner function out to some boundary, at what $d$ do I truncate my Hilbert space?

In general, the Wigner function is a linear combination of each density matrix element $|n\rangle\langle n'|$, where each component is given by [288]

$$W_{nn'}(q,p) = \begin{cases} \frac{(-1)^n}{\pi}\sqrt{\frac{2^{n'}n!}{2^n n'!}}(q-ip)^{n'-n}e^{-(q^2+p^2)}L_n^{n'-n}(2(q^2+p^2)) & n \leq n' \\ \frac{(-1)^{n'}}{\pi}\sqrt{\frac{2^n n'!}{2^{n'}n!}}(q+ip)^{n-n'}e^{-(q^2+p^2)}L_{n'}^{n-n'}(2(q^2+p^2)) & n \geq n' \end{cases}, \tag{7.23}$$

and $L_n^\alpha$ are the Laguerre polynomials. The furthest features in the Wigner function are determined just beyond the last zero of the largest contributing $n$ Laguerre polynomial, where the function can attain non-trivial values before the exponential modulation terminally brings the Wigner function to zero. Suppose I truncate my Hilbert space at $d = N+1$, such that the $|N\rangle\langle N|$ is the last contributing term. This component's contribution to the Wigner function can be seen in phase-space at least up to (and slightly beyond) the last zero of $L_N^0$. Although the exact location of the zeros $x_k$ of Laguerre polynomials remains an open question, there exist bounds on their locations [289], leading to the largest zero adhering to

$$x_1 > 2N - 2 + \sqrt{(N-1)^2+1}. \tag{7.24}$$

Using $x_1$ as the argument for the Laguerre poylnomial in Eq. 7.23, it is easy to see that the Wigner function must be mapped out to at least a phase-space radius of

$$q_{max} = \left[N - 1 + \frac{1}{2}\sqrt{(N-1)^2+1}\right]^{1/2}, \tag{7.25}$$

leading to a minimum phase space area

$$A = \pi(N-1) + \frac{\pi}{2}\sqrt{(N-1)^2+1}. \tag{7.26}$$

For large $N$, we have that $A^2 \approx \frac{9\pi^2}{4}N^2$. When comparing the number of required measurements between $d^2$ and $A^2$, we see that both the overlap and WVBW tomography methods scale as $N^2$; however, the larger scaling factor on $A^2$ should not be ignored as this could pose additional experimental difficulties for large $N$, and the increase in required measurements will only be exacerbated for multimode systems when the required measurement scaling will become $d^{2M}$ and $A^{2M}$, respectively, for an $M$-mode system.

**Cat-state example**

To explore a specific example, consider the four-component cat state (a.k.a compass state), given by

$$|\psi\rangle = |\alpha\rangle + |-\alpha\rangle + |i\alpha\rangle + |-i\alpha\rangle \tag{7.27}$$

with photon-number distribution

$$P(n) = \begin{cases} \frac{2|\alpha|^{2n}}{n!(\cosh|\alpha|^2+\cos|\alpha|^2)} & n \bmod 4 = 0 \\ 0 & \text{otherwise} \end{cases} \tag{7.28}$$

The Wigner function in Eq. 7.27 consists of Gaussians from each of the four classical coherent-state $|\alpha\rangle\langle\alpha|$ components located at corners of a square in phase-space, interference terms perpendicular to each edge of the square that resemble cat-state fringes, and interference terms between the Guassians along the diagonals of the square resulting in a checkerboard pattern. The checkerboard interference is governed by the oscillatory term

$$\cos(Pq) + \cos Lp = 2\cos\left(\tfrac{1}{2}(Pq + Lp)\right)\cos\left(\tfrac{1}{2}(Pq - Lp)\right), \tag{7.29}$$

where $L = P = 2\sqrt{2}|\alpha|$ is the the phase-space separation between opposite gaussian peaks. The periodic tiles formed each contain four smaller squares of area (see Ref. [287] for details)

$$a = \frac{\pi^2}{2|\alpha|^2}, \tag{7.30}$$

where I have used $\alpha = \frac{1}{\sqrt{2}}(q + ip)$ and set $\hbar = 1$. If I wish to measure the Wigner function out to where it decays to an arbitrarily chosen value of $10^{-2}$, then I will need to measure to a phase-space radius of $q_{max} = \sqrt{2}|\alpha| + 2$, which is the point at which the four Gaussian terms have decayed sufficiently. This value comes from the fact that each Gaussian is centered at distance of $\sqrt{2}|\alpha|$ with $\sigma = 1$, and the state is normalized so that each of the four peaks has a maximum value of $\frac{1}{4\pi}$. This leads to an area of $A = \pi q_{max}^2$ requiring

$$M_{WVBW} = \frac{A}{a} = \frac{2}{\pi}|\alpha|^2(\sqrt{2}|\alpha| + 2)^2 \approx \frac{4}{\pi}|\alpha|^4 \tag{7.31}$$

measurements with the WVBW method for large $|\alpha|$. Using $q_{max}$ in Eq. 7.25 to find $N$ results in the number of measurements required for overlap tomography to be

$$M_{\mathcal{O}} \approx \frac{16}{9}|\alpha|^4 \tag{7.32}$$

Note that the use of Eq. 7.25 relies on Eq. 7.24, which could result in an over-estimation for the required $N$ due to the inequality.

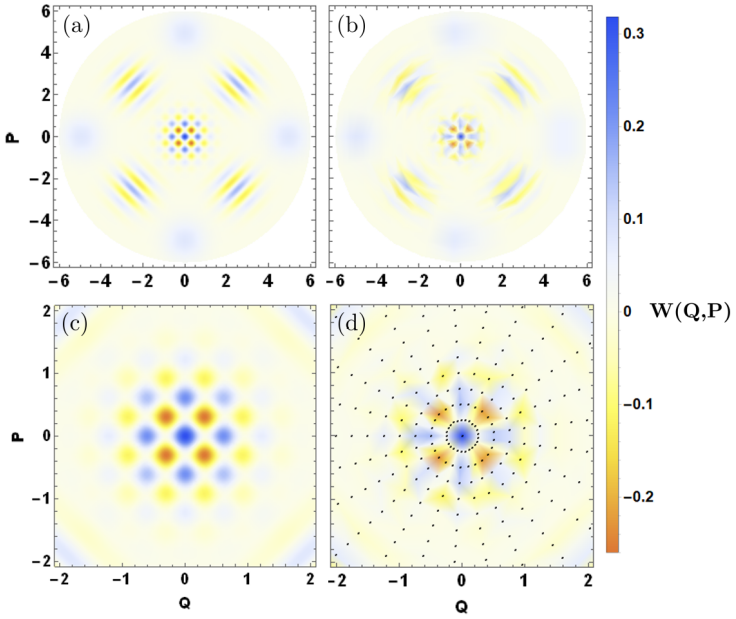



FIGURE 7.4: (a) Wigner function of compass state in the ideal case using overlap tomography. (b) Wigner function of the same state using WVBW tomography with the same number of measurements ($25^2$) as in (a). To achieve the same crisp resolution as depicted at left, the WVBW method would require $100^2$ measurements. (c) and (d) zoom in on the central interference fringes with (d) showing the actual measurement coordinates as black points.

Looking at Eq. 7.31 and 7.32, the overlap method actually scales slightly worse at first glance than the WVBW method as $\alpha$ becomes large. However, if one wishes to catch the fast oscillations of the Wigner function using the WVBW method, measuring one point in every area $a$ may not be sufficient if one is unlucky. Consider the Wigner function of a four-component cat state displaced by $D(K\frac{\sqrt{a}}{2})$, where $K$ is an odd integer. Now, if the WVBW measurement scheme covers a rectangular grid beginning at the origin with point spacing of $\sqrt{a}$, the measurements will always unluckily find the zeros of Eq. 7.29, leading to what appears to the observer to be a flat Wigner function around the origin with no quantum interference. To avoid cases like this, it would be necessary

to sample more points for each phase-space area $a$. Interestingly enough, this directly relates to the Nyquist-Shannon sampling theorem [290, 291], whereby it is necessary to have a sampling rate of at least twice the signal frequency in order to reconstruct it. In this case, the frequency of oscillation in each direction, from Eq. 7.29, is $f = \frac{\sqrt{2}}{\pi}|\alpha|$. Sampling at twice this frequency in each direction yields the new required single-point sample area

$$a_s = \frac{\pi^2}{8|\alpha|^2}, \tag{7.33}$$

leading to a new number of required measurements in the WVBW method to be

$$M'_{WVBW} \approx \frac{16}{\pi}|\alpha|^4. \tag{7.34}$$

Again, it is necessary to apply this fine-meshed sampling to the entirety of phase-space if little is known about the state *a priori*. With this consideration, the number of required measurements becomes approximately three times that of Eq. 7.32, and may be substantially more if one wishes to measure all the maximum and minimum Wigner function fringe values without applying any flavor of maximum-likelihood or other reconstruction algorithms. The comparison between the WVBW and overlap methods for the same number of measurements is displayed in Fig. 7.4, where the Wigner function at top left is the ideal compass state Wigner function reconstructed from $(N+1)^2$ overlap measurements, with $N = 24$ and $|\alpha| = 3.5$. The Wigner function plotted at the top right of the figure shows the results of the WVBW tomography method with the same number of measurements, where the Wigner function is sampled at 25 equally spaced amplitudes each at 25 phase values. The maximum amplitude to use was determined from Eq. 7.25 using the Fock-basis truncation of $N = 24$. The lower half of the figure shows the same plots as above but zoomed in to see the central interference pattern. It is clear that although WVBW successfully catches the negativity and general features of the distribution, it fails to display the finer features and at times misses local extrema. The 3D plots in Fig. 7.5 show more clearly how even though main features are caught with the same number of measurements using the WVBW technique, many of the finer features lose definition.

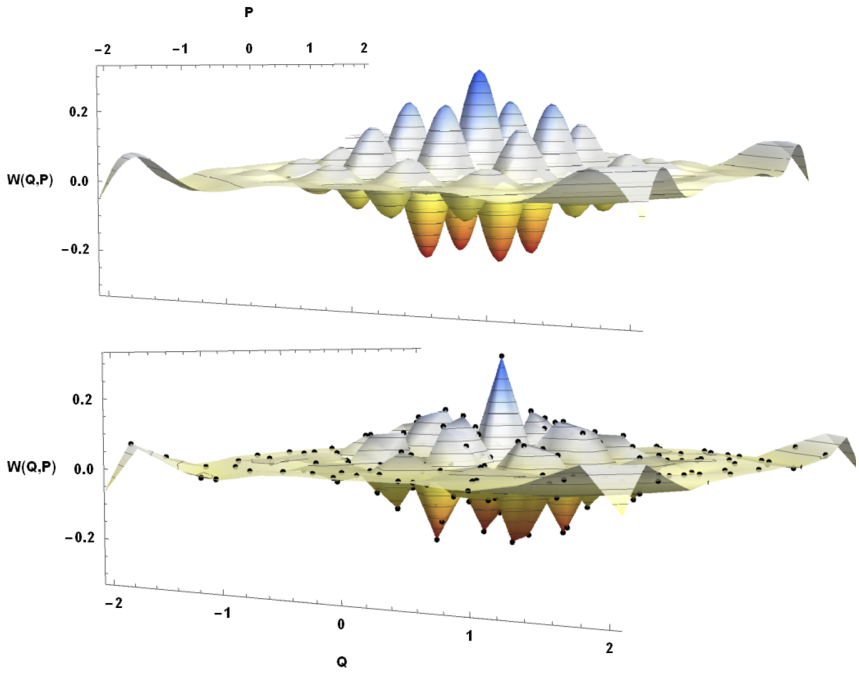



FIGURE 7.5: 3D representation of the central interference fringes in the Wigner function. Top: Wigner function from the overlap tomography method. Bottom: Wigner function from the WVBW method after sampling $d^2$ points.

## 7.3 Variations on Overlap Tomography

### 7.3.1 Click detection

The discussion of overlap tomography in Sec. 7.2 is based on the experimenter having access to parity measurements so that the Wigner function at the origin can be measured. However, this is only necessary to get the overlap measurements between the unknown Wigner function and the probing state. The real magic of the overlap method arises from the relation of the overlap measurement to the density operator given by Eq. 7.16. This overlap measurement can be achieved without parity measurements at all. In fact, even PNR measurements are not necessary, as the overlap can be obtained using a displacement and click detectors only as shown in Fig. 7.6.

This experimental implementation is exactly that of the WVBW method except that the parity detection is replaced by a detector that can only determine the difference between events where the measured number of photons was zero or nonzero. From this type of measurement, one can determine the probability to measure zero photons,

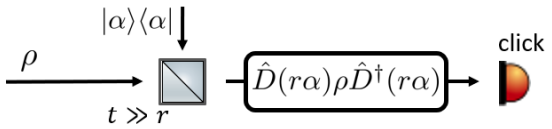



FIGURE 7.6: Experimental implementation of a displacement followed by detection with a click detector that can only differentiate between zero photons and one or more photons.

which is written as

$$P(0) = \mathrm{Tr}\left[|0\rangle\langle 0|\,\hat{D}^\dagger(\beta)\rho\hat{D}(\beta)\right], \tag{7.35}$$

where $\beta = -r\alpha$ for the experimental scheme shown in the figure.

By using the cyclic property of the trace and pushing the displacement operators into the zero photon bra and ket, one can see that this measurement is exactly that of the overlap measurement of $\rho$ with a coherent state given by the opposite of the displacement value:

$$P(0) = \mathrm{Tr}\left[|\beta\rangle\langle\beta|\,\rho\right] = \mathcal{O}(\rho, \beta). \tag{7.36}$$

From here, one can now perform several displacements and use the same SDP method with the overlap measurements to reconstruct the density matrix, and from there calculate the Wigner function.

On the surface, this method seems quite useful in performing tomography on a variety of states. In reality, this only works well for tomography of density matrices with low mean-photon number. In order to accurately determine the overlap measurement needed for reconstruction, one needs an accurate determination of the zero photon detection probabilities. However, for higher energy states which must also be displaced, the probability to detect zero photons for some measurement settings may be quite low, and it will thus be necessary to perform many measurements before an accurate overlap value is obtained.

### 7.3.2 Quadratic speed-up with full PNR

As yet another alternative take on overlap tomography, again consider using the same set-up as the WVBW method, but simply keep the photon-number measurements as they are without converting to parity (recall that current parity measurements are just obtained from PNR measurement distributions). This allows the experimenter to make

full use of the Fock-basis measurement without reducing the results to a binary even-or-odd measurement in post-processing. Now, displacing the unknown quantum state and obtaining photon-number probabilities allows for obtaining many different overlap measurements for a *single* measurement setting. In fact, the number of multiplexed overlap measurements obtained for a single 'measurement setting' is given by the upper limit on PNR detection allowed by the detector. This will lead to a quadratic speed-up in the number of measurements required to perform full state tomography.

The overlaps can be obtained by writing out the probabilities to detect a given number of photons, $k$, similarly to the zero photon probability in the previous section. Here,

$$\begin{aligned} P(k) &= \mathrm{Tr}\Big[\, |k\rangle \langle k| \, \hat{D}^{\dagger}(\beta)\rho\hat{D}(\beta)\Big] \\ &= \langle k,\beta| \, \rho \, |k,\beta\rangle = \mathcal{O}_k^{\beta}, \end{aligned} \tag{7.37}$$

where $|k,\beta\rangle = \hat{D}(\beta)\,|k\rangle$. This expression is simply the fidelity between $\rho$ and the displaced $k$-photon Fock state. Thus for a single displacement value, performing PNR detection and obtaining the photon-number probability distribution yields $N$ separate overlap measurements, where $N$ is the maximum number of photons the detector can resolve. We can now re-frame the SDP problem in terms of a new coefficient matrix.

We can write the overlaps as

$$\mathcal{O}_k^{\beta_j} = \sum_{n,m=0}^{N} c_{km}^{j} c_{kn}^{j*} \rho_{nm}, \tag{7.38}$$

where $c_{kn}^{j} = \langle m| \, \hat{D}(\beta) \, |k\rangle$ is the overlap coefficient which is given by [292]

$$\langle m| \, \hat{D}(\beta) \, |k\rangle = e^{-|\beta|^2/2} \left(\frac{k!}{m!}\right)^{1/2} \beta^{m-k} L_k^{(m-k)}(|\beta|^2), \tag{7.39}$$

and $L_k^j(x)$ are the associated Laguerre polynomials. By using several values of $\beta_j$ for displacements, we can write the matrix for a series of equations with the form of

Eq. 7.38 as

$$\begin{pmatrix} \mathcal{O}_0^0 \\ \mathcal{O}_1^0 \\ \vdots \\ \mathcal{O}_0^1 \\ \vdots \\ \mathcal{O}_N^j \end{pmatrix} = \begin{pmatrix} c_{00}^0 c_{00}^{0*} & c_{01}^0 c_{00}^{0*} & \dots & c_{0N}^0 c_{0N}^{0*} \\ c_{10}^0 c_{10}^{0*} & c_{11}^0 c_{10}^{0*} & \dots & c_{1N}^0 c_{1N}^{0*} \\ \vdots & \vdots & \ddots & \vdots \\ c_{00}^1 c_{00}^{1*} & c_{01}^1 c_{00}^{1*} & \dots & c_{0N}^1 c_{0N}^{1*} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N0}^j c_{N0}^{j*} & c_{N1}^j c_{N0}^{j*} & \dots & c_{NN}^j c_{NN}^{j*} \end{pmatrix} \begin{pmatrix} \rho_{0,0} \\ \rho_{0,1} \\ \vdots \\ \rho_{N,N} \end{pmatrix}, \tag{7.40}$$

which has the familiar expression for the matrix equation we found earlier of $\mathbf{O} = \mathbf{CP}$. Here, the dimensions will depend not only on the number of displacement values used but also on the resolution capability of the PNR detector. As before $\mathbf{P} \in \mathbb{C}^{(N+1)^2}$, but now we have that $\mathbf{O} \in \mathbb{R}^{j(N+1)}$ and $\mathbf{C} \in \mathbb{C}^{j(N+1)\times(N+1)^2}$. In order to have enough independent parameters to solve the matrix inversion problem, we now only require that the number of measurement settings is $j \geq (N+1)$, which substantially reduces the experimental burden for the number of required measurements compared to the previous version of overlap tomography, where $j \geq (N+1)^2$ coherent states were needed. Using more than the minimum required measurements simply increases the number of rows in the coefficient matrix which will help mitigate the effects of noise and random measurement errors when the SDP problem of Eq. 7.22 is solved. Note that there must be a minimum of three measurement settings, even in the case where the detector can measure many more photons that are possibly contained in $\rho$. The case where $\beta = 0$ will only give information about the main diagonal of $\rho$, and purely real $\beta$ will not yield information on the complex-valued parameters in $\rho$. Thus there must always be nonzero, real, and imaginary values for the displacement.

### Multi-mode systems

Suppose one has a multi-mode system and the ability to perform independent displacements and PNR measurements on each mode. By taking advantage of the multiplexed nature of the overlap measurements obtained by PNR detectors, one can achieve a quadratic reduction in the number of 'measurement settings' required to fully tomography an unknown multi-mode system. If each of the $M$ modes can be expressed in a Hilbert space with photon-number cutoff of $N$ (and each PNR detector has at least $N$-photon resolution), then the full multi-mode density matrix will have $(N+1)^{2M} - 1$

elements and can be expressed as

$$\rho = \sum_{\boldsymbol{n},\boldsymbol{m}}^{N} \rho_{\boldsymbol{n},\boldsymbol{m}} \left|\boldsymbol{n}\right\rangle \left\langle \boldsymbol{m}\right| , \tag{7.41}$$

where $\boldsymbol{n}, \boldsymbol{m} \in \mathbb{Z}^M$ where each element has a maximmum cutoff of $N$, and $|\boldsymbol{n}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes ... \otimes |n_M\rangle$. If each mode, $j$, has an associated PNR detector that yields a measurement result of $k_j$, then a given multimode detection probability for all $k_j$ is

$$P(\boldsymbol{k}) = \mathrm{Tr}\left[\left|\boldsymbol{k}\right\rangle \left\langle \boldsymbol{k}\right| \hat{D}(\boldsymbol{\alpha})\rho\hat{D}^{\dagger}(\boldsymbol{\alpha})\right] = \mathcal{O}_{\boldsymbol{k}}^{\boldsymbol{\beta}_j}, \tag{7.42}$$

where $\boldsymbol{\alpha} \in \mathbb{C}^M$ is the vector defining the $M$-mode displacement and $\boldsymbol{\beta} = -\boldsymbol{\alpha}$. For each displacement setting, assembling the probability distribution from the collected statistics will yield $(N+1)^M$ distinct probability measurements, which are exactly the overlap values necessary for tomography. Thus, instead of needing $(N+1)^{2M}$ measurement settings, only $(N+1)^M$ displacement settings are necessary, leading to a quadratic speed-up in tomography over prior methods. The origin of this speed-up can be seen as taking advantage of the quantum nature of the detector, which allows for a full Fock-basis measurement over all $M$ modes, which also scales exponentially in the number of detectors.

In requiring fewer measurement settings, the experimenter additionally gains a reduction in the number of calibration measurements needed. At each measurement setting, the values of the coherent state displacements must be precisely determined. This can be obtained by blocking the input signal and only allowing the displacement field to impinge upon the PNR detectors. The accuracy of this calibration directly contributes to the error in the tomographic reconstruction; thus it is important to take many samples to obtain precise values of $\alpha$. The number of samples needed for the actual overlap measurements will be discussed next.

**Finite sampling error**

In order to collect the overlap values, it is necessary to sample a probability distribution, which requires many single-shot measurements to acquire sufficient statistics and reduce measurement error. This is a well-known fact of experimental physics, but it is often neglected when considering analyses as to the theoretical efficiency of a given

protocol. Fortunately, the quadratic reduction in the number of measurement settings *also* translates to a quadratic reduction in the number of physical samples taken. This can be seen by realizing that in both cases of overlap from parity or overlap from probabilities, the measurements are obtained by post-processing PNR detection statistics.

Suppose we send $S$ laser pulses to a PNR detector that results in outcome $n$ with maximum photon-number detection of $M$. This can be seen as measuring $S$ independent and identically distributed (iid) random variables, $X$. The probability of event $k$, $P(X = k)$, can be written as a scaled sum of all outcomes of *different* iid random variables, $Y$, where the probability of outcome $k$ is the same, $P(X = k) = P(Y = k) \equiv P(k)$, but all other outcomes are zero, $P(Y \neq k) = 0$. In this way, we can write

$$P(X = k) = \frac{1}{S} \sum_{s}^{S} \frac{y_s}{k}, \tag{7.43}$$

where $y_s$ is the sampled result of the $s$-th iid random variable, $Y$. The variance of this expression is given by

$$\text{Var}[P(X = k)] = \text{Var}\left[\frac{1}{S} \sum_{s}^{S} \frac{y_s}{k}\right] = \frac{1}{S^2 k^2} \sum_{s}^{S} \text{Var}\left[y_s\right], \tag{7.44}$$

where the last line follows because each random sampling event is independent from all other events. Because the random variables $Y$ each have a distribution that is only either 0 or $k$, the variance of each $y_s$ is

$$\text{Var}\left[y_s\right] = \sum_{n=0}^{M} n^2 P(Y = n) - \left(\sum_{n=0}^{M} n P(Y = n)\right) \tag{7.45}$$

$$= k^2 \left(P(k) - P(k)^2\right)^2. \tag{7.46}$$

Since there are $S$ total events in the sum, the variance of the measured probability is

$$\text{Var}[P(X = k)] \equiv \Delta_k^2 = \frac{P(k) - P(k)^2}{S}. \tag{7.47}$$

Thus, the uncertainty $\Delta_k$ scales as $S^{-1/2}$, which is the same scaling as the standard error on the mean, which might have been expected. Since each probability is less than one, we can at worst estimate the error as $\Delta_k \approx S^{-1/2} = \Delta$. If one takes enough samples such that the measured probability is near the true probability (so we can use the measured

probability, $P_{meas}(k) \approx P(k)$ to calculate the variance), then the overlap measurements become $\mathcal{O}_{\boldsymbol{k}}^{\boldsymbol{\beta}_j} \pm \Delta$. For each measurement setting, $S$ measurement samples then result in $(N+1)^M$ total overlap measurements, each with error $\Delta$.

Now, consider the error to each overlap measurement obtained by calculating the parity from Eq. 7.9. In this case, we can write the variance of the expectation of parity as

$$\text{Var}\left[\langle(-1)^{\hat{n}}\rangle\right] = \text{Var}\left[\frac{1}{S}\sum_{s}^{S}(-1)^{x_s}\right], \tag{7.48}$$

where $x_s$ is a PNR measurement result independent of all other $x_s$. To leading order, the variance of a function of a random variable can be expanded as[293]

$$\text{Var}\left[f(X)\right] \approx \text{Var}[X]\left(f'(X)|_{\text{E}[X]}\right)^2, \tag{7.49}$$

where we can write the derivative as

$$f'(X) = \frac{d}{dX}(-1)^X = i\pi(-1)^X. \tag{7.50}$$

For $S$ total measurement samples, the variance on the parity is

$$\text{Var}\left[\langle(-1)^{\hat{n}}\rangle\right] \approx \frac{\sigma^2}{S}|i\pi(-1)^{\mu}|^2 = \frac{\pi^2\sigma^2}{S}, \tag{7.51}$$

where $\sigma$ is the standard deviation on the measured photon number and $\mu$ is the mean. Thus, $S$ samples with the parity method of Eq. 7.9 also yields an overlap $\mathcal{O} \pm \Delta$ where $\Delta \sim O(S^{-1/2})$, yet only a *single* overlap measurement is obtained for each measurement setting. The new overlap method that uses probabilities as per Eq. 7.42 results in a similar accuracy on each overlap measurement as in the parity case for the same number of measurements, $S$, but instead yield an exponentially increased $(N+1)^M$ number of measurements. Thus, the quadratic reduction in the number of measurement settings extends to the true number of measurements required to perform overlap tomography to a similar level of accuracy.

## 7.4 Loss Compensation

A major experimental annoyance stems from imperfect detectors as touched on in the previous chapter. Fortunately, some of these effects can be accounted for when performing tomography by characterizing the statistical effects of the imperfections and applying a map back to what would be a perfect state. An example of this is corrections for loss on a measured photon number distribution. Suppose one has a pure Fock state, $|n\rangle$, and sends it to a PNR detector where the only imperfection is efficiency $\eta < 1$. Instead of registering $n$ photons every time, the PNR detector will actually measure a distribution according to its POVM. Assuming the detector can resolve at least $n$ photons, the resultant probability distribution for measuring $k$ photons will be

$$P(k) = \binom{n}{k} \eta^k (1-\eta)^{n-k}. \tag{7.52}$$

When $\eta \to 1$, the detection probability goes to zero for any event other than $k = n$, and as $\eta \to 0$, only the $k = 0$ term survives, which corresponds to losing all of the $n$ photons. This is equivalent to sending $|n\rangle$ through a beamsplitter of reflectivity $r = \sqrt{1-\eta}$ (where the second mode is vacuum) and tracing out over the reflected mode while sending the transmitted mode to a perfect PNR detector.

Perhaps surprisingly, this pure loss is a linear mapping and thus unique. That is to say, for a given $\eta$, each photon number distribution before loss maps to a unique distribution after loss. Thus if one has accurately calibrated the efficiency of a PNR detector, this loss effect can be accounted for and 'undone' in post-processing by applying the inverse map as has been demonstrated previously [112, 294]. The calibration of $\eta$ could be achieved, for example, by sending a known coherent state, $|\alpha\rangle$ to the detector and measuring the mean photon number. Under pure loss, the coherent state becomes $|\alpha\rangle \to |\sqrt{\eta}\alpha\rangle$, so the experimenter will measure a mean photon number of $\eta|\alpha|^2$ and can directly extract $\eta$ to a high degree of accuracy.

For tomography with PNR detectors, one could in principle apply this statistical correction after each measured distribution. However, since many individual photon number distributions are measured, this may become cumbersome. Alternatively, as discussed next, one can apply a similar inverse mapping to account for loss on the entire density matrix at once.

### 7.4.1 Complete density matrix correction

We now wish to correct an arbitrary density matrix given a known loss. In this case, we have experimentally measured $\rho'$, but our goal is to reconstruct the density matrix before the loss, $\rho$. As shown in Fig. 7.7, this can be modeled by sending $\rho$ through a fictitious 'loss beamsplitter' with reflection and transmission coefficients of $r = \sqrt{1-\eta}$ and $t = \sqrt{\eta}$, where $\eta$ is the overall transmission efficiency.

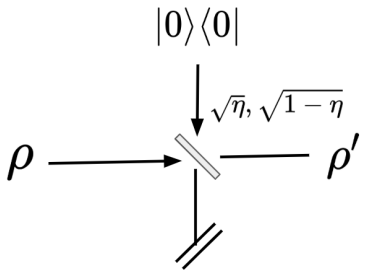



FIGURE 7.7: Lossy channel.

The general quantum state density matrix before the loss can be written as

$$\rho = \sum_{n,n'=0}^{\infty} \rho_{n,n'} |n\rangle\langle n'|$$
$$\rho = \sum_{n,n'=0}^{\infty} \frac{\rho_{n,n'}}{\sqrt{n!n'!}} a^{\dagger n} |0\rangle\langle 0| a^{n'}. \tag{7.53}$$

If this state enters into the loss beamsplitter in mode $\hat{a}$ with vacuum in mode $\hat{b}$, then the mode operators transform in the Heisenberg picture according to $\hat{a} \to t\hat{a} + r\hat{b}$ and $\hat{b} \to -r\hat{a} + t\hat{b}$ to yield an output density matrix

$$\rho_{out} = \sum_{n,n'=0}^{\infty} \rho_{n,n'} \frac{(t\hat{a}^\dagger + r\hat{b}^\dagger)^n}{\sqrt{n!}} |0\rangle_a |0\rangle_b \langle 0|_b \langle 0|_a \frac{(t\hat{a} + r\hat{b})^{n'}}{\sqrt{n'!}}. \tag{7.54}$$

Tracing out over mode $b$ yields the final state after loss, which is given by

$$\rho' = Tr_b[\rho_{out}] = \sum_{n,n'=0}^{\infty} \rho_{n,n'} \sum_{k=0}^{n} \sum_{k'=0}^{n'} A_{n,n',k,k'} |n-k\rangle\langle n'-k'| \langle k|k'\rangle \delta_{k,k'}, \tag{7.55}$$

where

$$A(n,n',k,k') = \sqrt{\binom{n}{k}\binom{n'}{k'}} r^{k+k'} t^{n+n'-k-k'}. \tag{7.56}$$

Substituting $n-k$ and $n'-k$ with $m$ and $m'$ allows us to rearrange the expression and write a sum over the Fock components in order, which can be written as

$$\rho' = \sum_{m,m',k=0}^{\infty} \rho_{(m+k),(m'+k)} A(m+k, m'+k, k, k) |m\rangle\langle m'|, \tag{7.57}$$

where it is easy to see that each element of the density matrix after loss is related to the original state by

$$\rho'_{m,m'} = \sum_{k=0}^{\infty} \rho_{(m+k),(m'+k)} \sqrt{\binom{m+k}{k}\binom{m'+k}{k}} r^{2k} t^{m+m'}. \tag{7.58}$$

This can be viewed as a generalized Bernoulli distribution [295], so can be inverted to read

$$\rho_{m,m'} = \sum_{k=0}^{\infty} \rho'_{(m+k),(m'+k)} \sqrt{\binom{m+k}{k}\binom{m'+k}{k}} (-1)^k \left(\frac{r}{t}\right)^{2k} t^{-m-m'}. \tag{7.59}$$

In practice, the sum over $k$ can be truncated to some value, $N_{max}$, beyond which the entries in the initial density matrix are negligible. We can then reformulate Eq. 7.58 as a series of $N_{max}$ linear maps from the $i^{th}$ diagonal of $\rho'$ to the $i^{th}$ diagonal of $\rho$, where the main diagonal is given by $i = 0$. This mapping can be visualized in Fig. 7.8. Each of these linear maps, $\mathbf{M}^{(i)}$, is an upper triangular matrix of dimension $N_{max} - i \times N_{max} - i$ with elements

$$\mathbf{M}^{(i)}_{jk}(\eta) = \begin{cases} 0 & j > k \\ \sqrt{\binom{k}{k-j}\binom{i+k}{k-j}} (1-\eta)^{(k-j)} \eta^{\frac{i}{2}+j} & \text{otherwise} \end{cases} \tag{7.60}$$

Since each $\mathbf{M}^{(i)}(\eta)$ is triangular with nonzero diagonal elements, the inverse mappings can be found by inverting the generalized Bernoulli transformation and are given by [295]

$$\text{Inv}[\mathbf{M}^{(i)}(\eta)] = \mathbf{M}^{(i)}(\eta^{-1}). \tag{7.61}$$

The existence of this inversion is due to the known well defined statistical nature of loss channel, which makes it possible to perfectly reconstruct any $\rho$ within a finite-dimensional Hilbert space when $\eta$ and $\rho'$ are precisely known [295]. However, the presence of any small deviations in an experimentally measured $\rho'$ can lead to unphysically large or non-positive diagonal density matrix elements in the reconstruction of $\rho$,

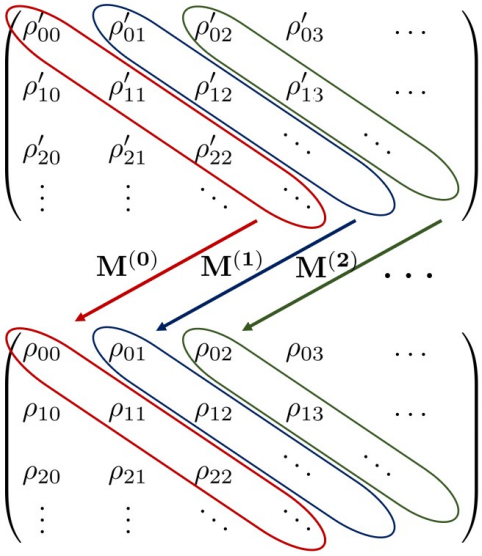



FIGURE 7.8: Each diagonal of a density matrix subjected to loss is uniquely mapped to a new, loss-degraded diagonal in $\rho'$. This mapping can be undone by applying the inverse mapping.

even while $\rho$ remains normalized, which is similar to the possible numerical instabilities that arise when using pattern-functions [274]. These errors become pronounced for low detector efficiencies at high photon-numbers as seen Fig. 7.9a for the specific case of a loss-compensated cat state. Therefore, it becomes extremely crucial to have *a priori* information about the energy of the quantum states under consideration.

Here, we are able to relax this issue by inverting each $\mathbf{M}^{(i)}(\eta)$ using semidefinite programming, where the optimization problem is defined as

$$\begin{aligned} &\underset{\rho}{\text{Minimize}} \quad \sum_{i=0}^{N_{max}} ||\rho'^{(i)}_{diag} - \mathbf{M}^{(i)}\rho^{(i)}_{diag}||_2 \\ &\text{Subject to} \quad \rho \geq 0, \ \text{Tr}[\rho] = 1, \text{ and } \ \rho_{m,m} \leq \eta^{-m}\rho'_{m,m}, \end{aligned} \tag{7.62}$$

where $\rho^{(i)}_{diag}$ denotes the $i^{th}$ diagonal of $\rho$ and the third constraint is obtained by noting that each element in the sum in Eq. (7.59) is positive for $m = m'$, leading to the inequality when the sum is truncated after the first term. Additionally, it is only necessary to sum over the upper diagonals of $\rho$ in the minimization (hence the sum starting at $i = 0$), due to the enforced hermiticity of $\rho$.

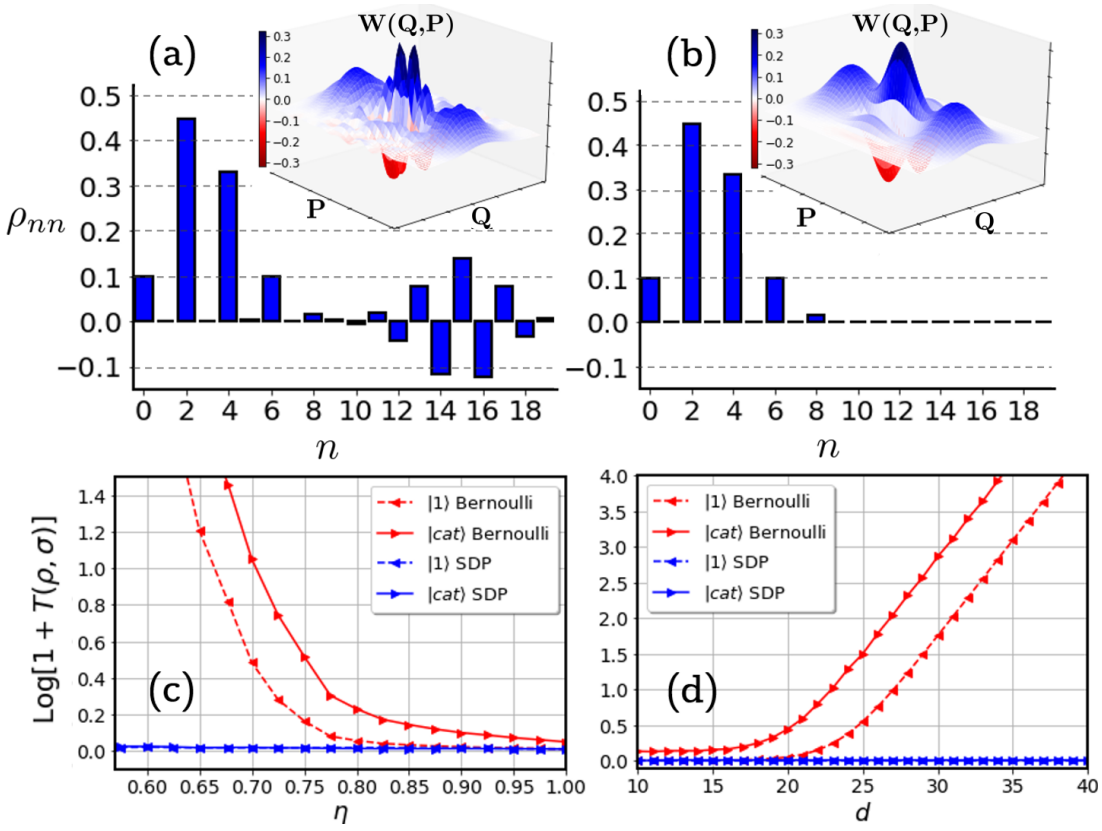



FIGURE 7.9: Loss-compensation for tomographed cat state of amplitude $\sqrt{3}$ after transmission of $\eta = 0.70$ and Hilbert space cut-off of $d = 20$, with (a) inversion using the generalized Bernoulli transformation and (b) inversion using SDP. The logarithm of the trace-distance between the reconstructed state, $\rho$, and the target state, $\sigma$, is plotted against $\eta$ [$d$] in (c) [(d)]. We note that $T(\rho, \sigma) > 1$ occurs due to the unphysical reconstruction of $\rho$ and large non-positive diagonal elements. The figure insets show the Wigner function for each state.

The application of these constraints enforces physicality and avoids the numerically unstable reconstruction that would result by using an exact expression for $\mathbf{M}^{(i)}(\eta)^{-1}$. We demonstrate in Fig. 7.9 how small errors on density matrix elements from performing the tomographic procedure on a loss degraded cat state and single-photon Fock state give rise to an unphysical loss-compensated state using the analytical matrix inversion from Ref. [295], whereas inversion using SDPs successfully reconstructs the state prior to loss. Although all errors in the tomographed density matrix elements prior to loss compensation are on the order of $10^{-3}$ (not depicted), the analytical matrix inversion drastically magnifies these slight deviations. In particular, Fig. 7.9c and Fig. 7.9d show that the validity of the loss-compensation can heavily depend on both

the overall loss and on the choice of Hilbert dimension cutoff. When comparing the reconstructed state, $\rho$, to the ideal state without loss, $\sigma$, using the trace distance defined by $T(\rho,\sigma) = \frac{1}{2}||\rho - \sigma||_1$, we see that the deviation of $\rho$ from $\sigma$ grows quickly as $\eta$ shrinks and $d$ increases in the case of analytic inversion. However, $T(\rho,\sigma)$ is both small and relatively independent of either $\eta$ or $d$ when using SDP. As a result, our method is significantly more robust to experimental noise.

### 7.4.2 Equivalence of photon-number distributions

To show the equivalence of the photon-number distributions measured in each configuration in Fig. 7.10, we adapt the approach originally introduced in [275]. The signal and LO modes are described by annihilation operators $\hat{a}$ and $\hat{b}$ respectively, and $\hat{c}_v$ and $\hat{d}_v$ are vacuum modes. For a perfect PNR detector, the probability of measuring $n$ photons

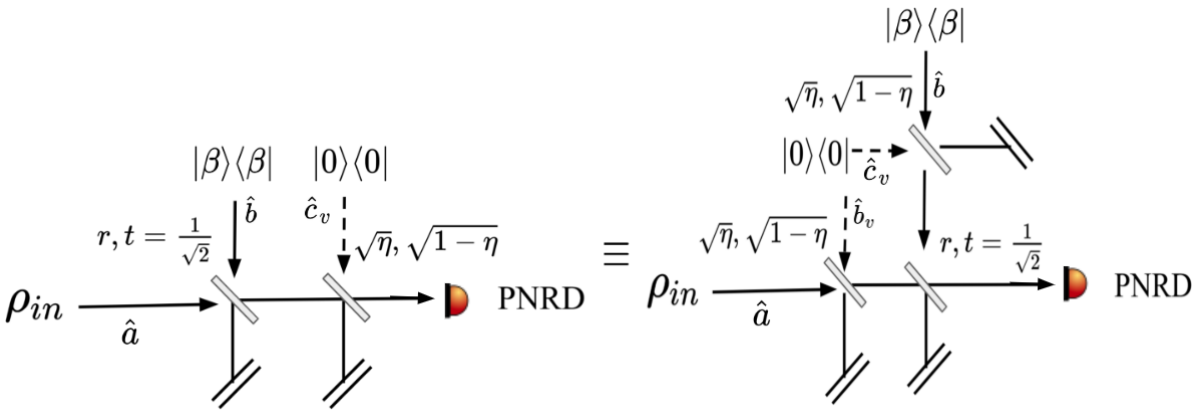



FIGURE 7.10: Schematic of the loss model. Left and right networks produce the same photon-number distribution.

is given by [296]

$$P(N=n) = \left\langle : \frac{\hat{N}^n}{n!} e^{-\hat{N}} : \right\rangle_{\rho_{in}}, \tag{7.63}$$

where $\hat{N} = \hat{d}^\dagger \hat{d}$ is the photon-number operator of the detection mode and the expectation value is calculated over the initial states, and :: is the normal ordering. By employing the Heisenberg picture, we first determine the detection mode in terms of input modes for the network on the left of Fig. 7.10. The input mode denoted by annihilation

operator, $\hat{a}$, evolves to

$$\text{After first beamsplitter: } \hat{a} \rightarrow \frac{\hat{a}+\hat{b}}{\sqrt{2}} \tag{7.64}$$

$$\text{After second beamsplitter: } \sqrt{\eta}\left(\frac{\hat{a}+\hat{b}}{\sqrt{2}}\right)+\sqrt{1-\eta}\hat{c}_v \tag{7.65}$$

Since the input states for mode $\hat{b}$ and $\hat{c}_v$ are coherent and vacuum states respectively, the normal ordering allows to treat them as complex numbers. As a result, the effective photon-number operator is given by

$$\hat{N}^L_{\text{eff.}} = \hat{d}^\dagger \hat{d}, \tag{7.66}$$

where the detection mode is

$$\hat{d}^L = \sqrt{\eta}\left(\frac{\hat{a}+\beta}{\sqrt{2}}\right) \tag{7.67}$$

Likewise, for the right network, we have

$$\text{After left beamsplitter: } \hat{a} \rightarrow \sqrt{\eta}\hat{a}+\sqrt{1-\eta}\hat{b}_v \tag{7.68}$$

$$\text{After top beamsplitter: } \hat{b} \rightarrow \sqrt{\eta}\hat{b}+\sqrt{1-\eta}\hat{c}_v \tag{7.69}$$

$$\text{After balanced beamsplitter: } \hat{d} = \frac{1}{\sqrt{2}}(\sqrt{\eta}\hat{a}+\sqrt{1-\eta}\hat{b}_v+\sqrt{\eta}\hat{b}+\sqrt{1-\eta}\hat{c}_v),$$

where $\hat{c}_v$, $\hat{b}_v$ are vacuum modes and $\hat{b}$ is a coherent state. We again utilize the fact that normal ordering allows coherent states to be represented by a complex number and the vacuum state can also be considered as a coherent state with zero amplitude. Thus, the detection mode can be further simplified as

$$\hat{d}^R = \sqrt{\eta}\left(\frac{\hat{a}+\beta}{\sqrt{2}}\right). \tag{7.70}$$

From Eq. 7.67 and Eq. 7.70, one can see that both networks have the same detection mode, therefore would produce the same photon-number distribution for a given quantum state under investigation.

## 7.5 Applications to Quantum Field Theory

Near the inception of using quantum systems to perform computations, Feynman proposed using quantum computers as simulators of quantum physics [22]. He rather bluntly, but succinctly, stated the problem as: "Nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy."

He was clearly right, and evidently, it hasn't been easy thus far. However, there has been extensive progress on the theoretical development of quantum algorithms for quantum simulation, and there have also been experimental advances such as applied to quantum chemistry [297–299]. As far as quantum field theory applications, Jordan, Lee, and Preskill (JLP) developed a quantum algorithm to compute relativistic scattering amplitudes in a massive quantum field theory with quartic interactions. The algorithm is polynomial in the number of scattering particles and has the potential to yield an exponential speed-up over classical algorithms in the strong-coupling and high precision regimes [300].

However, the JLP algorithm requires a qubit register than discretizes both space (position) and field amplitude. If one instead purses a CV implementation, such as with quantum optics, the field amplitude can be left as a continuous parameter and can be described by a single qumode instead of a register of qubits. The JLP algorithm was translated to CV in just this way by Marshall et al. [301]. This approach to simulating scattering amplitudes uses squeezing to effectively confer mass in a rotated massive bosonic particle basis. The CV algorithm as proposed is depicted in Fig. 7.11.

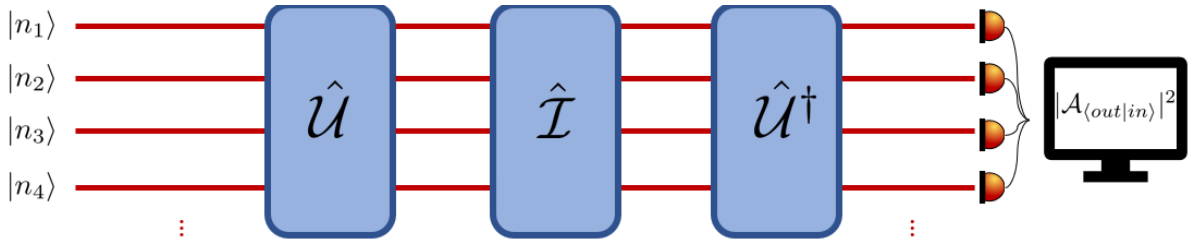



FIGURE 7.11: Quantum algorithm to simulate relativistic scattering of massive bosonic fields using a CV implementations.

One first prepares a series of Fock states $|n_1\rangle, |n_2\rangle, ..., |n_M\rangle$. These Fock states will eventually correspond to excited states of the massive bosonic field we wish to simulate. The unitary operator that maps the photonic basis to the QFT basis, $\hat{u}$, is

dependent upon the number of qumodes used in the simulation. This operator can be decomposed into three steps, $\hat{\mathcal{U}} \equiv \hat{U}_3\hat{U}_2\hat{U}_1$. Supposing that we have $M$ qumodes, the first operation acts to map the photon-number creation and annihilation operators according to

$$\begin{aligned}\hat{U}_1\hat{a}_0\hat{U}_1^\dagger &= \sum_{k=0}^{M-1}\hat{a}_k,\\ \hat{U}_1\hat{a}_m\hat{U}_1^\dagger &= \sum_{k=0}^{M-1}\cos\tfrac{2\pi mk}{M}\hat{a}_k,\\ \hat{U}_1\hat{a}_{M-m}\hat{U}_1^\dagger &= \sum_{k=0}^{M-1}\sin\tfrac{2\pi mk}{M}\hat{a}_k,\end{aligned} \tag{7.71}$$

where $1 \leq m \leq M/2$. This rotation, being Gaussian (and passive), can be decomposed into two-mode beamsplitter operations and phase shifts alone. Next, single-mode squeezing operations are applied to every mode to confer mass,

$$\hat{U}_2 = \prod_{m=0}^{M-1} e^{\frac{r_m}{2}\left(\hat{a}_m^2 - \hat{a}_m^{\dagger 2}\right)}, \tag{7.72}$$

where $e^{2r_m} = \omega_m$ for $m \leq M/2$ and $= e^{-2r_m} = \omega_m$ for $m > M/2$. Here, the oscillation frequency is related to the mass, $\mu$, in the scalar field theory, with $\omega_m^2 = \mu^2 + 4\sin^2\frac{m\pi}{M}$. Finally, the last operator is identity on the zeroth mode, and a series of balanced beamsplitters on all other modes whose actions is given by

$$\begin{aligned}\hat{U}_3\hat{a}_m\hat{U}_3^\dagger &= \tfrac{1}{\sqrt{2}}\left(\hat{a}_m + i\hat{a}_{M-m}\right),\\ \hat{U}_3\hat{a}_{M-m}\hat{U}_3^\dagger &= \tfrac{1}{\sqrt{2}}\left(i\hat{a}_m + \hat{a}_{M-m}\right).\end{aligned} \tag{7.73}$$

Together, these three operators form the Gaussian transformation that brings the photon-number creation operators of the Fock-basis inputs to the creation operators of excitations of the massive scalar field. From here, an interaction operation, $\hat{\mathcal{G}}$ can be applied to simulate scattering or other interactions. This operator is necessarily non-Gaussian for it to be of any interest for a quantum speed-up, and it involves the quartic phase gate as a means to simulate a quartic self-interaction in the implementation outlined in Ref. [301]. After this non-Gaussian operation is applied, whatever it may be, the operator $\hat{\mathcal{U}}^\dagger \equiv \hat{U}_1^\dagger\hat{U}_2^\dagger\hat{U}_3^\dagger$ must be applied to rotate back to the photon basis. Finally, performing a PNR detection on all modes and obtaining the probability distribution

from a given set-up allows one to experimentally measure the modulus-squared value of the scattering amplitudes. However, we would like to obtain the complex-valued scattering amplitude in each case. We detail below an extension of the method detailed by Marshall et al. that will overcome this difficulty and allow full access to the desired quantities.

### 7.5.1 Phase-sensitive simulation of scattering amplitudes

Following a quantum computation, we would like to determine the scattering amplitude, $\mathcal{A}$, that is given by projecting our evolved state onto a measurement operator. However, measuring the final state in the photon-number basis only yields statistics that give us probabilities, which lead to $|\mathcal{A}|^2$. To get around this difficulty, we can perform a series of weak displacements to the state before counting photons and use these measurements to construct $\mathcal{A}$.

Scattering amplitudes of interest to quantum field theory can potentially be simulated with continuous-variable quantum computation (CVQC) by preparing a series of Fock states as an input state, performing unitary operations, and then projecting the resultant state in the Fock basis with photon-number-resolving measurements [301]. The resultant scattering amplitudes appear as

$$\mathcal{A} = \langle \text{out} | \, \mathcal{U} \, | \text{in} \rangle \tag{7.74}$$

Where $|\text{in}\rangle$ is the input state to the scatting process, $\mathcal{U}$ is the collection of operations performed by the computation, and $|\text{out}\rangle$ is the scattered state. Fortunately, in the context of the proposal by Marshall et al., both the input and output states can be written as a tensor product of Fock states follow by a unitary transformation. In this case, the scattering amplitude can be rewritten as

$$\mathcal{A} = \langle n_N | \otimes \ldots \langle n_2 | \otimes \langle n_1 | \, \mathcal{U}' \, | m_1 \rangle \otimes | m_2 \rangle \ldots \otimes | m_N \rangle \,, \tag{7.75}$$

where $N$ is the number of physical modes that are input and then detected. Suppose we start with the simplest case, where there is only a single physical mode, in which case we have that

$$\mathcal{A}_n = \langle n | \, \psi \rangle, \tag{7.76}$$

where $|\psi\rangle = \mathcal{U}' |m\rangle$. By preparing many copies of $|\psi\rangle$ and performing PNR detection, we can collect statistics to determine the detection probabilities - in other words, determine $|\mathcal{A}_n|^2$ for each $n$ - but this does not give us access to the full complex number that is $\mathcal{A}$. To remedy this, we can examine the effects of displacing $|\psi\rangle$ before measurement.

The displacement operator, defined as

$$\hat{D}(\alpha) = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}}, \tag{7.77}$$

can be written as $\hat{D}(\alpha) = \hat{\mathbb{1}} + \alpha \hat{a}^\dagger - \alpha^* \hat{a} + \mathcal{O}(\alpha^2)$ for small displacements. The detection probabilities after displacement, which we do have access to, are given by

$$P_n(\alpha) = |\langle n| \hat{D}(\alpha) |\psi\rangle|^2. \tag{7.78}$$

Using the Heisenberg Picture, the displacement in the above equation can be applied to the detection result instead of the state $|\psi\rangle$, to be rewritten as

$$\begin{aligned} P_n(\alpha) =& |\langle n|\psi\rangle + \alpha\sqrt{n}\langle n-1|\psi\rangle - \alpha^*\sqrt{n+1}\langle n+1|\psi\rangle + \mathcal{O}(\alpha^2)|^2 \\ =& |\mathcal{A}_n|^2 + 2\sqrt{n}\mathrm{Re}[\alpha^* \mathcal{A}_n \mathcal{A}_{n-1}^*] - 2\sqrt{n+1}\mathrm{Re}[\alpha \mathcal{A}_n \mathcal{A}_{n+1}^*], \end{aligned} \tag{7.79}$$

where in the last line we have neglected $\mathcal{O}(\alpha^2)$ terms. Suppose we let $\alpha \in \mathcal{R}$. Now, with just three measurement settings, we can get:

$$P_n(0) = |\mathcal{A}_n|^2 \tag{7.80}$$

$$P_n(\alpha) = P_n(0) + 2\alpha(\sqrt{n}\mathrm{Re}[\mathcal{A}_n \mathcal{A}_{n-1}^*] - \sqrt{n+1}\mathrm{Re}[\mathcal{A}_n \mathcal{A}_{n+1}^*]) \tag{7.81}$$

$$P_n(i\alpha) = P_n(0) + 2\alpha(-\sqrt{n}\mathrm{Im}[\mathcal{A}_n \mathcal{A}_{n-1}^*] + \sqrt{n+1}\mathrm{Im}[\mathcal{A}_n \mathcal{A}_{n+1}^*]). \tag{7.82}$$

At first glance, it seems the above equations may be difficult to solve. Things are simplified greatly, however, when one writes out the zero photon detection probabilities for each displacement setting and makes use of the fact that the annihilation operator acting on vacuum is zero. Thus $P_0(0)$, $P_0(\alpha)$, $P_0(i\alpha)$, and $P_1(0)$ are sufficient to entirely determine $\mathcal{A}_0$ and $\mathcal{A}_1$, up to a global phase. For a maximum photon detection of $M$ photons, these $3M$ equations can be solved to obtain all $\mathcal{A}_0, \mathcal{A}_1, ....\mathcal{A}_M$ scattering amplitudes.

**Multi-mode Extention**

The above results can be easily generalized to an $N$-mode system, where local weak displacements of $0, \alpha$, and $i\alpha$ are applied in every combination to each mode before detection. In this case, the multi-mode detection probabilities can be written as

$$P_{\boldsymbol{n}}(\boldsymbol{\alpha}) = |\langle \boldsymbol{n}| \prod_j \hat{D}(\alpha_j) |\psi\rangle|^2, \tag{7.83}$$

where the PNR detector of the $j^{\text{th}}$ mode detects $n_j$ photons, $|\boldsymbol{n}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes ... \otimes |n_N\rangle$, and $\boldsymbol{\alpha} \in \mathbb{C}^M$ is the vector defining the $M$-mode displacement. If only one mode is displaced at a time, that is, $\boldsymbol{\alpha} = [0, ..., \alpha_j, ..., 0] \equiv \boldsymbol{\alpha}_j$, then to leading order in $\alpha_j$, the probability is given by

$$P_{\boldsymbol{n}}(\boldsymbol{\alpha}) \approx |\langle \boldsymbol{n}| \psi\rangle + \alpha_j \sqrt{n_j} \langle \boldsymbol{n} - \boldsymbol{e}_j| \psi\rangle - \alpha_j^* \sqrt{n_j + 1} \langle \boldsymbol{n} + \boldsymbol{e}_j| \psi\rangle|^2, \tag{7.84}$$

where $\boldsymbol{e}_j$ is the unit vector along component $j$. Similarly to the single mode case above, if each mode is now displaced one at a time by $0$, $\alpha$, and $i\alpha$ for $\alpha \in \mathbb{R}$, the measured probabilities will be (to leading order in $\alpha$)

$$P_{\boldsymbol{n}}(\boldsymbol{0}) = |\mathcal{A}_{\boldsymbol{n}}|^2 \tag{7.85}$$

$$P_{\boldsymbol{n}}(\boldsymbol{\alpha}_j) = P_{\boldsymbol{n}}(\boldsymbol{0}) + 2\alpha \left( \sqrt{n_j} \mathrm{Re}[\mathcal{A}_{\boldsymbol{n}} \mathcal{A}^*_{\boldsymbol{m}-\boldsymbol{e}_j}] - \sqrt{n_j + 1} \mathrm{Re}[\mathcal{A}_{\boldsymbol{n}} \mathcal{A}^*_{\boldsymbol{n}+\boldsymbol{e}_j}] \right) \tag{7.86}$$

$$P_{\boldsymbol{n}}(i\boldsymbol{\alpha}_j) = P_{\boldsymbol{n}}(\boldsymbol{0}) + 2\alpha \left( -\sqrt{n_j} \mathrm{Im}[\mathcal{A}_{\boldsymbol{n}} \mathcal{A}^*_{\boldsymbol{n}-\boldsymbol{e}_j}] + \sqrt{n_j + 1} \mathrm{Im}[\mathcal{A}_{\boldsymbol{n}} \mathcal{A}^*_{\boldsymbol{n}+\boldsymbol{e}_j}] \right). \tag{7.87}$$

For an $N$-mode system with maximum PNR resolution of $M$ photons on each mode, displacing in this way leads to $2N + 1$ measurement settings which give rise to $(2N + 1)^M$ equations that can be solved to determine the $2N^M$ real parameters associated with the possible scattering amplitudes. A point of note is that while the number of real parameters we wish to find scales exponentially with the number of modes, the number of measurement settings needed to determine these parameters is linear. Thus, we maintain the quantum speed-up of the original algorithm.

The ability to use displacements to access the full complex-valued scattering amplitudes can be thought of as a partial tomographic reconstruction. Although we do not wish to entirely reconstruct the multi-mode quantum state, we are interested in measurements with phase sensitivity. PNR detectors alone are unable to collect any phase information on a state as they project directly onto the Fock basis. Another way to see

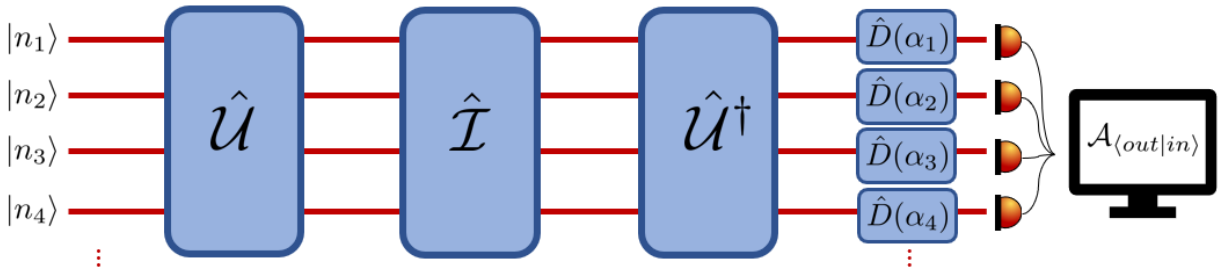



FIGURE 7.12: Phase-sensitive simulation of scattering amplitudes. Input Fock states are rotated to the massive bosonic field basis by operator $\hat{\mathcal{U}}$ and then subjected to interactions described by the operator $\hat{\mathcal{I}}$. After rotating back to the photon-number basis with $\hat{\mathcal{U}}^\dagger$, performing a series of single-mode displacement followed by PNR detection allows for access to the complex-valued scattering amplitude.

this is via the uncertainty principle: if I can directly measure photon number, then the conjugate parameter, the phase, must be entirely indeterminate. However, if we induce phase sensitivity through the use of displacements, which do have phase information, we are essentially performing a projection onto a displaced Fock basis. If many more displacements are used, the full overlap tomography scheme can be deployed to acquire full state information. Since this is invariably an exponentially scaling task, we wish to avoid full state tomography and only access limited phase information — thus the small number of displacements.

# Chapter 8

# Experimental Design

With the exception of Ch. 6, the majority of this work has focused on explaining a variety of theoretical methods to generate and measure non-Gaussian quantum states. This chapter will focus on the experimental design, both implemented and proposed, for realizing a variety of the quantum state engineering protocols discussed. I will focus specifically on progress made towards demonstrating photon catalysis and will discuss successes and current challenges. As measurement is a key portion of demonstrating an experiment, I will also discuss results from the tomography protocol of Ref. [264] in addition to related unpublished results. Finally, I will discuss a proposal to generate the two-mode coherent photon-subtracted state from Ch. 5 with only slight modifications to the current experimental setup.

As all experiments in this section make use of many of the same resources, the beginning of this chapter will be dedicated to presenting the current state of many experimental tools. Since the TES would saturate from continuous-wave (CW) laser operation, I will begin by discussing our implementation of converting the CW beam to pulsed operation through the use of acousto-optic modulators (AOMs). Additionally, as the workhorse of single-photon generation used to verify overlap tomography and for photon catalysis is a down-conversion source from an optical parametric oscillator (OPO), I will also discuss the OPO, cavity locking, temperature control, and additional spectral filtering mechanisms. Once the foundations are in place, I will discuss current experimental results.

## 8.1 Pulsing

As we saw in Ch. 6, the cooling tail of signal peaks from the TES is vital to achieve maximum photon-number resolution. Depending on the bias settings, this can be as long as $10 - 30\,\mu$s from detection onset. However, the arrival of the photons must be constrained to within the rising edge of the peak of $\sim$ 500 ns. If additional photons arrive before the end of the cooling tail, a second peak will begin in the middle of the previous tail. These events are known as pile-ups and were one of the largest limiting factors in previous TES experiments with CW lasers [114, 263, 264]. These pile-ups make it impossible to determine accurate photon-number measurements, so to avoid this, one must either use laser pulses or very weak CW light with average power less than about one photon per 10 $\mu$s, which corresponds to about 18.7 fW. In the latter case, we cannot achieve large enough photon numbers to fully utilize the TES, so pulsing is the ideal path forward. The pulse signal timing is discussed in Appx. D.

### 8.1.1 Acousto-optic modulation

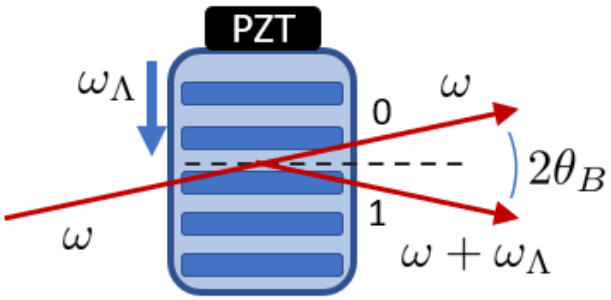



FIGURE 8.1: Light entering an AOM is diffracted at the Bragg angle as it travels through the crystal and acquires a frequency shift of $\omega_\Lambda$ due to the sound wave travelling through the crystal. The light can diffract into multiple orders, but only the first-order diffraction is used for the experiment.

Instead of using a pulsed source such as a mode-locked laser, we start with CW light and convert to pulsed operation by quickly switching the beam on and off to form a pulse. To perform this switching, we use an acousto-optic modulator (AOM) [302]. The AOM works on the principle of Bragg diffraction, where light diffracts from a crystalline lattice as a sound wave propagates through the crystal. The sound wave is a collective oscillation of the atoms in the lattice, so the process can be seen as an interaction between the lattice phonons and the optical photons. The momentum exchange

acts to alter the frequency of the light and impart an angular path deviation. This angular shift is given by the Bragg angle,

$$\sin\theta_B = \frac{\lambda}{\Lambda}, \tag{8.1}$$

where $\lambda$ is the wavelength of light in a vacuum and $\Lambda$ is the wavelength of the sound wave. Additionally, the light acquires a Doppler shift depending on the direction of the sound wave in the crystal and the diffraction order as shown in Fig. 8.1. For an input beam at frequency $\omega$, the output frequency will be

$$\omega^{(0)} = \omega, \qquad \omega^{(1)} = \omega + \omega_\Lambda, \tag{8.2}$$

where the superscripts indicate diffraction order and $\omega_\Lambda$ is the frequency of the sound wave in the material.

The sound wave in the AOM is driven by an RF signal applied to a piezoelectric actuator (PZT) that can be quickly switched on and off with rise time on the order of $\sim 10 - 20$ ns depending on the wavelength and how tightly the beam is focused in the crystal. When the RF drive is absent, there is zero light in the diffracted beam, meaning that the diffracted light can be entirely switched off. When properly aligned and focused, and with sufficient RF power, a majority of the incident light can be converted into the first-order diffraction output.

We implement pulsing using separate AOMs for each the pump beam at 532 nm and the LO at 1064 nm using the first-order diffracted beams. The two AOMs are synchronously pulsed with tunable pulse duration limited to a minimum of $\sim$ 30 ns due to the turn-on dynamics of the PZT within the AOM. Because the pump beam will be used to down-convert to half the frequency, it is essential that the RF drive to the pump AOM is exactly double the drive frequency of the infra-red AOM. Maximum power conversion of 90% to the first-order diffraction was achieved using a drive frequency of 95.000 MHz to the IR AOM. The pump AOM was then set to have a 190.000 MHz drive frequency where a diffraction efficiency of 80% was obtained.

It was verified that the frequency shift of the pump beam was exactly double the IR frequency shift through an optical interference measurement. After being sent through the AOM, the shifted IR light was sent through the OPO and up-converted through second-harmonic generation (SHG) to the pump frequency. This up-converted beam was then interfered with the pump AOM output and a DC interference fringe

was observed. If the frequencies of the two beams were not identical, there would be no static interference pattern.

An important experimental consideration concerning the turn-on dynamics of the AOMs should be noted: just as a laser should be given time to thermalize for power stability, so must the AOMs. Although they work immediately and the beams can be used for routine alignment, the AOMs should be turned on and left at desired operating conditions for approximately $60 - 90$ minutes before performing precision measurements with the output beams.

## 8.2 Single-Photon Generation

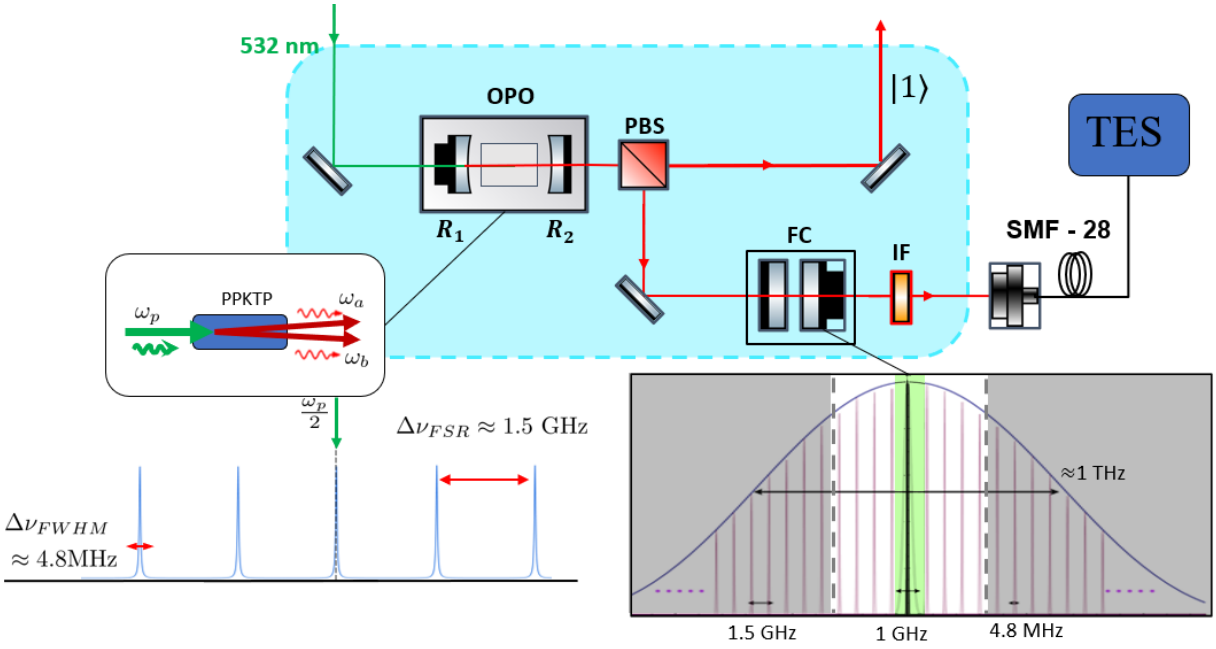



FIGURE 8.2: A type II periodically-poled KTP (PPKTP) crystal is placed in a one-sided cavity to form an optical parametric oscillator (OPO) with $R_1 \approx 1$ and $R_2 = 0.98$. The output frequencies of the down-conversion process are filtered with an external filter cavity (FC) and interference filter (IF) as explained in the main text. The filtered light is coupled to a fiber and sent to the TES, where a single-photon measurement heralds the presence of a single photon in the other channel.

To generate single photons for the experiment, we make use of the spontaneous parametric down-conversion (SPDC) process in nonlinear optics to produce pairs of photons. These pairs are first spatially separated into two paths, and then one of the

twin photons is sent to the TES for detection. When the TES registers one photon, it acts to herald the presence of the other photon of the pair in the undetected path. Provided the heralded single photon is spectrally and spatially pure (and not lost), it can be used for other types of quantum experiments such as photon catalysis.

The single-photon source is implemented using the process of SPDC from a quasi-phase matched type II nonlinear crystal [13]. The crystal is formed from periodically-poled $KTiOPO_4$ (PPKTP) and is placed into an optical cavity to form an optical parametric oscillator (OPO). The cavity is entirely transmissive to the pump beam at 532 nm while it acts as one-side for the down-converted field at 1064 nm as shown in Fig. 8.2. The cavity serves dual purposes: it acts to spectrally filter the broadband emission over the $\approx$ 1 THz phase-matching bandwidth of the crystal while also providing a cavity-enhancement effect for the resonant frequencies. This is a technique commonly used for generating single photons [303].

The OPO consists of two spherical mirrors with 5 cm focal lengths spaced 10 cm apart, where one mirror is highly reflective with $R_1 \approx 1$ and the other has reflectivity $R_2 = 0.98$ for 1064 nm light. This cavity design leads to a free-spectral range (FSR) given by

$$\Delta_{FSR} = \frac{c}{2L} = 1.5\text{GHz}. \tag{8.3}$$

The cavity finesse can then be calculated as [304]

$$F = \frac{\pi(R_1R_2)^{1/4}}{1-\sqrt{R_1R_2}} \approx 310. \tag{8.4}$$

For low-loss cavities, we can obtain the linewidth of each resonant peak from

$$\Delta_{FWHM} \approx \frac{\Delta_{FSR}}{F} \approx 4.8\text{MHz}. \tag{8.5}$$

Sending a weak pump beam at 532 nm through the OPO will now spontaneously produce pairs of cross-polarized, co-linear photons that can be spatially separated at a polarizing beamsplitter (PBS). Due to the cavity spectrum, these two photons may be separated in frequency by integer multiples of twice the FSR of the cavity; however, we will want pure, degenerate single photons at exactly half of the pump frequency. To achieve this, we must perform spectral filtering on the photons used for heralding. This will ensure that we only collect data on the other half of the experiment (the half using the heralded photon) when the TES registers a heralding photon at the desired

frequency. This filtering is achieved through the use of a passive filter cavity (FC) with a cavity linewidth of 1 GHZ and FSR of 300 GHz. When the FC and OPO are locked together onto the same reference frequency, the linewidth of the FC will serve to filter out all but the degenerate OPO cavity mode. Additionally, in order to eliminate any other simultaneously resonant modes within the large phase-matching bandwidth of the crystal, we utilize an interference filter (IF) to eliminate all but one of the FC resonant modes (gray region in inset of Fig. 8.2).

When the filtering is in place and the cavities are properly locked to a reference beam, all photons that make it through the heralding path to the TES are now certain to be the same frequency as the reference locking beam, which will be used as a local oscillator later. This ensures that the down-converted photon in the heralded path will *also* be the same frequency as the measured photon, ensuring a highly pure single photon. Further details on this specific cavity design and filtering can be found in the Ph.D. thesis by Reihaneh Shahrokhshahi at UVA [305]. Additionally, details regarding the limits of single-photon generation with SPDC can be found in Ref. [306], and a general model for heralding single photons included detector efficiency can be found in Rajveer Nehra's Ph.D. thesis [286].

### 8.2.1 On/off locking

In order to ensure that the filter cavity works properly to filter all but the degenerate OPO mode, both the FC and OPO must be locked to the same reference frequency. This is achieved by using the method of Pound-Drever-Hall locking [222]. In our implementation, we denote the frequency reference as the locking beam, which is derived from the local oscillator (LO) used for interference with the signal later. If the cavities are locked to the same frequencies as the LO, then we can be sure that the degenerate mode will properly interfere.

Because the OPO is approximately a one-sided cavity at 1064 nm, the locking beam is injected through the front of the cavity, and the weakly transmitted mode is sent to a photodiode and used to generate an error signal for locking. In this implementation of PDH locking, the cavity length is modulated through the use of a mirror-mounted PZT[1]. Simultaneously, the locking beam is sent through the FC to a second photodiode

[1]The PDH locking on the filter cavity also places modulation on the cavity length with a PZT.

to generate an error signal for locking it as well. Since the signal and locking beam follow the same path through the FC, polarization is used as a separate degree of freedom so that the locking beam path through the FC to the photodiode is independent of the signal path to the TES.

Unfortunately, the nature of the two-mirror cavities means that we cannot have counter-propagating cavity modes, such as in a three or four mirror cavity. This fact, paired with the highly sensitive nature of the TES, means that we cannot simultaneously lock the cavities and perform single-photon measurements with the TES. The SPDC light exiting the cavity must follow the same path as the locking beam reflected off of the front mirror of the cavity. Fortunately, the monolithic OPO cavity design and extremely short FC (0.5 cm length) mean that both cavities are stable for reasonable amounts of time ($\sim$ 1 second). Knowing this, we can implement an on/off locking scheme, where the cavities are locked by the frequency reference for some amount of time, and then the locking beam is shut off while data is collected within the stability time of the cavities. After a short period, the locking beam is re-engaged to ensure the cavities remain locked. By repeating this process and introducing optical shutters to protect the TES from saturation, data at the single-photon level can be collected in short bursts between periods of cavity locking.

The sequencing of when to turn on and off the locking beam, when to open the shutters, and when to pulse the laser must be designed carefully. The shutters to protect the TES have response times of $\sim 15 - 20$ ms, so it is important to build in additional dead time where the locking beam is shut off before the shutter is opened and only re-engaged once the shutter is fully closed. All signals are controlled and synchronized by an external Quantum Composers 9500 series pulse generator. The pulse schedule is displayed in Appx. D.

### 8.2.2 Mode-matching

Once the single-photon state is generated, the state must be properly mode-matched to achieve maximum interference with the local oscillator mode as well as the optical fiber coupled to the TES for measurement. And while losses on the heralding path only reduce state generation rates and do not degrade the purity of the heralded photon, it behooves the experimenter to limit losses as much as possible to maximize count rates. While spectral filtering is achieved through the use of cavities as discussed above, the

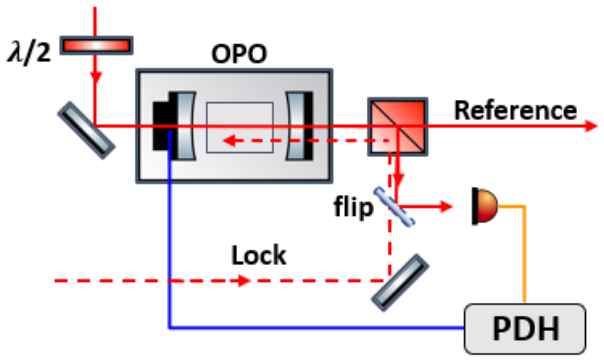



FIGURE 8.3: A reference beam fed through the back of the high-reflectivity mirror on the OPO is used for mode-matching with the cavity field.

mode-matching in this section refers to ensuring that the spatial wavefront of the cavity output matches other beams and optics further along the propagation path.

Previous implementations of this experiment achieved spatial mode-matching by reflecting the classical locking beam off of the front of the OPO and using this for alignment. While the beam path is identical in this case, the beam profile may deviate from the cavity mode due to interference with a locking beam that does not precisely match the cavity mode. To remedy this problem, we couple a reference beam through the high-reflectivity mirror at the back of the cavity as shown in Fig. 8.3. Because the cavity is asymmetric and the reflectivity of the mirror $R_1 \approx 1$, the maximum laser power transmitted through the cavity is only $\sim 30\,\mu$W. While small, this is enough for characterization and cavity locking.

Using the reference beam, a CW visibility of 97% was achieved with the LO used for the tomography experiment and visibility of 98% was obtained for use in the photon catalysis experiment. Additionally, the reference beam can be matched closely to the fiber mode by sending a counter-propagating beam back through the fiber and measuring the Gaussian beam parameters of the output fiber mode. By matching the reference beam to the output fiber mode, coupling efficiencies of 85% and 90% were achieved for respective detectors B and C shown in Fig. 8.6.

### Temperature stability

Due to the fact that KTP is a birefringent material and we are using type II SPDC, it is vital that the cross-polarized photons are both resonant simultaneously within the OPO. This can be achieved by modifying the refractive indices of the KTP crystal at the

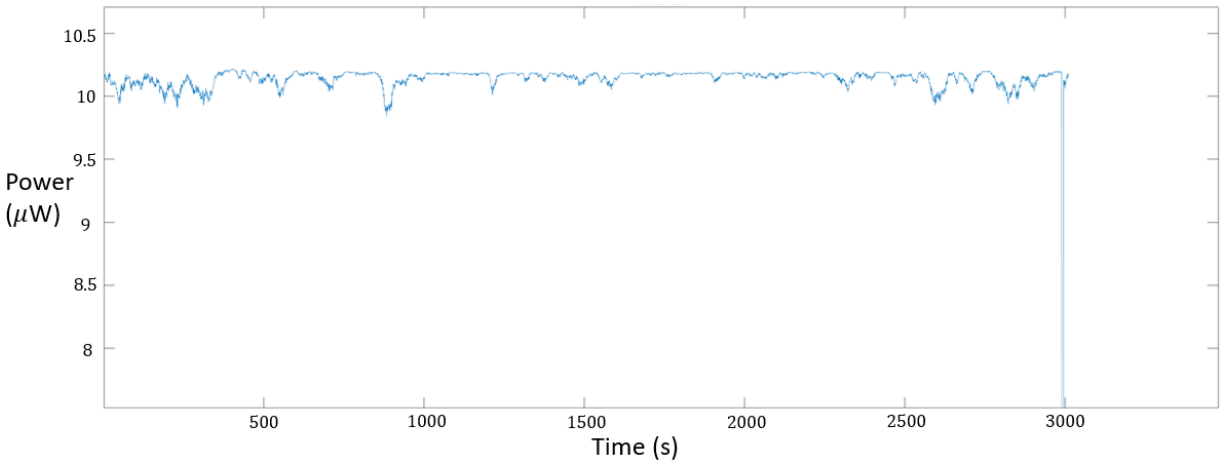



FIGURE 8.4: Stability of the horizontally polarized reference beam power output from OPO over the course of a 50 minute period while the cavity was locked on the vertical polarization. Small temperature fluctuations result in dips in the transmitted power to 96% of the peak power.

two polarizations by tuning the crystal temperature. At the correct temperature setting, the OPO becomes double-resonant for both down-converted photons in the pair.

The temperature was stabilized near 28°C to within one millidegree using the commercial HTC-500 temperature controller. By sending the reference beam through the back of the OPO and locking on one polarization, the transmitted power of the other polarization can be monitored over time to characterize the temperature stability of the crystal. This is shown in Fig. 8.4, where the cavity was continuously locked for a 50 minutes period and the transmitted power was monitored. Small thermal fluctuations across the crystal cause the beam power to drift as the cavity lock only follows one of the two polarizations, but the temperature controller acts to correct deviations before long. Over the measurement period, the beam power drifted to a minimum of 96% of the peak power while remaining above 98% of the peak power on average.

### Temporal synchronization

Beyond matching the spatial mode profile, it is also vital to overlap the LO pulse with the signal pulse in the time domain. Otherwise, there will be no interference. The first step is to ensure the path length is the same for each the LO and the signal after the AOM aperture. Although care was taken to ensure that the path lengths were the same to within several centimeters, the relatively long 200 ns pulse duration used means that a full meter of path-length difference would only degrade the temporal overlap by

less than 2%. However, the difference in turn-on dynamics of the two separate AOMs means that a delay may need to be introduced in one path to compensate.

The pulse timing could be verified by bringing the OPO above threshold and interfering the pulsed down-converted signal with the pulsed LO beam. Unfortunately, previous experiments conducted by Reihaneh Shahrokhshahi indicated that the high intra-cavity power of the OPO cavity above threshold damages the KTP crystal [305], rendering this method unfeasible. Alternatively, the TES can be used to measure arrival times of photons from each beam path. If the photons arrive at simultaneous times from the initiation of an RF trigger pulse common to each AOM and the path lengths are the same, then the two optical pulses will temporally overlap at the beamsplitter where the interference should occur.

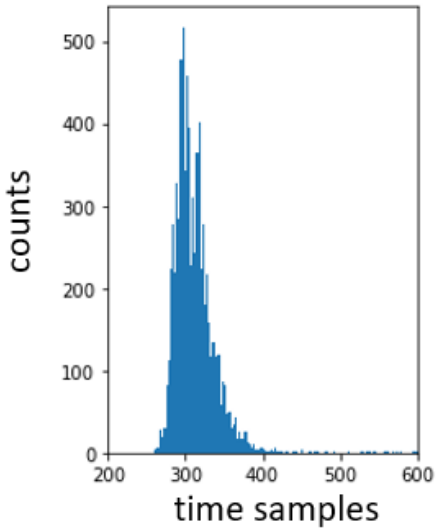



FIGURE 8.5: Histogram of the timestamp of single-photon peaks in terms of number of EFADC samples from initiation of a trigger pulse. The mean peak time was 317 samples after the trigger, which at 4 ns/sample corresponds to 1.268 $\mu$s for this particular TES channel.

To do this, one can start by blocking the OPO and sending a weak coherent state pulse following the path of the local oscillator to the TES. Fig. 8.5 shows the histogram of the peak times for single-photon events in EFADC samples from the trigger signal. For this data set, the mean peak-time occurred at 317 samples points, or 1.268 $\mu$s, after the arrival of the trigger. Next, one can block the LO path and collect data on the TES channel from SPDC photons from the pulsed pump traveling through the OPO. Because the timing information is the same whether or not the SPDC pair is frequency-degenerate, information from the heralding channel can be ignored and all

single-photon events can be collected and used for this calibration. By tuning the delay of the signal sent to pulse the pump AOM, this process can be repeated until the mean arrival times of the LO pulses match the mean arrival times of the SPDC events. Once completed, the mean arrival times differed by less than 2 EFADC sample points meaning that the centers of the pulses were within 8 ns of each other.

It is important to note that the pulse generator used has the ability to tune all outgoing pulses with respect to one master trigger. This master trigger is sent to the EFADC to prepare it for an incoming TES signal, but all pulses sent to turn on/off the AOMs and shutters can be tuned with respect to the master trigger. Another important point is that one must only use single-photon events for this calibration from both the SPDC and LO. The TES response rise-time varies slightly with photon number, so timing calibration should be done with the same photon-number events.

### 8.2.3 Full experimental design

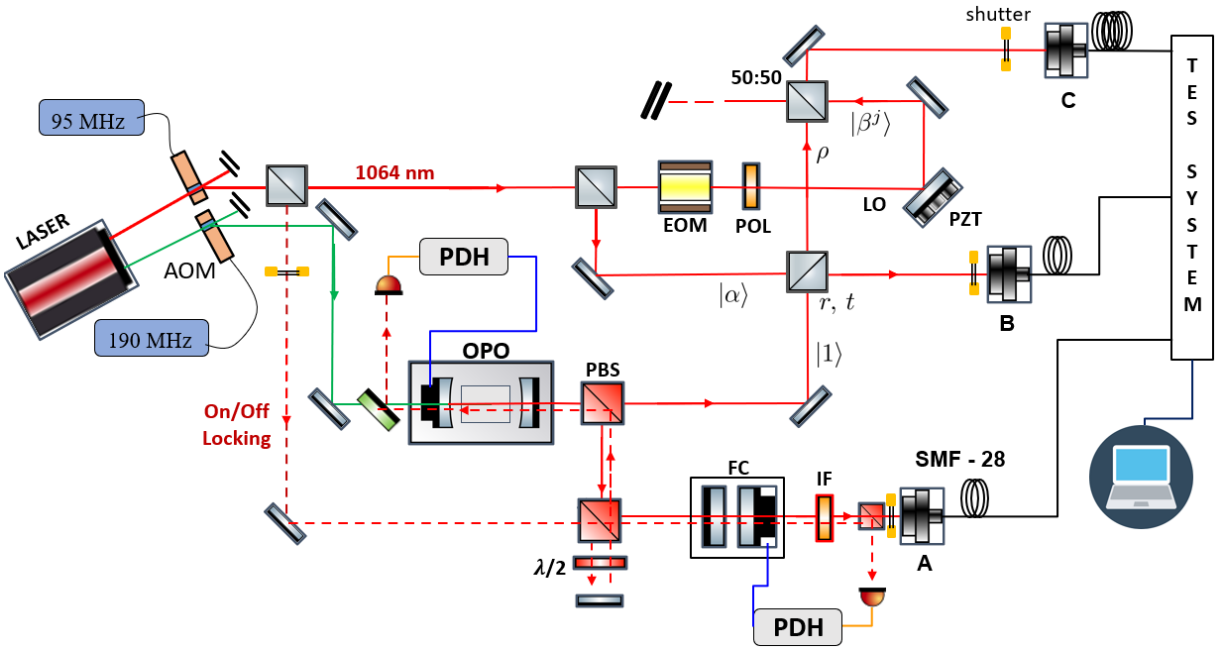



FIGURE 8.6: Full experimental design for photon catalysis, with the exception of the reference beam used for mode-matching. EOM: Electro-optic modulator. POL: Polarizer. PZT: Piezo-electric actuator. PDH: Pound-Drever-Hall cavity locking. $\frac{\lambda}{2}$: Half-wave plate. PBS: Polarising beam-splitter. LO: Local oscillator. SMF-28: Single-mode optical fiber[2].

The full experimental design is depicted in Fig. 8.6. The diagram shows the paths of the locking beam as dotted lines, which are operated in CW mode during the 'on' phase of the on/off locking scheme, meaning that the IR AOM is left active without pulsing. During the 'off' phase of locking, the shutter in the locking beam path is closed and the shutters in front of the detectors are opened. Additionally, pulsed mode for both the green and IR AOMS is engaged.

Once a single photon is registered by channel A of the TES, the presence of a single photon is heralded on the transmitted path through the PBS just after the OPO. This photon then interferes with the coherent state, $|\alpha\rangle$, at a beamsplitter of variable reflectivity $r, t$[3]. This step is omitted in the case of the single-photon tomography experiment. One output of this beamsplitter is sent to TES channel B, where a measurement of $n$ photons indicates the precise form of catalysis that has occurred. The other port output is now our engineered state, $\rho$, which is sent to the tomography protocol. Tomography is performed by interfering $\rho$ with a variable coherent state $|\beta^j\rangle$, with amplitude controlled by the polarizer placed after an electro-optic modulator used to rotate polarization, and the phase is controlled by a mirror-mounted PZT. The result of this interference is sent to TES channel C to measure the expectation of parity as discussed in Ch. 7. Note that this interference for tomography can be modified to a displacement if an observable other than parity is used for characterization, as discussed in Sec. 7.3.2 in the previous chapter.

## 8.3 Overlap Tomography

To verify that the new pulsed setup works as expected, the experiment from Ref. [264] was repeated. The results for the tomography of the single photon are shown in Fig. 8.7, where the photon number distributions are shown when all data is considered next to the phase-averaged data. The plotted Wigner functions show the results after loss-inversion is applied to recover Wigner function negativity. The results of this experiment successfully recreate the CW experiment from the previous work with minor but notable differences. In this case, less care was taken to optimize all aspects of the experiment, so the overall loss was approximately 10% higher. Taken together, all losses totaled to approximately 58%. Additionally, the pump power was not optimized for

[3]This beamsplitter is actually comprised of a PBSs and a waveplate used to tune the splitting ratio

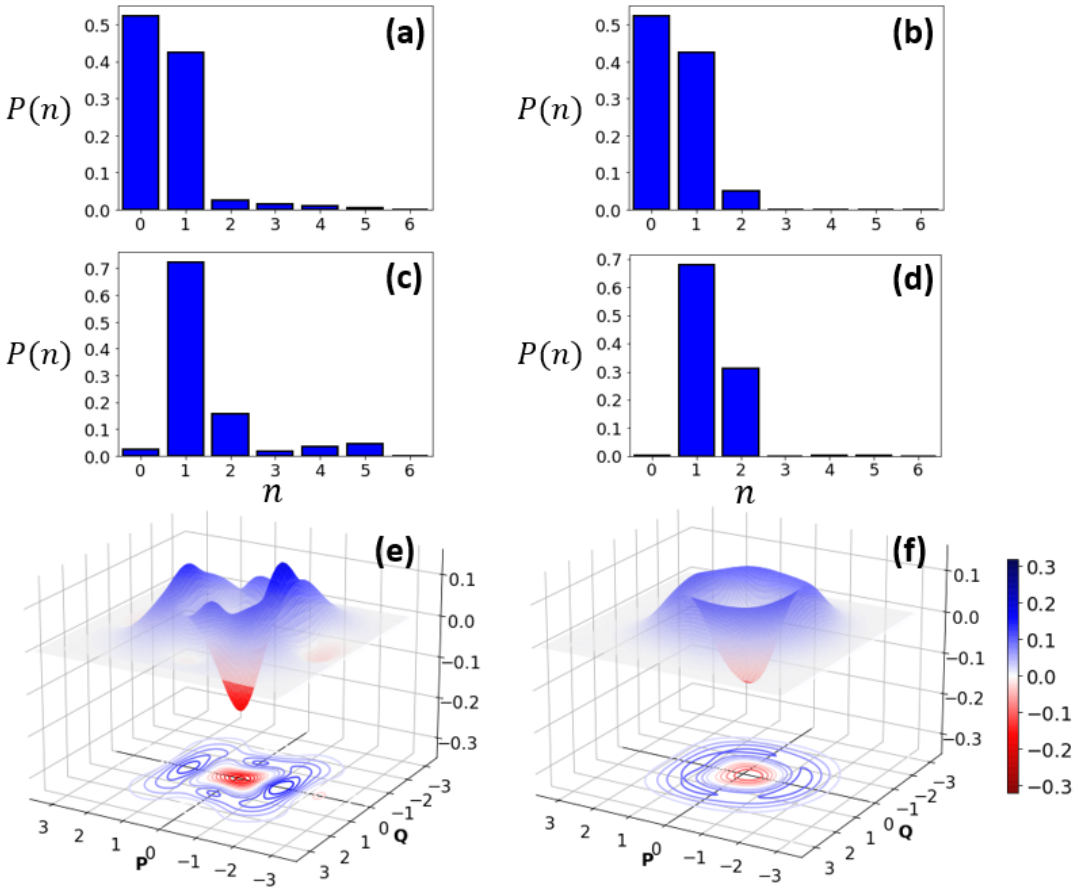



FIGURE 8.7: Left column: reconstructions with all data points. Right column: reconstruction assumming phase-space symmetry by averaging data at different LO phases. (a) and (b) show the photon-number distribution reconstruction from the direct data. (c) and (d) shows the photon-number distribution after performing statistical loss-inversion on the data, where the associated Wigner functions are shown in (e) and (f).

single-photon generation. As one can see by looking at the photon-number distributions in the figure, the two-photon contributions are quite high. This comes about when loss through the FC on the heralding path causes the TES channel to see a single photon, but two photons would have been detected in the lossless case. Thus, the signal side of the experiment now has a two-photon Fock state when the experimenter mistakenly thinks it is a single photon. Because the tomography process is agnostic to the type of input state, this is clearly seen by the reconstruction. In addition to the photon number distributions, the reconstructed Wigner functions are less negative than the single-photon state. This is due to the two-photon state's Wigner function spiking positive at the origin of phase-space.

Otherwise, this experiment is nearly identical to the results from before. The Wigner function reconstructed from the raw data clearly resembles that of a single photon but with extra ripples and noisy features. These additional ripples can be attributed mainly to a combination of slight amplitudes fluctuations in the LO probe over the course of the several hours of data collection experiment and the effects of finite sampling errors on the measured statistics. Additionally, the precise purity and composition of the state being measured will drift slightly over the course of the experiment as the pump power, crystal temperature, or OPO cavity locking change. However, enforcing symmetry by averaging measurements at different LO phases for the same LO amplitude effectively washes away these fluctuations by averaging over longer times and larger data sets.

In addition to the tomography of a single photon, tomography of a phase-averaged coherent state (PACS) was performed. In Ref. [264], care was taken to demonstrate the tomographic method worked for both a pure quantum state, the single photon, and for a classical-like state with asymmetric phase qualities — the coherent state. However, the procedure was not demonstrated on a statistical mixture. Here, overlap tomography was also demonstrated using the mixture of a PACS as shown in Fig. 8.8.

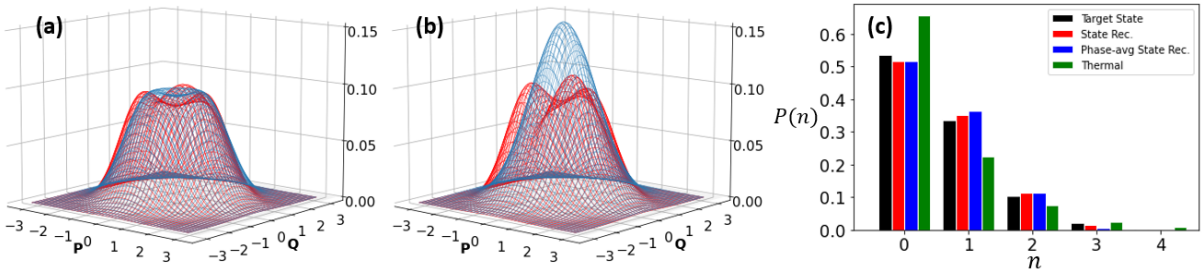



FIGURE 8.8: Wigner function and photon-number distribution for the tomographic reconstruction of a phase-averaged coherent state. (a) Comparison of the measured Wigner function (red) with the target phase-average coherent state (blue); the fidelity is 98%. (b) Comparison of the measured reconstruction (red) to a thermal mixture Wigner function with the same mean photon number (blue); the fidelity is only 90%. (c) Comparison of the photon number distribution for the target state (black) with reconstructions using all data (red), averaging data (blue), and the thermal distribution (green).

This is quite interesting because not only is the reconstructed Wigner function quite close the expected PACS in Fig. 8.8(a), but it is also clearly different from the nearest thermal mixture as shown in Fig. 8.8(b). The expected PACS, or target state, was

determined by finding the PACS having the same mean-photon number as the measured 'unknown' state. Similarly, the nearest thermal mixture was determined to be the thermal mixture with the same mean photon number as the measured state. Fig. 8.8(c) demonstrates the efficacy of the tomographic reconstruction by directly comparing the photon number distributions. The target state in black is compared to reconstructions using all measured data as-in (red) and the measured data assuming phase-space symmetry (blue). The distribution for the thermal mixture in green is clearly quite different; thus the method of overlap tomography is successful on mixtures as well as pure states.

## 8.4 Photon Catalysis

Chapter 2 introduced photon catalysis and showed that for a single-photon input, it is possible to filter out a Fock-basis component. Although the derivation presented there was instructive, it is less cumbersome if we treat the process of photon catalysis as a Kraus operator. Using a pure single photon at one port of the beamsplitter and a PNR detector on one output, the Kraus operator for an $m$ PNR detection can be derived to be ( A.6)

$$\hat{K}_m = \left(\frac{r}{t}\right)^m \frac{1}{r\sqrt{m!}} \left(mt^2 - r^2\hat{a}^\dagger\hat{a}\right) \hat{a}^{m-1} e^{-\beta\hat{a}^\dagger\hat{a}}, \tag{8.6}$$

where $r^2 + t^2 = 1$ are the beamsplitter parameters and $\beta = -\ln t$. In this form, it is easy to see that when the beamsplitter is balanced with $r = t$, the term in parentheses becomes $(m - \hat{n})$, so the $m$-th photon component must vanish from the superposition when applying $\hat{K}_m$ to any state. Noting that for a coherent state,

$$e^{-\beta\hat{a}^\dagger\hat{a}} \left|\alpha\right\rangle \propto \left|e^{-\beta}\alpha\right\rangle, \tag{8.7}$$

we have that photon catalysis applied to a coherent state leads to

$$\hat{K}_m \left|\alpha\right\rangle \propto \left(mt^2 - r^2 e^{-\beta}\alpha\hat{a}^\dagger\right) \left|e^{-\beta}\alpha\right\rangle, \tag{8.8}$$

which is now a superposition between a different coherent state and a photon-added coherent state. Setting $r = t$ for simplicity leads to the output state

$$\left|\phi\right\rangle \propto (m - \tfrac{1}{\sqrt{2}}\alpha\hat{a}^\dagger) \left|\tfrac{1}{\sqrt{2}}\alpha\right\rangle. \tag{8.9}$$

### 8.4.1 Realistic modeling

Ideally, Eq. 8.9 captures what we would like to generate experimentally. Unfortunately, the experimental setup shown in Fig. 8.6 contains many real-world imperfections that cannot be neglected. There are three main contributing imperfections that can be included to form a better theoretical model for what we are likely to see. These imperfections include:

- an impure heralded single photon, such that the true heralded state is a mixture of Fock states,
- photon loss and detector inefficiency,
- imperfect visibility, or mode-matching, between the heralded single photon and coherent state.

The majority of these influences can be extensively characterized and built into a model before running the experiment. Additional errors include finite sampling as discussed in Sec. 7.3.2, miscellaneous experimental drifts over the several-hour course of data collection, and TES resolution errors beyond detection efficiency. Overall, modeling the three main errors bullet-pointed above will be sufficient to capture the main features of the data. It should be noted that approximately 0.3% of pulses resulted in a single-photon detection on the heralding TES channel. Increasing the pump power further increases the count rates, but also increases the likelihood of either having two-photon events or having a pair of twin SPDC photons where one pair is not frequency degenerate.

#### Heralding

First, consider the true heralded distribution shown in Fig. 8.9. The distribution at left shows the raw distribution after heralding on a single photon. The distribution on the right then shows the distribution after correcting for losses, which is closer to the true distribution that will interfere with the coherent state at the catalysis step. Because these photons originate from an SPDC process, the quantum state that interacts with the rest of the experiment is just a diagonal density matrix that is a mixture of Fock states and thus is exactly characterized by the photon number distribution.

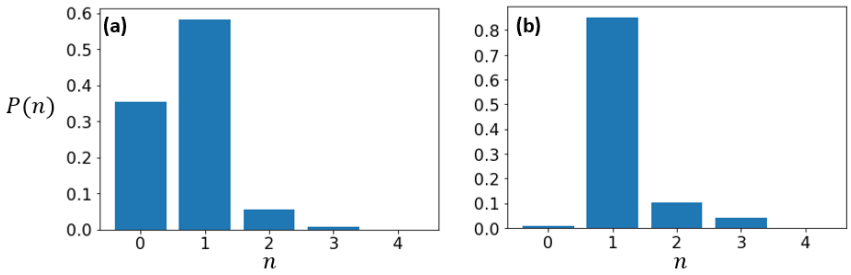



FIGURE 8.9: (a) Raw measured heralded photon number distribution when heralding on a single photon. (b) Heralded distribution after correcting for measured losses.

From the figure, the heralded state is primarily a single photon, with some additional two and three-photon components. These higher number photons come from events where losses on the heralding detection (FC and detector) caused a single photon to be measured while two or three would have been present without loss. The Faraday rotator, filter cavity, and other optics account for 16% loss, while fiber coupling on this arm accounts for an additional 25% loss.

### Coherent state calibration

In addition to calibrating the purity of the single photon from the heralded distribution, it is also important to calibrate the coherent state that interferes with the single photon. This is achieved by blocking the beam path from the OPO and just acquiring samples of the coherent state on the detector. The mean photon number of the measurement can then be used as the value for $|\alpha_\eta|^2$ *after* all losses have occurred. Fortunately, the coherent state is an eigenstate of the annihilation operator, thus a coherent state maps to a different, smaller coherent state under loss. With properly calibrated loss, one can determine the value of the coherent state that impinges on the interference beamsplitter to be

$$|\alpha_0\rangle = |\sqrt{\tfrac{2}{\eta}}\alpha_\eta\rangle\,, \tag{8.10}$$

where $|\alpha_0\rangle$ is the coherent state that interferes with the single-photon state, $|\alpha_\eta\rangle$ is the directly measured coherent state on a detector with efficiency $\eta$, and the additional factor of $\sqrt{2}$ comes from the fact that the beamsplitter used in this specific photon catalysis process is balanced with $r = t = \frac{1}{\sqrt{2}}$.

### Loss calibration

Losses are calibrated at several points throughout the experiment by measuring the transmitted power of a classical beam. Although the maximum fiber coupling achieved for some experiments reached 90%, it is estimated that the fiber coupling for the data presented here was reduced to approximately 80%. Additional losses on the heralded signal path were measured at 5% from various optical elements used. Finally, as discussed in Ch. 6, a value of approximately 90% can be used for the TES detection efficiency. This leads to an overall value of $\eta \approx 0.8 * 0.95 * 0.9 = 0.68$. Due to the relatively large uncertainty on detection efficiency of the TES, both beam paths for channels B and C in Fig. 8.6 are assumed to have the same overall efficiency of $\eta$.

Furthermore, by looking at the heralded signal distribution from Fig. 8.9, it is clear that more losses must be present elsewhere. It is likely that there is approximately 10% additional loss from temperature drift and on/off locking error that accrues on the OPO output. However, using the value of $\eta \approx 0.68$ for the remaining portions of the experiment is consistent with classical measurements and matches the measured PNR distributions for the catalysis experiment remarkably well, as we will see.

### Imperfect visibility

Although the classical visibility was measured to be quite high with 98% spatial mode-matching with CW beams, and care was taken to ensure the pulses arrived simultaneously as discussed in Sec. 8.2.2, we will see in the next section that the experimental data indicates substantially poorer visibility. Here, I will discuss a method to model such mode-mismatch.

When considering the interference of two density matrices, $\rho_1$ and $\rho_2$, a full treatment can be described by the diagram in Fib. 8.10(a), which is similar to the model developed in Ref. [307]. In this model, each of the two modes can be decomposed into an overlapping portion that interferes properly and a portion that does not interfere. The non-interfering portions can first be split by beamsplitters, which introduces vacuum. The interfering portion then overlaps perfectly at the beamsplitter, while the non-interfering portions are split and interfere only with vacuum. Thus, while in the ideal case there would have been two modes in and two modes out, now each output of the beamsplitter has three independent modes. The detector on an output now acts as a

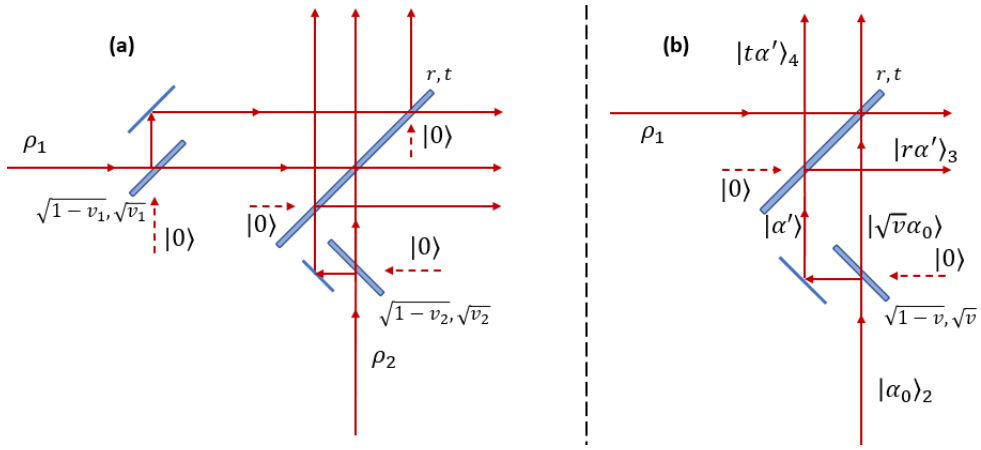



FIGURE 8.10: (a) Full model for including visibility imperfections. (b) Simplified model where $\rho_2$ is a coherent state and the majority of $\rho_1$ interferes with a portion of $\rho_2$.

bucket detector across all three modes. From this model, one can see that imperfect visibility is worse than loss alone, as vacuum gets in and the bucket detection introduces mixing through the inability to distinguish whether the measured photons originated from the interfering or non-interfering modes.

As opposed to this full model, we can consider a simplified model in Fig. 8.10(b) that while less accurate, acts to sufficiently capture the experimental results as will be verified in the next section. First, take the second density matrix, $\rho_2$, to be the coherent state as is the case for this experiment. Now for a coherent state $|\alpha_0\rangle$, imperfect visibility simply splits the state into different coherent states in separate modes. The second simplification in this model is that the majority of $\rho_1$ interferes with the coherent state, but some of the coherent state does not interfere with $\rho_1$. This is approximately valid if we consider $\rho_1$ as the detection mode, and if we note that the mean photon number of $\rho_1$ is considerably lower than the coherent state, we see that the majority of the photons contaminating the TES detection originate from the non-interfering portion of the coherent state.

The results of photon catalysis can be derived from Fig. 8.10(b) by recognizing that the state after the interference beamsplitter is given by

$$\rho' = \hat{B}_{12}(\rho_1 \otimes |\sqrt{v}\alpha_0\rangle \langle\sqrt{v}\alpha_0|)\hat{B}_{12}^\dagger \otimes (|r\alpha'\rangle \langle r\alpha'|)_3 \otimes (|t\alpha'\rangle \langle t\alpha'|)_4, \tag{8.11}$$

where $v$ is the visibility, $\alpha_0$ is the initial coherent state, and $\alpha' = \sqrt{1-v}\alpha$.

Suppose we now have a PNR detector on the reflected output. Instead of performing a perfect Fock-basis measurement on a single mode, each component of the projector will actually be a mixture, as for a measured photon count of $m$, these $m$ photons could have come in any combination from the two modes. The projection element for an $m$ photon detection will therefore become

$$|m\rangle\langle m| \rightarrow \sum_{k}^{m}(|k\rangle\langle k|)_1 \otimes (|m-k\rangle\langle m-k|)_3. \tag{8.12}$$

Because the third mode is a coherent state, we can derive the form of the projector acting on the remaining single mode to be

$$\begin{aligned}\hat{\Pi}_{m,v} &= \mathrm{Tr}_3\left[\sum_{k}^{m}(|k\rangle\langle k|)_1 \otimes (\langle r\alpha'|m-k\rangle\langle m-k|r\alpha'\rangle)_3\right] \\ &= \sum_{k}^{m} e^{-|r\alpha'|^2}\frac{|r\alpha'|^{2(m-k)}}{(m-k)!}(|k\rangle\langle k|)_1.\end{aligned} \tag{8.13}$$

In the above, the dependence on visibility is implicitly contained within the definition of $\alpha'$. However, this projector does not include detector inefficiency. By modifying each Fock-basis projector of Eq 6.4 from Ch. 6 with Eq. 8.13, we can arrive at the full single-mode projector of

$$\hat{\Pi}^{(\eta)}_{m,v} = \sum_{n=0}^{\infty}\binom{n}{m}\eta^m(1-\eta)^{n-m}\sum_{k}^{n} e^{-|r\alpha'|^2}\frac{|r\alpha'|^{2(n-k)}}{(n-k)!}|k\rangle\langle k|. \tag{8.14}$$

Thus for a measured photon number of $m$ on the catalysis detector (channel B in Fig. 8.6), the projector in Eq. 8.14 is actually applied to the quantum state.

Once this projector is applied, one can calculate the expected probability distribution of the other detector (channel C in Fig. 8.6). As discussed in Sec. 6.1, the obtained distribution for a bucket detector will really be a discrete convolution of the individual probability distributions. In this case, the result will be the mixed signal obtained after the catalysis detection convolved with the mode-mismatched portion of the coherent state.

In order to determine the true visibility from the measurement data, the catalysis detector is first ignored, so the full measured distribution is obtained from the detector in channel C, regardless of the measurement in channel B. This amounts to taking a partial trace over all modes that were sent to channel B. The reason the calibration was

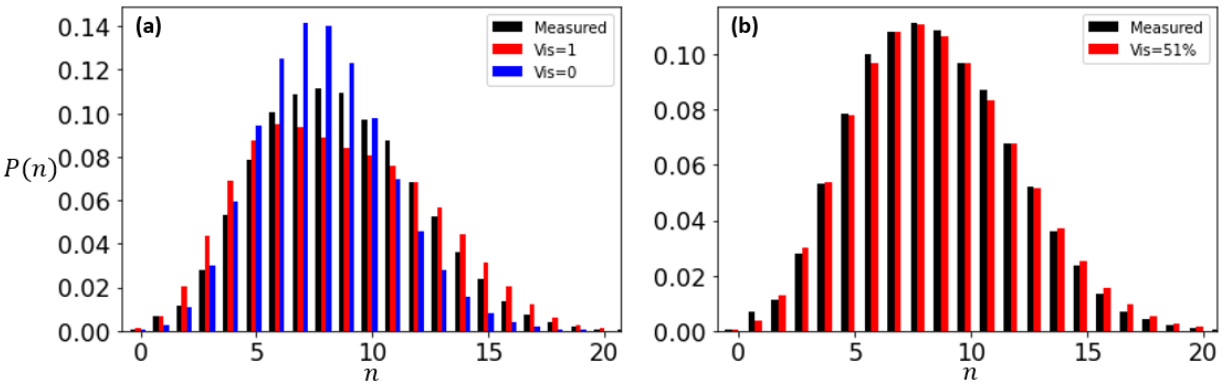



FIGURE 8.11: Photon number distributions for photon catalysis mixtures where the catalysis detector is ignored. (a) Measured distribution (black) compared to a theoretical model with ideal visibility (red) and zero visibility (blue) corresponding to no interference between the coherent state and single-photon input. (b) Measured distribution compared to the model with 51% visibility.

performed in this way was to ensure that all data was included and the distribution had the maximum statistical significance. The measured distribution (black) is shown in Fig. 8.11(a) comparing the cases for perfect interference (red) and a complete lack of interference between the heralded signal and the coherent state (blue). As is evident from the bar graph, the experimental data lies somewhere between the two extremes. By minimizing the sum-squared error between the model and the distribution, it can be found that the model matches the experimental data most closely when the visibility parameter is set to 51%.

### 8.4.2 Results

The results for the photon catalysis experiment with a balanced beamsplitter are shown in Fig. 8.12 compared to modeled results with various parameters set, where the catalysis PNR detection result is the number shown in red in each plot. More results beyond the displayed distribution were obtained, but the distributions for catalyzing on photon-number detections of $5-12$ each contained at least $5,000$ sampling events, so the distributions obtained have sufficient statistical significance.

First, column (i) of Fig. 8.12 shows the ideal results that could be expected for the input coherent state and heralded signal in the case of perfect interference visibility and no loss. Recall from Ch. 2 that for perfect photon catalysis with an input single-photon state and a balanced beamsplitter, we would expect the $m$-photon probability to vanish

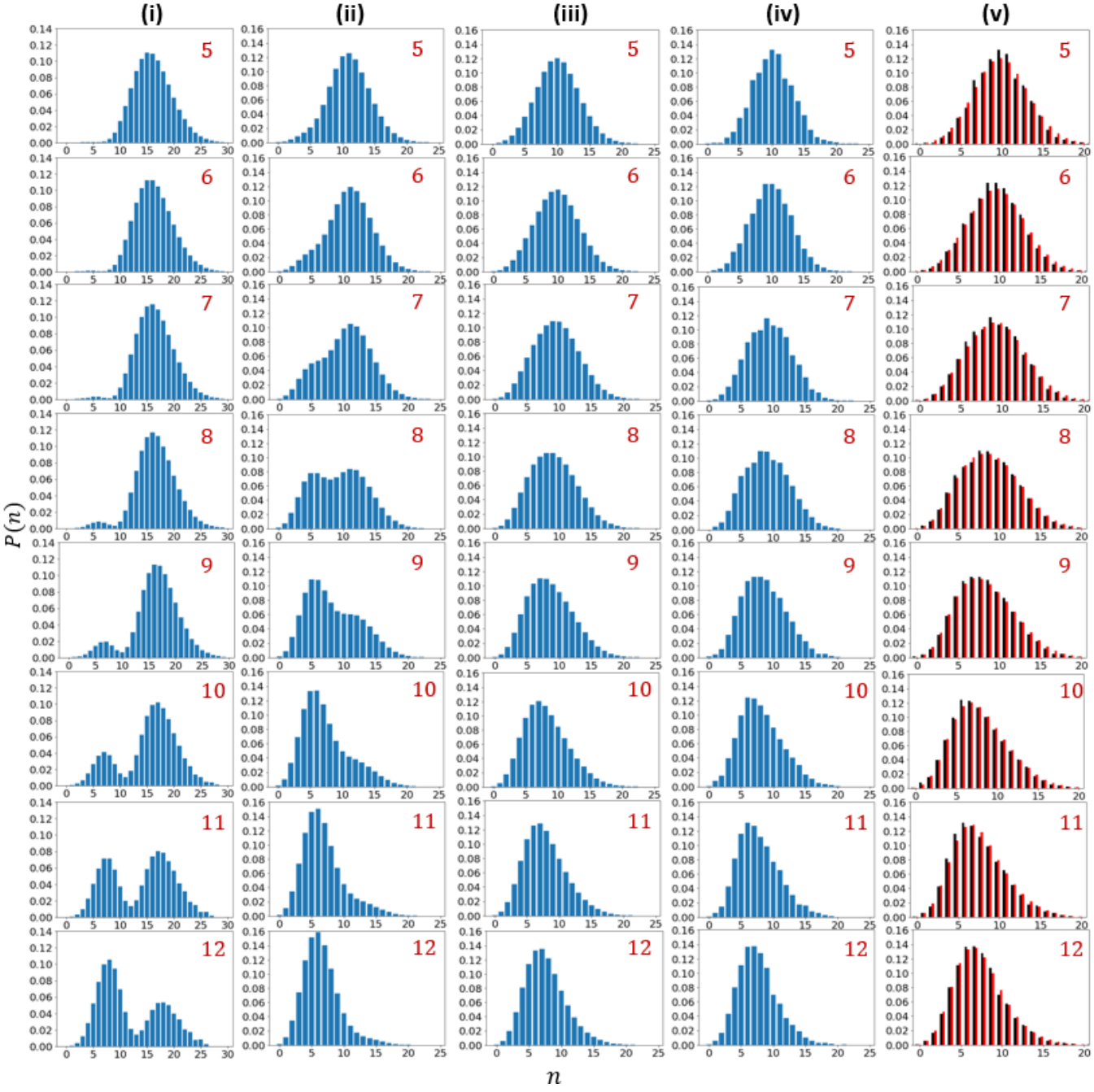



FIGURE 8.12: Photon-number distributions for photon catalysis, where the catalyzed photon detection is shown by the number in red. (i) Theoretical model for photon catalysis for input measured mixture of Fock states and coherent state ignoring losses and imperfect visibility. (ii) Model with loss and detector inefficiency included. (iii) Model including loss, detector inefficiency, and imperfect visibility. (iv) Raw experimental data. (v) comparison of experimental data (black) to full model (red).

for an $m$ photon detection on the catalysis detector. In this case, the heralded state is not a pure single-photon state but is the mixture of several Fock states shown in Fig. 8.9(b). This prevents any of the probabilities in the resultant distribution from vanishing entirely and imparts a shift to the minimum probability away from the exact catalysis detection. However, photon catalysis under these conditions still causes a strong bifurcation of the coherent-state Poissonian distribution depending on the measurement result.

Column (ii) of the figure adds in the overall experimental loss to the model, both on the catalysis projective measurement and on the final detector to determine the measured distribution. The loss on the projective measurement causes mixing of catalysis outputs for several different cases, which smears the bifurcation otherwise seen in column (i). Additionally, the loss on the final detector shifts the distribution towards the origin and further washes out defining features. Nonetheless, even with the measured experimental efficiency of $\eta \approx 0.68$ we would expect to see well-defined signatures of successful photon catalysis for a variety of cases, especially with catalysis detection events of $7 - 10$.

Column (iii) shows the full model including loss and the imperfect interference with visibility of $v = 0.51$, and column (iv) shows the experimental measurements. The model is then directly compared to the experimental data in column (v), where the measured data is shown in black and the model results are given in red. From the last column, it is clear that the relatively simple model captures the data extremely well, as the red and black distributions are nearly a perfect match.

Unfortunately, this brings to question a different issue: if the pulse timing and spatial mode interference were measured to be quite high, why does the model only fit the data when the overall visibility is lowered to 51%? One such probability is that the photons making it through the FC on the heralding channel are in fact not the same degenerate frequency as the heralded photons. However, as the OPO and FC are locked to the same reference frequency, the only likely way for this to happen is if two pairs of SPDC photons were actually emitted; one degenerate pair and one non-degenerate pair, where the non-degenerate photon was lost on the heralding channel and the degenerate photon was lost on the heralded channel. While possible, this is not likely to be the case as the losses in the heralded path were below 50%, so both photons would make it through to the detector more frequently than shown by the heralded distribution in Fig. 8.9.

Alternatively, while the spatial mode-matching and pulse arrival time can be correct, the pulse shape of the LO and OPO signal photons may be substantially different. As the pump and LO are pulsed by two separate AOMs, it is likely that the response of the two devices is not the same. In fact, this is clearly true, as the AOM timing must be tuned to ensure the pulses arrive at the detector simultaneously as discussed in Sec. 8.2.2. Additionally, consider the shape of each pulse. The AOM in the LO path is sent a square pulse, meaning that the spectral distribution of the LO will be the Fourier transform of the square pulse — a sinc function. The spectral output of the OPO, however, will follow a Lorentzian, which will not overlap perfectly with the sinc distribution of the LO. The experimental fix to this problem would likely be to carefully engineer pulse shaping of the RF drive to the AOMs to ensure the LO and signal have the same spectral and temporal shape. Additionally, implementing the pulsing through a single AOM placed before the IR beam is doubled to green may also solve this problem, as the pulse timing and shape between the two arms of the experiment will be derived from the same source.

Although this experiment has yet to reveal the nice results of photon catalysis by showing a strong reduction in certain Fock-state components of the output distribution, the results obtained do match the model quite well for experimentally measured parameters. If the visibility can be improved through careful pulse shaping, experimental results resembling column (ii) of Fig. 8.12 could be expected. Additionally, if losses could be reduced further, the results would likely lie somewhere between columns (i) and (ii).

## 8.5 Coherent-Photon Subtraction

Chapter 5 discussed a scheme where performing coherent photon subtraction from a two-mode squeezed state led to a resource that could be used for Heisenberg-limited interferometry. This section discusses how with minor modifications, the current experimental setup can be used to generate this state.

The main difficulty lies in adapting the current experiment to generate a pure, two-mode squeezed state at a single degenerate frequency. One arm of the experiment already uses a filter cavity to remove all non-degenerate frequency modes from that beam path, so an obvious solution would be to place a second filter cavity in the other beam path. However, in order to avoid building a second cavity, it is possible to play

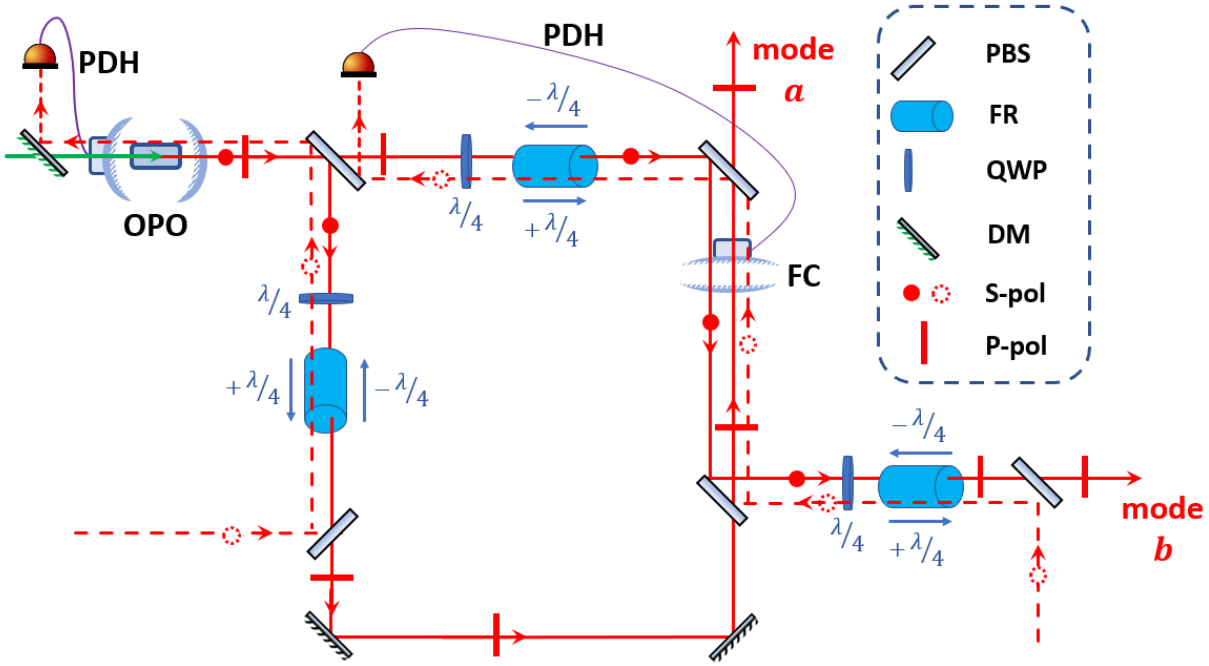



FIGURE 8.13: Proposed experimental schematic to produce frequency-degenerate two-mode squeezed light by multiplexing a single filter cavity. Careful use of Faraday rotators allows the simultaneous locking of both cavities without contaminating the squeezed light.

tricks with polarization and multiplex the same cavity. In addition to filtering out the non-degenerate frequencies, both beams would pass through the same cavity meaning that the two spatial modes would already be optimized for ideal interference as needed in the protocol in Fig. 5.2 from Ch. 5.

This could all be achieved through the addition of two Faraday rotators and additional polarization optics as shown in Fig. 8.13. This method takes advantage of the Faraday effect to rotate the polarization of a beam based on the direction of propagation as is already used in the current experiment. With the modification, both two-mirror cavities could be locked to the same LO, and the two filtered signal beams could be spatially separate from the rest of the experiment. These two modes could now be interfered and used for generating the coherent photon-subtracted resource state.

# Chapter 9

# Conclusion

In this Ph.D. dissertation, I have attempted to motivate a variety of ways to generate quantum states that are required for any type of quantum advantage in continuous-variable (CV) quantum information applications. Each method introduced has been based on experimental techniques that are currently available, and a realistic analysis of the success was performed in each case that included a majority of the expected imperfections. Additionally, I demonstrated improvements to experimental methods in non-Gaussian projections by substantially extending the limits of current photon-number-resolving (PNR) detection using highly efficient transition-edge sensors (TES) and fast, on-the-fly signal processing. With this modification, I have also made experimental progress towards realizing the quantum state engineering techniques discussed through preliminary photon catalysis experiments and quantum state tomography of benchmark states.

After introducing the basics of quantum optics and tying together ideas of qubit quantum computation (QC) with CV cluster states, I demonstrated the need for non-Gaussian resources to implement both a universal gate set and quantum error correction. Although much is dependent on the precise logical encoding of the bosonic Hilbert space, Gottesman-Kitaev-Preskill (GKP) states and rotation-symmetric states emerge as leading contenders for implementing fault-tolerant CVQC, as both encodings can correct against all error sets. The GKP codes are attractive in that a universal gate set in this encoding is Gaussian only, but the logical qubits for rotation-symmetric codes could be produced exactly without the finite-energy approximations the GKP states necessitate.

Although interesting in their own right for a variety of other quantum information applications, Schrödinger cat states are a highly desirable resource as these states

can be used either for certain rotation-symmetric encodings or as precursors to GKP states. Once one has a supply of cat states, many can be used to distill a logical GKP state. With this in mind, Ch. 2 discussed the method of photon catalysis, which is a promising method to generate large cat states in addition to other states with higher degrees of phase-space symmetry. While photon catalysis does not require squeezing resources, it does require single-photon states which are challenging to create on-demand. It is likely that photon catalysis will remain useful as a quantum state engineering technique, but it is the author's opinion that other methods will be more optimal for generating cat states both in terms of resource efficiency and current experimental capabilities.

With the limitations of photon catalysis in mind, Ch. 3 developed a method of performing photon subtraction on a cluster state. Surprisingly, this method acts to distill cat states, and can in fact proceed in the presence of Gaussian noise. This technique, which was dubbed the Photon-Assisted Node-Teleportation Method (PhANTM), was shown to near-deterministically produce cat states provided that a sufficient portion of the cluster state was used as an overhead resource. This method not only creates cat states but also embeds them within the cluster state where they can be directly used for QC algorithms, such as breeding for GKP state distillation.

Chapters 4 and 5 both relied on a similar technique to the photon subtraction in the previous chapter. Here, however, photon subtraction was performed coherently between a pair of modes. Unlike the PhANTM, the technique in these chapters can also be viewed as specific examples of a Gaussian boson sampling (GBS) device. While other methods beyond this work have extensively examined GBS devices and even devised machine learning methods to optimize certain state generation methods, the GBS device is classically inefficient to simulate. This, along with the propensity for complex optimization algorithms to be ill-defined or become stuck in local minima, indicates that specific examples of resource-efficient GBS devices are useful. In the first example in Ch. 4, coherent-photon subtraction from a two-mode squeezed state is demonstrated to have the capacity to generate rotation-symmetric states such as binomial code states when an additional PNR detector is used. Ch. 5 motivated a class of quantum states useful for Heisenberg-limited quantum sensing that could be demonstrated with a modified version of coherent-photon subtraction using only a single PNR detector.

After these chapters, the remainder of this dissertation transitioned to experimental techniques. Ch. 6 discussed PNR detection and current implementations with a

TES. By improving data collection and processing methods to measure output pulse area in real-time, the resolving capability for each TES channel was increased from roughly 8 photons to approximately 30. By multiplexing several channels, it is feasible to reach PNR detection over the entire range of 0-100 photons, which will be the subject of upcoming work. This has far-reaching implications for quantum state engineering, metrology, and many other quantum information applications including quantum random number generation.

Once any type of state is generated, it is vital to be able to fully characterize this state. In principle, one should have the ability to characterize even entirely unknown states. Ch. 7 discussed methods of using PNR detection to perform more efficient state characterization. A leading candidate for experimental use was overlap tomography using parity measurement, which has since been demonstrated. Additionally, overlap tomography was extended to fully use the PNR capabilities by looking at overlap measurements in the displaced Fock-state basis, which would yield a quadratic reduction in the number of measurement settings required to perform tomography.

Finally, Ch. 8 showcased current experiment progress on several fronts, including a successful verification of the previous single-photon tomography experiment with pulsed light. This chapter discussed the details of pure single-photon generation, including the use of an optical parametric oscillator for cavity enhancement and an external filter cavity for spectral filtering. Once generated through heralding detection, the single photon can be used for a variety of applications, including the demonstration of photon catalysis.

Photon catalysis was attempted for quantum state engineering, but there are several limiting factors to the current results. Due to the slow cool-down time of the TES, low single-photon heralding rates, and additional post-selection from the catalysis detection, collecting statistics for a single photon-number distribution of a particular output state required several hours of continuous measurement to achieve the desired accuracy due to finite-sampling errors. This imparted additional strain to experimental stability and ruled out the ability to perform full state tomography which would require the precise measurement of probability distributions at dozens of measurement settings. In addition to the sampling problem, it was determined that interference between the single-photon signal and coherent state was drastically diminished due to pulse shape mismatch, likely originating from the pair of acousto-optic modulators

used for pulsing. Despite these challenges, preliminary data indicates that if the inference can be improved, the photon catalysis experiment will yield the desired results. At the end of the chapter, I proposed a small modification to the current experimental set-up that may be used to generate a pure two-mode-squeezed state without additional cavities, which could then be immediately used for the protocols introduced in chapters 4 and 5.

Overall, this dissertation developed novel theoretical and experimental techniques for developing CV quantum information resources. It is the hope that future work will continue to bridge the gap between experimental resources and ideal theoretical concepts with the eventual goal of developing fault-tolerant universal quantum computation.

# Bibliography


1. Braunstein, S. L. & van Loock, P. Quantum information with continuous variables. *Rev. Mod. Phys.* **77,** 513 (2005).
2. Ferraro, A., Olivares, S. & Paris, M. G. Gaussian states in continuous variable quantum information. *arXiv preprint quant-ph/0503237* (2005).
3. Weedbrook, C. *et al.* Gaussian quantum information. *Rev. Mod. Phys.* **84,** 621–669 (2 2012).
4. Leonhardt, U. *Essential quantum optics: from quantum measurements to black holes* (Cambridge University Press, 2010).
5. Nielsen, M. A. & Chuang, I. L. *Quantum computation and quantum information* (Cambridge University Press, Cambridge, U.K., 2000).
6. Gerry, C., Knight, P. & Knight, P. L. *Introductory quantum optics* (Cambridge university press, 2005).
7. Bachor, H.-A. & Ralph, T. C. *A guide to experiments in quantum optics* (John Wiley & Sons, 2019).
8. Shankar, R. *Principles of quantum mechanics* (Springer Science & Business Media, 2012).
9. Mari, A. & Eisert, J. Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient. *Phys. Rev. Lett.* **109,** 230503 (23 2012).
10. D. Bartlett, S. D., C. Sanders, B. C., L. Braunstein, S. L. & Nemoto, K. Efficient classical simulation of continuous variable quantum information processes. *Phys. Rev. Lett.* **88,** 097904 (2002).
11. Hudson, R. L. When is the Wigner quasi-probability density non-negative? *Rep. Math. Phys.* **6,** 249–252. ISSN: 0034-4877 (1974).
12. Chabaud, U., Ferrini, G., Grosshans, F. & Markham, D. Classical simulation of Gaussian quantum circuits with non-Gaussian input states. *Physical Review Research* **3,** 033018 (2021).

13. Boyd, R. W. *Nonlinear Optics* 3rd Edition (Academic Press, 2008).

14. Vahlbruch, H., Mehmet, M., Danzmann, K. & Schnabel, R. Detection of 15 dB Squeezed States of Light and their Application for the Absolute Calibration of Photoelectric Quantum Efficiency. *Phys. Rev. Lett.* **117,** 110801 (11 2016).

15. M.S. Kim, W. Son, V. Bužek & P.L. Knight. Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement. *Phys. Rev. A* **65,** 032323 (2002).

16. Braunstein, S. L. Squeezing as an irreducible resource. *Phys. Rev. A* **71,** 055801 (2005).

17. Einstein, A., Podolsky, B. & Rosen, N. Can quantum-mechanical description of physical reality be considered complete? *Physical review* **47,** 777 (1935).

18. Walshe, B. W., Baragiola, B. Q., Alexander, R. N. & Menicucci, N. C. Continuous-variable gate teleportation and bosonic-code error correction. *Physical Review A* **102,** 062411 (2020).

19. Barnett, S. M. & Radmore, P. M. *Methods in Theoretical Quantum Optics* (Clarendon Press, Oxford, 1997).

20. Manzano, D. A short introduction to the Lindblad master equation. *AIP Advances* **10,** 025106 (2020).

21. Michael, M. H. *et al.* New Class of Quantum Error-Correcting Codes for a Bosonic Mode. *Phys. Rev. X* **6,** 031006 (3 2016).

22. Feynman, R. P. Simulating Physics With Computers. *Int. J. Theor. Phys.* **21,** 467–488 (1982).

23. Shor, P. W. *Algorithms for quantum computation: discrete logarithms and factoring* in *Proceedings, 35$^{th}$ Annual Symposium on Foundations of Computer Science* (ed S. Goldwasser) (IEEE Press, Los Alamitos, CA, Santa Fe, NM, 1994), 124–134.

24. Arute, F. *et al.* Quantum supremacy using a programmable superconducting processor. *Nature* **574,** 505–510 (2019).

25. Zhong, H.-S. *et al.* Quantum computational advantage using photons. *Science* **370,** 1460–1463. ISSN: 0036-8075. eprint: https://science.sciencemag.org/content/370/6523/1460.full.pdf (2020).

26. Wu, Y. *et al.* Strong quantum computational advantage using a superconducting quantum processor. *Physical review letters* **127,** 180501 (2021).

27. Zhong, H.-S. *et al.* Phase-programmable Gaussian boson sampling using stimulated squeezed light. *Physical review letters* **127,** 180502 (2021).

28. Preskill, J. Quantum computing in the NISQ era and beyond. *Quantum* **2,** 79 (2018).

29. Preskill, J. Quantum computing 40 years later. *arXiv preprint arXiv:2106.10522* (2021).

30. Aaronson, S. What Makes Quantum Computing So Hard to Explain? *Quantum Magazine* (2021).

31. Mooney, G. J., Hill, C. D. & Hollenberg, L. C. Entanglement in a 20-qubit superconducting quantum computer. *Scientific reports* **9,** 1–8 (2019).

32. Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Physical Review Letters* **86,** 5188 (2001).

33. Raussendorf, R. *Measurement-based quantum computation with cluster states* PhD thesis (Ludwig–Maximilians–Universität München, 2003).

34. Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86,** 5188 (2001).

35. Bell, J. S. in, 196–200 (Cambridge University Press, 1987).

36. M. Greenberger, D. M., A. Horne, M. A. & A. Zeilinger. in *Bell's Theorem, Quantum theory and Conceptions of the Universe* (ed Kafatos, M.) 69 (Kluwer, Dordrecht, The Netherlands, 1989).

37. Pfister, O. Continuous-variable quantum computing in the quantum optical frequency comb. *Journal of Physics B: Atomic, Molecular and Optical Physics* **53,** 012001 (2019).

38. Gu, M., Weedbrook, C., Menicucci, N. C., Ralph, T. C. & van Loock, P. Quantum computing with continuous-variable clusters. *Phys. Rev. A* **79,** 062318 (2009).

39. Ukai, R., Yoshikawa, J.-i., Iwata, N., van Loock, P. & Furusawa, A. Universal linear Bogoliubov transformations through one-way quantum computation. *Physical review A* **81,** 032315 (2010).

40. Alexander, R. N., Armstrong, S. C., Ukai, R. & Menicucci, N. C. Noise analysis of single-mode Gaussian operations using continuous-variable cluster states. *Phys. Rev. A* **90,** 062324 (6 2014).

41. van Loock, P., Weedbrook, C. & Gu, M. Building Gaussian cluster states by linear optics. *Phys. Rev. A* **76,** 032321 (2007).

42. Menicucci, N. C., Flammia, S. T. & van Loock, P. Graphical calculus for Gaussian pure states. *Phys. Rev. A* **83,** 042335 (2011).

43. Hein, M., Eisert, J. & Briegel, H. J. Multiparty entanglement in graph states. *Phys. Rev. A* **69,** 062311 (2004).

44. Menicucci, N. C. *et al.* Universal quantum computation with continuous-variable cluster states. *Phys. Rev. Lett.* **97,** 110501 (2006).

45. Gottesman, D. The Heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006* (1998).

46. García-Álvarez, L., Calcluth, C., Ferraro, A. & Ferrini, G. Efficient simulatability of continuous-variable circuits with large Wigner negativity. *Physical Review Research* **2,** 043322 (2020).

47. Sefi, S. & van Loock, P. How to Decompose Arbitrary Continuous-Variable Quantum Operations. *Phys. Rev. Lett.* **107,** 170501 (17 2011).

48. Gottesman, D., Kitaev, A. & Preskill, J. Encoding a Qubit in an Oscillator. *Phys. Rev. A* **64,** 012310 (2001).

49. Yoshikawa, J. *et al.* Demonstration of a quantum nondemolition sum gate. *Phys. Rev. Lett.* **101,** 250501 (2008).

50. Yukawa, M., Ukai, R., van Loock, P. & Furusawa, A. Experimental generation of four-mode continuous-variable cluster states. *Phys. Rev. A* **78,** 012301 (2008).

51. Pysher, M., Miwa, Y., Shahrokhshahi, R., Bloomer, R. & Pfister, O. Parallel generation of quadripartite cluster entanglement in the optical frequency comb. *Phys. Rev. Lett.* **107,** 030505 (2011).

52. Su, X. *et al.* Experimental preparation of eight-partite cluster state for photonic qumodes. *Opt. Lett.* **37,** 5178–5180 (2012).

53. Chen, M., Menicucci, N. C. & Pfister, O. Experimental realization of multipartite entanglement of 60 modes of a quantum optical frequency comb. *Phys. Rev. Lett.* **112,** 120505 (12 2014).

54. Yokoyama, S. *et al.* Ultra-large-scale continuous-variable cluster states multiplexed in the time domain. *Nat. Photon.* **7,** 982 (2013).

55. Yoshikawa, J.-i. *et al.* Invited Article: Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing. *APL Photonics* **1,** 060801 (2016).

56. Asavanant, W. *et al.* Generation of time-domain-multiplexed two-dimensional cluster state. *Science* **366,** 373–376 (2019).

57. Larsen, M. V., Guo, X., Breum, C. R., Neergaard-Nielsen, J. S. & Andersen, U. L. Deterministic generation of a two-dimensional cluster state. *Science* **366,** 369–372 (2019).

58. Asavanant, W. *et al.* One-hundred step measurement-based quantum computation multiplexed in the time domain with 25 MHz clock frequency. *arXiv preprint arXiv:2006.11537* (2020).

59. Wang, P., Chen, M., Menicucci, N. C. & Pfister, O. Weaving quantum optical frequency combs into hypercubic cluster states. *Phys. Rev. A* **90,** 032325 (2014).

60. Alexander, R. N., Yokoyama, S., Furusawa, A. & Menicucci, N. C. Universal quantum computation with temporal-mode bilayer square lattices. *Physical Review A* **97,** 032302 (2018).

61. Larsen, M. V., Guo, X., Breum, C. R., Neergaard-Nielsen, J. S. & Andersen, U. L. Deterministic multi-mode gates on a scalable photonic quantum computing platform. *Nature Physics,* 1–6 (2021).

62. Alexander, R. N. & Menicucci, N. C. Flexible quantum circuits using scalable continuous-variable cluster states. *Phys. Rev. A* **93,** 062326 (6 2016).

63. Wu, B.-H., Alexander, R. N., Liu, S. & Zhang, Z. Quantum computing with multidimensional continuous-variable cluster states in a scalable photonic platform. *Physical Review Research* **2,** 023138 (2020).

64. Fukui, K., Asavanant, W. & Furusawa, A. Temporal-mode continuous-variable three-dimensional cluster state for topologically protected measurement-based quantum computation. *Physical Review A* **102,** 032614 (2020).

65. Larsen, M. V., Chamberland, C., Noh, K., Neergaard-Nielsen, J. S. & Andersen, U. L. Fault-Tolerant Continuous-Variable Measurement-based Quantum Computation Architecture. *PRX Quantum* **2,** 030325 (2021).

66. Walshe, B. W., Alexander, R. N., Menicucci, N. C. & Baragiola, B. Q. Streamlined quantum computing with macronode cluster states. *Physical Review A* **104,** 062427 (2021).

67. Niset, J., Fiurášek, J. & Cerf, N. J. No-go theorem for Gaussian quantum error correction. *Phys. Rev. Lett.* **102,** 120501 (2009).

68. Grimsmo, A. L., Combes, J. & Baragiola, B. Q. Quantum Computing with Rotation-Symmetric Bosonic Codes. *Phys. Rev. X* **10,** 011058 (1 2020).

69. Baragiola, B. Q., Pantaleoni, G., Alexander, R. N., Karanjai, A. & Menicucci, N. C. All-Gaussian Universality and Fault Tolerance with the Gottesman-Kitaev-Preskill Code. *Phys. Rev. Lett.* **123,** 200502 (20 2019).

70. Hastrup, J., Larsen, M. V., Neergaard-Nielsen, J. S., Menicucci, N. C. & Andersen, U. L. Unsuitability of cubic phase gates for non-Clifford operations on Gottesman-Kitaev-Preskill states. *Physical Review A* **103,** 032409 (2021).

71. Tzitrin, I., Bourassa, J. E., Menicucci, N. C. & Sabapathy, K. K. Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes. *Physical Review A* **101,** 032315 (2020).

72. Mensen, L. J., Baragiola, B. Q. & Menicucci, N. C. Phase-space methods for representing, manipulating, and correcting Gottesman-Kitaev-Preskill qubits. *Physical Review A* **104,** 022408 (2021).

73. Grimsmo, A. L. & Puri, S. Quantum Error Correction with the Gottesman-Kitaev-Preskill Code. *PRX Quantum* **2,** 020101 (2021).

74. Menicucci, N. C. Fault-tolerant measurement-based quantum computing with continuous-variable cluster states. *Phys. Rev. Lett.* **112,** 120504 (12 2014).

75. Matsuura, T., Yamasaki, H. & Koashi, M. Equivalence of approximate Gottesman-Kitaev-Preskill codes. *Physical Review A* **102,** 032408 (2020).

76. Royer, B., Singh, S. & Girvin, S. Stabilization of finite-energy Gottesman-Kitaev-Preskill states. *Physical Review Letters* **125,** 260509 (2020).

77. Noh, K. & Chamberland, C. Fault-tolerant bosonic quantum error correction with the surface–Gottesman-Kitaev-Preskill code. *Physical Review A* **101,** 012316 (2020).

78. Walshe, B. W., Mensen, L. J., Baragiola, B. Q. & Menicucci, N. C. Robust fault tolerance for continuous-variable cluster states with excess antisqueezing. *Phys. Rev. A* **100,** 010301 (1 2019).

79. Fukui, K., Tomita, A., Okamoto, A. & Fujii, K. High-Threshold Fault-Tolerant Quantum Computation with Analog Quantum Error Correction. *Phys. Rev. X* **8,** 021054 (2 2018).

80. Fukui, K. High-threshold fault-tolerant quantum computation with the GKP qubit and realistically noisy devices. *arXiv:1906.09767 [quant-ph]* (2019).

81. Noh, K., Girvin, S. M. & Jiang, L. Encoding an Oscillator into Many Oscillators. *Phys. Rev. Lett.* **125,** 080503 (8 2020).

82. Vasconcelos, H. M., Sanz, L. & Glancy, S. All-optical generation of states for "Encoding a qubit in an oscillator". *Opt. Lett.* **35,** 3261–3263 (2010).

83. Weigand, D. J. & Terhal, B. M. Generating grid states from Schrödinger-cat states without postselection. *Phys. Rev. A* **97,** 022341 (2 2018).

84. Eaton, M., Nehra, R. & Pfister, O. Non-Gaussian and Gottesman-Kitaev-Preskill state preparation by photon catalysis. *New Journal of Physics* **21,** 113034 (2019).

85. Su, D., Myers, C. R. & Sabapathy, K. K. Conversion of Gaussian states to non-Gaussian states using photon-number-resolving detectors. *Phys. Rev. A* **100,** 052301 (5 2019).

86. Fukui, K. *et al.* Efficient backcasting search for optical quantum state synthesis. *arXiv preprint arXiv:2109.01444* (2021).

87. Flühmann, C. *et al.* Encoding a qubit in a trapped-ion mechanical oscillator. *Nature* **566,** 513–517 (2019).

88. Campagne-Ibarcq, P. *et al.* Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature* **584,** 368–372 (2020).

89. De Neeve, B., Nguyen, T.-L., Behrle, T. & Home, J. P. Error correction of a logical grid state qubit by dissipative pumping. *Nature Physics* **18,** 296–300 (2022).

90. Dakna, M., Knöll, L. & Welsch, D.-G. Quantum state engineering using conditional measurement on a beam splitter. *Eur. Phys. J. D* **3,** 295–308. ISSN: 1434-6060, 1434-6079 (Sept. 1998).

91. Ourjoumtsev, A., Tualle-Brouri, R., Laurat, J. & Grangier, P. Generating optical Schrödinger kittens for quantum information processing. *Science* **312,** 83–86 (2006).

92. Thekkadath, G., Bell, B., Walmsley, I. & Lvovsky, A. Engineering Schrödinger cat states with a photonic even-parity detector. *Quantum* **4,** 239 (2020).

93. Takase, K., Yoshikawa, J.-i., Asavanant, W., Endo, M. & Furusawa, A. Generation of optical Schrödinger cat states by generalized photon subtraction. *Physical Review A* **103,** 013710 (2021).

94. Motes, K. R., Baragiola, B. Q., Gilchrist, A. & Menicucci, N. C. Encoding qubits into oscillators with atomic ensembles and squeezed light. *Phys. Rev. A* **95,** 053819 (5 2017).

95. Shi, Y., Chamberland, C. & Cross, A. Fault-tolerant preparation of approximate GKP states. *New Journal of Physics* **21,** 093007 (2019).

96. Ourjoumtsev, A., Jeong, H., Tualle-Brouri, R. & Grangier, P. Generation of optical 'Schrödinger cats' from photon number states. *Nature (London)* **448,** 784 (2007).

97. Takahashi, H. *et al.* Generation of Large-Amplitude Coherent-State Superposition via Ancilla-Assisted Photon Subtraction. *Phys. Rev. Lett.* **101,** 233605 (Dec. 2008).

98. Gerrits, T. *et al.* Generation of optical coherent-state superpositions by number-resolved photon subtraction from the squeezed vacuum. *Phys. Rev. A* **82,** 031802 (3 2010).

99. Etesse, J., Bouillard, M., Kanseri, B. & Tualle-Brouri, R. Experimental generation of squeezed cat states with an operation allowing iterative growth. *Physical review letters* **114,** 193602 (2015).

100. Huang, K. *et al.* Optical synthesis of large-amplitude squeezed coherent-state superpositions with minimal resources. *Physical review letters* **115,** 023602 (2015).

101. Ulanov, A. E., Fedorov, I. A., Sychev, D., Grangier, P. & Lvovsky, A. Loss-tolerant state engineering for quantum-enhanced metrology via the reverse Hong–Ou–Mandel effect. *Nature communications* **7,** 1–6 (2016).

102. Sychev, D. V. *et al.* Enlargement of optical Schrödinger's cat states. *Nat. Photon.* **11,** 379–382. ISSN: 1749-4893 (June 2017).

103. Knill, E., Laflamme, R. & Milburn, G. A scheme for efficient quantum computation with linear optics. *Nature (London)* **409,** 46–52 (2001).

104. Bourassa, J. E. *et al.* Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer. *Quantum* **5,** 392. ISSN: 2521-327X (Feb. 2021).

105. Takeda, S. & Furusawa, A. Toward large-scale fault-tolerant universal photonic quantum computing. *APL Photonics* **4,** 060902 (2019).

106. Albert, V. V. *et al.* Performance and structure of single-mode bosonic codes. *Phys. Rev. A* **97,** 032346 (3 2018).

107. Pantaleoni, G., Baragiola, B. Q. & Menicucci, N. C. Modular bosonic subsystem codes. *Physical Review Letters* **125,** 040501 (2020).

108. Hillmann, T., Quijandría, F., Grimsmo, A. L. & Ferrini, G. Performance of teleportation-based error correction circuits for bosonic codes with noisy measurements. *arXiv preprint arXiv:2108.01009* (2021).

109. Knill, E. & Laflamme, R. Theory of quantum error-correcting codes. *Physical Review A* **55,** 900 (1997).

110. Eisert, J., Scheel, S. & Plenio, M. B. Distilling Gaussian States with Gaussian Operations is Impossible. *Phys. Rev. Lett.* **89,** 137903 (2002).

111. Lvovsky, A. I. *et al.* Quantum state reconstruction of the single-photon Fock state. *Phys. Rev. Lett.* **87,** 050402 (2001).

112. Laiho, K., Cassemiro, K. N., Gross, D. & Silberhorn, C. Probing the Negative Wigner Function of a Pulsed Single Photon Point by Point. *Phys. Rev. Lett.* **105,** 253603 (25 2010).

113. Morin, O., D'Auria, V., Fabre, C. & Laurat, J. High-fidelity single-photon source based on a Type II optical parametric oscillator. *Opt. Lett.* **37,** 3738–3740 (2012).

114. Nehra, R. *et al.* State-independent quantum state tomography by photon-number-resolving measurements. *Optica* **6,** 1356–1360 (2019).

115. Zavatta, A., Viciani, S. & Bellini, M. Quantum-to-Classical Transition with Single-Photon–Added Coherent States of Light. *Science* **306,** 660 (2004).

116. Wenger, J., Tualle-Brouri, R. & Grangier, P. Non-Gaussian Statistics from Individual Pulses of Squeezed Light. *Phys. Rev. Lett.* **92,** 153601 (2004).

117. Averchenko, V., Jacquard, C., Thiel, V., Fabre, C. & Treps, N. Multimode theory of single-photon subtraction. *New J. Phys.* **18,** 083042 (2016).

118. Lvovsky, A. I. & Mlynek, J. Quantum-Optical Catalysis: Generating Nonclassical States of Light by Means of Linear Optics. *Phys. Rev. Lett.* **88,** 250401 (25 2002).

119. Pegg, D. T., Phillips, L. S. & Barnett, S. M. Optical State Truncation by Projection Synthesis. *Phys. Rev. Lett.* **81,** 1604–1606 (8 1998).

120. Bartley, T. J. *et al.* Multiphoton state engineering by heralded interference between single photons and coherent states. en. *Physical Review A* **86.** ISSN: 1050-2947, 1094-1622 (Oct. 2012).

121. Hu, L.-Y., Wu, J.-N., Liao, Z. & Zubairy, M. S. Multiphoton catalysis with coherent state input: nonclassicality and decoherence. en. *Journal of Physics B: Atomic, Molecular and Optical Physics* **49,** 175504. ISSN: 0953-4075 (Aug. 2016).

122. Birrittella, R. J., Baz, M. E. & Gerry, C. C. Photon catalysis and quantum state engineering. *J. Opt. Soc. Am. B* **35,** 1514–1524 (2018).

123. Bartley, T. J. & Walmsley, I. A. Directly comparing entanglement-enhancing non-Gaussian operations. en. *New Journal of Physics* **17,** 023038. ISSN: 1367-2630 (Feb. 2015).

124. Lita, A. E., Miller, A. J. & Nam, S. W. Counting near-infrared single-photons with 95% efficiency. *Opt. Expr.* **16,** 3032–3040 (2008).

125. Alsing, P. M., Birrittella, R. J., Gerry, C. C., Mimih, J. & Knight, P. L. Extending the Hong-Ou-Mandel effect: The power of nonclassicality. *Physical Review A* **105,** 013712 (2022).

126. Caves, C. M. Quantum-Mechanical Radiation-Pressure Fluctuations in an Interferometer. *Phys. Rev. Lett.* **45,** 75 (1980).

127. Paris, M. G. A. Displacement operator by beam splitter. *Phys. Lett. A* **217,** 78–80. ISSN: 0375-9601 (July 1996).

128. Sharma, K. & Wilde, M. M. Characterizing the performance of continuous-variable Gaussian quantum gates. *Physical Review Research* **2,** 013126 (2020).

129. Johansson, J. R., Nation, P. D. & Nori, F. QuTiP: An open-source Python framework for the dynamics of open quantum systems. *Comp. Phys. Comm.* **183,** 1760–1772. ISSN: 0010-4655 (Aug. 2012).

130. Killoran, N. *et al.* Strawberry fields: A software platform for photonic quantum computing. *Quantum* **3,** 129 (2019).

131. Ralph, T. C., A. Gilchrist, Milburn, G. J., Munro, W. J. & S. Glancy. Quantum computation with optical coherent states. *Phys. Rev. A* **68,** 042319 (2003).

132. Neergaard-Nielsen, J. S., Nielsen, B. M., Hettich, C., Mølmer, K. & Polzik, E. S. Generation of a superposition of odd photon number states for quantum information networks. *Phys. Rev. Lett.* **97,** 083604 (2006).

133. Lund, A. P., Jeong, H., Ralph, T. C. & Kim, M. S. Conditional production of superpositions of coherent states with inefficient photon detection. *Phys. Rev. A* **70.** ISSN: 1050-2947, 1094-1622 (Aug. 2004).

134. Oh, C. & Jeong, H. Efficient amplification of superpositions of coherent states using input states with different parities. *J. Opt. Soc. Am. B* **35,** 2933–2939 (2018).

135. Etesse, J., Blandino, R., Kanseri, B. & Tualle-Brouri, R. Proposal for a loophole-free violation of Bell's inequalities with a set of single photons and homodyne measurements. *New Journal of Physics* **16,** 053001 (2014).

136. Walls, D. F. & Milburn, G. J. *Quantum Optics* (Springer, Berlin, 1994).

137. Aharonovich, I., Englund, D. & Toth, M. Solid-state single-photon emitters. *Nature Photonics* **10,** 631–641 (2016).

138. Loredo, J. C. *et al.* Scalable performance in solid-state single-photon sources. *Optica* **3,** 433–440 (2016).

139. Somaschi, N. *et al.* Near-optimal single-photon sources in the solid state. *Nature Photonics* **10,** 340–345 (2016).

140. Ding, X. *et al.* On-Demand Single Photons with High Extraction Efficiency and Near-Unity Indistinguishability from a Resonantly Driven Quantum Dot in a Micropillar. en. *Physical Review Letters* **116.** ISSN: 0031-9007, 1079-7114 (Jan. 2016).

141. Heuck, M., Pant, M. & Englund, D. R. Temporally and spectrally multiplexed single photon source using quantum feedback control for scalable photonic quantum technologies. en. *New Journal of Physics* **20,** 063046. ISSN: 1367-2630 (June 2018).

142. Dauler, E. A. *et al.* Review of superconducting nanowire single-photon detector system design options and demonstrated performance. *Optical Engineering* **53,** 081907. ISSN: 0091-3286, 1560-2303 (June 2014).

143. Nicolich, K. L. *et al.* Universal Model for the Turn-On Dynamics of Superconducting Nanowire Single-Photon Detectors. *Phys. Rev. Applied* **12,** 034020 (3 2019).

144. Hanschke, L. *et al.* Quantum dot single-photon sources with ultra-low multi-photon probability. *npj Quantum Information* **4,** 43 (2018).

145. Laghaout, A. *et al.* Amplification of realistic Schrödinger-cat-state-like states by homodyne heralding. *Physical Review A* **87,** 043826 (2013).

146. Noh, K., Albert, V. V. & Jiang, L. Quantum capacity bounds of Gaussian thermal loss channels and achievable rates with Gottesman-Kitaev-Preskill codes. *IEEE Trans. Inform. Theory*, 1–1. ISSN: 0018-9448 (2018).

147. Hastrup, J., Neergaard-Nielsen, J. S. & Andersen, U. L. Deterministic generation of a four-component optical cat state. *Optics letters* **45,** 640–643 (2020).

148. Glancy, S. & Knill, E. Error analysis for encoding a qubit in an oscillator. *Phys. Rev. A* **73,** 012325 (2006).

149. Arrazola, J. M. *et al.* Machine learning method for state preparation and gate synthesis on photonic quantum computers. *Quantum Science and Technology* **4,** 024004 (2019).

150. Eaton, M., González-Arciniegas, C., Alexander, R. N., Menicucci, N. C. & Pfister, O. Measurement-based generation and preservation of cat and grid states within a continuous-variable cluster state. *arXiv preprint arXiv:2112.10311* (2021).

151. Oh, C. & Jeong, H. Efficient amplification of superpositions of coherent states using input states with different parities. *Journal of the Optical Society of America B* **35,** 2933. ISSN: 0740-3224, 1520-8540 (Nov. 2018).

152. Gagatsos, C. N. & Guha, S. Impossibility to produce arbitrary non-Gaussian states using zero-mean Gaussian states and partial photon number resolving detection. *Phys. Rev. Research* **3,** 043182 (4 2021).

153. Walschaers, M., Sarkar, S., Parigi, V. & Treps, N. Tailoring Non-Gaussian Continuous-Variable Graph States. *Phys. Rev. Lett.* **121,** 220501 (22 2018).

154. Walschaers, M., Parigi, V. & Treps, N. Practical framework for conditional non-gaussian quantum state preparation. *PRX Quantum* **1,** 020305 (2020).

155. Marshall, K., Pooser, R., Siopsis, G. & Weedbrook, C. Repeat-until-success cubic phase gate for universal continuous-variable quantum computation. *Phys. Rev. A* **91,** 032321 (3 2015).

156. Arzani, F., Treps, N. & Ferrini, G. Polynomial approximation of non-Gaussian unitaries by counting one photon at a time. *Physical Review A* **95,** 052352 (2017).

157. Johansson, J. R., Nation, P. D. & Nori, F. QuTiP 2: A Python framework for the dynamics of open quantum systems. *Computer Physics Communications* **184,** 1234–1240 (2013).

158. Bromley, T. R. *et al.* Applications of near-term photonic quantum computers: software and algorithms. *Quantum Science and Technology* **5,** 034010 (2020).

159. Suzuki, S., Takeoka, M., Sasaki, M., Andersen, U. L. & Kannari, F. Practical purification scheme for decohered coherent-state superpositions via partial homodyne detection. *Physical Review A* **73,** 042304 (2006).

160. Walshe, B. W., Mensen, L. J., Baragiola, B. Q. & Menicucci, N. C. Robust fault tolerance for continuous-variable cluster states with excess antisqueezing. *Physical Review A* **100,** 010301 (2019).

161. Knill, E. Scalable quantum computing in the presence of large detected-error rates. *Physical Review A* **71,** 042322 (2005).

162. Svore, K. M., Hastings, M. & Freedman, M. Faster Phase Estimation. *Quantum Information and Computation* **14** (2018).

163. Terhal, B. & Weigand, D. Encoding a qubit into a cavity mode in circuit QED using phase estimation. *Physical Review A* **93,** 012315 (2016).

164. Campagne-Ibarcq, P. *et al.* A stabilized logical quantum bit encoded in grid states of a superconducting cavity. *arXiv preprint arXiv:1907.12487* (2019).

165. Morais, L. A. *et al.* Precisely determining photon-number in real-time. *arXiv:2012.10158 [physics.ins-det].* arXiv: 2012.10158 [physics.ins-det] (2020).

166. Cahall, C. *et al.* Multi-photon detection using a conventional superconducting nanowire single-photon detector. *Optica* **4,** 1534–1535 (2017).

167. Endo, M. *et al.* Quantum detector tomography of a superconducting nanostrip photon-number-resolving detector. *Optics Express* **29,** 11728–11738 (2021).

168. Fitch, M. J., Jacobs, B. C., Pittman, T. B. & Franson, J. D. Photon-number resolution using time-multiplexed single-photon detectors. *Phys. Rev. A* **68,** 043814 (4 2003).

169. Achilles, D., Silberhorn, C., Śliwa, C., Banaszek, K. & Walmsley, I. A. Fiber-assisted detection with photon number resolution. *Opt. Lett.* **28,** 2387–2389 (2003).

170. Nehra, R., Chang, C.-H., Yu, Q., Beling, A. & Pfister, O. Photon-number-resolving segmented detectors based on single-photon avalanche-photodiodes. *Opt. Express* **28,** 3660–3675 (2020).

171. Hastrup, J. & Andersen, U. L. All-optical cat-code quantum error correction. *arXiv preprint arXiv:2108.12225* (2021).

172. Nehra, R., Eaton, M., Pfister, O. & Marandi, A. All-optical Quantum State Engineering for Rotation-symmetric Bosonic States. *arXiv preprint arXiv:2105.11035* (2021).

173. Sabapathy, K. K., Qi, H., Izaac, J. & Weedbrook, C. Production of photonic universal quantum gates enhanced by machine learning. *Physical Review A* **100,** 012326 (2019).

174. O'Driscoll, L., Nichols, R. & Knott, P. A. A hybrid machine learning algorithm for designing quantum experiments. *Quantum Machine Intelligence* **1,** 5–15 (2019).

175. Bergmann, M. & van Loock, P. Quantum error correction against photon loss using multicomponent cat states. *Phys. Rev. A* **94,** 042332 (4 2016).

176. Reinhold, P. *et al.* Error-corrected gates on an encoded qubit. *Nature Physics* **16,** 822–826 (2020).

177. Hu, L. *et al.* Quantum error correction and universal gate set operation on a binomial bosonic logical qubit. *Nature Physics* **15,** 503–508 (2019).

178. Chuang, I. L., Leung, D. W. & Yamamoto, Y. Bosonic quantum codes for amplitude damping. *Physical Review A* **56,** 1114 (1997).

179. Li, L. *et al.* Cat codes with optimal decoherence suppression for a lossy bosonic channel. *Physical review letters* **119,** 030502 (2017).

180. Bacon, D. Operator quantum error-correcting subsystems for self-correcting quantum memories. *Physical Review A* **73,** 012340 (2006).

181. Kribs, D., Laflamme, R. & Poulin, D. Unified and generalized approach to quantum error correction. *Physical review letters* **94,** 180501 (2005).

182. Aliferis, P. & Cross, A. W. Subsystem fault tolerance with the Bacon-Shor code. *Physical review letters* **98,** 220502 (2007).

183. Hofmann, H. F. Generation of a highly-phase-sensitive polarization-squeezed N-photon state by collinear parametric down-conversion and coherent photon subtraction. *Physical Review A* **74,** 013808 (2006).

184. Eaton, M., Nehra, R., Win, A. & Pfister, O. Heisenberg-limited quantum interferometry with multiphoton-subtracted twin beams. *Physical Review A* **103,** 013726 (2021).

185. Hong, C. K., Ou, Z. Y. & Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59,** 2044 (1987).

186. Carranza, R. & Gerry, C. C. Photon-subtracted two-mode squeezed vacuum states and applications to quantum optical interferometry. *J. Opt. Soc. Am. B* **29,** 2581–2587 (2012).

187. Takagi, R. & Zhuang, Q. Convex resource theory of non-Gaussianity. *Phys. Rev. A* **97,** 062337 (6 2018).

188. Fukuda, D. *et al.* Titanium-based transition-edge photon number resolving detector with 98% detection efficiency with index-matched small-gap fiber coupling. *Optics express* **19,** 870–875 (2011).

189. Gerrits, T., Lita, A., Calkins, B. & Nam, S. W. in *Superconducting devices in quantum optics* 31–60 (Springer, 2016).

190. Yukawa, M. *et al.* Generating superposition of up-to three photons for continuous variable quantum information processing. *Optics express* **21,** 5529–5535 (2013).

191. Bonneau, D., Mendoza, G. J., O'Brien, J. L. & Thompson, M. G. Effect of loss on multiplexed single-photon sources. *New Journal of Physics* **17,** 043057 (2015).

192. Joshi, C., Farsi, A., Clemmen, S., Ramelow, S. & Gaeta, A. L. Frequency multiplexing for quasi-deterministic heralded single-photon sources. *Nature communications* **9,** 1–8 (2018).

193. Kaneda, F. & Kwiat, P. G. High-efficiency single-photon generation via large-scale active time multiplexing. *Science advances* **5,** eaaw8586 (2019).

194. Caves, C. M., Thorne, K. S., Drever, R. W. P., Sandberg, V. D. & Zimmermann, M. On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle. *Rev. Mod. Phys.* **52,** 341–392 (2 1980).

195. Braginsky, V. B. & Khalili, F. Y. *Quantum Measurement* (ed Thorne, K. S.) (Cambridge University Press, 1992).

196. Lévy-Leblond, J.-M. & Balibar, F. *Quantics: Rudiments of Quantum Physics* (North-Holland, 1990).

197. Demkowicz-Dobrzański, R., Jarzyna, M. & Kołodyński, J. in (ed Wolf, E.) 345–435 (Elsevier, 2015).

198. Caves, C. M. Quantum-mechanical noise in an interferometer. *Phys. Rev. D* **23,** 1693 (1981).

199. Lang, M. D. & Caves, C. M. Optimal Quantum-Enhanced Interferometry Using a Laser Power Source. *Phys. Rev. Lett.* **111,** 173601 (17 2013).

200. Grangier, P., Slusher, R. E., Yurke, B. & LaPorta, A. Squeezed-light-enhanced polarization interferometer. *Phys. Rev. Lett.* **59,** 2153 (1987).

201. Xiao, M., Wu, L. A. & Kimble, H. J. Precision measurement beyond the shot-noise limit. *Phys. Rev. Lett.* **59,** 278 (1987).

202. The LIGO Collaboration. A gravitational wave observatory operating beyond the quantum shot-noise limit. *Nat Phys* **7,** 962–965 (2011).

203. The LIGO Collaboration. Enhanced sensitivity of the LIGO gravitational wave detector by using squeezed states of light. *Nat. Photon.* **7,** 613 (2013).

204. B. Yurke, L. McCall, S. L. & R. Klauder, J. R. SU(2) and SU(1,1) interferometers. *Phys. Rev. A* **33,** 4033 (1986).

205. Luis, A. & Sánchez-Soto, L. Quantum phase difference, phase measurements and Stokes operators. *Prog. Opt.* **41,** 421 (2000).

206. J. Holland, M. J. & K. Burnett. Interferometric detection of optical phase shifts at the Heisenberg limit. *Phys. Rev. Lett.* **71,** 1355 (1993).

207. Bouyer, P. & Kasevich, M. A. Heisenberg-limited spectroscopy with degenerate Bose-Einstein gases. *Phys. Rev. A* **56,** 1083 (1997).

208. Kim, T., Pfister, O., Holland, M. J., Noh, J. & Hall, J. L. Influence of decorrelation on Heisenberg-limited interferometry using quantum correlated photons. *Phys. Rev. A* **57,** 4004 (1998).

209. Dowling, J. P. Correlated input-port, matter-wave interferometer: Quantum-noise limits to the atom-laser gyroscope. *Phys. Rev. A* **57,** 4736 (1998).

210. Feng, S. & Pfister, O. Quantum interference of ultrastable twin optical beams. *Phys. Rev. Lett.* **92,** 203601 (2004).

211. Feng, S. & Pfister, O. Sub-shot-noise heterodyne polarimetry. *Opt. Lett.* **29,** 2800 (2004).

212. Bollinger, J. J. ., Itano, W. M., Wineland, D. J. & Heinzen, D. J. Optimal frequency measurements with maximally correlated states. *Phys. Rev. A* **54,** R4649–R4652 (6 1996).

213. Boto, A. N. *et al.* Quantum Interferometric Optical Lithography: Exploiting Entanglement to Beat the Diffraction Limit. *Phys. Rev. Lett.* **85,** 2733 (2000).

214. Mitchell, M. W., Lundeen, J. S. & Steinberg, A. M. Super-resolving phase measurements with a multiphoton entangled state. *Nature* **429,** 161–164 (2004).

215. Walther, P. *et al.* De Broglie wavelength of a non-local four-photon state. *Nature* **429,** 158–161 (2004).

216. Afek, I., Ambar, O. & Silberberg, Y. High-NOON States by Mixing Quantum and Classical Light. *Science* **328,** 879–881. ISSN: 0036-8075 (2010).

217. Lang, M. D. & Caves, C. M. Optimal quantum-enhanced interferometry. *Phys. Rev. A* **90,** 025802 (2 2014).

218. Escher, B. M., de Matos Filho, R. L. & Davidovich, L. General framework for estimating the ultimate precision limit in noisy quantum-enhanced metrology. *Nat. Phys.* **7,** 406–411 (2011).

219. Pooser, R. C. & Pfister, O. Particle-number scaling of the phase sensitivity in realistic Bayesian twin-mode Heisenberg-limited interferometry. *Phys. Rev. A* **69,** 043616 (2004).

220. Dorner, U. *et al.* Optimal quantum phase estimation. *Physical review letters* **102,** 040403 (2009).

221. Aasi, J. *et al.* Advanced ligo. *Classical and quantum gravity* **32,** 074001 (2015).

222. Drever, R. W. P. *et al.* Laser phase and frequency stabilization using an optical resonator. *Appl. Phys. B* **31,** 97–105 (1983).

223. Vahlbruch, H. *et al.* Coherent Control of Vacuum Squeezing in the Gravitational-Wave Detection Band. *Phys. Rev. Lett.* **97,** 011101 (1 2006).

224. Tse, M. *et al.* Quantum-enhanced advanced ligo detectors in the era of gravitational-wave astronomy. *Physical Review Letters* **123,** 231107 (2019).

225. Acernese, F. *et al.* Increasing the astrophysical reach of the advanced virgo detector via the application of squeezed vacuum states of light. *Physical Review Letters* **123,** 231108 (2019).

226. Grote, H. *et al.* First long-term application of squeezed states of light in a gravitational-wave observatory. *Physical review letters* **110,** 181101 (2013).

227. McCuller, L. *et al.* Frequency-Dependent Squeezing for Advanced LIGO. *Physical Review Letters* **124,** 171102 (2020).

228. Heidmann, A. *et al.* Observation of Quantum Noise Reduction on Twin Laser Beams. *Phys. Rev. Lett.* **59,** 2555–2557 (1987).

229. Mertz, J., Debuisschert, T., Heidmann, A., Fabre, C. & Giacobino, E. Improvements in the observed intensity correlation of optical parametric oscillator twin beams. *Opt. Lett.* **16,** 1234 (1991).

230. Laurat, J., Longchambon, L., Fabre, C. & Coudreau, T. Experimental investigation of amplitude and phase quantum correlations in a type II optical parametric oscillator above threshold: from nondegenerate to degenerate operation. *Opt. Lett.* **30,** 1177 (2005).

231. Campos, R. A., Saleh, B. E. A. & Teich, M. C. Quantum-mechanical lossless beam splitter: SU(2) symmetry and photon statistics. *Phys. Rev. A* **40,** 1371 (1989).

232. Jing, J., Feng, S., Bloomer, R. & Pfister, O. Experimental continuous-variable entanglement from a phase-difference-locked optical parametric oscillator. *Phys. Rev. A* **74,** 041804 (4 2006).

233. Pysher, M., Miwa, Y., Shahrokhshahi, R. & Pfister, O. A new way to lock the signal-field phase difference in a type-II optical parametric oscillator above threshold. *Opt. Expr.* **18,** 27858 (2010).

234. You, C. *et al.* Scalable multiphoton quantum metrology with neither pre-nor post-selected measurements. *Applied Physics Reviews* **8,** 041406 (2021).

235. Snyder, J. J., Giacobino, E., Fabre, C., Heidmann, A. & Ducloy, M. Sub-shot-noise measurements using the beat note between quantum-correlated photon beams. *J. Opt. Soc. Am. B* **7,** 2132–2136 (1990).

236. Thekkadath, G. S. *et al.* Quantum-enhanced interferometry with large heralded photon-number states. *npj Quantum Information* **6,** 89 (2020).

237. Kacprowicz, M., Demkowicz-Dobrzański, R., Wasilewski, W., Banaszek, K. & Walmsley, I. Experimental quantum-enhanced estimation of a lossy phase shift. *Nature Photonics* **4,** 357–360 (2010).

238. Birrittella, R. & Gerry, C. C. Quantum optical interferometry via the mixing of coherent and photon-subtracted squeezed vacuum states of light. *JOSA B* **31,** 586–593 (2014).

239. Leonhardt, U. Quantum statistics of a two-mode SU (1, 1) interferometer. *Physical Review A* **49,** 1231 (1994).

240. Li, D. *et al.* Effects of loss on the phase sensitivity with parity detection in an SU (1, 1) interferometer. *JOSA B* **35,** 1080–1092 (2018).

241. Marino, A. M., Corzo Trejo, N. V. & Lett, P. D. Effect of losses on the performance of an SU(1,1) interferometer. *Phys. Rev. A* **86,** 023844 (2 2012).

242. Adhikari, S., Bhusal, N., You, C., Lee, H. & Dowling, J. P. Phase estimation in an SU (1, 1) interferometer with displaced squeezed states. *OSA Continuum* **1,** 438–450 (2018).

243. Anderson, B. E., Schmittberger, B. L., Gupta, P., Jones, K. M. & Lett, P. D. Optimal phase measurements with bright-and vacuum-seeded SU (1, 1) interferometers. *Physical Review A* **95,** 063843 (2017).

244. Gupta, P., Schmittberger, B. L., Anderson, B. E., Jones, K. M. & Lett, P. D. Optimized phase sensing in a truncated SU (1, 1) interferometer. *Optics express* **26,** 391–401 (2018).

245. Zhang, J.-D. *et al.* Super-sensitive angular displacement estimation via an SU (1, 1)-SU (2) hybrid interferometer. *Optics Express* **26,** 33080–33090 (2018).

246. Caves, C. M. Reframing SU (1, 1) interferometry. *Advanced Quantum Technologies,* 1900138 (2019).

247. Frascella, G., Agne, S., Khalili, F. Y. & Chekhova, M. V. Overcoming detection loss and noise in squeezing-based optical sensing. *npj Quantum Information* **7,** 1–6 (2021).

248. Cramér, H. *Mathematical methods of statistics* (Princeton university press, 1999).

249. Rao, C. R. in *Breakthroughs in statistics* 235–247 (Springer, 1992).

250. Helstrom, C. W. Quantum detection and estimation theory. *Journal of Statistical Physics* **1,** 231–252 (1969).

251. Jarzyna, M. & Demkowicz-Dobrzański, R. Quantum interferometry with and without an external phase reference. *Physical Review A* **85,** 011801(R) (2012).

252. Schwinger, J. On Angular Momentum. *U.S. Atomic Energy Commission Report. No. NYO–3071 (1952), reprinted in Quantum Theory of Angular Momentum, edited by L. C. Biedenharn and H. van Dam (Academic Press, New York, 1965), pp. 229–279.*

253. Braunstein, S. L. & Caves, C. M. Statistical distance and the geometry of quantum states. *Physical Review Letters* **72,** 3439 (1994).

254. Oh, C., Lee, S.-Y., Nha, H. & Jeong, H. Practical resources and measurements for lossy optical quantum metrology. *Physical Review A* **96,** 062304 (2017).

255. Kim, M., Park, E., Knight, P. & Jeong, H. Nonclassicality of a photon-subtracted Gaussian field. *Physical Review A* **71,** 043805 (2005).

256. Gerry, C. C. Heisenberg-limit interferometry with four-wave mixers operating in a nonlinear regime. *Physical Review A* **61,** 043811 (2000).

257. Plick, W. N., Anisimov, P. M., Dowling, J. P., Lee, H. & Agarwal, G. S. Parity detection in quantum optical metrology without number-resolving detectors. *New Journal of Physics* **12,** 113025 (2010).

258. Anisimov, P. M. *et al.* Quantum Metrology with Two-Mode Squeezed Vacuum: Parity Detection Beats the Heisenberg Limit. *Phys. Rev. Lett.* **104,** 103602 (10 2010).

259. Arakawa, Y. & Holmes, M. J. Progress in quantum-dot single photon sources for quantum information technologies: A broad spectrum overview. *Applied Physics Reviews* **7,** 021309 (2020).

260. Thomas, S. & Senellart, P. The race for the ideal single-photon source is on. *Nature Nanotechnology* **16,** 367–368 (2021).

261. Tao, X. *et al.* A high speed and high efficiency superconducting photon number resolving detector. *Superconductor Science and Technology* **32,** 064002 (2019).

262. Helstrom, C. W. *Quantum Detection and Estimation Theory* (Mathematics in Science and Engineering, 123, Academic Press, New York, 1976).

263. Sridhar, N. *et al.* Direct measurement of the Wigner function by photon-number-resolving detection. *J. Opt. Soc. Am. B* **31,** B34–B40 (2014).

264. Nehra, R. *et al.* Generalized overlap quantum state tomography. *Phys. Rev. Research* **2,** 042002 (4 2020).

265. Herrero-Collantes, M. & Garcia-Escartin, J. C. Quantum random number generators. *Reviews of Modern Physics* **89,** 015004 (2017).

266. Gerry, C. C. *et al.* Proposal for a quantum random number generator using coherent light and a non-classical observable. *JOSA B* **39,** 1068–1074 (2022).

267. Royer, A. Wigner function as the expectation value of a parity operator. *Phys. Rev. A* **15,** 449–450 (2 1977).

268. *Quantum State Estimation* (eds Paris, M. & Řeháček, J.) (Springer-Verlag Berlin Heidelberg, 2004).

269. Gerchberg, R. & Saxton, W. A practical algorithm for the determination ofthe phase from image and diffraction plane pictures. *Optik* **35** (1972).

270. Cramer, M. *et al.* Efficient quantum state tomography. *Nat. Commun.* **1,** 149 (2010).

271. Wigner, E. On the Quantum Correction For Thermodynamic Equilibrium. *Phys. Rev.* **40,** 749–759 (1932).

272. Hillery, M., O'Connell, R., Scully, M. & Wigner, E. Distribution functions in physics: Fundamentals. *Physics Reports* **106,** 121–167. ISSN: 0370-1573 (1984).

273. Smithey, D. T., Beck, M., Raymer, M. G. & Faridani, A. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: Application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70,** 1244 (1993).

274. Lvovsky, A. I. & Raymer, M. G. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.* **81,** 299 (2009).

275. Wallentowitz, S. & Vogel, W. Unbalanced homodyning for quantum state measurements. *Phys. Rev. A* **53,** 4528–4533 (6 1996).

276. Banaszek, K. & Wódkiewicz, K. Direct Probing of Quantum Phase Space by Photon Counting. *Phys. Rev. Lett.* **76,** 4344 (1996).

277. Leibfried, D. *et al.* Experimental Determination of the Motional Quantum State of a Trapped Atom. *Phys. Rev. Lett.* **77,** 4281–4285 (21 1996).

278. P. Bertet *et al.* Generating and Probing a Two-Photon Fock State with a Single Atom in a Cavity. *Phys. Rev. Lett.* **88,** 143601 (2002).

279. Vlastakis, B. *et al.* Deterministically Encoding Quantum Information Using 100-Photon Schrödinger Cat States. *Science* **342,** 607–610. ISSN: 0036-8075. eprint: http://science.sciencemag.org/content/342/6158/607.full.pdf (2013).

280. Husimi, K. Some formal properties of the density matrix. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* **22,** 264–314 (1940).

281. Lundeen, J. S. *et al.* Tomography of quantum detectors. *Nat. Phys.* **5,** 27–30 (2009).

282. Grandi, S., Zavatta, A., Bellini, M. & Paris, M. G. A. Experimental quantum tomography of a homodyne detector. *New J. Phys.* **19,** 053015 (2017).

283. Zhang, L. *et al.* Recursive quantum detector tomography. *New Journal of Physics* **14,** 115005 (2012).

284. Diamond, S. & Boyd, S. CVXPY: A Python-Embedded Modeling Language for Convex Optimization. *Journal of Machine Learning Research* **17,** 1–5 (2016).

285. Johansson, J. R., Nation, P. D. & Nori, F. QuTiP 2: A Python framework for the dynamics of open quantum systems. *Computer Physics Communications* **184,** 1234–1240 (2013).

286. Nehra, R. *Non-Gaussian States and Measurements in Quantum Information Science* PhD thesis (2020).

287. Zurek, W. H. Sub-Planck structure in phase space and its relevance for quantum decoherence. *Nature* **412,** 712–717 (2001).

288. Harland, D., Everitt, M. J., Nemoto, K., Tilma, T. & Spiller, T. Towards a complete and continuous Wigner function for an ensemble of spins or qubits. *Physical Review A* **86,** 062117 (2012).

289. Driver, K. & Jordaan, K. Bounds for extreme zeros of some classical orthogonal polynomials. *Journal of Approximation Theory* **164,** 1200–1204 (2012).

290. Nyquist, H. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers* **47,** 617–644 (1928).

291. Shannon, C. E. Communication in the presence of noise. *Proceedings of the IRE* **37,** 10–21 (1949).

292. Cahill, K. E. & Glauber, R. J. Ordered expansions in boson amplitude operators. *Physical Review* **177,** 1857 (1969).

293. Benaroya, H., Han, S. M. & Nagurka, M. *Probability models in engineering and science* (CRC press, 2005).

294. Achilles, D., Silberhorn, C. & Walmsley, I. A. Direct, loss-tolerant characterization of nonclassical photon statistics. *Phys. Rev. Lett.* **97,** 043602 (2006).

295. Kiss, T., Herzog, U. & Leonhardt, U. Compensation of losses in photodetection and in quantum-state measurements. *Phys. Rev. A* **52,** 2433–2435 (3 1995).

296. Kelley, P. & Kleiner, W. Theory of electromagnetic field measurement and photoelectron counting. *Physical Review* **136,** A316 (1964).

297. Huh, J., Guerreschi, G. G., Peropadre, B., McClean, J. R. & Aspuru-Guzik, A. Boson sampling for molecular vibronic spectra. *Nat. Photon.* **9,** 615–620 (2015).

298. Wang, C. S. *et al.* Efficient Multiphoton Sampling of Molecular Vibronic Spectra on a Superconducting Bosonic Processor. *Phys. Rev. X* **10,** 021060 (2 2020).

299. Arrazola, J. M. *et al.* Quantum circuits with many photons on a programmable nanophotonic chip. *Nature* **591,** 54–60 (2021).

300. Jordan, S. P., Lee, K. S. M. & Preskill, J. Quantum Algorithms for Quantum Field Theories. *Science* **336,** 1130–1133. ISSN: 0036-8075. eprint: http://science.sciencemag.org/content/336/6085/1130.full.pdf (2012).

301. Marshall, K., Pooser, R., Siopsis, G. & Weedbrook, C. Quantum simulation of quantum field theory using continuous variables. *Phys. Rev. A* **92,** 063825 (6 2015).

302. Korpel, A. Acousto-optics—a review of fundamentals. *Proceedings of the IEEE* **69,** 48–53 (1981).

303. Slattery, O., Ma, L., Zong, K. & Tang, X. Background and review of cavity-enhanced spontaneous parametric down-conversion. *Journal of Research of the National Institute of Standards and Technology* **124,** 1 (2019).

304. Siegman, A. *Lasers* (University Science Books, Sausalito, CA, 1986).

305. Shahrokhshahi, R. *A Cavity-enhanced Narrow-band Multiphoton Source for Applications in Quantum Information* PhD thesis (2014).

306. Christ, A. & Silberhorn, C. Limits on the deterministic creation of pure single-photon states using parametric down-conversion. *Phys. Rev. A* **85,** 023829 (2 2012).

307. Laiho, K., Avenhaus, M., Cassemiro, K. & Silberhorn, C. Direct probing of the Wigner function by time-multiplexed detection of photon statistics. *New Journal of Physics* **11,** 043012 (2009).

# Appendix A

# Photon catalysis

## A.1 Numerical modeling

We begin by defining a generic beamsplitter operation, projector, and input density matrices as

$$U_{ab} = e^{\theta(ab^\dagger - a^\dagger b)} \tag{A.1}$$

$$P_n = (|n\rangle\langle n|)_a \otimes \mathbb{1}_b \tag{A.2}$$

$$\rho_a^{in} = (|\psi\rangle\langle\psi|)_a \tag{A.3}$$

$$\rho_b^{in} = (|1\rangle\langle 1|)_b \tag{A.4}$$

where $r = \cos\theta$ is the reflection coefficient. After the beamsplitter, but before the detection event, the new density matrix is

$$\rho_{ab} = U_{ab}\rho_a^{in} \otimes \rho_b^{in} U_{ab}^\dagger. \tag{A.5}$$

We then apply the projective measurement of mode $a$ by an ideal detector and normalize to obtain the resulting density matrix of mode $b$, given by

$$\rho_{out} = \frac{Tr_a[\rho_{ab}P_n]}{Tr[\rho_{ab}P_n]}. \tag{A.6}$$

Imperfect detection can be modeled by including a loss beamsplitter followed by a perfect detector as depicted in Fig.A.1, where we trace out over the lost mode. In this case we now have a vacuum input mode to consider, so a new input state of $\rho_c^{in} = (|0\rangle\langle 0|)_c$, a loss beamsplitter $U_{ac} = e^{\xi(ac^\dagger - a^\dagger c)}$ where $\eta = \cos\xi$ is the detector efficiency,

and a modified projector, $P_n^\eta = (|n\rangle\langle n|)_a \otimes \mathbb{1}_b \otimes \mathbb{1}_c$. Eq. (A.5) becomes

$$\rho_{abc} = U_{ac}\rho_{ab} \otimes \rho_c^{in} U_{ac}^\dagger. \tag{A.7}$$

As before, we apply the projective measurement of the detector, but must now also trace out over the lost beamsplitter port in addition to the detected mode. Our final state is now

$$\rho_{out,\eta} = \frac{Tr_{ac}[\rho_{abc}P_n^\eta]}{Tr[\rho_{abc}P_n^\eta]}. \tag{A.8}$$

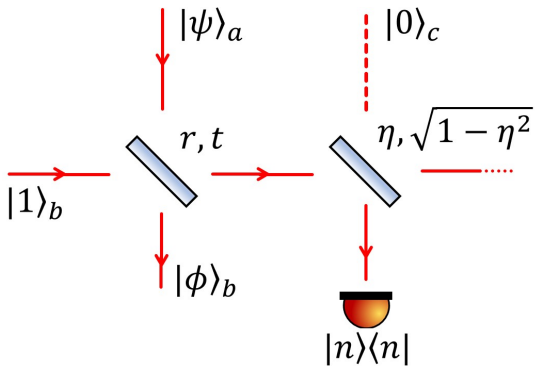



FIGURE A.1: Loss model. Photon catalysis with an imperfect detector can be modeled by placing a beamsplitter of reflectivity $\eta$ with vacuum input before an ideal detector, where the other output is lost via a partial trace.

In several sections of this work, we explore the possibility to perform cascaded photon catalysis by using the resulting density matrix from the $i^{\text{th}}$ step as the input state to the $i+1$ interference beamsplitter,

$$\rho_a^{in\,(i+1)} = \rho_{out}^{(i)} \tag{A.9}$$

This procedure allows us to independently vary the beamsplitter parameter $r_j$ and PNR detection $n_j$ at each step and obtain the intermediate resulting density matrix. After repeated the process as many times as desired, we can compute the fidelity between a target state, $\rho_T$, and the result of our process, $\rho_{out}$, which is defined by Eq. 2.7.

When performing cascaded photon catalysis calculations, we we limited the Hilbert space dimension to 50 as the tensor product between Hilbert spaces causes an exponential increase in computational expense with each additional mode added.

## A.2 Cascaded photon catalysis

Consider iterating the photon catalysis process by using the output from one step as the input to the next. Here, we derive the general quantum state at the output of an N-step photon catalysis procedure where the initial input is taken to be an arbitrary pure state, $\psi$, and each step uses a single-photon Fock state as the second input. The $i^{\text{th}}$ step is described by a beamsplitter of real reflectivity and transmissivity, $r_i$ and $t_i$, followed by a perfect PNR detector that detects $n_i$ photons. We start by giving the result and then proceed with mathematical induction as proof. After N steps, we posit that the final quantum state, up to a normalization, will be

$$|\phi_N\rangle \propto \sum_{m=0}^{\infty}\Bigg\{\psi_{m-N+\sum_{j=1}^{N}n_j}\left[\frac{[m-N+\sum_{j=1}^{N}n_j]!}{m!}\right]^{\frac{1}{2}}\prod_{k=0}^{N-1}\left[\frac{r_{k+1}^{n_{k+1}-1}}{\sqrt{n_{k+1}!}}t_{N-k}^{m-1+\sum_{j=0}^{k-1}(n_{N-j}-1)}\right.\\ \left.\times\left(n_{N-k}t_{N-k}^2-r_{N-k}^2(m+\sum_{j=0}^{k-1}(n_{N-j}-1))\right)\right]\Bigg\}|m\rangle\,. \tag{A.10}$$

By taking $N=1$, it is easy to see that this reduces to Eq. 2.2, thus A.10 is verified for a single photon catalysis step. For the proof, we now assume A.10 to be true and feed $\phi_N$ into an $(N+1)^{\text{th}}$ step. The resulting quantum state after the PNR detection of $n_{N+1}$ photons can be obtained from equation 2.2 by using the $m+n_{N+1}-1^{\text{th}}$ component of $\phi_N$ to arrive at

$${}_b\langle n_{N+1}|\,U_{ab}\,|\phi_{N+1}\rangle_a\otimes|1\rangle_b \propto \sum_{m=0}^{\infty}\Bigg\{\frac{\psi_{m+n_{N+1}-1-N+\sum_{j=1}^{N}n_j}}{\sqrt{n_{N+1}}}\binom{m+n_{N+1}-1}{m}^{\frac{1}{2}}\\ \times\left[\frac{[m+n_{N+1}-1-N+\sum_{j=1}^{N}n_j]!}{(m+n_{N+1}-1)!}\right]^{\frac{1}{2}}r_{N+1}^{n_{N+1}-1}t_{N+1}^{m-1}[n_{N+1}t_{N+1}^2-mr_{N+1}^2]\\ \times\prod_{k=0}^{N-1}\left[\frac{r_{k+1}^{n_{k+1}-1}}{\sqrt{n_{k+1}!}}t_{N-k}^{m+n_{N+1}-2+\sum_{j=0}^{k-1}(n_{N-j}-1)}\right.\\ \left.\times\left(n_{N-k}t_{N-k}^2-r_{N-k}^2(m+n_{N+1}-1+\sum_{j=0}^{k-1}(n_{N-j}-1))\right)\right]\Bigg\}|m\rangle\,. \tag{A.11}$$

Reordering the sum and product indices allows us to arrive at an expression identical to Eq. A.10, but with $N\to N+1$, and thereby complete the proof. From Eq. A.10, we can also extract the overall success probability for a particular set of detection

events:

$$P(n_1,\ldots,n_N) = \sum_{m=0}^{\infty}\left\{\psi^2_{m-N+\sum_{j=1}^{N} n_j}\left[\frac{[m-N+\sum_{j=1}^{N} n_j]!}{m!}\right]\right.$$
$$\left.\times\prod_{k=0}^{N-1}\left[\frac{r_{k+1}^{2n_{k+1}-2}}{n_{k+1}!}t_{N-k}^{2m-2+2\sum_{j=0}^{k-1}(n_{N-j}-1)}\left(n_{N-k}t_{N-k}^2-r_{N-k}^2(m+\sum_{j=0}^{k-1}(n_{N-j}-1))\right)^2\right]\right\}. \tag{A.12}$$

## A.3 Fidelity with SSV

We now derive an expression for the fidelity between a pure SSV state and the result from an $N$-step filtered coherent state. The SSV states are given by

$$|SSV_\pm\rangle = C_{SSV\pm}\left[D(\beta)\pm D(-\beta)\right]S(r)\,|0\rangle\,, \tag{A.13}$$

where $r$ is the single-mode squeezing parameter, $\beta$ is the displacement magnitude to create the superposition, and the normalization coefficient is

$$C_{SSV\pm} = \left[\frac{1}{2\pm 2e^{-2|\beta|^2e^{-2r}}}\right]^{\frac{1}{2}}. \tag{A.14}$$

As a first step, we derive the overlap coefficients between an n-photon Fock state and a displaced m-photon Fock state, which will be useful later. This is given by

$$A(n,m,\gamma) \equiv \langle n|\,D(\gamma)\,|m\rangle \tag{A.15}$$
$$= \langle n|\,D(\gamma)\frac{a^{\dagger m}}{\sqrt{m!}}D^\dagger(\gamma)\,|m\rangle \tag{A.16}$$
$$= \langle n|\,\frac{(a^\dagger-\gamma^*)^m}{\sqrt{m!}}\,|\gamma\rangle \tag{A.17}$$
$$= e^{-|\gamma|^2/2}\sum_{k=0}^{m}\binom{m}{k}\frac{(-1)^k\sqrt{n!}\,|\gamma|^{2k+n-m}}{m!(n-m+k)!}. \tag{A.18}$$

Because the photon-catalyzed coherent state must be displaced back to the origin of phase space before comparing to an SSV state (see main text), we are interested in a fidelity of the form

$$F = |\,\langle SSV_\pm|\,D^\dagger(\delta)\,|\phi_N\rangle\,|^2, \tag{A.19}$$

where $\delta$ here is the amplitude of displacement needed after the photon catalysis procedure. By substituting the coefficients of a coherent state of magnitude $\alpha$ for $\psi$ in Eq. A.10 and using the overlap coefficients given by A.18, we can arrive at the overall expression

$$F = \frac{C_{SSV\pm}^2 C_{|\phi_N\rangle}^2}{\cosh r}\left|\sum_{\ell,m=0}^{\infty}\left\{(\tanh r)^\ell \frac{(2\ell-1)!!\alpha^m}{\sqrt{(2\ell)!m!}}[A(2\ell,m,-\beta-\delta)-A(2\ell,m,\beta-\delta))]\right.\right.$$
$$\left.\left.\times\prod_{k=0}^{N-1}\left[t_{k+1}^m\left(n_{N-k}t_{N-k}^2 - r_{N-k}^2(m+\sum_{j=0}^{k-1}(n_{N-j}-1))\right)\right]\right\}\right|^2, \tag{A.20}$$

where $C_{|\phi_N\rangle}$ is determined by normalizing Eq. A.10 with a coherent state input. This equation for fidelity can now be numerically maximized for the experimentally accessible parameters of $r_i$, $\alpha$, and $\delta$, as well as the desired post-selected PNR measurements of $n_i$ for $N$ photon catalysis steps.

## A.4 Optimized parameters

The cascaded photon catalysis protocol for $N = 2, 3$, and 4 steps was numerically optimized to yield the high-fidelity SSV$^{(2)}$ states discussed in the main text. The full list of parameters for the protocols and resultant states are displayed in Table A.1.

| $N$ | 2 | 3 | 4 |
|---|---|---|---|
| Fidelity | $> 0.999$ | 0.984 | 0.977 |
| Success probability | $1.5 \times 10^{-2}$ | $1.8 \times 10^{-3}$ | $4.2 \times 10^{-5}$ |
| $\beta$ | 0.90 | 1.35 | 1.59 |
| Squeezing parameter | -0.22 | -0.48 | -0.52 |
| $\alpha$ | 1.20 | 3.54 | 4.66 |
| $\{r_1, r_2, r_3, r_4\}$ | $\{0.60, 0.85, -, -\}$ | $\{0.64, 0.49, 0.52, -\}$ | $\{0.58, 0.55, 0.70, 0.42\}$ |
| $\{n_1, n_2, n_3, n_4\}$ | $\{1, 2, -, -\}$ | $\{5, 2, 1, -\}$ | $\{6, 4, 2, 1\}$ |

TABLE A.1: Parameters of $N$-step photon-catalysis that optimize the fidelity with the nearest SSV$^{(2)}$ state. The fidelity was numerically optimized for all combinations of integer $n_i$ below 10, with all other parameters allowed to vary independently. The combinations with the highest fidelities and success probabilities are displayed. $\alpha$: input coherent state amplitude. $\beta$: target SSV state amplitude. $r_i$: real-valued reflectiviy of the $i^{\text{th}}$ beamsplitter in cascaded photon catalysis where $r_i^2 + t_i^2 = 1$. $n_i$: Measurement result of the $i^{\text{th}}$ PNR detector.

## A.5 SSV enlargement by PNR

As described in the main text, the breeding protocol for enlarging a general $M^{\text{th}}$ order symmetry SSV state consists of sending two identical $SSV^{(M)}$ states in respective modes $a$ and $b$ to a balanced beamsplitter, and performing a PNR measurement on the output mode $b$. Before the detection step, the state is given by Eq. 2.29 which can be rewritten as

$$\begin{aligned} |\phi\rangle \propto & \sum_{n=0}^{M-1} U_{ab} D_a(\beta_n) D_b(\beta_n) S_a(\xi_n) S_b(\xi_n) |0\rangle_a |0\rangle_b \\ & + \sum_{n_1 \neq n_2}^{M-1} U_{ab} D_a(\beta_{n_1}) D_b(\beta_{n_2}) S_a(\xi_{n_1}) S_b(\xi_{n_2}) |0\rangle_a |0\rangle_b , \end{aligned} \tag{A.21}$$

where we note that $|\beta_{n_1}| = |\beta_{n_2}|$ and $|\xi_{n_1}| = |\xi_{n_2}|$ for all values of $n_1$ and $n_2$. Using a balanced beamsplitter defined by $U_{ab} = e^{\frac{\pi}{4}(ab^\dagger - a^\dagger b)}$, it can be seen that identical displacers and squeezers on separate modes are transformed in the Heisenberg picture

according to

$$U_{ab}D_a(\beta_n)D_b(\beta_n)U_{ab}^\dagger = D_a(\sqrt{2}\beta_n) \tag{A.22}$$

$$U_{ab}S_a(\xi_n)S_b(\xi_n)U_{ab}^\dagger = S_a(\xi_n)S_b(\xi_n), \tag{A.23}$$

so after detecting zero photons in mode $b$, the first sum in Eq. A.21 becomes

$$\frac{1}{\sqrt{\cosh|\xi|}}\sum_{n=0}^{M-1} D_a(\sqrt{2}\beta_n)S_a(\xi_n)\left|0\right\rangle_a. \tag{A.24}$$

If we can show that the terms from the second sum in Eq. A.21 are negligible, then normalizing Eq. A.24 will give us the desired enlarged $SSV^{(M)}$ state. The beamsplitter acting on the two unequal displacers gives

$$U_{ab}D_a(\beta_{n_1})D_b(\beta_{n_2})U_{ab}^\dagger = D_a(\frac{\beta_+}{\sqrt{2}})D_b(\frac{\beta_-}{\sqrt{2}}) \tag{A.25}$$

where $\beta_+ = \beta_{n_1} + \beta_{n_2}$ and $\beta_- = \beta_{n_1} - \beta_{n_2}$. Applying Eq. (13) and (14) from Ref. [15] and writing the squeezing arguments as $\xi_{n1} = -re^{2i\varphi}$ and $\xi_{n2} = -re^{2i\theta}$, we see that the two unequal squeezers on vacuum are transformed by

$$\begin{aligned} &U_{ab}S_a(-re^{2i\varphi})S_b(-re^{2i\theta})\left|0\right\rangle_a\left|0\right\rangle_b = R_a(\varphi)R_b(\theta)S_a\Big(-\frac{r}{2}(1+e^{2i(\varphi-\theta)})\Big) \\ &\times S_b\Big(-\frac{r}{2}(1+e^{-2i(\varphi-\theta)})\Big)S_{ab}\Big(-2r\sinh(\varphi-\theta)\Big)\left|0\right\rangle_a\left|0\right\rangle_b, \end{aligned} \tag{A.26}$$

where $R(\Theta) = e^{i\Theta a^\dagger a}$ is the phase-shift operator. Detecting zero photon in mode $b$ leads to each term in the second sum of Eq. A.21 taking the form:

$$D_a(\beta_+)S_a(\varsigma)\,{}_b\langle 0|\,D_b(\beta_-)S_b(\varsigma^*)S_{ab}(\Delta)\left|0\right\rangle_b\left|0\right\rangle_a, \tag{A.27}$$

where $\varsigma = -\frac{r}{2}(1+e^{2i(\varphi-\theta)})$ and $\triangle = -2r\sinh(\varphi-\theta)$. Now, realizing that $\beta_- = \frac{|\beta|}{\sqrt{2}}(e^{\frac{2\pi i n_1}{M}} - e^{\frac{2\pi i n_2}{M}})$ is always nonzero for $n_1 \neq n_2$ when $n_1, n_2 < M$, we see that each of the terms having the form of Eq. A.27 consists of the overlap of vacuum with two and single-mode squeezed vacuum that has been displaced. For finite squeezing, it is clear that this overlap term can always be made arbitrarily small for a large enough displacement amplitude $|\beta_-|$. If this is the case, then the entire second sum in Eq. A.21 vanishes, and detecting zero photons in mode $b$ reduces A.21 to A.24.

The only question that remains is how large of an initial displacement amplitude, $|\beta|$, is considered large enough, as it can be seen that increasing $M$ leads to more terms in the sum that we would wish to neglect. In practice, this can be determined numerically for a desired fidelity threshold with the target enlarged $SSV^{(M)}$, and we see from Fig.2.9 in the main text that $|\beta| \geq \sqrt{2}$ for $M = 2$ and 3, and $|\beta| \geq \sqrt{5}$ for $M = 4$ is sufficient for high-fidelity breeding.

## A.6 Kraus Operator Formalism

The derivation for photon catalysis presented in the main text gives the Schrödinger evolution of the input states interfered at a beamsplitter followed by PNR measurement on one mode. However, this process can instead be represented as a Kraus operator acting on a single input. With this representation, alternative physical insight into the process can be gained. To derive the Kraus operator, first start from Eq. 2.1 given in the main text, and project mode 2 onto the $n$-photon Fock state to get

$$ {}_2\langle n|\phi\rangle_{12} = \sum_{m=0}^{\infty} \psi_m t^m \left( \frac{t\sqrt{m!n}}{\sqrt{(m-n+1)!(n-1)!}} \left(\frac{r}{t}\right)^{n-1} - \frac{r\sqrt{m!(m-n+1)}}{\sqrt{m!(m-n)!}} \left(\frac{r}{t}\right)^{n} \right) |m-n+1\rangle_1. \tag{A.28} $$

Next, make use of the two identities

$$ \hat{a}^{n-1}|m\rangle = \frac{\sqrt{m!}}{\sqrt{(m-n-1)!}} |m-n+1\rangle, \tag{A.29} $$

$$ \hat{a}^{\dagger}\hat{a}^{n}|m\rangle = \frac{\sqrt{m!(m-n+1)}}{\sqrt{(m-n)!}} |m-n+1\rangle \tag{A.30} $$

to arrive at

$$ {}_2\langle n|\phi\rangle_{12} = \sum_{m=0}^{\infty} \psi_m t^m \left( \frac{t\sqrt{n}}{\sqrt{(n-1)!}} \left(\frac{r}{t}\right)^{n-1} \hat{a}_1^{n-1} - \frac{r}{\sqrt{n!}} \left(\frac{r}{t}\right)^{n} \hat{a}_1^{\dagger}\hat{a}_1^{n} \right) |m\rangle_1. \tag{A.31} $$

By pulling the operators out of the sum and using Eq. 1.132, we can arrive at the Kraus operator form of photon catalysis, where

$$ {}_2\langle n|\hat{B}_{12}|\psi\rangle_1|1\rangle_2 = \hat{K}_1|\psi\rangle_1 \tag{A.32} $$

with

$$\hat{K} = \left(\frac{r}{t}\right)^n \frac{1}{r\sqrt{n!}}(nt^2 - r^2\hat{a}^\dagger\hat{a})\hat{a}^{n-1}e^{-\beta\hat{a}^\dagger\hat{a}} \tag{A.33}$$

and $\beta = -\ln t$.

The form of this Kraus operator more clearly illuminates the photon catalysis process when the other input is a coherent state, i.e., $|\psi\rangle = |\alpha\rangle$. Noting that the damping operator just decreases the value of $\alpha$,

$$e^{-\beta\hat{a}^\dagger\hat{a}}\,|\alpha\rangle \propto |e^{-\beta}\alpha\rangle\,, \tag{A.34}$$

we see that

$$\hat{K}\,|\alpha\rangle \propto (nt^2 - e^{-\beta}r^2\alpha\hat{a}^\dagger)\,|e^{-\beta}\alpha\rangle\,. \tag{A.35}$$

Thus, the single-photon catalysis process on a coherent state creates a superposition between a slight smaller coherent state and the same smaller coherent state with a single photon added to it.

As yet another way to look at it, consider the coherent state as displaced vacuum and commute the coherent state to the back of the equation. Using $\hat{D}^\dagger(\gamma)\hat{a}^\dagger\hat{D}(\gamma) = \hat{a}^\dagger + \gamma^*$, we can rewrite Eq. A.35 as

$$\begin{aligned}\hat{K}\,|\alpha\rangle &\propto \hat{D}(e^{-\beta}\alpha)(nt^2 - e^{-2\beta})r^2|\alpha|^2 - e^{-\beta}r^2\alpha\hat{a}^\dagger)\,|0\rangle \\ &\propto \hat{D}(e^{-\beta}\alpha)\left[(nt^2 - e^{-2\beta})r^2|\alpha|^2)\,|0\rangle - e^{-\beta}r^2\alpha\,|1\rangle\right],\end{aligned} \tag{A.36}$$

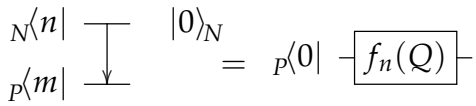
which is now just the displaced superposition $\psi_0\,|0\rangle + \psi_1\,|1\rangle$. Evidently, when $nt^2 = e^{-2\beta})r^2|\alpha|^2$, we are left with a perfectly displaced single photon as derived otherwise in the main text of Ch. 2.

# Appendix B

# PhANTM

## B.1 PhANTM Derivation

We fully derive the effects of photon subtraction proceeded by teleportation in the presence of finite squeezing. First, we derive the circuit identity given by the following lemma:




(B.1)

where the arrow in the diagram represents a beamsplitter, $\hat{B}_\theta$, with reflectivity $r = \sin\theta$ and transmissivity $t = \cos\theta$. The lefthand side can be written as

$$ {}_{N_1}\langle n|_{p_2}\langle m|\hat{B}_\theta|0\rangle_N \otimes I, \tag{B.2} $$

and in the $Q$-basis, the bras can be expressed as

$$ {}_N\langle n| = \int du\, {}_q\langle u|\, \psi_n(u) \tag{B.3} $$

$$ {}_P\langle m| = \frac{1}{\sqrt{2\pi}} \int dv\, {}_q\langle v|\, e^{-imv}, \tag{B.4} $$

where $\psi_n(x)$ is the wavefunction of the harmonic oscillator:

$$ \psi_n(x) = \frac{1}{\pi^{1/4}\sqrt{2^n n!}} e^{-x^2/2} H_n(x). \tag{B.5} $$

Using the above expressions, and temporarily neglecting the $(2\pi)^{-1/2}$ factor, Eq. B.2 becomes

$$\int du dv {}_{q_1}\langle u|{}_{q_2}\langle v|\psi_n(u)e^{-imv}\hat{B}_\theta|0\rangle_{N_1} \tag{B.6}$$

$$= \int du dv {}_{q_1}\langle 0|{}_{q_2}\langle 0|\hat{B}_\theta\hat{B}_\theta^\dagger e^{iu\hat{P}_1+iv\hat{P}_2}\psi_n(u)e^{-imv}\hat{B}_\theta|0\rangle_{N_1} \tag{B.7}$$

$$= \int du dv {}_{q_1}\langle 0|{}_{q_2}\langle 0|e^{iu(t\hat{P}_1-r\hat{P}_2)+iv(t\hat{P}_2+r\hat{P}_1)}\psi_n(u)e^{-imv}|0\rangle_{N_1} \tag{B.8}$$

where in the last line we have used the fact that a beamsplitter applied to a pair of $q = 0$ states has no effect. Expressing the vacuum state in the $Q$-basis as well yields

$$\int du dv dz {}_{q_1}\langle tu + rv|z\rangle_{q_1 q_2}\langle tv - ru|\psi_n(u)\psi_0(z)e^{-imv} \tag{B.9}$$

$$= \int du dv {}_{q_2}\langle tv - ru|\psi_n(u)\psi_0(tu + rv)e^{-imv} \tag{B.10}$$

$$= \frac{1}{\sqrt{2\pi}} \int dx {}_Q\langle x|f_n(x) = {}_p\langle 0|f_n(\hat{Q}) \tag{B.11}$$

where we reintroduced the missing prefactor in the final two lines. With a change of the integration variables to $x = tv - ru, y = tu + rv$, one can see that

$$f_n(x) = \int dy\, \psi_n(ty - rx)\psi_0(y)e^{-im(tx+ry)}. \tag{B.12}$$

Changing variables again, this can be rewritten as

$$\begin{aligned} f_n(x) &= \frac{1}{t}\int dy\, \psi_n(y)\psi_0(\frac{y + rx}{t})e^{-\frac{im}{t}(ry+x)} \\ &= \tfrac{e^{-\imath/2}}{t}e^{-i\zeta x}\int dy\, {}_N\langle n|y\rangle_Q\, {}_Q\langle y|\hat{D}(\alpha)\hat{S}(\imath)|0\rangle_N \\ &= \tfrac{e^{-\imath/2}}{t}e^{-i\zeta x}\,\langle n|\,\hat{D}(\alpha)\hat{S}(\imath)\,|0\rangle \end{aligned} \tag{B.13}$$

where here, $\alpha = -2^{-1/2}(rx + i\frac{r}{t}m)$, $\imath = \ln t$, and $\zeta = \frac{m}{t}(1 + \frac{1}{2}t^2)$.

At this point, it becomes necessary to derive an identity for the overlap coefficients of the Fock states with displaced, squeezed vacuum. This overlap can be written as

$$\langle n|\,\hat{D}(\alpha)\hat{S}(\imath)\,|0\rangle = e^{-\frac{i}{2}\gamma\beta}\,\langle n|\,\hat{Z}(\gamma)\hat{X}(\beta)\hat{S}(\imath)\,|0\rangle, \tag{B.14}$$

where $\gamma = \sqrt{2}\text{Im}[\alpha]$ and $\beta = \sqrt{2}\text{Re}[\alpha]$. Transforming to the $Q$-basis, we have

$$\begin{aligned}
&e^{-\frac{i}{2}\gamma\beta}\int dsdt\psi_n(s)\psi_0(t)_q\langle s|\hat{Z}(\gamma)\hat{X}(\beta)\hat{S}(\tau)|t\rangle_q \\
=&e^{\frac{\tau}{2}-\frac{i}{2}\gamma\beta}\int dsdt\psi_n(s)\psi_0(t)e^{is\gamma}\delta(s-\beta-te^{-\tau}) \\
=&\frac{e^{\frac{\tau}{2}-\frac{i}{2}\gamma\beta}}{\sqrt{2^n n!\pi}}\int dse^{-s^2/2}H_n(s)e^{-(s-\beta)^2e^{2\tau}/2}e^{is\gamma} \\
=&\frac{e^{\frac{\tau}{2}-i\gamma(\beta/2-B)+C}}{A\sqrt{2^n n!\pi}}\int dse^{-s^2/2}H_n(\tfrac{s}{A}+B)e^{i\frac{s}{A}\gamma},
\end{aligned} \tag{B.15}$$

where we have defined

$$A = \sqrt{1+e^{2\tau}} \tag{B.16}$$

$$B = \frac{\beta e^{2\tau}}{1+e^{2\tau}} \tag{B.17}$$

$$C = \frac{\beta^2 e^{4\tau}}{2(1+e^{2\tau})} - \frac{1}{2}\beta^2 e^{2\tau}. \tag{B.18}$$

We can now make use of the pair of Hermite polynomial identities

$$H_n(x+y) = 2^{-\frac{n}{2}}\sum_{k=0}^{n}\binom{n}{k}H_{n-k}\left(x\sqrt{2}\right)H_k\left(y\sqrt{2}\right) \tag{B.19}$$

$$H_n(\gamma x) = \sum_{i=0}^{\lfloor\frac{n}{2}\rfloor}\gamma^{n-2i}(\gamma^2-1)^i\binom{n}{2i}\frac{(2i)!}{i!}H_{n-2i}(x), \tag{B.20}$$

along with the property that Hermite polynomials are eigenfunctions of the Fourier transform, i.e.,

$$\textstyle\int_{-\infty}^{\infty}dxe^{ikx}e^{-x^2/2}H_n(x) = \sqrt{2\pi}(i)^n e^{-\frac{1}{2}k^2}H_n(k), \tag{B.21}$$

to take Eq. B.15 and arrive at the result of

$$\begin{aligned}
&\langle n|\hat{D}(\alpha)\hat{S}(\tau)\,|0\rangle = \\
&\kappa\sum_{k=0}^{n}\sum_{j=0}^{\lfloor\frac{n-k}{2}\rfloor}\chi_{j,k}H_k(\sqrt{2}B)H_{n-k-2j}(\tfrac{\gamma}{A})
\end{aligned} \tag{B.22}$$

where

$$\kappa = \frac{i^n e^{\frac{\zeta}{2} - i\gamma(\beta/2 - B) + C}}{A^{n+1}\sqrt{2^{n-1}n!}} e^{-\frac{\gamma^2}{2A^2}} \tag{B.23}$$

$$\chi_{j,k} = \binom{n}{k}\binom{n-k}{2j}\frac{(2j!)A^{2j+k}}{j!\sqrt{-2^{2j+k}}}(\tfrac{2}{A^2} - 1)^j. \tag{B.24}$$

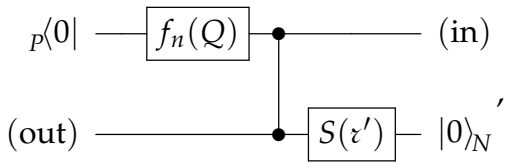
We can now put everything back together and see that $f_n(x)$ is a polynomial of degree $n$ with an exponential envelope. Using Eq. B.13 along with the derived overlap in Eq. B.22 and the intermediary definition of Eqs. eqs. (B.16) to (B.18), (B.23) and (B.24) lead to

$$f_n(x) = e^{-imx(\frac{2t}{1+t^2})} e^{-x^2\frac{t^2r^2}{2+2t^2}} f'_n(x), \tag{B.25}$$

where $f'_n(x)$ is the polynomial portion given by

$$f'_n(x) = \frac{i^n}{\sqrt{2^{n-1}n!(1+t^2)^{n+1}}} \sum_{k=0}^{n} \sum_{j=0}^{\lfloor \frac{n-k}{2} \rfloor} \binom{n}{k}\binom{n-k}{2j} \frac{(2j)!r^{2j}}{2^j j!} \left(\frac{1+t^2}{2}\right)^{k/2} (-i)^{k+2j} H_k(\tfrac{-x\sqrt{2}rt^2}{1+t^2}) H_{n-k-2j}(\tfrac{-mr}{t\sqrt{1+t^2}}) \tag{B.26}$$

Subtraction-assisted teleportation can thus be recognized as a circuit of the form

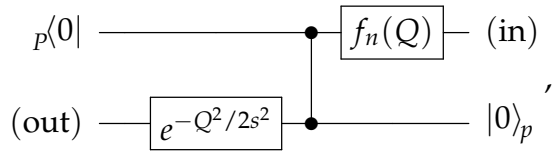



(B.27)

where $f$ is the resulting function from Eq. B.25. The operator in the top wire is a function of $\hat{Q}$ only so commutes with the $\hat{C}_Z$, and we can make use of the fact that we can write the input momentum-squeezed state as

$$\begin{aligned}
\hat{S}(\iota')|0\rangle_N &= \pi^{-1/4}\int dt e^{-t^2/2s^2}|t\rangle_q \\
&= \pi^{-1/4} e^{-\hat{Q}^2/2s^2}\int dt|t\rangle_q \\
&= \pi^{1/4}\sqrt{\frac{2}{s}} e^{-\hat{Q}^2/2s^2}|0\rangle_p,
\end{aligned} \tag{B.28}$$

where $s = e^{r'}$. In this form, the exponentiated function of $\hat{Q}$ commutes through the $C_Z$ gate to give us




(B.29)

which is easily recognizable as the teleportation circuit with additional operators applied bother before and after the teleportation. The overall Kraus operator representation of the action of the circuit acting on the input is given by

$$\hat{K} = \pi^{1/4}\sqrt{\frac{2}{s}}e^{-\hat{Q}^2/2s^2}\hat{R}(\tfrac{\pi}{2})f_n(\hat{Q}). \tag{B.30}$$

However, if we note that there is a $Q$-quadrature shift within $f_n(\hat{Q})$ that we may wish to undo with feed-forward operations, we can commute this to the back of the Kraus operator. This leads to an alternate form of

$$\begin{aligned}\hat{K} =& \tfrac{\pi^{1/4}\sqrt{2}}{\sqrt{s}}e^{\frac{m^2}{s^2}\sigma^2}\hat{X}(m\sigma)\hat{R}(\tfrac{\pi}{2}) \\ &\times e^{-\frac{1}{2s^2}(\hat{P}-m\sigma)^2}e^{-\hat{Q}^2\frac{\sigma t(1-t^2)}{4}}f_n'(\hat{Q}),\end{aligned} \tag{B.31}$$

where we define $\sigma = \frac{2t}{1+t^2}$. For weak beamsplitter reflectivity, we have that $t \to 1$ and $\sigma \to 1$, and in this limit, we can recover the form of the idealized case discussed in the main text. Taking the large squeezing limit $s \to \infty$ and weak beamsplitter reflectivity $r \to 0$, we have to leading order in $r$ that

$$\begin{aligned}f_n'(\hat{Q}) \to& \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor}\frac{\sqrt{n!}2^{-j}r^{2j}}{2^n j!(n-2j)!}\times \\ &\sum_{k=0}^{n-2j}\binom{n-2j}{k}H_k(\tfrac{-Qr}{\sqrt{2}})i^{n-k-2j}H_{n-k-2j}(\tfrac{-mr}{\sqrt{2}}).\end{aligned} \tag{B.32}$$

Defining the generalized probabilist's Hermite polynomial $H_{e_n}^{[\alpha]}(x)$ as

$$H_{e_n}^{[\alpha]}(x) = \alpha^{\frac{n}{2}}H_{e_n}\left(\frac{x}{\sqrt{\alpha}}\right) = \left(\tfrac{\alpha}{2}\right)^{\frac{n}{2}}H_n\left(\frac{x}{\sqrt{2\alpha}}\right) \tag{B.33}$$

and making using of the identity

$$H_{e_n}^{[\alpha+\beta]}(x+y) = \sum_{k=0}^{n} \binom{n}{k} H_{e_k}^{[\alpha]}(x) H_{e_{n-k}}^{[\beta]}(y), \tag{B.34}$$

along with the special case of

$$H_{e_n}^{[0]}(x) = (x)^n \tag{B.35}$$

allows us to simplify Eq. B.32 to

$$f_n'(\hat{Q}) \to \frac{r^n\sqrt{n!}}{2^n\sqrt{2^n}} \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \frac{\left(\sqrt{2}(Q+im)\right)^{n-2j}}{j!(n-2j)!}. \tag{B.36}$$

We can now simplify the above expression by using the expansion of the Hermite polynomial given by

$$H_n(x) = n! \sum_{m=0}^{\lfloor \frac{n}{2} \rfloor} \frac{(-1)^m}{m!(n-2m)!}(2x)^{n-2m}, \tag{B.37}$$

which leads to the limiting case Kraus operator of

$$\hat{K} \propto \hat{X}(m)\hat{R}(\tfrac{\pi}{2}) H_n\left(\frac{i\hat{Q}-m}{\sqrt{2}}\right). \tag{B.38}$$

## B.2 Hermite Polynomial

We prove the relation

$$H_{e_n}(Q) = \sum_{k=0}^{n} \binom{n}{k} (-iP)^k (Q+iP)^{n-k}. \tag{B.39}$$

This is easily verified to be true for the first few terms, so we proceed by induction. Assuming Eq. B.39 is true for $n$, we can make use of the relation $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ to

prove the validity for $n+1$ as follows:

$$\sum_{k=0}^{n+1} \binom{n+1}{k} (-iP)^k (Q+iP)^{n+1-k} = \tag{B.40}$$

$$\sum_{k=0}^{n} \binom{n}{k} (-iP)^k (Q+iP)^{n-k}(Q+iP) + \sum_{k=1}^{n} \binom{n}{k-1} (-iP)^k (Q+iP)^{n+1-k} + (-iP)^{n+1} \tag{B.41}$$

$$= H_{e_n}(Q)(Q+iP) + \sum_{k=0}^{n-1} \binom{n}{k} (-iP)^{k+1} (Q+iP)^{n-k} + (-iP)^{n+1} \tag{B.42}$$

$$= H_{e_n}(q)(Q+iP) + (-iP) \left( \sum_{k=0}^{n} \binom{n}{k} (-iP)^k (Q+iP)^{n-k} \right) \tag{B.43}$$

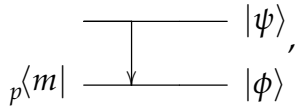
$$= H_{e_n}(Q)(Q+iP) + (-iP)\, H_{e_n}(Q) = QH_{e_n}(Q) + \underbrace{[-iP, H_{e_n}(Q)]}_{-\frac{dH_{e_n}(Q)}{dQ}} \tag{B.44}$$

$$= QH_{e_n}(Q) - \frac{dH_{e_n}(Q)}{dQ} \equiv H_{e_{n+1}}(Q) \tag{B.45}$$

QED.

## B.3 Beamsplitter to $\hat{C}_Z$ Breeding

Several other works have discussed breeding cat states for enlargement or to make grid states in the context of beamsplitters followed by a homodyne measurement. Here, we show how this is translated to the canonical cluster state measurements through the use of the beamsplitter unitary decomposition. Breeding cat states is achieved by using the circuit




(B.46)

where $|\psi\rangle$ and $|\phi\rangle$ and the input cat states used for breeding. Up to an overall phase, the balanced beamsplitter can be decomposed as [18]

$$\hat{B}_{12} = e^{-i\hat{Q}_1\hat{P}_2} \left( \hat{S}_1^\dagger(\tfrac{1}{2}\ln 2)\hat{S}_2(\tfrac{1}{2}\ln 2) \right) e^{i\hat{P}_1\hat{Q}_2}. \tag{B.47}$$

Performing a homodyne measurement in the $P$-basis to mode two after applying a beamsplitter can thus be written as

$$\begin{aligned} {}_{2P}\langle m|\hat{B}_{12} &= {}_{2P}\langle m|\hat{Z}_1^\dagger(m)\left(\hat{S}_1^\dagger(\ln\sqrt{2})\hat{S}_2(\tfrac{1}{2}\ln 2)\right)e^{i\hat{P}_1\hat{Q}_2} \\ &= \hat{Z}_1^\dagger(m)\hat{S}_1^\dagger(\ln\sqrt{2})\,{}_{2P}\langle m'|e^{i\hat{P}_1\hat{Q}_2}, \end{aligned} \tag{B.48}$$

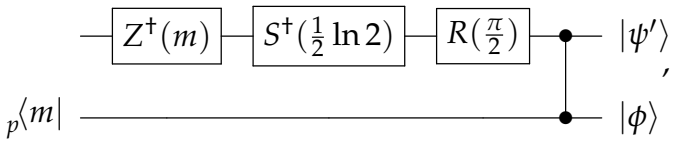
where we have used that

$$\begin{aligned} \hat{S}^\dagger(r)|m\rangle_P &= \hat{Z}(m')\hat{S}^\dagger(r)|0\rangle_p \\ &= |m'\rangle_p \end{aligned} \tag{B.49}$$

with $m' = 2^{-1/2}m$. Finally, we can arrive at the $\hat{C}_Z$ gate we want by rewriting

$$e^{i\hat{P}_1\hat{Q}_2} = \hat{R}_1(\tfrac{\pi}{2})\hat{C}_{Z_{12}}\hat{R}_1^\dagger(\tfrac{\pi}{2}). \tag{B.50}$$

With the help of Eqs. B.48 and B.50, the starting circuit in Eq. B.46 can be converted to




(B.51)

where $\psi$ has been Fourier transformed to $|\psi'\rangle = \hat{R}(\frac{\pi}{2})|\psi\rangle$. Thus, if the input cat states are properly rotated, breeding with the beamsplitter is exactly equivalent to breeding with a $\hat{C}_Z$ gate up to Gaussian operations that can be undone with further cluster state processing and feed-forward displacements.

# Appendix C

# Quantum Sensing

## C.1 Experimental State Derivation

Input state:

$$|\phi_{in}\rangle = |0,\varepsilon\rangle_{ab}\,|0\rangle_c = \frac{1}{\cosh\varepsilon}\sum_n^{\infty}(\tanh\varepsilon)^n\,|n\rangle_a\,|n\rangle_b\,|0\rangle_c \tag{C.1}$$

The overall state is transformed by the beamsplitters and projective PNR measurement to become

$$|\phi_{out}\rangle = {}_c\langle Z|\,U_{ac}U_{bc}\,|\phi_{in}\rangle\,, \tag{C.2}$$

where $U_{bc} = exp\big[\theta_1(bc^\dagger - b^\dagger c)\big]$, and $U_{ac} = exp\big[\theta_2(ac^\dagger e^{-i\varphi} - a^\dagger c e^{i\varphi})\big]$ to allow for an additional phase of $\varphi$ between mode $a$ and the reflected mode $c$ from the first beamsplitter, where reflectivities and transmissivities $r_i = \sin\theta_i$, $t_i = \cos\theta_i$. Note that the second beamsplitter operation, $U_{ac}$, will act between mode $a$ and the *transformed* mode $c$ from the output of the first beamsplitter operation. With this in mind, the first beamsplitter transforms the input state to

$$\begin{aligned}
U_{bc}\,|\phi_{in}\rangle &= \sum_n^{\infty}\frac{(\tanh\varepsilon)^n}{\sqrt{n!}\cosh\varepsilon}U_{bc}{b^\dagger}^n U_{bc}^\dagger\,|n\rangle_a\,|0\rangle_b\,|0\rangle_c \\
&= \sum_n^{\infty}\frac{(\tanh\varepsilon)^n}{\sqrt{n!}\cosh\varepsilon}(t_1 b^\dagger + r_1 c^\dagger)^n\,|n\rangle_a\,|0\rangle_b\,|0\rangle_c \\
&= \sum_n^{\infty}\frac{(\tanh\varepsilon)^n}{n!\cosh\varepsilon}\sum_k^{n}\binom{n}{k}t_1^k r_1^{n-k}{a^\dagger}^n{b^\dagger}^k{c^\dagger}^{n-k}\,|0\rangle_a\,|0\rangle_b\,|0\rangle_c\,.
\end{aligned} \tag{C.3}$$

Now, applying the second beamsplitter yields the state

$$\begin{aligned}
&\sum_n^\infty \frac{(\tanh\mathfrak{r})^n}{n!\cosh\mathfrak{r}} \sum_k^n \binom{n}{k} t_1^k r_1^{n-k} b^{\dagger k} (t_2 a^\dagger + e^{i\varphi} r_2 c^\dagger)^n (t_2 c^\dagger - e^{-i\varphi} r_2 a^\dagger)^{n-k} |0\rangle_{abc} \\
=&\sum_n^\infty \frac{(-r_1 r_2^2 \tanh\mathfrak{r})^n}{n!\cosh\mathfrak{r}} \sum_k^n \binom{n}{k} \left(\frac{t_1}{r_1 r_2}\right)^k b^{\dagger k} \sum_j^n \binom{n}{j} \left(\frac{t_2}{r_2}\right)^j \\
&\times \sum_m^{n-k} \binom{n-k}{m} (-1)^{m+k} \left(\frac{t_2}{r_2}\right)^m (e^{i\varphi})^{k+m-j} a^{\dagger j+n-k-m} c^{\dagger n-j+m} |0\rangle_{abc} . \qquad \text{(C.4)}
\end{aligned}$$

Projecting output mode $c$ onto the PNR detection event of $Z$ photons means that the only terms in the above sums that survive occur when $j = n + m - Z$. It is also important to note that the maximum value of $j$ is $n$, so the remaining sum over $m$ goes from zero to $m_{max} = \text{Min}(n-k, Z)$. These substitutions lead to

$$\begin{aligned}
|\phi_Z\rangle =& \frac{r_2^Z \sqrt{Z!}}{t_2^Z \cosh\mathfrak{r}} \sum_{n=0}^\infty (-t_2 r_1 r_2 e^{-i\varphi} \tanh\mathfrak{r})^n \sum_{k=0}^n \frac{\sqrt{(2n-k-Z)!}}{\sqrt{k!}} \left(\frac{-e^{i\varphi} t_1}{r_1 r_2}\right)^k \\
&\times \sum_{m=0}^{m_{max}} \binom{n}{n+m-Z} \left(\frac{t_2}{r_2}\right)^{2m} \frac{(-1)^m}{m!(n-k-m)!} |2n-k-Z\rangle_a |k\rangle_b \\
&\propto \sum_{n=0}^\infty c_n \sum_{k=k_{min}}^n d_{n,k} |n+k-Z\rangle_a |n-k\rangle_b , \qquad \text{(C.5)}
\end{aligned}$$

where $k_{min} = \text{Max}(0, Z-n)$, $c_n = (t_1 t_2 \tanh\mathfrak{r})^n$, and the coefficient $d_{n,k}$ is

$$d_{n,k} = \frac{\sqrt{(n+k-Z)!}}{\sqrt{(n-k)!}} \left(\frac{r_1 r_2}{-e^{i\varphi} t_1}\right)^k \sum_{m=0}^{\text{Min}(Z,k)} \binom{n}{n+m-Z} \left(\frac{t_2}{r_2}\right)^{2m} \frac{(-1)^m}{m!(k-m)!}. \qquad \text{(C.6)}$$

The form of the general case in Eq. C.5 shows that after all is said and done, we have a superposition over $n$ of two-mode superpositions that are desirable for quantum enhanced interferometry, where the $c_n$ terms depends on a new effective squeezing, which is reduced from the originial value by the transmissivity of the two beamsplitters. The success probability to create this state after a given $Z$ PNR detection is given by

$$P(Z) = \frac{Z!}{(\cosh\mathfrak{r})^2} \left(\frac{r_2}{t_2}\right)^{2Z} \sum_{n=0}^\infty \sum_{k=k_{min}}^n c_n^2 d_{n,k}^2 \qquad \text{(C.7)}$$

## C.2 Highly unbalanced beamsplitters

If we consider the case where both beamsplitters are identical and highly transmissive with $r_1 = r_2 \equiv r$ and $r \ll 1$, then examining Eq. C.6 shows that only the term with $m = k$ contributes to the coefficient $d_n$ to leading order. Furthermore, since the sum over $m$ is truncated at $m_{max} = \mathrm{Min}(k, Z)$, the sum over $k$ in Eq. C.5 can be effectively truncated at $Z$ to the same order of approximation. The output state then becomes

$$|\phi_Z\rangle \propto \sum_{n=0}^{\infty} (t^2 \tanh \mathfrak{r})^n \sum_{k_{min}}^{Z} \sqrt{\binom{n}{k}\binom{n}{Z-k}\binom{Z}{k}} e^{ik\varphi} |n+k-Z\rangle_a |n-k\rangle_b , \tag{C.8}$$

with an approximate success probability of

$$P(Z) \approx \frac{r^{2Z}}{(\cosh \mathfrak{r})^2} \sum_{n=0}^{\infty} (1-2nr^2)(\tanh \mathfrak{r})^{2n} \sum_{k_{min}}^{Z} \binom{n}{k}\binom{n}{Z-k}\binom{Z}{k} \tag{C.9}$$

## C.3 Coefficients

Here we verify several properties of the coefficients of the experimental state. In the case of a highly unbalanced beamsplitter, then for a given $n$, the experimental state takes the form of Eq. 5.19 with coefficients having the form shown in Eq. C.8 to be

$$c_k \propto \sqrt{\binom{n}{k}\binom{n}{Z-k}\binom{Z}{k}} e^{ik\varphi}, \tag{C.10}$$

where all $c_k$ have the same proportionality constant from normalization. Setting $\varphi = 0$ ensures that all $c_k$ are real. When writing the state in the Schwinger representation, these coefficients become

$$c_m \propto \sqrt{\binom{n}{m+s}\binom{n}{s-m}\binom{2s}{m+s}}, \tag{C.11}$$

where $s = \frac{Z}{2}$ and $m = k - s$, from which it is clear to see that

$$c_m = c_{-m}. \tag{C.12}$$

Now, how do neighboring coefficients relate? Calculating the ratio between neighbors yields

$$\frac{c_{m+1}}{c_m} = \frac{s-m}{s+m+1}\left(\frac{n-s-m}{n-s+m+1}\right)^{1/2}, \tag{C.13}$$

which leads to the bounds of

$$\frac{1}{2s} \overset{(i)}{\leq} \frac{c_{m+1}}{c_m} \overset{(i)}{\leq} 2s, \tag{C.14}$$

where $(i)$ denotes the use of the approximation that $n \gg s$. This ratio is useful when calculating terms that appear in expectation values, and can be used to determine a bound on the sum of all pairs of neighboring coefficients to be

$$\frac{1}{2s} \overset{(i)}{\leq} \sum_{m=-s}^{s-1} c_m c_{m+1} \leq 1, \tag{C.15}$$

where the upper bound can be derived from the Cauchy–Schwarz inequality. Similarly, the ratio between next neighboring coefficients has the bound

$$\frac{c_{m+2}}{c_m} \overset{(i)}{\geq} \frac{1}{s(2s-1)}, \tag{C.16}$$

which leads to bounding the sum of next-nearest neighboring coefficients to be

$$\frac{1}{2s^2-s} \overset{(i)}{\leq} \sum_{m=-s}^{s-2} c_m c_{m+2} \leq 1. \tag{C.17}$$

## C.4 General Multi-Photon Subtracted State

Here we derive the Fisher information for the general case of the state given by Eq. 5.19. All of the relevant terms are:

$$\begin{aligned}
\langle a^\dagger b\rangle &= \sum_{k=0}^{Z-1} c_k c_{k+1}^* \sqrt{(n-k)(n-Z+k+1)} \\
\langle a b^\dagger\rangle &= \sum_{k=0}^{Z-1} c_k^* c_{k+1} \sqrt{(n-k+1)(n-Z+k)} \\
\langle a^\dagger a b b^\dagger\rangle &= \sum_{k=0}^{Z} |c_k|^2 (n-Z+k)(n-k+1)) \\
\langle a a^\dagger b^\dagger b\rangle &= \sum_{k=0}^{Z} |c_k|^2 (n-Z+k+1)(n-k)) \\
\langle a^\dagger a^\dagger b b\rangle &= \sum_{k=0}^{Z-2} c_k c_{k+2}^* \sqrt{(n-k)(n-k-1)(n-Z+k+1)(n-Z+k+2)} \\
\langle a a b^\dagger b^\dagger\rangle &= \sum_{k=0}^{Z-2} c_k^* c_{k+2} \sqrt{(n-k+1)(n-k+2)(n-Z+k)(n-Z+k-1)}.
\end{aligned}$$

If we take a case where the number of photons removed from the state is small compared to the total, i.e., $n \gg Z$, then denoting the use of this approximation as $(i)$, the Fisher information is

$$\begin{aligned}
\mathcal{F} \overset{(i)}{=} 2\sum_{k=0}^{Z} n^2 |c_k|^2 - \sum_{k=0}^{Z-2} n^2 (c_k c_{k+2}^* + c_k^* c_{k=2}) + \left(\sum_{k=0}^{Z-1} n c_k c_{k+1}^*\right)^2 \\
+ \left(\sum_{k=0}^{Z-1} n c_k^* c_{k+1}\right)^2 - \sum_{k=0}^{Z-1}\sum_{k'=0}^{Z-1} n^2 (c_k c_{k+1}^* c_{k'}^* c_{k'+1} + c_k^* c_{k+1} c_{k'} c_{k'+1}^*).
\end{aligned} \tag{C.18}$$

If we now make assumption $(ii)$ that all of the coefficients are the same, $c_k = \frac{1}{\sqrt{Z}}$, then

$$\mathcal{F} \overset{(i,ii)}{=} \frac{4n^2}{Z}, \tag{C.19}$$

so the Cramér-Rao inequality leads to

$$(\Delta\phi_d)^2 \geq \frac{Z}{N^2}, \tag{C.20}$$

where $N = 2n - Z$.

## C.5 MZI fringe and phase sensitivity

Using the Schwinger representation to determine the observables when the general Z-subtracted state from Eq. 5.26 is input into a MZI, we start by finding the expectation value of $J_x$ and $J_z$. We find that

$$\langle J_z \rangle = 0 \tag{C.21}$$

$$\begin{aligned}\langle J_x \rangle &= \frac{1}{2}\sum_{m=1-s}^{s} c_{m-1}^* c_m \sqrt{(j+m)(j-m+1)} + \frac{1}{2}\sum_{m=-s}^{s-1} c_{m+1}^* c_m \sqrt{(j-m)(j+m+1)} \\ &= \sum_{m=-s}^{s-1} c_{m+1} c_m \sqrt{(j-m)(j+m+1)} \\ &\overset{(i)}{=} j \sum_{m=-s}^{s-1} c_{m+1} c_m,\end{aligned}$$

where we have used that all $c_m \in \mathbb{R}$, $c_m = c_{-m}$, and have used approximation $(i)$ that $j \gg s$. From here, we can estimate the value of the remaining sum by making use of the bounds on $\frac{c_{m+1}}{c_m}$, denoted by $(ii)$, for all of the coefficients derived in Appx C.3. This leads to the result that

$$\frac{j}{2s} \overset{(i,ii)}{\leq} \langle J_x \rangle \overset{(i)}{\leq} j \tag{C.22}$$

The observable fringe is given by the expectation value of $2J_z$ at the output, where

$$\begin{aligned}|\langle J_z' \rangle| &= |\cos\phi \langle J_z \rangle - \sin\phi \langle J_x \rangle| \\ \frac{j}{2s}|\sin\phi| &\overset{(i,ii)}{\leq} |\langle J_z' \rangle| \overset{(i)}{\leq} j|\sin\phi|.\end{aligned} \tag{C.23}$$

The end result above shows that the measurable fringe scales with the mean photon number of the state. When calculating $\Delta\phi$, the Heisenberg transformations yield

$$\frac{\partial d}{\partial d\phi}\langle J_z' \rangle = -\sin\phi \langle Jz \rangle - \cos\phi \langle Jx \rangle \tag{C.24}$$

$$(\Delta J_z')^2 = (\cos\phi \Delta J_z)^2 + (\sin\phi \Delta J_x)^2 - \sin\phi\cos\phi(\langle \{J_z, J_x\} \rangle - 2\langle Jz \rangle \langle Jx \rangle). \tag{C.25}$$

Deriving the quantities individually, we have

$$(\Delta J_z)^2 = \langle J_z^2 \rangle = \sum_{m=-s}^{s} m^2 |c_m|^2 < s^2, \tag{C.26}$$

where the inequality is obtained by replacing all $m^2$ with the maximum value of $s^2$. We also find that

$$\begin{aligned}
\langle J_x^2 \rangle &= \frac{1}{4}\langle J_+^2 + J_-^2 + 2(J^2 - J_z^2)\rangle \\
&= \frac{1}{2}\sum_{m=-s}^{s} |c_m|^2 (j^2 + j - m^2) \\
&\quad + \frac{1}{2}\sum_{m=-s}^{s-2} c_m c_{m+2} \sqrt{(j+m+2)(j-m-1)(j+m+1)(j-m)} \\
&\overset{(i)}{=} \frac{j(j+1)}{2} - \frac{s^2}{2} + \frac{1}{2}(j^2 - s^2) \sum_{m=-s}^{s-2} c_m c_{m+2} \\
&\frac{j^2}{2}\left(1 + \frac{1}{s(2s+1)}\right) + \frac{j}{2} \overset{(i,ii)}{\leq} \langle J_x^2 \rangle \overset{(i)}{\leq} 1,
\end{aligned} \tag{C.27}$$

where the sum in the second-to-last line was bounded by the ratio of next-nearest neighboring coefficients derived in Appx C.3. This leads to the variance

$$(\Delta J_x)^2 \sim \mathcal{O}(j^2). \tag{C.28}$$

The other necessary terms are

$$\langle \{J_z, J_x\} \rangle = \sum_{m=-s}^{s-1} (2m+1) c_{m+1} c_m \sqrt{(j-m)(j+m+1)} \approx \langle J_x \rangle \tag{C.29}$$

$$2\langle Jz \rangle \langle Jx \rangle = 0, \tag{C.30}$$

which when combined with Eqs. C.23, C.26, and C.28 lead to an overall value for the phase difference estimator given by Eq. 5.29 to be

$$(\Delta\phi)^2 \overset{(i)}{\leq} \frac{(s\cos\phi)^2 + (\Delta J_x \sin\phi)^2 + \sin\phi\cos\phi\langle J_x \rangle}{(\frac{j}{2s}\cos\phi)^2} \tag{C.31}$$

This takes on a minimum value when $\phi = 0$ to yield the upper bound

$$\Delta\phi_{min} \overset{(i,ii)}{\leq} \frac{2s^2}{j}, \tag{C.32}$$

which scales with the HL up to a constant factor. By taking the upper limit of $\langle J_x \rangle$ from Eq. C.22, the lower bound on $\Delta\phi$ when $\phi = 0$ is

$$\frac{\sqrt{s}}{j} \overset{(i,ii)}{\leq} \Delta\phi_{min}. \tag{C.33}$$

## C.6 Phase sensitivity for a general twin-beam input

In this section, we show that our photon subtraction protocol also works for the most general twin-beam statistical mixture, e.g. as produced by an OPO above threshold. The density operator in the Fock basis is given by

$$\rho_{ab} = \sum_{n,n'} \rho_{n,n'} \left| nn \right\rangle \left\langle n'n' \right| . \tag{C.34}$$

The two beamplitter operations are given by

$$\begin{aligned} U_{ac'}U_{bc} &= exp\left[\theta_2(a^\dagger c' - ac'^\dagger)\right] exp\left[\theta_1(b^\dagger c - bc^\dagger)\right] \\ &= \sum_{j,k} \frac{\theta_1^k \theta_2^j}{j!k!} \left(a^\dagger c' - ac'^\dagger\right)^j \left(b^\dagger - bc^\dagger\right)^k, \end{aligned} \tag{C.35}$$

where $c' = c\cos\theta_1 - b\sin\theta_1$ is the transformed vacuum mode from the first beamsplitter input. Because the input state, $\rho_{ab}$, consists solely of vacuum in the input mode $c$, and we are post-selecting the transformed mode $c$ on a detection of $Z$ photons, we need only consider terms with of the form $c^x c^{\dagger(x+Z)}$ and $c^{\dagger(x+Z)}c^x$. To further simplify, we can assume the highly unbalanaced beamsplitter regime, where both $\theta_1 \ll 1$ and $\theta_2 \ll 1$, in which case we need only consider terms with $c^{\dagger Z}$. Thus, to leading order in $\theta$, we have $j + k = Z$ and

$$U_{ac'}U_{bc} \approx \sum_{k=0}^{Z} \frac{r_1^k r_2^{Z-k}}{k!(Z-k!)} \left(-at_1\right)^{Z-k} \left(-b\right)^k c^{\dagger Z}, \tag{C.36}$$

where $r_1 = \sin\theta_1 \approx \theta_1$ and $r_2 = \sin\theta_2 \approx \theta_2$. Sending the twin-beam input through the unbalanced beamsplitters and detecting $Z$ photons in mode $c$ leads to

$$\rho_{out} = \mathrm{Tr}_c\left[(|Z\rangle\langle Z|)_c U_{ac'} U_{bc} \rho_{ab} \otimes |0\rangle\langle 0| U_{bc}^\dagger U_{ac'}^\dagger\right] \tag{C.37}$$

$$= \mathrm{Tr}_c\left[(|Z\rangle\langle Z|)_c \sum_{k,k'}^{Z} \frac{r_1^{k+k'} r_2^{2Z-k-k'} t_1^{2Z-k-k'}}{k!k'!(Z-k)!(Z-k')!} \sum_{n,n'}^{\infty} \rho_{n,n'} a^{Z-k} b^k c^{\dagger Z} |n,n,0\rangle\langle n',n',0| a^{\dagger(Z-k')} b^{\dagger k'} c^Z\right] \tag{C.38}$$

$$\propto (r_2 t_1)^{2Z} \sum_{n,n'}^{\infty} \rho_{n,n'} \sum_{k,k'}^{Z} C_{k,k'} |n-Z+k, n-k\rangle\langle n'-Z+k', n'-k'|, \tag{C.39}$$

where $C_{n,n'}$ contains all of the remaining binomial coefficients,

$$C_{n,n'} = \left[\binom{n}{k}\binom{Z}{k}\binom{n}{Z-k}\binom{n'}{k'}\binom{Z}{k'}\binom{n'}{Z-k'}\right]^{1/2} \left(\frac{r_1}{t_1 r_2}\right)^{k+k'}. \tag{C.40}$$

Writing the output in the Schwinger representation, we have

$$\rho_{out} = \sum_{j,j'}^{\infty} \rho_{j,j'} \sum_{m,m'=-s}^{s} C'_{m,m'} |j,m\rangle\langle j',m'|, \tag{C.41}$$

where $s = \frac{Z}{2}$. From here, the calculations for the Schwinger operators follow the form of C.5 for each of the superpositions within the mixture, leading to the finding that about the interferometric phase $\phi = 0$, we have that $(\Delta J_z')^2 \leq s^2$. Additionally, following the arguments of C.3, we find that $\langle J_x\rangle \geq \frac{j_{avg}}{2s}$, where $j_{avg}$ is the average value of $j$ in the statistical mixture. These values lead to the determination that about $\phi = 0$,

$$\Delta\phi = \frac{(\Delta J_z')}{\left|\frac{\partial}{\partial\phi}\langle J_z'\rangle\right|}\Big|_{\phi=0} \overset{(i,ii)}{\leq} \frac{2s^2}{j_{avg}}, \tag{C.42}$$

and hence the general twin-beam source is sufficient to achieve phase-sensitivity scaling proportionally with the HL.

# Appendix D

# Pulse Timing

| *Channel* | DELAY | WIDTH | Mode | Gate |
|---|---|---|---|---|
| Trigger | 79.90 $\mu$s | 100 ns | Active High | Disabled |
| Pump AOM | 79.97 $\mu$s | 100 ns | Active High | Disabled |
| IR AOM | 0 | 80.10 $\mu$s | Active Low | Disabled |
| Lock shutter | 40 ms | 182 ms | Active High | Active Low |
| TES shutter | 373 ms | 227 ms | Active Low | Disabled |
| PDH lock freeze | 420 ms | 150 ms | Active Low | Active Low |
| Gate | 400 ms | 200 ms | Active High | N/A |

TABLE D.1: Global settings: pulse period of 80.2 $\mu$s and a duty cycle of 5,000 pulses on followed by 2,500 pulses off. All channels initialize at the same time and are INACTIVE for duration DELAY. The signal is then ACTIVE for duration WIDTH. The mode determines whether the signal is high (5V) or low (0V) when ACTIVE. Channels can additionally be gated, such that they follow the prescribed settings only when the gate is enabled or disabled.

This section discusses the pulse timing sequence used to drive the AOMs, trigger data collection, perform on/off cavity locking, and operate the shutters. The pulsing was achieved using the Quantum Composers 9500 series pulse generator. The pulse generator has a fixed internal clock as a reference, and 8 signals can be independently set with respect to the same reference. Each channel can emit a square pulse from 0 to 5 volts with tunable delay and width, while all channels adhere to a global pulse period and duty cycle if specified. The timings for each channel are displayed in Table D.1 and the outputs are shown in Fig. D.1. The global pulse settings give an $\mu$s period to allow time for $100 - 200$ ns width pulses followed by an 80 $\mu$s gap during which the TES can completely re-cool before the next event. Additionally, the pulses were sent in trains of

5,000 pulses followed by a duration of 2,500 periods without pulses during which the cavity can be locked.

The specifics of the individual timing for each channel are relatively arbitrary, but the results can be seen more clearly in the figure. First, Fig. D.1(a) shows the individual pulses for the pump (green trace) and IR (magenta) AOMs. Each pulse is 100 ns in width, but the IR pulse must be delayed so that the pump pulse occurs first. The delay was chosen by tuning the pulse separation and performing TES measurements on photon arrival times as discussed in Sec. 8.2.2.

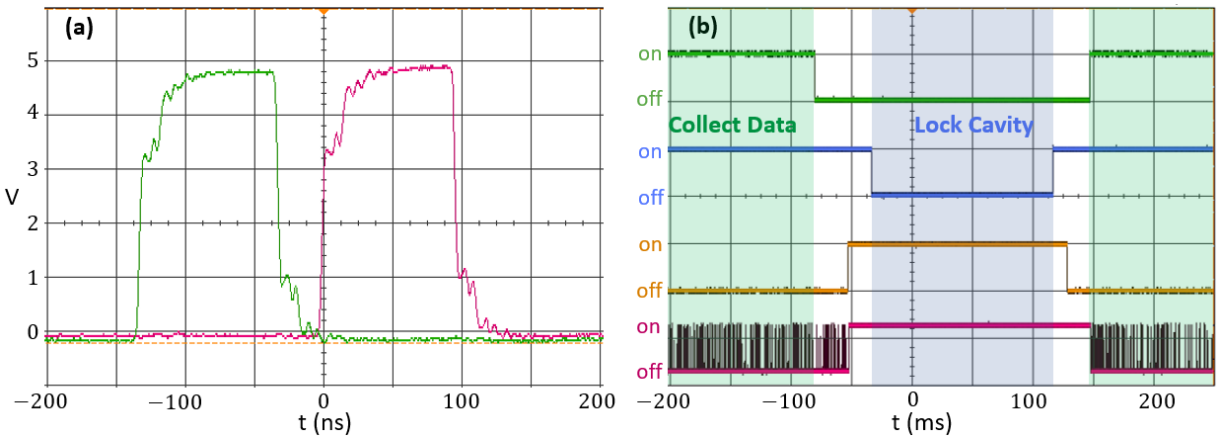



FIGURE D.1: Green trace: TES shutter. Blue trace: signal to freeze PDH locking servo. Yellow trace: Locking beam shutter. Magenta trace: IR AOM pulsing. The green shaded region is when data collection occurs while the blue shaded region is when the cavities are locked. The region between is used as a buffer to ensure signals are turned on in the proper order.

Zooming out, the pulse schedule for the entire data collection schedule can be seen in Fig. D.1(b). Here, the green trace is the signal enabling the shutters in front of the TES channels, the yellow trace shows the signal to the shutter in front of the locking beam, the blue trace is a signal to freeze the PDH cavity locking, and the magenta is the IR AOM pulse schedule. During the green shaded area, the experiment is in the data collection phase. Here, the TES shutters are open, the locking shutter is closed, the PDH locking is frozen to prevent it responding to a lack of signal, and both AOMs are pulsing. During the blue shaded region, the experiment is in the cavity locking phase. Here, the shutter positions are reversed, the signal to freeze the cavity lock is disabled, and the IR beam is turn to CW so that there is sufficient signal to lock the cavities.

These two regions are relatively clear, but the region between, the unshaded region, is actually the most important. This region acts as a 'buffer' between the data collection phase. Here, it is important to ensure that the signals turn on and off in the precise order, taking into account physical delays from the mechanical shutters. When transitioning from data collection to locking, the TES shutter must be the first to close to prevent the detector from saturating when the locking beam is engaged. Second, the locking beam shutter must open before the PDH locking is enabled. Otherwise, the PDH electronics will register no signal on the photodiode used for generating the error signal, so the servo will throw the cavity out of lock if one is not careful. When transitioning from the locking phase to the data collection phase, everything must proceed in the precise reverse order to that just described.

# Appendix E

# Quantum Simulation Tutorial

# Introuction to Quantum Simulation with QuTiP

Miller Eaton

January 26, 2022

## 1 Intro and Install Instructions

Quantum Toolbox in Python (QuTiP) is an open-source software package useful for simulation a variety of quantum systems. Although it is well-documented online, it has many applications and can quickly turn into a deep rabbit hole if one isn't careful. In order to have a more directed focus, I attempt to give a brief introduciton as it applies to simulating quantum optical systems. To begin, it will be necessary to install Python locally along with the QuTiP libraries. I suggest installing Anaconda and using Jupyter Notebooks with Anaconda Navigator. Using Jupyter Notebooks allows for one to easily interact with the code and quickly edit and replot different segments.

After installing Anaconda, run the Anaconda Prompt to lauch a terminal. Next, if desired, create an environment for QuTiP using 'conda create –name myenv' where 'myenv' is the environment name. Activate this environment with 'conda activate myenv'. Once one has an Anaconda environment active (or simply in the base environment), one can install QuTip with pip using the command 'pip install qutip'. Install Anaconda navigator with 'conda install anaconda-navigator'. After the install, launch Anaconda Navigator, and then Jupyter Notebook can be installed using the user interface. Alternatively, it can be installed via the terminal.

Because the quantized electromagnetic field can be described as a quantum harmonic oscillator as discussed briefly in Thesis Chapter 1, any quantum-optical mode (qumode) can be written in the Fock basis and described in terms of eigenstates of the harmonic oscillator. This is ammenable to computation with linear algebraic computing methods since all states can be represented by a discrete basis. Additionally, nearly all physical quantum states can be accurately described with a finite-dimensional Hilbert space - we impose an energy cutoff that limits the maximum Fock-basis component. This is equivalent to saying that the probability to find any number of photons above some maximum value in a given quantum state vanishes.

Of course, states in real life can have many orders of magnitude of photons. However, if a system has $N$ photons, then quantum effects generally only persist if the losses are on the order of $\frac{1}{N}$ or lower (see the Thesis Chapter on Quantum Sensing). Thus, for quantum effects, a realistic Hilbert space for a single qumode is generally not too large. However, the Hilbert space scales exponentially when considering multiply qumodes. The point of quantum computing is to take advantage of this exponential scaling, so obviously it will be difficult to simulate classically, but the point here is that the Hilbert space scales as $d^n$ where $d$ is the Hilbert space dimension of a single mode and $n$ is the number of modes. With qumodes, the base itself, $d$, can be quite large, which leads to running into a classical wall much faster than with qubits.

For example, supposed I have a computer with 16 GB of memory which is $\sim 1.3 \times 10^{11}$ bits. To perform one simple simulation, the computer will at minimum have to hold one matrix describing the state, one matrix describing the operator to be applied, and one matrix describing the output

state. This is even very optimistic, as I am ignorning any intermediate memory needed to calculate floating-point operations of the matrix multiplication. Regardless, for a Hilber space of $D = d^n$, the state will have $\sim 2D$ real parameters (since probability amplitudes can be complex), and the operator will need memory allocation for $\sim (2D)^2$ parameters. If each parameter is a float data type containing 32 bits, then to simulate this simple evolution we need $32 \times (4D + 4D^2)$ bits of memory at minimum. This leads to a maximum total Hilbert space of $D \approx 32,000$, which means one could not quite simulate 15 qubits. However, for an optical system with 1% overall loss, which is not unreasonable experimentally, one could have up to $N \approx 100$ photons in an optical mode with quantum effects. Setting $d = 100$ means we can only simulate 2 qumodes with the same 16 GB machine! This, again, is the power of continuous-variable quantum optics we would like to leverage.

I digress - back to simulating quantum optics. The first order of business is to understand how quantum optical states can be represented in a way that a machine could easily perform calculations. The natural way is already prescribed by linear algebra. We simply assign each state a complex vector that gives the Fock-basis representation. As an example, the single-mode Fock state $|n\rangle$ can be described by the unit vector $(0, 0, ..., 1, ..., 0)$, where the 1 is located in the $n$-th possition of the $d$-dimensional vector. Similarly, operators are represented by $d^2$-dimensional square matrices. Multi-mode states and operators are just created by taking the tensor product between single-mode state vectors and operator matrices, respectively. This will become more clear with some examples, so let's start by importing some other Python packages used for mathematics and plotting, and they we'll import everything from qutip. Generally, it saves memory and computation time by only importing specific functions from the QuTip library, but for now, we won't worry about that and will just import everything.

## 2 Basics

[1]:
```
import numpy as np
import  matplotlib.pyplot as plt
import math, cmath, random, functools, scipy
from qutip import*
```

Let's start by defining a small Hilbert space, just large enough to see what's going on without become overly cumbersome. We'll then define the bosonic annihilation operator, $\hat{a}$, along with a single photon Fock state, $|1\rangle$.

[2]:
```
Dim=4 #Define the Hilbert Space dimension
a1=destroy(Dim) #the single-mode annihilation operator
photon=fock(Dim,1) # a Fock state existing in Hilbert space size Dim with␣
 ↪excitation 1
```

Start by looking at exactly what the variable 'photon' is. Notice that in addition to being a vector, the QuTiP library makes it a Quantum object from the class Qobj(), which has several useful features in addition to keeping track of the quantum properties (such as being a bra or ket).

[3]:
```
photon
```

[3]:

Quantum object: dims = [[4], [1]], shape = (4, 1), type = ket

$$\begin{pmatrix} 0.0 \\ 1.0 \\ 0.0 \\ 0.0 \end{pmatrix}$$

The above vector is a ket, but we can get the bra by taking the dagger of the quantum object. In fact, the .dag() method can be applied to any quantum object to get the daggered state or operator.

[4]:
```
photon.dag()
```

[4]:
Quantum object: dims = [[1], [4]], shape = (1, 4), type = bra

$$\begin{pmatrix} 0.0 & 1.0 & 0.0 & 0.0 \end{pmatrix}$$

Note that instead of using the fock() function, we could also define an array and later pass it to the Qobj() class. This is a useful way to initialize any input state vector. The .full() method converts a quantum object back to an array.

[5]:
```
my_array=np.array([0,1,0,0])
print('To a quantum object: ', Qobj(my_array))
print('And back to an array... ',Qobj(my_array).full())
```

```
To a quantum object:  Quantum object: dims = [[4], [1]], shape = (4, 1), type =
ket
Qobj data =
[[0.]
 [1.]
 [0.]
 [0.]]
And back to an array…  [[0.+0.j]
 [1.+0.j]
 [0.+0.j]
 [0.+0.j]]
```

I can also use the .norm() method to find the norm of any vector. In this way I can normalize an otherwise non-normalized state. I can also use the .unit() method to turn any vector into a unit vector. Suppose I want to input the superposition $\frac{1}{\sqrt{2}}(|1\rangle + |2\rangle)$ but am lazy and neglect the normalizing factor.

[6]:
```
my_array=np.array([0,1,1,0])
superpos=Qobj(my_array)
normed_superpos=superpos.unit()
normed_superpos
```

[6]:

Quantum object: dims = [[4], [1]], shape = (4, 1), type = ket

$$\begin{pmatrix} 0.0 \\ 0.707 \\ 0.707 \\ 0.0 \end{pmatrix}$$

Next, let's look at an operator, the annihilation operator defined above.

[7]:
```
a1
```

[7]:
Quantum object: dims = [[4], [4]], shape = (4, 4), type = oper, isherm = False

$$\begin{pmatrix} 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.414 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.732 \\ 0.0 & 0.0 & 0.0 & 0.0 \end{pmatrix}$$

The annihilation operator is define such that it will map every $n$ Fock state to the $n-1$ Fock state along with the necessary coefficients as expected: $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$. Similarly, one can use the creation operator simply by taking the dagger of annihilation operator. Let's apply $\hat{a}^{\dagger 2}$ to the vacuum to generate the two-photon Fock state.

[8]:
```
vac=fock(Dim,0)
two_photon=(a1.dag()**2*vac).unit()
two_photon
```

[8]:
Quantum object: dims = [[4], [1]], shape = (4, 1), type = ket

$$\begin{pmatrix} 0.0 \\ 0.0 \\ 1.0 \\ 0.0 \end{pmatrix}$$

Superpositions can be defined by turning an array into a quantum object, but they can also be formed by summing Fock states. For example, lets make the state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|2\rangle)$.

[9]:
```
superpos2=(fock(Dim,0)+1j*fock(Dim,2)).unit()
superpos2
```

[9]:
Quantum object: dims = [[4], [1]], shape = (4, 1), type = ket

$$\begin{pmatrix} 0.707 \\ 0.0 \\ 0.707j \\ 0.0 \end{pmatrix}$$

Because all real quantum state simulations will need to include imperfects such as loss, we need a way to deal with mixtures. Fortunately, density matrices are equally easy to deal with. However,

note that the density matrix takes up more memory and thus takes longer to perform calculations with since the computer needs to put a complex number at each element in the square matrix. We can get the density matrix in two ways: either use the ket2dm() function or take a product of the ket and the bra. We can check that the two density matrices are the same by using the fidelity() function.

[10]:
```
dm1=ket2dm(superpos2)
dm2=superpos2*superpos2.dag()
print(dm1)
print('The fidelity is: ', fidelity(dm1,dm2)**2)
```

```
Quantum object: dims = [[4], [4]], shape = (4, 4), type = oper, isherm = True
Qobj data =
[[0.5+0.j  0. +0.j  0. -0.5j 0. +0.j ]
 [0. +0.j  0. +0.j  0. +0.j  0. +0.j ]
 [0. +0.5j 0. +0.j  0.5+0.j  0. +0.j ]
 [0. +0.j  0. +0.j  0. +0.j  0. +0.j ]]
The fidelity is:  1.0000000149011607
```

The fidelity is 1, so they are the same state. The fractial deviations are due to round-off errors when using the built in functions and a finite Hilbert space. Generally, numerical calculations aren't reliable beyond the fourth or fifth decimal. Note that this fidelity function returns

$$\mathrm{Tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right]$$

($\mathrm{Tr}\left[\sqrt{\sigma\rho}\right]$ if either state is pure), so we also need to square it to get the regular definition. Without squaring, we get the overlap of the absolute value of the probability amplitudes, but because we want to know how close two states are, comaring probabilites makes more sense. This is clearer if we take the state dm1 which is an even superposition of $|0\rangle$ and $|2\rangle$, so we would expect the fidelity with either consistuent part to be $1/2$ and not $\frac{1}{\sqrt{2}}$.

[11]:
```
print('Without squaring: ',fidelity(dm1,fock(Dim,2)))
print('With squaring: ', fidelity(dm1,fock(Dim,2))**2)
```

```
Without squaring:  0.7071067811865475
With squaring:  0.4999999999999999
```

Now, let's try to calculate some expectation values. For a given operator $\hat{O}$, this can be done either as $\langle\psi|\hat{O}|\psi\rangle$ or by taking the trace $\mathrm{Tr}[\rho\hat{O}]$. Let's do the expectation of photon number, $\langle\hat{a}^\dagger\hat{a}\rangle$. Notice that while the answer is the same, the data type isn't. Using the latter method of the trace, we just get a number. Using the former, however, we get a $1\times 1$ quantum object.

[12]:
```
print(superpos2.dag()*a1.dag()*a1*superpos2)
print((dm1*a1.dag()*a1).tr())
```

```
Quantum object: dims = [[1], [1]], shape = (1, 1), type = bra
Qobj data =
[[1.]]
1.0
```

Occassionally it is useful to time code segments, as occassionally diffect methods achieving the same calculation may take different times. In general, if one can perform a calculation with pure state, it is generally easier (faster or less memory) to perform calculations without density matrices. Of course, if one later wishes to introduce loss, the density matrix formalism will become necessary.

[13]:
```
Q=(1/np.sqrt(2))*(a1.dag()+a1)
import time
t1=time.time()
var=(dm1*Q**2).tr()-(dm1*Q).tr()**2
t2=time.time()
print('Variance of Q is ', var, ' and took ', t2-t1)
```

```
Variance of Q is  1.4999999999999996  and took  0.0009970664978027344
```

Next, consider how we treat multi-mode systems. If one has two qumodes, $|\psi\rangle|\phi\rangle$, this state can be encode by taking the tensor product of the two. Similarly, I could make a two mode operator out of a tensor product of two single-mode operators. Remember that taking a tensor product is essentially how the Hilbert spaces merge, so we always have to be careful to make the dimensions of everything match.

As an example, let's take $|\psi\rangle = |0\rangle$ and $|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. Now, apply the creation operator to just the first mode to add in one photon, and calculate the mean photon number, which can be done on the individual mode separately and then later add the result, so the tensor products aren't really necessary. Next, apply a superposition of a creation operator to each mode: $\frac{1}{\sqrt{2}}(\hat{a}_1^\dagger + \hat{a}_2^\dagger)$. Now taking the tensor products is absolutely necessaray. One might naively think that the expectation of final photon number will be the same in both cases, but since $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$, the coefficient will change the ratio of the compenents of the superpostion. Thus, applying creation operators is not the same as 'adding' a photon. Note that qeye($Dim$) is the identity matrix as a quantum object for a $Dim$ dimensional Hilbert space. I'll also reduce the size of the Hilbert space for easier visualization, and define an annihilation operator for each mode. This isn't strictly necessary, but it is helpful as a bookkeeping method.

[14]:
```
Dim=3; a1=destroy(Dim); a2=destroy(Dim)
psi=fock(Dim,1)
phi=(fock(Dim,0)+fock(Dim,1)).unit()
in_state=tensor(psi,phi)
print('The initial state is, ', in_state)
out1=(tensor(a1.dag(),qeye(Dim))*in_state).unit()
print()
print('Applying the creation operator to the first mode and normalizing yields␣
 ↪', out1)
exp1=out1.dag()*(tensor(a1.dag()*a1,qeye(Dim))+tensor(qeye(Dim),a2.
 ↪dag()*a2))*out1
print('Which has photon-number expectation value of ', exp1)
out2=((tensor(a1.dag(),qeye(Dim))+tensor(qeye(Dim),a2.dag()))*in_state).unit()
print('Applying a superposition of creation operators to both modes and␣
 ↪normalizing yields ', out2)
```

```
exp2=out2.dag()*(tensor(a1.dag()*a1,qeye(Dim))+tensor(qeye(Dim),a2.
 ↪dag()*a2))*out2
print('Which has photon-number expectation value of ', exp2)
```

```
The initial state is,  Quantum object: dims = [[3, 3], [1, 1]], shape = (9, 1),
type = ket
Qobj data =
[[0.        ]
 [0.        ]
 [0.        ]
 [0.70710678]
 [0.70710678]
 [0.        ]
 [0.        ]
 [0.        ]
 [0.        ]]

Applying the creation operator to the first mode and normalizing yields  Quantum
object: dims = [[3, 3], [1, 1]], shape = (9, 1), type = ket
Qobj data =
[[0.        ]
 [0.        ]
 [0.        ]
 [0.        ]
 [0.        ]
 [0.        ]
 [0.70710678]
 [0.70710678]
 [0.        ]]
Which has photon-number expectation value of  Quantum object: dims = [[1], [1]],
shape = (1, 1), type = bra
Qobj data =
[[2.5]]
Applying a superposition of creation operators to both modes and normalizing
yields  Quantum object: dims = [[3, 3], [1, 1]], shape = (9, 1), type = ket
Qobj data =
[[0.        ]
 [0.        ]
 [0.        ]
 [0.        ]
 [0.37796447]
 [0.53452248]
 [0.53452248]
 [0.53452248]
 [0.        ]]
Which has photon-number expectation value of  Quantum object: dims = [[1], [1]],
shape = (1, 1), type = bra
```

```
Qobj data =
[[2.57142857]]
```

## 3 Functions and Operators

QuTiP comes with a lot of predefined functions that can be directly used to make or manipulate quantum states. For example, instead of using fock($Dim, n$) to initialize states, one could begin with a coherent state, $|\alpha\rangle$ with by using coherent($Dim, \alpha$). QuTiP also has built in operators for squeezing, displacements, spins, commutators, and several more. However, differences in definitions (factors of 2, $\hbar$, etc...) can occasionally cause problems if one is inconsistent. As QuTiP does not have all operators predefined, and mixing my own definitions with their may sometimes be incompatible, I generally prefer defining all of my own functions from scratch.

To demonstrate this, I will define a displacement operator, a single-mode rotation operation, and a squeeze operator. After generating a state we can visualize its photon number distribution, Wigner function, and calculate expectation values. For example to apply an optical phase (rotation in phase space), I would need to create the operator

$$\hat{R} = e^{i\theta\hat{a}^\dagger\hat{a}}.$$

Because the creation and annihilation operators are already matrices, I must use the method .expm() which exponentiates an existing operator.

[15]:
```
def D(alpha, Dim):
    #alpha is the complex displacement amplitude
    #Dim is the Hilbert space dimension
    a1=destroy(Dim)
    op=alpha*a1.dag()-np.conj(alpha)*a1;
    return op.expm() #exponentites an operator

def R(theta, Dim):
    #theta is the rotation angle
    #Dim is the Hilbert space dimension
    a1=destroy(Dim)
    op=1j*theta*a1.dag()*a1;
    return op.expm() #exponentites an operator

def Sq(amp,phase,Dim):
    #amp is the strength of the squeeze (same as r)
    #phase is the squeezing angle
    #Dim is the Hilbert space dimension
    a1=destroy(Dim)
    op=(1/2)*(amp*np.e**(1j*phase)*a1.dag()**2-amp*np.e**(-1j*phase)*a1**2)
    return op.expm() #exponentites an operator
```

[16]:
```
Dim=30 #Hilbert dimension
a1=destroy(Dim) # annihilation operator
Q=(a1.dag()+a1)/np.sqrt(2)
P=1j*(a1.dag()-a1)/np.sqrt(2)
```

```
vac=fock(Dim,0) #vacuum state

alpha=(1+2j)/np.sqrt(2) #displacement value
coher=D(alpha,Dim)*vac #coherent state

print('Expectation value of Q is ', expect(Q,coher))
print('Variance of Q is ', variance(Q,coher))
print('Expectation value of P is ', expect(P,coher))
print('Variance of P is ', variance(P,coher))
```

```
Expectation value of Q is  0.9999999999999997
Variance of Q is  0.49999999999999956
Expectation value of P is  1.9999999999999991
Variance of P is  0.5000000000000009
```

The code above starts by making a vacuum state and then uses the displacement function we defined to generate the coherent state. I then calculated the expectation values and variance of the quaradure operators $\hat{Q}$ and $\hat{P}$ using inbuilt QuTiP functions. These function return the expectation value and variance, respectively, of an operator for a certain state, eg,

$$\text{expect}(\hat{O}, |\psi\rangle) = \langle\psi|\hat{O}|\psi\rangle.$$

Note that these functions work if the input state is a ket or a density matrix.

We see that $\langle\hat{Q}\rangle = \frac{1}{\sqrt{2}}\text{Re}(\alpha)$ and $\langle\hat{P}\rangle = \frac{1}{\sqrt{2}}\text{Im}(\alpha)$ as expected. Also, the variance for both $\hat{Q}$ and $\hat{P}$ is $1/2$, which is the same as vacuum ($\hbar = 1$).

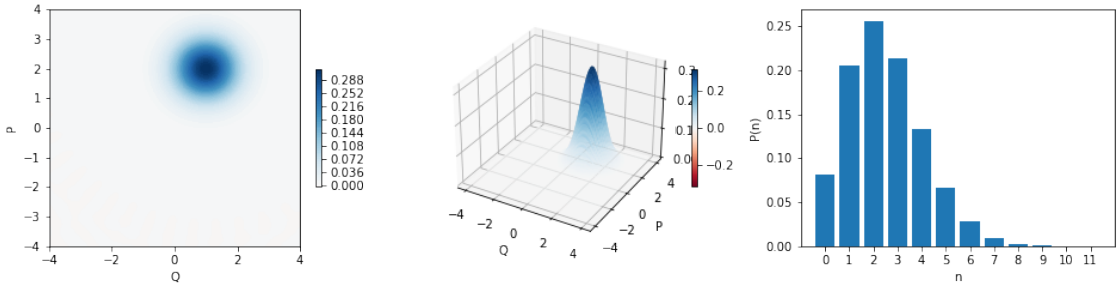
Next, let's define a plotting function to easily visualize the photon-number distribution and the Wigner function of the states.

[17]:
```
import matplotlib.pyplot as plt
from matplotlib import cm
import matplotlib as mpl



def state_plot(state,Qmax,points,numcut):
    #Qmax: highest quadrature value at the plot edge
    #points: number of points in one axis from 0 out
    #numcut: define the maximum photon number to show for the probability␣
 ↪distribution
    dx=Qmax/points
    xvec = np.arange(-points,points+1)*dx
    X,Y = np.meshgrid(xvec, xvec)
    W=wigner(state,xvec,xvec)

    fig = plt.figure(figsize=(16,8))
    # `ax` is a 3D-aware axis instance, because of the projection='3d' keyword␣
 ↪argument to add_subplot
    ax = fig.add_subplot(2, 3, 1)
```

```
    #define a surface plot and use vmin and vmax to set the limits of the␣
↪colorscale to ensure the middle is at the origin
    p = ax.contourf(X, Y, W, 60, cmap=cm.RdBu, vmin=-1/np.pi,vmax=1/np.pi)
    ax.set_ylabel('P')
    ax.set_xlabel('Q')
    cb = fig.colorbar(p, shrink = 0.5) # add a colorbar
    # surface_plot with color grading and color bar
    ax = fig.add_subplot(2, 3, 2, projection='3d')
    p = ax.plot_surface(X, Y, W, rstride=1, cstride=1, cmap=cm.RdBu, vmin=-1/np.
↪pi,vmax=1/np.pi, linewidth=0.5)
    ax.set_ylabel('P')
    ax.set_xlabel('Q')
    cb = fig.colorbar(p,shrink = .5)
    ax=fig.add_subplot(2, 3, 3)
    if isket(state)==True:
        #if the state is a ket, take the square magnitude of each element
        p=ax.bar(range(numcut),(np.abs(state.full().flatten()))**2)[0:
↪numcut],tick_label=range(numcut))
    else:
        #if density matrix, just use the diagonal
        p=ax.bar(range(numcut),state.diag()[0:numcut],tick_label=range(numcut))
    ax.set_ylabel('P(n)')
    ax.set_xlabel('n')

    return W, fig
```

[18]:
```
Wig_plot_coh=state_plot(coher,4,50,12)
print('Mean photon number is: ', expect(a1.dag()*a1,coher))
```

```
Mean photon number is:  2.4999999999999982
```

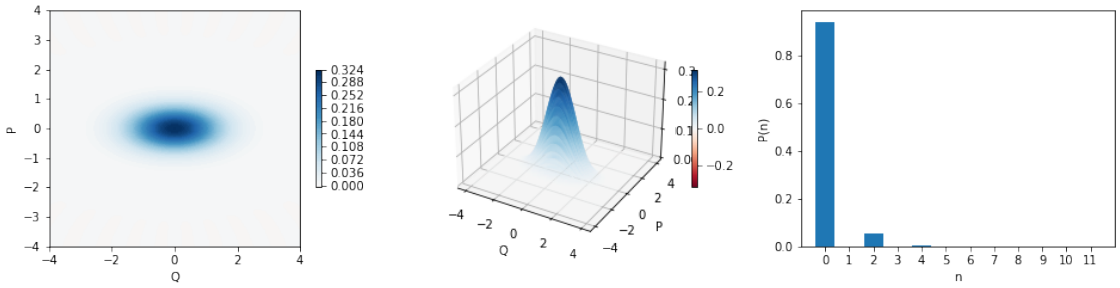



We can now see the nice Wigner function of our state along with the Poissonian photon number distribtuion of $\langle \hat{a}^\dagger \hat{a} \rangle = |\alpha|^2$.

Next, do the same thing but with a squeezed vacuum state. Then we can also apply the rotation operator after squeezing the vacuum to change the squeezing angle.

[19]:
```
sqz_param=0.345 #About 3 dB of squeezing
sqz_vac=Sq(sqz_param, 0, Dim)*vac
Wig_plot_sqz=state_plot(sqz_vac,4,50,12)
print('Mean photon number is: ', expect(a1.dag()*a1,sqz_vac))
print('Expectation value of Q is ', expect(Q,sqz_vac))
print('Variance of Q is ', variance(Q,sqz_vac))
print('Expectation value of P is ', expect(P,sqz_vac))
print('Variance of P is ', variance(P,sqz_vac))
```

```
Mean photon number is:  0.12382290057728355
Expectation value of Q is  0.0
Variance of Q is  0.9968577666215033
Expectation value of P is  0.0
Variance of P is  0.25078803453306353
```



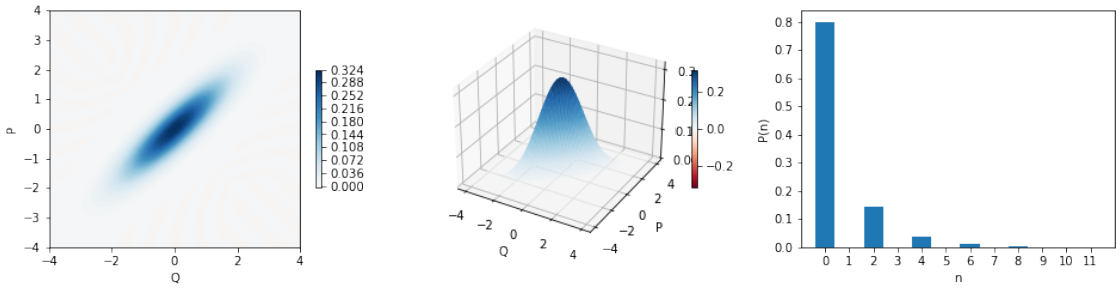

Above, we can see that the uncertainty in the $\hat{P}$ quadrature has been reduced, while the $\hat{Q}$ uncertainty is increased. Since our squeezing parameter was equivalent to 3 dB of squeezing, we see that the uncertainty doubled from the vacuum in $\hat{Q}$ and is reduced by half for $\hat{P}$. Another interesting quality of squeezed states, showcased here, is that they only contain *even* photon number probabilities.

Let's apply the same squeeze operator again to squeeze twice as hard, then rotate using the phase operator.

[20]:
```
phase=np.pi/4
sqz_vac2=R(phase,Dim)*Sq(sqz_param, 0, Dim)*sqz_vac
Wig_plot_sqz2=state_plot(sqz_vac2,4,50,12)
print('Mean photon number is: ', expect(a1.dag()*a1,sqz_vac2))
print('Variance of Q is ', variance(Q,sqz_vac2))
print('Variance of P is ', variance(P,sqz_vac2))
print('Variance of (Q+P)/sqrt(2) is ', variance((Q+P)/np.sqrt(2),sqz_vac2))
print('Variance of (Q-P)/sqrt(2) is ', variance((Q-P)/np.sqrt(2),sqz_vac2))
```

```
Mean photon number is:  0.5566199519267636
Variance of Q is  1.0566199519267634
Variance of P is  1.0566199519267634
Variance of (Q+P)/sqrt(2) is  1.9874490939334797
```

```
Variance of (Q-P)/sqrt(2) is  0.12579080992004604
```




The above state is squeezed twice as hard, so we should expect the minimum uncertainty to be $\frac{1}{4}$ of the vacuum uncertainty (so about $\frac{1}{8}$). However, we also rotated the state by an angle of $\frac{\pi}{4}$. Thus $\hat{Q}$ and $\hat{P}$ are no longer the squeezed and antisqueezed quadratures, but rather $\frac{(\hat{Q}-\hat{P})}{\sqrt{2}}$ and $\frac{(\hat{Q}+\hat{P})}{\sqrt{2}}$.

Plots of the Wigner function are helpful because they give a general idea about both quadratures simultaneous. They are also useful as they easily indicate the difference between Gaussian and non-Gaussian state (non-Gaussian pure states have negativity in phase space). It is important to note that for a statistical mixture, the Wigner function may be positive everywhere yet still by non-Gaussian. On occassion, however, it is also nice to directly plot the probability distribution of the quadratures. This can be done by integrating the Wigner function over the conjugate quadrature.

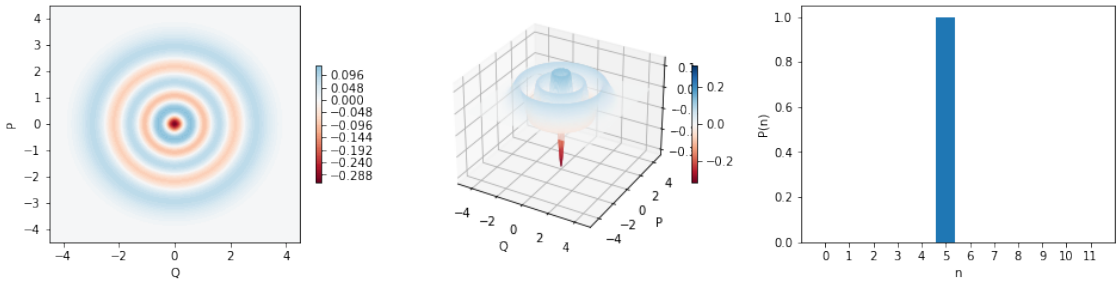
Let's try this with some Fock states. We know that

$$|n\rangle = \frac{\pi^{-1/4}}{\sqrt{2^n n!}} \int dx e^{-x^2/2} H_n(x) |x\rangle_q$$

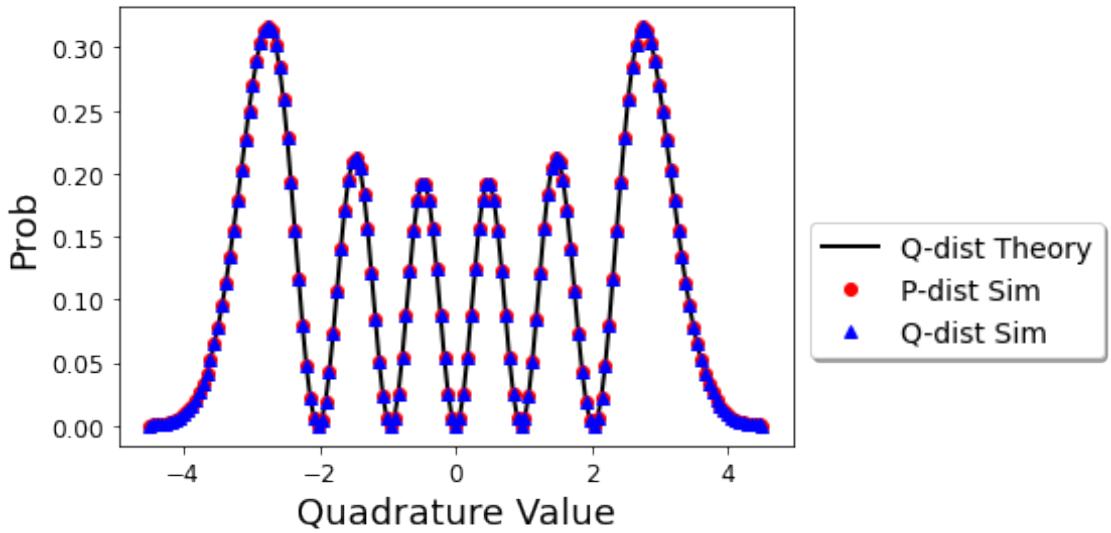
where $H_n(x)$ are the $n$-th order Hermite polynomials and $|x\rangle_q$ is an eigenstate of the $\hat{Q}$-quadrature. Let's verify this by calculating the Wigner function of a Fock state, integrating it, and comparing the resultant probability to the Hermite polynomials.

[21]:
```
from scipy.integrate import simps
from scipy.special import hermite
from scipy.special import factorial as fac

five_photon=fock(Dim,5)
#increase the number of points in the plot so the integral will be more accurate
Wig_plot_fock=state_plot(five_photon,4.5,80,12)
print('Minimum value of the Wigner function is ', np.min(Wig_plot_fock[0]))
```

```
Minimum value of the Wigner function is  -0.3183098861837905
```




[22]:
```
dx=4.5/80
xvec = np.arange(-80,81)*dx #Use same xvec as above
W=wigner(five_photon,xvec,xvec) #Wigner function as a 2D array
pdist=simps(W,xvec,dx) #integrate over the q axis to get the p distribution
qdist=simps(W.T,xvec,dx) #integrate over the p axis to get the q distribtuion

###Define the theoretical 5-photon Fock state distribution in terms of the␣
 ↪Hermite polynomial###
theory_dist=(np.e**(-xvec**2)*(hermite(5)(xvec))**2)/(2**5*fac(5)*np.pi**(1/2))

plt.plot(xvec,theory_dist,'-k', xvec,pdist,'or',xvec,qdist,'^b', linewidth=2)
plt.legend(('Q-dist Theory','P-dist Sim','Q-dist Sim', ),
           loc=(1.025,0.2), shadow=True, fontsize=14)
plt.xlabel('Quadrature Value',fontsize=18)
plt.ylabel('Prob',fontsize=18)
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)
plt.show()
```



From above, we get great agreement between all of the different calculations. Note that to get the probability distribution along any generalized quadrature, it is usually easier computationally to simply rotate the state in phase-space first (with the rotation operator), and then calculate and integrate the Wigner function along the 'new' $\hat{Q}$ and $\hat{P}$ axes.

## 4 Interactions and Detection

Next, we'll go through how to simulate basic interactions between qumodes and simulate projective measurement. To do this, let's start by simulating the Hong-Ou-Mandel effect. First, we'll need to define a beamsplitter operator along with a function for detection.

[23]:
```
def BS(theta,Dim):
    a1=destroy(Dim)
    a2=destroy(Dim)
    op=((theta)*(tensor(a1,a2.dag()) - tensor(a1.dag(),a2)))
    return op.expm()
def PNR_m1(n,Dim): #pure PNR measurement on mode 1 of two modes
    proj=tensor(fock(Dim,n)*fock(Dim,n).dag(),qeye(Dim))
    return proj
```

The Hong-Ou-Mandel effect occurs when a single photon is input on each input of a balanced beamsplitter. Due to the quantum interference, the photons will perfectly 'bunch' in that they will both exit the beamsplitter together on the same output port. To show this, let's look at the probability to measure a single photon on one of the beamsplitter outputs as a function of beamsplitter reflectivity. When the beamsplitter is perfectly transmittive, the photons don't interact and the detector should always measure one photon. Once the beamsplitter becomes balanced with $r = t$, we expect the photons to bunch, so the detector would either see zero or two photons. Thus the probability to measure a single photon drops to zero.

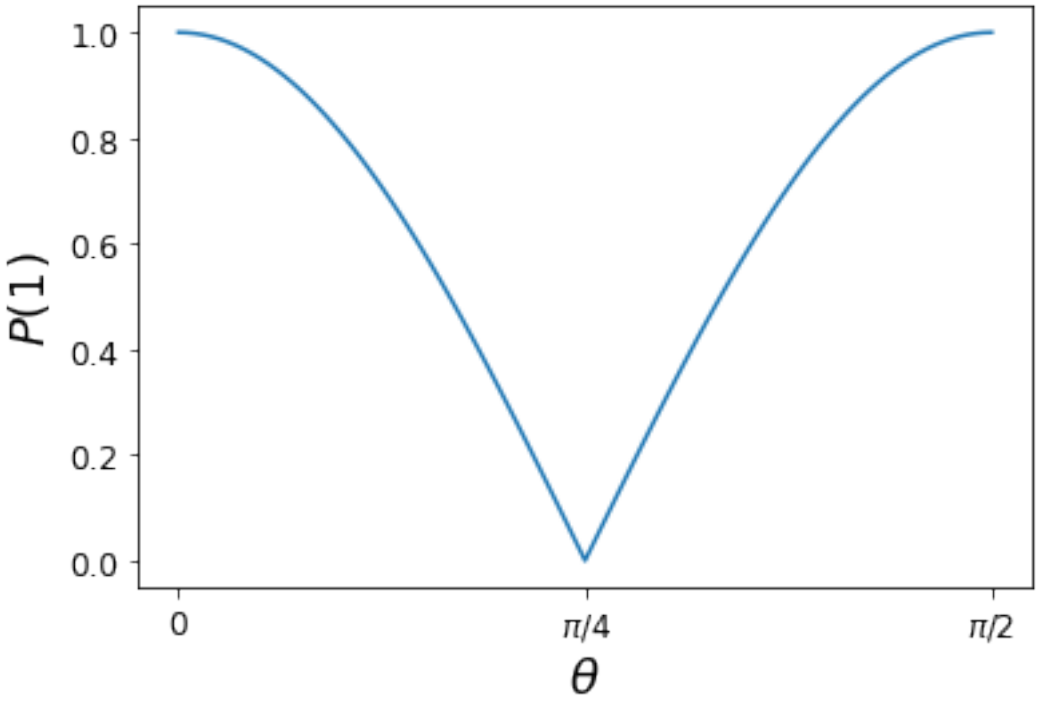
To show this, we'll but a single photon in each input and apply the beamsplitter unitary,

$$\hat{B} = e^{\theta(\hat{a}_1\hat{a}_2^\dagger - \hat{a}_1^\dagger\hat{a}_2)}.$$

Then we'll plot the probability to measure a single photon on output one as a function of $\theta$.

[24]:
```
Dim=5
photon1=fock(Dim,1)
photon2=fock(Dim,1)
in_state=tensor(photon1,photon2) #two-mode input state
theta_vec=np.linspace(0,np.pi/2,101) #values for beamsplitter angle
single_photon_prob=[] #empty list to hold our probabilities
for i in theta_vec: #Loop through each theta value
    out_state=BS(i,Dim)*in_state #state after the beamsplitter
    prob=(PNR_m1(1,Dim)*out_state).norm() #Project the output mode 1 on a␣
 ↪single photon and get the coefficient
    single_photon_prob.append(prob)
plt.plot(theta_vec,single_photon_prob)
```

```
plt.xlabel(r'$\theta$',fontsize=18)
plt.ylabel(r'$P(1)$',fontsize=18)
plt.xticks([0,np.pi/4,np.pi/2],['0',r'$\pi/4$',r'$\pi/2$'],fontsize=12)
plt.yticks(fontsize=12)
plt.show()
```




## 5 Photon Catalysis

Building on the idea of the Hong-Ou-Mandel effect, let's go on to simulate Photon Catalysis, which can be thought of as a generalized Hong-Ou-Mandel effect. With this, a single photon, $|1\rangle$ is input into one port of the beamsplitter while the other port is allowed to be any quantum state, $|\psi\rangle$. Next, the user measures one output mode with a photon-counting detector to measure $k$ photons. The output state can thus be expressed as

$$|\phi\rangle = {}_1\langle k|\hat{B}(\theta)|1\rangle_1|\psi\rangle_2.$$

Regardless of the initial state, $|\psi\rangle$, any particular fock-basis component of the output state may be filtered out depending on the beamsplitter parameter and post-selected detection result.

To simulate this, let's start with a single photon in mode one and a coherent state with mean photon number $|\alpha|^2 = 4$ in mode two, and then interfere them at a balanced beamsplitter. To leave things slightly more general, let's write everything in terms of density matrices instead of kets. This way, we can easily extent our code to mixtures as well as pure states. To do this, we can use the QuTiP function ket2dm() to turn an input ket into a density matrix, or we can immediately start from a fock state and coherent state density matrix using '_dm' as per built-in QuTiP utilities.

Next, remember that density matrices evolve as $\rho' = \hat{U}\rho\hat{U}^\dagger$, and we can get mode two alone by performing a partial trace over mode one. Thus our output density matrix, $\sigma$, for photon catalysis will look like:

$$\sigma = \frac{\mathrm{Tr}_1[|k\rangle\langle k|\hat{B}\rho_{12}\hat{B}^\dagger]}{P(k)}$$

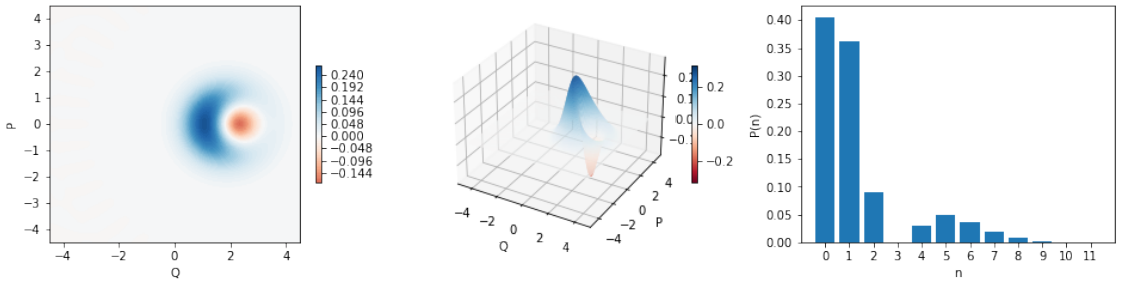
where

$$P(k) = \mathrm{Tr}[|k\rangle\langle k|\hat{B}\rho_{12}\hat{B}^\dagger]$$

is the probability of detecting $k$ photons in mode one.

[25]:
```
Dim=30
alpha=2
k=3
mode1=fock_dm(Dim,1) #single photon density matrix
mode2=coherent_dm(Dim,alpha) #coherent state density matrix
bs5050=BS(np.pi/4,Dim) #balanced beamsplitter
rho_in=tensor(mode1,mode2) #state into the beamsplitter

P_k=(PNR_m1(k,Dim)*bs5050*rho_in*bs5050.dag()).tr() #probability of measuring k␣
 ↪photons in mode 1
rho_out=(PNR_m1(k,Dim)*bs5050*rho_in*bs5050.dag()).ptrace(1)/P_k

#NOTE: The operation .ptrace(m) takes the partial trace over every mode
#EXCEPT m, where the numbering startes at 0.  So .ptrace(1) means you
#keep the second mode and trace out over the first

print('Probability to detect ', k, ' photons in mode one is: ', P_k)
print('Verify the output is normalized: trace of ', r'$\rho$', ' is ', rho_out.
 ↪tr())
print('Probability of ', k, ' photons to now also be in mode two: ', rho_out.
 ↪diag()[3])
Wig_plot=state_plot(rho_out,4.5,80,12)
```

```
Probability to detect  3  photons in mode one is:  0.13533528323576166
Verify the output is normalized: trace of  $\rho$  is  1.0
Probability of  3  photons to now also be in mode two:  0.0
```



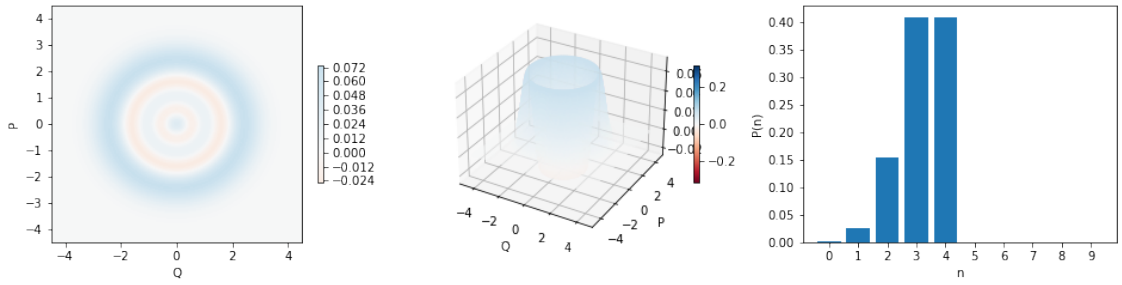
Interestingly, we see that for the balanced beamsplitter and a 3 photon detection in mode one, it is impossible to subsequently detect 3 photons in mode two as well! It turns out that the input state can be of any form, and that it doesn't matter what photon number was detected in mode one at all. Provided one input to the beamsplitter is a single-photon and the beamsplitter is balanced, a detection of $k$ photons in mode one ensures the $k$ photon probability in mode two is exactly zero. This can also be extended to a variety of beamsplitter angles, and in fact, to larger Fock-state inputs.

## 6 Simulating Loss

Loss is one of the largest sources of experimental error in any quantum optics experiment, and thus it is important to simulate. Loss degrades the purity of a state, so we must do everything with density matrices in this case (occassionally there are ways around this if all loss is commuted to the detection). Loss can be modeled by interfering a state with vacuum at a beamsplitter and tracing out over the reflected mode. Let's try this with a 5-photon Fock state and 20% loss.

[26]:
```
Dim=10
fivephotondm=fock_dm(Dim,4) #five-photon density matrix
vac=fock_dm(Dim,0) #vacuum density matrix
loss=0.2
loss_theta=np.arcsin(np.sqrt(loss)) #cos(theta)^2=loss
loss_bs=BS(loss_theta,Dim) #define the beamsplitter
bs_output=loss_bs*tensor(fivephotondm,vac)*loss_bs.dag() #two-mode state after␣
↪the beamsplitter
lossy_state=(bs_output.ptrace(0))/bs_output.tr() #trace out over reflected mode␣
↪and normalize

Wig_plot=state_plot(lossy_state,4.5,80,Dim)
print('Minimum value of the Wigner function is ', np.min(Wig_plot[0]))
```

```
Minimum value of the Wigner function is  -0.02584281790869087
```




Already at 20 loss, we see that the negativity in the Wigner function has nearly vanished, dropping from $-\frac{1}{\pi} \approx -0.32$ to $-0.025$. Although the negativity won't entirely vanish until the loss reaches 50%. Similarly, the five photon component is no longer even the dominant value. This demonstrates why quantum experiments can be quite challenging - loss must be very low to see quantum effects.

Suppose we weren't interested in getting the full quantum state, but only simulate the probability distribution. Instead of introducing the ancillary vacuum mode and needing to simulate a full two-mode system, we can instead use a detector POVM with imperfect detection efficieny to describe the loss. This is exeedingly useful when loss needs to be simulated for multi-mode systems and adding additional vacuum modes is infeasible to simulate use to the exponential scaling of the Hilbert space.

Detector inefficiency can be exactly described by a loss channel (beamsplitter and partial trace) before a perfect detector. Similarly, one could commute all of the losses in an experiment to the detection, and then just model the imperfect detector. To do this, we will project our pure state onto a detector described by a weighted sum of Fock-basis projectors that has the information about efficiency already accounted for.

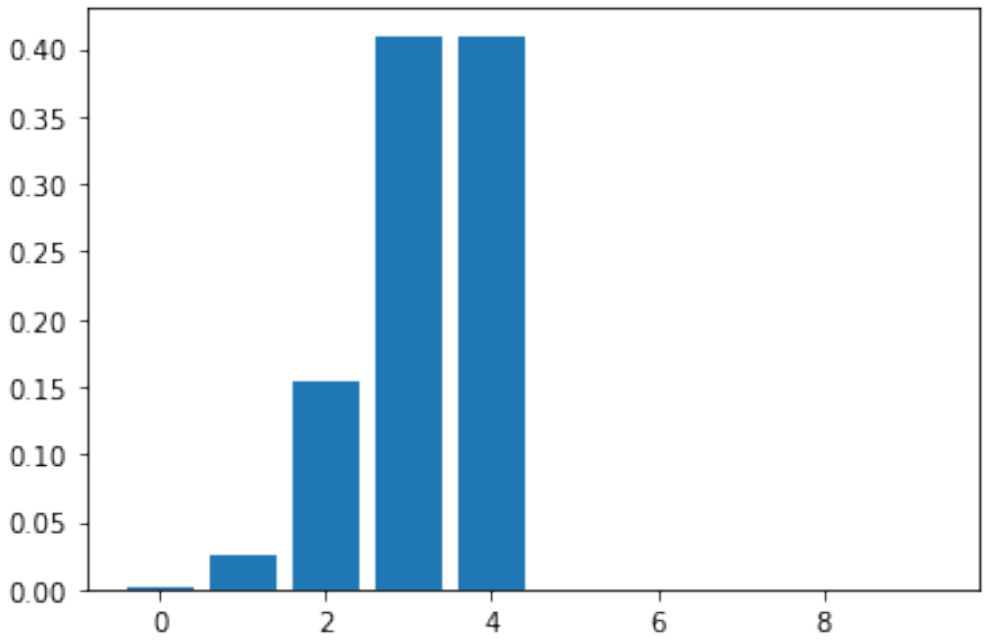
[27]:
```
from scipy.special import binom #import binomial coefficient calculator

def pnr_det_eta(eta, click, n_trunc,Dim):
    #eta: detector efficiency
    #click: number of photons measured
    #n_trunc: maximum number of photons resolvable by the detector
    #Dim: hilbert space dimension
    pi_n = 0;
    vec = np.arange(click,n_trunc)
    #Build the POVM sum for a detector with quantum efficiency eta<1
    for i in vec:
          pi_n +=  binom(i,click)*math.pow((1-eta),(i-click))*math.
 ↪pow(eta,click)*fock(Dim,i)*fock(Dim,i).dag()
    return Qobj(pi_n)
```

Now that this POVM is defined, I can calculate the probability to measure any given photon number of a quantum state, $\rho$, using

$$P(n) = \mathrm{Tr}[\hat{\Pi}_n \rho],$$

where $\hat{\Pi}_n$ is the POVM element associated with an $n$ photon detection event. This probability is exactly the same as the probability to measure $n$ photons with a perfect detector after $\rho$ has gone through a lossy channel where the transmission of the channel is given by $\eta$. Let's check this with the five-photon Fock state and look at the probability distribution for the same 20% loss ($\eta = 0.8$).

[28]:
```
photon_probs=[]
eta=0.8
for n in range(Dim):
    photon_probs.append((fivephotondm*pnr_det_eta(eta,n,Dim,Dim)).tr())
plt.bar(range(Dim),photon_probs)
plt.show()
```




The above distribution is exactly, the same as before, but we didn't need to include the ancillary vacuum mode! However, this methods will only yield the diagonal of the density matrix as opposed to the whole state. Alternatively, since loss is a well-defined statistical process, it is possible to define a linear map to map the density matrix before loss to one after loss, but we won't get into that here. The method shown above with detector POVMs is particularly useful in two mode cases where one would like to project on a particular photon number, such as in photon catalysis, but one doesn't want to use an ancillary mode.

The catalysis function defined below wraps everything we did with photon catalysis before into a single function, where the user inputs two states, a beamsplitter reflectivity, and the number of photons to project on with an imperfect detector.

[29]:
```
def catalysis(in1,in2,refl,num_det,eta,Dim):
    #'refl' is the beamsplitter reflectivity such that 1-refl=transmittance␣
 ↪(r^2+t^2=1 where refl=r^2)
    #first, turn the inputs into density matrices if they aren't already
    if isket(in1)==True:
        rho1=ket2dm(in1)
    else:
        rho1=in1
    if isket(in2)==True:
        rho2=ket2dm(in2)
    else:
        rho2=in2
    a1=destroy(Dim)
    a2=destroy(Dim)
```

```
    Projector = tensor(pnr_det_eta(eta, num_det, Dim,Dim),qeye(Dim)); #Detector␣
↪POVM used to project mode 1
    Initial_state=tensor(rho1,rho2); #tensor product of two input density␣
↪matrices
    theta=np.arccos(np.sqrt(refl)); #beamsplitter parameter

    BS1= ((theta)*(tensor(a1,a2.dag()) - tensor(a1.dag(),a2))).expm() #Unitary␣
↪beamsplitter operator

    Rho=BS1*Initial_state*BS1.dag(); #evolve the density matrix through the␣
↪beamsplitter

    P=(Rho*Projector).tr() #Find the probability to detect 'num' photons in␣
↪mode 1
    print('The probability of a sucessful detection is:',P)

    Rho_filtered = ((Rho*Projector).ptrace(1))/((Rho*Projector).tr()) #find the␣
↪normalized output state
    '''The operation .ptrace(m) takes the partial trace over every mode
    EXCEPT m, where the numbering startes at 0.  So .ptrace(1) means you
    keep the second mode'''
    print('BS has reflectivity',refl,' and I am detecting the |',num_det,
          '> state.')
    return Rho_filtered
```

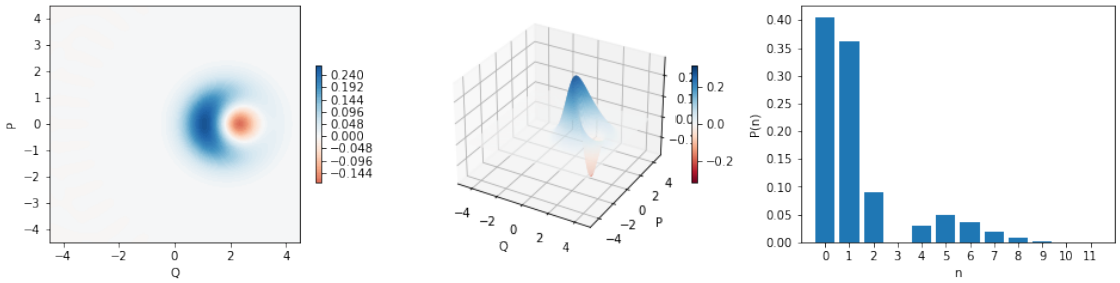
First, replicate what we did before with perfect detector efficiency, single photon on one input, coherent state on the other, balanced beamsplitter, and a projection of 3 photons on output mode one. Then plot the resulting state.

Next, do the same considering 10% loss on the projection detector and compare the results.

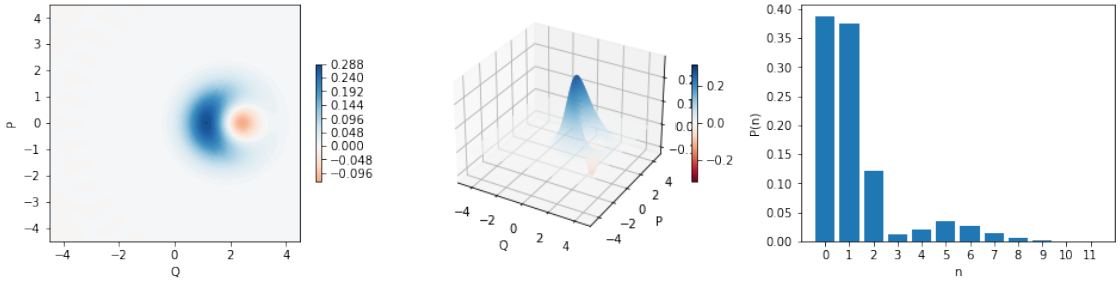
[30]:
```
Dim=30
alpha=2
k=3
eta1=1
state1=catalysis(fock_dm(Dim,1),coherent_dm(Dim,alpha),0.5,k,eta1,Dim)
Wig_plot1=state_plot(state1,4.5,80,12)
print('Minimum value of the Wigner function without loss is ', np.
 ↪min(Wig_plot1[0]))
plt.show()

eta2=0.9
state2=catalysis(fock_dm(Dim,1),coherent_dm(Dim,alpha),0.5,k,eta2,Dim)
Wig_plot2=state_plot(state2,4.5,80,12)
print('Minimum value of the Wigner function with 10% loss is ', np.
 ↪min(Wig_plot2[0]))
```

```
The probability of a sucessful detection is: 0.13533528323576166
```

```
BS has reflectivity 0.5  and I am detecting the | 3 > state.
Minimum value of the Wigner function without loss is  -0.17903979403201714
```



```
The probability of a sucessful detection is: 0.14621017260273436
BS has reflectivity 0.5  and I am detecting the | 3 > state.
Minimum value of the Wigner function with 10% loss is  -0.1224461516440973
```



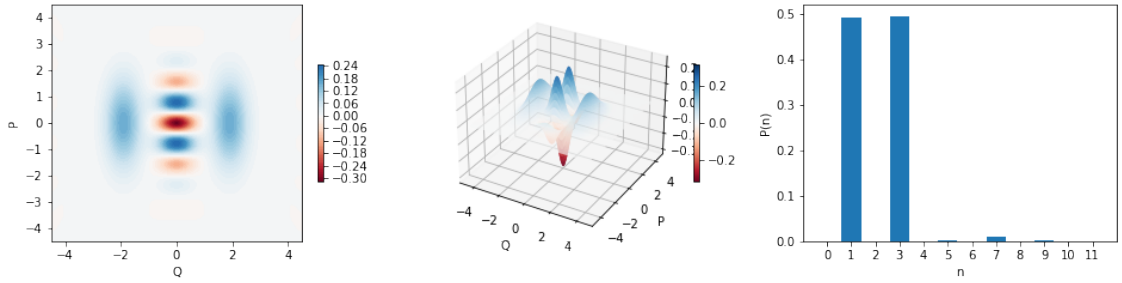
We can now directly see the effects of loss for this process. The amount of Wigner negativity has decreased, but perhaps more importantly, the probability to measure 3 photons in the new output is no longer zero.

## 7 Example: Generating Cat States

I will now show an example using photon catalysis to make cat states. I'll first define a superposition of squeezed vacuum (SSV) state, otherwise known as a squeezed cat state, to use as a benchmark. Then, using single photons, a coherent state, and PNR detectors, I will show how repeating photon catalysis can generate the same SSV state. More details can be found in M Eaton, R Nehra, O Pfister, NJP 21 (11), 113034 (2019).

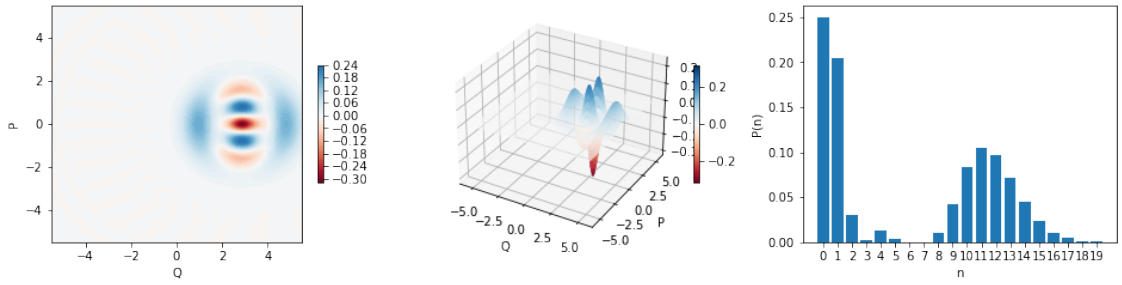
[31]:
```
def SSV_minus(r,alpha,Dim):
    state = ket2dm((displace(Dim,alpha)-displace(Dim,-alpha))
                   *squeeze(Dim,r)*fock(Dim,0))
    norm_state = state/state.tr()
    return norm_state
Dim=30
```

```
ssv=SSV_minus(0.476595,1.349,Dim) #benchmark squeezed cat-state
Wig_plot_cat=state_plot(ssv,4.5,80,12)
```

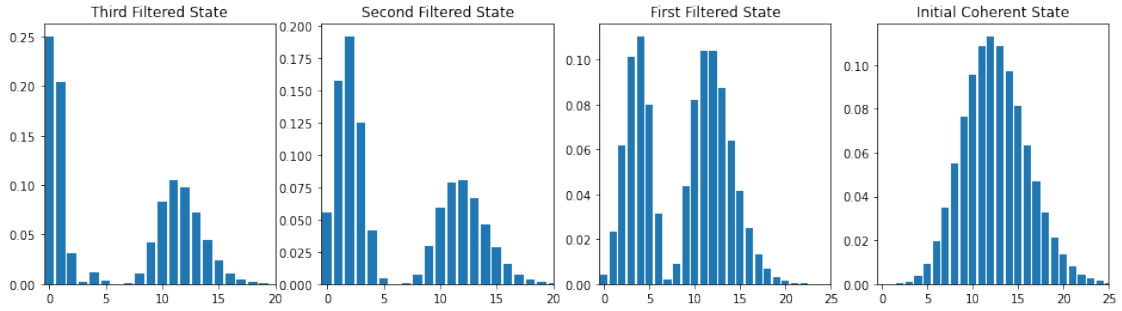



[32]:
```
Dim = 30 ##Dimenstion of the Hilbert spac
eta=1 #detector efficiency -- set eta=1 for ideal case


#parameters from Table D1 of the NJP 2019 paper
refl1=0.637566**2; refl2=0.490869**2; refl3=0.520135**2;
delta=3.53709; beta=2.03846; sq=0.476595; alpha=1.349; n1=5;n2=2;n3=1;

fock1=fock_dm(Dim,1) #input fock state at each state
coher_in=coherent_dm(Dim,delta) #input coherent state

step_one=catalysis(coher_in,fock1,refl1,n1,eta,Dim) #stage one, detect n1␣
 ↪photons on output
step_two=catalysis(step_one,fock1,refl2,n2,eta,Dim); #stage two, detect n2␣
 ↪photons
step_three=catalysis(step_two,fock1,refl3,n3,eta,Dim); #stage three, detect n3␣
 ↪photons

W_step3=state_plot(step_three,5.5,80,20)
```

```
The probability of a sucessful detection is: 0.07142454603455682
BS has reflectivity 0.406490404356  and I am detecting the | 5 > state.
The probability of a sucessful detection is: 0.15531158677247464
BS has reflectivity 0.240952375161  and I am detecting the | 2 > state.
The probability of a sucessful detection is: 0.1619065393887992
BS has reflectivity 0.270540418225  and I am detecting the | 1 > state.
```



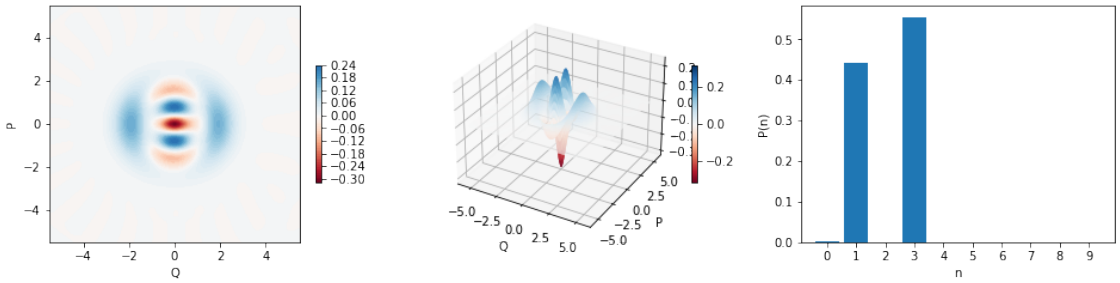

[33]:
```
#plot photon number distributions
fig, axes = plt.subplots(1, 4, figsize=(16,4))
bar0 = axes[0].bar(range(Dim), step_three.diag())
lbl0 = axes[0].set_title("Third Filtered State")
lim0 = axes[0].set_xlim([-.5, 20])
bar1 = axes[1].bar(range(Dim), step_two.diag())
lbl1 = axes[1].set_title("Second Filtered State")
lim1 = axes[1].set_xlim([-.5, 20])
bar2 = axes[2].bar(range(Dim), step_one.diag())
lbl2 = axes[2].set_title("First Filtered State")
lim2 = axes[2].set_xlim([-.5, 25])
bar3 = axes[3].bar(range(Dim), ket2dm(coherent(Dim,delta)).diag())
lbl3 = axes[3].set_title("Initial Coherent State")
lim3 = axes[3].set_xlim([-.5, 25])
plt.show()
```




As a last step, the state must be dispalced back to the origin. Then we can calculate the fidelity of our state with our target squeezed cat state.

[34]:
```
final_state=D(-beta,Dim)*step_three*D(-beta,Dim).dag() #displace final state␣
 ↪back to the origin
print('Fidelity with the target SSV state is: ', fidelity(ssv,final_state)**2)
W_fin=state_plot(final_state,5.5,80,10)
```

```
Fidelity with the target SSV state is:  0.9828867072962356
```


p
Q
0.24
0.18
0.12
0.06
0.00
−0.06
−0.12
−0.18
−0.24
−0.30
Q
P
P(n)
n

# Appendix F

# TES Data Processing

# TES_data_processing

December 22, 2021

## 1 Processing TES Peak Data into Photon Counts

The EFADC saves all measured data as .bin files, and the python script EFADC_decode3.py converts the binary data into human readable .csv files with all of the necessary peak data at each time step. At this point, the user has the freedom about the usage of the data. For the the current work, focus remains on the peak height and peak area to determine photon number counts at each time step. Occassionally, peak time is important to ensure that signal timing align, and to eliminate noise photons that may impinge on the detector at the incorrect time.

When determining what peak paramters to use when determining binning for photon counts, there is some freedom about how the data can be used. Several approaches were attempted for this work, but the analysis focused on those methods that optimized the separation between groupings in a given histogram so as to minimize overlap between potential photon-number binning and thus given better distinction to events. The methods that yield the best results were relative peak height and relative area. The relative peak height is defined as the absoulte recorded peak height minus the pedestal value (the value of the signal voltage exactly at the trigger). This works quite well at differentiating events, but only at low photon numbers (0-4). Depending on TES bias and photon flux, the absolute peak height or relatvie peak height may work better. Relative area is defined as the measured area minus the area of the rectangled defined by the duration of the peak multiplied by TETHi. This value gave the least noise out of all methods considered and allowed for decent peak differentiation all the way out to 30 photons. The relative peak height can be used to verify the efficacy of the area method at low photon number.

[4]:
```
import numpy as np
from qutip import* #need to download
import struct
import os, glob
import scipy.io
import bitarray as b #need to download
import matplotlib.pyplot as plt
import csv
from scipy.special import factorial as fac
from scipy.optimize import curve_fit
from scipy.signal import find_peaks
from matplotlib.ticker import (MultipleLocator, AutoMinorLocator)
from numpy import array
def flatten(t):
    #use to flatten lists
```

```
    return [item for sublist in t for item in sublist]
```

First, the .csv files must be loaded in. In this example, there is one data file for every 250,000 triggers at each channel where data was recorded. Thus, 'data4' has 2 channels (Ch1-Ch2) and 2,500,000 triggers (10 subfiles). The data is stored in outdat and indexed as outdat[i][j][k], where i is the file number, j is the channel number, and k is the type of data indexed by header, which is printed out below. As an example, outdat[0][1][4] gives an array of all the peak times in order for channel 2 for the first data files, data4_0Ch1.csv and data4_0Ch2.csv.

[3]:
```
filenames=[]
numfiles=10
basename='/data4_'
for i in range(numfiles):
    filenames.append(os.getcwd()+basename+str(i))

outdat=[]
for k in range(numfiles):
    chdata=[]
    for i in range(8):
        fpath=str(filenames[k])+'Ch'+str(i+1)+'.csv'
        if os.path.exists(fpath):
            file=open(fpath)
            datafile=csv.reader(file)
            header=next(datafile)
            rows=[]
            for row in datafile:
                rows.append(row)
            chdata.append(np.asarray(rows,dtype=float).T)
            file.close()

    outdat.append(chdata)


print('Data ordering: ', header)
```

```
Data ordering:  ['Trigger time', 'Channel number', 'Area', 'Peak Height', 'Peak
Time', 'End Time', 'Start Time', 'Pedestal', 'overflow', 'underflow']
```

The below section takes the outdat variable above and reorganizes it into separate variables for each channel. The values of TEThi for each channel should be included, as these are used to aid in the cacluation of relative peak area.

Additionally if one is using shutters and on/off locking, there is the option to remove pulses on each channel surrounding the shutter. This is necessary because there are several miliseconds of delay when the shutters open and/or close where signal may be different between the TES channels due to variations in the mechanical shutters and slight timing mismatches. Discarding data from about 225 of the pulses before the shutter closes and after it opens ensures the rest of the data is consistent. Also, the data at the edges has different TES behavior as it transitions to no photon flux to very high photon flux, meaning the peak information for the first few events after the shutter

opens is unreliable.

The code to do this looks for how long it has been since an event occurred, and if it is too long (greater than max_timesep number of triggers), then the code assumes a shutter had been closed, and it removes the trim_fr events from each channel before the shutter and trim_bk events after the shutter. Additionally, one should inclue the approximate separation in terms of trigger clock as pulse_sep. Note that the EFADC samples one data point every 4 ns.

***Note:*** Even if the shutters are on, this removing of peaks should *__only__ be used when calibrating a coherent state or calculating the histograms used for binning. When heralding, the timestamps of the single photons allow for determining which events to keep.

[20]:
```
TEThi0=500; TEThi1=350; TEThi2=250; #TET settings in the runinit file for each␣
↪channel

#coherent state calibration variables
ch0area=[]; ch0pheight=[]; ch0Rarea=[]; ch0Rpheight=[]; ch0trigtime=[];␣
↪ch0ptime=[];
ch1area=[]; ch1pheight=[]; ch1Rarea=[]; ch1Rpheight=[]; ch1trigtime=[];␣
↪ch1ptime=[];
ch2area=[]; ch2pheight=[]; ch2Rarea=[]; ch2Rpheight=[]; ch2trigtime=[];␣
↪ch2ptime=[];

#pulse_sep=20050 #clock timestamp difference between pulses. Here, 20050␣
↪corresponds to 20050 x 4ns = 80.2 microseconds
#max_timesep=400 #max number of triggers allowed between events before assuming␣
↪it was from a shutter being closed
#trim_fr=225 # amount of triggers to remove before shutter opens
#trim_bk=225 # amount of triggers to remove after shutter closes

for i in range(numfiles):
        for k in range(2):
            trigtime=outdat[i][k][0]
            sort_times=np.argsort(trigtime) #sorts timestamps correctly
            good_idx=np.arange(0,len(sort_times),1)

            '''#This section finds breaks in the data from shutters closing and
            #calculates a "good index" array that is for the timestamps I want
            #to keep. I can then use this for all data

            scal_dif=np.diff(trigtime[sort_times])/pulse_sep
            breaks=np.where(scal_dif>max_timesep)[0]
            size=len(scal_dif)+1
            if len(breaks>0):
                if breaks[0]==0:
                    good_idx=np.arange(breaks[0]+trim_fr,breaks[1]-trim_bk,1)
                    offset=1
                else:
```

```
                    good_idx=np.arange(0,breaks[0]-trim_bk,1)
                    offset=0
                for m in range(len(breaks)-offset-1):
                    good_idx=np.append(good_idx,np.
↪arange(breaks[m+offset]+trim_fr,breaks[m+1+offset]-trim_bk,1))
                good_idx=np.append(good_idx,np.arange(breaks[-1],size,1))'''

            height=((outdat[i][k][3])[sort_times])[good_idx]
            area=((outdat[i][k][2])[sort_times])[good_idx]
            eval('ch'+str(k)+'area').append(area) #channel areas
            #eval('ch'+str(k)+'Rarea').
↪append(area-(((outdat[i][k][5]-outdat[i][k][6])
            #                                          ␣
↪*eval('TETlow'+str(k)))[sort_times])[good_idx])

            eval('ch'+str(k)+'Rarea').
↪append(area-(((outdat[i][k][5]-outdat[i][k][6])
                                                       ␣
↪*eval('TEThi'+str(k)))[sort_times])[good_idx])
            #Alternative method for determining relative areas using the␣
↪pedestal values. This does not seem to work as well.
            #eval('ch'+str(k)+'Rarea').
↪append(area-(((outdat[i][k][5]-outdat[i][k][6])
            #                                          ␣
↪*outdat[i][k][7])[sort_times])[good_idx])


            eval('ch'+str(k)+'pheight').append(height) #channel peak heights
            eval('ch'+str(k)+'Rpheight').
↪append(height-((outdat[i][k][7])[sort_times])[good_idx]) #channel relative␣
↪peak heights
            eval('ch'+str(k)+'ptime').
↪append(((outdat[i][k][4])[sort_times])[good_idx]) #channel peak times
            eval('ch'+str(k)+'trigtime').
↪append((trigtime[sort_times])[good_idx]) #channel trigger times

all_areas=[np.array(flatten(ch0area)),np.array(flatten(ch1area)),np.
↪array(flatten(ch2area))] #raw areas
all_Rareas=[np.array(flatten(ch0Rarea)),np.array(flatten(ch1Rarea)),np.
↪array(flatten(ch2Rarea))] #Relative area, right now subtracts a rectangle␣
↪from start time to end time at TEThi
all_pheight=[np.array(flatten(ch0pheight)),np.array(flatten(ch1pheight)),np.
↪array(flatten(ch2pheight))] #raw peak height
all_Rpheight=[np.array(flatten(ch0Rpheight)),np.array(flatten(ch1Rpheight)),np.
↪array(flatten(ch2Rpheight))] #Rpheight = relative peak height, takes peak␣
↪height and subtracts the pedestal
```

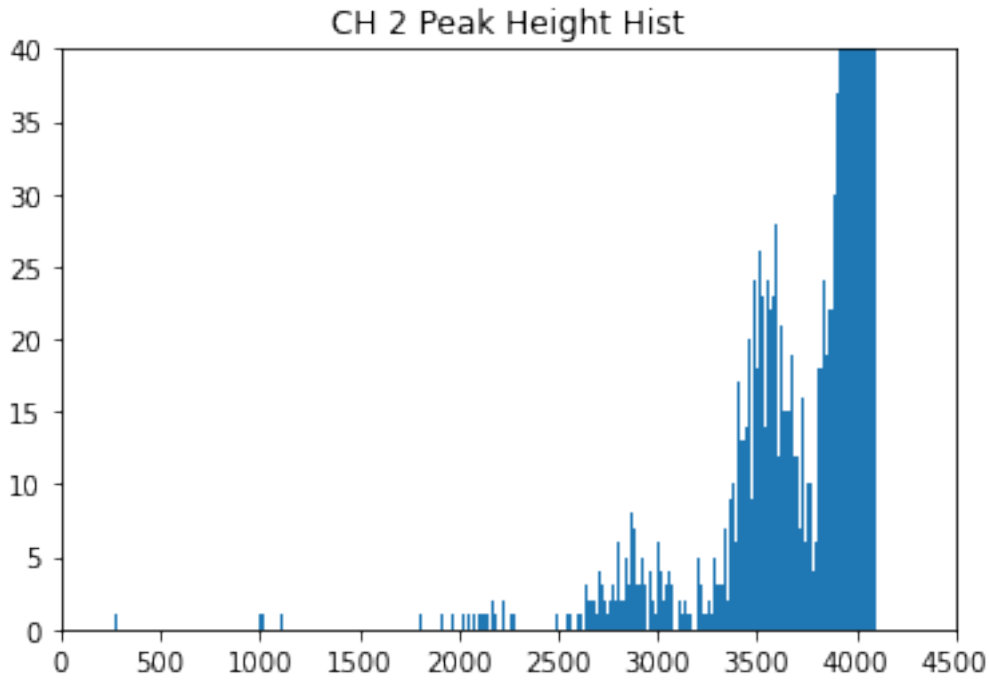
```
#all_RhRa=[np.array(flatten(ch0RhRa)),np.array(flatten(ch1RhRa)),np.
 ↪array(flatten(ch2RhRa))]
all_ptime=[np.array(flatten(ch0ptime)),np.array(flatten(ch1ptime)),np.
 ↪array(flatten(ch2ptime))]  #peak times
all_trigtime=[np.array(flatten(ch0trigtime)),np.array(flatten(ch1trigtime)),np.
 ↪array(flatten(ch2trigtime))] #all trigger times

print('The number of recorded area events on Ch1 is given by:',␣
 ↪len(all_areas[0]))
```

```
The number of recorded area events on Ch1 is given by: 2500000
```

The above variables are indexed based on channel number, e.g., all_areas[0] gives the areas for all files for Channel 1.

For the rest of the file, let's focus on the data from Channel 2. First, let's plot a histogram of the peak heights.

[65]:
```
histvals=np.hstack(all_pheight[1])
plt.hist(histvals,bins=1000);
plt.xlim([0,4500])
plt.ylim([0, 40])
plt.title('CH 2 Peak Height Hist')
plt.show()

twophot_idx=(all_pheight[1]<2400)*(all_pheight[1]>1600)
print('The proportion of two photon peaks is about ',␣
 ↪len(all_pheight[1][twophot_idx])/len(all_pheight[1]))
```

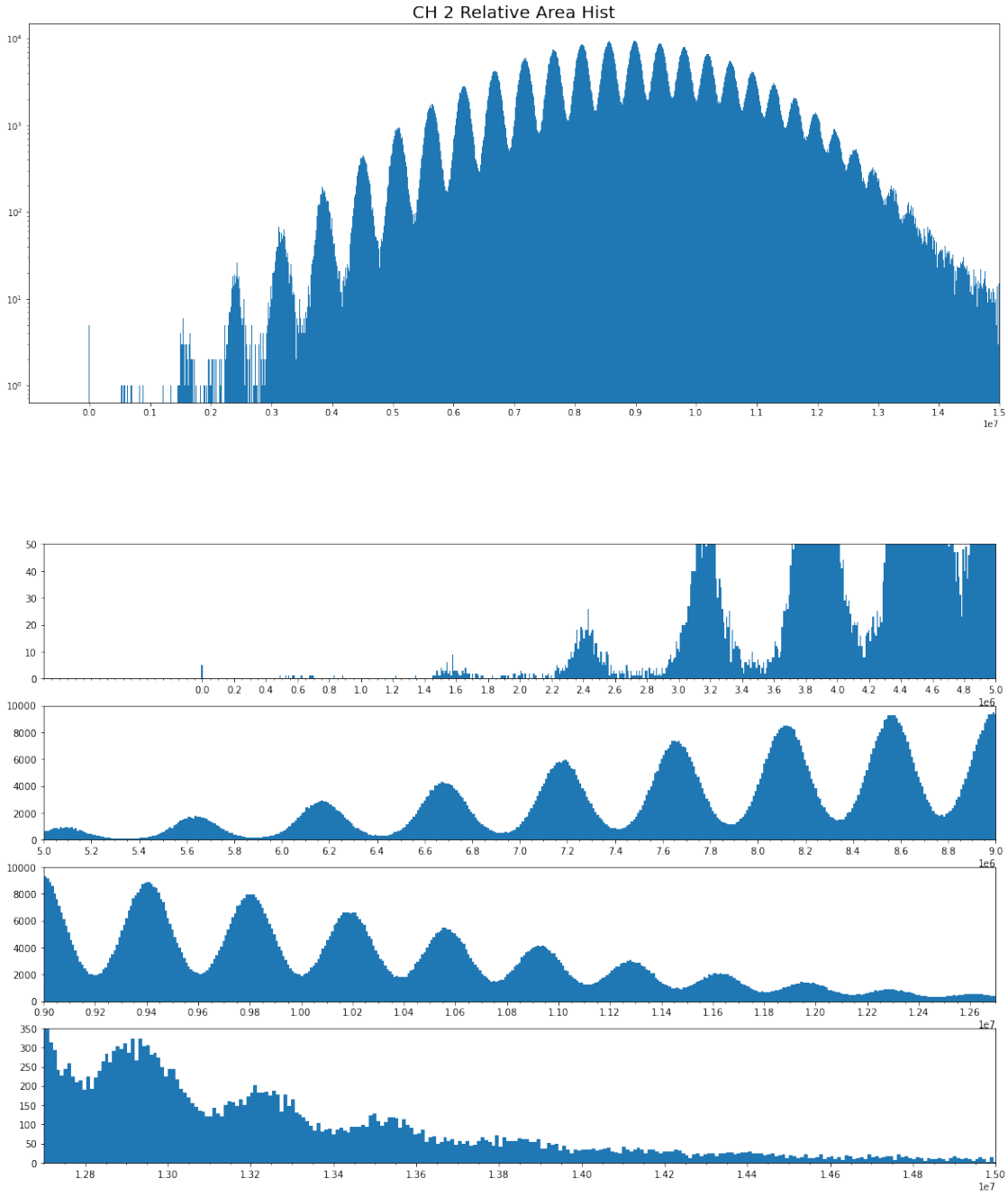


```
The proportion of two photon peaks is about  3.4e-05
```

In this example, the photon flux is very high (mean photon number ~14), so there is a low chance of having low photon numbers. With the bias settings here, the peak height saturates quickly around the 5 photon mark, thus only 0-4 can be resolved with the height method. However, one can clearly see the distinction between the 0 photon events (less that 500 height), the 1 photon events (heights 500-1500), the 2 photon events (heights 1500-2400), and 3 photon events (heights 2450-3200). These values can be used as a self-consistency method to check binning obtained from another method, such as areas.

Below, I show the relative area histogram for the data. The first is log scale and the second is regularly scaled, which is spread out over 4 separate graphs with continuation in area to account for scales. We can now compare the number of two photon peaks from this method and see if it agrees with the previous peak height methodology.

[482]:
```
histvals=np.hstack(all_Rareas[1])
fig = plt.figure(figsize=(20, 8))
ax = fig.add_subplot(1, 1, 1)
ax.set_yscale('log')
ax.hist(histvals,bins=2500);
ax.set_xticks(np.linspace(0,2e7,21))
ax.set_xlim([-1e6,15e6])


plt.title('CH 2 Relative Area Hist',fontsize=20)
plt.show()

xlims=[-1e6,5e6,9e6,12.7e6,15e6]
ylims=[50,10000,10000,350]
fig, ax = plt.subplots(4, sharey='row',figsize=(18, 12))
for k in range(4):
    ax[k].hist(histvals,bins=2500);
    ax[k].set_xticks(np.linspace(0,2e7,101))
    ax[k].xaxis.set_minor_locator(MultipleLocator(5e4))
    ax[k].set_xlim([xlims[k],xlims[k+1]])
    ax[k].set_ylim([0,ylims[k]])

plt.show()


twophot_idx=(all_Rareas[1]<2e6)*(all_Rareas[1]>1.2e6)
print('The proportion of two photon peaks is about ',␣
 ↪len(all_Rareas[1][twophot_idx])/len(all_Rareas[1]))
```




```
The proportion of two photon peaks is about  3.64e-05
```

From the two different methods, the two photon contributions agree quite well, so this indicates our method is working.

One can now determine threshold markers from the minimum between each of the histogram peaks. These values can be put in 'by hand' for quick results, or one can fit each peak to a Gaussian distribution and find the interesection with neighboring peaks to determine the binning. This is demonstrated at the end of this code, but surprisingly, the 'by hand' results are quite robust and

nearly as accurate as the results obtained by numerically performing the fits. The results of the binning can also be verified by checking the sum-squared error of the measured distribution with that of an expected coherent state. For now, let's just put in the values by hand.

***Note:*** These binning thresholds must be verified and changed for each and every data set. This is because these values are dependent upon the specific TES detector used, TES bias settings, TES temperature, EFADC settings, and input coherent state power. As a side point, the dependence on the coherent state power has more to do with the time between pulses and the electronic amplification and High-pass filtering after the TES measures the events.

[139]:
```
ch2Rathresh=[0.4e6,1.2e6,2e6,2.8e6,3.5e6,4.15e6,4.8e6,5.35e6,5.9e6,6.4e6,6.
↪925e6,7.4e6,7.9e6,8.35e6,8.775e6,9.2e6,9.6e6,1.00e7,1.0375e7,1.075e7,1.
↪1115e7,1.145e7,1.18e7,1.215e7,1.2475e7,1.28e7,1.31e7,1.34e7,1.37e7,1.4e7,1.
↪425e7,1.46e7]

'''The below threshold are from Gaussian fits, the above are 'by eye'. Both␣
↪methods are surprisingly similar'''
#ch2Rathresh=[0.4e6, 1109850.,  1976700.,  2750850.,  3497850.,  4163400., ␣
↪4775100.,
#        5339250.,  5884950.,  6408000.,  6915300.,  7402800.,  7874250.,
#        8328450.,  8765550.,  9187050.,  9591750.,  9994650., 10371750.,
#        10744050., 11103900., 11455050., 11796300., 12137400., 12453150.,
#        12781200., 13087050., 13410300., 13695300., 13997850., 14289150.,1.
↪455e7]
```

The next section accounts for the events that ***aren't*** recorded by the EFADC at all. This occurs when the signal never crosses TEThi, which means we should count it as a 0 photon event. However, it is still important to keep track of these events, as simply discarding them will alter the statistics. In this case, TEThi was set slightly into the noise floor, thus there were no events that weren't caught, and all zero photon events were seen directly. However, the next section will be necessary if this isn't the case.

[480]:
```
pulse_sep=22500
zeroidx_insertch1=[]
zeroidx_insertch2=[]
new_areas=[]
new_trigtime=[]
for k in range(2):
    skips=np.where(np.diff(all_trigtime[k+1])>pulse_sep+1000) #skips should be␣
↪greater than pulse_sep but less than 2 X pulse_sep
    c=0
    for i in range(len(skips[0])-1):
        #The below if statement is necessary to remove events that are␣
↪separated by exactly 1000. I am unsure why this happens, but it seems like␣
↪there is some systematic fictitious skipping that occurs, either due to␣
↪pulse-generator reset or extra data timestamps by the DAC. Either way, these␣
↪events are not real.
        if (skips[0][i+1]-skips[0][i])!=1000:
```

```
            c+=1
            print(skips[0][i])
            if np.mod(c,2)==1 and c!=1:
                eval('zeroidx_insertch'+str(k+1)).append(skips[0][i]+1)
    #print(eval('zeroidx_insertch'+str(k+1)))
    new_areas.append(np.insert(all_Rareas[k],eval('zeroidx_insertch'+str(k+1)),
                               np.
 ↪zeros(len(eval('zeroidx_insertch'+str(k+1))))))
    new_trigtime.append(np.
 ↪insert(all_trigtime[k],eval('zeroidx_insertch'+str(k+1)),
                                    ␣
 ↪(all_trigtime[k][eval('zeroidx_insertch'+str(k+1))]-22500)))

print('size of new areas, including zero events. This should the same as the␣
 ↪number of actual triggers sent', len(new_areas[1]))
```

```
size of new areas, including zero events. This should the same as the number of
actual triggers sent 2500000
```

The code segment below now uses the thresholds for binning defined above, along with the new areas arrays and determines the probability of a photon to be in each bin. For this example, I make the peak thresholds for Ch1 the same as for Ch2 (so the code could still be left general for multiple channels). Note that this will not give the correct results for channel 1, as the binning must be done independently for each channel. Additionally, I removed all events with areas larger than the highest threshold, as these have indeterminate photon number, although certainly equal to or larger than the highest photon number resolvable. Depending on the instance, it may be useful to keep track of these separately. For the current illustration, I just ignore them, which can safely be done as they account for less that above 0.3% of events.

While calculating the probabilities, the below code also calculates a bit string based on the parity of the measured event (0 for even and 1 for odd), in addition to a string containing the results of the measured photon number modulo 4. These can both be used for unbiased random number generation as discussed further in the main document.

[141]:
```
ch1Rathresh=np.array(ch2Rathresh)
numtrigs=len(new_areas[1])

coher_probs1=[];
coher_probs2=[];

for k in range(2):
    #pthresh=eval('ch'+str(k)+'pthresh')
    athresh=eval('ch'+str(k+1)+'Rathresh')
    start_A=new_areas[k]
    bad_idx=start_A>athresh[-1] #gets rid of all events that can't be resolved:␣
 ↪those with areas larger than the highest threshold
    tmpA=start_A[~bad_idx]
    numevents=len(tmpA)
```

```
    vars()['bitstring'+str(k+1)]=np.zeros(numevents);
    vars()['mod4_bitstring'+str(k+1)]=np.zeros(numevents);
    #numevents=len(tmpA)
    for i in range(len(athresh)+1):
        if i==0:
            idx=tmpA<athresh[i]
            #prob=(len(tmpA[idx])+numtrigs_approx-len(tmpA))/numevents
            prob=len(tmpA[idx])/numevents
            eval('bitstring'+str(k+1))[idx]='00'
        elif i==len(athresh):
            #idx=tmpA>athresh[i-1]
            #prob=len(tmpA[idx])/numevents
            prob=0
        else:
            lowlim=athresh[i-1]
            uplim=athresh[i]
            idx=(lowlim<tmpA)*(tmpA<uplim)
            prob=len(tmpA[idx])/numevents
            if np.mod(i,2)==1:
                eval('bitstring'+str(k+1))[idx]=1

            if np.mod(i,4)==0:
                eval('mod4_bitstring'+str(k+1))[idx]=0
            elif np.mod(i,4)==1:
                eval('mod4_bitstring'+str(k+1))[idx]=1
            elif np.mod(i,4)==2:
                eval('mod4_bitstring'+str(k+1))[idx]=2
            else:
                eval('mod4_bitstring'+str(k+1))[idx]=3

        eval('coher_probs'+str(k+1)).append(prob)

print('The numbe of events within the calibrated thresholds for Channel 2 is:␣
 ↪', numevents)

print('probability of odd on Ch2: ', np.dot(np.
 ↪mod(range(len(coher_probs2)),2),coher_probs2))

'''CH2 probs of 0-3'''
print('probability of mod4=0: ', len(mod4_bitstring2[mod4_bitstring2==0])/
 ↪len(mod4_bitstring2))
print('probability of mod4=1: ', len(mod4_bitstring2[mod4_bitstring2==1])/
 ↪len(mod4_bitstring2))
print('probability of mod4=2: ', len(mod4_bitstring2[mod4_bitstring2==2])/
 ↪len(mod4_bitstring2))
print('probability of mod4=3: ', len(mod4_bitstring2[mod4_bitstring2==3])/
 ↪len(mod4_bitstring2))
```

```
The numbe of events within the calibrated thresholds for Channel 2 is:  2492860
probability of odd on Ch2:  0.5014854424235617
probability of mod4=0:  0.24905249392264306
probability of mod4=1:  0.2506779361857465
probability of mod4=2:  0.24946206365379525
probability of mod4=3:  0.2508075062378152
```

The below code can be used to save data. It if the .open() command on a file is used, it is ***always*** necessary to have the corresponding .close()

[142]:
```
#bitstringdata=bitstring1.astype(int)
#file=open(os.getcwd()+basename+"bitarray_testdata"+".bin","wb")
#num=bytearray(list(bitstringdata))
#file.write(num)
#file.close()
```

Now, let's examine how well the binned data corresponds to a coherent state. After plotting the measured distribution, we can check the sum-squared error (SSE) with the expected distribution. Becase we are sending a coherent state,

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}},$$

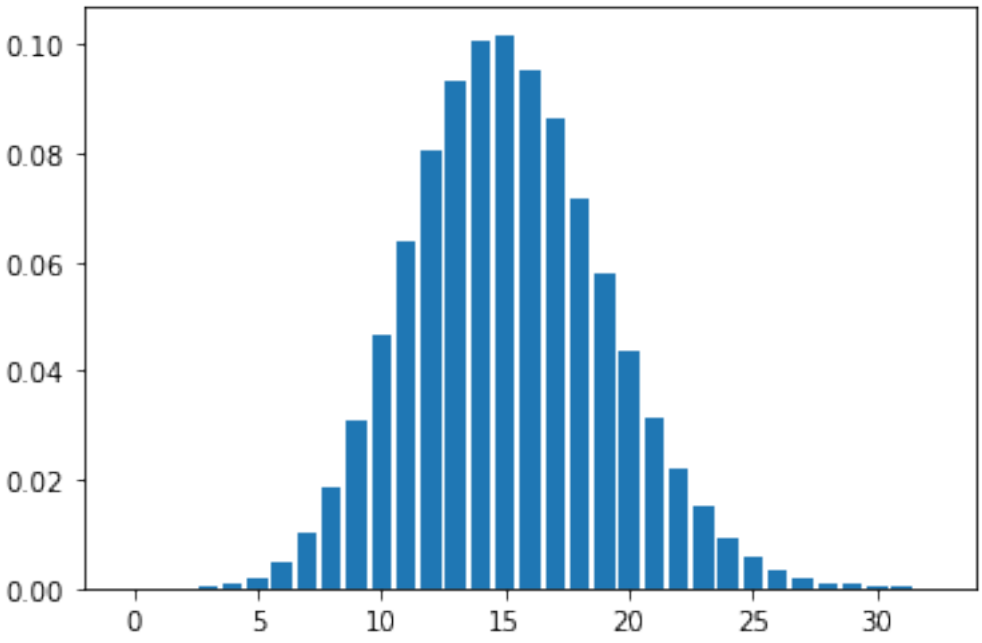
we see that the expected photon number distribution should follow the Poissonian distribution of

$$P(n) = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}$$

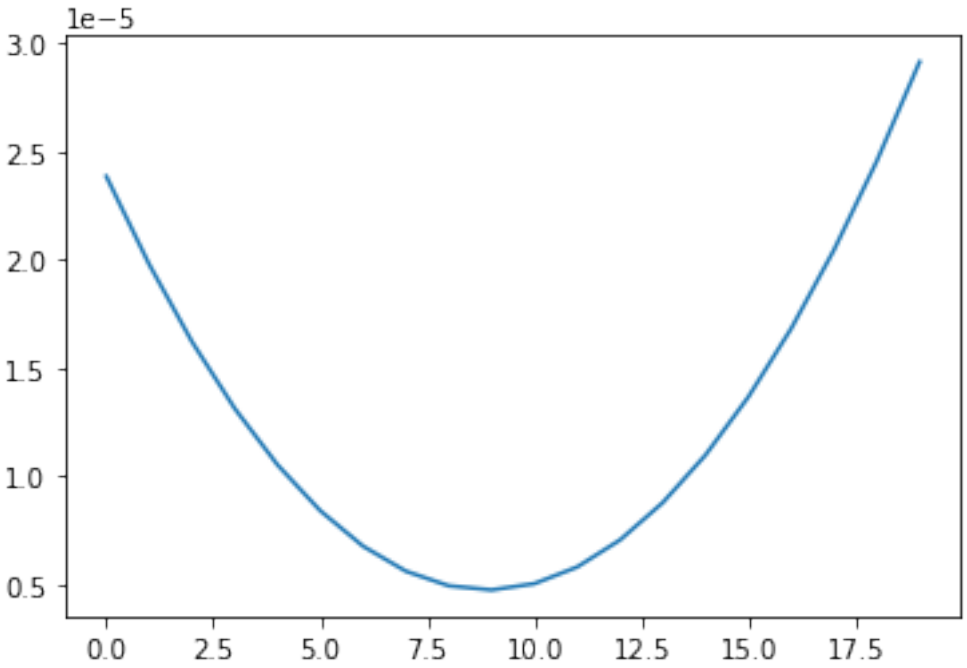
where $|\alpha|^2$ is the mean photon number. However, if there is some error, we can assume that perhaps the measured mean photon number is not precisely the same as the photon number of the best fit coherent state (due to amplitude noise or binning error). Instead, we can simply plot the SSE as a function of $\alpha$ and find the minium value.

[491]:
```
plt.bar(range(len(coher_probs2)),coher_probs2)
plt.show()
meas_meanch2=np.dot(range(len(coher_probs2)),coher_probs2)
print('The measured mean photon number is given by: ', meas_meanch2)
print('The photon number probabilities for this coherent state are: ',␣
 ↪coher_probs2)
#Try fitting to a coherent state
meas_dist=np.asarray(coher_probs2)
sse=[]
N_dim=40
for k in range(20): #set to determine number of loop iterations
    #mu=np.sqrt(24.7+0.01*k)
    mu=np.sqrt(meas_meanch2-0.1+0.01*k) #coherent state amplitude
    test=coherent_dm(N_dim,mu) #theory coherent state
    sse.append(np.sum(np.abs((meas_dist[0:30]/(np.sum(meas_dist[0:30])))
                             -(test.diag()[0:30])/(np.sum(test.diag()[0:
 ↪30])))**2))
```

```
plt.plot(sse)
plt.show()
print(np.min(sse)) #try to minimize sse
```




```
The measured mean photon number is given by:  15.14062923710116
The photon number probabilities for this coherent state are:
[2.005728360196722e-06, 4.412602392432788e-06, 3.610311048354099e-05,
0.00018252128077790168, 0.0006622915045369575, 0.0019194820407082628,
0.004695410091220526, 0.00998010317466685, 0.018671325305071283,
0.030967242444421266, 0.04676516130067473, 0.06395425334756064,
0.08044896223614643, 0.09347536564428006, 0.10081512800558395,
0.10178429594923101, 0.0954995467053906, 0.0862756031225179,
0.07154232487985687, 0.05773769886796691, 0.04362338839726258,
0.031581797613985545, 0.021903355984692282, 0.01491058462970243,
0.009027382203573405, 0.005775695386022 4804, 0.003313864396717024,
0.0020101409625891548, 0.0011155861139414166, 0.0006783373314185313,
0.0003907158845663214, 0.0002479080253203148, 0]
```



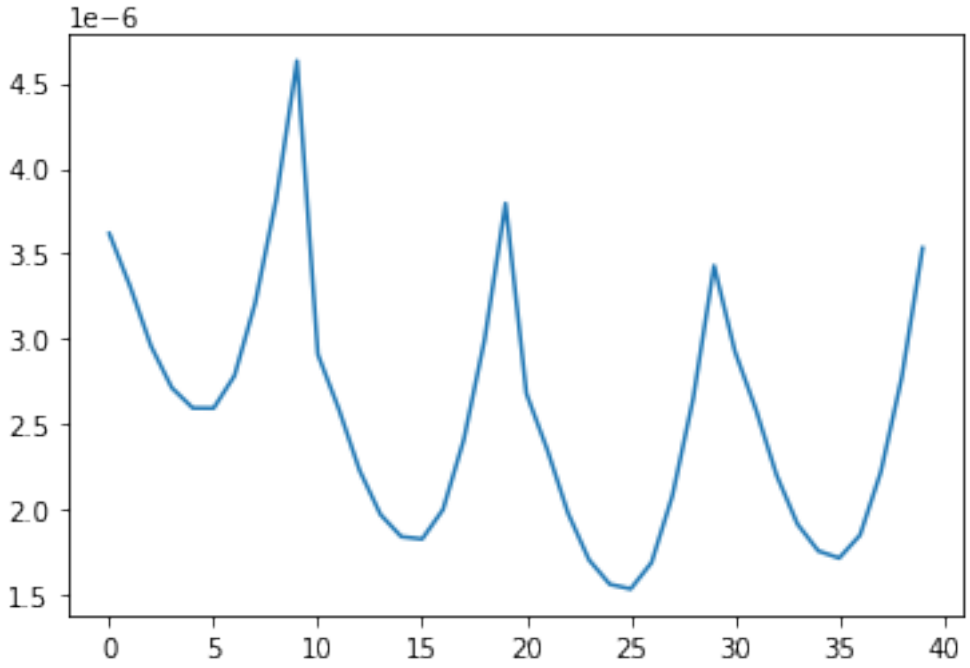

```
4.761659602312153e-06
```

The second plot from the above codes plots the SSE vs the index of the loop, which has a minimum value on about the 8th iteration. This corresponds to an alpha value of 15.11, which is very near the calculated mean photon number. Furthermore, the SSE is quite low, indicating that the binning was performed properly. Without checking the SSE fit with a true coherent state, it may be difficult to ascertain that the binning is in fact correct due to low sampling at some photon numbers. For example, in the histogram of areas shown several cells above, there are very few events with near zero area (below 0.2e6). One might be tempted to erroneously group all events with relative areas less than 1.0e6 as a single bin to call all of them zero photons. However, one can easily check that this would be incorrect, as re-running the above few cells by doing this (after removing the first value in ch2Rathresholds) would give a distribution that qualitatively looks like a coherent state, but the SSE will in fact be about an order of magnitude worse. Thus, the true distribtuion binning follows the thresholds used above.

Once the bin fitting is assured, the same process with SSE can be performed with amplitude noise or thermal noise to determine the quality of the measured coherent state. While the above test was more to determine if the data processing was performed correctly, the below is more to diagnose the measured state.

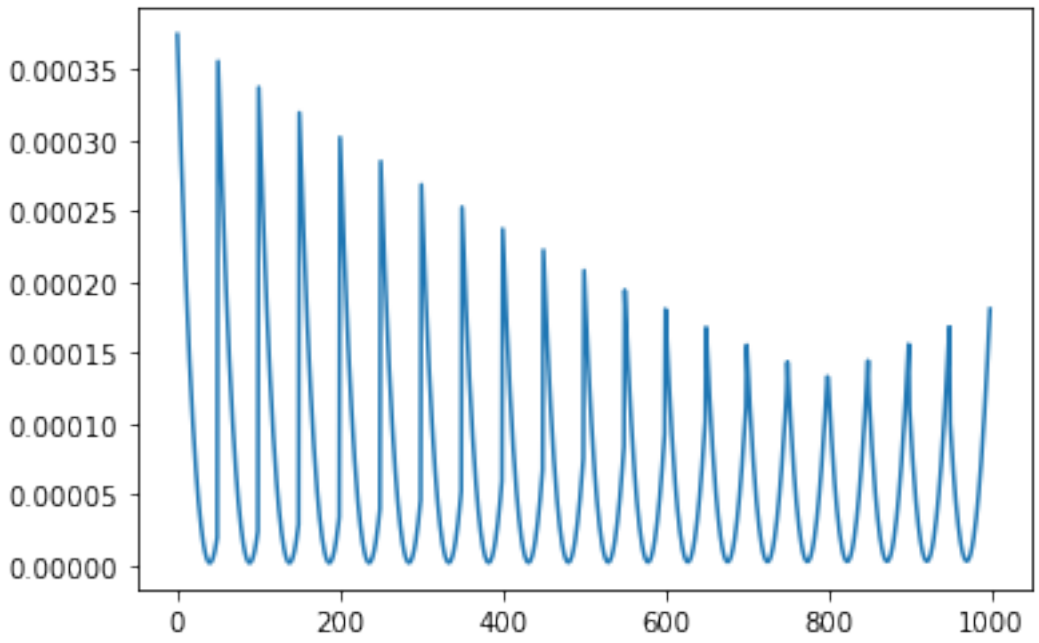
[195]:
```
#Try a mixture of gaussian distributed coherent states

meas_dist=np.asarray(coher_probs2)
sse=[]
for k in range(4):
    #mu=np.sqrt(24.7+0.01*k)
```

```
    mu=np.sqrt(meas_meanch2-0.03+0.01*k)
    for j in range(10):
        sigma=0.04+0.005*j
        test=0
        dx=0.01
        steps=6*round(sigma/(dx))
        xvec=np.linspace(mu-steps*dx/2,mu+steps*dx/2,steps)
        for x in xvec:
            test=test+np.e**(-(mu-x)**2/(2*sigma**2))*coherent_dm(N_dim,x)
        test=test/test.tr()
        sse.append(np.sum(np.abs((meas_dist[0:30]/(np.sum(meas_dist[0:30])))
                          -(test.diag()[0:30])/(np.sum(test.diag()[0:30])))**2))
plt.plot(sse)
plt.show()
print(np.min(sse))
minloc=np.where(sse==np.min(sse))[0]
print('Optimal parameters occur when the state is a mixture of coherent states␣
 ↪with a central mean photon number of ', meas_meanch2-0.03+0.01*np.
 ↪floor(minloc/10), ' and is mixed with other coherent states having a␣
 ↪Gaussian standard deviation of sigma=', 0.04+0.005*np.mod(minloc,10))
```




```
1.5291066076581972e-06
Optimal parameters occur when the state is a mixture of coherent states with a
central mean photon number of  [15.13062924]  and is mixed with other coherent
states having a Gaussian standard deviation of sigma= [0.065]
```

[196]:
```
#Try a mixture of a coherent state convolved with a thermal state
meas_dist=np.asarray(coher_probs2)
sse=[]
for k in range(20):
    #mu=np.sqrt(24.7+0.01*k)
    mu=np.sqrt(meas_meanch2-0.5+0.01*k)
    coher=coherent_dm(N_dim,mu).diag()
    for j in range(50):
        thrmamp=0.1+0.01*j
        dist=np.convolve(coher,thermal_dm(N_dim,thrmamp).diag())
        sse.append(np.sum(np.abs((meas_dist[0:30]/(np.sum(meas_dist[0:30])))
                          -(dist[0:30])/(np.sum(dist[0:30])))**2))
plt.plot(sse)
#plt.ylim([0.01,0.02])
plt.show()
print(np.min(sse))
minloc=np.where(sse==np.min(sse))[0]
print('Optimal parameters occur when the coherent state has mean photon number␣
↪', meas_meanch2-0.5+0.01*np.floor(minloc/50), ' and is convolved with␣
↪thermal noise with mean photon number, ', 0.1+0.01*np.mod(minloc,50))
```




```
1.4896181285011333e-06
Optimal parameters occur when the coherent state has mean photon number
[14.64062924]  and is convolved with thermal noise with mean photon number,
[0.5]
```

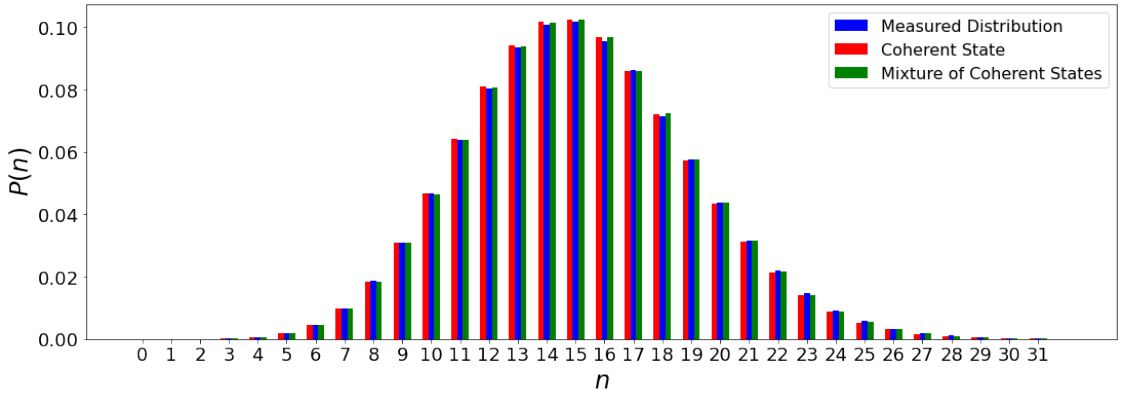
Both of the above fits (extra thermal noise vs amplitude noise) have roughly the same SSE so would

both be good model fits. However, considering that the fits are about the same, we can consider the fact the the TES has almost negligible dark count, and without any external light, the chances of having thermal noise this large is quite small. Thus, we can conclude that the likely explanation is in fact amplitude noise, either on the laser or due to fiber effects. This is shown in the direct comparison plot below.

[202]:
```
N_dim=40
coher_appx=coherent_dm(N_dim,np.sqrt(15.11))
#sqcoher_appx=ket2dm(squeeze(N_dim,-0.0692)*coherent(N_dim,np.sqrt(5.579)))

mu=np.sqrt(15.13)
sigma=0.065
mix=0
dx=0.01
steps=6*round(sigma/(dx))
xvec=np.linspace(mu-steps*dx/2,mu+steps*dx/2,steps)
xvec=np.linspace(mu-sigma/2,mu+sigma/2,2*round(sigma/(dx)))
for x in xvec:
    mix=mix+np.e**(-(mu-x)**2/(2*sigma**2))*coherent_dm(N_dim,x)

#plt.subplots(figsize=(10, 5))
plt.figure(figsize=(18,6))
ax = plt.subplot(111)
x=range(N_dim)
#ax.bar(np.array(range(N_dim))-0.2, (test/test.tr()).diag(), width=0.2,␣
 ↪color='b', align='center')
ax.bar(np.array(range(len(coher_probs2))),coher_probs2, width=0.2, color='b',␣
 ↪align='center',label='Measured Distribution')
ax.bar(np.array(range(N_dim)[0:32])-0.2, coher_appx.diag()[0:32], width=0.2,␣
 ↪color='r', align='center',label='Coherent State')
ax.bar(np.array(range(N_dim)[0:32])+0.2,(mix/mix.tr()).diag()[0:32], width=0.2,␣
 ↪color='g', align='center', label='Mixture of Coherent States')
#ax.bar(np.array(range(N_dim)[0:30])+0.2,dist[0:30], width=0.2, color='g',␣
 ↪align='center', label='Mix with Thermal')
ax.legend(fontsize=16)
ax.set_xticks(np.array(range(32)))
ax.tick_params(axis='both',labelsize=18)
ax.set_xlabel('$n$',fontsize=24)
ax.set_ylabel('$P(n)$',fontsize=24)
#ax.bar(np.array(range(len(coher_probs2)))+0.3,coher_probs2, width=0.4,␣
 ↪color='r', align='center')
#print('Squeezing in db: ',20*np.log10(np.e**-0.0692))
#print('Fidelity with coherent state: ', fidelity(coher_appx,meas_coher)**2)
#print('Fidelity with squeezed coherent state: ',␣
 ↪fidelity(sqcoher_appx,meas_coher)**2)
print('Mean photon number of ideal coherent state: ', 15.11)
```

```
print('Spread in mean photon number: ',(mu-sigma/2)**2, ' to ',(mu+sigma/2)**2)

plt.show()
```

```
Mean photon number of ideal coherent state:  15.11
Spread in mean photon number:  14.878223795595902  to  15.3838887044041
```




## 2 Gaussian Binning

The next set of code will show how to systematically determine the photon binning by taking the area histograms and fitting a Gaussian function to each peak, and setting the bin threshold at the intersection point between the two Gaussians. This method is the most accurate way to bin the histograms, although it surprisingly differs little from binning manually.

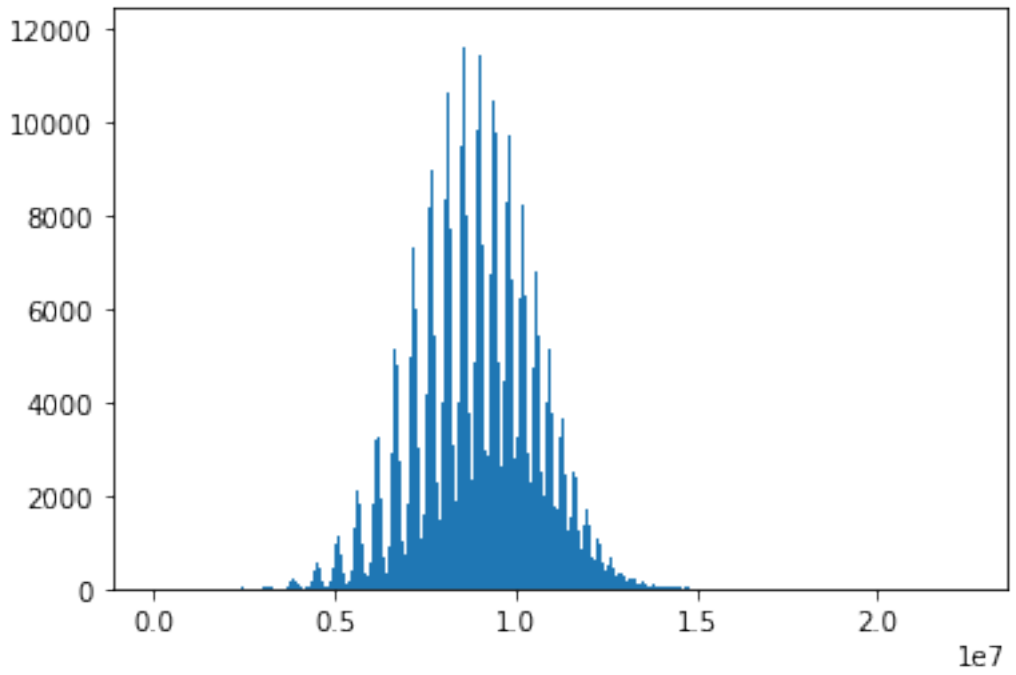
First, define two functions that will compute and return an array that is a Gaussian sum. Both of these functions take in an array of Gaussian parameters grouped in 3-tuples and an array-like x vector. The Gaussian parameters are input such that in each triple, the first parameter is the center of the Gaussian, the second parameter is the amplitude, and the third parameter is the standard deviation.

[203]:
```
def gauss_sum(params,x):
    f = np.zeros_like(x)
    for k in range(round(len(params)/3)):
        ctr = params[3*k]
        amp = params[3*k+1]
        wid = params[3*k+2]
        f = f + amp * np.exp( -((x - ctr)/wid)**2/2)
    return f

def gauss_sum_normed(params,x):
    f = np.zeros_like(x)
    for k in range(round(len(params)/3)):
        ctr = params[3*k]
```

```
        amp = params[3*k+1]
        wid = params[3*k+2]
        f = f + np.exp( -((x - ctr)/wid)**2/2)/np.sqrt(2*np.pi*wid**2)
    return f
```

Take the the histogram used earlier, but now turn it into a function with defined bin heights and bin centers. Note that the number of bins determines the size of the xvector as used later.

[213]:
```
histvals=np.hstack(all_Rareas[1])
bin_heights, bin_borders, _ = plt.hist(histvals, bins=2000, label='histogram')
bin_centers = bin_borders[:-1] + np.diff(bin_borders) / 2
```




Now, define a method that will fit an input function to a sum of Gaussian functions. This is achieved by the below function, which takes several parameters. The first two inputs are the two arrays that define the function to be fit (the bin_centers and bin_heights). The next parameters, pkprom, determines the minimum peak prominence the peakfinder should use when assigning Gaussian peaks to the data. The start and stop are xvector values that are used to determine which subset of the data is being fit, while num_Gauss tells the function how many Gaussians should be used to fit the data. Widthguess is the starting peak standard deviation in terms of xvector indices use for the guess for the optimizer. Finally, maxiters limits how many cycles the optimizer can run through.

Because the histogram contains peaks on vastly different scales, it is impossible with the current method to catch each and every peak all at once. Thus some cross between a manual and automated fitting process is need to catch all possible peaks. The histogram can be sectioned into groupings of peaks with same order-of-magnitude amplitude, and each of these peaks can be fit with the

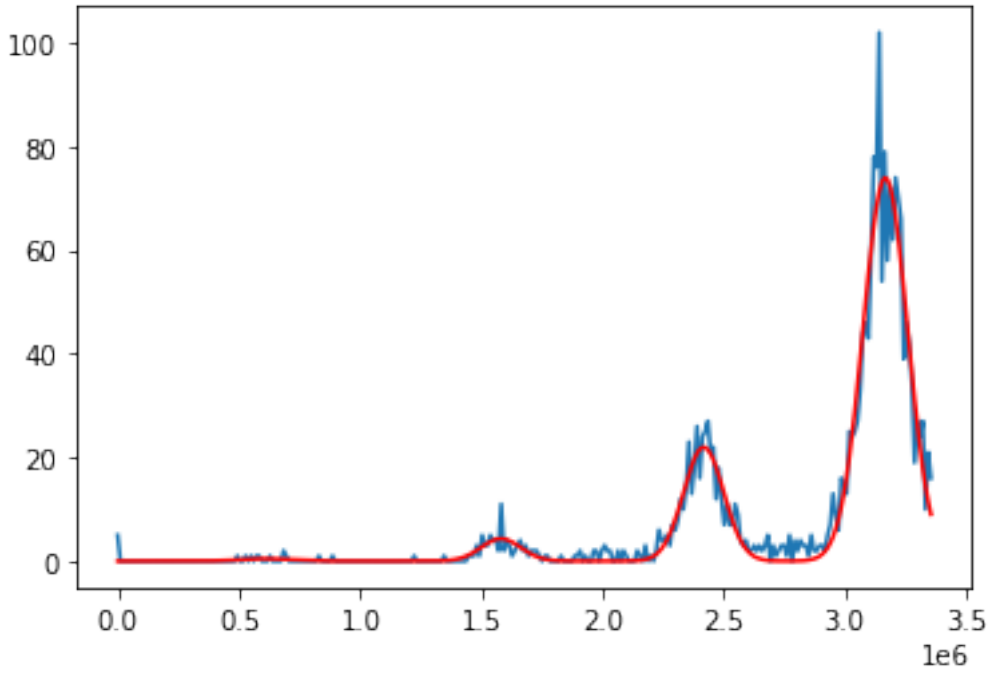
optimizer. After all sections of the histogram have been fit, they can all be stitched together at the end.

[373]:
```
'''attempt at curve fitting codes'''
from scipy.optimize import curve_fit
from scipy.signal import find_peaks

def func(x, *params):
    y = np.zeros_like(x)
    for i in range(0, len(params), 3):
        ctr = params[i] #center of Gaussian
        amp = params[i+1] #amplitude of Gaussian
        wid = params[i+2] #width of Gaussian
        #constshift=params[i+3] #constant offset to data
        y = y + amp * np.exp( -((x - ctr)/wid)**2/2)#+constshift
    return y
def␣
↪gauss_curve_fit(bin_centers,bin_heights,pkprom,start,stop,num_Gauss,widthguess,maxiters):
↪

    x=bin_centers[start:stop]
    y=bin_heights[start:stop]
    peaks=find_peaks(y, distance=15,prominence=pkprom) #finds the peaks and␣
↪peak locations with scipy function

    #plt.plot(x,y)
    #plt.show()
    peaklocs=(peaks[0]+start)*bin_centers[-1]/len(bin_heights) #peak locations
    #wid_guess=10*bin_centers[-1]/100
    wid_guess=widthguess
    #guess = [peaklocs[0],2,0.1e6,peaklocs[1],4,0.1e6,peaklocs[2],9,1e6]
    guess=[]
    #the number of how many Gaussian to use in the fit
    for i in range(num_Gauss):
        #guess += [peaklocs[i+3], bin_heights[peaks[0][i+3]],␣
↪20*bin_centers[-1]/1000]
        guess += [peaklocs[i], bin_heights[peaks[0][i]+start],wid_guess ]␣
↪#guess vals for Gaussain, center of G, height of G, and Std. dev. of G

    popt, pcov = curve_fit(func, x, y, p0=guess,bounds=(0,np.inf),maxfev =␣
↪maxiters)
    #print(popt)
    fit = func(x, *popt)

    plt.plot(x,y)
    plt.plot(x, fit , 'r-')
    plt.show()
```

```
    return fit,popt
```

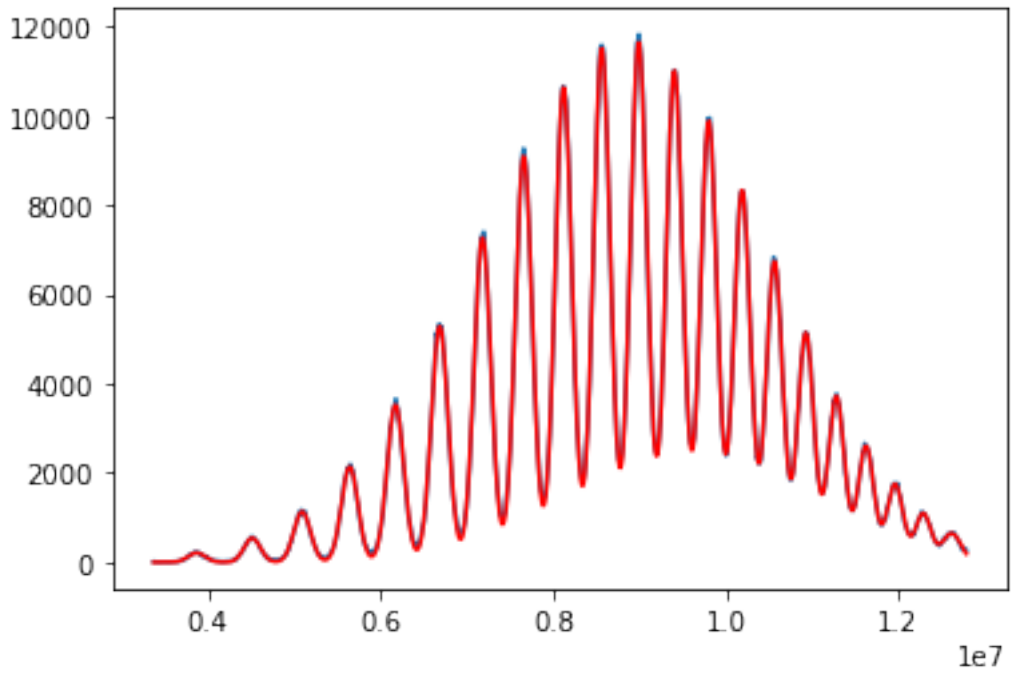
In this example, I start by fitting the 1-4 photon peaks first. Note that the zero peak is too sharp for a Gaussian fit to work for several reasons. First, the zero photon peak occurs for recorded events as well as non-recorded events, meaning that there may be less data which compounds with the fact that zero photon events are already quite rare in this data set. Additionally, the peak would be skewed right, as the left half is cut off at zero. Thus, while the other peaks can be fit with the optimizer, the zeroth-order cuttoff is determined by half the distance from zero to the first peak.

[374]:
```
fit1=gauss_curve_fit(bin_centers,bin_heights,2,0,300,4,2e5,100000)
print('Gaussian fit parameters are given by ', fit1[1])
```



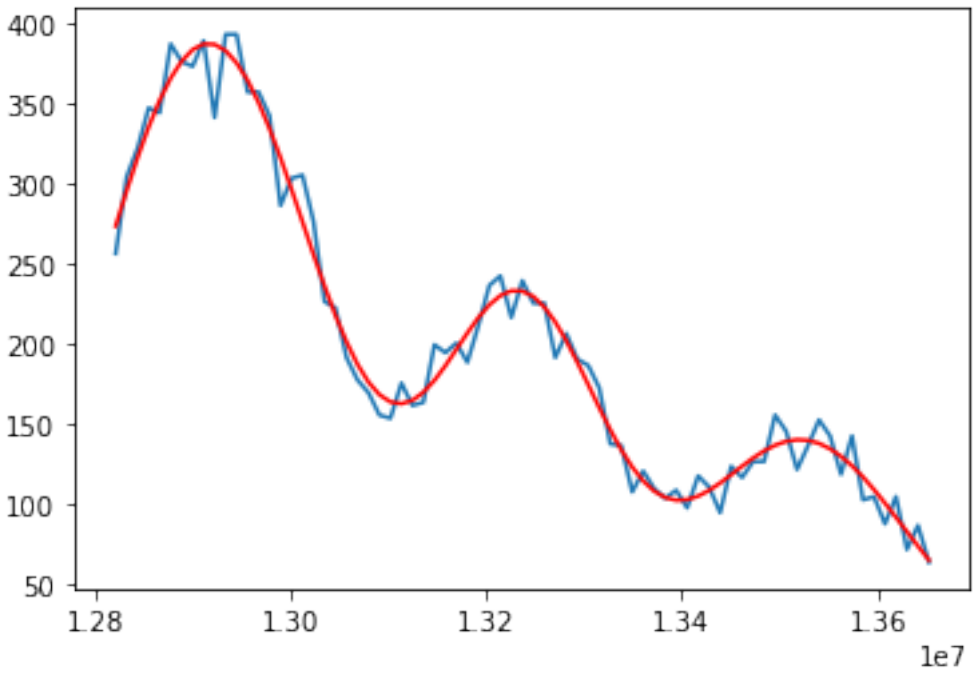

```
Gaussian fit parameters are given by  [6.04001084e+05 4.60611019e-01
9.94996845e+04 1.57928168e+06
 4.31243203e+00 8.23284133e+04 2.41674599e+06 2.19823036e+01
 8.25616420e+04 3.16664210e+06 7.40006767e+01 9.18434976e+04]
```

[430]:
```
fit2=gauss_curve_fit(bin_centers,bin_heights,200,300,1140,21,0.3e5,100000)
print('Gaussian fit parameters are given by ', fit2[1])
```



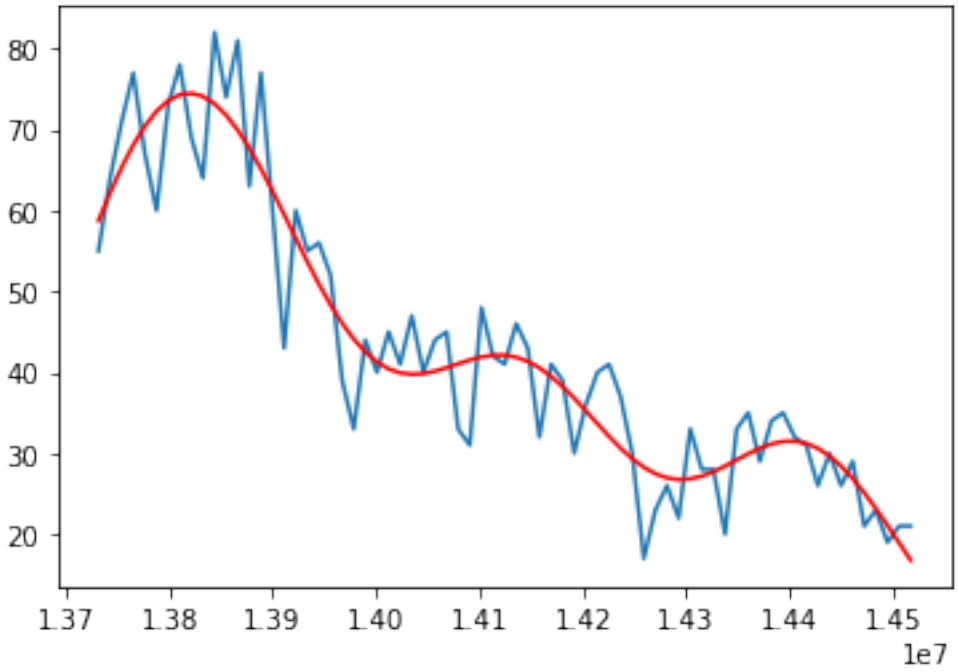

```
Gaussian fit parameters are given by  [3.86327020e+06 2.11696709e+02
9.81853096e+04 4.50898195e+06
 5.42158825e+02 9.47704394e+04 5.08723614e+06 1.13339910e+03
 9.74796763e+04 5.63882344e+06 2.14023786e+03 9.66295090e+04
 6.16943035e+06 3.52592695e+03 9.76375196e+04 6.68308308e+06
 5.29137764e+03 9.84758549e+04 7.17881799e+06 7.27936897e+03
 9.79727601e+04 7.65702086e+06 9.09967249e+03 9.87779600e+04
 8.11823309e+06 1.06392778e+04 9.81042276e+04 8.56181451e+06
 1.15191332e+04 9.78291815e+04 8.98965294e+06 1.16530548e+04
 9.76223221e+04 9.40312187e+06 1.10216200e+04 9.70047816e+04
 9.80161978e+06 9.88222373e+03 9.64529366e+04 1.01886964e+07
 8.33741350e+03 9.64219152e+04 1.05633727e+07 6.73575535e+03
 9.52499188e+04 1.09271043e+07 5.13865811e+03 9.44273492e+04
 1.12803901e+07 3.71501507e+03 9.48725115e+04 1.16246545e+07
 2.59933168e+03 9.32132540e+04 1.19613181e+07 1.74611643e+03
 9.49308624e+04 1.22867768e+07 1.08828345e+03 9.05731504e+04
 1.26163966e+07 6.47513413e+02 1.10101529e+05]
```

[470]:
```
fit3=gauss_curve_fit(bin_centers,bin_heights,20,1142,1217,3,0.3e5,100000)
print('Gaussian fit parameters are given by ', fit3[1])
```



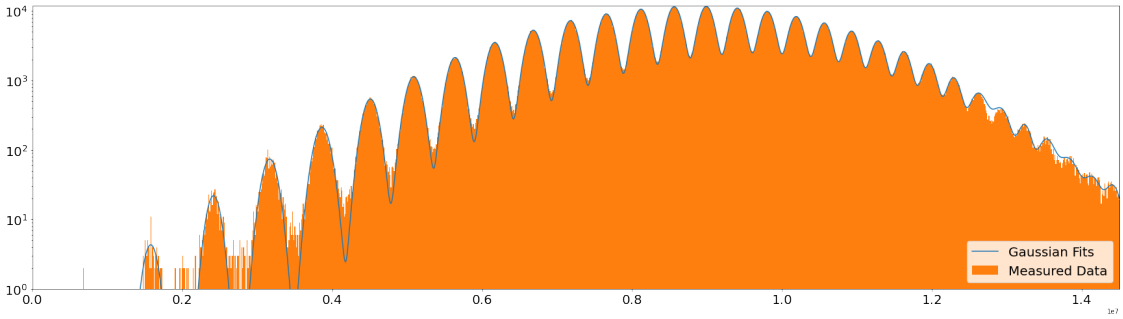

```
Gaussian fit parameters are given by  [1.29162212e+07 3.87191385e+02
1.13141100e+05 1.32348811e+07
 2.22284577e+02 8.35333280e+04 1.35226247e+07 1.39018462e+02
 1.04863667e+05]
```

[476]:
```
fit4=gauss_curve_fit(bin_centers,bin_heights,10,1223,1294,3,1e5,100000)
print('Gaussian fit parameters are given by ', fit4[1])
```




```
Gaussian fit parameters are given by  [1.38170074e+07 7.41552810e+01
1.25393136e+05 1.41387644e+07
 3.84607371e+01 1.03965308e+05 1.44128779e+07 3.01402491e+01
 9.63476991e+04]
```

[487]:
```
'''Now stitch together all of the fit parameters '''
gauss_params_fin=flatten([fit1[1],fit2[1],fit3[1],fit4[1]])

xvec=np.linspace(0,1.6e7,10000)
fig = plt.figure(figsize=(30, 8))
ax = fig.add_subplot(1, 1, 1)
ax.plot(xvec,gauss_sum(gauss_params_fin,xvec), label='Gaussian Fits')
ax.set_yscale('log')
ax.hist(histvals, bins=2000, label='Measured Data')
ax.set_xlim([0,1.45e7])
ax.set_ylim([1,1.2e4])
ax.tick_params(axis='both', labelsize=20 )
plt.legend(loc='lower right', fontsize=20)

plt.show()
```




As seen above, the Gaussian fits do a reasonable job of fitting the data with the exception of the zero peaks, which must be included manually. One may notice that some of the peaks, especially near the end, do not fit the distribution as well as one might like. This is due to the fact that the data set only contained 2.5e6 data points. Although this is already a substantial amount, the coherent state used is quite large, and thus more data will more accurately round out the Gaussian fits. When the entirety of this data set is used, the peaks fit the data more accurately and can be used to bin the data as desired. From here, one can rerun the data set with a larger amount of measurements and then perform the same analysis for each channel where data was collected. Once that has been achieved the Gaussian fit parameters can be used to determine binning threshold as per the code in the next section.

# Appendix G

# Histogram Bin Overlaps

# bin_error_checks

December 22, 2021

## 1 Bin Overlap Error

This code takes Gaussian fit parameters from a previously obtained histogram, and then examines how well those Guassians can be used to define bins for counting photon numbers. The first cell has several lists of Gaussian parameters that together desribe one full histrogram. In this case, 15 million events were sampled from a coherent state distribution with mean photon number of about 14. After using the fitted Gaussian bits, the areas of the events in the histogram were used to calculate the photon number distribution of the coherent state, which is given by the list 'coher_dist' below.

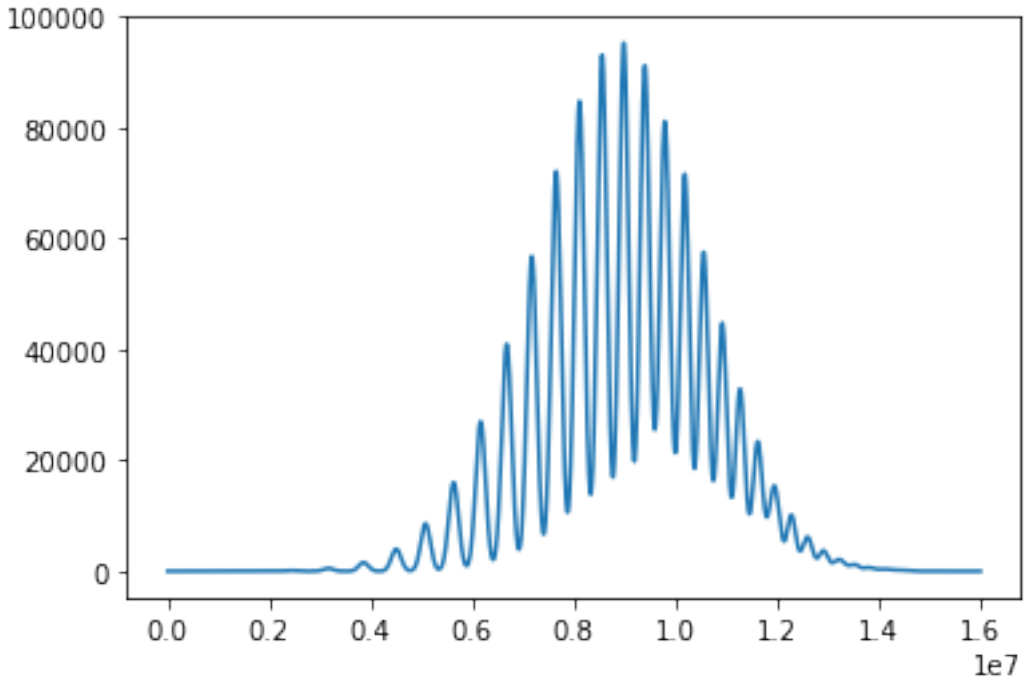
[18]:
```
import numpy as np
from qutip import* #need to download
import struct
import os, glob
import scipy.io
import bitarray as b #need to download
import matplotlib.pyplot as plt
import csv
from scipy.special import factorial as fac
from scipy.optimize import curve_fit
from scipy.signal import find_peaks
from scipy.integrate import simps

from matplotlib.ticker import (MultipleLocator, AutoMinorLocator)

def flatten(t):
    #use to flatten lists
    return [item for sublist in t for item in sublist]
'''Calculated Gaussian parameters fits from previous codes on data'''

#Measured distribution for this data
```

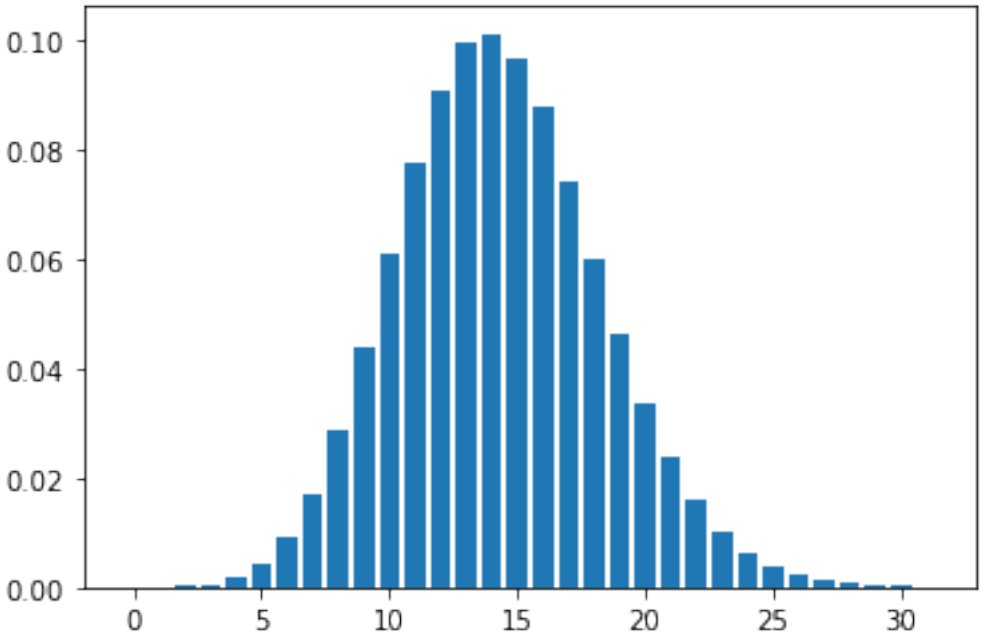
```
coher_dist=[5.414203621781381e-06, 3.054680315005051e-05, 0.
 ↪00015534085453111025, 0.0005753761083493103, 0.0017126930790235103, 0.
 ↪0042276240823509725, 0.00925721872092581, 0.016997792007552612, 0.
 ↪02933308575049117, 0.044284977209210645, 0.06089582076285594, 0.
 ↪078098617245816, 0.091234544704578, 0.10093425759804939, 0.
 ↪09962609242419898, 0.0986438489399758, 0.08790494316122248, 0.
 ↪07356946881114583, 0.06002533713610954, 0.04516682397936079, 0.
 ↪03475851883163625, 0.02164097292838031, 0.016245685598265155, 0.
 ↪010442461419955779, 0.0058813624997150855, 0.003949628121079507, 0.
 ↪0017987187587918145, 0.0012519510350119168, 0.0007123354073743725, 0.
 ↪00016255979269348542, 0.0002529970457832411, 0.00022218287455310262, 0]
#coher_dist=[5.34736160175938 85e-06, 3.001206698987457e-05, 0.
 ↪0001510629652497027 4, 0.0005797876816707617, 0.0016944452075575063, 0.
 ↪0042095767369450 35, 0.0090595668677207 79, 0.01712980499709605, 0.
 ↪0288435347958500 98, 0.04407469221422146, 0.060947222276252856, 0.
 ↪0775440290056935 4, 0.09112311905720133, 0.0994908710176945, 0.
 ↪1013187997392359 1, 0.0966113167951470 6, 0.08820840593212233, 0.
 ↪0740697813310104 4, 0.0601246643778622 15, 0.04619786213820002, 0.
 ↪0338324226442315 44, 0.0237120060767417 25, 0.01606494477612569, 0.
 ↪0099632709784181 15, 0.0064683691195482 23, 0.0036993715981171672, 0.
 ↪0022858634007120 95, 0.0011974079466739711, 0.0007202896077569896, 0.
 ↪0004086721104144 6127, 0.0002323428615964 4544, 0]

'''Gaussian fit parameters for binning. Each Gaussian has three parameters, and␣
 ↪they occur in order in each list. The first
is the location of the ceneter of the Gaussian in arbitrary area units used for␣
 ↪the histogram, the second is the Gaussian
peak amplitude, and the third is the standard deviation.'''


gauss_params_fin=[6.08268770e+05,  2.45024365e+00,  1.12708544e+05, 1.
 ↪56300607e+06, 2.37937662e+01, 9.05904304e+04,
                2.41043719e+06, 1.41715463e+02, 8.73496397e+04, 3.16019593e+06,␣
 ↪5.19109750e+02, 9.72316856e+04,
                3.85652682e+06, 1.58079006e+03, 9.40180788e+04, 4.49960453e+06,␣
 ↪3.99563251e+03, 9.44425103e+04,
                5.07787115e+06, 8.57688496e+03, 9.47181494e+04, 5630506.
 ↪09640559,   16016.10272244,   97158.56749569,
                6161137.05076052,   27001.54183779,   97167.60939734, 6674973.
 ↪14828196, 40964.75701783,   98093.59459127,
                7170491.77639382,   56800.26467025,  97911.73929927, 7648979.
 ↪23851249 ,  72124.33278633,   99040.31382459,
                8108023.50008762,   84787.72518092,   98904.57569745, 8552144.
 ↪03989489,   93112.12210683,   97695.19977713,
                8979752.00179682,   95255.7521997 ,   97227.98508008, 9393836.
 ↪44675478,   91158.58222094,   98409.72811466,
```

```
              9788275.76988778,   81123.92324075 ,  103752.39055317, 10179787.
↪99087714,    71545.20150504 ,   93673.74158892,
              10552643.30368328,  57519.17636692,    96039.84687626, 10916503.
↪80586508,    44685.73830857,    94439.4994612,
              11269516.16300907,    32828.38105522,    94419.97121048,␣
↪11616137.93309524,    23230.73362739,    95283.32052594,
              1.19428338e+07, 1.53362445e+04, 1.05822141e+05, 1.22804839e+07,␣
↪1.00113073e+04, 8.69557886e+04,
              1.25960704e+07, 6.17215490e+03, 1.00249661e+05, 1.29103100e+07,␣
↪3.62335316e+03,8.80348198e+04,
              1.32202881e+07, 2.11691440e+03, 1.13544341e+05, 1.35193111e+07,␣
↪1.15662164e+03, 8.65090415e+04,
              1.38045792e+07, 6.64399822e+02, 1.16787485e+05,1.40957079e+07, 3.
↪14143831e+02, 8.58188067e+04,
              1.43757675e+07, 2.25769568e+02, 1.60925503e+05]
```

Next, define a few functions. The functions below take in a list of Gaussian parameters and an x vector, and then return the sum of gaussians. The first function gives the sum directly, but the second function first normalizes each Gaussian in the sum before returning the result.

[17]:
```
def gauss_sum(params,x):
    f = np.zeros_like(x)
    for k in range(round(len(params)/3)):
        ctr = params[3*k] #center of Gaussian on x-axis
        amp = params[3*k+1] #amplitude of peak
        wid = params[3*k+2] #standard deviation of the peak
        f = f + amp * np.exp( -((x - ctr)/wid)**2/2)
    return f

def gauss_sum_normed(params,x):
    f = np.zeros_like(x)
    for k in range(round(len(params)/3)):
        ctr = params[3*k] #center of Gaussian on x-axis
        amp = params[3*k+1] #amplitude of peak
        wid = params[3*k+2] #standard deviation of the peak
        f = f + np.exp( -((x - ctr)/wid)**2/2)/np.sqrt(2*np.pi*wid**2)
    return f

xvec=np.linspace(0,1.6e7,10000)
g_sum=gauss_sum(gauss_params_fin,xvec)

plt.plot(xvec,g_sum)
plt.show()
plt.bar(range(len(coher_dist)),coher_dist)
plt.show()
```



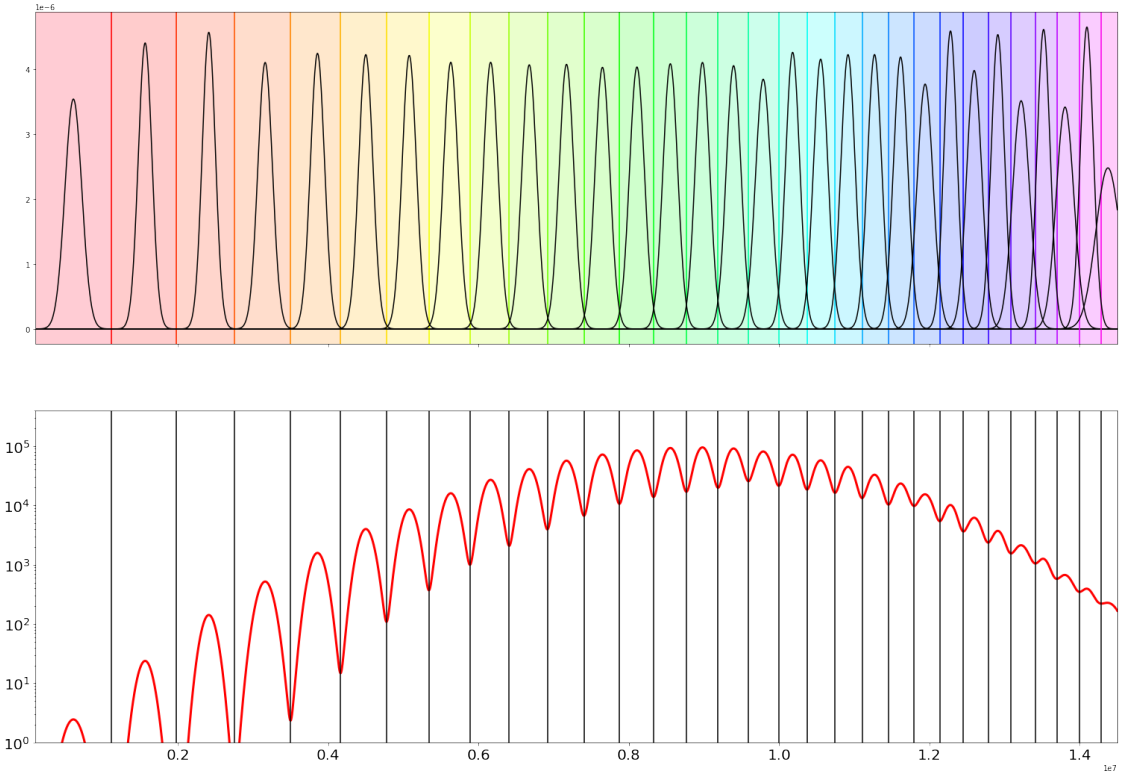

Next, find the intersection points between the neighboring Gaussians to determine the bins for the data. These locations will occur where there is a minima in the Gaussian sum.

[19]:
```
'''Instead of the intersection points between neighboring Gaussians, use the␣
 ↪minima of the sum of Gaussians'''
#This function takes in the gaussian parameters, the maximum cuttoff for the x␣
 ↪value, and the number
#of points in the x-axis array corresponding to histogram areas
def xbin_locs2(params,xmax,points):
    x=np.linspace(0,xmax,points)
    f = np.zeros_like(x)
    for k in range(round(len(params)/3)):
        #print(k)
        ctr = params[3*k]
        amp = params[3*k+1]
        wid = params[3*k+2]
        f=f+amp*np.exp( -((x - ctr)/wid)**2/2)
    pks=find_peaks(-f)[0]*xmax/points #these are the locations of the minima
        #diffs=-np.abs(f1-f2)
        #loc=np.where(diffs==np.max(diffs[round(points*ctr1/xmax):
 ↪round(points*ctr2/xmax)]))[0]*xmax/points
        #vec.append(loc)
    return pks

#Calculate the location of the bins for the photon counts
photonbins=xbin_locs2(gauss_params_fin,1.5e7,100000)
```

Now, plot the results, both the normalized Gaussians with bins and the normal sum.

[20]:
```
params=gauss_params_fin
NUM_COLORS = 31

#cm = plt.get_cmap('turbo')
cm = plt.get_cmap('gist_rainbow')
#ax = fig.add_subplot(111)
#ax.set_color_cycle([cm(1.*i/NUM_COLORS) for i in range(NUM_COLORS)])
fig, ax = plt.subplots(2, sharex=True,figsize=(26, 18))

#ax.plot(xvec,gauss_sum_normed(gparams2,xvec),color='k',linewidth=2)

for k in range(len(photonbins)+1):
    if k==0:
        lend=0
    else:
        lend=photonbins[k-1]
    if k==len(photonbins):
        rend=1.6e7
    else:
        rend=photonbins[k]
    ax[0].axvline(rend, color=cm((1.*k+1)/NUM_COLORS))
```

```
    ax[0].axvspan(lend, rend, alpha=0.2,color=cm(1.*k/NUM_COLORS) )
    ctr = params[3*k]
    amp = params[3*k+1]
    wid = params[3*k+2]
    f =  np.exp( -((xvec - ctr)/wid)**2/2)/np.sqrt(2*np.pi*wid**2)
    ax[0].plot(xvec,f,color='k')
#ax.plot(xvec,gauss_sum(gauss_params_fin,xvec),color='r',linewidth=3)
plt.xlim([1e5,1.48e7])
#plt.show()

#fig, ax = plt.subplots(figsize=(22, 6))
ax[1].plot(xvec,gauss_sum(gauss_params_fin,xvec),color='r',linewidth=3)
#ax.hist(histvals, bins=1500, label='histogram')


for k in range(len(photonbins)+1):
    if k==0:
        lend=0
    else:
        lend=photonbins[k-1]
    if k==len(photonbins):
        rend=1.6e7
    else:
        rend=photonbins[k]
    ax[1].axvline(rend, color='k')
    #ctr = params[3*k]
    #amp = params[3*k+1]
    #wid = params[3*k+2]
    #f =  amp * np.exp( -((xvec - ctr)/wid)**2/2)
    #ax.plot(xvec,f,color='r')
    #ax.axvspan(lend, rend, alpha=0.15,color=cm(1.*k/NUM_COLORS) )
plt.yscale("log")
plt.ylim([1,4e5])
plt.xlim([1e5,1.45e7])
ax[1].tick_params(axis='both', which='major', labelsize=20)
ax[1].tick_params(axis='both', which='minor', labelsize=20)
plt.show()
```




The error for an individual event will be the overlap of each Gaussian with neighboring bins. The next function takes in the bins found previously, the left limit cuttoff, and the Gaussian parameters.

[21]:
```
'''take those bins I just found, and integrate the gaussian in the bin over␣
↪limits of the bin boundaries'''
def bin_occupancy(bins,llim,gparams):
    points=100000
    xmax=1.6e7
    x=np.linspace(0,xmax,points)
    ovlp=[]
    right_error=[]
    left_error=[]
    for k in range(len(bins)+1):
        if k==0:
            lend=llim
        else:
            lend=bins[k-1]
        if k==len(bins):
            rend=1.46e7
        else:
            rend=bins[k]

        ctr = gparams[3*k]
```

```
        amp = gparams[3*k+1]
        wid = gparams[3*k+2]
        f = np.exp( -((x - ctr)/wid)**2/2)/np.sqrt(2*np.pi*wid**2)

        lidx=round(lend*points/xmax)
        ridx=round(rend*points/xmax)
        ldat=np.array(f)
        ldat[lidx:]=0
        rdat=np.array(f)
        rdat[:ridx]=0
        f[:lidx]=0
        f[ridx:]=0
        ovlp.append(simps(f,x,1))
        left_error.append(simps(ldat,x,1))
        right_error.append(simps(rdat,x,1))
    return ovlp,left_error,right_error


'''lerror gives a list of the overlap of each gaussian with everything to the␣
 ↪left of its bin, which is to a good
approximation just the overlap with the neighboring bin to the left. rerror the␣
 ↪gives the overlap with the bin to the
right. good_ovlp is the overlap of the gaussian with its own bin.'''
overlaps=bin_occupancy(photonbins,0,gauss_params_fin)
rerror=np.asarray(overlaps[2])
lerror=np.asarray(overlaps[1])
good_ovlp=np.asarray(overlaps[0])


np.set_printoptions(precision=4)
print('Probability of a photon count to be in the assigned bin: ',good_ovlp)

print('Approximate probability of a photon count to be miscounted and assigned␣
 ↪to the bin to the left: ',lerror)

print('Approximate probability of a photon count to be miscounted and assigned␣
 ↪to the bin to the right: ',rerror)
```

```
Probability of a photon count to be in the assigned bin:  [1.     1.     1.
0.9997 0.9994 0.998  0.9964 0.9942 0.9922 0.9896
 0.9866 0.9821 0.978  0.9745 0.9697 0.96   0.9476 0.9556 0.9471 0.9425
 0.9356 0.9252 0.8839 0.9264 0.8906 0.9065 0.8325 0.8753 0.7762 0.8605
 0.623 ]
Approximate probability of a photon count to be miscounted and assigned to the
bin to the left:  [0.0000e+00 2.8334e-07 3.4015e-07 1.2755e-05 6.8171e-05
1.8492e-04
 6.9307e-04 1.3564e-03 2.2381e-03 3.2466e-03 4.5824e-03 6.4794e-03
```

```
 9.0457e-03 1.1029e-02 1.3815e-02 1.7809e-02 2.9066e-02 2.4108e-02
 2.9782e-02 3.3901e-02 3.9686e-02 4.5479e-02 8.3153e-02 5.0032e-02
 7.7006e-02 7.1196e-02 1.2038e-01 1.0380e-01 1.7497e-01 1.2740e-01
 2.9527e-01]
Approximate probability of a photon count to be miscounted and assigned to the
bin to the right:  [4.2883e-06 2.4937e-06 4.8722e-05 2.5740e-04 5.5091e-04
1.7722e-03
 2.9013e-03 4.4128e-03 5.5359e-03 7.1334e-03 8.8129e-03 1.1470e-02
 1.2906e-02 1.4445e-02 1.6499e-02 2.2186e-02 2.3297e-02 2.0242e-02
 2.3144e-02 2.3628e-02 2.4693e-02 2.9285e-02 3.2929e-02 2.3525e-02
 3.2420e-02 2.2317e-02 4.7123e-02 2.0891e-02 4.8860e-02 1.2084e-02
 8.1688e-02]
```

We can assume that the error is limited to nearest-neighbor bins, as the next-nearest-neighbor overlap will be considerably smaller. Now, for even photon number outcomes, the overlap to the left and right is the probability to be miscounted as an odd, and the same holds true for odd photons being miscounted as even. Also, the parity error will depend on the specific distribution, as the Gaussian overlap errors are not the same for every photon-number bin.

[22]:
```
'''Find the parity error, which will depend on the measured distribution as␣
↪well'''
ldist_error=np.multiply(lerror,coher_dist[:len(lerror)])
rdist_error=np.multiply(rerror,coher_dist[:len(rerror)])
tot_err=rdist_error+rdist_error


odd_error=np.sum(tot_err[::2]) #total error prob of even to be counted as an␣
↪odd (sums all error over even counts)
even_error=np.sum(tot_err[1::2]) #total error prob of an odd to be counted as␣
↪an even (sums all error over odd counts)

print('Average error on the parity for this distribution is given by: ',␣
↪odd_error-even_error)
```

```
Average error on the parity for this distribution is given by:
0.00043309896413963947
```

[ ]:
```

```

# Appendix H

# Overlap Tomography SDP Reconstruction

# Fock_state_overlaptomo

March 23, 2022

## 1 Overlap Tomography Code

This code performs the tomographic reconstruction from the overlap tomography experiment. Specifically, this code is for the overlap tomography of a single-photon Fock state. The basis for this reconstruction is a convex-optimization semidefinite programming (SDP) algorithm used to solve the matrix equation $M = CP$, where $M$ is a column vector containing all of the overlap measurements, $C$ is a coefficient matrix that can be calculated from the coherent state calibration measurements, and $P$ is the unknown density operator in Liouville vector representation. Instead of directly inverting $C$, which could cause problems due to even small errors in the calibration blowing up upon matrix inversion, the more robust minimization of the $L2$ norm, $||M - CP||_2$ can be performed. Minimizing this quantity yields a global minimum that this computationally efficient to solve. Additionally, we can use physicality constraints on the density matrix $\rho = \rho^\dagger$ and $\mathrm{Tr}[\rho] = 1$ to ensure a physical reconstruction. This code uses the Python libraries CVXPY and CVXOPT to perform the optimization.

Once the reconstruction is performed, a second optimization problem can be solved to compensate for calibrated loss, as the loss process is just a different matrix mapping from the pure state to the loss-degraded state.

[1]:
```
import time
from matplotlib import rc
import math
from qutip import *
#from qutip.ipynbtools import plot_animation
#import qutip

import numpy as np
%matplotlib inline
import matplotlib.pylab as plt
import matplotlib as mpl
from mpl_toolkits.mplot3d import Axes3D
from matplotlib import cm
from IPython.display import display, Math, Latex

import cmath
from mpl_toolkits.axes_grid1 import AxesGrid
from scipy.special import factorial
from scipy.special import binom
```

```
from scipy import linalg
import scipy as scp
import cvxpy as cp
import cvxopt

global N_dim
global N_trun
N_dim = 10;
N_trun = 6;
```

First we'll calculate the coefficient matrix from the measured coherent states. We could use directly measured results, but we can improved our results by including some additional calibrations such as imperfect visibility between the signal photon and the coherent state. If we assume that the fiber collects everything, then the only mode we should care about is the portion of the signal that correctly interferes with the coherent state, but in actuality, we also measure the non-interferring photons. Additionally, when calibrating the coherent state, we collect more light than actually interferes with the signal, as some of this coherent state will not overlap. Thus, when calculating the coefficient matrix, the value of the coherent state used should actually be smaller than what is measured by a factor of the visibility. Below, we have the measured amplitudes of the five non-zero coherent state values which were calibrated before and after each experimental run. The values used for the reconstruction is the average of the measurements scaled by the visibility.

[2]:
```
'''Coherent State calibrations'''
vis=0.9
M=vis**2
alpha_LO1 = np.sqrt(2)*np.array([0,0.2392,0.2577, 0.2877, 0.3158, 0.50774])␣
↪#Measured values at start
alpha_LO2 =np.sqrt(2)*np.array([0,0.2653,0.2728, 0.293 ,0.3198, 0.5054])  ␣
↪#Measured values at end
alpha_LO=(alpha_LO1+alpha_LO2)/2
alpha_LO_new=alpha_LO*vis
```

[3]:
```
np.sqrt(1-M)
```

[3]:
```
0.4358898943540673
```

The measured amplitudes are only part of the calibrations. Additionally, we need the phase of each measurement. In this experiment, I used a mirror-mounted PZT to vary the phase of the overlapping coherent states. Each shift to the piezo applied a 0.58 radian phase. Additionally, because the EOM was used to control the amplitudes of the coherent states, the phase also depends on the amplitude since changing the voltage on the EOM will change the phase. The phases were all calibrated with classical beams before the experimental data collection. At the end of the cell below, the variable 'Final_prob' is now a 2D array that contains all of the coefficients needed to solve the SDP problem.

[4]:
```
'''Create a matrix of the coefficients of all of the
coherent states used to probe the state to tomograph'''
```

```
prob_matrix_tomo=[]
phase_step=0.58 #the PZT changes the phase by 0.58 rads
EOM_phase_step=0.277 #phase change due to EOM in rads

'''In this set of data, I did i=24;16;8;1 for the EOM values, so I need to␣
 ↪assign EOM values to negative phases.
Hence the minus sign on the EOM_phase_step in line 22 of this cell'''
amp_num=5 #number of coherent state probe amplitudes (not including vacuum)

for i in range(amp_num):
    amp=alpha_LO_new[i+1];
    if (i+1)>4:
        EOM_phase=np.e**(-EOM_phase_step*(i-5)*1j)
    elif 5>(i+1)>0:
        EOM_phase=np.e**(-EOM_phase_step*(i)*1j)
    else:
        EOM_phase=1
    for k in range(10): #(number of phases used)
        phase=np.e**(k*1j*phase_step)
        B = coherent_dm(N_dim,amp*phase*EOM_phase).full()
        ele = B.flatten()
        prob_matrix_tomo.append(ele)

'''Add the vacuum component last'''
B_vac=coherent_dm(N_dim,0).full()
prob_matrix_tomo.append(B_vac.flatten())

Final_prob = np.array(prob_matrix_tomo);
```

The cell below contains the measured photon-number probability distributions from the TES for each overlap setting for five values of coherent state amplitude each at 10 phases from 0 to $2\pi$. The first set of 'amp_0' is when the coherent state field is block, thus the signal is simply overlapped with vacuum.

[5]:
```
'''experimental number distributions go here'''
refl_tot=.4 #This is the loss we determined from the heralding ratio

amp_0_1=[0.8134,0.1791,0.0075,0,0,0,0]
amp_0_2=[0.7974,0.2026,0,0,0,0,0]
amp_0_3=[0.7869,0.2131,0,0,0,0,0]
amp_0=np.zeros(7)
for i in range(len(amp_0_1)): #average the 3 measurements
    amp_0[i]=(amp_0_1[i]+amp_0_2[i]+amp_0_3[i])/3

amp_1_p1 = [0.6619, 0.3094,0.0288,0,0,0,0]
amp_1_p2 = [ 0.6917,0.2667,0.0417,0,0,0,0]
amp_1_p3 = [ 0.6771,0.2396,0.0833,0,0,0,0]
```

```
amp_1_p4 = [ 0.6579,0.2895,0.0439,0.0088,0,0,0]
amp_1_p5 = [0.66950,0.3136,0.0169,0,0,0,0]
amp_1_p6 = [0.7254,0.2042,0.0634,0.007,0,0,0]
amp_1_p7 = [0.7109,0.25,0.0391,0,0,0,0]
amp_1_p8 = [0.7284,0.2407,0.0309,0,0,0,0]
amp_1_p9 = [0.6855,0.2642,0.044,0.0063,0,0,0]
amp_1_p10 = [ 0.7414,0.2241,0.0345,0,0,0,0]

amp_2_p1 = [0.7143,0.2476,0.0381,0,0,0,0]
amp_2_p2 = [ 0.7596,0.2212,0.0192,0,0,0,0]
amp_2_p3 = [ 0.7042,0.2535,0.0352,0.007,0,0,0]
amp_2_p4 = [ 0.731,0.2164,0.0526,0,0,0,0]
amp_2_p5 = [0.7345,0.2373,0.0226,0.0056,0,0,0]
amp_2_p6 = [0.6842,0.25,0.0658,0,0,0,0]
amp_2_p7 = [0.6698,0.283,0.0472,0,0,0,0]
amp_2_p8 = [0.7068,0.2408,0.0419,0.0105,0,0,0]
amp_2_p9 = [0.7341,0.2197,0.0405,0.0058,0,0,0]
amp_2_p10=[0.6587,0.3077,0.024,0.0096,0,0,0]

amp_3_p1 = [0.6863,0.2598,0.049,0.0049,0,0,0]
amp_3_p2 = [ 00.7037,0.2361,0.0602,0,0,0,0]
amp_3_p3 = [ 0.6364,0.298,0.0606,0.0051,0,0,0]
amp_3_p4 = [ 0.6524,0.2866,0.061,0,0,0,0]
amp_3_p5 = [0.6946,0.2635,0.0419,0,0,0,0]
amp_3_p6 = [0.7113,0.232,0.0515,0.0052,0,0,0]
amp_3_p7 = [0.7468,0.1962,0.038,0.019,0,0,0]
amp_3_p8 = [0.6526,0.3105,0.0368,0,0,0,0]
amp_3_p9 = [0.7086,0.2514,0.04,0,0,0,0]
amp_3_p10=[0.6527,0.3054,0.0359,0.006,0,0,0]

amp_4_p1 = [0.6866,0.2687,0.0398,0.005,0,0,0]
amp_4_p2 = [ 0.6842,0.269,0.0409,0.0058,0,0,0]
amp_4_p3 = [0.6258,0.2945,0.0736,0.0061,0,0,0]#check this point - might be wrong
amp_4_p4 = [ 0.6299,0.3052,0.0519,0.013,0,0,0]
amp_4_p5 = [0.6848,0.2446,0.0707,0,0,0,0]
amp_4_p6 = [0.6377,0.3261,0.029,0.0072,0,0,0]
amp_4_p7 = [0.671,0.2645,0.0581,0.0065,0,0,0]
amp_4_p8 = [0.6918,0.2453,0.0503,0.0126,0,0,0]
amp_4_p9 = [0.7299,0.1971,0.073,0,0,0,0]
amp_4_p10=[0.7071,0.2357,0.0571,0,0,0,0]

amp_5_p1 = [0.5798,0.3109,0.1008,0.0084,0,0,0]
amp_5_p2 = [ 0.5783,0.2771,0.0904,0.0482,0.006,0,0]
amp_5_p3 = [0.5333,0.3222,0.1111,0.0167,0.0167,0,0]
amp_5_p4 = [ 0.5839,0.2919,0.0932,0.0311,0,0,0]
amp_5_p5 = [0.549,0.3464,0.0915,0.0131,0,0,0]
amp_5_p6 = [0.6557,0.2842,0.0546,0.0055,0,0,0]
```

```
amp_5_p7 = [0.6481,0.2099,0.1049,0.037,0,0,0]
amp_5_p8 = [0.6051,0.2484,0.1274,0.0127,0,0.0064,0]
amp_5_p9 = [0.5988,0.2515,0.0958,0.0359,0.012,0.006,0]
amp_5_p10=[0.6175,0.2568,0.1038,0.0219,0,0,0]

'''Restructure data slightly'''
Amp0=[]; Amp1=[]; Amp2=[]; Amp3=[]; Amp4=[]; Amp5 = []
for k in range(6):
    for i in range(10):
        if k==0:
            temp_val=np.asarray(eval('amp_0'))
        else:
            temp_val=np.asarray(eval('amp_'+str(k)+'_p'+str(i+1)))
        eval('Amp'+str(k)).append(temp_val)
```

Earlier, it was mentioned that a correction needed to be applied to the coherent state amplitudes to account for the imperfect visibility. Similarly, the extra photons from the coherent state that don't properly interfere with the signal add extra photons to the measured probabilities. Since we know the visibility ahead of time, we can use the measured value and deconvolve the erroneous photons from the data to recover the true measured distribution. This is done by the function 'Amp_correction' below, which simply performs matrix multiplication to map the raw distribution to the deconvoluted distribution. The matrix needed is calculated in the rest of the cell. Note that this matrix is actually a ragged matrix of many submatrices, each designed to correct the mismatched visibility for the different coherent state amplitudes used for the experiment.

First, 'matrix_M' is calculated, which contains all of the matrix maps that convolve the true distributions with non-interfering coherent state photons. These matrices must then be inverted to 'M_inv' which will be applied to the actually measured distributions to attain the true distributions that we can use for further tomographic reconstruction.

[6]:
```
def Amp_correction(num_dist,Matrix):
    leng=len(num_dist)
    new=[]
    for i in range(leng):
        new.append(np.dot(Matrix,num_dist[i]))
    return new

#Matrices that describes the convolution of a density matrix with coherent␣
 ↪state probabilities
matrix_M=[]
dim_out=N_dim #Hilbert space dimentsion cuttoff of the output (measured) state
leng=len(alpha_LO)
for i in range(leng):
    Prob_vec=[]
    state=coherent_dm(N_dim,np.sqrt(1-M)*alpha_LO[i])
    for j in range(7):
        row=[]
        for k in range(7):
```

```
        #for k in range(dim_in):
            if k > j: #ensures the result is upper triangular
                val=0
            else:
                #val=(binom(j+k,k)*binom(j+i+k,k))**(1/2)*r**(2*k)*t**(j+(i+j))
                #The above line is correct for the values as they start, but I␣
↪forgot
                #that there is an offset, so I need to replace k with (k-j)
                val=state.diag()[j-k]
            row.append(val) #appends each value to the k^th position in the␣
↪j^th row

        #for n in range(j):
        #    row[n]=row[j-1-n]
        Prob_vec.append(row) #appends the j^th row to the i^th matrix
    matrix_M.append(Prob_vec)

M_inv=[] #Invert the above matrix
for i in range(len(alpha_LO)):
    inverse=np.linalg.inv(matrix_M[i])
    M_inv.append(inverse)

#New amplitudes that are now the corrected measured probability distributions
Amp0_new=Amp0
for k in range(5):
   ␣
↪vars()['Amp'+str(k+1)+'_new']=Amp_correction(eval('Amp'+str(k+1)),M_inv[k+1])
```

Next, take array of probability distributions and turn it into an array of parities. Remeber that the parity for each distribution is directly related to the overlap fidelity of the unknown density matrix with the calibrated coherent states.

The capital P# variables (P0, P1, etc.) hold the parities for all measured data points. The lowercase p# variables hold the averaged measured parity accross all phases for a given amplitude. The final array, 'Fid_array_fin', is the vector of overlap measurements that can go into our SDP optimization algorithm along with the coefficient matrix to solve for the unknown density operator.

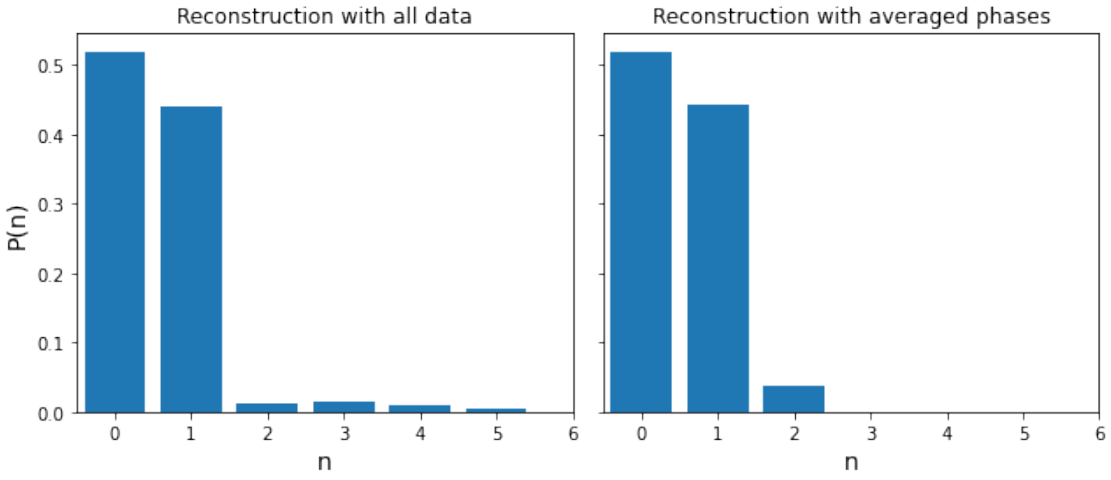
[7]:
```
def par_array(num_dist_matrix, length, n_trun):
    P_array=[]
    for k in range(length):
        P_n=num_dist_matrix[k]
        Parity = 0;
        for i in range(n_trun):
            Parity = Parity + math.pow(-1,i)*P_n[i]
        P_array.append(Parity)
    return P_array

for k in range(6):
```

```
    vars()['P'+str(k)]=par_array(eval('Amp'+str(k)+'_new'),10,N_trun)

for i in range(5):
    temp=eval('P'+str(i+1))
    avg=np.average(temp)
    vars()['p'+str(i+1)]=[avg,avg,avg,avg,avg,avg,avg,avg,avg,avg]
p0=P0

#sets the final distsributions used, P# has data for amplitude# at all phases
Fid_array=np.ndarray.flatten(np.array([P1,P2,P3,P4,P5]))

#This is the array that now averages the measurements at each amplitude overall␣
 ↪all phases instead
Fid_array_averaged=np.ndarray.flatten(np.array([p1,p2,p3,p4,p5]))

'''append the vacuum measurement data'''
Fid_array_fin=np.append(Fid_array,P0[1])
Fid_array_fin_averaged=np.append(Fid_array_averaged,p0[1])
```

The function below is designed to use the measurements ('Fid_array_fin') and the calibrated coefficients from the coherent state probe matrix ('Final_prob') to solve the semidefinite programming opimization problem. This function minimizes the error function, which is the L2 norm mentioned above, $||CP - M||_2$. In addition, a small parameter $\gamma$ is used as a regularizer that penalizes large elements of the reconstruction. This is a small overall effect but helps minimize the spurious effects of noise that could lead to unrealistically large coherences in the density matrix from experimental errors.

[8]:
```
def Convex_optimization_state_tomography(Matrix_prob, Measurements, gamma):
    #gamma: small regularizing parameter that helps with noise. Aribitrary, but␣
 ↪should be small, order of 0.01 or less
    C = Matrix_prob; #This is the input coefficient matrix
    Meas = Measurements; #These are the measured overlap fidelities
    P = cp.Variable((N_dim,N_dim), PSD = True) #unknown density matrix
    #P = cp.Variable((N_dim,N_dim), Hermitian = True) #set hermitian true when␣
 ↪I include phase

    Error = cp.norm(C@cp.vec(P) - Meas,2) #second paramter gives norm type

    Obj_detect = cp.Minimize(Error + gamma*cp.norm(cp.vec(P),2))
    constraints = [cp.trace(P)==1]# physicality constraint
    for i in range(N_dim):
        constraints.append(cp.real(P[i][i]) >= 0) #ensure diagonals are real
        if i>5:
            #based on direct photon counting (without overlap tomo), we know␣
 ↪the probability distribution for the state is zero beyond 4 photons
            constraints.append(cp.real(P[i][i]) == 0)
    Prob_detect = cp.Problem(Obj_detect,constraints)
```

```
    Prob_detect.solve(verbose = False)
    #Prob_detect.solve(cp.CVXOPT)  #can choose a different solver
    p_values = (P.value)
    return p_values

P1 = Convex_optimization_state_tomography(Final_prob, Fid_array_fin, .01) #All␣
↪data points used
P2 = Convex_optimization_state_tomography(Final_prob, Fid_array_fin_averaged, .
↪01) #averaged phases

P_arr1 = np.array(P1).reshape(N_dim,N_dim)
P_arr2 = np.array(P2).reshape(N_dim,N_dim)


#Plots show the reconstructed photon number distributions
fig, ax = plt.subplots(1,2, sharey=True,figsize=(9,4))

ax[0].bar(range(N_dim),P_arr1.diagonal(0))
ax[0].set_title('Reconstruction with all data')
ax[0].set_ylabel('P(n)',fontsize=14)
ax[0].set_xlabel('n',fontsize=14)
ax[0].set_xlim([-0.5,6])
#ax1.xlabel('n',fontsize=12)
ax[1].bar(range(N_dim),P_arr2.diagonal(0))
ax[1].set_title('Reconstruction with averaged phases')
ax[1].set_xlabel('n',fontsize=14)
ax[1].set_xlim([-0.5,6])
plt.tight_layout()
plt.show()
```

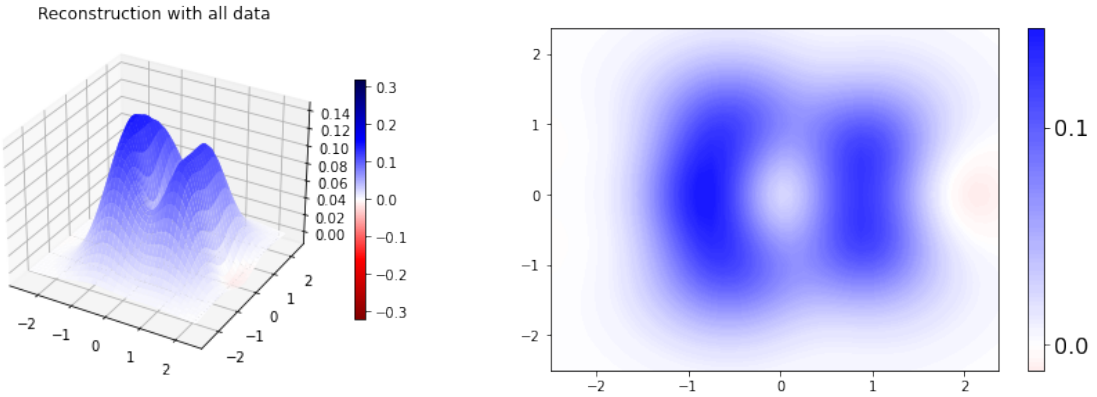



Looking at the photon-number distributions, we see that for the most part we just have vacuum

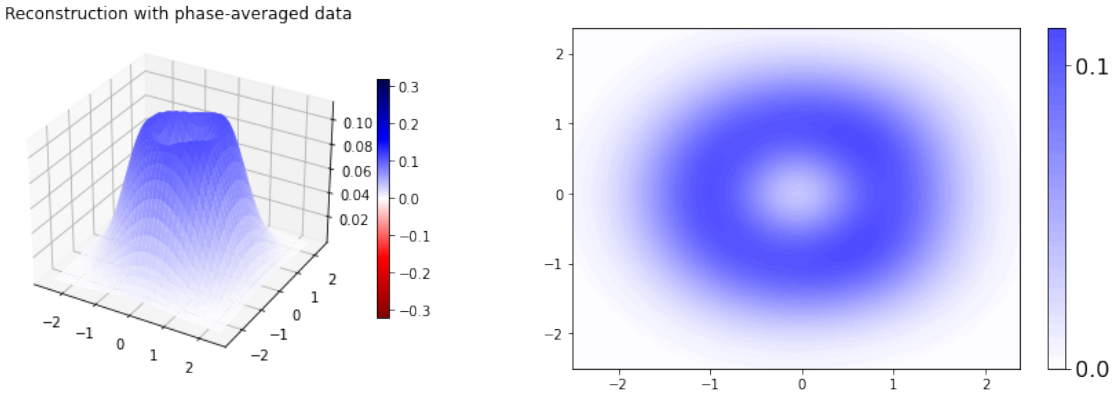
and single photon probabilities. This makes sense, since we send a single photon through a lossy channel. Now let's plot the Wigner functions for both reconstructions.

[9]:
```
xvec = np.arange(-20.,20.)*5./40
yvec = np.arange(-50.,50)*5/40
X,Y = np.meshgrid(xvec, xvec)
X1,Y1 = np.meshgrid(yvec,yvec)

q_tomo1=Qobj(P_arr1)
q_tomo2=Qobj(P_arr2)
W1=wigner(q_tomo1,xvec,xvec)
W2=wigner(q_tomo2,xvec,xvec)

fig = plt.figure(figsize=(16,10))
# `ax` is a 3D-aware axis instance, because of the projection='3d' keyword␣
 ↪argument to add_subplot
ax = fig.add_subplot(2, 2, 2)
p = ax.contourf(X, Y, W1, 80, cmap=cm.seismic_r,vmin=-0.32,vmax=0.32)
#cb = fig.colorbar(p, shrink = 0.7)
cb = fig.colorbar(p, shrink = 1)
cb.ax.tick_params(labelsize=16)
cb.set_ticks([-0.3,-0.2,-0.1, 0,0.1,0.2,0.3]);

# surface_plot with color grading and color bar
ax = fig.add_subplot(2, 2, 1, projection='3d')
p = ax.plot_surface(X, Y, W1, rstride=1, cstride=1, cmap=cm.seismic_r,vmin=-0.
 ↪32,vmax=0.32, linewidth=0.5)
cb = fig.colorbar(p,shrink = .7)
plt.title('Reconstruction with all data')
plt.show()
print('Fidelity with a single photon is ',fidelity(q_tomo1,fock(N_dim,1))**2)

fig = plt.figure(figsize=(16,10))
# `ax` is a 3D-aware axis instance, because of the projection='3d' keyword␣
 ↪argument to add_subplot
ax = fig.add_subplot(2, 2, 2)
p = ax.contourf(X, Y, W2, 80, cmap=cm.seismic_r,vmin=-0.32,vmax=0.32)
#cb = fig.colorbar(p, shrink = 0.7)
cb = fig.colorbar(p, shrink = 1)
cb.ax.tick_params(labelsize=16)
cb.set_ticks([-0.3,-0.2,-0.1, 0,0.1,0.2,0.3]);

# surface_plot with color grading and color bar
ax = fig.add_subplot(2, 2, 1, projection='3d')
p = ax.plot_surface(X, Y, W2, rstride=1, cstride=1, cmap=cm.seismic_r,vmin=-0.
 ↪32,vmax=0.32, linewidth=0.5)
cb = fig.colorbar(p,shrink = .7)
```

```
plt.title('Reconstruction with phase-averaged data')
plt.show()
print('Fidelity with a single photon is ',fidelity(q_tomo2,fock(N_dim,1))**2)
```




```
Fidelity with a single photon is  0.441096175995693
```




```
Fidelity with a single photon is  0.4430140560643
```

The above Wigner functions do not dip negative since the overall loss in this run of the experiment slightly exceed 50%. While the phase-averaged data more closely resembles the symmetry of the Fock state, the whole data set is not too bad. The asymmetries originate from experimental imperfections such as finite sampling and amplitude fluctuations over the course of data collection (~6 hours).

Fortunately, loss is a statistical process that uniquely maps one density operator to another loss-degraded density operator. As discussed in the main text, another SDP optimiztion problem can be solved to determine the true density matrix before loss. First, define 'matrix_L' which is the matrix that applies loss to the density matrix. Here, loss was as high as 60%.

[10]:
```
refl_tot=0.6
'''Define the Matrix to correct the loss on the reconstructed \rho' '''
#dim_in=N_dim  #Hilbert space dimension cuttoff of the input quantum state
dim_out=N_dim #Hilbert space dimentsion cuttoff of the output (measured) state
r=np.sqrt(refl_tot) #loss BS reflection coefficient
t=np.sqrt(1-refl_tot) #loss BS transmission coefficient

#Define my 'ragged' matrix that gives me the transformation between my initial␣
 ↪state
#and my final state after the BS and trace.  Here, each layer gives a matrix␣
 ↪that
#maps ones of the diagonals of my new, measured density matrix to the same
#diagonal of my original matrix.

matrix_L=[]
for i in range(N_dim):
    M=[]
    for j in range(dim_out-i):
        row=[]
        for k in range(dim_out-i):
            if k < j: #ensures the result is upper triangular
                val=0
            else:
                val=(binom(k,k-j)*binom(i+k,k-j))**(1/
 ↪2)*(refl_tot)**((k-j))*(1-refl_tot)**(j+(i/2))
            row.append(val) #appends each value to the k^th position in the␣
 ↪j^th row
        M.append(row) #appends the j^th row to the i^th matrix
    matrix_L.append(M)

#For example, the main diagonals map according to
#rho_out.diagonal(0)=np.matmul(matrix_M[0],rho_in.diagonal(0))

#Note, this is NOT YET NORMALIZED!  I still need to do that in the optimization␣
 ↪algorithm

#rearrange the varibles into the same form that works with the SDP problem
rho_temp1=np.zeros((dim_out,dim_out),complex) #all data points
rho_temp2=np.zeros((dim_out,dim_out),complex) #phase-averaged data
for i in range(N_dim):
    for j in range(N_dim):
        rho_temp1[i][j]=P_arr1[i][j]
        rho_temp2[i][j]=P_arr2[i][j]
```

The next cell has the SDP algorithm used to correct for loss. As a different matrix must map each diagonal (not just the main diagonal) of the density matrix into a loss-degraded density matrix, this SDP problem actually optimizes over several matrix inversions. It is thus a bit more messy

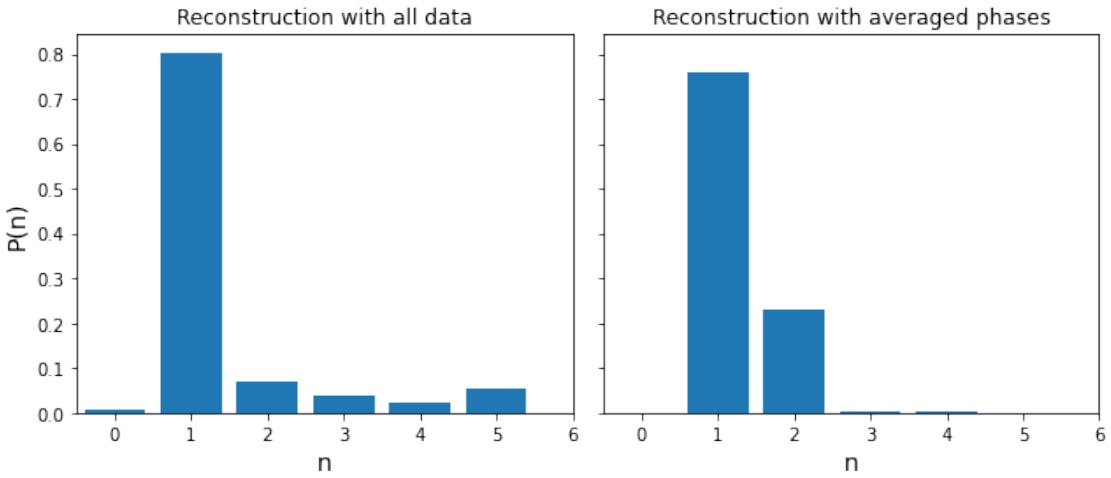
than before.

[15]:
```
'''Method using SDPs to reconstruct for loss, which uses the first matrix,␣
 ↪matrix_L as the loss matrix
which we need to find the inverse of'''
from cvxopt import blas, lapack, solvers

def Convex_optimization_loss_reconstruct(Matrix_map, Rho_measured, gamma,delta):
    M = np.asarray(Matrix_map);
    Rho = Rho_measured;
    Error1=0
    Error2=0
    P = cp.Variable((N_dim,N_dim), PSD = True) #set hermitian true when I␣
 ↪include phase
    #P = cp.Variable((N_dim,N_dim), complex = True)
    #P = cp.Variable((N_dim,N_dim), hermitian=True)
    norm=0

    shifted=[]
    for i in range(N_dim):
        row=[]
        for j in range(N_dim-i):
            row.append(P[j][i+j])
        shifted.append(row)
    #The above loop rearranges the variable matrix P into a new matrix where
    #the new rows are now the different diagonals of the original P matrix
    Error_vec=[]
    for k in range(N_dim):
        vec=[]
        for i in range(N_dim-k):
            num=0
            for j in range(N_dim-k):
                num=num+M[k][i][j]*shifted[k][j]
                #inner loop performs matrix multiplication between
                #one of the rows from the new shifted P matrix and
                #one of the matrices that needs to be inverted, M[k].
            vec.append(num)
            #This inner multiplications results in the vector vec
            #I also use this number to subract element-wise the corresponding
            #values of the diagonals of the measured Rho
            temp_error=cp.abs(num-np.diagonal(Rho,k)[i])
            Error1=Error1+cp.power(temp_error,2)



    shifted2=[]
    for i in range(N_dim):
```

```
        row=[]
        for j in range(N_dim-i):
            row.append(P[j+i][j])
        shifted2.append(row)

    for k in range(N_dim):
        vec=[]
        for i in range(N_dim-k):
            num=0
            for j in range(N_dim-k):
                num=num+M[k][i][j]*shifted2[k][j]
                #inner loop performs matrix multiplication between
                #one of the rows from the new shifted P matrix and
                #one of the matrices that needs to be inverted, M[k].
            vec.append(num)
            #This inner multiplications results in the vector vec
            #I also use this number to subract element-wise the corresponding
            #values of the diagonals of the measured Rho
            temp_error=cp.abs(num-np.diagonal(Rho,k)[i].conj())
            Error2=Error2+cp.power(temp_error,2)

    Obj_detect = cp.Minimize(Error1+Error2+ gamma*cp.norm(P,2))

    positive_diag=[]

    constraints = [cp.trace(P)==1,cp.diag(cp.real(P))>=0] #constrains the␣
↪density matrix to be physical
    for i in range(N_dim):
        for j in range(N_dim):
            if i==j:
                constraints.append(cp.abs(P[i][j]) <= (1/np.
↪sqrt(1-refl_tot)**(i+j))*cp.abs(Rho[i][j])+delta)
    Prob_detect = cp.Problem(Obj_detect,constraints)
    Prob_detect.solve(verbose = False) #set verbose = True to see output logs
    #Prob_detect.solve(cp.CVXOPT,verbose = False)

    p_values = (P.value)
    return p_values

Rho1=Convex_optimization_loss_reconstruct(matrix_L,rho_temp1,0.0,0.0) #all data
Rho2=Convex_optimization_loss_reconstruct(matrix_L,rho_temp2,0.0,0.0)␣
↪#phase-averaged data
```

[16]:
```
#Plots show the reconstructed photon number distributions
fig, ax = plt.subplots(1,2, sharey=True,figsize=(9,4))

ax[0].bar(range(N_dim),Rho1.diagonal(0))
```

```
ax[0].set_title('Reconstruction with all data')
ax[0].set_ylabel('P(n)',fontsize=14)
ax[0].set_xlabel('n',fontsize=14)
ax[0].set_xlim([-0.5,6])
#ax1.xlabel('n',fontsize=12)
ax[1].bar(range(N_dim),Rho2.diagonal(0))
ax[1].set_title('Reconstruction with averaged phases')
ax[1].set_xlabel('n',fontsize=14)
ax[1].set_xlim([-0.5,6])
plt.tight_layout()
plt.show()
```

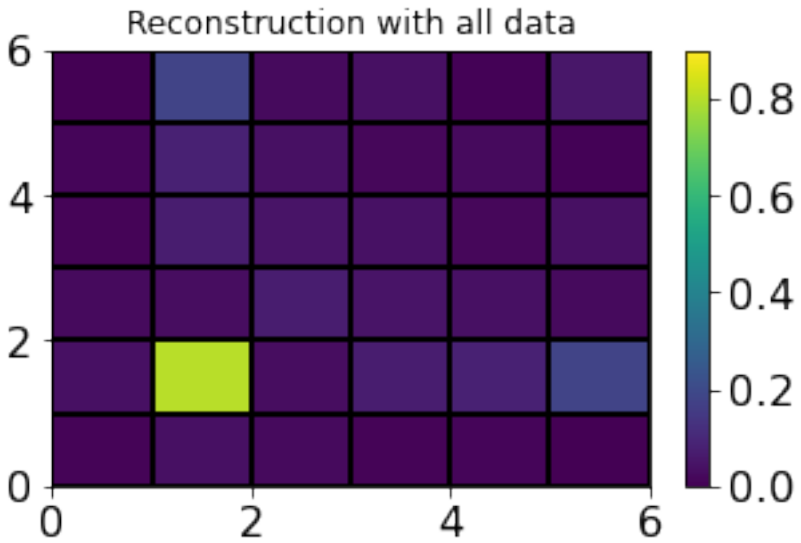


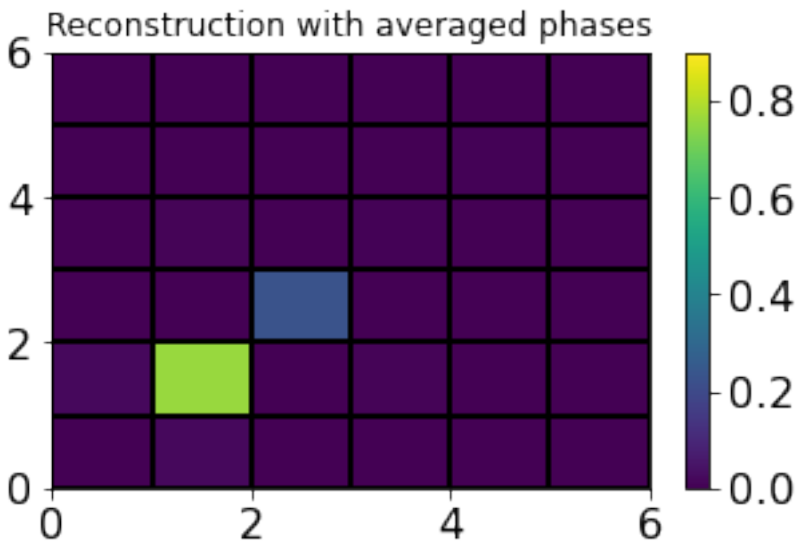

From the above probability distributions, it is clear that the loss reconstruction removed the vacuum component and produced a state that is considerably closer to the pure single-photon we expect. However, the relatively large 2 photon component may not be erroneous after all. In fact, the pump power was too high, so the 2 photon probability in reality was likely non-negligible. Thus, the tomography protocol was able to provide us information on the generated state.

We can also plot the density matrix elements, followed by Wigner functions of the full reconstruction with loss compensation.

[17]:
```
s1=np.abs(Rho1)
temp1=np.delete(s1,np.s_[6:],0)
state1=np.delete(temp1,np.s_[6:],1)

s2=np.real(Rho2)
temp2=np.delete(s2,np.s_[6:],0)
state2=np.delete(temp2,np.s_[6:],1)

fig = plt.figure(figsize=(5,3))
plt.pcolormesh(state1, vmin=0.0, vmax=.9,edgecolor='k', linewidth=1)
```

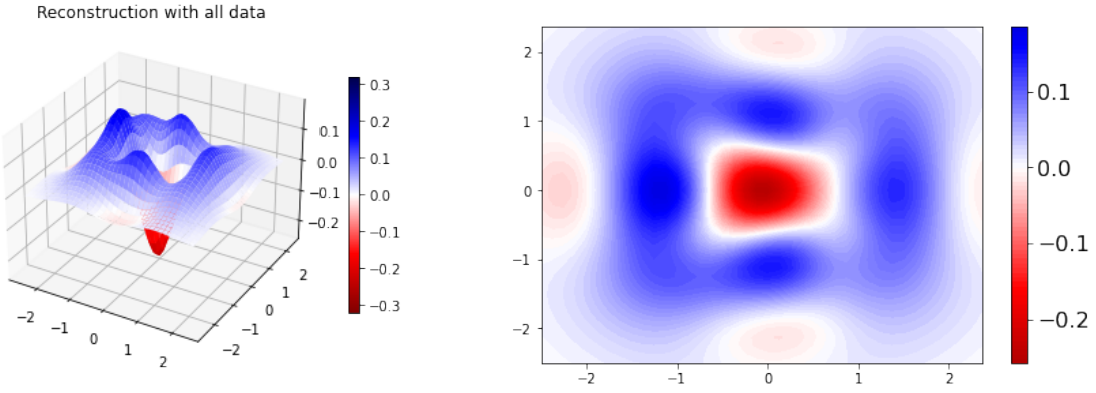
```
cbar=plt.colorbar()
cbar.ax.tick_params(labelsize=16)
plt.xticks([0,2,4,6])
plt.tick_params(labelsize=16)
plt.title('Reconstruction with all data')
plt.show()

#plt.imshow(state)
fig = plt.figure(figsize=(5,3))
plt.pcolormesh(state2, vmin=0.0, vmax=.9,edgecolor='k', linewidth=1)
cbar=plt.colorbar()
cbar.ax.tick_params(labelsize=16)
plt.xticks([0,2,4,6])
plt.tick_params(labelsize=16)
plt.title('Reconstruction with averaged phases')
plt.show()
```



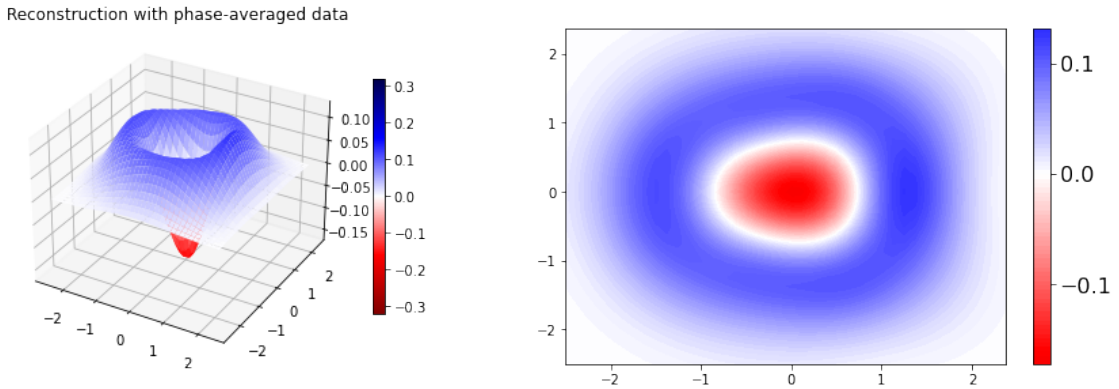



[18]:
```
xvec = np.arange(-20.,20.)*5./40
yvec = np.arange(-50.,50)*5/40
X,Y = np.meshgrid(xvec, xvec)
X1,Y1 = np.meshgrid(yvec,yvec)

q_tomo1=Qobj(Rho1)
q_tomo2=Qobj(Rho2)
W1=wigner(q_tomo1,xvec,xvec)
W2=wigner(q_tomo2,xvec,xvec)

fig = plt.figure(figsize=(16,10))
# `ax` is a 3D-aware axis instance, because of the projection='3d' keyword␣
 ↪argument to add_subplot
ax = fig.add_subplot(2, 2, 2)
p = ax.contourf(X, Y, W1, 80, cmap=cm.seismic_r,vmin=-0.32,vmax=0.32)
#cb = fig.colorbar(p, shrink = 0.7)
cb = fig.colorbar(p, shrink = 1)
cb.ax.tick_params(labelsize=16)
cb.set_ticks([-0.3,-0.2,-0.1, 0,0.1,0.2,0.3]);

# surface_plot with color grading and color bar
ax = fig.add_subplot(2, 2, 1, projection='3d')
p = ax.plot_surface(X, Y, W1, rstride=1, cstride=1, cmap=cm.seismic_r,vmin=-0.
 ↪32,vmax=0.32, linewidth=0.5)
cb = fig.colorbar(p,shrink = .7)
plt.title('Reconstruction with all data')
plt.show()
print('Fidelity with a single photon is ',fidelity(Qobj(Rho1),fock(N_dim,1))**2)
```

```
fig = plt.figure(figsize=(16,10))
# `ax` is a 3D-aware axis instance, because of the projection='3d' keyword␣
 ↪argument to add_subplot
ax = fig.add_subplot(2, 2, 2)
p = ax.contourf(X, Y, W2, 80, cmap=cm.seismic_r,vmin=-0.32,vmax=0.32)
#cb = fig.colorbar(p, shrink = 0.7)
cb = fig.colorbar(p, shrink = 1)
cb.ax.tick_params(labelsize=16)
cb.set_ticks([-0.3,-0.2,-0.1, 0,0.1,0.2,0.3]);

# surface_plot with color grading and color bar
ax = fig.add_subplot(2, 2, 1, projection='3d')
p = ax.plot_surface(X, Y, W2, rstride=1, cstride=1, cmap=cm.seismic_r,vmin=-0.
 ↪32,vmax=0.32, linewidth=0.5)
cb = fig.colorbar(p,shrink = .7)
plt.title('Reconstruction with phase-averaged data')
plt.show()
print('Fidelity with a single photon is ',fidelity(Qobj(Rho2),fock(N_dim,1))**2)
```




```
Fidelity with a single photon is  0.8034190879882358
```



```
Fidelity with a single photon is  0.7594397813576096
```