Strategies for noisy photonic quantum technologies: quantum computation to quantum key distribution

by

Joseph Eli Bourassa

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

Abstract

Strategies for noisy photonic quantum technologies: quantum computation to quantum key distribution

Joseph Eli Bourassa
Doctor of Philosophy
Graduate Department of Physics
University of Toronto
2021

This thesis focuses on photonic implementations of two quantum technologies: quantum computers and quantum key distribution (QKD). Part I concentrates on photonic quantum computation using a qubit encoding for continuous-variable (CV) systems such as photonic modes known as Gottesman-Kitaev-Preskill (GKP) states. GKP states would be an advantageous encoding for photonics due to their error-correcting properties in the CV space and due to the experimental accessibility of gate and measurement implementations. This thesis advances the prospect of photonic quantum computation with GKP states on three fronts: first, to address the outstanding issue of how to prepare these qubits in optics, we analyze a proposal for their generation using two experimentally-accessible resources, multimode squeezed states and photon-number resolving detectors. Next, to overcome the difficulty of simulating GKP states in the Fock basis, a phase space simulation technique is introduced to understand realistic noise on these qubits. Lastly, to determine the potential of GKP qubits for fault-tolerant photonic quantum computation, a measurement-based architecture with GKP states is analyzed to determine thresholds for the quality of state required.

Part II shifts focus to QKD, and how security proofs can account for realistic device imperfections. We first examine entanglement-based QKD protocols that employ high-dimensional photonic degrees of freedom (e.g. time-frequency or electric field quadrature) to produce more bits of key per optical signal. While entanglement-based protocols circumvent security assumptions about the quantum state source, they come at the cost of having to trust one's detectors. This thesis analyzes a previously-identified detector loophole for these schemes, and provides a modified security proof that succeeds in establishing security when encoding in electric field quadrature, with limited success for time-frequency encoding. Luckily, an alternative protocol known as measurement-device-independent (MDI) QKD requires no security assumptions about detectors. Instead, MDI QKD requires trusted sources of quantum states, which are nonetheless subject to their own imperfections that can compromise security. In the remainder of the thesis, we advance security proofs for MDI QKD in the presence of two potential source flaws: noise that introduces mixedness to the quantum states, and side-channels that leak information about the encoding choices of states.

Acknowledgements

There are many people who guided my path to this specific point in my education.

Thank you to my parents and grandparents for creating an environment in which education was a virtue and academic success celebrated, instilling that obtaining a PhD would be an achievable and worthwhile pursuit. Thank you to my brother for paving the way through our school, setting an incredibly high bar for me to attempt to match. Thank you to my maternal grandparents—a professor of engineering and biologist—who sparked a curiosity for science (and art), whether it was gifting me a microscope, explaining the ocean tides with a plate of water, or guiding me for countless hours through museums to teach me beyond the classroom.

In grade 10, I thought I wanted to be a medical doctor, likely from watching too many TV hospital dramas. Thank you to my physics and math teacher, Mr. Gaalaas, for presenting an alternative option; I learned from you the frustration and incomparable reward of breaking through puzzles that don't simply have an answer at the back of the book. Thank you to my chemistry teacher, Ms. Sunil, for introducing me to some of the first concepts in quantum physics, like the Heisenberg uncertainty principle; I think I get it now[1]. Thank you to my English and philosophy teacher, Mr. Tucker, for teaching me to write and to think critically about how knowledge becomes established; I apologize for all the typos in this thesis.

In my first year of undergrad, my fate was sealed after taking Professor Paul Brumer's seminar, The Quantum World and its Classical Limit. Thank you to Professor Brumer for opening this world, for mentoring me and investing in my academic development throughout my undergraduate degree via our many blackboard meetings, accompanied by an espresso (or an instant coffee, in your case). Thank you to Professor John Sipe and Dr. Zachary Vernon for supervising my undergraduate thesis in physics and supplying a solid foundation in quantum optics. Thank you to Kyle and Stepan for the many hours studying together; with three (already successful or soon-to-be) PhDs, it paid off.

Throughout graduate school, I have had the opportunity to learn from the best. First and foremost, thank you to my supervisor, Professor Hoi-Kwong Lo, for introducing me to the potentials of quantum technology. I am incredibly grateful for your support and guidance throughout my research projects, and for your wisdom and sage advice as I navigate this field. Thank you to my supervisory committee, Professor Li Qian and Professor Daniel James, and the many colleagues (special thanks to Amita, Mike, Tom, Aaron, Xiaoqing, and Mattia) for the insightful discussions. Thank you to the third floor team at McLennan for all the behind-the-scenes work to ensure the academic program ran smoothly. Thank you to Professor Charles Lim for teaching me new ways to think about QKD, and to you and your research group for welcoming me to CQT/Singapore. Thank you to Christian Weedbrook for bringing me onto the team at Xanadu. Thank you to Krishna for introducing me to bosonic qubits, but especially for your mentorship and all the time you invested in my growth as a researcher (despite time invested, the same cannot also be said of my growth as a ping pong player). Thank you to the many colleagues at Xanadu (special thanks to Ish, Nico, and Guillaume) for providing a fruitful environment for the exciting pursuit of a quantum computer.

Thank you to Ilan Tzitrin, a colleague, officemate and friend, who gets his own paragraph. Thank you for being such a delight to work with on numerous research projects and papers, for making our offices at U of T and Xanadu an enjoyable place to go to each day (when we could), and for countless hours spent joking (and doing some physics) on the phone during the past year of lockdowns to maintain a sense of normalcy.

[1] Although Feynman might disagree.

While I spent the last five years pursuing a PhD, there are people to thank for the support that they would have provided regardless of how I spent my time. Thank you to my friends for providing much needed breaks and distractions from work, be it BBQ feasts, bike rides, flag football or tennis games, obscure films at TIFF, trips to Vancouver, or many phone and Zoom calls.

Thank you to my family and to my partner's family for the care packages of food; the time spent around the dinner table (when we could and when we will again); the reunions in Victoria; asking "How's Science?"; the jam sessions; the games of Catan, Scrabble, and cryptic crosswords; the time by the pool and lake; and for the constant love and encouragement.

Finally, my partner, Steph, deserves a list of thanks at least as long as this thesis, so I will simply say: thank you for being the source of so much happiness in my life.

Contents

List of Tables

List of Figures

List of Commonly Used Acronyms

CPTP	Completely positive, trace-preserving
CV	Continuous-variable
GBS	Gaussian boson sampling
GKP	Gottesman-Kitaev-Preskill
IM	Intensity modulator
MBQC	Measurement-based quantum computation
MDI	Measurement-device independent
MWPM	Minimum weight perfect matching
PM	Phase modulator
PNR	Photon number resolving detector
POVM	Positive operator-valued measure
PSD	Positive semidefinite
QEC	Quantum error correction
QKD	Quantum key distribution
QPU	Quantum processing unit
RHG	Raussendorf, Harrington, Goyal
RSA	Rivest, Shamir, and Adleman
SDP	Semidefinite programming
SPD	Single photon detector
WCP	Weak coherent pulse

List of papers and presentations

Publications

- **J. E. Bourassa***, R. N. Alexander*, M. Vasmer, A. Patil, I. Tzitrin, T. Matsuura, D. Su, B. Q. Baragiola, S. Guha, G. Dauphinais, K. K. Sabapathy, N. C. Menicucci, and I. Dhand, "Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer", *Quantum*, vol. 5, p. 392, Feb. 2021, *These two authors contributed equally.

- **J. E. Bourassa**, I. W. Primaatmaja, C. C. W. Lim, and H.-K. Lo, "Loss-tolerant quantum key distribution with mixed signal states", *Phys. Rev. A*, vol. 102, p. 062 607, 6 Dec. 2020.

- I. Tzitrin*, **J. E. Bourassa***, N. C. Menicucci, and K. K. Sabapathy, "Progress towards practical qubit computation using approximate Gottesman-Kitaev-Preskill codes", *Physical Review A*, vol. 101, no. 3, Mar. 2020, *These two authors contributed equally.

- **J. E. Bourassa** and H.-K. Lo, "Entropic uncertainty relations and the measurement range problem, with consequences for high-dimensional quantum key distribution", *Journal of the Optical Society of America B*, vol. 36, no. 3, B65, Feb. 2019, Editor's Pick for feature issue *Quantum Key Distribution and Beyond*.

Preprints

- **J. E. Bourassa***, A. Gnanapandithan*, L. Qian, and H.-K. Lo, "Measurement device-independent quantum key distribution with passive, time-dependent source side-channels." arXiv:2108.08698, 2021, *These two authors contributed equally.

- I. Tzitrin, T. Matsuura, R. N. Alexander, G. Dauphinais, **J. E. Bourassa**, K. K. Sabapathy, N. C. Menicucci, and I. Dhand. "Fault-tolerant quantum computation with static linear optics." arXiv:2104.03241, 2021.

- **J. E. Bourassa**, N. Quesada, I. Tzitrin, A. Száva, T. Isacsson, J. Izaac, K. K. Sabapathy, G. Dauphinais, and I. Dhand. "Fast simulation of bosonic qubits via Gaussian functions in phase space." arXiv:2103.05530, 2021.

Talks

- "Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer", AWS Center for Quantum Computing, Spring 2021, co-presented with Ilan Tzitrin.
- "Fast simulation of bosonic qubits via Gaussian functions in phase space", AWS Center for Quantum Computing, Spring 2021.
- "The loss tolerant protocol with a twist", IQC Theory talks, Institute for Quantum Computing, Spring 2021.
- "The loss tolerant protocol with a twist", Workshop on Security Proofs in QKD, Institute for Quantum Computing, Summer 2020.
- "Finite-energy GKP grid states for CV quantum computing", Xanadu-University of Arizona Workshop on CV Photonic Quantum Computing, Winter 2020, co-presented with Ilan Tzitrin.
- "Measurement-range loophole in high-dimensional QKD", Workshop on Security Proofs in QKD, Institute for Quantum Computing, Summer 2018.

Posters

- "Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer", Bristol Quantum Information Technologies Workshop, online conference, Spring 2021.
- "The loss tolerant protocol with a twist", QCRYPT, online conference, Summer 2020.
- "Measurement-range loophole in high-dimensional QKD", Quantum Internet Workshop, University of Toronto, Summer 2018.
- "Security implications of pre-measurement filters in time-frequency QKD", Quantum Information Processing, TU Delft, Winter 2018.

Chapter 1

Introduction

Quantum technologies refer to devices and protocols which control and preserve the quantum properties of a physical system so they can be leveraged to perform tasks inaccessible to traditional classical technologies. In this thesis, we will concern ourselves with two prominent quantum technologies: quantum computers and quantum key distribution (QKD). At a high level, quantum computers prepare quantum data registers or qubits, apply transformations or gates, and perform measurements with probabilistic outcomes to read out classical data [2]. While it is difficult to pinpoint the exact confluence of quantum mechanical properties which lead to the power of quantum computers, one often invokes entanglement, superposition, and interference: entanglement allows for the combined state-space of qubits to grow exponentially as more qubits are added; transformations in the space of entangled qubits create tuned superpositions of multi-qubit states; interference between probability amplitudes in the superposition allows one to efficiently extract measurement outcomes that would otherwise be inefficient to sample with a classical computer. If realized, quantum computers could outperform their classical counterparts in tasks such as search algorithms [3], machine learning [4], simulation of quantum systems relevant to material and pharmaceutical design [5], and prime factorization [6], the last of these having significant repercussions for conventional cryptography.

A class of cryptographic protocols such as the RSA (Rivest, Shamir, and Adleman) cryptosystem [7] underpin current infrastructure for encryption, and rely on the difficulty of performing prime factorization with classical computers [2]. As prime factorization would be efficient on a quantum computer using Shor's algorithm [6], an alternative scheme for encryption is required. This is true not just for future communications, but even for present communications that require long-term secrecy, since an eavesdropper could harvest encrypted communication now and decrypt the information at a later date when quantum computers are available. While quantum computers could break traditional cryptography, quantum mechanics also offers a solution in the form of QKD. Broadly speaking, in a QKD protocol, two trusted but spatially separated parties, Alice and Bob, share a quantum system in the presence of an eavesdropper, Eve. After measurements are performed on the quantum system, Alice and Bob communicate classically about a subset of their results. Ultimately, Alice and Bob either distill a shared, secret key from their measurement results that they can then use to encrypt messages, or detect tampering by Eve and abort the protocol. We can appeal to several principles of quantum mechanics to justify the security of QKD [8]: first, there is the notion that measurement of a quantum system, given its probabilistic outcome, disturbs the state of the system, so tampering by Eve will leave traces that Alice and Bob can detect. Second, a consequence of quantum superpositions and unitary evolution is the no-cloning theorem [9], which states it is impossible to make a perfect copy of an unknown

quantum state, placing limitations on Eve's ability to replicate the quantum system from which Alice and Bob are extracting their secret key. Finally, if, from the eavesdropper's perspective, Alice and Bob share an entangled quantum system, then from the concept of monogamy of entanglement, there are limitations on how entangled Eve can be with that system, so she will not be privy to all the measurement results extracted from it [10–12].

While we have so far discussed these quantum technologies in the abstract, the question now becomes which physical platform capable of exhibiting and preserving quantum properties do we employ to build quantum computers and QKD systems? This thesis will focus on photonic quantum technologies, i.e. implementations which employ modes of light and their quantum optical properties. For QKD, photonics is currently the only choice, since photons travel at the speed of light and can retain their encoding even after travelling hundreds of kilometers through fiber or free space, which makes them ideal for distributing a quantum system between spatially-separated Alice and Bob. Even so, a mode of light is a rich physical system capable of encoding information in various degrees of freedom, so the game of designing a protocol becomes choosing encodings robust to noise, namely loss, and designing security proofs that account for various optical source and measurement device imperfections and loopholes.

For quantum computers, the choice of photonics as a platform is not as immediately obvious. To justify this choice, we must understand the terms of the game: not only do we want a *universal* quantum computer capable of performing any user-provided algorithm, we want the quantum computer to be *fault-tolerant*, meaning it can still provide accurate readout in the presence of qubit, gate and measurement noise. While one could continue refining the quality of each physical qubit, the likelier path to fault tolerant quantum computing is to redundantly encode a smaller number of high-quality logical qubits into many noisy physical qubits, using error correction to detect and correct errors in the physical qubits to yield low error rates in the logical qubits. Here, photonics presents several advantages. First, photonics opens the door to room-temperature computation, since photons can maintain their encoding even at room-temperature, as opposed to other platforms which may require cryogenics; this can prove advantageous when scaling to millions of physical qubits, as will likely be required to perform fault-tolerant quantum computation. Second, photonics provides a rich physical system for encoding qubits, both due to the many ways to encode a two-dimensional system within the inherently infinite-dimensional space of an optical mode, and due to the networkability of optical modes to encode many qubits into error-correcting codes; flexibility of error-correcting encoding can prove useful for combating noise and errors that naturally creep into quantum systems. Finally, as was the case for QKD, photonics is a natural platform for communication technology, so networking photonic quantum computers is likely to be simpler than having to transduce quantum information between two types of platforms.

More specifically, the first part of this thesis explores the prospect of performing photonic quantum computation using a choice of qubit encoding known as Gottesman-Kitaev-Preskill (GKP) states [13]. While we will dive into the details of GKP states in the chapters that follow, here I briefly summarize two reasons GKP qubits are a desirable choice of encoding for photonics:

- A wide class of important quantum gates and measurements—Clifford gates and Pauli measurements—correspond to a set of operations and measurements—linear optics, squeezing, and homodyne—that are experimentally accessible in the photonics context [13][1]. This means that two-qubit entangling gates (among others) are deterministic for these qubits, a clear distinction as compared to encoding a qubit in a dual-rail or single-photon polarization-based scheme [14]. There are also many error-correcting

[1] More generally stated, Clifford transformations correspond to Gaussian transformations.

codes, i.e. multi-qubit entangled states, that only require Clifford gates and Pauli measurements for encoding and error-correction, which makes the prospect of encoding many noisy physical qubits into a logical qubit more tenable.

- Ideally, GKP states allow one to correct for small displacement errors in phase space [13], and since the displacement operators form a basis to decompose any operator in the continuous-variable (CV) Hilbert space of the mode, this means even noise channels like loss can be decomposed into errors GKP states are designed to correct. In fact, GKP states are better at correcting for photon loss than encodings specifically designed to correct for that error [15]. Larger displacement errors only result in bit and phase flips, and can be handled by encoding many GKP qubits into an error-correcting code.

While GKP states present these advantages, the most onerous task for their use in photonics is their preparation. To start, their ideal forms are non-normalizable, infinite energy states, so we must first consider realistic, finite-energy approximations to the ideal state[2]. Even so, these approximate states are highly non-Gaussian[3], meaning a strong non-linear electric field interaction (generated from a Hamiltonian more than quadratic in the mode's quadrature operators [4]) would be required for their deterministic preparation, and unfortunately no such interactions are readily-available in optics. Instead, all-photonic GKP preparation schemes involve probabilistic but heralded state generation devices. To this end, Chapter 2, which is largely based on [16], provides a scheme for preparing GKP states using Gaussian boson sampling (GBS), along with tools for assessing the quality of the states prepared. Briefly, the GBS method consists of constructing a multimode Gaussian state by sending squeezed vacuum states into a linear interferometer, and then measuring all but one of the modes with photon-number-resolving detectors (PNRs) [17–20]. The chapter presents extensive analysis of the trade-off between generating a GKP state with a certain fidelity and the probability of it being produced.

Chapter 3 (based on [21]) introduces a formalism and method for simulating GKP states, along with other bosonic qubits like cat states and Fock states, as linear combinations of Gaussian functions in phase space. This framework leverages the rich Gaussian formalism for CV quantum systems, while extending it to states that are useful for quantum computing. The framework allows us to study realistic gates and noise that would be more cumbersome to simulate and understand using competing methods that leverage the Fock basis.

Building on the understanding developed in those two chapters, Chapter 4, which is based on [22], presents a blueprint for a scalable, fault-tolerant photonic quantum computer that leverages GKP states. A key result of the proposed architecture is that it is compatible with probabilistic sources of GKP states, such as those seen in Chapter 2 and which are to be expected in the near-term. The main focus of the chapter is the

[2] A keen reader can skip ahead to Figs. 2.1a and 2.1b to see what the wavefunctions of GKP states look like.

[3] Non-Gaussian refers to the fact that their Wigner quasiprobability distributions are not Gaussian functions, meaning that they cannot simply be constructed using Gaussian states (e.g. coherent, squeezed states), Gaussian transformations (e.g. linear optics, squeezing), and Gaussian measurements (e.g. homodyne, heterodyne).

[4] A CV quantum system can be characterized by two quadrature operators, typically labelled position $\hat{q}$ and momentum $\hat{p}$, nomenclature inherited from the mechanics of a point particle. These operators respect the Heisenberg uncertainty relation $[\hat{q}, \hat{p}] = i\hbar$. Additionally, in the same spirit as classical mechanics, one can define a phase space representation of the CV system (such as the Wigner function) where the two quadrature operators of Hilbert space correspond to two quadrature variables of phase space; for the mathematical details of the mapping, refer to Section 3.2.1. A main difference from the classical phase space representation is that the quantum phase space representation is not a probability distribution, stemming from the fact that $\hat{q}$ and $\hat{p}$ must respect the uncertainty relation. For a photonic mode, the CV system corresponds to the electromagnetic field. Since the field strength in an electromagnetic wave propagating in linear media oscillates periodically, once quantized, the field can be expressed in terms of two operators, effectively corresponding to the electric and magnetic fields. These are often termed the quadratures of the electromagnetic field, with dynamics analogous to those of position and momentum for a quantum harmonic oscillator.

noise analysis and quantum error correction scheme for the architecture. The chapter provides threshold values for noise and state generation probability required for fault-tolerance, setting important milestones and identifying areas of improvement for a road map to photonic quantum computing.

With the security threat posed by quantum computers as motivation, the second part of this thesis provides my contributions to the field of QKD. Specifically, it will explore how security proof techniques can be used to mitigate noise in quantum key distribution protocols employing realistic experimental devices. Chapter 5 (based on [23]) examines large-alphabet, entanglement-based QKD protocols. "Large-alphabet" refers to schemes which employ a high-dimensional photonic degree of freedom for encoding key, such as time-frequency or electric field quadrature, while "entanglement-based" denotes protocols which employ an untrusted source for distributing entangled quantum systems but trusted measurement devices for characterizing the entanglement and extracting key. Specifically, the chapter studies the measurement-range problem, wherein realistic detectors can only measure a finite range of the degree of freedom being employed, a loophole which compromises one of the best tools—the entropic uncertainty relations—for quantifying security in large-alphabet protocols. The chapter provides a modified security proof to address this problem, remedying the loophole for protocols encoding in electric field quadrature; however, even with the improved security bound, we find that loss hampers the feasibility of entanglement-based time-frequency QKD.

Having seen the potential for detector loopholes to compromise the security of entanglement-based QKD in Chapter 5, in Chapters 6 and 7 we move to study measurement-device-independent (MDI) QKD protocols. In MDI QKD, Alice and Bob send quantum states associated with key values to an untrusted central node which makes an announcement about the result of a Bell state measurement it may or may not have faithfully executed [24]. In Chapter 6, which is based on [25], we study the case when the states which Alice and Bob prepare are trusted but noisy qubit signals, i.e. mixed states for which they hold the purification. We adapt a leading proof technique [25] to treat this scenario, determining that Alice and Bob can invoke fictitious shield systems and perform virtual—i.e. entirely software-based—"twisting" operations on these shields to decrease Eve's knowledge of the secret key. We use an optimization technique known as semidefinite programming to determine the optimal twisting operation Alice and Bob should apply.

In Chapter 7, which is based on [26], we study the issue of non-trivial, passive[5] side-channels in the optical sources Alice and Bob use for MDI QKD which leak information about their choice of encoding to Eve. We identify a time-dependent side-channel in a common polarization-based QKD source that employs a Faraday mirror for phase stabilization. We use it as a representative case example, applying a recently-developed numerical proof technique [27] to quantify the sensitivity of the secret key rate to the quantum optical model for the side-channel, and to develop strategies to mitigate the information leakage.

Each chapter is relatively self-contained, including relevant background material and conclusions. Nonetheless, in Chapter 8, I provide a summary of key results and an outlook on what research avenues this thesis opens.

[5]Here, "passive" means not controlled by the eavesdropper.

Part I

Photonic Quantum Computing Using Gottesman-Kitaev-Preskill (GKP) States

Chapter 2

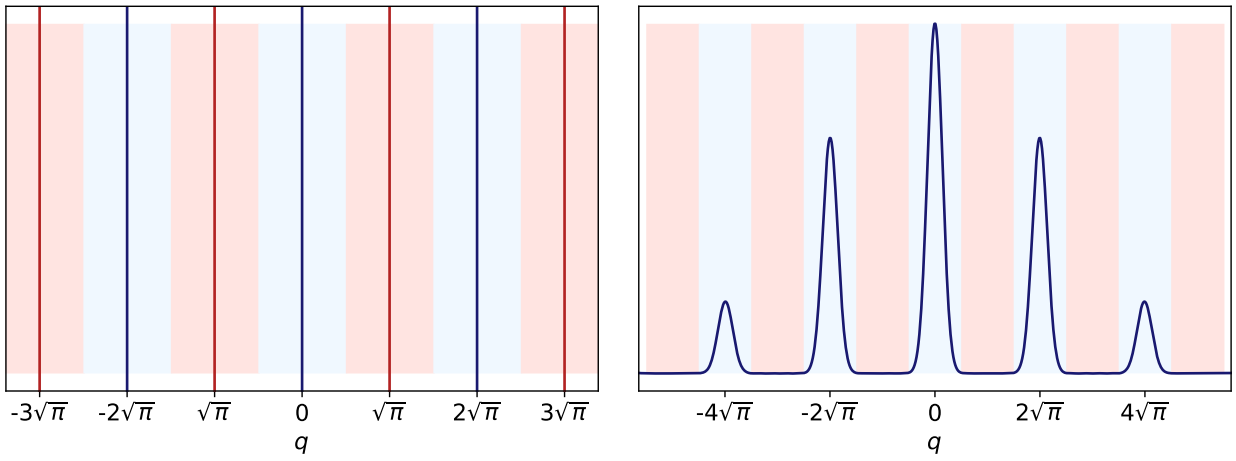
A Photonic State Preparation Method for Approximate GKP States

This chapter is based on [16], co-authored with Ilan Tzitrin, Nicolas C. Menicucci and Krishna K. Sabapathy. Krishna K. Sabapathy initiated and supervised the project. The work was collaborative and published in *Physical Review A*. Ilan Tzitrin and I shared first authorship. My main contributions were to the numerical simulations throughout the paper, analysis of the results, and to writing the third part of the published paper (Section 2.3 of this chapter) and the appendices. Additionally, I have condensed the second part (Section 2.2) of the work as compared to the publication, and rewritten the introduction presented here to better reflect the most recent literature on the subject, since the paper was written in 2019 and it has been a very active area of research. The work benefited from helpful discussions with Hoi-Kwong Lo, Saikat Guha, Arne L. Grimsmo, Victor V. Albert, and colleagues at Xanadu.

2.1 Introduction

Bosonic qubits refer to encoding two-dimensional quantum systems into the states of continuous-variable (CV) systems such as photonic modes. Among the most well-known bosonic qubits are Gottesman-Kitaev-Preskill (GKP) [13] states (occasionally referred to as grid states). Originally constructed to correct displacement errors in phase space with the help of inherent translation symmetry, GKP states are viewed as a promising qubit encoding for various quantum technologies such as quantum communication [15, 28–30], and quantum sensing [31, 32], and our main interest for this part of the thesis: fault-tolerant quantum computation [13, 33–46].

While we will provide an in-depth review of the formalism of GKP states in the next section, we first survey the major milestones in their study. In [13], the authors introduced the GKP encoding: a wide class of states for quantum harmonic oscillators, with wavefunctions defined as infinite superpositions or combs of quadrature eigenstates, as illustrated in Fig. 2.1a. They provided a set of universal gates for these states to be used for quantum computation, along with a protocol for how such states could correct small displacements in phase space or reduce larger displacements to errors that can be corrected by concatenating many GKP states into a qubit error-correcting code. They demonstrated that the states could be generalized to encode qudits, and tailored to correct either symmetric or biased displacement errors in phase space. Finally, the authors noted that while the ideal GKP states are elegant in their properties, they have infinite energy and



(a) Ideal GKP wavefunctions

(b) Finite-energy $|0\rangle_{\text{gkp}}$ wavefunction

Figure 2.1: In (a), ideal GKP qubit wavefunctions are presented. They consist of combs of quadrature eigenstates, with $|0\rangle_{\text{gkp}}$ ($|1\rangle_{\text{gkp}}$) given in blue (red). The shading represents the maximum tolerable displacement for GKP error correction; larger displacements introduce a bit flip that can be corrected by concatenating GKP qubits into a qubit error-correcting code. In (b), a version of the finite energy $|0\rangle_{\text{gkp}}$ is given, where each quadrature eigenstate in the superposition is replaced by a squeezed state, along with an overall envelope. The shading associated with the ideal states is shown to emphasize that a small but inherent error of mistaking 0 for 1 exists when using finite-energy GKP states.

are therefore unphysical; however, they defined finite-energy versions of GKP states, as shown in Fig. 2.1b, by replacing the quadrature eigenstates in the comb with narrow Gaussian functions (i.e. squeezed states that are often referred to as peaks), along with applying an overall Gaussian envelope. Finite-energy GKP states are certainly noisier than the ideal, but are still sufficient for fault-tolerant quantum computing. A thorough analysis of the GKP error correction protocol using approximate states was later provided in [33].

A resurgence of interest in the GKP encoding stems from the demonstration that finite-energy GKP states could be used for fault-tolerant quantum computing, where ~ 20 dB was determined to be a sufficient (but not necessary) level of squeezing for each squeezed state in the comb [34], and from the demonstration that GKP states are more robust at protecting against errors incurred by loss than encoding in single photons or even in states specifically designed to protect against loss [15]. Also demonstrated in [34] was the ability for GKP states to interface with CV cluster states—entangled, multimode squeezed states that are experimentally feasible to produce in photonics—meaning CV cluster states could be used as infrastructure to teleport GKP qubits for the purposes of measurement-based quantum computation, a prospect that elevated the potential of GKP states for photonic quantum computing.

Several studies have investigated the performance of GKP states concatenated into qubit error-correcting codes [22, 35, 38, 39, 41, 42, 44–48], with the most recent works demonstrating that a per-peak squeezing close to ~ 10 dB is likely sufficient for fault-tolerance. While some of these works may be tailored to different platforms—mainly photonics and superconducting cavities—a recurring theme is that since the qubits are encoded in a CV system, the measurements performed on states during error correction collect analog data which can be used to augment the standard qubit error correction procedures that would normally only use binary data. As a consequence, GKP states provide an extra layer of protection for encoding a qubit, since one can exploit both the redundancy of the CV space, as well as the redundancy of many qubits in an

error correcting code. Other notable works have focused on reducing the resource overheads for universal computation to just GKP states and Gaussian resources [36, 49] (i.e. Gaussian states including vacuum, coherent and squeezed states; linear optics; and homodyne measurements), and on the rich mathematical structure of GKP states [21, 50–53].

With GKP states holding such promise, a major push has been to devise and implement schemes for their preparation in various platforms. While GKP qubits have been realized experimentally in the motional degree of freedom of trapped ions [54–56] and in microwave cavities coupled to superconducting qubits [40], GKP states have yet to be produced experimentally in the optical domain. Notably, optical systems currently lack strong non-linear (third order or greater) interactions that have been central to the deterministic generation schemes for GKP states in these other platforms. As will be made more clear, GKP states are highly non-classical states, so they cannot be prepared simply using the Gaussian states (coherent and squeezed states), linear optical transformations, and homodyne/heterodyne measurements that are readily-accessible in optics. Instead, near-term optical demonstrations of GKP states will likely rely on heralded, probabilistic generation protocols leveraging the non-Gaussian resource provided by threshold and/or photon number-resolving measurements (PNRs), ever-maturing devices with high efficiency [57].

An early proposal [58] for optical GKP states relied on steady availability of cat states—superpositions of coherent states—an arguably simpler, but still highly non-classical state that would itself require non-Gaussian resources to generate. The authors of [59] improved on this proposal, offering a probabilistic method to make cat states by interfering a single photon and coherent state at a beamsplitter and measuring one mode with a PNR detector. The authors of [60] proposed creation of optical GKP states based on nonlinear interaction of squeezed states with atomic ensembles; however, the scheme stops short of providing candidate atomic systems, and it is unclear if such a proposal could ultimately be scaled to produce many GKP qubits for a fault-tolerant device. More recently, a series of works [17–20] introduced a state preparation method based on Gaussian Boson Sampling (GBS). By passing N squeezed vacuum states through a linear optical interferometer consisting of beamsplitters and phase shifters, and using PNRs to measure $N-1$ of the modes, one can in general prepare bosonic qubits in the remaining mode. The level of squeezing, interferometer angles, and PNR outcome form optimization parameters for the output state, with the measurement outcome heralding the desired result. In [18–20], the authors found a proof-of-principle demonstration that GBS state preparation could be used to produce GKP states. The GBS method is closest in spirit to that from [59], but offers a more general framework that does not require the interim step of needing to first produce single photons and cat states, and can instead directly produce the desired GKP states.

In this chapter, we first focus on understanding finite-energy, approximate GKP states. After reviewing the formalism of GKP states, we explore how various platform-independent figures of merit can be used to benchmark the quality of GKP states for near-term devices. In particular, we use the recently-introduced modular subsystem decomposition [61], which has proven to be very useful in the study of GKP qubits [16, 62–64], to monitor and mitigate errors on the logical qubit. We then provide thorough analysis of the preparation of approximate GKP states in optics based on GBS state preparation, providing resource requirements and relating the resulting states to the figures of merit discussed in the first part of the chapter.

The chapter is organized as follows. In Sec. 2.2, we review the formalism of the GKP code independently of the platform for its implementation. This includes a synthesis of results for the ideal (non-normalizable) GKP states in 2.2.1 and the finite-energy GKP states in 2.2.2; an overview of the modular decomposition in 2.2.3 and its relevance to analyzing GKP states; a set of prescriptions for tracking, quantifying, and alleviating errors in computation induced by imperfect GKP states in 2.2.4; and a discussion of error correction and

recovery with the normalizable states in 2.2.5.

Sec. 2.3 is dedicated to the preparation of approximate GKP states in the optical domain. In 2.3.1 we review the GBS method for the generation of general non-Gaussian states; in 2.3.2 we discuss a framework for approximating GKP states; in 2.3.3 we characterize the approximate states using the formalism described in Sec. 2.2; in 2.3.4 we present our strategy for the preparation of these approximate GKP states with optical circuits and provide our numerical results for the optimization of these circuits; and finally, in 2.3.5, we give comments on incorporating experimental imperfections into our scheme.

2.2 Formalism of GKP Qubits

Our physical landscape consists of continuous-variable (CV) systems whose Hilbert space is $L^2(\mathbb{R}^n)$, the square-integrable functions defined over the space of n real variables corresponding to n-mode systems. Examples of continuous-variable systems include modes of an electromagnetic field, harmonic oscillator chains, phonon modes in materials, continuous modes of ion traps and superconducting circuits. Since we will not use extensive elements of the CV formalism in this chapter, we leave a more complete review to the next chapter in Section 3.2, where it will be more relevant. In what follows we work in units where $\hat{q} = \frac{1}{\sqrt{2}}(\hat{a} + \hat{a}^\dagger)$ and $\hat{p} = \frac{-i}{\sqrt{2}}(\hat{a} - \hat{a}^\dagger)$, so that $[\hat{q}, \hat{p}] = i$, and $\hbar = 1$. This means the measured variance of a vacuum state is $\langle \hat{q}^2 \rangle = \langle \hat{p}^2 \rangle = \frac{1}{2}$. For more details on our conventions, see App. A.1.

2.2.1 Ideal GKP states

The ideal GKP states [13] are defined as the simultaneous eigenstates of the continuous-variable stabilizer elements

$$S_q = e^{i\left(2\sqrt{\pi}\right)(S_{11}\hat{q} + S_{21}\hat{p})}, \text{ and } S_p = e^{i\left(2\sqrt{\pi}\right)(S_{12}\hat{q} + S_{22}\hat{p})} \tag{2.1}$$

such that $\boldsymbol{S} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \in Sp(2, \mathbb{R})$, the real symplectic group in two dimensions. The Wigner function (see App. A.2.4 for the definition) of the ideal GKP states is a sum of delta functions located at the points of a lattice in phase space; the shape and spacing of this lattice is determined by $\boldsymbol{S}$. For example, rectangular GKP states are associated with the matrix

$$\boldsymbol{S}_{\text{rect}} = \begin{bmatrix} \sqrt{\pi}/\alpha & 0 \\ 0 & \alpha/\sqrt{\pi} \end{bmatrix}, \tag{2.2}$$

for $\alpha \in \mathbb{R}_{\neq 0}$, so that the unit cell of the lattice has dimensions $\frac{\sqrt{\pi}}{\alpha} \times \frac{\alpha}{\sqrt{\pi}}$ and an area of 1, while hexagonal GKP states [13, 28] have

$$\boldsymbol{S}_{\text{hex}} = \left(\frac{2}{\sqrt{3}}\right)^{1/2} \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}. \tag{2.3}$$

We will be focusing on square-lattice states – that is, rectangular GKP states with $\alpha = \sqrt{\pi}$. For these states, the stabilizer elements become

$$S_q^{\text{square}} = D\left(i\sqrt{2\pi}\right) \equiv Z\left(2\sqrt{\pi}\right), \text{ and } S_p^{\text{square}} = D\left(\sqrt{2\pi}\right) \equiv X\left(2\sqrt{\pi}\right), \tag{2.4}$$

where $D(\beta) \equiv e^{\beta\hat{a}^\dagger - \beta^*\hat{a}}$ is the displacement operator, $X(q) \equiv D(q/\sqrt{2})$ is a displacement in position by q and $Z(p) \equiv D(ip/\sqrt{2})$ is a displacement in momentum by p. After identifying the logical Z and X operations

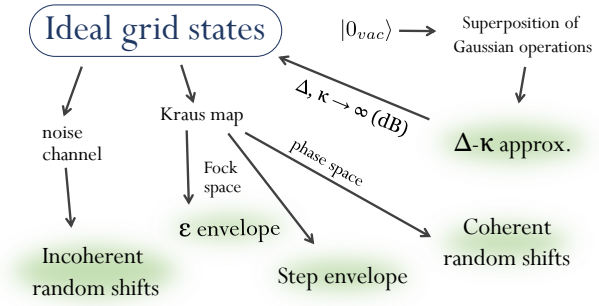


Figure 2.2: Various ways to obtain physical states starting from ideal GKP or grid states. From left to right, the approximations in green clouds correspond to discussions around Eq. (2.19), (2.11), (2.15), (2.14), and (2.8).

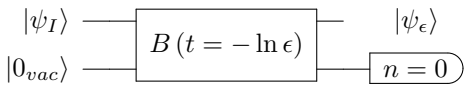


Figure 2.3: A schematic of how to obtain the normalizable GKP state $|\psi_\epsilon\rangle \equiv e^{-\epsilon\hat{n}}\,|\psi_I\rangle$ through an optical circuit, as noted in [1] and shown in App. A.2.1. An ideal GKP state and the vacuum state pass through the first and second mode of a beamsplitter $B(\theta,\phi)$ with transmissivity $-\ln\epsilon$ (see Eq. (2.56)); the second mode is measured and post-selected on a vacuum state.

for this code (logical operators are denoted by an overline),

$$\bar{Z} = Z\left(\sqrt{\pi}\right) \text{ and } \bar{X} = X\left(\sqrt{\pi}\right), \tag{2.5}$$

one can infer that the ideal encoded logical square GKP states, which we label with the subscript I, are infinite superpositions of infinitely squeezed states – delta spikes arranged like a comb – spaced by $2\sqrt{\pi}$ in position:

$$|0_I\rangle \equiv \sum_{n=-\infty}^{\infty} \left|2n\sqrt{\pi}\right\rangle_q. \tag{2.6}$$

The encoded logical 1 state is then just a q-displacement by $\sqrt{\pi}$ of the 0 state: $|1_I\rangle \equiv X\left(\sqrt{\pi}\right)|0_I\rangle$. The periodicity of the delta spikes implies that the ideal GKP code can correct displacements of position and momentum of up to $\sqrt{\pi}/2$, i.e., those displacements that do not confuse a logical 0 for a logical 1 [13]. The $\bar{X}$ eigenstates are then just

$$|+_I\rangle \equiv \sum_{n=-\infty}^{\infty} \left|n\sqrt{\pi}\right\rangle_q, \tag{2.7}$$

and $|-_I\rangle \equiv Z\left(\sqrt{\pi}\right)|+_I\rangle$.

2.2.2 Normalizable GKP states

In an experimental setting one will be dealing with finitely squeezed states. The canonical way to generate normalizable GKP states [1], which we label with a subscript corresponding to the normalization scheme, is to

[1] What we call normalizable is generally referred to as approximate in the literature. We prefer this nomenclature because there are several different approximations to ideal GKP states we will be considering.

replace the delta functions with Gaussians of width Δ and then introduce an overall Gaussian envelope of width κ^{-1}:

$$|\mu_{\Delta,\kappa}\rangle \propto \sum_{n=-\infty}^{\infty} e^{-\frac{1}{2}\kappa^2\left[(2n+\mu)\sqrt{\pi}\right]^2} X\left[(2n+\mu)\sqrt{\pi}\right]|\Delta\rangle_q\,, \mu = 0,1 \tag{2.8}$$

where $|\Delta\rangle$ is defined so that

$$\langle q\,|\,\Delta\rangle = \left(\frac{1}{\pi\Delta^2}\right)^{\frac{1}{4}} e^{-\frac{q^2}{2\Delta^2}}. \tag{2.9}$$

This is just one prescription to transition from infinite to finite energy states, as shown in Fig. 2.2. Recently, three conventional approximations of the GKP codes were analytically shown to be equivalent in [50].

It is common to work in the regime where $\Delta = \kappa$, so that in the $\Delta \to 0$ limit the number of spikes increases while the width of each spike decreases, and the normalizable states approach the ideal ones. In this case we omit the κ from the subscript and write simply

$$|\psi_\Delta\rangle \equiv |\psi_{\Delta,\kappa=\Delta}\rangle\,. \tag{2.10}$$

A less cumbersome way than Eq. (2.8) of expressing $|\psi_\Delta\rangle$, as pointed out in [28, 34], is to apply a non-unitary envelope operator $E(\epsilon) \equiv e^{-\epsilon\hat{n}}$ to the ideal state, where $\hat{n} \equiv \hat{a}^\dagger\hat{a} = \frac{1}{2}(\hat{q}^2+\hat{p}^2)$ is the number operator:

$$|\psi_\epsilon\rangle \equiv E(\epsilon)\,|\psi_I\rangle\,. \tag{2.11}$$

One can think of this operator as being the result of interfering an ideal GKP state and a vacuum state at a beamsplitter, measuring one of the modes, and postselecting on the vacuum [1]. We provide an illustration of this in Fig. 2.3 and a derivation in App. A.2.1. Note that $|\psi_\epsilon\rangle \approx |\psi_\Delta\rangle$ whenever $\Delta = \kappa$ and ϵ are small. From [47], the more precise regime is whenever

$$\tanh\frac{\Delta}{2} \approx \frac{\Delta}{2}. \tag{2.12}$$

Since $\tanh x = x - \frac{1}{3}x^3 + \ldots$, this will occur whenever $\frac{1}{24}\Delta^3$ is negligible.

Since we require non-Gaussian resource states for universal quantum computation, it is important to quantify the non-Gaussianity of the normalizable GKP states. For this, we plot the Wigner logarithmic negativity W_N (see App. A.2.4 for the definition) of the $|0_\epsilon\rangle$ state as a function of ϵ in Fig. 2.4. We see that W_N decreases for increasing ϵ, as expected.

We can, in principle, come up with other ways of approximately normalizing the ideal states. In general, we have that

$$|\psi_G\rangle \equiv G\,|\psi_I\rangle\,, \tag{2.13}$$

for some operation G, which we can regard as an error or a single Kraus operator. Note that G will not be trace-preserving in general; even though we have written the left-hand-side of (2.13) as a ket, it is understood that the state will need to be normalized. For example, G can be defined as the operator

$$G = \sqrt{\frac{2}{\pi\Delta^2}}\int d^2\alpha\, e^{-|\alpha|^2/\Delta^2} D(\alpha), \tag{2.14}$$

in other words a Gaussian distribution of displacements, as in [13, 15, 50].

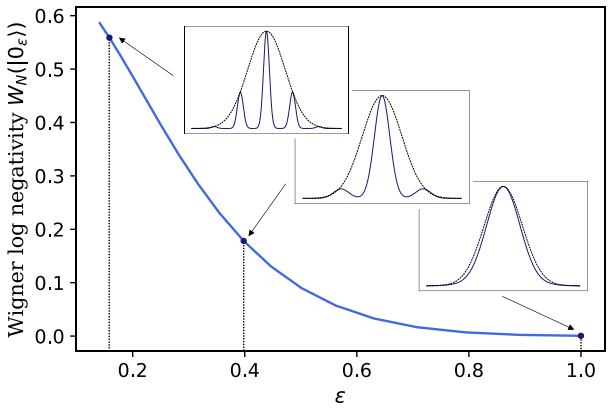


Figure 2.4: The Wigner logarithmic negativity W_N (defined in App. A.2.4) of the normalizable GKP state $|0_\epsilon\rangle$ from Eq. (2.11) as a function of ϵ shown as solid (blue) line. **Insets:** Highlighted from left to right are wavefunctions corresponding to ϵ values of 0.158, 0.398 and 1 (Δ values of 8, 4, and 0 dB). The overall Gaussian envelope of width $\kappa = \Delta^{-1}$ is also drawn. The Gaussian in the third inset noticeably does not perfectly envelop the $\epsilon = 1$ state because here there is an error of order 0.04 in the approximation between Δ and ϵ (as shown in Eq. (2.12)). W_N decreases with increasing ϵ, i.e. as the number of peaks in our GKP state decreases, the width of each one increases, and the state becomes more Gaussian.

G can also be chosen so that the normalizable state is a weighted superposition of Fock states, as in

$$|\mu_G\rangle = \sum_{n=0}^{\infty} g(n) \langle n | \mu_I \rangle |n\rangle , \tag{2.15}$$

for some envelope function $g(n)$. For example, we can set g to be a step function; or, to obtain (2.11), we can make $g(n) = e^{-\epsilon n}$ [2]. We will explore in Sec. 2.3 which envelope functions are best to target in the state preparation scheme we consider.

There are several advantages to using Eq. (2.11) to denote the normalizable states. Beyond its compactness, it provides us with a convenient way to explore the effects of deviating from ideal GKP states on Gaussian operations, as required for implementing Clifford gates, and on the logical content of the states, as we will see. It is also important to point out that the displacement and envelope operators will, in general, not commute (see App. A.2.3 for explicit commutation and conjugation relations with $E(\epsilon)$). This means that the logical state obtained by displacing $|0_G\rangle$ by $\sqrt{\pi}$ in position will not be the same as the state obtained by applying an envelope to $|1_I\rangle$. For example,

$$X\left(\sqrt{\pi}\right)|0_\epsilon\rangle = X\left(\sqrt{\pi}\right) E(\epsilon) |0_I\rangle \tag{2.16}$$

$$= E(\epsilon)\, e^{\sqrt{\pi}\left(e^{\epsilon}\hat{a}^\dagger - e^{-\epsilon}\hat{a}\right)} |0_I\rangle \tag{2.17}$$

$$\neq E(\epsilon)\, X\left(\sqrt{\pi}\right)|0_I\rangle = E(\epsilon)\,|1_I\rangle . \tag{2.18}$$

To avoid ambiguity, we will use the prescription implied by Eq. (2.13) for our normalizable states and explicitly write, e.g., $X(\sqrt{\pi})|0_G\rangle$ where necessary. We will investigate in the following sections what impact this difference will have on practical considerations and on the logical information encoded in the states.

Note that a more general way of writing down normalizable states is through some noise channel, $\mathcal{K}$,

[2] Note that we avoid ambiguity in notation because the subscript of an ideal state $|\psi_I\rangle$ can be interpreted as a normalizable state with an identity envelope

acting on the ideal states:

$$\rho_\kappa = \mathcal{K}\left(|\mu_I\rangle\langle\mu_I|\right) \tag{2.19}$$

This approach is taken in [1] with $\mathcal{K}$ corresponding to random Gaussian displacement errors, also known as the Gaussian classical noise channel.

2.2.3 Modular subsystem decomposition

As one deviates from the ideal GKP states, it becomes less obvious what the logical state is, where the information resides, and how to address and access it. Answers to these questions are facilitated by an important tool for analyzing states wherein a qubit is encoded periodically in a infinite-dimensional Hilbert space, the *modular subsystem decomposition*, investigated by Pantaleoni et al. [61]. Here we briefly review its formalism and discuss its application to approximate GKP states.

Given some real number α corresponding to the spacing between the logical basis states in position, established in Sec. 2.2.1, we can decompose any position eigenket $|s\rangle_q$ in an infinite-dimensional Hilbert space $\mathcal{H}$ as

$$|s\rangle_q = |\alpha m + u\rangle \equiv |m, u\rangle\,, \tag{2.20}$$

where $m \in \mathbb{Z}$ and $u \in [-\alpha/2, \alpha/2)$. We call αm the *integer part* of s (mod α) and u the *fractional part* of s (mod α). We can subsequently decompose the physical space $\mathcal{H}$ into $\mathcal{H} = \mathcal{L} \otimes \mathcal{G}$: a two-dimensional *logical* space $\mathcal{L}$ corresponding to our qubit and another (virtual) infinite-dimensional *gauge* space $\mathcal{G}$ corresponding to everything else. Position eigenkets break down as

$$|s\rangle_q = |\mu\rangle_{\mathcal{L}} \otimes |\tilde{m}, \tilde{u}\rangle_{\mathcal{G}}\,, \tag{2.21}$$

where $\mu = \text{parity}\,(m)$, $\tilde{m} = \frac{1}{2}(m - \mu)$, and $\tilde{u} = u$. Effectively, this decomposition amounts to stitching together the wavefunction sitting within position bins corresponding to the logical μ. In this subsystem picture, the ideal GKP states can be written

$$|\psi_I\rangle = |\bar{\psi}\rangle_{\mathcal{L}} \otimes |+_I\rangle_{\mathcal{G}}\,, \tag{2.22}$$

so that the logical mode is completely separable from gauge mode and we can recover the logical information perfectly through a trace over $\mathcal{G}$. Physically, we can access the logical information with a binned homodyne measurement, as shown in App. A.2.2. For the normalizable states, however, we see that

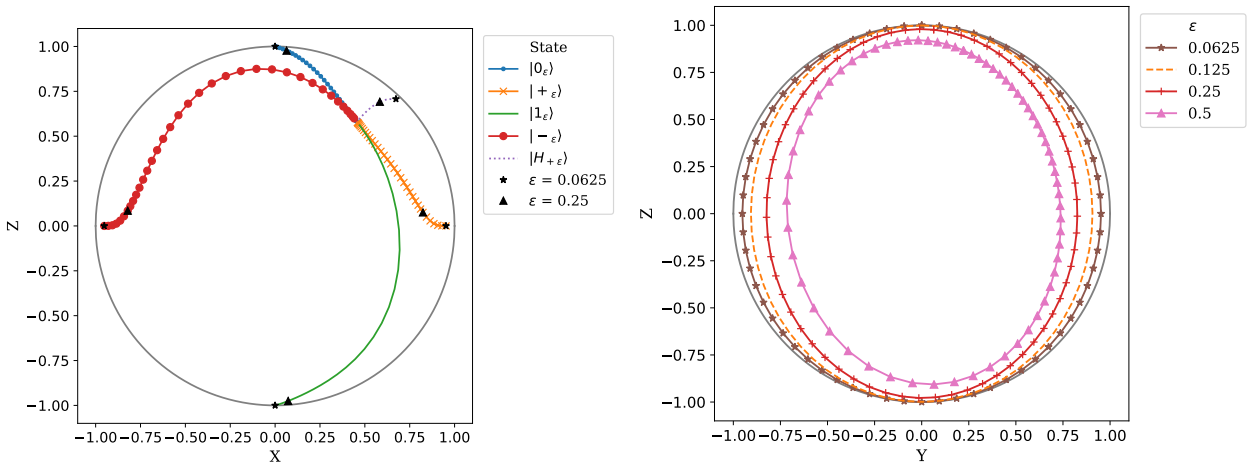
$$|\psi_\epsilon\rangle = E(\epsilon)\left(|\bar{\psi}\rangle_{\mathcal{L}} \otimes |+_I\rangle_{\mathcal{G}}\right), \tag{2.23}$$

for example. Since $E(\epsilon)$ acts on the full space $\mathcal{H}$, it will generally entangle the logical and gauge modes, leaving our logical qubit in a mixed state

$$\rho^{\mathcal{L}}(\epsilon) = \text{Tr}_{\mathcal{G}}\left[E(\epsilon)\,|\psi_I\rangle\langle\psi_I|\,E(\epsilon)\right]. \tag{2.24}$$

As ϵ grows, $\rho^{\mathcal{L}}$ will find itself somewhere inside the Bloch sphere; we plot its "trajectory" as a function of ϵ for the $\bar{Z}$ and $\bar{X}$ basis states in Fig. 2.5a. Notice that, for $\epsilon \gg 0$, the states converge at the same point: this is the vacuum state, the only state picked out by the envelope in this regime [3]. We can also see how the

[3] Note that in Pantaleoni et al., Fig. 2, the Bloch sphere trajectory for the $|+_G\rangle$ state differs form ours. This is because a



(a) (b)

Figure 2.5: Effects of the normalization envelope $E(\epsilon) = e^{-\epsilon\hat{n}}$ on the logical content of GKP states. In (a), trajectories of $|0_\epsilon\rangle, |1_\epsilon\rangle, |\pm_\epsilon\rangle$, and $|H_{+_\epsilon}\rangle$ as a function of ϵ confined to the X-Z plane of the Bloch sphere (here H_+ is the $+1$ eigenstate of the Hadamard operator; see (2.30)). For large ϵ, all Fock states but $|0\rangle$ are exponentially suppressed, causing all the trajectories to converge at the logical subsystem point corresponding to the physical vacuum state. We additionally show locations for reasonable values of ϵ: $\epsilon = 0.0625 \approx \Delta = 12$ dB, and $\epsilon = 0.25 \approx \Delta = 6$ dB are denoted by a star and triangle, respectively (see Eq. 2.12 and App. A.1 for conversions). We see that the required ϵ (dB) values to achieve higher purity states are lower (higher) towards the equator of the Bloch sphere. In (b), the modification of the X-Z plane of the ideal Bloch sphere as a function of ϵ. We see that points towards the equator are more distorted.

Bloch sphere itself changes as a function of ϵ, as depicted in Fig. 2.5b.

We calculate this efficiently by first noting that every unnormalized qubit operator can be decomposed as $\rho^{\mathcal{L}}(\epsilon) = \frac{1}{2}\sum_{i=0}^{3} s_i\sigma_i$, where σ_i correspond to the identity and the Pauli matrices, and $\boldsymbol{s}$ is the Stokes vector for the state (see for e.g. [65]). If $s_0 = 1$ then (s_1, s_2, s_3) corresponds to a Bloch vector. Thus, we can find the matrix that transforms the Stokes vectors corresponding to the ideal GKP qubits under application of the completely positive (but not trace preserving) map $E(\epsilon)$. Given the Stokes vectors corresponding to the states before renormalizing, we can now individually renormalize them by simply dividing by the first coefficient of each Stokes vector, yielding Bloch vectors in the remaining three components.

2.2.4 Operations on normalizable GKP states and error-tracking

The implementation of Clifford gates is a critical step for universal quantum computation. The Clifford group on n qubits, $\mathcal{C}_n$, is defined through its action on the Pauli group, $\mathcal{P}_n$, which consists of n-fold tensor products of Pauli gates. Any $U \in \mathcal{C}_n$ maps the Pauli group to itself under conjugation:

$$A \in \mathcal{P}_n \implies UAU^\dagger \in \mathcal{P}_n. \tag{2.25}$$

different approximation is used there, effectively corresponding to a different envelope operator.

$\bar{U}$	U physical (symbol)	U physical (name)
$\bar{X}$	$D(\sqrt{\pi/2}) = X(\sqrt{\pi})$	q displacement
$\bar{Z}$	$D(i\sqrt{\pi/2}) = Z(\sqrt{\pi})$	p displacement
$\bar{H}$	$F = R(\pi/2) = e^{i\frac{\pi}{2}\hat{n}}$	Fourier gate; $\frac{\pi}{2}$ rotation
$\bar{P}$	$P = e^{i\frac{1}{2}\hat{q}^2}$	Phase gate
$\overline{\text{CNOT}}$	$\text{SUM} = e^{-i\hat{q}_1 \otimes \hat{p}_2}$	SUM gate

Table 2.1: Conventional association between logical qubit operations and physical Gaussian transformations for ideal GKP encoding. Note that the physical gates are not unique due to fact that the stabilizers (2.4) and a π phase shift $e^{i\pi\hat{n}}$ act trivially on the code space. This means that any displacement by a Gaussian-integer multiple of $\sqrt{\pi/2}$ acts as a Pauli operator; $F^\dagger$ also represents $\bar{H}$; and a SUM gate of any odd-integer weight is also a CNOT. (A SUM gate of weight g is $e^{-ig\hat{q}_1 \otimes \hat{p}_2}$. The one shown in the table is weight 1.)

In principle, it is enough to consider a set of generators of the Clifford group, for example the Hadamard gate,

$$\bar{H} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{2.26}$$

and the phase gate,

$$\bar{P} = \sqrt{\bar{Z}} = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, \tag{2.27}$$

along with the two-qubit CNOT gate (also known as a CX or controlled-X gate),

$$\overline{\text{CNOT}} = |\bar{0}\rangle\langle\bar{0}| \otimes I + |\bar{1}\rangle\langle\bar{1}| \otimes \bar{X}. \tag{2.28}$$

Any Clifford element can then be obtained, in principle, from applications of the above gates in the generator set. In practice, and certainly with GKP states, one should also explicitly define other fundamental Clifford gates like X and Z rather than relying on compositions of the minimal set of generators.

In addition to these, one also requires an operation beyond the Clifford group (referred to as non-Clifford element), such as the $\pi/8$ gate or the T gate,

$$\bar{T} = \sqrt{\bar{P}} = \frac{1}{\sqrt{2}} e^{i\frac{\pi}{8}} \begin{bmatrix} e^{-i\frac{\pi}{8}} & 0 \\ 0 & e^{i\frac{\pi}{8}} \end{bmatrix}, \tag{2.29}$$

for fault-tolerant universal quantum computation.

Alternatively, non-Clifford gates can be effected by preparing special resource states called *magic states* that can be used in a type of gate teleportation circuit [66]. We can implement the $\bar{T}$ gate, for example, with a supply of *H-type* magic states, which are equivalent through Clifford unitaries to the Hadamard eigenstates. Refer to Fig. 5 of [13] for an example gate teleportation circuit. The eigenstate corresponding to the $+1$ eigenvalue is

$$\left|\bar{H}_+\right\rangle \equiv \cos\frac{\pi}{8}|\bar{0}\rangle + \sin\frac{\pi}{8}|\bar{1}\rangle. \tag{2.30}$$

Here we will only consider Gaussian gates, because non-Clifford gates that correspond to non-Gaussian gates can be realized via gate teleportation with magic states and Gaussian gates. This means all the gate resources remain Gaussian, and the non-Gaussian resources are imposed on the state preparation (of logical and magic

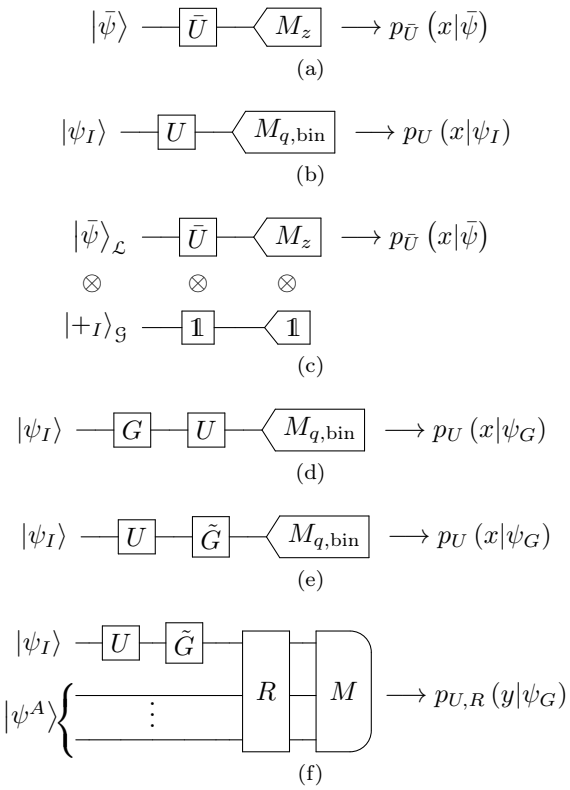


Figure 2.6: (a) The ideal logical circuit: a Clifford unitary, $\bar{U}$, is applied to a qubit, $|\bar{\psi}\rangle$, followed by Pauli Z measurements resulting in a bit, x, with probability $p_{\bar{U}}\left(x|\bar{\psi}\right)$. (b) The ideal encoded physical circuit: a Gaussian unitary, U, which implements $\bar{U}$ is applied to an ideal GKP state, followed by (binned) homodyne measurements to produce a bit with probability $p_U(x|\psi_I)$; (c) the circuit in (b) redrawn with a modular subsystem decomposition (note that we are only able to decompose the U in this way for an input ideal GKP state). The probability distributions in (a), (c), and (b) are identical: $p_{\bar{U}}(x|\bar{\psi}) = p_U(x|\psi_I)$; (d) The realistic GKP circuit, where the initial states are subject to a finite energy envelope, G. The unitary remains the same and a homodyne measurement produces a bit with probability $p_U\left(x|\psi_G\right)$; (e) the circuit in (d) but we highlight that the envelope can be conjugated with the Gaussian with appropriate modifications. (f) A general GKP circuit: the same as (e), but now we include a unitary recovery operation on the data states and ancillae $|\psi^A\rangle$. The measurement operator – which could be homodyne or PNR – is generalized to be on all modes, now producing a bit string of some composite variable, y. See the remark in the main text for a generalization of this schematic to a multimode circuit.

GKP states), which we will discuss in Sec. 2.3.

Suppose we would like to apply a gate $\bar{U} \in \mathcal{C}$ to a logical qubit $|\bar{\psi}\rangle$. If we encode $|\bar{\psi}\rangle$ in an ideal GKP state $|\psi_I\rangle$, $\bar{U}$ will correspond to a Gaussian operation U on the physical space – that is, a combination of a symplectic transformation on the quadrature operators and a displacement. We recall that not all Gaussian operations correspond to Clifford gates on the ideal GKP codes, just particular non-unique ones. Writing $|\psi_I\rangle$ in its modular subsystem decomposition, we have that

$$U|\psi_I\rangle = \left(\bar{U}|\bar{\psi}\rangle_{\mathcal{L}}\right) \otimes |+_I\rangle_{\mathcal{G}} . \tag{2.31}$$

We can see that this is true by applying U to both sides of Eq. (2.22): on the right-hand side we must obtain an ideal-GKP-encoded state (since U implements a Clifford unitary $\bar{U}$). Thus the resulting state can also be written in the form (2.22), and the action on the logical subsystem must be that of $\bar{U}$. Importantly, however, note that this does *not* mean that the physical operation can be modelled as $\bar{U}_{\mathcal{L}} \otimes I_{\mathcal{G}}$ since the decomposed operator is entangling in general [61]. The reason the gauge mode is unchanged in this particular case is that the entangling pieces of the decomposed gate do nothing when the gauge mode is exactly $|+_I\rangle$. In general, there are many physical operations U corresponding to a given logical gate $\bar{U}$; the standard mapping between $\bar{U}$ and U is given in Table 2.1.

The formalism above can be generalized to multimode states, which is necessary for generating entanglement and implementing gate teleportation. In the ideal case, a logical gate $\bar{U} \in \mathcal{C}_n$ acting on n qubits can be implemented as a Gaussian operation U on n oscillator modes. Otherwise, U will cause the logical and gauge subsystems of the different modes to interact [61]. Note further that we ought to allow for non-terminal measurements and classical feedforward within the computation. Although the unitaries will then generally depend on the result of these measurements, they will remain Gaussian.

For normalizable states, we have

$$U|\psi_\epsilon\rangle = UE(\epsilon)|\psi_I\rangle = \tilde{E}(\epsilon)U|\psi_I\rangle , \tag{2.32}$$

where $\tilde{E}(\epsilon) \equiv UE(\epsilon)U^\dagger$. Therefore, applying a Gaussian operation to a normalizable state can be viewed as applying this operation to an ideal state followed by a modified envelope. If we write the operation U as some function u of the creation and annihilation operators, $U = u\left(\hat{a}^\dagger, \hat{a}\right)$, we can also see, as shown in App. A.2.3, that

$$u\left(\hat{a}^\dagger, \hat{a}\right) E(\epsilon) = E(\epsilon)\, u\left(e^{\epsilon}\hat{a}^\dagger, e^{-\epsilon}\hat{a}\right) . \tag{2.33}$$

Expressions (2.32) and (2.33) show that the interplay between perfect Gaussian operations and imperfect GKP states causes injury to both. The damage increases with increasing ϵ – corresponding to fewer and broader peaks in the GKP state – and with increasing powers of quadrature operators that feature in U.

In this section we explore ways to quantify the severity of this damage and approaches to mitigating it. For this, we consider models for computational circuits with increasing complexity, as in Fig. 2.6. First, we will overview figures of merit for normalizable GKP states before any kind of recovery operation. This will be important for ensuring that the state input to the recovery is the best possible, thereby easing the requirements on the error correction that follows. Armed with these figures of merit, we will analyze the effects of an important class of Gaussian operations on the normalizable GKP states. In Sec. 2.2.5, we will

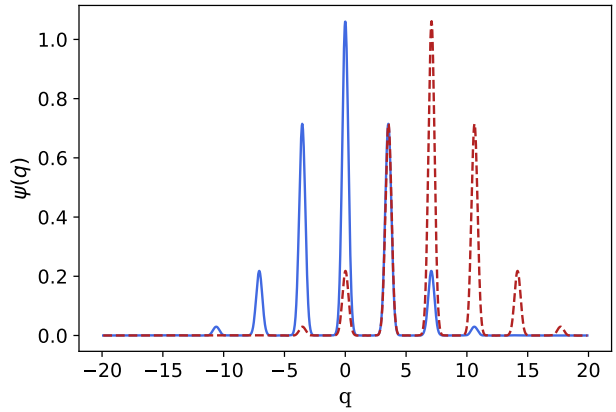


Figure 2.7: The normalizable GKP state $|0_\epsilon\rangle$ (blue, solid) and the same state after two applications of the $X(2\sqrt{\pi})$ gate (red, dashed) for $\epsilon = 0.063$ ($\Delta \approx 12$ dB). Because finite-energy GKP states have finite support in position space and hence lack a translational symmetry, the physical fidelity between the displaced state and the original state decreases with each application of $X(2\sqrt{\pi})$ even though this gate preserves the logical content.

discuss the notion of error correction with approximate GKP states and revisit the metrics.

Prerecovery figures of merit

Here we define several metrics to probe the quality and usefulness of the normalizable GKP states. For each metric, we will use as a sanity check a toy circuit consisting of an even number, $2k$, of logical $\bar{X}$ operations:

$$|\bar{0}\rangle \;\boxed{\bar{X}}\cdots\boxed{\bar{X}} \;\equiv\; \boxed{\bar{I}}$$

This circuit is equivalent to a logical identity. If we would like to implement it on an oscillator, we can encode $|\bar{0}\rangle \to |0_I\rangle$ and use the standard map $\bar{X} \to X(\sqrt{\pi})$ from Table 2.1. The ideal oscillator circuit is thus

$$|0_I\rangle \;\boxed{X(\sqrt{\pi})}\cdots\boxed{X(\sqrt{\pi})} \;\equiv\; \boxed{I}$$

Thanks to the complete translational symmetry of the ideal GKP state, this circuit is also equivalent to an identity on the oscillator space (and hence the logical space), despite it effecting a net translation of $n\alpha$ in position. In a more practical setting we have

$$|0_G\rangle \;\boxed{X(\sqrt{\pi})}\cdots\boxed{X(\sqrt{\pi})} \;\equiv\; \boxed{X(n\sqrt{\pi})}$$

Because the normalized GKP states break the translational symmetry of the ideal states, the latest circuit is no longer an identity on the state space (see Fig. 2.7). The extent of the problems that this causes will be made clearer using the following figures of merit. Note that we always treat U as a unitary operation on the physical space that effects $\bar{U}$ on the logical mode of an ideal GKP state.

Physical Fidelity. A straightforward figure of merit to consider is to compare physical states, for which we use the *physical fidelity*

$$F^{\mathcal{P}}\left(|\psi\rangle, |\phi\rangle\right) \equiv |\langle\phi\,|\,\psi\rangle|^2 . \tag{2.34}$$

We will want to compare the state before a transformation to the state after, and so we can use the shorthand notation

$$F_U^{\mathcal{P}}\left(|\phi\rangle\right) \equiv F^{\mathcal{P}}\left(U|\phi\rangle, |\phi\rangle\right). \tag{2.35}$$

If $U_{\mathbb{1}}$ is a unitary for which $\bar{U} = \bar{\mathbb{1}}$, then $F_U^{\mathcal{P}}\left(|\psi_I\rangle\right) = 1$ for an ideal GKP state $|\psi_I\rangle$. We might therefore demand for a normalizable state that

$$F_{U_{\mathbb{1}}}^{\mathcal{P}}\left(|\psi_G\rangle\right) \approx 1. \tag{2.36}$$

However, this requirement is too stringent: to see this, we can refer back to our toy circuit, where $U_{\mathbb{1}} = X\left(2k\sqrt{\pi}\right)$. In this case

$$F_{X\left(2k\sqrt{\pi}\right)}^{\mathcal{P}}\left(|\psi_\epsilon\rangle\right) = \langle\psi_\epsilon| X\left(2k\sqrt{\pi}\right)|\psi_\epsilon\rangle \tag{2.37}$$

$$= \langle\psi_I| E\left(\epsilon\right)\tilde{E}\left(\epsilon\right)|\psi_I\rangle, \tag{2.38}$$

where $\tilde{E}\left(\epsilon\right) = e^{-\epsilon\left[\left(\hat{q}-2k\sqrt{\pi}\right)^2+\hat{p}^2\right]}$ (see Table A.3 for more conjugation relations). For k high enough, our normalizable state can be displaced so much that the physical fidelity vanishes; however, the functional form of the wavefunction has not changed save for a rigid translation in position. This indicates that, while potentially useful for some applications, the physical fidelity does not adequately reveal the presence of information encoded in our normalizable state.

Logical Fidelity. Since we are interested in the logical content of the state rather than the content of the gauge mode, we might instead consider *logical fidelity,* i.e., the fidelity between the reduced logical states:

$$F^{\mathcal{L}}\left(|\psi\rangle, |\phi\rangle\right) \equiv F\left[\mathrm{Tr}_{\mathcal{G}}\left(|\phi\rangle\langle\phi|\right), \mathrm{Tr}_{\mathcal{G}}\left(|\psi\rangle\langle\psi|\right)\right], \tag{2.39}$$

where F on the right-hand-side is the fidelity for mixed states, given by

$$F\left(\rho,\sigma\right) \equiv \mathrm{Tr}\left|\sqrt{\rho}\sqrt{\sigma}\right|^2. \tag{2.40}$$

Again, we will make use of the notation

$$F_U^{\mathcal{L}}\left(|\phi\rangle\right) \equiv F^{\mathcal{L}}\left(U|\phi\rangle, |\phi\rangle\right), \tag{2.41}$$

and if $U_{\mathbb{1}}$ is such that $\bar{U} = I$, then $F_U^{\mathcal{L}}\left(|\psi_I\rangle\right) = 1$, and we might require

$$F_{U_{\mathbb{1}}}^{\mathcal{L}}\left(|\psi_G\rangle\right) \approx 1. \tag{2.42}$$

Returning to our toy circuit, if $U_{\mathbb{1}} = X\left(2k\sqrt{\pi}\right)$, this amounts to a displacement of the gauge mode only:

$$X(2k\sqrt{\pi}) = \mathbb{1}_{\mathcal{L}} \otimes X_{\mathcal{G}}(2k\sqrt{\pi}), \tag{2.43}$$

so the logical fidelity is expected to be close to unity. This shows the utility of the modular subsystem decomposition in an analysis of imperfect GKP states.

Distribution Distance. Since we are ultimately interested in the result of the computation being accurate, we can also define a post-readout metric. Returning to Fig. 2.6, we can compare the output of the ideal

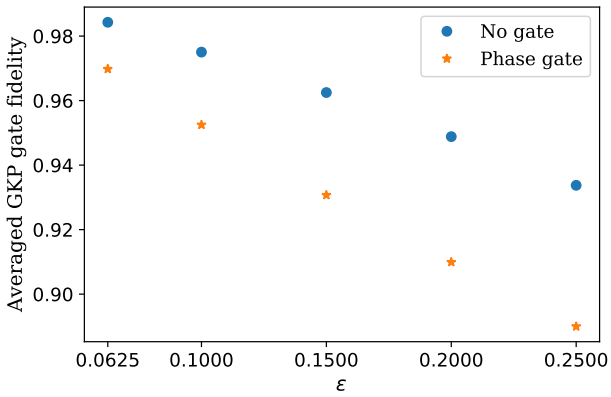


Figure 2.8: Averaged GKP gate fidelity $\bar{F}^{\mathcal{L}}_{U,G}\left(\bar{U}\right)$ from (2.45) evaluated for the logical and physical identity $\{\bar{U}=\bar{\mathbb{1}}, U=\mathbb{1}\}$ (blue), and logical and physical phase gate $\{\bar{U}=\bar{P}, U=P\}$ (orange) assuming a normalizable GKP state with $G=E(\epsilon)$ as a function of ϵ. This plot shows that even without the application of any gate, the error envelope E creates a deviation, on average, from unity fidelity with the ideal data qubit state. The deviation is even greater in the case of a phase gate, the reasons for which are given in Sec. 2.2.4. In both cases the fidelity becomes worse as ϵ becomes bigger, as expected.

circuit 2.6b with input state $|\psi_I\rangle$, which will be a bit string following some probability distribution $p_U(x|\psi_I)$, with the output $p_U(x|\psi_G)$ of the circuit 2.6d, initialized with a normalizable GKP state $|\psi_G\rangle$. For this we define the *distribution distance* through

$$\mathscr{D}^p_U\left(|\phi_G\rangle\right) = d[p_U(x|\phi_G), p_U(x|\phi_I)], \tag{2.44}$$

where d is a statistical distance, i.e., a generalized metric on the space of probability distributions. For a listing and comparison of probability metrics, see, for example, [67].

In general, we want that $\mathscr{D}^p_U(|\psi_G\rangle) \approx 0$ for any U. By focusing on the probability distribution, which we find after the readout, we do not require the state before the measurement in circuit 2.6f to be the same as the state before the measurement in circuit 2.6b. For example, in circuits 2.6d, any operation that modifies the ideal GKP states but keeps the probability distribution within the bin structure unchanged, e.g. a displacement less than $\sqrt{\pi}/2$, will still yield the same measurement statistics at readout.

With our toy circuit, if $U = X\left(2k\sqrt{\pi}\right)$, we see that a binned homodyne measurement (see A.2.2) will extract the logical information from $|\psi_G\rangle$, implying that $\mathscr{D}^p_U$ will be invariant with k.

Extremized and averaged measures. The figures of merit just described are single-shot measures; for a more complete picture, one can minimize, maximize, or average these measures over a set of a states. Let us attempt to do so for the logical fidelity:

$$\bar{F}^{\mathcal{L}}_{U,G}\left(\bar{U}\right) \equiv \int d\bar{\phi} F^{\mathcal{L}}\left(U\left|\phi_I\right\rangle, U\left|\phi_G\right\rangle\right) \tag{2.45}$$

$$= \int d\bar{\phi} \left\langle \bar{\phi}\right| \bar{U}^\dagger \mathcal{E}_{U,G}\left(\bar{\phi}\right) \bar{U} \left|\bar{\phi}\right\rangle, \tag{2.46}$$

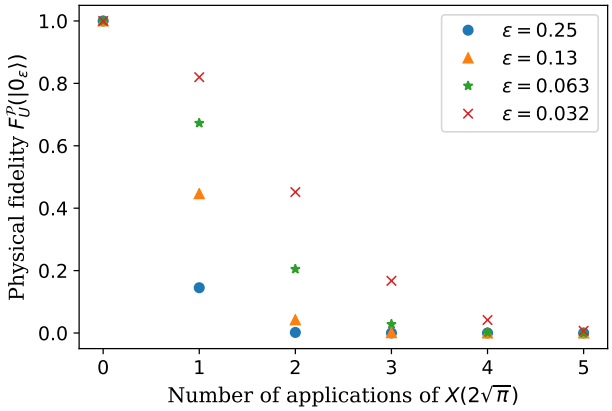


Figure 2.9: Physical fidelity of the normalizable GKP state $|0_\epsilon\rangle$ after sequential applications of the $X\left(2\sqrt{\pi}\right)$ gate, which in the ideal case implements a logical identity. The lower the ϵ, the better the approximation to the ideal GKP state, the better the translational symmetry of the state, and the slower the physical fidelity falls as a function of the number of gate applications.

where

$$|\phi_I\rangle = \left|\bar{\phi}\right\rangle_{\mathcal{L}} \otimes |+_I\rangle_{\mathcal{G}} \tag{2.47}$$

$$|\phi_G\rangle = G\left(\left|\bar{\phi}\right\rangle_{\mathcal{L}} \otimes |+_I\rangle_{\mathcal{G}}\right), \tag{2.48}$$

we define the channel

$$\mathcal{E}_{U,G}\left(\bar{\phi}\right) \equiv \mathrm{Tr}_{\mathcal{G}}\left[U\frac{G\left(\left|\bar{\phi}\right\rangle_{\mathcal{L}}\left\langle\bar{\phi}\right| \otimes |+\rangle_{\mathcal{G}}\langle+|\right)G^\dagger}{N\left(G,\bar{\phi}\right)}U^\dagger\right] \tag{2.49}$$

with normalization factor $N(G,\bar{\phi})$, and the integral is taken over the normalized uniform Haar measure on the two-dimensional state space. We call the quantity (2.45) the *averaged GKP gate fidelity* (of a physical gate U); it measures the average logical fidelity of the imperfect implementation $U|\phi_G\rangle$ to the perfect implementation $U|\phi_I\rangle$ [4]. Setting $U = \mathbb{1}$ gives us an application-independent metric of the quality of our GKP states. In Fig. 2.8 we plot $\bar{F}^{\mathcal{L}}_{U,G}\left(\bar{U}\right)$ for the identity gate and phase gate with the choice of envelope $G = E(\epsilon)$. Varying ϵ produces the expected behaviour.

Displacements in position

A displacement in position by $\sqrt{\pi}$ – half the period of $|0_I\rangle$ – is the standard way to implement a logical $\bar{X}$ operation on GKP-encoded qubits. However, we saw above that a physical X gate acting on a normalizable state quickly lowered the physical fidelity of the initial state while preserving the logical fidelity and the distribution distance. See Fig. 2.9 for the decay in physical fidelity with the number of gate applications for various envelope strengths. But there are reasons wanting to preserve the physical fidelity. For one, large displacements to a state from the origin require large amounts of energy , and higher energy can make the states more susceptible to loss. We propose several approaches for dealing with this.

First, we might wish to monitor the changes to the mean of our wavefunction caused by the gates in the circuit. Every Gaussian operation effects the map $\bar{\boldsymbol{r}} \to \boldsymbol{S}\bar{\boldsymbol{r}} + \boldsymbol{d}$ on the vector of means, $\bar{\boldsymbol{r}}$, where $\boldsymbol{S}$ is

[4] In [68] there is a simplified, easy-to-compute expression for averaged gate fidelity. However, it assumes that the error channel is trace-preserving, so we can only use it in our setting with appropriate modifications.

a symplectic matrix and $\boldsymbol{d}$ is some displacement. Finding the updated mean following an application of k Gaussian unitaries on n modes is therefore equivalent to multiplying k matrices of size $2n \times 2n$. After the circuit has been specified, we can thus classically compute the changes in the mean of our state. With this information in hand, we can then reoptimize the circuit to minimize the maximal displacement. For purely Gaussian circuits we suspect this will be an inexpensive compilation that can be done prior to the computation. The details are left to future research.

Second, we can modify the standard logical-to-physical mapping $\bar{U} \to U$; a naive approach is through

$$\bar{X} \to FPPF, \tag{2.50}$$

where F is the Fourier gate featured in Table 2.1. While this new prescription has no explicit displacements, it requires a larger number of physical operations, including a shearing of position and momentum that will damage the imperfect GKP state in a worse way, as we will see shortly. A better mapping is

$$\bar{X} \to \begin{cases} X\left(\sqrt{\pi}\right) \\ X\left(-\sqrt{\pi}\right) \end{cases} \tag{2.51}$$

This mapping could be *probabilistic*, for example, alternating between the physical gates with probability $\frac{1}{2}$. In this case one randomizes forward and backward displacement by $\sqrt{\pi}$. This is an example of a one-dimensional simple random walk: at the end of the circuit, our state is expected to remain undisplaced with a standard deviation of $\sqrt{n\sqrt{\pi}}$. Roughly, this means that if there is a threshold displacement k after which the computation becomes practically untenable, on average a circuit of depth at most $k^2/\sqrt{\pi}$ can accommodate the computation. A better mapping is *deterministic*, that is, ensuring that a forward displacement is always followed by a backward displacement and vice versa.

We see that the new prescriptions for the physical implementation of the $\bar{X}$ gate maintain both the logical and physical fidelities while minimizing the energy cost. In fact, there is evidence here that demanding that energy costs be minimized generally results in a prescription for the physical implementation of a gate that also improves the other figures of merit.

Displacements in momentum

Normally, the logical $\bar{Z}$ is physically realized on GKP states by a $\sqrt{\pi}$ displacement in momentum. The impact of momentum displacements $Z\left(\sqrt{\pi}\right)$ on the normalizable GKP states will be equivalent to that of the X gate. This can be seen in two different ways: First, the Z gate is the Fourier transform of the X gate, $Z = FXF^\dagger$, and F commutes with the envelope operator $E\left(\epsilon\right)$ as they are both exponentials of $\hat{n}$. Second, the X and Z gate are both linear in the quadrature operators, and so the damage inflicted by the envelope will be on the order of e^ϵ in both cases (both points are verified in App. A.2.3). Thus the physical fidelity seen in Fig. 2.9 will be the same in both cases.

As with position, we can monitor the changes to the average momentum and recompile the circuit. We ought also to apply the mapping

$$\bar{Z} \to \begin{cases} Z\left(\sqrt{\pi}\right) \\ Z\left(-\sqrt{\pi}\right) \end{cases} \tag{2.52}$$

deterministically. We do not consider $\bar{Z} \to PP$, as this prescription requires two applications of the quadratic phase gate and is thus a worse solution.

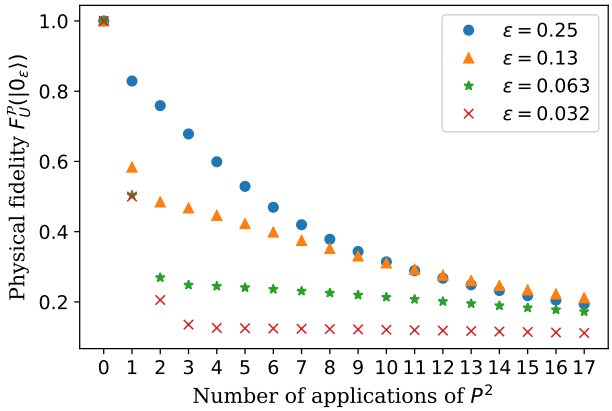


Figure 2.10: Physical fidelity of normalizable GKP state $|0_\epsilon\rangle$ after sequential applications of P^2 gates for various values of ϵ. In the ideal case this gate implements the logical $\bar{Z}$, i.e., an identity on the $|0_I\rangle$ state. The physical fidelity falls more slowly when ϵ is higher, that is, when $|0_\epsilon\rangle$ is further from the ideal state. In this regime the state has a lower squeezing and hence a wider peak at $q = 0$, where the phase gate introduces a close-to-trivial phase $e^{i\frac{1}{2}q^2}$.

Shearing operations

The logical phase gate $\bar{P}$ is effected on GKP states by the physical phase gate P, which shears the wavefunction by introducing a position-dependent phase. But the state $|0_I\rangle$ only has support on $q = 2k\sqrt{\pi}$ for $k \in \mathbb{Z}$, where the phase gate introduces a trivial factor. Since $P^2 = Z(\sqrt{\pi})$, this means that an application of P^n to $|0_I\rangle$ is an identity on both the logical and physical space for any n. On the other hand, the gate P^n applied to $|0_G\rangle$ will introduce a phase for every position value, and in particular a complex phase away from $q = k\sqrt{\pi}$. The damage inflicted on the state $E(\epsilon)|0_G\rangle$ will be on the order of $e^{2\epsilon}$ (App. A.2.3).

From Fig. 2.10, it looks like better states – those with higher squeezing – actually do worse under the phase gate when considering physical fidelity. This is because states with lower squeezing have a wider central peak, which is located in a region ($q \approx 0$) where the phase gate has little effect. The warping of the Bloch sphere as a result of the phase gate is shown in Fig. 2.11, where we see that states with higher squeezing better preserve logical information. As with displacements, we can deal with this problem by noting the flexibility in the logical-to-physical mapping:

$$\bar{P}^k = \begin{cases} \bar{I} & k \equiv 0 \,(\text{mod } 4) \\ \bar{P} & k \equiv 1 \,(\text{mod } 4) \\ \bar{Z} & k \equiv 2 \,(\text{mod } 4) \\ \bar{P}^{-1} & k \equiv 3 \,(\text{mod } 4) \end{cases} \rightarrow \begin{cases} I & k \equiv 0 \,(\text{mod } 4) \\ P & k \equiv 1 \,(\text{mod } 4) \\ Z(\sqrt{\pi}) & k \equiv 2 \,(\text{mod } 4) \\ P^{-1} & k \equiv 3 \,(\text{mod } 4)\,. \end{cases} \tag{2.53}$$

Thus, in any circuit recompilation, we ought to treat even applications of $\bar{P}$ as the identity or a $\bar{Z}$ gate, which are easier to implement. In the latter case, we can rely on the prescription for the Z gate we have provided. Similarly, 3 (mod 4) successive applications of the $\bar{P}$ gate should be replaced with a single application of $\bar{P}^{-1}$, which corresponds to the physical gate $P^{-1} \equiv e^{-i\frac{1}{2}\hat{q}^2}$.

In analogy with monitoring the position mean, we ought to keep track of how many shearing operations (positive exponents of P) and anti-shearing (negative exponents of P) we have used. For example, if the computation calls for four non-successive applications of $\bar{P}$, one might consider the physical pattern

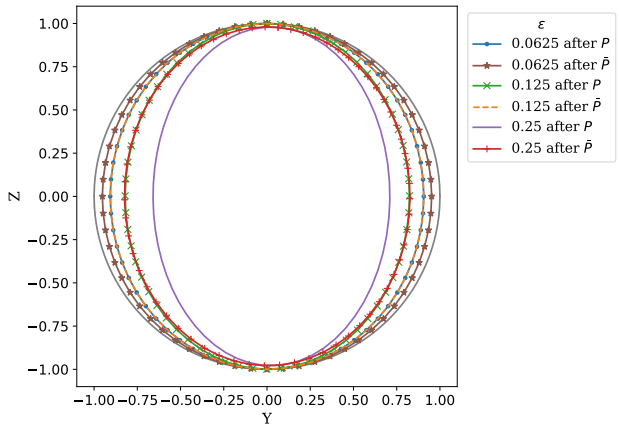


Figure 2.11: Warping of the logical Bloch sphere after application of the physical phase gate P to the normalizable GKP states $E(\epsilon)|0_I\rangle$, as compared to the action of a true logical phase gate $\bar{P}$. We show the result for various values of ϵ (for conversion to dB see App. A.1). We apply the gate to states in the $X-Z$ plane, so the $Y-Z$ plane is depicted since the ideal $\bar{P}$ gate rotates the sphere about the Z axis. Whereas $\bar{P}$ simply applies the rotation to an already distorted Bloch sphere, P additionally warps the Bloch sphere, notably near the equator where states lose more purity.

(P, P, P, P^{-3}) so that the net shear is 0. Another possibility is $(P, Z(\sqrt{\pi})P^{-1}, P, Z(-\sqrt{\pi})P^{-1})$, so that no more than one unit of shear is ever applied to the state. The best optimization strategy for any particular circuit is left to future research.

Rotations

The phase gate (a counter-clockwise rotation),

$$R(\phi) = e^{i\frac{\phi}{2}\left(\hat{q}^2+\hat{p}^2\right)}, \tag{2.54}$$

which implements the logical Hadamard operation whenever $\phi = \pi/2$ (and is called in this case the Fourier gate) commutes with the envelope operator $E(\epsilon)$. Therefore repeated applications of a perfect rotation will not damage the GKP state.

Squeezing and beamsplitters

Although they do not effect Clifford operations per se, quadrature squeezers

$$S(z) \equiv e^{\frac{1}{2}\left(z^*\hat{a}^2 - z\hat{a}^{\dagger 2}\right)} \tag{2.55}$$

and beamsplitters

$$B(\theta,\phi) \equiv e^{\theta\left(e^{i\phi}\hat{a}_1\hat{a}_2^\dagger - e^{-i\phi}\hat{a}_1^\dagger\hat{a}_2\right)}, \tag{2.56}$$

are used in the implementation of gates, such as the SUM gate below, in the continuous-variable optical domain. For later use, note that $t = \cos\theta$ is the beamsplitter transmissivity. Since the squeezing gate has terms $\hat{a}^2$ and $\hat{a}^{\dagger 2}$, its impact on normalizable GKP states will be comparable to that of the phase gate. On the other hand, beamsplitter gates have only products of $\hat{a}$ and $\hat{a}^\dagger$, which means they commute with the

envelope operator $E(\epsilon)$ (see App. A.2.3) and will not be harmful to GKP states, like the rotation gates described above.

SUM gates

For universal quantum computation we require at least one kind of two-qubit entangling operation, for example the $\overline{\text{CNOT}}$ gate. For ideal GKP states, this can translate to the continuous-variable SUM gate,

$$\text{SUM}(g) \equiv e^{-ig\hat{q}_1 \otimes \hat{p}_2}, \tag{2.57}$$

with $g = 1$ being the standard weight. However, one is free to choose any odd-integer weight to effect a $\overline{\text{CNOT}}$:

$$\overline{\text{CNOT}} \to \text{SUM}(2k+1) \text{ for } k \in \mathbb{Z}. \tag{2.58}$$

This can be seen by noting that the SUM gate effects the following quadrature transformations:

$$\hat{q}_1 \to \hat{q}_1 \tag{2.59}$$

$$\hat{p}_1 \to \hat{p}_1 - g\hat{p}_2 \tag{2.60}$$

$$\hat{q}_2 \to \hat{q}_1 + g\hat{q}_2 \tag{2.61}$$

$$\hat{p}_2 \to \hat{p}_2. \tag{2.62}$$

When the control mode is $|0_I\rangle$, then $\hat{p}_1$ and $\hat{q}_2$ are both shifted by an even multiple of $\sqrt{\pi}$ for any integer weight, meaning neither the control nor target mode are changed, as desired. On the other hand, when the control is $|1_I\rangle$, then $\hat{p}_1$ and $\hat{q}_2$ are shifted by an odd multiple of $\sqrt{\pi}$ for odd-integer weight; since $|0_I\rangle$ and $|1_I\rangle$ are $\sqrt{\pi}$-periodic in momentum, this implements a bit flip on the target. Therefore, like for the previous gates, one should update the weight of the SUM gate in a computational circuit depending on the weight of the previously applied SUM gate. To see how this translates to physical requirements, consider the decomposition of the SUM gate into squeezers and beamsplitters:

$$\text{SUM(g)} = B(\pi/2 + \theta, 0)\,(S(r,0) \otimes S(-r,0))\,B(\theta,0), \tag{2.63}$$

where $\sin(2\theta) = -\operatorname{sech}(r)$, $\cos(2\theta) = \tanh(r)$, and $\sinh(r) = -g/2$. From the discussion above, the harmful element is not the beamsplitter but the squeezer. As expected, r is a monotonic function of g; and positive values of g correspond to negative values of r.

One benchmark for how the normalizable states perform under the SUM gate is how entangled the modes become in the logical subsystem. Ideally, we have that $\overline{\text{CNOT}}\,|\bar{+}\rangle\,|\bar{0}\rangle$ is a maximally-entangled state. In Fig. 2.12, we plot the entanglement negativity of a two-qubit system in the logical subspace, a well-known measure for determining if two systems are entangled [69]. Entanglement negativity is defined to be the sum of the absolute values of the negative eigenvalues of the partial transpose of a bipartite density matrix with respect to one of the subsystems. For two-qubit systems, negativity ranges from 0 (no entanglement) to 1/2 (maximal entanglement). We initialize the two modes in the $|+_\epsilon\rangle\,|0_\epsilon\rangle$, and then apply the CNOT [SUM(1)] gate. Interestingly, we find a threshold corresponding to ~ 4.5 dB of squeezing is required to produce entanglement in the logical subsystem.

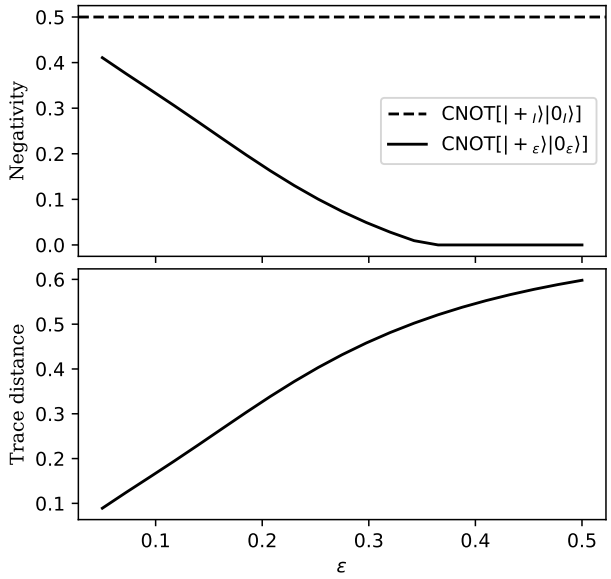


Figure 2.12: (a) Entanglement negativity (defined in main text) of a two-qubit system initialized in $|+_\epsilon\rangle\,|0_\epsilon\rangle$ and then subjected to the application of a CNOT [SUM(1)] gate, as compared to the ideal (maximal) negativity for perfect qubits. We see that to generate entanglement in the logical subsystem, we require $\epsilon \approx 0.36$, which corresponds to roughly $\Delta \approx 4.5$ dB of squeezing; however, to generate maximal entanglement, significantly more stringent thresholds are required, with $\epsilon \approx 0.05$ only generating 80% of the maximal entanglement negativity. (b) The logical subsystem trace distance between CNOT$[|+_\epsilon\rangle\,|0_\epsilon\rangle]$ and CNOT$[|+_I\rangle\,|0_I\rangle]$. Again, we see that to achieve the ideal distance of 0, smaller values of ϵ are required.

2.2.5 Error correction with normalizable GKP states

As a continuous limit of shift-resistant qudit codes, GKP encoding allows one to correct for small displacement errors in the encoded data state. In fact, the GKP codes accommodate arbitrary errors on the oscillator, since displacements – i.e., the Weyl-Heisenberg operators $X(\alpha)$ and $Z(\beta)$ – form a complete operator basis. The error syndrome measurement in the Steane approach [70] is shown and described in Figs. 2.13 and 2.14 and in the Knill teleportation approach [71, 72] in Fig. 2.15. The main difference between the two pictures is that, in the former, the data qubit interacts with two separable ancillae, whereas in the latter, the data qubit interacts with only one of two entangled ancillae. If one can guarantee a supply of high-quality ancillae, the Knill approach could be advantageous, as it precludes two applications of the SUM gate to a noisy data state. There may exist other scenarios involving GKP states in which one or the other circuit is preferable; for our

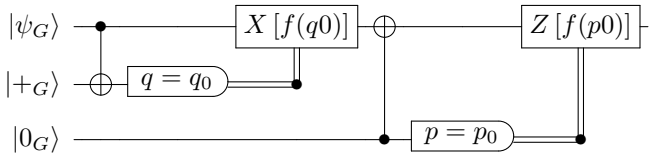


Figure 2.13: Error syndrome measurement with normalizable GKP states following the Steane approach. First, shifts in q are corrected: an encoded data qubit $|\psi_G\rangle$ and an ancilla $|+_G\rangle$ are sent through a SUM gate, and $|\psi_G\rangle$ is displaced according to the result of a homodyne q measurement on the ancilla. A similar procedure follows for shifts in p.

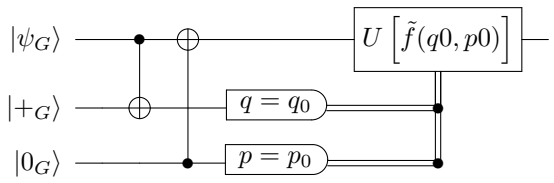


Figure 2.14: The circuit in Fig. 2.13 but with the measurements pushed to the end. Now a unitary is applied that encompasses both the correcting position and momentum shifts determined by a new function, $\tilde{f}$, of the homodyne measurement results q_0 and p_0. This unitary shifts the projected state back onto the GKP grid. When the error is too big, these shifts are mistakenly performed in the wrong direction, and a logical error results.

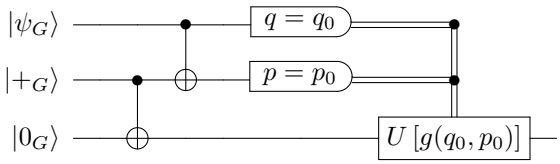


Figure 2.15: GKP error correction using the Knill approach. Here, the ancillary states $|+_G\rangle$ and $|0_G\rangle$ are entangled and then the data state $|\psi_G\rangle$ is sent through a SUM gate with $|+_G\rangle$ as the target. Homodyne position and momentum measurements are conducted on $|\psi_G\rangle$ and $|+_G\rangle$, respectively; this teleports the logical data to $|0_G\rangle$ up to a unitary, applied to $|0_G\rangle$ depending on the measurement results. Notice that, unlike in the Steane approach (Fig. 2.14), the output state after the projection is already on the GKP grid. The purpose of this unitary, therefore, is to correct the logical-Pauli byproduct operators that result from the teleportation. When the error is too big, these byproduct operators are misidentified, and a logical error results.

purposes, we focus on the Steane approach.

In the Steane error correction circuit, the SUM gate preserves the position of the data qubit and transforms the ancilla A through $q_A \to q_\psi + q_A$. This means that a measurement of the ancilla will yield $n\sqrt{\pi} + \delta q_\psi + \delta q_A$ for $n \in \mathbb{N}$, where the δq's denote the position shift errors. Therefore the function f that ought to be applied to the measurement outcomes before the correcting shift is

$$f(r) = -\mathrm{mod}_{\sqrt{\pi}}(r), \tag{2.64}$$

where $\mathrm{mod}_t(x) \in \left[-\frac{t}{2}, \frac{t}{2}\right)$. Whenever δ_ψ and $\delta_A < \frac{\sqrt{\pi}}{2}$, we have that $f(q_{\mathrm{net}}) = -\delta_\psi - \delta_A$, and we have corrected our q displacement errors. However, we have now introduced a new shift error in momentum, since the CNOT gate also effects $p_\psi \to p_\psi + p_A$. Given a perfect ancilla, we would be able to correct displacements of magnitude $\sqrt{\pi}/2$ in position and momentum. But the finite energy envelope introduces errors on the ancilla that restrict the range of correctable errors.

In this setting, since we accumulate three errors after a complete round of error correction, Glancy and Knill [33] found that error correction will only be successful if the magnitude of all shifts is less than $\sqrt{\pi}/6$. This ensures that the total error on the data qubit is within correctable region. For a position wavefunction

$\psi(q)$ of any of the noisy states, the probability of successful error correction is thus:

$$P_{\text{no error}} = \frac{\pi}{3} \sum_{s,t} \text{sinc}\left(\frac{\pi t}{3}\right) \times \int_{\sqrt{\pi}(2s-\frac{1}{6})}^{\sqrt{\pi}(2s+\frac{1}{6})} du \psi^*(2t\sqrt{\pi}+u)\psi(u). \tag{2.65}$$

where we provide a short derivation of this expression in App. A.2.5 and explain its meaning there. Note that although we are assuming approximate GKP ancillae, the formalism just described is general enough to accommodate arbitrary ancillary states, with their usefulness quantified by Eq. (2.65). For error correction to succeed with high probability, $P_{\text{no error}}$ must be high; this is satisfied by close-to-ideal GKP states. Conversely, a generic state with a high $P_{\text{no error}}$ must have little modular spread in both position and momentum, implying that it approaches the form of an ideal GKP state. In this sense, the Glancy-Knill condition is both necessary and sufficient for correcting displacement errors using the standard Steane and Knill schemes described above. We explore this point further in 2.3.3.

We note that the above GKP error correction cannot be expected to correct all errors. If during the q (p) error correction the total magnitude of position shifts is greater than $\sqrt{\pi}/2$, the procedure results in a bit-flip (phase flip) error on the logical qubit. However, one can aim to construct more sophisticated codes built on the GKP qubits, such as surface codes, which can protect more general errors as considered by a few references mentioned in Sec. 2.1 [35, 39, 47, 73]. We believe that the modular decomposition picture could play an important role in developing this further. Having introduced various figures of merit to track and understand the way error propagates due to state preparation errors, we now set up the tools to prepare explicit optical circuits to produce the realistic GKP states.

2.3 Photonic State Preparation and Characterization

2.3.1 Preparation of non-Gaussian states

GKP states are highly non-Gaussian (see, for example, the trend in Fig. 2.4), so their preparation requires non-Gaussian resources. For optical platforms, one such resource that is already experimentally accessible is the photon number-resolving (PNR) detector. A Gaussian multimode state can be prepared by applying a general interferometer, $U(\bar{\mathbf{\Theta}})$, with N^2 independent beam splitter and phase shift parameters $\bar{\mathbf{\Theta}}$, to a multimode input of displaced squeezed vacuum states $D(\boldsymbol{\alpha})S(\boldsymbol{z})|\mathbf{0}\rangle$. Next, by making PNR measurements of $N-M$ of the N modes, and obtaining results other than zero photon detections across the detectors, one prepares an M-mode non-Gaussian state [17–20]. As the architecture is analogous to Gaussian Boson Sampling (GBS), we will refer to the technique as state preparation with GBS devices. The circuit for generating a single-mode non-Gaussian state conditioned on measuring the remaining modes using PNR detectors is depicted in Fig. 2.16. This general framework encompasses other state preparation schemes, such as photon subtraction and addition (see for example, Fig. 1 of [19]).

When using GBS circuits for state preparation, the number of modes, the initial squeezing and displacement parameters, the interferometer beamsplitter angles, and the PNR measurement patterns can all become parameters for tailoring an output state according to a predefined cost function, such as closeness to a given target state and probability of successfully preparing the given state. Extensive analysis of this framework has been performed for the preparation of Fock, cat, NOON and weak-cubic-phase states [17–20], and

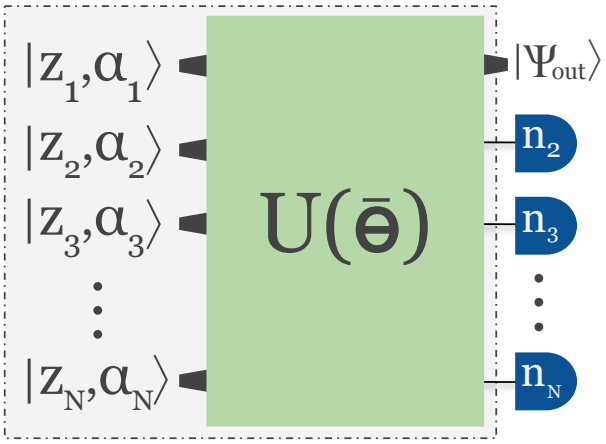


Figure 2.16: The model for the GBS-like device used for state preparation. Gaussian states consisting of squeezed, displaced vacuum states $|z_i, \alpha_i\rangle$ are sent through an N-mode interferometer parametrized by $U(\bar{\mathbf{\Theta}})$, followed by a PNR measurement on all but one of the modes. Given a PNR outcome of $\bar{\mathbf{n}}$, the desired output state, $|\psi_{\text{out}}\rangle$, is produced in the unmeasured mode. Our task is to optimize the circuit components ($\boldsymbol{\alpha}$, $\boldsymbol{z}$, and $\bar{\mathbf{\Theta}}$) for a given $\bar{\mathbf{n}}$ to produce a desired approximate GKP state. We loop over $\bar{\mathbf{n}}$ (subject to constraints) to find the best N-mode circuit for the task.

proof-of-concept calculations for the technique have been made for GKP states [19]. Here, we provide a thorough analysis for preparing GKP states with GBS circuits.

There are a few important results that guided our search for optimal state preparation using GBS circuits. First, if one wishes to prepare a Fock superposition of up to N photons by measuring pure Gaussian states, then the number of photons detected in PNR measurements should sum to N [19, 74]; this allows us to significantly restrict our search over post-selection patterns. Second, Su et. al. [18, 19] conjectured that measuring $(N-1)$ modes with PNR detectors in a GBS circuit outputs a Fock superposition with at most $(N+2)(N-1)/2$ independent coefficients. This implies that, by tuning $\boldsymbol{\alpha}, \boldsymbol{z}$, and $\bar{\mathbf{\Theta}}$ and choosing a suitable photon number post-selection pattern $\bar{\mathbf{n}}$, one can always, in principle, prepare with perfect fidelity a single-mode target state of at most $(N+2)(N-1)/2$ photons using an N-mode circuit. For a given target state, this allows us to set an upper bound on the number of modes in the circuit for which we search for optimal $\boldsymbol{\alpha}, \boldsymbol{z}, \bar{\mathbf{\Theta}}$ and $\bar{\mathbf{n}}$.

2.3.2 Core states for GKP

We now discuss a framework for approximating GKP states. In this section, we describe a formulation of GKP states that is platform-independent, so it can be applied to GKP state preparation in superconducting circuits or ion traps [40, 54, 75–77], for example.

An arbitrary single-mode quantum state can be constructed from a core superposition of Fock states followed by Gaussian operations: squeezing, displacement and rotation [78, 79]. This is sometimes referred to as the stellar representation of the state [79]. An approximation to a given state can be found by truncating the core state with a suitable $n_{\max}$:

$$|\psi\rangle \approx S(\zeta)D(\beta)\underbrace{\sum_{n=0}^{n_{\max}} \frac{c_n}{N(\boldsymbol{c}, n_{\max})}|n\rangle}_{\text{truncated core state}}, \tag{2.66}$$

where $N(\boldsymbol{c}, n_{\max})$ is the normalization constant. The exact state can be recovered by taking $n_{\max} \to \infty$.

The approximate representation (2.66) is particularly useful when using GBS circuits for state preparation. First, to prepare a target state with a given $n_{\max}$, we know how many modes and what set of PNR measurement patterns are required to guarantee production of a state with perfect fidelity to the target. Thus, we can find a circuit that optimally produces the truncated core state. Next, as the circuit consists of Gaussian operations on Gaussian states, the additional Gaussian operations, $S(\zeta)D(\beta)$, that we apply to the core state after it is produced can simply be absorbed into the Gaussian circuit, $U(\bar{\boldsymbol{\Theta}})$; moreover, the circuit can be re-decomposed, yielding new $\boldsymbol{\alpha}', \boldsymbol{z}', U(\bar{\boldsymbol{\Theta}}')$ [80]. Thus, operations, such as the squeezing on the core state, are implemented at the start of the circuit, eliminating the need for inline squeezing that is comparatively harder to implement [81]. Finally, if $n_{\max}$ is constrained by the available physical resources, such as the number of circuit modes, by targeting a core state from Eq. (2.66) rather than a truncation of the Fock expansion of $|\psi\rangle$ directly, we are generally able to attain higher fidelities between the prepared state and $|\psi\rangle$. Some of the extended support in Fock space is captured by the displacement and squeezing of the core state and may be unnecessarily discarded if one truncates the Fock expansion of $|\psi\rangle$.

The approximate representation (2.66) is additionally valuable if one wants to prepare GKP states with a different lattice symmetry, such as the hexagonal GKP states. As noted in 2.2.1, the hexagonal GKP is related to the square GKP via a symplectic transformation, which can be decomposed into Gaussian operations. Thus, if one has a GBS circuit which can prepare an approximate square GKP state, one can use it prepare the hexagonal GKP state by appending the Gaussian operations to the end of circuit (or by redecomposing the circuit into a new one).

Before we examine how to prepare approximations to the GKP X, Z and Hadamard eigenstates, let us clarify some nomenclature. *Normalizable GKP states*, reviewed in Sec. 2.2.2, are ideal GKP states reduced to a finite energy state by the application of an envelope operator. Our first task is to find an approximation, in the form Eq. (2.66), to a choice of normalizable GKP states, i.e., for a specific choice of envelope; approximations in the form (2.66) will be referred to as *approximate GKP states*. Keeping with our notation, we can denote these states as $|\psi_A\rangle$. Our second task is to find a GBS circuit that can optimally prepare the core state corresponding to approximate GKP states, and then to redecompose the circuit to include the Gaussian operations applied to the core state. We call the states output by the final circuit the *circuit GKP states*. If the circuit GKP states, for which we can define a clear experimental prescription, have high enough fidelity to the approximate GKP states, and these states have high enough fidelity to the normalizable GKP states, which for a good enough choice of envelope capture the properties of the ideal GKP states, then the circuit GKP states will share the desired properties of the ideal states.

As they have been the most commonly studied form of normalizable GKP states, we choose the $|\mu_\Delta\rangle$ states with $\mu \in \{0, 1, +, H_+\}$ as the states for which we want to find approximate GKP states (such as $\Delta = \kappa$ case in Eq. (2.8)). We could have also chosen to target (2.15) with a step function g, that is, some finite cutoff of the ideal GKP states expressed directly in the Fock basis. However, we found that, for equal cutoffs, these states had greater support than the $|\mu_\Delta\rangle$ in regions where the wavefunctions were supposed to vanish.

We note that since the wavefunctions of the above set of $|\mu_\Delta\rangle$ are all real, we do not need to apply complex squeezing to the core state, meaning $\zeta = r \in \mathbb{R}$. As the wavefunctions are also symmetric, we do not need to displace the core state, so $\alpha = 0$, and only the even Fock coefficients will contribute, meaning $c_n = 0$ for n odd. Thus, for a given $n_{\max}$, we want to find r and c_n (n even) such that the fidelity between the squeezed,

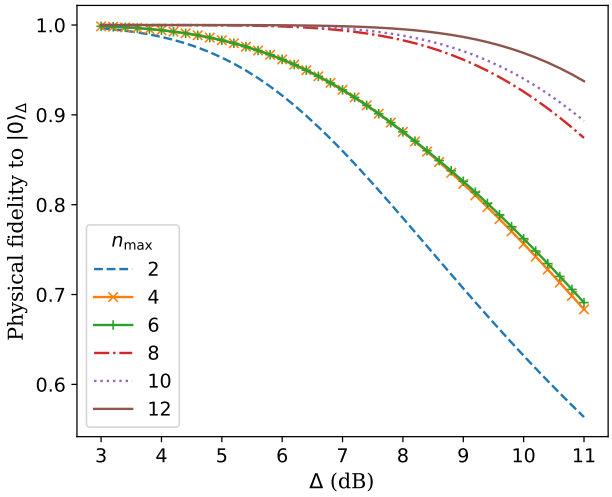


Figure 2.17: Physical fidelity of our approximate states, $|0_A\rangle$, to the target state, $|0_\Delta\rangle$, as Δ (dB) is varied (see Eq. (2.8) for the definition of $|0_\Delta\rangle$ and App. A.1 for conversion of dB to other conventions). The $|0_A\rangle$ states are constructed by applying squeezing to a core superposition of Fock states, so different line colours and styles correspond to different values of $n_{\max}$ for the core state. Note that, for different $n_{\max}$, the optimal squeezing parameter and Fock coefficients are generally different. The optimal $|0_A\rangle$ states for different $n_{\max}$ are found by maximizing the fidelity to $|0_\Delta\rangle$.

truncated core state is maximized with $|\mu_\Delta\rangle$. We summarize our method for finding optimal parameters in App. A.3.1. We found the scipy basinhopping global optimization package [82, 83] to be particularly useful.

In Fig. 2.17, we plot the fidelity between the normalizable and approximate GKP states for $\mu = 0$ as a function of Δ from 3 to 11 dB for even values of $n_{\max}$ from 2 to 12 photons. We provide a comment on the near identical results for $n_{\max} = 4$ and 6 in App. A.4.1. Our results for the other $\mu = 1, +, H_+$ states are available in App. A.4.2. As can be expected, for a fixed Δ, the fidelity improves monotonically with increasing $n_{\max}$, and for a fixed $n_{\max}$ the fidelity worsens monotonically with increasing Δ.

2.3.3 Characterization of approximate states

We can now look at various properties of the approximate states we have considered, namely the average photon number, orthogonality relations, the Glancy-Knill error correction condition, and behaviour in the modular decomposition. In App. A.5, we examine even more qualitative features of the approximate states, such as the projectors and quantum error correction matrices.

Average energy

The average energy of the approximate states will have repercussions for the resources required for making the state; for circuit GKP states, this translates to a demand on the initial squeezing applied to each mode. In Fig. 2.18, we plot the average photon numbers of $|0_A\rangle$ for different $n_{\max}$ as a function of Δ. We see that the average energy of the states is not too high; it increases with Δ and $n_{\max}$, but never exceeds five photons. We have already shown that increasing $n_{\max}$ is required for producing higher fidelities to the target GKP states, and we know additionally that states with higher Δ values (in dB) provide better error correction and encoding properties; thus, we learn a simple trend for the resources required, even in this method of preparing approximate GKP states: better states require more energy. Average photon number can also be

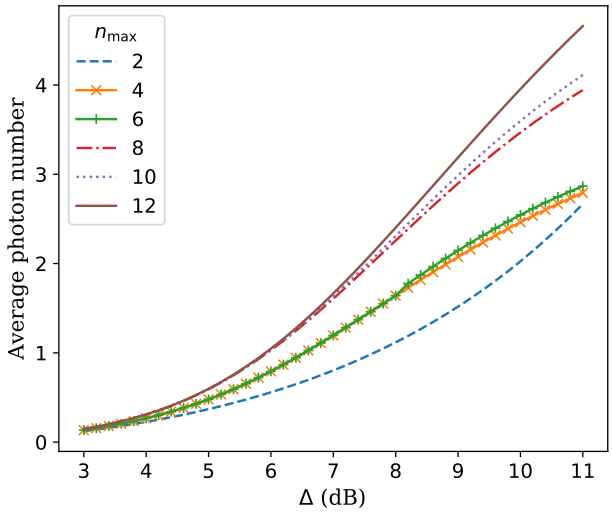


Figure 2.18: Average photon number of the approximate states $|0_A\rangle$ as a function of Δ (dB), the parameter that characterizes the target state $|0_\Delta\rangle$ that $|0_A\rangle$ is designed to approximate. Line colours/styles reflect different values of $n_{\max}$ for the core state used to construct $|0_A\rangle$. As expected, higher quality states – those with larger Δ in dB – require more energy.

used as a fundamental property to compare various bosonic codes; notably, when considering codes with average photon number less than five, GKP codes were shown to outperform other bosonic codes, including cat and binomial codes, for protecting against loss errors [15].

In App. A.4.2, we see that the average photon numbers for the $|1_A\rangle$, $|+_A\rangle$ and $|H_{+A}\rangle$ states are all on the same order and follow similar trends. Therefore, given our approach for preparing states with optical circuits, we expect that this will mean preparing different approximate GKP states on the Bloch sphere will require comparable resources, i.e., the same required order of magnitude for squeezing, and the same number of interferometer elements, and PNR detectors (see also Ref. [49]). This means one has the options when designing a computation of only preparing $|0_A\rangle$ states and $|H_{+A}\rangle$ states for non-Clifford gates, or of preparing a collection of states on the Bloch sphere from the outset: the resource requirements will be similar, and the number of gate applications in the circuit will be reduced. In other words, there might be a tradeoff between the number of different state preparation devices and the number of gate elements in the computational circuit.

Orthogonality

For the ideal GKP states, the logical 0 and 1 are orthogonal; however, for physically realizable GKP states, $|\psi_G\rangle$, they have nonzero overlap due to the tails of the wavefunction existing outside the bins. We can check the orthogonality of the approximate 0 and 1 GKP states to see how close it is to the orthogonality between the normalizable 0 and 1 states. This is valuable even when the approximate states do not have high fidelity to the normalizable states; if their overlap is small, they can still be used to encode a qubit, although they may not preserve the error-correcting properties of the GKP states nor the simplicity of the canonical gate implementation. A small overlap is also the most basic necessary condition for error correction, as we will require a low probability of mistaking a 0 for a 1.

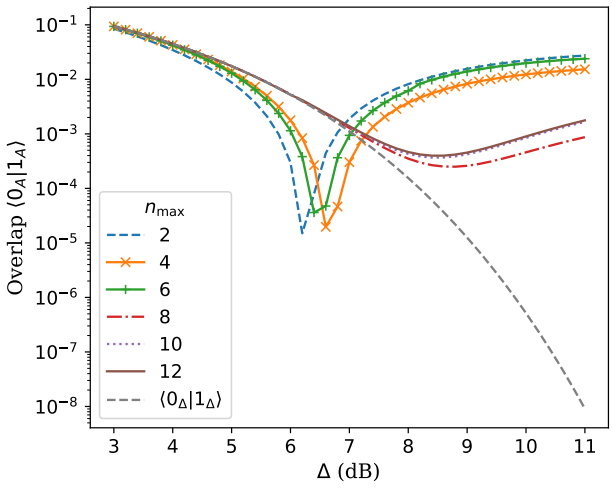


Figure 2.19: Overlap $\langle 1_A \,|\, 0_A\rangle$ between the approximate logical states as a function of Δ (dB), the parameter that characterizes the GKP states that the approximate states are targeting. Line colours/styles reflect different values of $n_{\max}$ for the core state used to construct $|0_A\rangle$. For comparison, we also plot the overlap $\langle 1_\Delta \,|\, 0_\Delta\rangle$ of the target states (grey/dashed). A small overlap is a minimal condition for being able to distinguish logical states.

In Fig. 2.19 we plot the overlap $\langle 1_A \,|\, 0_A\rangle$ as a function of Δ, for different $n_{\max}$. We see that all the states roughly follow the trend of the target overlap for low dB values of Δ, but as the squeezing of the target states increases, the approximate states become less orthogonal. This is because as Δ increases, the fidelity gets worse for both $|0_A\rangle$ and $|1_A\rangle$ relative to their target states; specifically, while for larger Δ the $|\mu_\Delta\rangle$ states consist of smooth, narrow peaks, the truncated core of the $|\mu_A\rangle$ states mean that their wavefunctions have additional oscillation between the peaks, where they should instead be close to zero (see Fig. 2.21 for an example wavefunction). This now increases the overlap $\langle 1_A \,|\, 0_A\rangle$ because they gain non-zero contributions at q values between the peak locations.

Glancy-Knill property

Using the Glancy-Knill condition in Eq. (2.65), we compute $P_{\text{no err}}$ for $|0_A\rangle$ with results depicted in Fig. 2.20. Even though, for example, $|0_A\rangle$ with $n_{\max} = 12$ has higher than 93% fidelity for all Δ considered, this can still translate to a drop in $P_{\text{no err}}$ from a target value of 74% to 58%. One must therefore be cautious in how one constructs approximate GKP states, as their error correcting properties may differ significantly. While physical fidelity depends on the specific choice of targeted finite-energy GKP state, there may be some global properties, such as the Glancy-Knill condition or logical fidelity, that are more valuable than the choice of representation.

This motivates examining an additional question: given an $n_{\max}$, what is the best $P_{\text{no err}}$ one can achieve? This question does not require defining a target state, since the cost function in Algorithm 7 is simply replaced with $P_{\text{no err}}$. We found that, using a core state with $n_{\max} = 12$ and a squeezing of $r \approx 1.87$ dB, we could modestly increase $P_{\text{no err}}$ from 57% to 61%. In Fig. 2.21, we plot three wavefunctions corresponding to three different values of $P_{\text{no err}}$: $|0_\Delta\rangle$ with $\Delta = 11$ dB, which yields $P_{\text{no err}} = 74\%$; $|0_A\rangle$ with $n_{\max} = 12$ that achieved $P_{\text{no err}} = 57\%$ by maximizing fidelity to $|0_\Delta\rangle$ with $\Delta = 11$ dB; and the state obtained from optimizing directly with respect to $P_{\text{no err}}$ and employing core states with $n_{\max} = 12$, giving $P_{\text{no err}} = 61\%$.

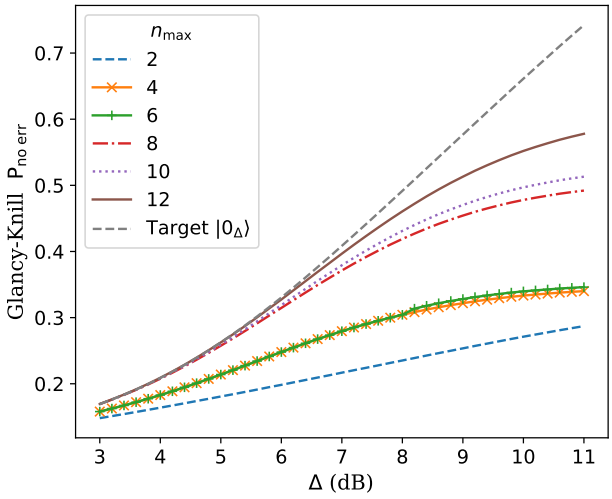


Figure 2.20: The Glancy-Knill property, $P_{\text{no err}}$, which characterizes the probability a state would not yield a logical error if used as an ancilla for Steane error correction with GKP states (see Sec. 2.2.5 for definition and details). The line colours/styles reflect $P_{\text{no err}}$ for the various $|0_A\rangle$, each constructed with a different $n_{\max}$ in the core state. We vary $|0_A\rangle$ with Δ, which parametrizes $|0_\Delta\rangle$, the state $|0_A\rangle$ approximates. Additionally, we provide $P_{\text{no err}}$ for $|0_\Delta\rangle$ (grey/dashed). For $\Delta = 11$ dB and $n_{\max} = 12$, we see that, even though $|0_A\rangle$ can have 93% fidelity to $|0_\Delta\rangle$ (see Fig. 2.17), $P_{\text{no err}}$ drops by 16%.

One interesting feature of the last wavefunction is that it has more support outside of the logical bins than the $|0_A\rangle$ state we plotted. This means that, although $|0_A\rangle$ is worse for error correction, it is better at encoding information in the logical subsystem of the modular decomposition. An additional issue with only using $P_{\text{no err}}$ as the cost function is that $P_{\text{no err}}$ can be perfectly satisfied by ideal GKP states, so the optimization procedure may be overly demanding and push towards the ideal states which we know to be non-normalizable, while we know that there exist finite energy GKP states that are suitable for computation [34].

It should also be noted that the Glancy-Knill condition is specifically a benchmark for using GKP states in error correction in quantum computation. It does not, for example, address the utility of GKP in quantum communication, where the values of Δ required for useful states are much lower. In [15] the authors showed that GKP states worked best for correcting noise resulting from a loss channel. For states with an average photon number of less than two photons, square lattice GKP states with $\Delta = 6.4$ dB were shown to be better than codes designed to correct errors due to loss. For such a Δ, with $n_{\max} = 12$, we found the fidelity of $|0_A\rangle$ to $|0_\Delta\rangle$ to be greater than 99.9%.

Logical subsystem Bloch sphere

While the Glancy-Knill condition quantifies the error-correcting capability of the approximate GKP states, examining the logical subsystem in the modular decomposition provides a benchmark for the encoding properties of the states. These are two distinct characteristics: for example, if one takes an ideal GKP state and blurs it such that the delta peaks now become distributions over position, as long as those distributions are confined to the original modular bins of the delta functions of width $\sqrt{\pi}$, then the logical information has not been disturbed, since the results of a binned homodyne measurement will be the same. However, the states might be inadequate for error correction, since the blurring could easily extend beyond the $\sqrt{\pi}/6$ threshold.

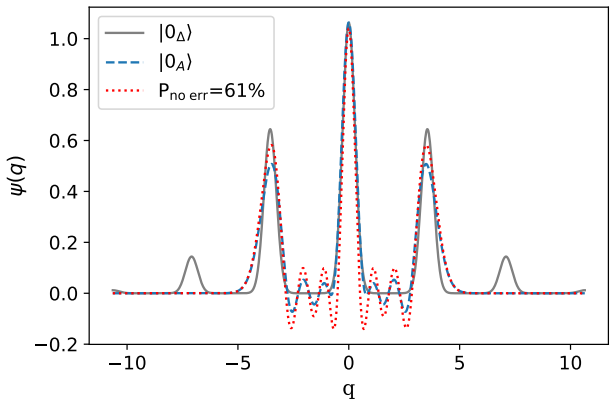


Figure 2.21: Examples of different normalizable and approximate GKP wavefunctions and their performance under the Glancy-Knill property (2.65): $|0_\Delta\rangle$ with $\Delta = 11$ dB, which yields $P_{\text{no err}} = 74\%$ (grey/solid); $|0_A\rangle$ with $n_{\max} = 12$ that achieves $P_{\text{no err}} = 57\%$ by maximizing fidelity to $|0_\Delta\rangle$ with $\Delta = 11$ dB (blue/dashed); and the state obtained from optimizing directly with respect to $P_{\text{no err}}$ with core states of $n_{\max} = 12$, giving $P_{\text{no err}} = 61\%$ (red/dotted). Although the red/dotted wavefunction has higher $P_{\text{no err}}$, the blue/dashed wavefunction is better confined to the bin structure of the GKP states, meaning it has better logical encoding when using the modular subsystem decomposition which only depends on the bin structure of the wavefunction.

We obtain the logical subsystem state from Eq. (2.24), except instead of $E(\epsilon)$ the error operators correspond to the envelopes associated with each approximate state. In Fig. 2.22, we plot the trajectories of the logical subsystems of the $|\mu_A\rangle$ states on the Bloch sphere as the target Δ is varied for different $n_{\max}$. We see that, in some cases, the logical information can be relatively close to the target position on the Bloch sphere even when the fidelity is quite low. For instance, with $n_{\max} = 2$, for $|0_A\rangle$ and $|+_A\rangle$, the approximate states are basically squeezed states in q and p. While this may not compromise the logical information in those states, we know that applications of Gaussian operations will keep the states almost Gaussian, and so we would not expect them to be suitable for universal computation. Combining the Bloch sphere picture with the Glancy-Knill results, we find that the better approximate states are provided, unsurprisingly, by increasing $n_{\max}$ and the target Δ.

2.3.4 Circuits for GKP state preparation

Algorithm for optimal circuits

Given the representation and characterization of approximate GKP states as single-mode Gaussian operations applied to truncated core states, we can now design GBS devices for producing the approximate states. As described in Sec. 2.3.1, our state preparation framework is to apply an interferometer, $U(\bar{\mathbf{\Theta}})$, to N modes of displaced, squeezed vacuum states, $D(\boldsymbol{\alpha})S(\boldsymbol{z})|\mathbf{0}\rangle$, perform a PNR measurement on $N-1$ modes, and postselect on a specific photodetection pattern, $\bar{\mathbf{n}}$, to obtain the target state in the N^{th} mode.

As we summarize in App. A.3.2, for a given truncated core state, we train $\boldsymbol{\alpha}, \boldsymbol{z}, \bar{\mathbf{\Theta}}$ and $\bar{\mathbf{n}}$ for fixed numbers of modes using machine learning algorithms, the Strawberry Fields [84] and the walrus simulators [85, 86]. Even once we find the optimal circuit for producing the truncated core state, we still have to include the squeezing, $S(r)$; to avoid the need for inline squeezing, we can take the Gaussian unitary $\tilde{U} = S(r)U(\bar{\mathbf{\Theta}})D(\boldsymbol{\alpha})S(\boldsymbol{z})$ applied to vacuum states on all the modes and express it according to the Euler or Bloch-Messiah [80] decomposition as an equivalent Gaussian unitary of the form $U(\bar{\mathbf{\Theta}}')D(\boldsymbol{\alpha}')S(\boldsymbol{z}')$, where now all the squeezing

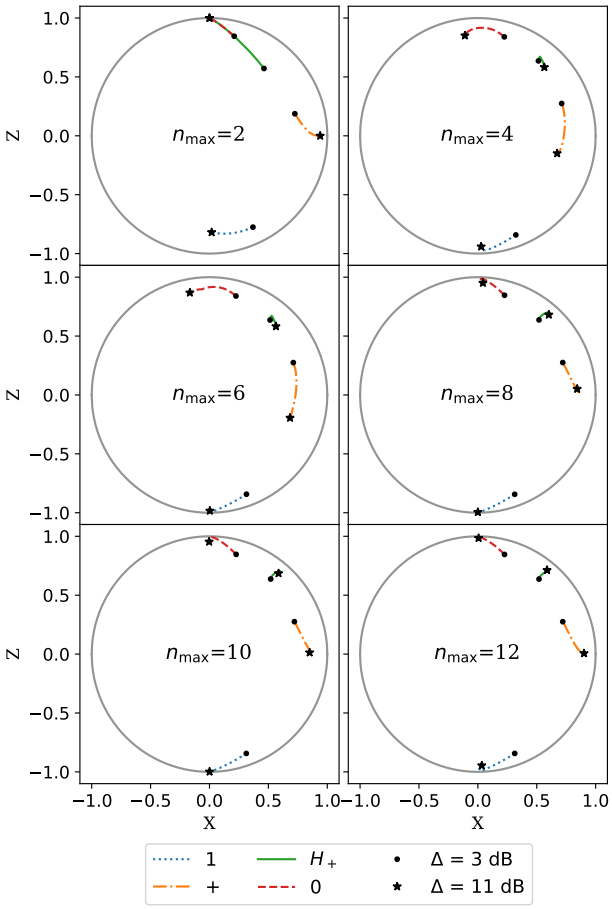


Figure 2.22: Location of the logical subsystem of the approximate states $|\mu_A\rangle$, where $\mu = 0, 1, +, H_+$, on the X-Z slice of the Bloch sphere. The trajectories are formed as the $|\mu_A\rangle$ vary with Δ, which parametrizes the states $|\mu_\Delta\rangle$ that the $|\mu_A\rangle$ approximate. We denote where the trajectories begin and end using a point for $\Delta = 3$ dB and a star for $\Delta = 11$ dB. We repeat the plot for different choices of $n_{\max}$ for the core states of $|\mu_A\rangle$. While in certain cases $|0_A\rangle$ and $|+_A\rangle$ can be mapped correctly on the Bloch sphere even with $n_{\max} = 2$, if we want all μ to be well-mapped, we need core states with higher $n_{\max}$.

is relegated to the source. This framework for training the circuits is analogous to the process described in [17].

Since the truncated core states only have even Fock coefficients, we do not need to implement the initial displacement on each of the modes. We use the rectangular decomposition from Strawberry Fields to decompose $U(\bar{\mathbf{\Theta}})$ as an N-mode interferometer with $N(N-1)$ independent beamsplitter and phase-shifter parameters. While a completely general interferometer will also include rotations on each of the modes after the application of all the beamsplitters, since we are performing PNR measurements on $N-1$ of the modes, we neglect these rotations.

Circuit optimization results

We now employ Algorithm 8 from Appendix A.3.2 to find the circuits to produce $|0_A\rangle$ for $\Delta = 10$ dB, for $n_{\max} = 4$, 8 and 12 photons. The minimum number of circuit modes we examined was 2, while for $n_{\max} = 4$, 8, and 12, we examined circuits with up to 3, 4, and 5 modes, respectively. For each circuit, we checked all PNR measurement patterns $\bar{\mathbf{n}}$, such that the number of photodetections summed to $n_{\max}$. We restricted our squeezing parameter search to, at most, 12 dB of squeezing, which is within the state of the art [87]. In nearly all cases, we found that the optimal fidelities were achieved by saturating the squeezing in at least one of the modes to $z \sim \pm$ 12 dB. We can increase the search over larger squeezing values in a straight-forward manner.

In Table 2.2, we provide our results for the best fidelities for different circuit sizes and values of $n_{\max}$. Additionally, we provide some other numerical results that still returned fidelities above 99%, but with modestly higher probabilities. Even though increasing N yields more independent parameters to tune to the target state, we see from our results that the increase in fidelity gained beyond three modes is marginal, with the corresponding probability of success vanishing. The exact parameters for the level of initial squeezing per mode and for the linear optical interferometer are available at [88], along with all the code used to implement the numerical results of the paper.

In Fig. 2.23, we plot the Wigner function of $|0_\Delta\rangle$ with $\Delta = 10$ dB, as well as the Wigner functions of the optimal states (highest fidelity) output by the three-mode circuits designed to produce $|0_A\rangle$ with $n_{\max} = 4$, 8 and 12 photons. These correspond to the starred results for $N = 3$ in Table 2.2. We see that, with increasing $n_{\max}$ – that is, as the core state resource improves – the number and sharpness of peaks approaches to that of $|0_\Delta\rangle$. The difference is smallest near the origin in phase space.

2.3.5 Experimental imperfections

Stability analysis of optical elements

The stability of the numerically computed optimal circuit parameters is important for experiment since, for example, one might have a given uncertainty in tuning the initial squeezing parameters and beamsplitter angles. As a benchmark of solution stability, we can find the worst-case fidelity within a small region in parameter space about the optimal solution. To illustrate this, let us take from Table 2.2 as an initial guess the optimal solution for a three-mode circuit designed to produce the approximate state $|0_A\rangle$ with a core state of $n_{\max} = 12$. We can then modify Algorithm 8 to minimize the fidelity and to only search within a region set by stability parameters. That is, given the optimal squeezing and beamsplitter parameters, $r_{\text{opt}}, \theta_{\text{opt}}, \phi_{\text{opt}}$, we search for the worst fidelity in a region $r_{\text{opt}} \pm \delta r/r_{\text{opt}}, \theta_{\text{opt}} \pm \delta\theta/\theta_{\text{opt}}, \phi_{\text{opt}} \pm \delta\phi/\phi_{\text{opt}}$. In Fig. 2.24 we show how much fidelity could change as a function of squeezing stability, $\delta r/r_{\text{opt}}$, and beamsplitter stability,

(a) $n_{\max} = 4$

N	1−(Fidelity to $\|0_A\rangle$)	Probability	$\bar{\mathbf{n}}$
2	0.33*	6.8%	(4)
3	1×10^{-5}*	2.1%	(1,3)
	3×10^{-4}	2.2%	(2,2)

(b) $n_{\max} = 8$

N	1−(Fidelity to $\|0_A\rangle$)	Probability	$\bar{n}$
2	0.34*	4.7%	(8)
3	1×10^{-3}*	0.41%	(4,4)
4	2×10^{-6}*	0.14%	(2,2,4)
	5×10^{-6}	0.19%	(1,3,4)

(c) $n_{\max} = 12$

N	1−(Fidelity to $\|0_A\rangle$)	Probability	$\bar{n}$
2	0.35*	2.3%	(12)
3	3×10^{-3}*	0.11%	(5,7)
4	4×10^{-8}*	5.5×10^{-5}	(3,3,6)
	2×10^{-5}	2.3×10^{-4}	(2,4,6)
5	7×10^{-9}*	6.5×10^{-5}	(1,1,3,7)
	7×10^{-8}	7.2×10^{-5}	(1,2,3,6)

Table 2.2: Results for the GBS circuits optimized to produce an approximate GKP state $|0_A\rangle$ constructed to approach $|0_\Delta\rangle$ with $\Delta = 10$ dB. We present, as a function of number of circuit modes N, the best fidelities (starred) along with other points of comparably high fidelity and probability found using Algorithm 8, with corresponding probabilities and PNR measurement patterns $\bar{\mathbf{n}}$. We examine $|0_A\rangle$ with core states of $n_{\max} =$ (a) 4, (b) 8, and (c) 12 photons.

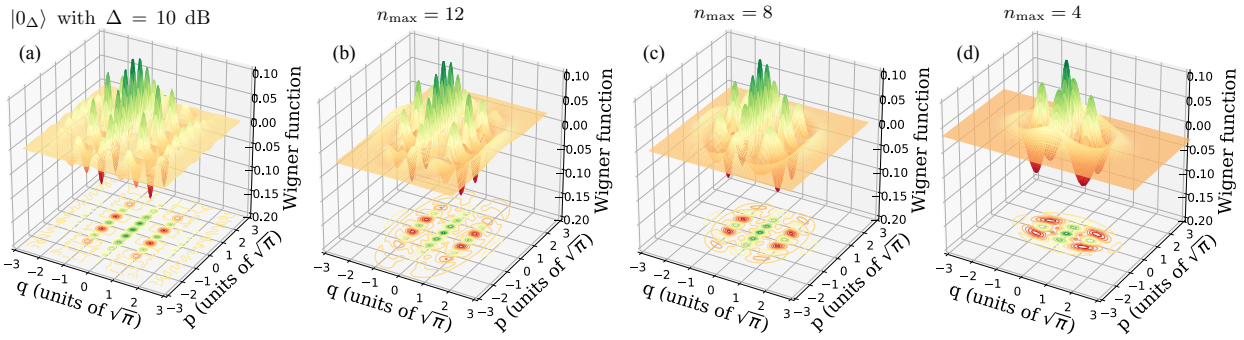


Figure 2.23: Wigner functions for (a) $|0_\Delta\rangle$ with $\Delta = 10$ dB, as well as for the optimal states output by the three-mode GBS devices designed to produce $|0_A\rangle$ with $n_{\max} =$ (b) 12, (c) 8, and (d) 4 photons. These correspond to the starred results for $N = 3$ in Table 2.2. We see the peak structure gets better with increasing $n_{\max}$, but differences to $|0_\Delta\rangle$ are still apparent further from the origin in phase space.

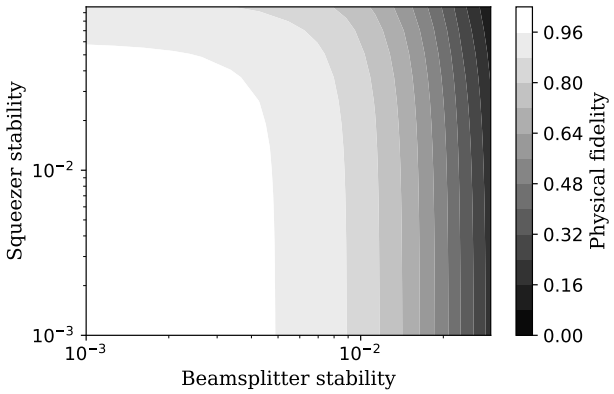


Figure 2.24: Stability analysis for the optimal 3-mode circuit designed to produce $|0_A\rangle$ with a core state of $n_{\max} = 12$ targeting $|0_\Delta\rangle$ with $\Delta = 10$ dB. We find the worst case fidelity between the circuit output and $|0_A\rangle$ as a function of the circuit beamsplitter and squeezer stability parameters, which are defined as the relative error to the ideal parameters.

$\delta\theta/\theta_{\text{opt}} = \delta\phi/\phi_{\text{opt}}$. We see that a much larger instability in the squeezing parameters can be tolerated compared to the beamsplitter parameters.

Effects of photon loss

In an experimental implementation of a GBS circuit device for state preparation, there will inevitably be loss in the optical components. Here we examine how such loss affects the results of the optimal circuit GKP states. We employ a simple loss model for our circuit: after each squeezer and each beamsplitter (with a complex transmissivity) of the interferometer in the rectangular decomposition, we apply a loss channel with loss parameter η. As the squeezers act on vacuum, they represent the only source of input light, so we capture the effect of lossy sources. Additionally, at the end of each mode, we apply a circuit out-coupling loss, η; this is followed by loss, also of magnitude η, before each PNR detector to account for detector inefficiency. The loss channel is modelled by coupling a beamsplitter of transmissivity $\sqrt{\eta}$ to an ancillary mode and then tracing out the mode. In Fig. 2.25, we plot how the fidelity, probability, and Wigner log negativity of the optimal $N = 3$ mode solutions from Table 2.2 change as a function of the single loss parameter η. Notably, we see that with increasing $n_{\max}$, loss becomes increasingly detrimental to fidelity, as it affects higher-photon number components. The probability remains relatively stable, while the Wigner log negativity also decreases with higher loss.

As a proof-of-concept, we also reoptimize some of the circuits in the presence of loss. For the three-mode circuit designed to produce the $n_{\max} = 12$ core state for $\Delta = 10$ dB, we re-ran a modified version of Algorithm 8 to find optimal circuits in the presence of loss. One change we made was to include Wigner log negativity in the cost function to ensure the states had non-Gaussian properties. Additionally, we skip the step of redecomposing the circuit, which we discuss in the next paragraph. In Fig. 2.25, we plot the reoptimized results for $n_{\max} = 12$ for three values of η. We stop at $\eta = 0.06$ dB because by that point we have dropped below the fidelity that can be achieved with only a Gaussian state, a threshold which we indicate with a black dashed line. Although the state still has nonzero Wigner log negativity, being in the regime where Gaussian states are approximating the state just as well as non-Gaussian ones significantly hampered our search for optimal states.

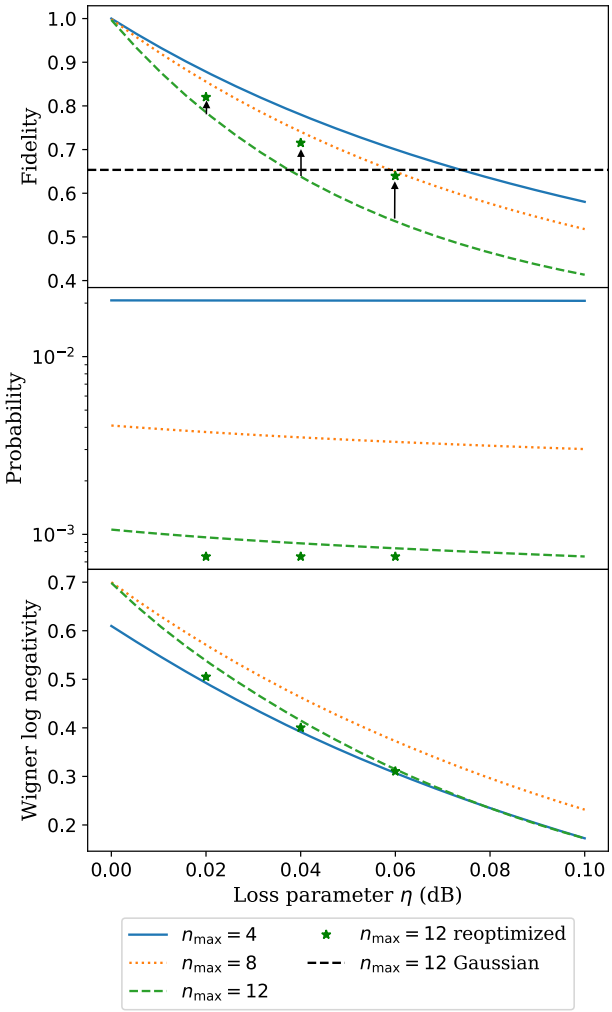


Figure 2.25: The effect of lossy optical components on the optimal GBS devices for GKP state preparation. We examine how the optimal circuits for $N = 3$ modes from Table 2.2 for core states of $n_{\max} = 4, 8, 12$ perform as a function of a single loss parameter, η. Our loss model consists of applying a loss channel parametrized by η to each optical component, as well as to the mode outcoupling and PNR detectors. We plot how the fidelity, probability and Wigner log negativity change as loss increases. Here, probability refers to the probability the GBS state preparation device outputs the state with corresponding fidelity and Wigner negativity in the other subplots. Additionally, for a few values of loss, we plot reoptimized results for lossy circuits designed to produce $n_{\max} = 12$ (green stars with arrows indicating which values were reoptimized). We see an increase in fidelity as a result. We stop reoptimization once fidelity is on the same order as that of the Gaussian state with highest fidelity to the target, a threshold we plot with a dashed black line.

There are several complications to our strategy for constructing GKP states with GBS devices that arise in the presence of loss. First, as already alluded to, there is the question of how to include the squeezing, r, applied to the core state, since the Euler decomposition is only valid for circuits without loss, so it cannot be applied perfectly in this situation. One option is to employ inline squeezing after the core state is produced by the lossy circuit, but this can be difficult experimentally [89]. Another (suboptimal) strategy is to find the optimal circuit for preparing the core state in the presence of loss, remove the loss channels, add the squeezing r, redecompose the circuit with the Euler decomposition, and then finally reinsert the loss channels. Since the r values applied to the core states of $|0_A\rangle$ with $n_{\max} = 4, 8$, and 12 are relatively small (less than 3 dB), we do not expect a large drop in fidelity using this strategy. For example, we used this strategy for the reoptimized result in Fig. 2.25 with $\eta = 0.02$ dB, and found that the fidelity only decreased from 82% to 81%.

A second complication to our strategy is the need to revisit the restriction that $n_{\max}$ photodetections ought to be detected to prepare a Fock superposition up to $n_{\max}$ with high fidelity. In the presence of loss, we might expect fewer photodetections are required as some photons could have been lost in the circuit. This leads to many more detection patterns to check, which we leave open to future study.

2.4 Conclusions

In the first part of this chapter, we reviewed the GKP encoding and various tools for assessing the quality and logical content of their finite energy versions, such as the modular subsystem decomposition and Glancy-Knill condition, showing care must be taken when determining how gates update the qubits. In the second part, we focused on obtaining the explicit resources required to produce a class of approximate GKP states using a photonic platform based on squeezed states, linear optics, and photon number resolving detectors—Gaussian Boson Sampling (GBS) state preparation. While [19] had conjectured that such GBS schemes could be used to produce non-Gaussian states in principle, we answered the next most immediate set of practical questions for such a state preparation protocol. Specifically, for single-mode approximate GKP states, we found the required number of optical modes, the level of initial squeezing, the interferometer parameters, and the number of photons that need to be measured to obtain the desired output state. We observed a fidelity-probability trade-off, and explored various properties of the output states, including their average energy and logical content under the modular subsystem decomposition. The stability analysis of these numerical optimization schemes suggested that the states output by the circuit were more robust to uncertainty in squeezing than beamsplitter angles. Finally, we investigated how loss decreased the quality of our approximate GKP states, and showed how circuits can be reoptimized in the presence of loss.

A recurring theme in the first part of this chapter was the difficulty of tracking noise in GKP qubits. With this as motivation, in the next chapter we introduce a novel mathematical framework for simulating GKP states along with other important bosonic qubits. The framework allows for a microscopic understanding of the noise that can arise in GKP states in the process of undergoing the gates required for quantum computation. Then, in Chapter 4, we build on our understanding of GKP states and our state preparation protocol to propose and analyse a photonic architecture for fault-tolerant quantum computing based on GKP states.

Chapter 3

Fast Simulation of Bosonic Qubits via Gaussian Functions in Phase Space

This chapter is based on [21], co-authored with Nicolás Quesada, Ilan Tzitrin, Antal Száva, Theodor Isacsson, Josh Izaac, Krishna K. Sabapathy, Guillaume Dauphinais, and Ish Dhand. The work was collaborative, and I was supervised mainly by Krishna K. Sabapathy, Guillaume Dauphinais, and Ish Dhand. I was first author for this work. My main contributions were to the formulation of the general formalism in Section 3.3, the analysis of GKP states in Section 3.4 and 3.5.3, to the numerical methods and simulations in Sections 3.6 and 3.7, and to writing of these sections in the manuscript along with the introduction, parts of the background material, and conclusion. The work benefited from helpful discussions with Rafael Alexander, Giacomo Pantaleoni, Daiqin Su, and Barbara Terhal.

3.1 Introduction

Photonics and superconducting cavities are leading platforms for building a scalable fault-tolerant quantum computer [22, 35, 40, 44, 90–95]. As continuous-variable (CV) quantum systems, these platforms rely on encoding qubits (two-level quantum systems) into the state of the CV system via so-called bosonic qubits, among which Gottesman-Kitaev-Preskill (GKP) states [13], cat states [96], and Fock states [14, 97] are the primary. Bosonic qubits are especially favourable for quantum computing because of their ability to correct physical errors within the CV space due to loss [15], random displacements [13], and rotations [98].

While concrete quantum computing architectures based on bosonic qubits have been proposed, the analysis and simulation of these qubits is challenging because of the infinite-dimensional Hilbert space that they occupy. This impedes the development and implementation of these architectures since determining fault-tolerance thresholds and overheads is limited by our ability to simulate these physical systems in realistic situations. The current most flexible method for simulating bosonic qubits relies on the Fock basis. Simulations in the Fock basis can be cumbersome, especially for CV states with large energy; in particular, high-quality and therefore high-energy cat and GKP states require a high photon-number cutoff, incurring large memory loads and processing times. Moreover, determining how states change under CV channels and measurements is computationally expensive in the Fock basis representation [99] as the energy of a given state can increase significantly under paradigmatic CV transformations such as squeezing and displacements.

Here, we overcome the challenge of studying bosonic qubits by introducing a novel formalism that

enables their analysis and simulation. Specifically, we present a mathematical framework for simulating a class of CV states, transformations, and measurements using linear combinations of Gaussian functions in phase space. This framework allows us to simulate the transformation of useful bosonic qubits such as GKP, cat, and Fock states under Gaussian channels and measurements, as well as under a class of valuable non-Gaussian channels effected through gate teleportation. Motivated as a tool to facilitate the design of quantum computing architectures, our framework can model important sources of decoherence (such as optical loss in photonics or dissipation in superconducting cavities) as well as transformations and measurements that are readily-implementable in photonics, such as linear optics, squeezing operations, homodyne and photon-counting detection. We accomplish this by leveraging the most convenient aspects of the Gaussian CV formalism—namely, the ability to regard the transformation of a state as a transformation of means vectors and covariance matrices—while providing the capability to simulate non-Gaussian systems, which is necessary to the construction of a quantum computer [100].

Informed by our formalism, we provide a method for the fast and accurate simulation of bosonic qubits. We find the scaling of the memory and processing time required for our simulator are vastly more favourable than current state-of-the-art Fock basis simulators [84]. We use this to conduct an in-depth numerical study of bosonic qubits in useful physical circuits under inevitable physical imperfections, including loss in optical components; finite-energy effects in resource states; and finite squeezing in the ancillae of measurement-based squeezing operations, the workhorse of inline squeezing [101]. Specifically, we analyze GKP states in three situations that are relevant for quantum computation. First, we examine GKP states passing through a qubit phase gate (introduced in [13] and studied in Sec. II D of [16]), a Clifford gate typically used in the universal gate set for GKP qubits and whose CV implementation we simulate being performed with measurement-based squeezing. Second, we consider the teleportation of a GKP state into a CV cluster state, a scenario present in proposals for measurement-based quantum computation with GKP qubits [22, 34, 44, 63]. Third, we study applying a qubit T gate to GKP states via gate teleportation with finite-energy GKP magic states and realistic entangling operations, which is an important scenario because the T gate, in conjunction with experimentally-accessible Gaussian operations, is a standard prerequisite for unlocking universal computation with GKP qubits [13]. While gate teleportation is expected to be favourable over other methods [62], simulation of the technique in the presence of realistic gate noise, to the best of our understanding, has not been performed, likely due to numerical challenges. Thus, by enabling the study of situations that were hitherto intractable, our work provides a valuable toolkit for the analysis and simulation of quantum computation based on bosonic qubits.

The structure of this chapter is as follows. In Section 3.2 we provide background on the Gaussian CV formalism and on bosonic qubits. In Section 3.3 we introduce our new formalism for simulating a wide class of states which can be expressed as linear combinations of Gaussian functions in phase space. We provide rules for how these states evolve under Gaussian and a class of non-Gaussian transformations and measurements. With the framework in hand, in Section 3.4 we show how GKP states can be written in our formalism. (Other bosonic qubits like cat and Fock states are presented written in our formalism in Appendix B.2). In Section 3.5, we detail how the formalism can be used for modelling loss, measurement-based squeezing, and useful non-Gaussian GKP qubit operations. Given the formalism, in Section 3.6, we provide our simulation methods, along with a comparison to simulation techniques in the Fock basis. Our formalism and methods allow us to present results from novel simulations of bosonic qubits in Section 3.7. We discuss additional areas of application for our simulator, and open research problems in Section 3.8.

For a more hands-on introduction to the simulations that are enabled by our formalism, we invite the

reader to consult the open-source code available in `Strawberry Fields` [84, 102] and an accompanying set of tutorials available online [103–105].

3.2 Background

In this section we provide overviews and pointers to the relevant literature of three different threads that are unified in the later parts of the manuscript. In Section 3.2.1 we provide a brief survey of the Gaussian formalism for CV quantum systems, including quantum phase-space, the symplectic formalism, and Gaussian states, channels, and measurements. In Section 3.2.2 we provide an overview of `Strawberry Fields`, the programming library in which we implement the formalism and methods developed in the rest of the chapter. Finally, in Section 3.2.3 we review the GKP and cat qubit encodings, which leverage the large Hilbert space of a CV mode.

3.2.1 Continuous-Variables and the Gaussian Formalism

Multimode continuous-variable systems are best described using the canonical position $\hat{q}_j$ and momentum $\hat{p}_j$ operators acting on the infinite-dimensional Hilbert space associated with the system. It is often convenient to group these operators, representing N modes, into a vector:

$$\hat{\boldsymbol{\xi}}^T = (\hat{q}_1, \hat{p}_1, \hat{q}_2, \hat{p}_2, \ldots, \hat{q}_N, \hat{p}_N). \tag{3.1}$$

The commutation relations these operators satisfy can be expressed succinctly in terms of the symplectic form

$$\boldsymbol{\Omega} = \bigoplus_{i=1}^{N} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{3.2}$$

as

$$[\hat{\xi}_j, \hat{\xi}_k] \equiv \hat{\xi}_j \hat{\xi}_k - \hat{\xi}_k \hat{\xi}_j = i\hbar\Omega_{j,k}. \tag{3.3}$$

As for any quantum mechanical system, a complete description of its state can be obtained by specifying its density matrix $\hat{\rho}$. For CV systems it is often useful to introduce the characteristic function [106]

$$\chi(\boldsymbol{r}; \hat{\rho}) = \mathrm{tr}\left(\hat{D}(\boldsymbol{r})\hat{\rho}\right), \tag{3.4}$$

where $\hat{D}(\boldsymbol{r}) = \exp\left(i\hat{\boldsymbol{\xi}}^T\boldsymbol{\Omega}\boldsymbol{r}\right)$ is the Weyl or displacement operator and $\boldsymbol{r} \in \mathbb{R}^{2N}$ is a real vector in phase-space.

Gaussian States

We recall Gaussian states [106] as the ones whose characteristic function takes the form

$$\chi(\boldsymbol{r}; \hat{\rho}) = \exp\left(-\tfrac{1}{2}\boldsymbol{r}^T\boldsymbol{\Sigma}\boldsymbol{r} - i\boldsymbol{\mu}^T\boldsymbol{\Omega}\boldsymbol{r}\right), \tag{3.5}$$

where we introduced the vector of means $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ of the state $\hat{\rho}$ with elements

$$\mu_i = \langle \hat{\xi}_i \rangle \,, \tag{3.6}$$

$$\Sigma_{i,j} = \tfrac{1}{2} \langle \hat{\xi}_i \hat{\xi}_j + \hat{\xi}_j \hat{\xi}_i \rangle - \mu_i \mu_j, \tag{3.7}$$

where the expectation value of an operator $\hat{A}$ is defined as

$$\langle A \rangle = \mathrm{tr}(\hat{A}\hat{\rho}). \tag{3.8}$$

The covariance matrix of a valid quantum state, whether or not it is Gaussian, satisfies the uncertainty relation

$$\boldsymbol{\Sigma} + i\tfrac{\hbar}{2}\boldsymbol{\Omega} \geq 0. \tag{3.9}$$

A different characterization of Gaussian pure states can be obtained by noting that they can be prepared by applying a Gaussian unitary $\hat{U}$ to the multimode vacuum state $|0\rangle^{\otimes N}$; that is, $|\psi\rangle_G = \hat{U}\,|0\rangle^{\otimes N}$, where the unitary is generated by a Hamiltonian that is at most quadratic in the quadrature operators [107, 108]. The vacuum state is the unique state that is mapped to zero by its respective destruction operator

$$\hat{a}_j\,|0_j\rangle = 0, \quad \hat{a}_j = \tfrac{1}{\sqrt{2\hbar}}(\hat{q}_j + i\hat{p}_j), \tag{3.10}$$

and has vector of means and covariance matrix

$$\boldsymbol{\mu}_{\mathrm{vac}} = \mathbf{0} \text{ and } \boldsymbol{\Sigma}_{\mathrm{vac}} = \tfrac{\hbar}{2}\mathbb{1}. \tag{3.11}$$

Finally, we introduce the Wigner function, which is the Fourier transform of the characteristic function:

$$W(\boldsymbol{\xi};\hat{\rho}) = \int \frac{d^{2N}\boldsymbol{r}}{(2\pi)^{2N}} \exp(-i\boldsymbol{\xi}^T\boldsymbol{\Omega}\boldsymbol{r})\ \chi(\boldsymbol{r};\hat{\rho}). \tag{3.12}$$

The Wigner function of a Gaussian state is a Gaussian function of the phase-space variables $\boldsymbol{\xi}$. Later, we introduce a class of states with Wigner functions that can be expressed as a linear combination of Gaussian functions in phase space.

Gaussian Transformations and Measurements

A Gaussian unitary transformation $\hat{U}$ is equivalent to a homogeneous linear phase-space transformation $\boldsymbol{\xi} \to \boldsymbol{S}^T\boldsymbol{\xi}$, followed by a phase-space displacement $\boldsymbol{\xi} \to \boldsymbol{\xi} + \boldsymbol{d}$. Here the symplectic map $\boldsymbol{S}$ (satisfying $\boldsymbol{S}\boldsymbol{\Omega}\boldsymbol{S}^T = \boldsymbol{\Omega}$) takes $\chi(\boldsymbol{r};\hat{\rho}) \to \chi(\boldsymbol{S}^T\boldsymbol{r};\hat{\rho})$, which, for Gaussian states, is equivalent to transforming the covariance matrix as $\boldsymbol{\Sigma} \to \boldsymbol{S}\boldsymbol{\Sigma}\boldsymbol{S}^T$ and mean as $\boldsymbol{\mu} \to \boldsymbol{S}\boldsymbol{\mu}$ [108]. Finally, the displacement transforms the mean as $\boldsymbol{S}\boldsymbol{\mu} \to \boldsymbol{S}\boldsymbol{\mu} + \boldsymbol{d}$. As examples, the single-mode displacement and squeezing operators are defined respectively by

$$\hat{D}(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a}), \tag{3.13a}$$

$$\hat{S}(\zeta) = \exp\left(\tfrac{\zeta}{2}\hat{a}^2 - \tfrac{\zeta^*}{2}\hat{a}^{\dagger 2}\right), \tag{3.13b}$$

but in phase-space are represented as

$$\boldsymbol{S}_{\text{disp.}} = \mathbb{1}, \quad \boldsymbol{d}^T_{\text{disp.}} = \sqrt{2\hbar}(\Re(\alpha), \Im(\alpha)), \tag{3.14a}$$

$$\boldsymbol{S}_{\text{sq.}} = \begin{pmatrix} e^{-r} & 0 \\ 0 & e^{r} \end{pmatrix}, \quad \boldsymbol{d}^T_{\text{sq.}} = \boldsymbol{0}, \tag{3.14b}$$

where in the last line we assumed for simplicity that $\zeta = \zeta^* = r$ is real.

A Gaussian channel is a linear completely-positive trace-preserving map from Gaussian states to Gaussian states. Gaussian channels can be described by a pair of real matrices $(\boldsymbol{X}, \boldsymbol{Y})$ with $\boldsymbol{Y} + i\frac{\hbar}{2}\boldsymbol{\Omega} \geq i\frac{\hbar}{2}\boldsymbol{X}\boldsymbol{\Omega}\boldsymbol{X}^T$ [109]. The action of the Gaussian channel described on the characteristic function is

$$\chi(\boldsymbol{r}; \hat{\rho}) \to \chi'(\boldsymbol{r}; \hat{\rho}) = \chi(\boldsymbol{X}\,\boldsymbol{r}; \hat{\rho}) \exp\left(-\frac{1}{2}\boldsymbol{r}^T\,\boldsymbol{Y}\,\boldsymbol{r}\right).$$

It follows that the transformation on the covariance matrix and mean is given by

$$\boldsymbol{\Sigma} \to \boldsymbol{X}^T\boldsymbol{\Sigma}\boldsymbol{X} + \boldsymbol{Y}, \;\; \boldsymbol{\mu} \to \boldsymbol{X}\boldsymbol{\mu}. \tag{3.15}$$

In addition to the mapping provided by $(\boldsymbol{X}, \boldsymbol{Y})$, a displacement can always be added to a Gaussian channel and the transformation will remain deterministic.

One can also consider the characteristic and Wigner functions of an arbitrary operator $\hat{A}$. This is the so-called Weyl transform of a general Hilbert space operator $\hat{A}$, obtainable by replacing $\hat{\rho}$ with $\hat{A}$ in Eq. (3.4) and then taking its Fourier transform as in Eq. (3.12). The expectation value of the operator can now be written

$$\langle \hat{A} \rangle = (2\pi\hbar)^N \text{tr}(\hat{\rho}\hat{A}) = \int d^{2N}\boldsymbol{\xi}\; W(\boldsymbol{\xi}; \hat{\rho}) W(\boldsymbol{\xi}; \hat{A}). \tag{3.16}$$

We allow that $\hat{A}$ is a measurement operator, in which case $\langle \hat{A} \rangle$ is the probability of the outcome associated with the operator. In the case the measurement operator describes the detection of a Gaussian state parametrized by $\boldsymbol{r}_M$ and $\boldsymbol{\Sigma}_M$, this is referred to as a *general-dyne measurement*, and the Weyl transform matches the Wigner function for that state [109]. A homodyne measurement is a limiting case of general-dyne measurement corresponding to projection onto an eigenstate of a quadrature operator, which can be treated as an infinitely squeezed state along that quadrature. The probability distribution for the outcome $\boldsymbol{r}_M$ of a general-dyne measurement on a Gaussian state, as can be deduced from Eq. (3.16), is itself a Gaussian distribution since the integration is between two Gaussian functions.

Gaussian measurement motivates a transformation beyond Gaussian channels, namely conditional Gaussian dynamics, i.e., an update to a subset of modes of a multimode Gaussian state conditioned on the outcome of a Gaussian measurement that is performed on the remaining modes. Following [109], the covariance and mean of the multimode state can be written as:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_A & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}^T_{AB} & \boldsymbol{\Sigma}_B \end{pmatrix} \text{ and } \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{pmatrix}, \tag{3.17}$$

where A denotes the modes that will remain active and B those that will be measured. If the Weyl transform of the measurement operator $\hat{M}$ corresponds to a Gaussian state with mean $\boldsymbol{r}_M$ and covariance $\boldsymbol{\Sigma}_M$, then the partial trace $\text{tr}_B(\hat{\rho}_{AB}\hat{M})$ corresponds to a Gaussian integral in phase space over the quadrature variables

of modes B, yielding a Gaussian state in modes A with the following covariances and means:

$$\begin{aligned}\boldsymbol{\Sigma}_A &\to \boldsymbol{\Sigma}_A - \boldsymbol{\Sigma}_{AB}(\boldsymbol{\Sigma}_B + \boldsymbol{\Sigma}_M)^{-1}\boldsymbol{\Sigma}_{AB}^T, \\ \boldsymbol{\mu}_A &\to \boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}(\boldsymbol{\Sigma}_B + \boldsymbol{\Sigma}_M)^{-1}(\boldsymbol{r}_M - \boldsymbol{\mu}_B).\end{aligned} \tag{3.18}$$

As we show in Section 3.3, we can take inspiration from the Gaussian phase space mathematical framework we have reviewed to introduce a class of states and measurements that can be expressed as a linear combination of Gaussian functions in phase space, along with how such states transform. Importantly, we find in Section 3.4 and Appendix B.2 that common bosonic qubit encodings fall within this formalism. Next, we review the definitions and properties of those bosonic qubits.

3.2.2 Continuous-Variable Simulation with Strawberry Fields

`Strawberry Fields` is a full-stack Python library for programming, designing, simulating, and optimizing continuous-variable quantum optical circuits [84, 110]. The library has a unified frontend that allows to write CV quantum circuits and programs at a high-level. Moreover, it allows users with basic knowledge about CV and quantum photonics to access an application layer that can be used to solve practical problems in graph theory, point processes and chemistry. The frontend also provides functionality for gate decomposition and program compilation and verification. Once programs are verified and compiled they are passed to a software backend or directly to cloud-available hardware [111]. The software or quantum photonic hardware can return a number of useful results to the frontend, including batches of samples, cost functions for further numerical optimization or representations of the quantum state. The frontend provides further functionality for exploration such as sample processing and plotting. The three software backends handle the actual simulation using different internal numerical representations of quantum states, each with their own unique advantages and weaknesses. The `gaussian` backend simulates Gaussian states undergoing Gaussian and non-Gaussian (threshold and photon-number-resolving) measurements [86, 112]. The `fock` and `tf` backends use a Fock basis truncation to represent CV quantum states and operations as high-dimensional tensors [99]. They differ in the tools they rely on for the numerical implementation of the tensor operations: the former uses the `NumPy` package [113], while the latter employs `TensorFlow` [114].

We implement the results of this manuscript as a new, fourth backend of `Strawberry Fields` [102]. This `bosonic` backend can be regarded as a generalization of the `gaussian` backend, and integrates directly into the rest of the `Strawberry Fields` stack. For a series of beginner to advanced tutorials implemented by the authors on using the new backend, see [103–105]. The `bosonic` backend implements much of the formalism and methods we discuss in Sections 3.3 through 3.6, enabling new simulation capabilities while benefiting from the unified high-level frontend functionality of the rest of the library.

3.2.3 Bosonic Qubits

Residing in a two-dimensional subspace of the infinite-dimensional Hilbert space of a CV mode, the bosonic qubit is robust unit for quantum computation in platforms such as photonics. Several classes of bosonic qubits can moreover correct errors within the CV space, adding an additional level of protection against physical noise. Here, we review the definitions and main properties of two promising encodings: GKP and cat states. Understanding how these states behave in realistic settings is especially valuable as their use becomes more widespread in quantum technologies. As we show in Section 3.4 and Appendix B.2, we are able to

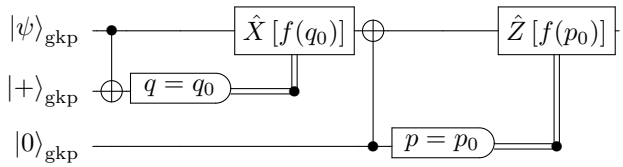


Figure 3.1: Error syndrome measurement with normalizable GKP states following the approach from [13]. First, shifts in q are corrected: an encoded data qubit $|\psi\rangle_{\text{gkp}}$ and an ancilla $|+\rangle_{\text{gkp}}$ are sent through a SUM gate, and $|\psi\rangle_{\text{gkp}}$ is displaced according to the result of a homodyne q measurement on the ancilla, mod $\sqrt{\pi\hbar}$. A similar procedure follows for shifts in p.

express these qubits as linear combinations of Gaussian functions in phase space, allowing us to model them under realistic conditions such as loss, finite-energy, and noisy gate teleportations.

GKP States

The ideal square-lattice GKP logical states are defined as infinite combs of Dirac delta functions spaced by $2\sqrt{\pi\hbar}$ in the position quadrature:

$$|k\rangle_{\text{gkp}} = \sum_{s=-\infty}^{\infty} |\sqrt{\pi\hbar}(2s+k)\rangle_q\,,\ k = \{0,1\}, \tag{3.19}$$

where $|\cdot\rangle_q$ denotes an eigenstate of the position quadrature and k denotes the logical value. Rectangular and hexagonal lattice encodings are related to the square lattice via symplectic transformations [13], a mapping that we show falls neatly within our formalism. One advantage of the GKP encoding is that Clifford gates and measurements correspond to Gaussian transformations [13], which are experimentally accessible in the photonics context, as we review in Appendix B.5. Pauli X and Z gates correspond to displacements by $\sqrt{\pi\hbar}$ along the q and p quadratures, respectively. The Hadamard gate is a rotation by $\pi/2$ in phase space. The qubit phase, CX, and CZ gates correspond to a CV quadratic phase, CX, and CZ gates, which are active Gaussian transformations, in the sense of requiring a squeezing component. In practice, (measurement-based) inline squeezers are challenging to implement but are nonetheless feasible and deterministic, as we discuss in Section 3.5.2. Pauli X and Z measurements correspond to homodyne measurements along q and p quadratures. A universal gate set can be completed with the qubit T gate; in the GKP encoding, this gate can be implemented through gate teleportation with a magic state, a process we review and align with our formalism in Section 3.5.3.

GKP states can correct small displacement errors in phase space, and and can reduce larger displacement errors to qubit-level Pauli errors. A qubit error-correction code concatenated with GKP states can then be used to correct these discrete errors [13]. In Fig. 3.1 we review the GKP error-correction circuit from [13], noting that various other decompositions of the circuit exist [33, 115]. Briefly, two ancillary GKP states are entangled with the data mode to be corrected and measured with homodyne detectors; the outcomes of these measurements determine the displacement that is then applied to the data mode to correct for the error (up to a logical Pauli error). We discuss in Section 3.5.3 how our formalism can treat this circuit.

To the chagrin of experimentalists, ideal GKP qubits have infinite energy; to the chagrin of theorists, we must consider their finite-energy, normalizable forms. One such form is obtained by replacing each Dirac delta with a Gaussian peak corresponding to a squeezed state of variance $\Delta^2/2$, and then applying an overall

Gaussian envelope of width $1/\Delta^2$ so that peaks further from the origin are suppressed [13]:

$$|k^{\Delta}\rangle_{\text{gkp}} \equiv \frac{1}{\mathcal{N}_k} \int_{-\infty}^{+\infty} dx \sum_s e^{-\Delta^2[(2s+k)\sqrt{\pi}]^2/2} e^{-[x-(2s+k)\sqrt{\pi\hbar}]^2/2\hbar\Delta^2} |x\rangle_q \,. \tag{3.20}$$

This normalization process is not symmetric in phase space because the peaks are constrained to remain centred at the initial positions of the ideal state [50]. A related normalization process, which has the Fock damping operator $E(\epsilon) = e^{-\epsilon\hat{n}}$ applied to the ideal state, is symmetric since the number operator $\hat{n}$ acts symmetrically in phase space; we denote such states as $|k^{\epsilon}\rangle_{\text{gkp}}$. In [50] the authors provide a thorough review of the connections and mappings between these finite energy forms of GKP states, noting that they can be related to each other by a simple squeezing operation; in [51] the authors explore alternative normalization envelopes to Gaussians, demonstrating sufficient conditions for the normalization process to yield physical states. Yet another option for finite energy GKP states are comb states [116]; these correspond to taking only a finite superposition of evenly-weighted q-squeezed states centred at the location of the peaks in the ideal state.

Cat States

To define cat states, one starts with coherent states $|\alpha\rangle = \hat{D}(\alpha)|0\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle$, $\alpha \in \mathbb{C}$, which are Gaussian states with covariance matrix equal to the vacuum covariance matrix $\boldsymbol{\Sigma}_{\text{vac}} = \frac{\hbar}{2}\mathbb{1}$ and displacement vector $\boldsymbol{\mu} = [\sqrt{2\hbar}\Re(\alpha), \sqrt{2\hbar}\Im(\alpha)]$. Ideal (two-lobe) cat states are superpositions of coherent states [96]:

$$|k^{\alpha}\rangle_{\text{cat}} = \sqrt{\mathcal{N}}(|\alpha\rangle + e^{i\pi k}|-\alpha\rangle), \;\; k = \{0, 1\}, \tag{3.21}$$

with normalization

$$\mathcal{N} = \frac{1}{2(1 + e^{-2|\alpha|^2}\cos(\pi k))}. \tag{3.22}$$

There is some freedom in the choice of logical basis states. For example, for resource-efficient preparation of Pauli X eigenstates, one can identify them with coherent states $|\pm\alpha\rangle$, which are approximately orthogonal as α increases in magnitude. Then, the Pauli Z gate is given by a rotation in phase space by π. Pauli Z measurements become photon number parity measurements, since the wavefunction for $k = 0$ (1) is symmetric (antisymmetric) and therefore only contains even (odd) photon numbers. A cat-qubit Bell state can be prepared by splitting a higher-energy cat state $|k^{\sqrt{2}\alpha}\rangle_{\text{cat}}$ at a 50:50 beam-splitter. Bell state measurements on cat qubits can be performed by interacting two states at a beam-splitter, then measuring photon number patterns at the output. The Pauli X gate, small single qubit rotations, and two-qubit entangling gates can all be applied deterministically via gate teleportation, conditional on the availability of cat Bell states [96]. In Appendix B.2.1, we show how cat states can be written as a linear combination of Gaussian functions in phase space, and in Appendix B.2.2 we show the same for Fock states. This means that our formalism enables simulation of the states, teleportation-based gates and measurements required for quantum computation with cat states.

3.3 Linear Combinations of Gaussians in Phase Space

Having reviewed the relevant background material on CV Gaussian formalism in Section 3.2.1, we present a new formalism for simulating a wide class of CV states, transformations and measurements. Specifically, in

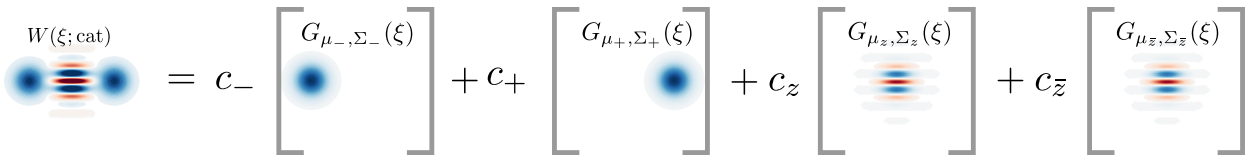


Figure 3.2: Gaussian decomposition of the Wigner function of the (non-Gaussian) cat state. We use the labels $\mathcal{M} = \{+, -, z, \bar{z}\}$ for the four Gaussians needed. Note that the last two terms have complex coefficients and means and are also complex conjugates of each other; thus the imaginary part of the Gaussians cancel out and we only plot the real part. The details of this particular decomposition are provided in Appendix B.2.1. For an introductory tutorial implemented by the authors on using the `bosonic` backend of `Strawberry Fields` to obtain this figure, see [103].

Section 3.3.1, we introduce states with Wigner functions that can be expressed as a linear combination of Gaussian functions in phase space. Next, in Section 3.3.2, we provide the framework for describing Gaussian and a class of non-Gaussian measurements on the aforementioned states. Finally, in Section 3.3.3, we detail how the states in our formalism transform under deterministic and conditional Gaussian maps, as well as under a class of non-Gaussian transformations.

3.3.1 States in the Wigner Representation

In this chapter, we consider n-mode states whose Wigner function can be written as a linear combination of Gaussian functions in phase space:

$$W(\boldsymbol{\xi};\hat{\rho}) = \sum_{m\in\mathcal{M}} c_m G_{\boldsymbol{\mu}_m,\boldsymbol{\Sigma}_m}(\boldsymbol{\xi}), \tag{3.23}$$

where $\mathcal{M}$ is the set of indices which can in general be multiparametered, $\boldsymbol{\xi} \in \mathbb{R}^{2n}$ is the phase space variable for an n-mode CV quantum system, and $\hat{\rho}$ is the corresponding density matrix operator in Hilbert space. Each Gaussian in the linear combination is associated with a weight c_m, a $2n$-dimensional mean $\boldsymbol{\mu}_m$ and a $2n \times 2n$ covariance matrix $\boldsymbol{\Sigma}_m$, all complex-valued in general. The normalized multivariate Gaussian distribution G is defined as

$$G_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(\boldsymbol{\xi}) \equiv \frac{\exp\left[-\frac{1}{2}(\boldsymbol{\xi}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\xi}-\boldsymbol{\mu})\right]}{\sqrt{\det(2\pi\boldsymbol{\Sigma})}}. \tag{3.24}$$

If $\hat{\rho}$ is a physical density matrix—in particular, if it has unit trace—then its corresponding Wigner function is normalized through

$$\sum_{m\in\mathcal{M}} c_m = 1. \tag{3.25}$$

At this point, we do not restrict our means, covariances and weights to be real, as imaginary components from these quantities can be required to produce interference fringes and negativity in the Wigner function; however, Hermiticity of the density matrix implies, at least, that the total Wigner function is real. Furthermore, we require the real part of the covariance matrices to be positive-definite so that the distribution remains bounded; however, the covariance matrices need not always respect the uncertainty relation from Eq. (3.9).

As an example of a wide class of states that fall into this framework, consider a pure state that consists of

a superposition of Gaussian pure states (i.e. displaced squeezed vacuum states):

$$|\psi\rangle = \sum_n \kappa_n |\gamma_n, \zeta_n\rangle = \sum_n \kappa_n \hat{D}(\gamma_n)\hat{S}(\zeta_n)|0\rangle . \tag{3.26}$$

Here, γ_n and ζ_n are the complex displacement and squeezing parameters of the n^{th} term in the superposition; in this context $\hat{D}(\gamma)$ is the displacement operator and $\hat{S}(\zeta)$ is the squeezing operator (cf. Eq. (3.13)). The density matrix for this state is a sum of terms of the form $\kappa_m\kappa_n^* |\gamma_m, \zeta_m\rangle \langle\gamma_n, \zeta_n|$. Since the Weyl transform is linear, the Wigner function for the state is a linear combination of functions in phase space, each associated with an operator $|\gamma_m, \zeta_m\rangle \langle\gamma_n, \zeta_n|$. For $m = n$, the phase space functions are simply products of $|\kappa_n|^2$ and the Wigner function for $|\gamma_n, \zeta_n\rangle$, which is a Gaussian state and hence has a Gaussian Wigner function. For $m \neq n$, the phase space function is still Gaussian, albeit with complex weight, means and covariances (see Appendix A of [21]). Thus, states of the form (3.26) can be expressed in the form (3.23).

As we discuss in Section 3.4 and Appendix B.2, the representation in Eq. (3.23) is useful for describing salient families of continuous variable states that can act as bosonic qubits. Using the derivation from the previous paragraph, as well as additional, tailored derivations, we show how to write GKP and cat qubits as linear combinations of Gaussian functions in phase space. Moreover, we show how Fock states, as well as superpositions of Fock states created by performing photon-number-resolving measurements on some modes of a multimode Gaussian state can be expressed in the form of Eq. (3.23). In addition to giving us the ability to write down states useful for quantum computing, our formalism is well-suited to subjecting such states to general Gaussian transformations and measurements, as well as certain non-Gaussian transformations via gate teleportation, as we explore in the next few sections.

3.3.2 Gaussian and a Class of Non-Gaussian Measurements

Given a state Wigner function written as a linear combination of Gaussians in phase space, we now describe the formalism for Gaussian measurements on such states. A general-dyne Gaussian measurement on n modes is characterized by the $2n \times 2n$ covariance matrix $\boldsymbol{\Sigma}_M$ of the Gaussian state onto which one projects. The outcome of a general-dyne measurement is a point in phase space, $\boldsymbol{r}_M$ [109]. Given a state that can be described by Eq. (3.23), the probability of outcome $\boldsymbol{r}_M$ is

$$p(\boldsymbol{r}_M; \hat{\rho}, \boldsymbol{\Sigma}_M) = (2\pi\hbar)^N \int d\boldsymbol{\xi}\, W(\boldsymbol{\xi}; \hat{\rho})\, G_{\boldsymbol{r}_M, \boldsymbol{\Sigma}_M}(\boldsymbol{\xi}) = \sum_{m\in\mathcal{M}} c_m G_{\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_M + \boldsymbol{\Sigma}_m}(\boldsymbol{r}_M) . \tag{3.27}$$

A special case of general-dyne measurement is homodyne measurement. Without loss of generality, we consider a measurement of the $\boldsymbol{q}$ quadrature, for which

$$\boldsymbol{\Sigma}_M = \lim_{\epsilon\to 0} \frac{\hbar}{2} \begin{pmatrix} \epsilon\mathbb{1} & \mathbf{0} \\ \mathbf{0} & \epsilon^{-1}\mathbb{1} \end{pmatrix} , \tag{3.28}$$

and the measurement outcome becomes $\boldsymbol{r}_M \to \boldsymbol{q}_M$. Since the $\boldsymbol{q}$-homodyne distribution can be retrieved by integrating out the $\boldsymbol{p}$ quadrature, and since the Wigner function is a linear combination of Gaussians, we have

$$p(\boldsymbol{q}_M; \hat{\rho}_M) = \sum_{m\in\mathcal{M}} c_m G_{\boldsymbol{\mu}_m^{(\boldsymbol{q})}, \boldsymbol{\Sigma}_m^{(\boldsymbol{qq})}}(\boldsymbol{q}_M) , \tag{3.29}$$

where $\boldsymbol{\mu}_m^{(\boldsymbol{q})}$ and $\boldsymbol{\Sigma}_m^{(\boldsymbol{qq})}$ denote the $\boldsymbol{q}$-quadrature components of the means and covariances. Importantly, the

distribution is yet another linear combination of Gaussians, this time of a variable in n rather than $2n$ dimensions. Later, in Section 3.6, we provide a tailored simulation method for sampling outcomes of Gaussian measurements from states in our formalism.

The mathematical method for calculating the probability distribution of a Gaussian measurement can be straightforwardly generalized to a class of non-Gaussian measurements for which the measurement operator can itself be represented in phase space as a linear combination of Gaussian functions. Assuming the Weyl transform of a measurement operator $\hat{M}$ associated with outcome M is of the form

$$\Phi(\boldsymbol{\xi}; \hat{M}) = \sum_{j \in \mathcal{J}} d_j G_{\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j}(\boldsymbol{\xi}), \tag{3.30}$$

the probability of obtaining M is simply given by:

$$p(M; \hat{\rho}) = (2\pi\hbar)^N \int d\boldsymbol{\xi} \Phi(\boldsymbol{\xi}; \hat{M}) W(\boldsymbol{\xi}; \hat{\rho}) = \sum_{m \in \mathcal{M}} \sum_{j \in \mathcal{J}} c_m d_j G_{\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m + \boldsymbol{\Sigma}_j}[\boldsymbol{\mu}_j]. \tag{3.31}$$

Note that, for measurement operators, we do not necessarily have the unit trace condition, so $\sum_{j \in \mathcal{J}} d_j \neq 1$ in general. In Appendix B.2.2, we show, as a pertinent example, how Fock states can be expressed as a linear combination of Gaussians in phase space, which means photon-number-resolving measurements can be described with this formalism.

3.3.3 Gaussian and a Class of Non-Gaussian Transformations in a Gaussian-Inspired Framework

As we saw in Section 3.2.1, Gaussian transformations in phase space map Gaussian states to Gaussian states. As we show next, these maps motivate a wider class of Gaussian and non-Gaussian transformations for states with Wigner functions that can be represented as linear combinations of Gaussian functions, as in Eq. (3.23).

The first class of transformations we consider are deterministic Gaussian completely positive and trace preserving (CPTP) maps, that is, transformations that are not conditioned on any probabilistic measurement outcomes. As we reviewed in Section 3.2.1, deterministic Gaussian transformations acting on an n-mode Gaussian state can be parametrized by two $2n \times 2n$ matrices, $\boldsymbol{X}$ and $\boldsymbol{Y}$, and a length-$2n$ vector $\boldsymbol{d}$ that together transform the covariance matrix and mean of the Wigner function [109]. Since the mapping is linear over phase space variables, it generalizes straightforwardly for a linear combination of Gaussian functions in phase space:

$$\boldsymbol{\Sigma}_m \to \boldsymbol{X}\boldsymbol{\Sigma}_m\boldsymbol{X}^T + \boldsymbol{Y}, \boldsymbol{\mu}_m \to \boldsymbol{X}\boldsymbol{\mu}_m + \boldsymbol{d}. \tag{3.32}$$

When $\boldsymbol{X}$ is a symplectic matrix and $\boldsymbol{Y} = \boldsymbol{0}$, this corresponds to a Gaussian unitary transformation.

Deterministic Gaussian transformations modify neither the number nor the weighting of the peaks in the linear combination, which makes them easy to apply to states of the form (3.23). A wide class of CV operations—displacement, squeezing, rotation, and beam-splitters—as well as common noise models—loss and Gaussian random displacements—fall under this umbrella. By contrast, simple transformations such as displacements and squeezing can quickly push Fock distributions beyond the energy cutoff, an important limitation of the Fock representation.

The second class of transformations we consider are conditional dynamics: how a measurement of some modes updates the remaining active modes. In this case, our formalism opens the door to a class of non-Gaussian transformations of the Wigner function; if we take a set of target modes of the form (3.23) and

interact them with a set of non-Gaussian ancillae, also of the form (3.23), and then perform a non-Gaussian measurement on the ancillary modes of the form (3.30), then the effective transformation on the target modes is non-Gaussian in general. This conclusion applies even if the ancillary modes or the measurement—but not both—are Gaussian. Effecting a non-Gaussian transformation on a state through an interaction with non-Gaussian ancillary modes or through a non-Gaussian measurement has been studied extensively in the context of CV gate teleportation and state preparation [13, 16, 17, 19, 117–121]. However, we show next that we can represent such non-Gaussian transformations in the spirit of conditional Gaussian dynamics reviewed in Section 3.2.1.

To understand these Gaussian-inspired but nonetheless non-Gaussian transformations, consider two sets of modes: the set A of active modes, described initially by a Wigner function of the form (3.23) with weights a_ℓ, means $\boldsymbol{\mu}_\ell$ and covariances $\boldsymbol{\Sigma}_\ell$, for $\ell \in \mathcal{L}$; and the set B of modes that will eventually be measured, with corresponding parameters b_k, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$, for $k \in \mathcal{K}$. If the two sets of modes are entangled via a deterministic Gaussian transformation parametrized by $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{d})$, the weights, means and covariances become:

$$\begin{aligned} c_m &= a_\ell b_k, \;\; m = (\ell, k) \in \mathcal{M} = (\mathcal{L}, \mathcal{K}), \\ \boldsymbol{\mu}_m &= \boldsymbol{X}(\boldsymbol{\mu}_\ell \oplus \boldsymbol{\mu}_k) + \boldsymbol{d}, \\ \boldsymbol{\Sigma}_m &= \boldsymbol{X}(\boldsymbol{\Sigma}_\ell \oplus \boldsymbol{\Sigma}_k)\boldsymbol{X}^T + \boldsymbol{Y}. \end{aligned} \tag{3.33}$$

In turn, we can express the means and covariances as

$$\boldsymbol{\Sigma}_m = \begin{pmatrix} \boldsymbol{\Sigma}_{m,A} & \boldsymbol{\Sigma}_{m,AB} \\ \boldsymbol{\Sigma}_{m,AB}^T & \boldsymbol{\Sigma}_{m,B} \end{pmatrix}, \;\; \boldsymbol{\mu}_m = \begin{pmatrix} \boldsymbol{\mu}_{m,A} \\ \boldsymbol{\mu}_{m,B} \end{pmatrix}, \tag{3.34}$$

where A and B indicate the active and measured modes, respectively.

Consider now a measurement $\hat{M}$ with outcome M on modes B with a phase space representation of the form from Eq. (3.30), parametrized by weights d_j, means $\boldsymbol{\mu}_j$ and covariances $\boldsymbol{\Sigma}_j$, with $j \in \mathcal{J}$. Since the partial trace is linear, the corresponding phase space integral is a sum of many partial Gaussian integrals. Thus the covariances and means of modes A update as [109]:

$$\begin{aligned} \boldsymbol{\Sigma}_{m,A} &\to \boldsymbol{\Sigma}_{m,A} - \boldsymbol{\Sigma}_{m,AB}(\boldsymbol{\Sigma}_{m,B} + \boldsymbol{\Sigma}_j)^{-1}\boldsymbol{\Sigma}_{m,AB}^T, \\ \boldsymbol{\mu}_{m,A} &\to \boldsymbol{\mu}_{m,A} + \boldsymbol{\Sigma}_{m,AB}(\boldsymbol{\Sigma}_{m,B} + \boldsymbol{\Sigma}_j)^{-1}(\boldsymbol{\mu}_j - \boldsymbol{\mu}_{m,B}). \end{aligned} \tag{3.35}$$

Until this point, the update rules have been inspired by the conditional dynamics for Gaussian states. However, our consideration of multiple Gaussians instead of just one necessitates a novel rule: an expansion and re-weighting of the peaks in phase space. While outcome M occurs with probability $p(M; \hat{\rho}_{AB})$, as in Eq. (3.31), each peak from modes B and from the measurement operator contribute differently to this probability, with a weight given by:

$$w(M|\boldsymbol{\mu}_{m,B}, \boldsymbol{\Sigma}_{m,B}, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j) = c_m d_j G_{\boldsymbol{\mu}_{m,B}, \boldsymbol{\Sigma}_{m,B} + \boldsymbol{\Sigma}_j}(\boldsymbol{\mu}_j). \tag{3.36}$$

Therefore, given the result M, the weights update as:

$$c_m \to \gamma_{m,j} = \frac{w(M|\boldsymbol{\mu}_{m,B}, \boldsymbol{\Sigma}_{m,B}, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}{p(M; \hat{\rho}_{AB})}, \tag{3.37}$$

where we include a normalization by all the new weights. As a result of this process, the total number of

Gaussians in the Wigner function for modes A, as well as their associated weights, means and covariances, is been grown both by modes B and by the measurement M:

$$\ell \xrightarrow{B} m = (\ell, k) \xrightarrow{M} (m, j) = (\ell, k, j) \in (\mathcal{L}, \mathcal{K}, \mathcal{J}), \tag{3.38}$$

such that the final number of Gaussian functions required to describe mode A is the product of the initial number of functions in modes A and B, and in the Weyl transform of $\hat{M}$.

While we have only been tracking transformations of Gaussian functions in phase space, it is worth emphasizing that these conditional dynamics increase the number of Gaussian functions initially in modes A, thereby necessarily describing non-Gaussian transformations. The cost of modelling non-Gaussian transformations with our formalism is that the number of peaks one needs to consider can grow; still, we show that this is a worthwhile trade-off compared to alternative methods for simulating certain classes of bosonic qubits.

In the case that modes B and the measurement are truly Gaussian states, as opposed to linear combinations of Gaussian functions, then $b_k = d_j = 1$, and $(\mathcal{L}, \mathcal{K}, \mathcal{J}) \to \mathcal{L}$. This means the initial number of weights does *not* increase, and the initial means $\boldsymbol{\mu}_\ell$ and covariances $\boldsymbol{\Sigma}_\ell$ follow the traditional Gaussian conditional update rule reviewed in Section 3.2.1:

$$\begin{aligned} \boldsymbol{\Sigma}_{m,A} &\to \boldsymbol{\Sigma}_{m,A} - \boldsymbol{\Sigma}_{m,AB}(\boldsymbol{\Sigma}_{m,B} + \boldsymbol{\Sigma}_M)^{-1}\boldsymbol{\Sigma}_{m,AB}^T, \\ \boldsymbol{\mu}_{m,A} &\to \boldsymbol{\mu}_{m,A} + \boldsymbol{\Sigma}_{m,AB}(\boldsymbol{\Sigma}_{m,B} + \boldsymbol{\Sigma}_M)^{-1}(\boldsymbol{r}_M - \boldsymbol{\mu}_{m,B}), \end{aligned} \tag{3.39}$$

where $\boldsymbol{\Sigma}_M$ and $\boldsymbol{r}_M$ parametrize the Gaussian measurement as in Eq. (3.27). The peak reweighting from Eq. (3.37) is still required; however, it is simpler because the total number of weights does not increase and $d_j = 1$.

In Section 3.5, we explore how various transformations such as loss, Fock damping, measurement-based squeezing, and gate teleportation onto bosonic qubits can be described in this formalism. Before this, in the following section, we present the description of GKP states as linear combinations of Gaussians in phase space.

3.4 GKP States Expressed as Linear Combinations of Gaussians

Having provided a general formalism in Section 3.3, we now demonstrate how GKP qubits can be written in the form of Eq. 3.23. In Appendix B.2, we provide the details for how other bosonic qubits, namely cat and Fock states, can also be written as linear combinations of Gaussians.

Care must be taken when dealing with the finite-energy forms of GKP states. Before we tackle those, let us recall the Wigner representation of the ideal states.

3.4.1 Ideal GKP States

Any pure GKP state can be expressed in the Bloch sphere representation as

$$|\psi\rangle = \cos\tfrac{\theta}{2}|0\rangle_{\text{gkp}} + e^{-i\phi}\sin\tfrac{\theta}{2}|1\rangle_{\text{gkp}}. \tag{3.40}$$

The Wigner function of an ideal single-mode square-lattice GKP state can be expressed in terms of the parameters of Eq. (3.23) in the following way [13]:

$$\begin{aligned} \mathcal{M} &= \{m \equiv (k\sqrt{\pi\hbar}/2, \ell\sqrt{\pi\hbar}/2) \,|\, k, \ell \in \mathbb{Z}\}, \\ \boldsymbol{\Sigma}_m &= \lim_{\delta \to 0+} \delta \mathbb{1}, \ \boldsymbol{\mu}_m = m. \end{aligned} \tag{3.41}$$

The weights $c_m(\theta, \phi)$—which encode the logical content of the state—are presented in Appendix B.1 as given in Ref. [53]. Recall each Gaussian peak in the Wigner function is a function of $\boldsymbol{\Sigma}_m^{-1}$, so we cannot directly evaluate the limit as $\delta \to 0+$. The Gaussian distributions for the ideal GKP Wigner function are Dirac delta functions located at the lattice points enumerated in $\mathcal{M}$. This means that the ideal GKP state has infinite energy and cannot be normalized, rendering it unphysical. However, the covariances matrices do not vary with the index and are proportional to the identity, a feature that is put to use in the following section. Note that GKP states corresponding to alternative lattice spacings, such as rectangular or hexagonal GKP states, can be related to square-lattice states by symplectic transformations in phase space [13]. As we showed in Section 3.3.3, symplectic transformations are straightforward to apply to states that can be expressed as linear combinations of Gaussians in phase space, so it is easy to transform between lattices in our framework. Additionally, GKP qudits simply correspond to a different linear combination of δ-functions in phase space, so our formalism can be employed to treat those states as well.

3.4.2 Finite-Energy GKP States

There are different ways to obtain finite-energy versions of the GKP states. A summary of these is presented in Fig. 1 of Ref. [16]. Here, we focus on a commonly-used model of finite-energy GKP states: a Fock damping operator

$$\hat{E}(\epsilon) = e^{-\epsilon \hat{n}}, \ \epsilon > 0, \tag{3.42}$$

applied to the ideal GKP state in Eq. (3.19). Because $\hat{E}(\epsilon)$ is a single non-unitary Kraus operator, it is non-trace-preserving, meaning one needs to carefully account for the normalization of the resulting states.

A derivation motivated by the dilation of the operator $\hat{E}(\epsilon)$ (provided in Appendix B.3.1) shows that the Wigner function corresponding to the $\hat{E}(\epsilon)$ operator applied to an ideal state can be represented in the same form as Eq. (3.23). With the set $\mathcal{M}$, ideal coefficients $c_m(\theta, \phi)$, and ideal means $\boldsymbol{\mu}_m$ given in Eq. (3.41), we have that:

$$\begin{aligned} c_m(\epsilon; \theta, \phi) &= \frac{c_m(\theta, \phi)}{\mathcal{N}_\epsilon} \exp\left[-\frac{1 - e^{-2\epsilon}}{\hbar(1 + e^{-2\epsilon})} \boldsymbol{\mu}_m^T \boldsymbol{\mu}_m\right], \\ \boldsymbol{\mu}_m(\epsilon) &= \frac{2e^{-\epsilon}}{1 + e^{-2\epsilon}} \boldsymbol{\mu}_m, \ \boldsymbol{\Sigma}_m(\epsilon) = \frac{\hbar}{2} \frac{1 - e^{-2\epsilon}}{1 + e^{-2\epsilon}} \mathbb{1}. \end{aligned} \tag{3.43}$$

Here $\mathcal{N}_\epsilon$ is chosen such that $\sum_{m \in \mathcal{M}} c_m(\epsilon; \theta, \phi) = 1$. In this representation, the weights and means are real. In practice, one can take a finite number of terms from $\mathcal{M}$—a numerical cutoff—depending on the desired precision. The value of the cutoff would depend on the strengths of the weights and the normalization of the Gaussian distributions in the Wigner expansion.

We compare the model for finite-energy GKP states we have derived in Eq. 3.43 to the approach from [34]. There, the Dirac delta spikes of the ideal GKP Wigner function are replaced with Gaussians of non-zero variance. While that approach can capture the broadening of peaks due to finite-energy and other noise effects, and while the covariance matrix of each peak can be updated as Gaussian channels are applied, the

state does not have an envelope, so it still has infinite energy and remains unphysical. Moreover, in that model, the locations of peaks are not tracked, so their contraction towards the origin due to finite energy effects and any shifts under Gaussian channels are ignored. This can have an impact on the estimation of qubit-level errors. By contrast, the mathematical form for finite-energy GKP states provided here offers a rich noise model that captures these effects.

For an alternative phase-space representation of finite-energy GKP states, one can apply $\hat{E}(\epsilon)$ to the wavefunction to obtain a superposition of squeezed states, and then use this form to calculate the Wigner function directly. This representation differs from the one previously presented in two ways. First, the weights and means of the resulting Wigner function are complex rather than real numbers. Second, the covariances of the individual Gaussian functions now respect the uncertainty relation. Since in this representation, each peak is directly mappable to squeezed states in the wave function, it can be more useful for converting effects observed at the wavefunction level to phase space transformations. The derivation of this alternative representation is provided in Appendix B.3.2.

Let us denote a logical GKP qubit by $|\psi(\boldsymbol{a})\rangle = a_0|0\rangle_{\text{gkp}} + a_1|1\rangle_{\text{gkp}}$. We find that the parameters from Eq. (3.23) become

$$\begin{aligned}
&\mathcal{M} = \{m \equiv (k,\ell,s,t)\,|\,s,t \in \{0,1\}\,\&\,k,\ell \in \mathbb{Z}\},\\
&\boldsymbol{\mu}_m(\epsilon) = -\frac{\beta\sqrt{\pi\hbar}}{2}\begin{pmatrix}\frac{(t+s+2\ell+2k)}{\alpha}\\ i(s-t+2\ell-2k)\end{pmatrix},\\
&\boldsymbol{\Sigma}_m(\epsilon) = \frac{\hbar}{2}\begin{pmatrix}1/\alpha & 0\\ 0 & \alpha\end{pmatrix},\ (\alpha,\beta) = (\coth(\epsilon), -\operatorname{csch}(\epsilon)),\\
&c_m = \frac{a_s a_t^*}{\mathcal{N}(\boldsymbol{a},\epsilon)}\exp\left\{-\frac{\alpha\pi}{2}\left[(s+2\ell)^2+(t+2k)^2\right]\right\}\exp\left[\frac{\beta^2\pi(t+s+2\ell+2k)^2}{4\alpha}\right].
\end{aligned} \tag{3.44}$$

Here $\mathcal{N}(\boldsymbol{a},\epsilon)$ is an overall normalization constant such that $\sum_{m\in\mathcal{M}} c_m = 1$; it depends on the choice of the GKP state and the strength of the Fock-damping operator applied to it. This time, the Gaussians all have the same covariance matrix, and this matrix respects the uncertainty relation [107]. While the coefficients can be complex, this only stems from the initial qubit coefficients a_0, a_1. Finally, we can see that the means $\boldsymbol{\mu}_m(\epsilon)$ are complex; the resulting sinusoidal oscillations in phase space are what generate the interference fringes and negative regions of the Wigner function in this representation.

3.4.3 Probabilistically prepared GKP states

We recall the Gaussian Boson Sampling-type state preparation devices from Chapter 2, wherein Fock measurements on all but one mode of a multimode Gaussian state herald a superposition of Fock states in the remaining mode. The exact form of the superposition can be tuned using the means and covariance matrix of the multimode Gaussian, as well as the choice of which Fock measurements to postselect upon [16, 17, 19, 20]. As the input state is Gaussian, and we can approximate Fock measurements to arbitrarily high accuracy as linear combinations of Gaussians (see Appendix B.2, the output state can be calculated using the partial trace form of Eq. (3.16); this materializes as a linear combination of Gaussian integrals over all but the two phase space quadratures of the output mode, yielding a linear combination of Gaussian functions in phase space for that mode. Notably, this state preparation device can be used as a means to prepare bosonic qubits, including GKP and cat states.

3.5 Useful Transformations and Measurements in a Gaussian-Inspired Framework

Given our formalism from Section 3.3, and examples from Section 3.4 and Appendix B.2 for writing bosonic qubits as linear combinations of Gaussians, we now examine how our formalism can be used to treat useful transformations for studying quantum computing with bosonic qubits, namely loss, measurement-based squeezing and non-Gaussian gate teleportation.

3.5.1 Loss Channel and Fock Damping

Loss is a physical process modeled as an interaction with a thermal environment through a beam-splitter transformation, resulting in a bosonic Gaussian channel. The loss channel is closely related to the additive random noise channel by composition with an amplifier channel [122–125]. The strength of the loss parameter of the channel is set by the beam-splitter angle. The pure loss channel is easily described in the form (3.32) with the matrix pair

$$(\boldsymbol{X}_\eta, \boldsymbol{Y}_\eta) = (\sqrt{\eta}\mathbb{1}, (1-\eta)\hbar\mathbb{1}/2), \tag{3.45}$$

where η is sometimes referred to as the transmission or transmissivity parameter, and where the environment is assumed to have zero temperature. Thermal loss—interaction with a thermal environment with covariance matrix $\hbar(\bar{n}+\frac{1}{2})\mathbb{1}$—can also be incorporated into this description by multiplying $\boldsymbol{Y}_\eta$ by $(2\bar{n}+1)$. The Kraus operators and other details of the pure loss channel are provided in great detail in [125]. Loss is the dominant imperfection in the photonics context.

The Fock damping or finite-energy operator is given by (3.42). It is valuable for converting infinite-energy states, such as ideal position and momentum eigenstates and GKP states, into their normalizable, finite-energy forms. The operator is directly linked to the loss channel as it forms a leading-order Kraus operator of the channel (see Eq. 4.6 of [125]), while maintaining the purity of the state. In [28], in the context of GKP states, it was shown that $\hat{E}(\epsilon)$ can be derived by passing a state through a beam-splitter of transmissivity $\cos\theta = e^{-\epsilon}$ with an ancillary vacuum state, and postselecting the ancillary mode on vacuum (see also Appendix B1 of [16] for a simple derivation of this result).

As we show in Appendix B.4, the use of only Gaussian states and operations in deriving $\hat{E}(\epsilon)$ makes the operator fit neatly into the formalism for Gaussian transformations developed in Section 3.3.3. If the initial means and covariances of the mode are $(\boldsymbol{\mu}_{m,0}, \boldsymbol{\Sigma}_{m,0})$, then we can simply use Eq. (3.34) with:

$$\boldsymbol{\Sigma}_m = \boldsymbol{S}_\theta \begin{pmatrix} \boldsymbol{\Sigma}_{m,0} & \mathbf{0} \\ \mathbf{0} & \hbar\mathbb{1}/2 \end{pmatrix} \boldsymbol{S}_\theta^T, \ \boldsymbol{\mu}_m = \boldsymbol{S}_\theta \begin{pmatrix} \boldsymbol{\mu}_{m,0} \\ \mathbf{0} \end{pmatrix}, \tag{3.46}$$

where $\boldsymbol{S}_\theta = \begin{pmatrix} \cos\theta\mathbb{1} & \sin\theta\mathbb{1} \\ -\sin\theta\mathbb{1} & \cos\theta\mathbb{1} \end{pmatrix}$ is the symplectic matrix for a beam-splitter assuming a mode-wise ordering (q_1, p_1, q_2, p_2). From there one can proceed with the update rule in Eq. (3.39), as well as the per-peak reweighting in Eq. (3.37) with $d_j = 1$, $\boldsymbol{r}_M = \mathbf{0}$, and $\boldsymbol{\Sigma}_M = \hbar\mathbb{1}/2$, since the projection is onto vacuum.

3.5.2 Squeezed Ancilla-Assisted Gates

While passive Gaussian transformations can be effected in optics using phase shifters and beam-splitters, and while squeezing applied to vacuum can be achieved with nonlinear cavity resonators [126] or waveguides [127],

inline squeezing—squeezing applied directly to an arbitrary input state—poses a greater challenge. A feasible approach to inline squeezing is possible by consuming an ancillary squeezed vacuum state [101]. To effect squeezing in q (p), one first interferes the target state with a highly q-squeezed (p-squeezed) vacuum on a beam-splitter parametrized by angle θ. Next, one performs a p-homodyne (q-homodyne) measurement on the ancillary mode with a detector of efficiency η, producing result p_M. Finally, one applies a feedforward p-displacement (q-displacement) by $\sqrt{\eta^{-1}}\tan\theta p_M$ to the target mode. As we discuss next, the resulting average map effects a transformation equivalent to squeezing the target mode by $\cos\theta$ in q, along with some noise dependent on the level of squeezing of the ancillary state and the efficiency of the homodyne detection. Squeezing along any quadrature can be achieved by preceding and following inline squeezing with rotations.

Since the ancillary state, the transformation between the target and ancilla, the measurement on the ancilla, and the feedforward are all Gaussian, inline squeezing using a squeezed ancillary state—commonly referred to as measurement-based squeezing—falls within the formalism developed in Section 3.3.3. In Appendix B.5, we derive the resulting transformation on the covariance matrices, means and modes of the linear combination of Gaussians in phase space that form the target state. On average (integrating over p_M), the map due to this type of squeezing is a deterministic Gaussian CPTP map parametrized by:

$$\boldsymbol{X}_s^{(q)} = \begin{pmatrix} \cos\theta & 0 \\ 0 & \frac{1}{\cos\theta} \end{pmatrix}, \; \boldsymbol{Y}_{s,r,\eta}^{(q)} = \frac{\hbar}{2}\begin{pmatrix} \sin^2\theta e^{-2r} & 0 \\ 0 & \tan^2\theta\frac{1-\eta}{\eta} \end{pmatrix}, \tag{3.47}$$

with $\cos\theta = e^{-s}$ and r is the squeezing parameter for the ancilla. The superscript q denotes squeezing in the q quadrature. This result almost matches the result from [101]; however, we avoid any extra displacement on the target state by using knowledge of the homodyne detection efficiency in the feedforward displacement.

Importantly, while measurement-based squeezing produces noisy squeezed states on average, in the single-shot case (not averaging over p_M), the covariances, means and weights of the Gaussian functions of the target state transform in a non-linear fashion that precludes description with a deterministic CPTP map. We show this fact in Appendix B.5. For initial means and covariances $(\boldsymbol{\mu}_{m,0}, \boldsymbol{\Sigma}_{m,0})$, we can set Eq. (3.34) to be:

$$\begin{aligned} \boldsymbol{\Sigma}_m &= \boldsymbol{X}_{\eta,2}\boldsymbol{S}_\theta\boldsymbol{S}_{r,2}\begin{pmatrix} \boldsymbol{\Sigma}_{m,0} & \boldsymbol{0} \\ \boldsymbol{0} & \hbar\mathbb{1}/2 \end{pmatrix}\boldsymbol{S}_{r,2}^T\boldsymbol{S}_\theta^T\boldsymbol{X}_{\eta,2}^T + \boldsymbol{Y}_{\eta,2}, \\ \boldsymbol{\mu}_m &= \boldsymbol{X}_{\eta,2}\boldsymbol{S}_\theta\boldsymbol{S}_{r,2}\begin{pmatrix} \boldsymbol{\mu}_{m,0} \\ \boldsymbol{0} \end{pmatrix}, \\ \boldsymbol{S}_{r,2} &= \mathbb{1}\oplus\boldsymbol{S}_r, \boldsymbol{S}_r = \begin{pmatrix} e^{-r} & 0 \\ 0 & e^r \end{pmatrix} \\ \boldsymbol{X}_{\eta,2} &= \mathbb{1}\oplus\boldsymbol{X}_\eta, \; \boldsymbol{Y}_{\eta,2} = \boldsymbol{0}\oplus\boldsymbol{Y}_\eta, \end{aligned} \tag{3.48}$$

where $\boldsymbol{S}_{r,2}$ models squeezing of the ancillary vacuum, $\boldsymbol{S}_\theta$ is the beam-splitter symplectic, and $(\boldsymbol{X}_{\eta,2}, \boldsymbol{Y}_{\eta,2})$ model inefficiency in the homodyne detector for the second mode. Next, we use the update rule in Eq. (3.39), as well as the per-peak reweighting in Eq. (3.37) with $d_j = 1$, $\boldsymbol{r}_M^T = (0, p_M)$ and $\boldsymbol{\Sigma}_M = \lim_{\epsilon\to 0}\frac{\hbar}{2}\begin{pmatrix} \epsilon^{-1} & 0 \\ 0 & \epsilon \end{pmatrix}$, as we are modelling p-homodyne measurements. Finally, to that result we apply a displacement in p-quadrature to the means by $\sqrt{\eta^{-1}}\tan\theta p_M$.

Inline squeezing is an especially valuable operation to be able to model, as it is required to perform the most general Gaussian unitary operations [80]. Moreover, for GKP states specifically, inline squeezing is required to perform certain Clifford qubit gates such as the phase gate and the controlled-NOT and

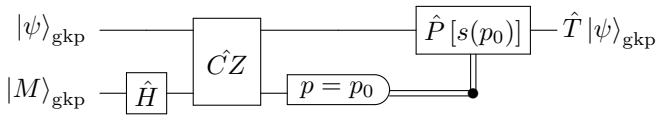


Figure 3.3: Optical implementation of the GKP qubit T gate up to global phase, following the method from [13]. Here, in the ideal limit, $|M\rangle_{\text{gkp}} = \frac{1}{\sqrt{2}}(e^{-i\pi/8}|0\rangle_{\text{gkp}} + e^{i\pi/8}|1\rangle_{\text{gkp}})$. The GKP Hadamard gate (with operator $\hat{H}$) is a $\pi/2$ phase shift, and the GKP CZ gate is discussed in Appendix B.5. The output p_0 of the p-homodyne measurement is processed by the function $s(p_0)$, which rounds the value to the nearest $n\sqrt{\pi\hbar}$ and returns the parity of n. If n is even (odd), no phase gate (a phase gate $\hat{P}(s) = e^{is\hbar\hat{q}^2/2}$ with $s = 1$) is applied.

controlled-Z gates, as we review in Fig. B.3 of Appendix B.5. Using our formalism, we now can model realistic noise effects from inline squeezing applied to bosonic qubit states in both the average and single-shot case. This is useful, for example, when quantifying the build-up of noise in stitching together cluster states of bosonic qubits using active entangling operations [22, 34], although there exist methods for tailored clusters that do-away with inline squeezing [128].

3.5.3 GKP T gate and Error Correction

Just as inline squeezing operations can be teleported onto a mode via the consumption of a squeezed vacuum ancilla, non-Gaussian gates can be effected via gate teleportation using non-Gaussian ancillary states or measurements [13, 16, 17, 19, 117–121]. As a pertinent example, we consider the implementation of the T gate for GKP states. A single-qubit T gate is a non-Clifford gate that, in conjunction with Clifford gates, completes the set of universal operations. Physically, in the GKP encoding, it is implemented through a non-Gaussian transformation, the only one strictly required in the universal set. To effect a T gate via gate teleportation, an ancillary GKP magic state $|M\rangle_{\text{gkp}} = \frac{1}{\sqrt{2}}(e^{-i\pi/8}|0\rangle_{\text{gkp}} + e^{i\pi/8}|1\rangle_{\text{gkp}})$ is first rotated in phase space by $\pi/2$ and entangled with the target mode at a CZ gate. Next, a p-homodyne measurement is performed on the ancilla; the measurement outcome is rounded to the nearest $n\sqrt{\pi\hbar}$, and the parity of n determines whether a feedforward phase gate is applied on the target mode [13]. The non-Gaussianity of the gate enters via the ancillary GKP resource state, which has significant Wigner negativity [53]. See Fig. 3.3 for a summary, and Fig. B.3 of Appendix B.5 for decompositions of the CZ and phase gate.

We emphasize that the ancillary GKP magic state and the target GKP state can both be described as Wigner functions of the form (3.23). The CZ and phase gates—implemented directly or via ancilla-assisted squeezing—can be described through an average map or as a single-shot transformation, as we explored generally in Section 3.3.3 and specifically to these transformations in Appendix B.5. Finally, a p-homodyne measurement of the ancillary mode containing the magic state falls neatly into the formalism from Sections 3.3.2 and 3.3.3. Unlike measurement-based squeezing, we cannot write the average transformation effected by this circuit as a Gaussian CPTP map on the original state, but we can simulate the single-shot transformation of the gate teleportation by selecting or sampling an outcome for the last homodyne measurement and using Eqs. (3.35) and (3.37).

While we can now implement single-shot non-Gaussian operations via gate teleportation, a simple average map for T gate teleportation is untenable for several reasons. First, the ancillary state is itself described by a linear combination of Gaussian functions in phase space, in contrast to the single Gaussian peak of the ancillary squeezed state used for inline squeezing. In the latter case, when the ancillary mode is added to

the description, the weights $\{c_m\}$ in the Wigner function of Eq. (3.23) do not change, since they are each multiplied by 1, while new rows and columns for the additional mode are added to the covariances and means of the Gaussian functions. When the ancillary magic state is added, however, in addition to new rows and columns being added to the covariances and means, the weights and the set of indices $\mathcal{M}$ of the Wigner function *also* change, since we must consider all the cross-terms of different Gaussian peaks in the target and ancillary state. This means that the final resulting state after the teleportation operation has a different set of weights than what it started with, which cannot be captured by a simple average map on each Gaussian function. Second, when applying the feedforward phase gate, the homodyne data is processed by a highly non-linear function because of the binning and parity check, so there would again be no opportunity for the integration over homodyne outcomes to yield a simple Gaussian CPTP map as in Eq. (3.32).

Notice that the GKP error-correction circuit in Fig. 3.1 is just another example of a gate teleportation circuit that can be treated by our formalism. The initial ancillary modes are GKP states, and hence can be expressed as a linear combination of Gaussian functions; the entangling operations are Gaussian CX gates; and the homodyne measurements and feedforward displacements are also all Gaussian. Thus, single-shot error correction, with appropriate noise sources such as loss and Fock damping incorporated into the circuit, can be studied using our formalism.

3.6 Simulation Methods

Having provided our general formalism in Section 3.3, and useful states and transformations in Sections 3.4 and 3.5, and Appendix B.2, we now describe our simulation methods. In Section 3.6.1, we summarize how to track the weights, means and covariance matrices associated with the linear combination of Gaussian functions in phase space that describes the state of the modes in a photonic circuit. In Section 3.6.2, we detail a technique for sampling outcomes of Gaussian measurements on states in our formalism. Finally in Section 3.6.3, we compare our simulation method to a state-of-the-art method the employs the Fock basis, demonstrating the advantage of our formalism and method as compared to this alternative leading technique.

3.6.1 Tracking Weights, Means and Covariances

Our method for simulating states and transformations from our formalism is relatively straightforward. It can be summarized as follows:

1. Initialize the weights, vectors of means, and covariance matrices for each mode at the start of the circuit. For example, Eq. (B.2) provides these parameters for the cat state.

2. Combine the initial weights, means and covariances for each mode into weights, means and covariances for the multimode state of the whole circuit as follows:

$$\begin{array}{lccc} & \text{Mode 1} & \text{Mode 2} & & \text{Multimode} \\ \text{Weights} & a_\ell & b_k & \to & a_\ell b_k \\ \text{Means} & \boldsymbol{\mu}_\ell & \boldsymbol{\mu}_k & & \boldsymbol{\mu}_\ell \oplus \boldsymbol{\mu}_k \\ \text{Covariances} & \boldsymbol{\Sigma}_\ell & \boldsymbol{\Sigma}_k & & \boldsymbol{\Sigma}_\ell \oplus \boldsymbol{\Sigma}_k \end{array} \tag{3.49}$$

   This procedure can be applied recursively to build the weights, means and covariances for the full multimode initial state.

3. For each gate or measurement in the circuit, apply the associated transformation to the weights, means and covariances of the state. For example, to apply the loss channel, the matrices that parametrize the channel from Eq. (3.45) are used in Eq. (3.32). Alternatively, for a given measurement outcome on a subset of modes, the other modes are updated using Eqs. (3.35) and (3.37).

4. The output of the simulation consists of the weights, means and covariances of all the modes after the application of transformations and measurements in the circuit, along with any of the measurement outcomes collected along the way. Since these weights, means and covariances can be used to construct the Wigner function of the state as in Eq. (3.23), we have access to full state information.

For some states, such as GKP and cat states, all the Gaussian peaks in the linear combination representing that mode have the same covariance matrix. In those cases, to save memory, we need not initialize the same copy many times, and can simply track a single covariance matrix.

In the next part, we detail our method for sampling outcomes of Gaussian measurements for states in our formalism.

3.6.2 Sampling Outcomes of a Gaussian Measurement

To further the utility of our formalism and method for performing simulations, we now provide an algorithm for sampling outcomes of Gaussian measurements on states that can be expressed as linear combinations of Gaussian functions in phase space. In Section 3.3.2, we detailed in Eq. (3.27) how to calculate probability distributions for Gaussian measurements within our formalism. An efficient algorithm exists for sampling non-Gaussian yet positive Wigner functions under Gaussian transformations and Gaussian measurements [100, 129]. The technique consists of first sampling a point in phase space $\boldsymbol{u}$ from the initial Wigner function (which is possible since the positivity of the Wigner function means it can be treated as a probability distribution), then applying a linear transformation associated with the Gaussian operation to the sampled point $\boldsymbol{u}' = S\boldsymbol{u} + \boldsymbol{x}$, and finally sampling from a Gaussian distribution associated with the measurement operator, centred at $\boldsymbol{u}'$. Such a technique could be useful for states in Eq. (3.23) with positive-valued weights, and real-valued means and covariances; however, if the Wigner function has negativity, a different technique is required. A general approach to sampling from non-trivial Wigner functions can be found in Appendix D of [130]; however, we find that a more tailored approach one can use is a rejection-sampling technique [131]. In [132], the authors consider a rejection-sampling method for a distribution composed of real-valued Gaussian functions with (positive or negative) real weights. Here, we extend the method to include Gaussian functions with complex weights and means. We summarize the technique in Algorithm 1. First, we separate the peaks into those which have negative weights and real-valued means $\mathcal{M}^- = \{m|c_m < 0\} \cap \{m|\Im(\boldsymbol{\mu}_m) = 0\}$, and everything else $\{m \notin \mathcal{M}^-\}$:

$$p\left(\boldsymbol{r}_M; \hat{\rho}, \boldsymbol{\Sigma}_M\right) = \sum_{m\in\mathcal{M}^-} c_m G_{\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_m}\left(\boldsymbol{r}_M\right) + \sum_{m'\notin\mathcal{M}^-} c_{m'} G_{\boldsymbol{\mu}_{m'}, \boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m'}}\left(\boldsymbol{r}_M\right). \tag{3.50}$$

By writing $\boldsymbol{\mu}_{m'} = \Re(\boldsymbol{\mu}_{m'}) + i\Im(\boldsymbol{\mu}_{m'})$ we can obtain

$$G_{\boldsymbol{\mu}_{m'}, \boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m'}}\left(\boldsymbol{r}_M\right) = e^{\frac{1}{2}\Im(\boldsymbol{\mu}_{m'})^T(\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m'})^{-1}\Im(\boldsymbol{\mu}_{m'})} G_{\Re(\boldsymbol{\mu}_{m'}), \boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m'}}\left(\boldsymbol{r}_M\right) e^{i[\boldsymbol{r}_M-\Re(\boldsymbol{\mu}_{m'})]^T(\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m'})^{-1}\Im(\boldsymbol{\mu}_{m'})}. \tag{3.51}$$

This allows us to define an upper-bounding function to the distribution of interest:[1]

$$p\left(\boldsymbol{r}_M;\hat{\rho},\boldsymbol{\Sigma}_M\right)\leq g(\boldsymbol{r}_M)=\sum_{m\notin\mathcal{M}^-}\tilde{c}_m G_{\Re(\boldsymbol{\mu}_m),\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_m}\left(\boldsymbol{r}_M\right),$$
$$\tilde{c}_m=|c_m|e^{\frac{1}{2}\Im(\boldsymbol{\mu}_m)^T(\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_m)^{-1}\Im(\boldsymbol{\mu}_m)}. \tag{3.52}$$

Next, we note that $g(\boldsymbol{r}_M)$ is a scalar multiple of a probability distribution $q\left(\boldsymbol{r}_M\right)$:

$$g(\boldsymbol{r}_M)=\mathcal{N}q\left(\boldsymbol{r}_M\right),\ \mathcal{N}=\sum_{m\notin\mathcal{M}^-}\tilde{c}_m. \tag{3.53}$$

Importantly, one can sample from $q\left(\boldsymbol{r}_M\right)$ more straightforwardly: first, one samples a value $m_0\notin\mathcal{M}^-$ according to the distribution $p_m=\tilde{c}_m/\mathcal{N}$; since the weights $\tilde{c}_m$ from $g(\boldsymbol{r}_M)$ are now all positive, we can associate them with probabilities after proper normalization by $\mathcal{N}$. Next, one samples a phase space value $\boldsymbol{r}_0$ according to $G_{\Re(\boldsymbol{\mu}_{m_0}),\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m_0}}\left(\boldsymbol{r}\right)$. Given $\boldsymbol{r}_0$, one then samples y_0 uniformly form $[0,g(\boldsymbol{r}_0)]$. If $y_0\leq p\left(\boldsymbol{r}_0;\hat{\rho},\boldsymbol{\Sigma}_M\right)$, then $\boldsymbol{r}_0$ is kept as the sampled value of $p\left(\boldsymbol{r}_M;\hat{\rho},\boldsymbol{\Sigma}_M\right)$; otherwise, it is rejected and the process is restarted until a sample is drawn.

This sampling technique leverages the form of the Wigner function of our states: we use the fact that our distribution can be bounded above by a mixture of Gaussian functions with all-positive weights, and that it is computationally easy to sample a value from a single Gaussian distribution.

Algorithm 1 Simulating Gaussian measurement outcomes using rejection sampling

Input: Weights c_m, means $\boldsymbol{\mu}_m$, and covariances $\boldsymbol{\Sigma}_m$ of the initial state. Covariance $\boldsymbol{\Sigma}_M$ of the measurement.
Output: Sampled phase-space position outcome $\boldsymbol{r}_0$

$\mathcal{M}^-=\{m|c_m<0\}\cap\{m|\Im(\boldsymbol{\mu}_m)=0\}$
$\tilde{c}_m=|c_m|e^{\frac{1}{2}\Im(\boldsymbol{\mu}_m)^T(\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_m)^{-1}\Im(\boldsymbol{\mu}_m)}$
drawn $\leftarrow$ False
while not drawn **do**
  Sample $m_0\notin\mathcal{M}^-$ according to $p_m=\tilde{c}_m/\mathcal{N}$
  Sample $\boldsymbol{r}_0$ according to $G_{\Re(\boldsymbol{\mu}_{m_0}),\boldsymbol{\Sigma}_M+\boldsymbol{\Sigma}_{m_0}}\left(\boldsymbol{r}\right)$
  Sample y_0 uniformly from $[0,g(\boldsymbol{r}_0)]$
  if $y_0\leq p\left(\boldsymbol{r}_0;\hat{\rho},\boldsymbol{\Sigma}_M\right)$ **then**
    drawn $\leftarrow$ true
  end if
end while
return $\boldsymbol{r}_0$

With the simulation methods in hand, we now compare our simulation method to simulation methods in the Fock basis.

3.6.3 Comparison to Fock Basis Simulations

It is well-known that representing a Gaussian state in terms of its vector of means and its covariance matrix is a more efficient representation than writing the state in the Fock basis, especially considering the energy of the state increases under displacements and squeezing. Here, we draw a similar conclusion for the states we

[1] Note that we do not include any weights from $\mathcal{M}^-$ in this bound since bounding a real-valued Gaussian with a negative weight by a Gaussian with the absolute value of the negative weight is a looser upper bound than simply bounding it by 0.

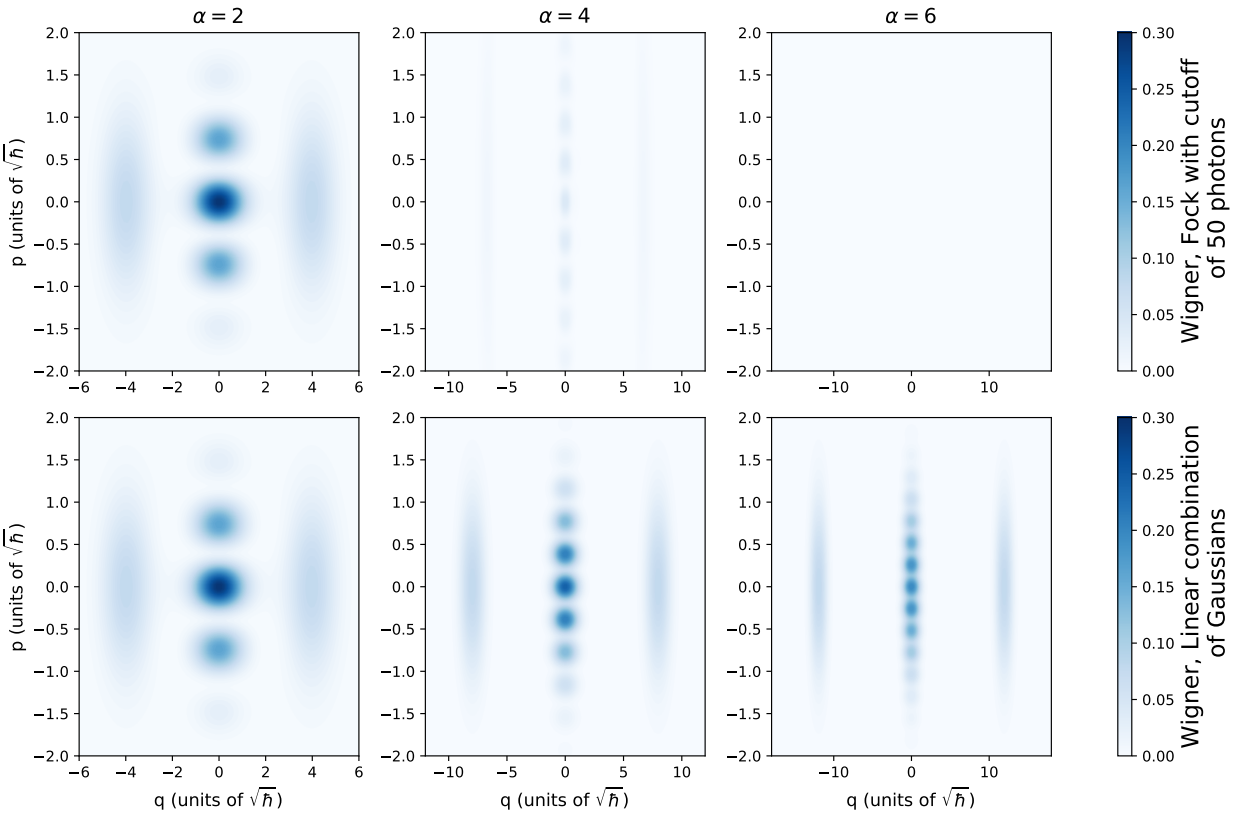


Figure 3.4: Simulation of two cat states passing through a lossy beam-splitter, with the Wigner function of the first mode displayed for increasing values of α. Note the difference in scale between q and p. We compare the output of the simulation using the `fock` backend of `Strawberry Fields` with a cutoff of 50 photons (top row), and using a linear combination of Gaussians in the `bosonic` backend (bottom row). While the number of photons in the state increases with α, saturating the cutoff for the Fock basis simulation and leading to an incorrect output, the number of Gaussian peaks in phase space that require tracking remains constant. Note additionally that the mixing introduced by loss necessitates the full density matrix being tracked in the Fock basis, while loss does not increase the complexity of simulation for the linear combination of Gaussians.

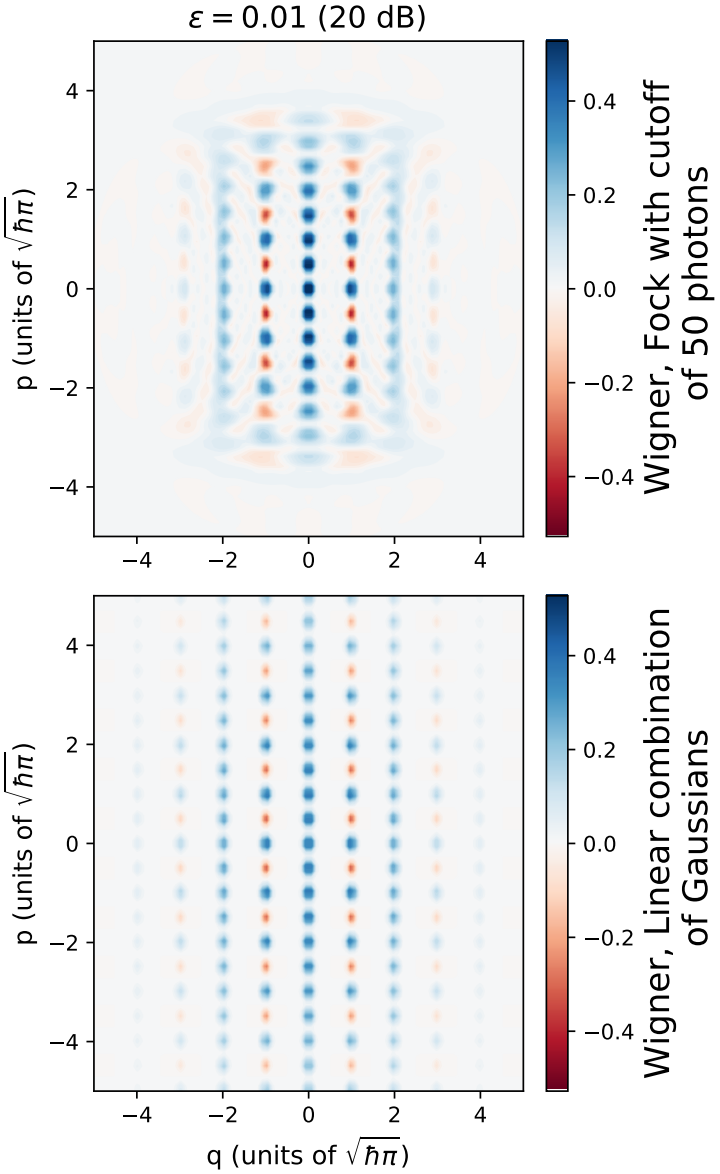


Figure 3.5: Wigner functions for a $|+^{\epsilon}\rangle_{\text{gkp}}$ state with $\epsilon = 0.01$ (20 dB of per-peak squeezing) teleported onto a 15 dB squeezed state. We entangle the states via a CV CZ gate and and subject them 1% loss in each mode; next, we measure the mode originally containing the GKP state in the p-basis, postselecting on $p = 0$. In the top panel, we perform the simulation using the `fock` backend in `Strawberry Fields` with a cutoff of 50 photons. In the bottom row, we employ the `bosonic` backend. We see that the `bosonic` backend can maintain a correct description of the state even as the number of photons becomes very high due to the large per-peak squeezing.

have studied under our formalism. In this and upcoming sections, we implement our simulation technique in the new `bosonic` backend of `Strawberry Fields`.

Consider that the Hamiltonian for a single mode is given by $\hat{H} = \frac{1}{2}(\hat{q}^2 + \hat{p}^2) = \frac{1}{2}\hat{r}^2 = \hbar(\hat{n} + \frac{1}{2})$. This means that the Wigner function of a Fock state $|n\rangle$ has an associated radius in phase space of roughly $|\boldsymbol{r}_n| = \sqrt{\hbar(2n+1)}$, beyond which the function decays monotonically. This fact can also be demonstrated rigorously using an analytic representation of Fock Wigner functions and the properties of Laguerre polynomials [133, 134]. Thus, to determine the Fock representation for a state which has a phase-space Gaussian peak at a point $\boldsymbol{r}_0$, a conservative estimate has that we would need a photon number of at least

$$n(\boldsymbol{r}_0) \sim \frac{|\boldsymbol{r}_0|^2}{2\hbar} \tag{3.54}$$

to reach the required radius in phase space. Furthermore, Fock states beyond this value may be necessary, for example, to shape the phase space peak for a desired level of squeezing. In the presence of realistic noise sources like loss, the state additionally becomes mixed, requiring a density matrix rather than a state vector representation in the Fock basis, squaring the number of elements which one needs to track. From these considerations, we see that for a state with a phase space peak at $\boldsymbol{r}_0$, the number of elements one needs to track in the Fock basis to have a faithful representation of the state scales like $\frac{|\boldsymbol{r}_0|^4}{\hbar^2}$.

Let us compare this Fock-basis scaling to what we would obtain by expressing our states of interest as linear combinations of Gaussian functions in phase space. The trivial case is a Gaussian state; regardless of its position, orientation, and level of squeezing in phase space, one need only track a 2-component vector of means and a 2×2 covariance matrix. For an N mode Gaussian state, the number of elements to track scales like $4N^2 + 2N$—no exponential scaling since all the modes are represented by a single Gaussian function. Adding a mode with a Gaussian state to a series of other modes with states represented by a linear combination of Gaussian functions only increases the dimension of the covariance matrices and means by 2, but does not change the number of weights, means or covariances one needs to track. This is especially useful for the GKP encoding which, when combined with Gaussian states, enables universal quantum computation [36, 49].

We have shown in Appendix B.2.1 that single-mode two-lobe cat states can be represented using one 2×2 covariance matrix, along with four weights and 2-component vectors of means—two real-valued and two complex-valued. Crucially, the size of the mathematical objects required for this representation is independent of the energy of the cat state associated with α, and is invariant under Gaussian transformations. While the covariance for N modes encoded as cat qubits scales like $4N^2$, the number of weights still scales exponentially, like 4^N, and the number of elements in the means grows like $2N(4^N)$. This is still preferable to the Fock representation, for which the number of density matrix elements scales like $|\alpha|^{4N}$; notably, the scaling for the linear combination of Gaussians representation does not depend on the energy of the state or modifications under Gaussian transformations. This can be valuable since, for example, the error rates for some qubit gates on cat states can scale like $1/|\alpha|$ [90], so to examine regimes of low error, one must increase the energy of the cat state. In Fig. 3.4, we examine the case of two cat states sent through a lossy beam-splitter. We trace out one of the modes and plot the Wigner function of the remaining mode for increasing values of energy, as parametrized by α. We compare the results using the `fock` backend of `Strawberry Fields` using a photon number cutoff of 50 photons per mode— going beyond this value saturates the memory of the standard desktop terminal used for simulation—and using the `bosonic` backend, representing states as linear combinations of Gaussian functions in phase space. We find that, for $\alpha = 2$, the Fock representation still works well, albeit requiring more memory and running more slowly when simulating the lossy beam-splitter.

For $\alpha = 4$ and 6, the Fock representation with 50 photons quickly becomes insufficient. This is unsurprising: for these values of α, the action of the beam-splitter leads to Gaussian peaks in phase space at distances, respectively, of $8\sqrt{\hbar}$ and $12\sqrt{\hbar}$ from the origin, requiring, by the conservative estimate in Eq. (3.54), greater than 32 and 72 photons for each case. In fact, although $n = 32$ is comfortably within the cutoff of 50 photons, the Fock representation cannot accurately construct the Wigner function for $\alpha = 4$, confirming that Eq. (3.54) is an underestimation of the required photon number cutoff.

As seen in Appendix B.2.2, even Fock states of n photons can be approximated by n real-valued weights, n 2×2 covariance matrices, and one 2-component vector of means. This is perhaps more memory-intensive than representing a pure number state in the Fock basis; however, under Gaussian transformations such as displacements or squeezing, or the common loss channel, representing the state by a linear combination of Gaussians becomes advantageous.

Finally, while GKP states have an infinite number of peaks, we can consider only a finite number since the weights in Eq. (3.43) decay exponentially under the Fock damping operator. Since the Fock damping operator decays exponentially in Fock space, an analogous procedure can be executed to establish a reasonable photon number cutoff. If a single-mode GKP state is approximated by a finite number of peaks, with the furthest peak located near $\frac{\sqrt{\pi}}{2}(k, \ell)$, then the number of peaks one needs to track scales like $k^2 + \ell^2$, and consequently so does the number of elements for the vectors of means; by contrast, only a single 2×2 covariance matrix is required. Compare this with the Fock representation, where, by Eq. (3.54), the number of density matrix elements scales like $(k^2 + \ell^2)^2$. For N modes encoded as GKP states, we track $(k^2 + \ell^2)^N$ weights and means, and a single $2N \times 2N$ covariance matrix, an improvement over the $(k^2 + \ell^2)^{2N}$ scaling for the number of density matrix elements in the Fock basis. As a demonstration of the advantage of the `bosonic` backend, in Fig. 3.5 we plot the Wigner function for the output mode of a simulated CV teleportation of a high-energy GKP state ($\epsilon = 0.01$, corresponding to 20 dB of per-peak squeezing) onto a p-squeezed state with 15 dB of squeezing using a lossy CZ gate, homodyne measurement, and feedforward displacement. We see that the `fock` backend is limited in the radius of phase space that it can accurately capture, while the `bosonic` backend produces all peaks correctly. Moreover, the all-Gaussian nature of the teleportation circuit allows for more efficient computation when representing the states as linear combinations of Gaussians over the Fock representation, which requires many tensor contractions for large density matrices.

Having compared our scaling to the memory requirements for the Fock basis, we now provide a short comment on the potential connections between our framework and techniques used for simulating qubit circuits using linear combinations of stabilizer states. Analogously to how Gaussian states are efficient to simulate in phase space under Gaussian transformations [135] or to how ideal, inifinite-energy GKP states can be efficiently simulated undergoing certain Gaussian transformations [34, 52], two-dimensional qubits stabilized by Pauli operators (i.e. eigenstates of the operators with +1 eigenvalues, also termed "stabilizer states") transforming under Clifford gates and being measured in the computational basis can be efficiently simulated [136, 137]. This is because the Clifford gates map Pauli operators to Pauli operators, so one can efficiently track how the stabilizer operators update, rather than the full state. Such circuits are not sufficient for universal quantum computation, however; for that, one must supplement the circuits with a non-Clifford element, typically chosen to be the T-gate, and implemented in the simulation using magic states and Clifford elements (see Fig. 3.3 for a decomposition of the gate). The difficulty of simulating such circuits can therefore be tied directly to the number of magic states used.

A clever technique for improving the simulation cost is to decompose the initial tensor product of magic states into either an exact or an approximate linear combination of stabilizer states and then tracking how each

of the operators that defines the linear combination evolves under Clifford gates [138–140]. The simulation does not become efficient, however, since the number of operators that defines the linear combination grows exponentially in the number of magic states (past a certain point, at least). Our formalism is reminiscent of this technique, in that we break down our states into a linear combination of functions (albeit not necessarily states) that can each be efficiently simulated in phase space under Gaussian transformations. Analogously to how magic states can be combined with Clifford gates to apply non-Clifford operations, linear combinations of Gaussian functions in phase space can be combined with Gaussian operations to effect non-Gaussian gates; however, the memory costs of the simulation currently scale exponentially with the number of modes that are initialized as linear combinations of Gaussian functions, analogous to the exponential scaling in the qubit case with the number of magic states. An interesting observation from [138] was that for $m \leq 6$ magic states, one needs to consider only 7 instead of 2^m stabilizer states in the linear combination; this motivates investigating whether decompositions of multimode states in our formalism directly into multimode Gaussians can yield a fewer number of Gaussian functions than single mode decompositions followed by the combination technique given in (3.49). While a deeper comparison between these two techniques is outside the scope of this chapter, we point to some additional avenues for future research in Section 3.8.

3.7 Numerical Simulations

Having established our simulation method and its advantage over competing techniques in Section 3.6, we now put it to use in simulations of bosonic qubits, leveraging a numerical implementation of our method available through the `bosonic` backend of `Strawberry Fields` [102]. In Section 3.7.1, we provide a basic test-run of our simulator, plotting Wigner functions of and sampling from bosonic qubits. In Section 3.7.2, we provide novel simulations of bosonic qubits undergoing realistic gate implementations, with a focus on GKP states. For an advanced tutorial implemented by the authors on using the `bosonic` backend of `Strawberry Fields` to simulate realistic GKP qubit gates, see [105].

3.7.1 Basic Examples: Wigner Plots and Homodyne Sampling

As a first simple demonstration of simulations with our technique, in Fig. 3.6 we plot the Wigner functions for $|0^\epsilon\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB of squeezing per peak of the wavefunction), for $|0^\alpha\rangle_{\text{cat}}$ with $\alpha = 2$, and for the single-photon Fock state. For the GKP state, we can identify the following features: the Gaussian functions centred near integer and half-integer multiples of $\sqrt{\pi}$; the positive and negative peaks determined by the weighting function; and the per-peak variance, which is smaller than that of the vacuum. For the cat state, we see a similar distribution to Fig. 3.2, where the interference fringes are produced by the Gaussians with complex weights and means. Finally, for the Fock state, we see how two Gaussians, both with zero means but with slightly different covariance matrices, can combine to form a rotationally symmetric Wigner function with a region of negativity in the centre.

Algorithm 1 offers a straightforward method for sampling the results of general-dyne measurements of states with Wigner functions that are expressed as linear combinations of Gaussian functions. We use the algorithm to simulate 2000 samples of q and p quadrature homodyne measurements of a $|0^\alpha\rangle_{\text{cat}}$ with $\alpha = 2$, and of a $|0^\epsilon\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB per peak), and plot the output in Fig. 3.7. We see that the results align with the asymptotic marginal distributions for these quadratures. The marginals can also be easily attained in our formalism: integrating out one quadrature for a linear combination of Gaussian functions amounts to

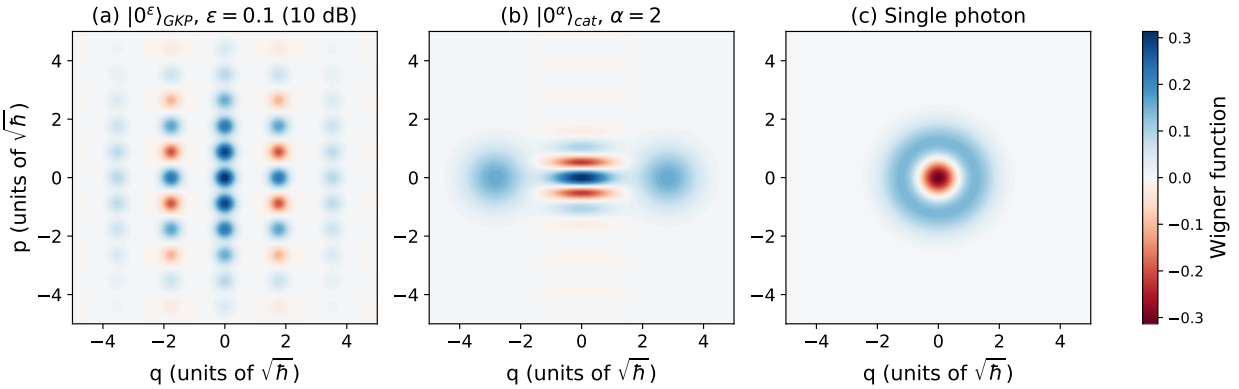


Figure 3.6: Wigner functions for GKP, cat, and Fock states. All figures are produced using linear combinations of Gaussians in phase space, as presented in Section 3.4 and Appendix B.2. For the GKP state, negative regions correspond directly to Gaussian peaks with negative weights, while for the cat state negativity is produced by the sinusoidal oscillations of complex-valued Gaussian peaks. In the single photon case, negativity is produced by the difference of two zero-mean Gaussians with slightly different covariance matrices.

dropping the entries associated with that quadrature from each mean and covariance matrix, as we saw in Eq. (3.29).

For the cat state, we see the q quadrature distribution is the same as for mixture of coherent states centred at $\pm\sqrt{2\hbar}\alpha$, since the interference fringes cancel out, while they are prominent in the p quadrature. For the GKP state, we can clearly identify the peaks at even (all) integer multiples of $\sqrt{\pi\hbar}$ in q (p) quadrature.

3.7.2 Novel Simulations of Useful CV Circuits

Measurement-Based Squeezing

In Section 3.5.2 and in Appendix B.5, we examine how to use our formalism to describe measurement-based squeezing, a feasible method for applying in-line squeezing operations. Later subsections examine gates that employ measurement-based squeezing in more complicated circuits with bosonic qubits, but we restrict ourselves to a simpler case to develop a clear a picture of the action of the transformation. In Fig. 3.8 we plot the the Wigner function for measurement-based squeezing applied to vacuum. We assume that the ancillary state has a fixed level of $r = 1.2$ (~ 10.5 dB) of squeezing, and that the efficiency of the homodyne detection is 0.99. We consider three different target levels of squeezing, $r_{\text{target}} = 0.3, 1,$ and 2, to apply to the vacuum state, and calculate the Wigner functions in the case of ideal, direct inline squeezing; the average map of measurement-based squeezing; and a single-shot occurrence of measurement-based squeezing.

There are a few key observations we can make from the simulation. First, if r_{target} is comparable to or greater than the level of resource squeezing r, then the variance in the q quadrature exceeds the desired value. This is to be expected: Eq. (3.47) tell us that, while the level of noise in the q quadrature is modulated by e^{-2r} (constant across the simulations), it grows with r_{target}. Next, we see that, in the average case, the level of anti-squeezing essentially matches the ideal case. This is due to the high efficiency of the homodyne measurement, which forces the level of p quadrature noise—proportional to $\frac{1-\eta}{\eta}$, per Eq. (3.47)—to be small. Additionally, the mean in the average case matches the ideal case, as guaranteed by Eq. (3.47). Finally, in the single-shot case, the location at which the state is centred along the p quadrature can vary on the order

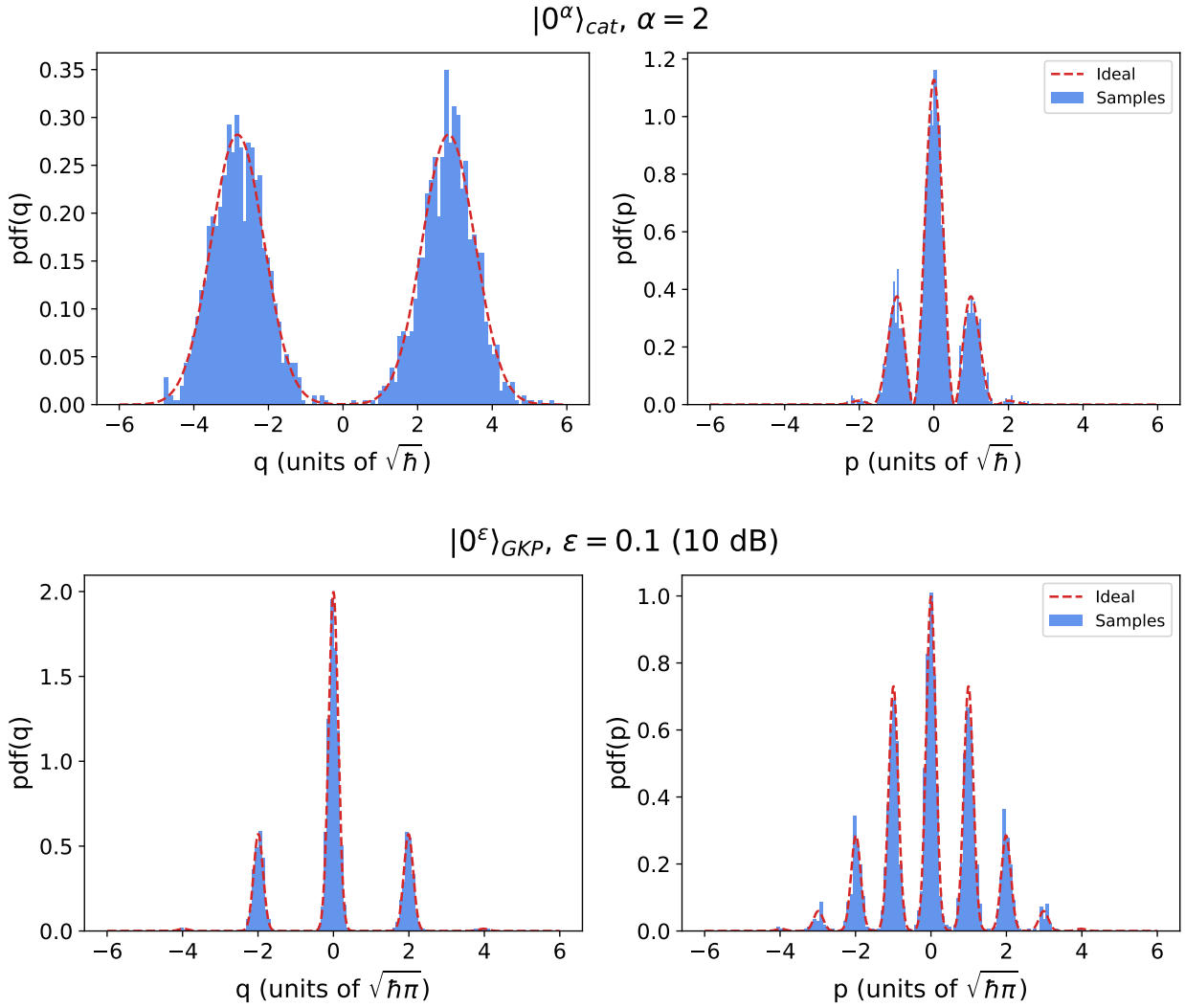


Figure 3.7: 2000 homodyne measurement samples in both quadratures for a cat (top row) and a GKP (bottom row) state obtained using Algorithm 1. We compare histograms obtained from finite sampling to the asymptotic distributions, obtained easily from the linear combination of the marginals of each Gaussian function in the Wigner function, as in Eq. (3.29). For an intermediate tutorial implemented by the authors on how these results can be generated with the `bosonic` backend of `Strawberry Fields`, see [104].

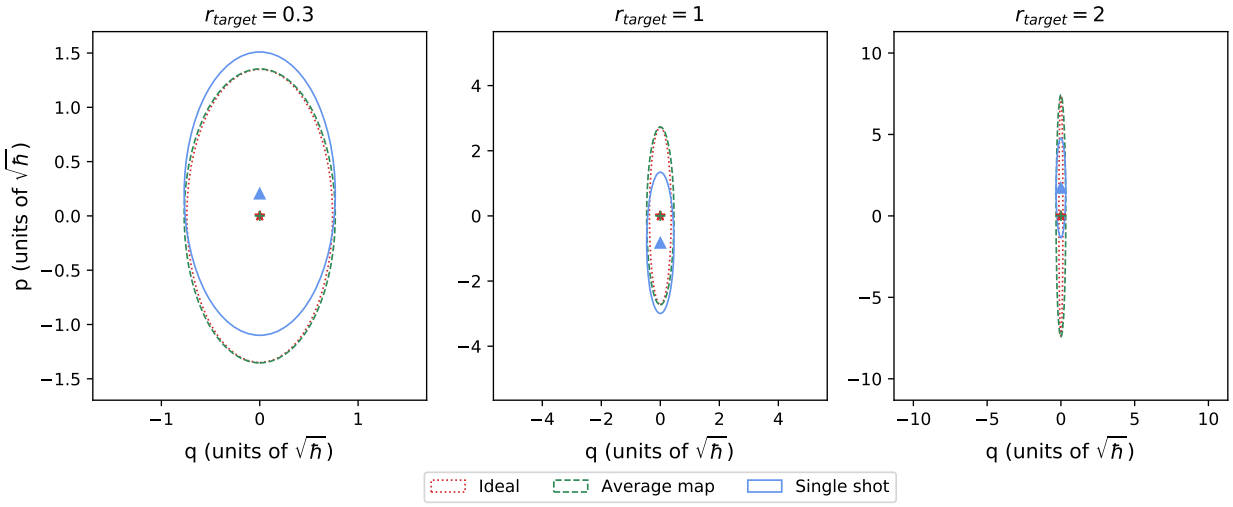


Figure 3.8: For different levels of target squeezing on vacuum r_{target}, a comparison of the states produced from ideal squeezing (red), the average map from measurement-based squeezing (green), and a single-shot occurrence of measurement-based squeezing (blue). The centre and axis lengths are set by the mean and variances of the output state. We simulate using an ancillary squeezed state of $r = 1.2$ ($\sim$10.5 dB), and a homodyne detection efficiency of 0.99. A key takeaway is that the single-shot instances have less antisqueezing and a distribution of locations for where they can be centred in phase space, indicating the level of antisqueezing and the correct mean value observed in the average map are a result of the averaging procedure itself.

of the ideal anti-squeezing; moreover, the level of anti-squeezing in the single-shot case becomes significantly lower than ideal as r_{target} becomes comparable to or greater than the level of resource squeezing r. This indicates that the accurate level of anti-squeezing and the correct mean observed on average is an artifact of the averaging process itself: the states that factor into the average have less anti-squeezing but a broad distribution of where they are centred along the p quadrature, resulting in a broadened average output state. The differences between ideal, average, and single-shot instances of measurement-based squeezing are crucial to understand for realistic implementations of bosonic codes; our formalism and simulations are valuable for such analysis.

GKP Phase Gate

The phase gate is a pertinent example of a single-mode GKP gate that employs inline squeezing, an operation that may be performed, in practice, through measurement-based methods. In Appendix B.5.2, we provide details for such an optical implementation. Here, we present results of a simulation of a realistic phase gate applied to $|+^{\epsilon}\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB). In Fig. 3.9, we plot the Wigner functions of the average output states, varying the level of squeezing of the ancillary squeezed state and the efficiency of the homodyne detection. We compare the graphs to those for $|+i^{\epsilon}\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB), which is the finite energy version of the ideal application $\hat{P}\,|+\rangle_{\text{gkp}}$. We find several differences. First, while the finite-energy $|+i^{\epsilon}\rangle_{\text{gkp}}$ has a symmetric envelope in phase space, and each Gaussian peak has an isotropic spread about its mean, the application of a phase gate to a finite-energy $|+^{\epsilon}\rangle_{\text{gkp}}$ results in an asymmetric envelope and anisotropic Gaussian peaks about each mean in phase space. This would be the case even if the phase gate was applied perfectly to a finite energy state: the finite-energy $|+^{\epsilon}\rangle_{\text{gkp}}$ is non-periodic due to the envelope, and each peak has finite width;

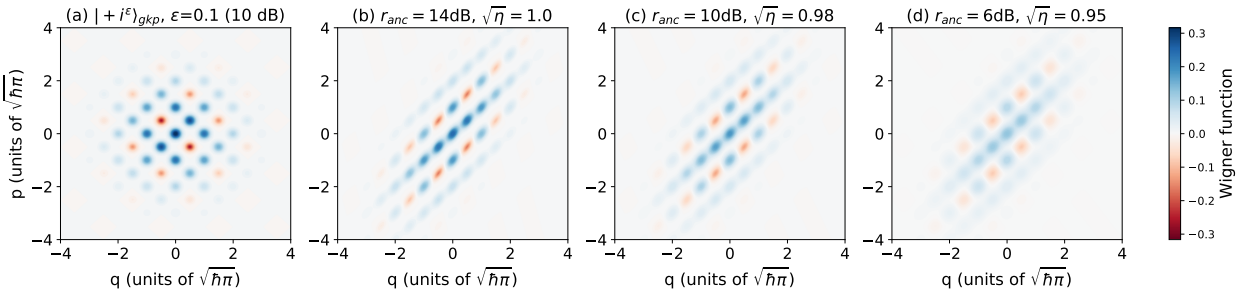


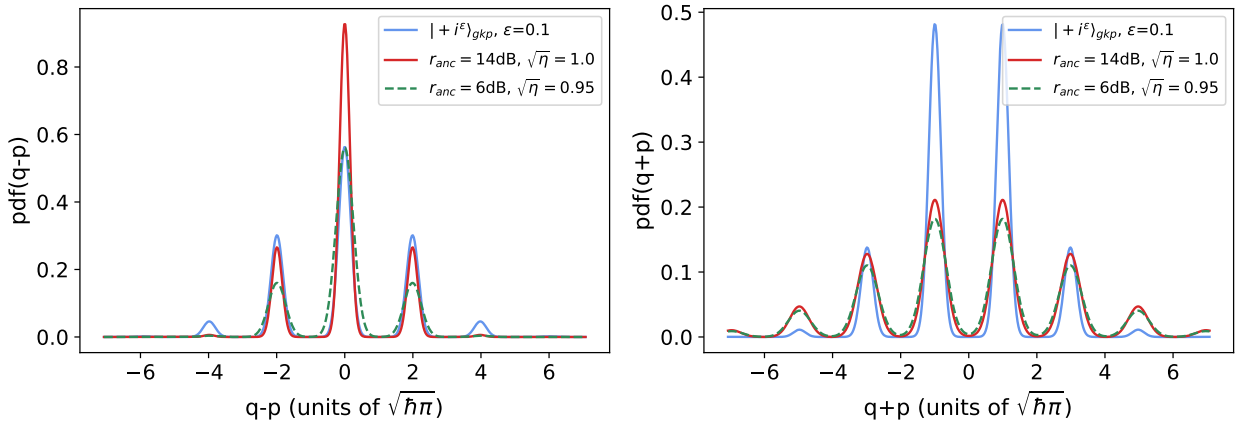
Figure 3.9: (a) Wigner function for $|+i^{\epsilon}\rangle_{\text{gkp}}$ simulated directly with $\epsilon = 0.1$ (10 dB). We compare this to a method for ideally preparing the same state by generating $|+^{\epsilon}\rangle_{\text{gkp}}$ with $\epsilon = 0.1$, then applying a GKP phase gate $\hat{P} = e^{i\hbar\hat{q}^2/2}$. Since the phase gate requires inline squeezing (see Appendix B.5), we simulate measurement-based squeezing with various levels of ancillary squeezing and ancilla detector efficiency, which we label in (b)-(d). We see that phase gates apply a shearing effect to the envelope and peaks of the finite energy GKP state, with lower quality measurement-based squeezing adding additional broadening to the peaks. This simulation was performed with the `bosonic` backend of `Strawberry Fields`.

even a perfect phase gate would cause shearing effects in both the envelope and the individual peaks, an effect not seen in ideal states with perfect periodicity and infinitely narrow peaks. Second, as the quality of the squeezed ancilla resource and homodyne efficiency worsen, we find that the peaks in phase space broaden, albeit asymmetrically, decreasing the height of each peak.

The differences in the Wigner function do not tell the full story, however; we are also interested in how they affect readout of the state. In Fig. 3.10, we plot the marginal distribution along $q-p$ and $q+p$ of three of the states in Fig. 3.9. Recall that binning the outcome along $q-p$ to the nearest $n\sqrt{\pi}$ and taking the parity of n corresponds, ideally, to the a measurement in the qubit Pauli Y basis; alternatively, one can bin the outcome along $q+p$ to the nearest $n\sqrt{\pi}$ and take the parity of $n+1$. This choice comes from the fact that $\hat{H}\hat{\sigma}_y\hat{H}^\dagger = -\hat{\sigma}_y$ for qubits, where $\hat{\sigma}_y$ is the Pauli Y operator and $\hat{H}$ is a qubit Hadamard gate operator, effected for GKP states by a CV phase space rotation by $\pi/2$. In practice, these two CV operators can be measured by performing homodyne detection along $\frac{q\pm p}{\sqrt{2}}$ and then rescaling the outcome by $\sqrt{2}$. We find that the choice of readout quadrature affects the probability of obtaining the correct qubit result. Because the Wigner function is sheared in phase space, measuring along $q-p$ results in a narrower envelope and narrower peaks. For the purpose of qubit readout, the peak width is what matters since the likelihood of falling outside a 0-bin does not depend on the number of peaks. The results are presented in Table 3.1. There is a sizable drop in readout quality going from $q-p$ to $q+p$, even though both represent the same ideal qubit measurement. This is further confirmation that the one-to-many mapping of qubit-to-CV operators affords GKP states an advantageous flexibility [16]. Using our simulator, we identify how to adapt the readout strategy based on the noise to obtain a more faithful measurement of the binary qubit outcomes.

GKP to CV Cluster Teleportation

CV cluster states—multimode Gaussian states constructed by stitching together momentum-squeezed states with CV CZ gates—are a valuable resource for measurement-based quantum computing with GKP states. In particular, a GKP state can be added to and teleported along the cluster using the same operations as would be used for a momentum-squeezed state [34]. The teleportation circuit begins with a GKP state and a momentum-squeezed state interacting via a CZ gate; next, a p-homodyne measurement is applied to the



(a) (b)

Figure 3.10: The marginal distribution for a homodyne measurement along (a) $\frac{q-p}{\sqrt{2}}$ and (b) $\frac{q+p}{\sqrt{2}}$, each followed by a rescaling of the results by $\sqrt{2}$. Binning the result to the nearest $n\sqrt{\pi}$ and taking the parity of (a) n or (b) $n+1$ effects a GKP qubit Y measurement. We present the marginals for Wigner functions (a), (b), and (d) from Fig. 3.9. While marginal distributions along either $q-p$ or $q+p$ could be sampled for a Pauli Y measurement, the narrower peaks along $q-p$ yield higher likelihood of falling within the correct bin, emphasizing the value of tracking shearing effects to optimize readout fidelity.

	p(reading out a qubit 0)	
Simulation parameters	$q-p$	$q+p$
$\|+i^\epsilon\rangle_{\text{gkp}}$, $\epsilon = 0.1$	99.5%	99.5%
$r_{\text{anc}} = 14$ dB, $\sqrt{\eta} = 1$	99.9%	92.2%
$r_{\text{anc}} = 6$ dB, $\sqrt{\eta} = 0.95$	95.2%	87.2%

Table 3.1: For the marginal distributions in Fig. 3.10, we calculate the probability of correctly identifying a qubit 0 readout from a Pauli Y measurement, depending on whether the CV homodyne measurement is performed along $q-p$ or $q+p$. $q-p$ provides better readout, due to narrowed peaks in the marginal distribution.

GKP mode, yielding outcome p_0; finally, a shift in q by $-p_0$ is applied to the remaining mode. Here, we seek to understand the dynamics of a realistic optical implementation of GKP teleportation into a CV cluster. As discussed in Appendix B.5.2, we break apart the CZ gate into beam-splitters and inline squeezing operations effected using measurement-based squeezing, and then investigate the effects of finite squeezing and loss in the teleportation circuit. The circuit requires three squeezed states: the ancilla mode onto which the GKP is teleported and the resource states for measurement-based squeezing within the decomposition of the CZ gate. We vary the level of squeezing of these Gaussian modes, while fixing the GKP state to be $|0^{\epsilon}\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB). Additionally, we vary loss within the circuit; we apply a loss channel of transmissivity η to each mode after initialization and at the output of each of the four beam-splitters in the circuit (two for the decomposition of the CZ gate and one for each inline squeezing operation). For the three homodyne detectors in the circuit, we fix an efficiency of $\eta_{\text{det}} = 99\%$.

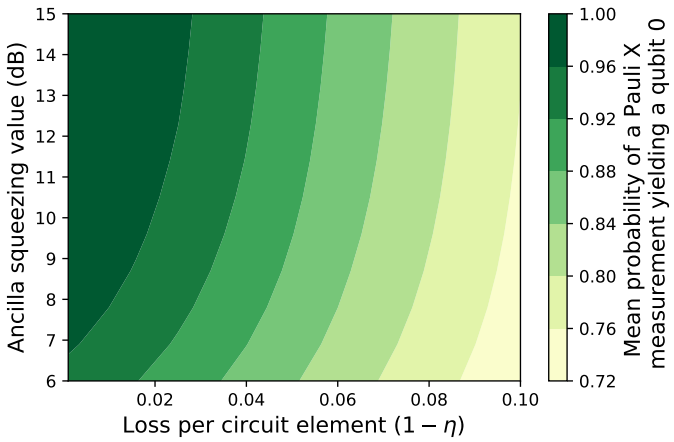
For each value of squeezing and η, we run 500 simulations, the teleported state output by each run depending on the probabilistic feedforward value. Then, given the final state in each run, we measure the qubit Pauli X and Z operators, achieved by performing homodyne measurements along p and q, respectively, binning the outcome to the nearest $n\sqrt{\pi}$, then taking the parity of n. Ideally, teleportation of $|0\rangle_{\text{gkp}}$ onto a CV cluster yields $|+\rangle_{\text{gkp}}$, which would return a qubit 0 outcome 100% of the time when measuring Pauli X, and 50% of the time when measuring Pauli Z. In Fig. 3.11 (a), we plot—as a function of squeezing and loss—the mean probability of obtaining the outcome 0 from a Pauli X measurement. As expected, we see that high squeezing and low loss provide the best readout. Interestingly, we also find that above 10 dB of squeezing (an amount comparable to the per-peak squeezing of the teleported state), loss is the major determinant of readout quality for this Pauli measurement. We are also interested in how far individual runs can deviate from the mean probability of correct readout, so in Fig. 3.11 (b), we plot the spread of the probability from (a) as a function of squeezing and loss. We find that the spread is quite small (only a tenth of a percentage in the worst case), meaning that individual runs yield values close to the correct readout.

The story differs for the Pauli Z measurement. Fig. 3.11 (c), where we plot the spread for the probability of reading out a qubit 0 outcome in the Pauli Z measurement, shows that there can be significant deviation from the 50% mean value in the presence of finite squeezing and loss. While in the ideal case, every instance of the teleportation circuit yields a teleported state that has a 50-50 chance of generating outcomes 0 or 1 in a Pauli Z measurement, finite squeezing and loss can distort the readout odds: certain instances of the teleportation circuit (heralded by homodyne measurements of the ancillae) can cause one outcome to be favoured by several extra percentage points. Interestingly, in contrast to the Pauli X measurement, the spread of results in the Pauli Z measurement are more sensitive to the level of squeezing than to the level of loss.

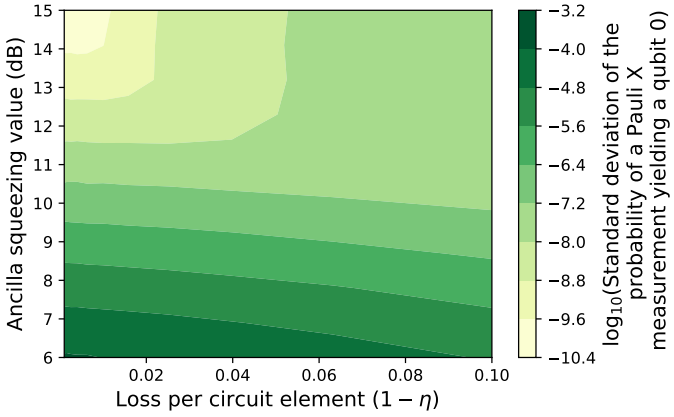
While it is intuitive that low loss and high squeezing should yield qubit outcomes closer to ideal, it is interesting—for this choice of initial GKP states—that that loss (squeezing) is the greater noise source for Pauli X (Z) measurements. It is worth investigating, however, whether a different feedforward function or postselection of outcomes can improve the faithfulness of qubit readout, and whether the homodyne outcomes can be used to assign a confidence rating to the qubit readout of the teleported mode. Such strategies have been explored in the context of fault-tolerant quantum computing architectures with GKP qubits [22, 35, 39].

GKP T Gate Teleportation

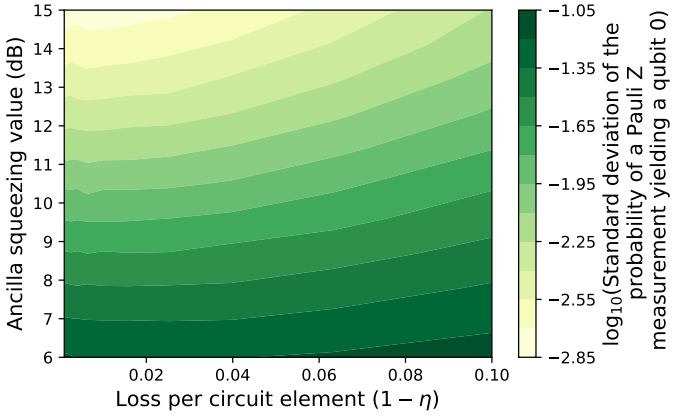
While Clifford gates can be performed using Gaussian resources in the GKP encoding, one requires a non-Clifford—for GKP states, non-Gaussian—gate to achieve universality. As we discussed in Section 3.5.3, the qubit T gate can be applied via gate teleportation using a GKP magic state $|M\rangle_{\text{gkp}}$. A realistic T gate



(a)



(b)



(c)

Figure 3.11: Teleportation of $|0^{\epsilon}\rangle_{\text{gkp}}$ onto a p-squeezed state. Ideally, the teleported state should be $|+\rangle_{\text{gkp}}$, with Pauli X (Z) measurement of the state yielding a qubit 0 outcome with 100% (50%) probability. However, with noisy teleportation circuits (loss and measurement-based squeezing in the CZ gate), each instance of the circuit, conditioned on the probabilistic value of the teleportation feedforward, yields a slightly different state. We simulate the noisy teleportation circuit 500 times. In (a) and (b), we plot the mean and the (log of the) standard deviation for the probability of a Pauli X measurement yielding 0, as a function of circuit loss and squeezing. Having found the mean probability of a Pauli Z measurement yielding 0 to be roughly 50% across squeezing and loss levels, in (c) we provide the (log of the) standard deviation for the probability a Pauli Z measurement yields 0.

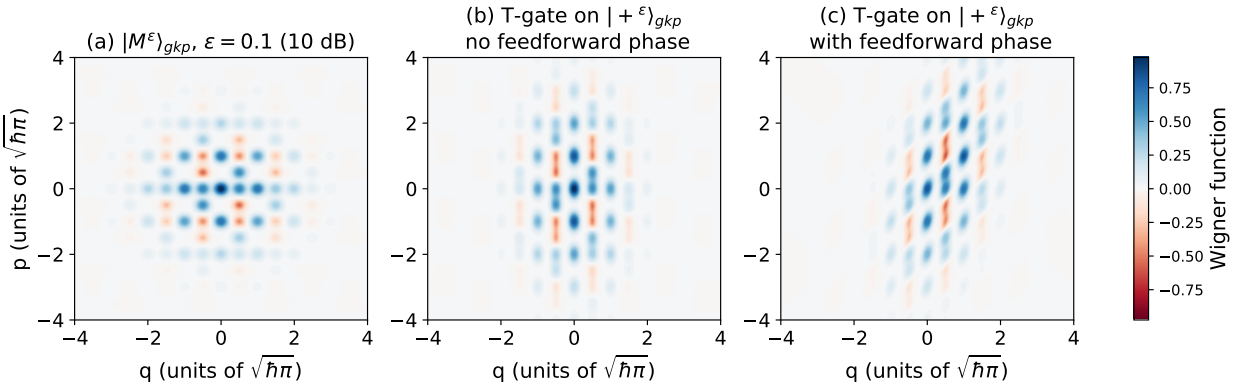


Figure 3.12: (a) Wigner function for a GKP magic state $|M^\epsilon\rangle_{\text{gkp}} = \hat{E}(\epsilon)[\frac{1}{\sqrt{2}}(e^{-i\pi/8}|0\rangle_{\text{gkp}} + e^{i\pi/8}|1\rangle_{\text{gkp}})]$ with $\epsilon = 0.1$ (10 dB). (b) Wigner function for T gate teleportation applied to $|+^\epsilon\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB) using the circuit from Fig. 3.3 and postselecting on an outcome that does not require the feedforward phase gate. (c) Output from the same circuit, but postselecting on an outcome that does require the feedforward phase gate. The gate teleportation circuit for the T gate with realistic components is involved, but now simulatable using the `bosonic` backend of `Strawberry Fields`.

teleportation is no small feat: an ancillary GKP state is required in addition to the three ancillary squeezed states for the CZ gates and the feedforward phase gate, and the multiple beam-splitters each incur loss. As far as we are aware, no investigation of realistic T gate application has been carried out, likely due to the difficulty of simulation. Let us close this gap.

First, in Fig. 3.12, we present the result of a T gate applied to $|+^\epsilon\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB); in the ideal limit ($\epsilon \to 0$), this should return the state $|M\rangle_{\text{gkp}}$. For sake of clarity, in this figure we assume that the CZ gate and the phase gate are applied perfectly, but employ a finite-energy resource state

$$|M^\epsilon\rangle_{\text{gkp}} = \hat{E}(\epsilon)[\frac{1}{\sqrt{2}}(e^{-i\pi/8}|0\rangle_{\text{gkp}} + e^{i\pi/8}|1\rangle_{\text{gkp}})], \tag{3.55}$$

also with $\epsilon = 0.1$ (10 dB). In (a) we present the Wigner function for a finite energy $|M^\epsilon\rangle_{\text{gkp}}$ to use for comparison, while in (b) we provide the Wigner function for the output of the T gate circuit given a measurement outcome on the ancillary GKP mode that does not require the feedforward phase. In (c), we present the same as (b) but for an instance of the T gate circuit that does require feedforward phase. We see that the gate works to some extent: the positive and negative peaks in (b) and (c) are in the same places as positive and negative peaks in (a), and we recall from Section 3.4 that the logical information of the state is encoded in the weights. However, even with the CZ and feedforward phase gate implemented perfectly, we still see some differences between the Wigner functions. First, notice that the p quadrature variance of each peak is now larger than the q quadrature variance, even though they began as symmetric. This is because the CZ gate adds the q variance of the resource magic state to the data state. Second, in (c), there is a shearing effect of both the overall envelope and the individual peaks due the phase gate, as also seen in Section 3.7.2.

Next, we examine the effect of finite-energy magic states on the quality of the T gate teleportation, keeping the rest of the circuit ideal. We again simulate the T gate circuit applied to $|+^\epsilon\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB), this time varying the value of ϵ for $|M^\epsilon\rangle_{\text{gkp}}$. For each ϵ, we perform 500 simulations, each time producing a different output state based on the probabilistic measurement of the ancilla. For the output state

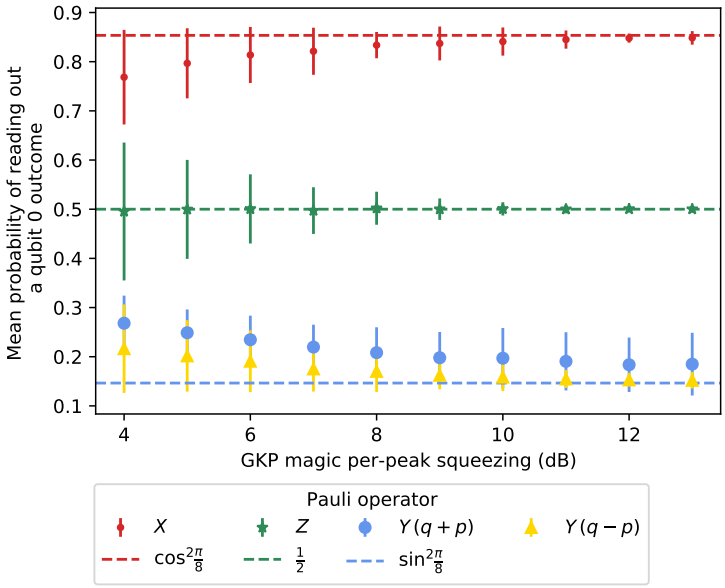


Figure 3.13: Mean probability of obtaining a qubit 0 outcome from different Pauli operator measurements after applying a T gate to $|+^{\epsilon}\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB). We vary the per-peak squeezing of the GKP magic resource state and see how well the qubit readout matches the ideal readout values for the Pauli operators (dashed lines). We see that a per-peak squeezing on the order of 11 dB (close to that of the data state) is sufficient to retrieve correct readout to within 1%. The simulation provides another example of the usefulness of having two methods for performing Pauli Y measurements, as seen in Fig. 3.10; homodyne detection and binning along $q - p$ yields a more faithful outcome.

of each simulation, we calculate marginal distributions in phase space to determine the probability of reading out a qubit 0 outcome for the four Pauli measurements (three Pauli operators, and two ways of measuring Pauli Y). In Fig. 3.13, we plot how the mean probability of obtaining a qubit 0 readout value for the Pauli measurements varies as a function of the magic state's finite energy parameter ϵ. We additionally provide what the ideal probabilities for readout should be for the Pauli measurements. We find that, with smaller epsilon (increased per-peak squeezing), the statistics become more closely aligned with the ideal values, along with less variation in the output state of each instance of the circuit, with ϵ on the order of 11 dB seeming to be sufficient to achieve correct readout to within 1%. It is good news that the quality of magic state required is not significantly higher than that of the data state. We also note that measuring the Pauli Y operator by performing homodyne along $q - p$ again yields more faithful readout of the qubit outcomes.

Finally, we perform the same simulation of a GKP T gate applied to $|+^{\epsilon}\rangle_{\text{gkp}}$, this time implementing the CZ and phase gates using realistic components: lossy beam-splitters, inefficient homodyne measurements, and measurement-based squeezing. We fix the magic state to have $\epsilon = 0.1$ (10 dB); we set the efficiency for the four homodyne measurements (three for measurement-based squeezing in the CZ and feedforward phase gate, and one on the ancillary magic state) to be 99%; and we choose a squeezing level of 12 dB for the three squeezed states used for measurement-based squeezing. We investigate how the Pauli operator readout varies with respect to a loss parameter η, which we apply in multiple places throughout the circuit: right after each state (GKP or squeezed) is initialized and after each beam-splitter.

A few comments are in order. First, in the case of no loss, we already see that the readout is not as good as when the CZ and phase gates are applied perfectly; this is undoubtedly due to the change from idealized

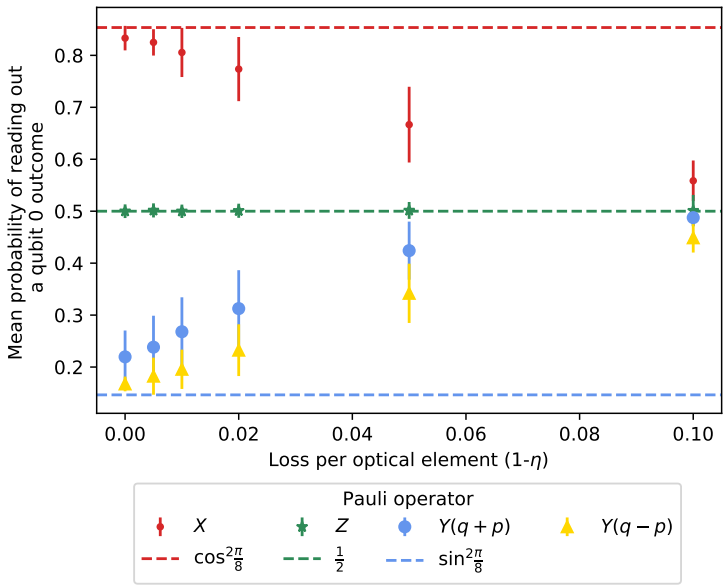


Figure 3.14: Mean probability of obtaining a qubit 0 outcome from different Pauli operator measurements after applying a T gate to $|+^{\epsilon}\rangle_{\text{gkp}}$ with $\epsilon = 0.1$ (10 dB). Here, the gate teleportation is effected using imperfect optical elements (see Figs. 3.3 and B.3 for circuit breakdowns). We fix the ancillary magic state to have the same ϵ as the data state; we employ 12 dB squeezed states for measurement-based squeezing, and set the homodyne efficiency to 99%. We see how readout changes as we vary η, which parametrizes the loss channel we apply after each of the five states are initialized and after each of the five beam-splitters.

squeezing to measurement-based squeezing; even with squeezed resource states of 12 dB, the readout has room for improvement, motivating deployment of even more highly squeezed states. Second, as loss per optical element is increased, the readout quality quickly deteriorates. This is likely due to the number of loss channels the data state undergoes, since it must pass through four beam-splitters to perform the teleportation. Notably, only the Pauli Z readout is still close to the desired value, but this is coincidence: a broad, noisy distribution is essentially uniform over the bins applied to the quadrature outcomes, resulting in a 50% chance of reading out a logical 0 from a Pauli measurement. The sensitivity of the optically-teleported T gate to loss, due to its many components, is additional motivation for pursuing passive optical implementations of computing with GKP states [128]; by potentially eliminating in-line squeezing, these architectures reduce opportunities for loss and noise brought in by measurement-based squeezing.

3.8 Summary and Open Problems

In this chapter, we have introduced a formalism for simulating continuous-variable quantum states. At its heart is a representation of these states as linear combinations of Gaussian functions in phase space. This novel framework can be used to analyze and simulate valuable classes of CV states, namely bosonic qubits like GKP, cat, and Fock states, under realistic transformations. Our mathematical framework inherits the simplicity and convenience of the Gaussian CV formalism, requiring only to track how transformations and measurements update the weights, means, and covariance matrices of the Gaussian functions in the linear combination. In addition to the straightforward inclusion of Gaussian channels in our formalism, a salient class of non-Gaussian transformations and measurements—photon-number-resolving measurements and gate

teleportation for GKP and cat qubits—can also be simulated using our framework.

Enabled by our formalism, we provided simulation methods that outperformed state-of-the-art CV simulators that employ the Fock basis. Using our new method, we were able to perform novel simulations of bosonic qubits in realistic settings, informing future design decisions for quantum computers based on bosonic qubits. We focused on GKP qubits, examining how they transform under gates that leverage measurement-based squeezing, how they interface with CV clusters, and how they transform under non-Clifford gates implemented via gate teleportation. While we only investigated how the readout of Pauli operators was affected in these settings, we emphasize that our simulator outputs complete state information in the form of the elements required to construct the Wigner function, which can be used as input to other simulators or to more advanced analytical tools, such as the modular subsystem decomposition [61, 63].

Numerical simulations leveraging our formalism and methods were performed in the new `bosonic` backend of `Strawberry Fields`, an existing feature-rich Python library for simulating CV optical circuits. The backend, implemented by the authors, is accessible through a user-friendly frontend of the public repository [102], along with some tutorials on its use [103–105]; we hope this encourages exploration by those interested in CV quantum information.

We anticipate several useful applications of our work. For one, our new formalism will prove valuable in the design of circuit modules for quantum computers that employ bosonic qubits, as it can identify, for example, where losses need to be reduced in a circuit, or what level of ancillary squeezing is required to attain a target gate fidelity. While our focus in the simulations has been on GKP qubits, similar analysis of realistic gate application for cat states can be performed. Second, it is known that in optical settings where deterministic high-order nonlinearities are currently lacking, bosonic qubits will need to be produced probabilistically in the near term. There are promising heralded state generation techniques that involve photon-number-resolving measurements on some modes of a multimode Gaussian state [16, 59], a scenario which we have already discussed falls neatly within our formalism. Previous investigations and optimizations of this technique were limited by the speed and memory limitations of the Fock basis; a straightforward extension would be to use our new formalism to develop these bosonic state preparation methods even further. Yet another state preparation proposal for GKP states relies on interacting or "breeding" cat states with beam-splitters and homodyne measurements on some of the modes [58, 141]. Since the circuits in those protocols consist of Gaussian transformations and measurements applied to cat states, our formalism applies here too.

Finally, there are several avenues for future research. First, it is worth investigating the connection between other analytical representations of CV states (such as Riemann-Theta functions for GKP qubits [36, 50], or the shifted Fock basis representation of cat states [90]) and the ones we have presented here to identify any additional opportunities for faster or more accurate simulation. Second, for the class of states and maps that we have considered, it would be of interest to determine the exact mathematical conditions for physicality in terms of the parameters of the states (weights, vectors of means and covariances matrices). Moreover, while we considered a wide class of transformations which fall under this formalism, it would be valuable to determine the most general class of transformations which can be reduced to manipulating the weights, means and covariances of the Gaussian functions in the linear combination. Third, it would be fruitful to conduct a deeper analysis of connections between our framework, the stabilizer formalism [136, 137] and the sum of stabilizer states technique [138–140] that are used to speed up simulations of qubit circuits. Specifically, it would be worthwhile to determine whether it is possible for finite-energy GKP stabilizer states to be efficiently simulated under GKP Clifford operations—we already know it to be possible for their infinite-energy

counterparts [52]. Additionally, we chose to express our states as sums of Gaussian functions, each individually easy to simulate in phase space; however, it is also known that some Gaussian states can be distilled into GKP magic states [36], meaning those Gaussian states must lie outside the qubit space stabilized by GKP Pauli operators. Therefore, it would be interesting to determine how Gaussian states can be expressed as linear combinations of GKP qubit stabilizer states and compare the simulation scaling in that case. Lastly, large-scale simulation of bosonic qubits under realistic noise models is challenging due to exponential scaling. Our formalism is a step towards numerically tractable models capturing some of those effects, either by developing suitable approximations of the states based on the mathematical framework provided, or by using the simulator to develop more detailed models for the evolution of qubit-level information or noise, then leveraging existing work in the field of quantum fault-tolerance for qubits.

Armed with a deepened intuition of the phase space behaviour of GKP states, and with the state preparation technique provided in Chapter 2, we now move to introduce and analyse a photonic architecture for fault-tolerant quantum computing based on GKP states and Gaussian resources.

Chapter 4

Noise Analysis for a Fault-Tolerant Photonic Quantum Computer

This chapter is based on [22], co-authored with Rafael N. Alexander, Michael Vasmer, Ashlesha Patil, Ilan Tzitrin, Takaya Matsuura, Daiqin Su, Ben Q. Baragiola, Saikat Guha, Guillaume Dauphinais, Krishna K. Sabapathy, Nicolas C. Menicucci, Ish Dhand. The work was collaborative, and published in *Quantum*. My work was mainly supervised by Krishna K. Sabapathy, Nicolas C. Menicucci and Ish Dhand. Rafael N. Alexander and I shared first authorship. My main contributions were in performing derivations for the noise model discussed in Sections 4.3 and C.1, designing the inner decoder discussed in 4.3, determining weight assignments for the outer decoder as given in C.3, and writing various sections and the appendices of the paper. The numerical simulations in 4.4, based on these decoding strategies I contributed, were performed by Michael Vasmer and Guillaume Dauphinais. I have included summaries of sections of the paper where I was not a main contributor as part of the background material in the chapter (Section 4.2), since my contributions would lack important context without review of these sections. The work benefited from helpful discussions with Xanadu colleagues.

4.1 Introduction

As discussed in Chapter 1, photonics is a promising platform for quantum computation due to its potential to scale and network devices, and due to the flexibility in choice of error correcting code. In this chapter, we consider a photonic architecture for fault-tolerant quantum computation relying on Gottesman-Kitaev-Preskill (GKP) and Gaussian states of light. The architecture relies on the probabilistic but heralded GKP state preparation scheme from Chapter 2, and the noise analysis is informed by the representation of GKP states as linear combinations of Gaussian functions as presented in Chapter 3.

Current architectures for scalable and universal photonic quantum computing live on two extremes. The first type of architectures [34, 142] aim to use a division of labor between Gaussian and non-Gaussian resources (see Table 4.1). The Gaussian resource is provided by easy-to-generate and scalable CV cluster states, which are multi-mode Gaussian states stitched out of squeezed vacuum states [143]. There has been substantial progress in designing and deterministically generating CV cluster states in one [144–146], two [147–151], and higher dimensions [95, 152, 153]. In each of these architectures, the quantum information is encoded in a bosonic qubit introduced by Gottesman, Kitaev and Preskill [13], due to its ability to

Table 4.1: Examples and implications of Gaussianity and non-Gaussianity in the context of measurement-based quantum computing with GKP qubits. Note that only the generation of GKP qubits requires cryogenic temperatures in our architecture. Qubit Clifford gates are effected with CV Gaussian transformations; qubit Pauli measurements are performed with CV Gaussian measurements; and non-Clifford gates require ancillary GKP magic states plus Gaussian transformations and measurements.

	Gaussian	**Non-Gaussian**
States	q-/p-squeezed; CV cluster states	GKP computational and magic states
Transformations	squeezing; displacement; linear optics	None
Measurements	homodyne	PNRs
Used to implement	Clifford gates	Non-Clifford gates
Experimental Characteristics	"Easy"; room temp.; deterministic	"Hard"; cryogenic temp.; probabilistic

interface deterministically with the CV cluster state via the same entangling operations that generate the cluster. Clifford circuits—which make up the majority of operations required for a fault-tolerant quantum computer—can be implemented via measurement-based quantum computation (MBQC) on the CV cluster state. While the entangled resource state need not be composed entirely of bosonic qubits, a truly on-demand supply of GKP encoded states is still required; they provide the necessary non-Gaussianity, implement non-Clifford gates, and correct CV errors. Thus far, prior work has required that such qubits can be supplied and coupled to the cluster state deterministically at regular intervals.

The second type of architectures includes the schemes developed for the cat-basis encoding [94, 154], the GKP encoding [35, 38, 44], and the dual-rail encoding [14, 91, 155]. While these architectures can provide an extra resiliency to noise, they must contend with the non-deterministic generation of individual qubit states, particularly in the former two cases where the states have a complicated structure. The latter case has the added challenge of non-deterministic entangling or "fusion" gates, which are required to grow a cluster state. Each qubit gate is eventually implemented by consuming probabilistically generated photons, which imposes formidable multiplexing requirements for cluster state generation—unlike schemes for generating CV cluster states.

In light of these two approaches, it is important to devise a scheme that combines the best of both worlds. In this chapter, we consider a *hybrid* CV/bosonic qubit cluster state where each mode is probabilistically encoded as a GKP qubit or as a squeezed state. This approach leverages the deterministic generation of squeezed states and their ability to entangle with either another squeezed state or a GKP state via the same deterministic operation [34], as well as the error-correcting capabilities of GKP states, while accounting for the current state of theory and technology for their preparation in the optical domain. The numerous procedures for generating optical GKP states that have been proposed tend either to be non-deterministic, as they rely on post-selected measurements directly [16, 17, 19, 20, 74, 156] or indirectly [58, 59, 141]; or require the experimentally challenging conditions of coherent interactions with matter [60, 157] or extremely strong optical nonlinearity [157]. Recent advances in photon-number-resolving (PNR) detectors [57, 158–160] have substantially improved the viability of the post-selection approach in the near term, with methods based on Gaussian boson sampling (GBS) [16, 17, 19, 20] now within reach of state-of-the-art optical devices. Low-probability sources can be improved with the help of multiplexing at the cost of an increased overhead.

Here, we propose an architecture for measurement-based quantum computing that possesses the advantage

of CV-based schemes and yet is compatible with probabilistic GKP qubit sources. We consider a *hybrid* CV cluster state where each mode is substituted with a GKP qubit at random and with probability $(1-p_0)$—or said the other way, a qubit cluster state where each node is substituted with a squeezed state with *swap-out probability* p_0. The precise state we consider is the lattice from the Raussendorf, Harrington, Goyal (RHG) model [161–163], but our scheme can readily accommodate other error-correcting codes. Our use of CV resources affords us an important alternative over existing approaches, wherein a qubit that failed to be produced must be erased from the lattice. Instead, we replace the no-show qubit with a squeezed vacuum state: it can still encode logical information (albeit not as well as a GKP state [61]) but has the distinction of being Gaussian and thus easily producible [1]. This approach—one of the main innovations in this chapter—propels us beyond existing fault tolerance methods, such as those that rely on lattice renormalization to deal with defects [164–167]. To characterize the robustness of our architecture as a function of p_0, we perform Monte Carlo simulations of our architecture operating as a quantum memory. We observe a minimum required squeezing of 10.5 dB or a maximum tolerable swap-out probability of $p_0 \approx 0.236$; for an experimentally accessible squeezing value of 15 dB [87], our simulations suggest a swap-out threshold of $p_0 \approx 0.133$, which translates to substantially reduced multiplexing requirements for GKP generation. In part these result stem from a tailored decoding procedures that we present and which allow us to perform fault-tolerant computation on our hybrid resource state.

This chapter is structured as follows. Section 4.2 provides the necessary background, a summary of a planar architecture that could be used to generate the entangled resource state, and a review of how fault-tolerant logical-level computation can be performed. The method to implement quantum error correction, including a specialized decoder for the hybrid lattice, is presented in Section 4.3. Section 4.4 presents the fault-tolerance thresholds for our architecture. We discuss open challenges in Section 4.5.

4.2 Background

In this section, we review the pieces required for the analysis of our photonic fault-tolerant quantum computing architecture. In Section 4.2.1, we review GKP qubits, a proposal for their probabilistic generation via Gaussian Boson Sampling devices, and their ability to interface with CV cluster states. This leads to Section 4.2.2 where we review the measurement-based quantum computation formalism, and summarize Section V of [22] detailing how logical gates can be performed in our architecture. In Section 4.2.3, we provide a summary of Section III from [22] which outlines how modular planar chips can be used to generate the hybrid CV/GKP resource state of our architecture.

4.2.1 Qubits Encoded Into Bosonic Modes

Gottesman-Kitaev-Preskill (GKP) qubits

Bosonic qubit encodings are two-dimensional subspaces within the infinite-dimensional Hilbert space of a bosonic mode. Good choices of this two-dimensional subspace allow for experimentally convenient ways of preparing the encoded qubit states, implementing desired unitary gates, and faithfully performing measurement readout. In some cases, the redundancy of the full infinite-dimensional Hilbert space can even be leveraged to detect and correct CV errors – random Gaussian displacements, rotations, and photon loss, for a few –

[1] We cannot replace all modes with squeezed vacua because that would negate the error-correction benefits of the encoded qubits [135].

without destroying the encoded information. Examples of bosonic codes include GKP [13], dual-rail [14, 97], cat [96, 168], hypercat [169, 170], binomial [171], and general rotation-symmetric codes [98].

For reasons we will describe, our architecture exploits the GKP encoding, which we have extensively reviewed in Section 2.2 and Section 3.2.3. For notational convenience we restrict our discussion to square-lattice GKP encoding but the results can be generalized to GKP states on other lattices. In their ideal form, the GKP qubit states $|0\rangle_{\text{gkp}}$ and $|1\rangle_{\text{gkp}}$ are defined as Dirac delta combs with a spacing of $2\sqrt{\pi}$ in position space [13]:

$$|\mu\rangle_{\text{gkp}} = \sum_n |(2n+\mu)\sqrt{\pi}\rangle_q \,, \; \mu = 0, 1. \tag{4.1}$$

As previously reviewed, the chief advantage of GKP states is that qubit Clifford operations map to CV Gaussian operations, which can feasibly be implemented *deterministically* using linear optical elements, homodyne detection, and Gaussian states of light [13]. In Appendix C.2, we provide optical circuits for the application of GKP qubit gates in detail. Deterministic all-optical entangling gates are a distinct advantage of the GKP encoding over dual-rail encoding schemes [14]. Moreover, since qubit Pauli measurements correspond simply to homodyne measurements, the computational module has the potential to operate at faster speeds and higher efficiencies than architectures relying on photon-counting detectors [87, 172]. Non-Clifford operations require a non-Gaussian resource. Unitary implementations of the T gate can be achieved using a cubic-phase interaction, though this is difficult in practice [120], and does not perform well for finite-energy states [62]. As an alternative, T gates can be performed through gate teleportation by preparation of a GKP magic state [13], which we also review in Appendix C.2. Unfortunately, ideal GKP states are non-normalizable states with infinite energy, so we must consider their approximate, finite-energy versions as reviewed in Section 2.2.2 and given in Eq. (3.20). While finite-energy effects will be ever-present in any real implementation, the noise they introduce does not preclude GKP states from being useful for error correction and fault-tolerant quantum computation [15, 33, 34].

Using bosonic codes as the physical qubits for a fault-tolerant quantum computing architecture provides two tiers of protection from noise. The first comes from the bosonic code itself. GKP qubits possess a degree of intrinsic robustness to those bosonic noise channels that result in small displacements (relative to the $\sqrt{\pi}$ lattice spacing) in phase space. This includes weak levels of photon loss, the dominant error mode for quantum communication [15, 28]. Any shift that is less than half the lattice spacing ($\sqrt{\pi}/2$) can be corrected by non-destructively measuring the GKP stabilizers—which are $2\sqrt{\pi}$ shifts in either position or momentum. This CV error-correction procedure outputs continuous syndrome data, which can be used to undo the displacement with the help of a decoder. Noise that leads to larger displacements can result in errors that are undetectable by measuring only the GKP qubit stabilizers. In the fault-tolerant regime, these larger displacements are much less likely to occur, and so occasional errors on GKP qubits can be corrected by applying the second layer of protection: a qubit quantum error-correcting code. Implementing these codes requires only Clifford gates and Pauli measurements, both of which are easy (Gaussian) for the GKP qubit encoding.

An essential part of any quantum error correction procedure is the *decoder*, which specifies the recovery operation that has to be applied for given syndrome data. The two-layer structure of the error correction described above requires two stages of decoding. The first of these stages translates continuous GKP-stabilizer syndrome data into the operations required to return to the GKP code subspace housed in each mode, up to qubit-level errors. It can also provide some detailed information about the relative likelihood of different discrete qubit-level errors [35]. The second stage maps syndrome data obtained from measuring

the higher-level qubit-code stabilizers to a qubit-level recovery operation. To avoid confusion, we refer to the former as the *inner decoder*, and the latter as the *outer decoder*. Decoders that are tailored for our architecture are described in more detail in Sections 4.3.2 and 4.3.3, respectively.

Measurement-Based Quantum Computing with CV Cluster States and GKP Qubits

Modes of light are not well suited to serving as stationary quantum data registers. Optical modes can interact with only a few optical elements before they must be measured or else lost. Fortunately, this constraint is compatible with the measurement-based model for quantum computing, where each quantum data register is entangled to a constant number of others, and then measured at a detector—the entire computation being specified by which measurements are chosen. Here we review relevant details and terminology on CV photonic measurement-based quantum computation.

Measurement-based quantum computation in the qubit setting involves preparing an entangled resource state (most commonly, a cluster state [173]), and performing a sequence of single-site adaptive measurements [174]. These cluster states are specified by a graph; for each node a qubit is prepared in the $|+\rangle$ state and for each edge a CZ gate is applied. Cluster states are said to be *universal resources* if they enable universal quantum computation when given access to adaptive single-site measurements.

This paradigm can be generalized to the CV degrees of freedom present in a bosonic mode: CV cluster states are Gaussian entangled states that enable CV measurement-based quantum computation via local measurements [143]. The simplest of these are referred to as *canonical* CV cluster states [175], which are constructed by applying controlled-Z gates $e^{i\hat{q}\otimes\hat{q}}$ to momentum-squeezed vacuum states. CV cluster states with any graph can be generated on demand since both the controlled-Z gates, and the preparation of momentum-squeezed vacuum states can be implemented deterministically.

Ideal CV cluster states cannot be normalized and correspond to unphysical infinite-energy states. In a similar way to the GKP qubit, the approximate nature of physical CV cluster states can be captured by a finite-width Gaussian envelope structure of the state's position space wavefunction:

$$\begin{aligned} |C(\mathbf{V},\epsilon)\rangle &= e^{i\hat{\boldsymbol{q}}^{\mathrm{T}}\mathbf{V}\hat{\boldsymbol{q}}/2}\left[\frac{\sqrt{\epsilon}}{\pi^{1/4}}\int_{-\infty}^{\infty} ds\, e^{-s^2\epsilon/2}\,|s\rangle_q\right]^{\otimes N} \\ &= \left[\frac{\epsilon}{\sqrt{\pi}}\right]^{\frac{N}{2}}\int_{\mathbb{R}^N} d^N\boldsymbol{s}\, e^{i\boldsymbol{s}^{\mathrm{T}}\mathbf{V}\boldsymbol{s}/2}e^{-\boldsymbol{s}^{\mathrm{T}}\boldsymbol{s}\epsilon/2}\,|\boldsymbol{s}\rangle_{\boldsymbol{q}}\,, \end{aligned} \tag{4.2}$$

where $\mathbf{V}$ is a real symmetric adjacency matrix corresponding to the cluster state's graph and $\epsilon/2$ is the variance of each momentum-squeezed vacuum state in the momentum quadrature. The case of $\epsilon \to 0$ corresponds to the infinite squeezing limit.

Note that the CV controlled-Z gate $e^{i\hat{q}\otimes\hat{q}}$ is common to both the GKP qubit encoding and canonical CV cluster state generation. Therefore it is possible to generate a *hybrid* cluster state with nodes comprising momentum-squeezed states, GKP qubits, and their common CZ gates [34]. This is one of the key concepts that enables computation with our hybrid resource state.

4.2.2 Cluster States: Fault Tolerance and Logical Computation

That quantum information can be processed reliably in the presence of noise is key achievement of quantum error correction [70, 176]. Given a logical circuit to be implemented, the idea is to redundantly encode the information content of the logical qubits into larger collections of physical qubits and perform computation on

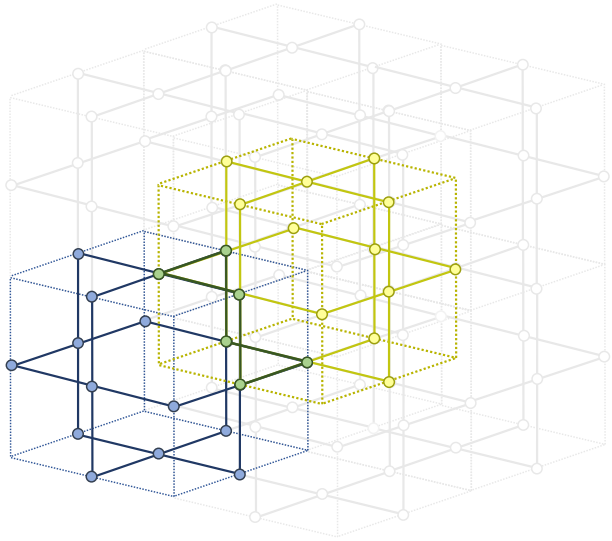


Figure 4.1: A primal cell of the RHG lattice, i.e., a $2 \times 2 \times 2$ stack of unit cells (blue), with the dual cell identified in the middle (yellow). The nodes and edges which overlap are highlighted in green. The faces of the cells are also called primal or dual, and comprise primal or dual boundaries. In surface code terms, smooth (rough) boundaries end on the faces of primal (dual) cubes. All the links are CZ gates.

these collections. If physical qubits of sufficient quality are available and if sufficiently precise operations and measurements can be performed on these qubits such that the noise strength remains below the threshold of the specific code used, then the logical quantum circuit can, in principle, be applied with arbitrarily high precision [177–179].

In practice, it is often desirable to choose a quantum error-correcting code capable of tolerating a high error rate and not requiring long-ranged connectivity between physical qubits. The *surface code* [180, 181] is a commonly used code because it has both these properties, enjoying a high threshold of $\sim 1\%$ [182–184] and only needing nearest-neighbor connectivity of qubits in two spatial dimensions. However, optical quantum computing architectures are better suited to measurement-based approaches to quantum error-correcting codes, often implemented using cluster states corresponding to foliated (i.e., layered) lattice sheets. Perhaps the best studied cluster-state based error-correcting code is the RHG lattice [161–163]. In the foliated picture, it can be thought of as alternating so-called primal and dual sheets of 2D cluster states that encode the surface code. This special topology is what leads to fault tolerance: error detection and decoding involve multiple mutually entangled layers at a time. The RHG lattice serves as a good first candidate to study for our architecture because not only is it universal for MBQC, but it has high fault-tolerant computational error thresholds ($\lesssim 1\%$). A full fault-tolerant computation can be performed on the RHG lattice with the help of the RHG scheme [161–163, 185]. A schematic of the RHG lattice is presented in Fig. 4.1.

A subject of growing interest is to consider a two-layer encoding: bosonic qubits concatenated into a qubit cluster [35, 38, 39, 41, 42, 47, 48]. The main motivation here is that CV errors larger than those the inner bosonic code can handle are picked up and corrected by the outer qubit code. Gate-based models for the concatenated GKP-surface code catering to a superconducting platform were considered in Refs. [39, 41, 47, 48]. Measurement-based quantum computation using GKP qubits encoded in a cluster state compatible with a photonic architecture was considered in Refs. [35, 38]. The approach was to generate a 3D cluster state to implement a measurement-based analogue of the surface code using GKP qubits. The cluster state was

generated using a post-selection (fusion-based) approach that is non-deterministic by nature. This work also introduced an analogue quantum-error-correction scheme, where the real-valued measurement outcome from the homodyne measurement was explicitly used in the decoding procedure. The errors considered in Ref. [35] are finite squeezing effects in the GKP state preparation along with a Gaussian random displacement noise. More realistic noise, such as loss in the entangling gates and the homodyne measurements, was also considered in Ref. [38]. Furthermore, the GKP state preparation is considered to be deterministic and the final state generated is a connected cluster populated only by GKP states.

While the focus of this chapter is on error correction for a quantum memory (i.e. no logical gates are applied), for completeness, we briefly review how logical qubits are encoded and gates are performed in the RHG lattice using the lattice surgery method [186–191]. For complete details, we refer to Section V of [22]. The essential elements of lattice surgery in our architecture are [22, 190]:

1. *Logical qubits*: patches of physical qubits, disentangled from the rest of the lattice via Pauli Z measurements, define a logical qubit, with the code distance given by the shortest width of the patch. Patch boundaries ending on primary (dual) lattice faces are termed smooth (rough) boundaries. Initializing a logical $|\bar{+}\rangle$ ($|\bar{0}\rangle$) corresponds to measuring the first temporal layer of a patch in X (Z). Recall that in the GKP encoding, Pauli X (Z) measurements correspond simply to homodyne measurements in p (q) quadrature.

2. *Logical $\bar{Z}$ and $\bar{X}$ gates and measurements*: chains of bit flips, applied in software, between smooth (rough) edges implement logical Pauli X (Z) gates. Measuring primal/dual sheets in Z/X (X/Z) implements a Pauli Z measurement, with the outcome given by parity along the same chains that define logical operators.

3. *Entangling gates*: patches can be merged by switching all the Z measurements between patches to X measurements, yielding a single logical qubit out of two logical qubits, with the state given by the two preceding logical qubits. This operation can be used to construct a logical CNOT gate [186, 190].

4. *State injection*: non-Clifford logical gates can be implemented by injecting a physical magic state into the RHG lattice, using patch mergers to augment the single magic state up to a whole logical qubit magic state of correct code distance, then using gate teleportation to apply a non-Clifford gate.

4.2.3 An Architecture for Photonic Quantum Computing With Hybrid Resource States

This section summarizes the different components of our architecture. For complete information, we refer to Section III of [22]. The architecture comprises four modules, which we now survey. The first two modules are entirely dedicated to the preparation of single-mode GKP and squeezed states; entanglement into the cluster state and measurement are relegated to the third and fourth modules.

State preparation

The state preparation module generates high-quality GKP qubits, albeit with low probability. GBS devices (see Fig. 4.2) are promising probabilistic sources of GKP qubits, among other non-Gaussian states [16, 17, 19, 20] (see also Chapter 2). GBS state preparation consists of sending N displaced squeezed vacuum states into a general interferometer on N modes, followed by PNR detectors on $N-1$ of the modes, as depicted in Fig. 4.2.

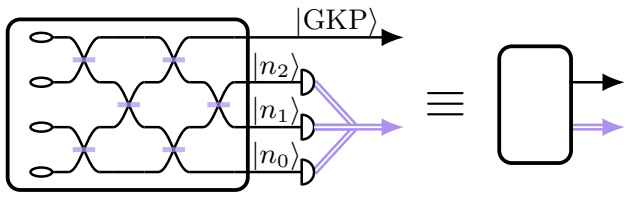


Figure 4.2: **GBS devices for state preparation.** (left) A single integrated photonic device implementing GBS-based preparation of non-Gaussian states based on the schemes presented in Refs. [16, 17, 19, 20] and in Chapter 2. Here, the ellipses on the left represent on-chip sources of squeezed light (such as ring resonators); the intersecting lines emanating from the sources represent the waveguides of the interferometer with purple rectangles at intersections between lines denoting beamsplitters; the half-circles on the right represent PNR detectors. The emitted light from one output port is in a chosen non-Gaussian state subject to obtaining the correct click pattern $\{n_i\}$ at the PNR detectors connected to the remaining output ports. The double purple lines represent classical logic, which is used to trigger a switch on the emitted port. (right) A simplified representation of a single GBS device.

The number of modes, the displacement, squeezing, and inteferometer parameters, as well as the photon number pattern at the PNR detectors, can all be tuned so that the device can herald the desired high-fidelity non-Gaussian output state. This procedure exploits the non-Gaussianity of PNR detectors and can generate arbitrary logical single-qubit states for a variety of bosonic encodings, including the GKP and cat encodings. Since the generation of the desired state requires a particular pattern of photon number detection outcomes to be observed, the generation is non-deterministic but heralded.

As a concrete example, consider that small-scale GBS devices made up of 3-mode interferometers, two PNR detectors registering up to 7 photons, and three momentum-squeezed vacuum states with up to 12 dB of squeezing have the potential of producing $|0^{\Delta}\rangle_{\text{gkp}}$ GKP states with $\Delta^2 = 0.1$ ($\Delta_{\text{dB}} = 10\,\text{dB}$, for $\Delta = 10^{-\Delta_{\text{dB}}/20}$) with a fidelity of 76% (92%, and 96%) and heralding success probability of 2.1% (0.4% and 0.1%) [16]. We see here a fidelity-probability trade-off for a fixed number of modes. Comparable results are observed for the preparation of finite-energy GKP magic states. As discussed next, these sources can be multiplexed to obtain higher rates of generation, as we detail in Section 4.2.3 for our architecture.

Multiplexing

The multiplexing module boosts the qubit generation rates and, in the event of a qubit generation failure, substitutes in a momentum-squeezed vacuum mode. For a fixed required fidelity, the generation rate for GKP states can be boosted to arbitrary desired probability values $1 - p_0$ by using spatial multiplexing, as illustrated in Fig. 4.3. Multiplexing requires active feed-forwarding of PNR detector outcomes, which can be implemented using 2×2 "crossbar" switches that effect either the identity or SWAP gate. A binary tree of these kinds of switches is sufficient to move a successfully prepared state into the correct output port [192]. In the event that no GBS device successfully produces a GKP state, we include an additional switch at the output of the multiplexing device which can swap in a momentum-squeezed state to replace the output of the multiplexed GBS devices, as shown in Fig. 4.3.

Computational

The computational module implements the deterministic entangling operations, thereby enabling universal and fault-tolerant measurement-based quantum computation. The first step in the generation of the resource states is to create one-dimensional hybrid cluster states that extend in the temporal direction. Recall from

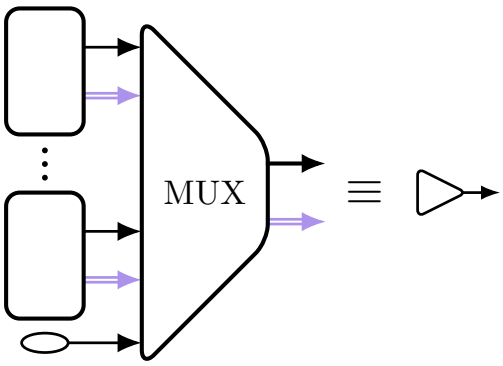


Figure 4.3: **Multiplexed state generation.** Multiplexed GBS devices for increased rate of state preparation. The multiplexer consists of a binary tree of 2×2 switches that either implements an identity or SWAP gate on each optical mode, moving a successfully generated GKP state to the correct output port. If no GBS device produces a GKP state, we swap the output of the multiplexing device for a deterministically generated momentum-squeezed state (depicted by the ellipse on bottom left). The right-hand side shows the simplified diagram for the hybrid quantum light source. Note that the classical information wire is suppressed.

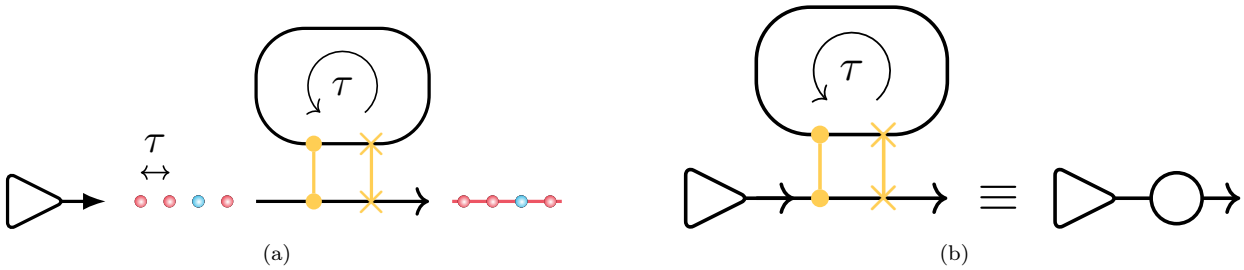


Figure 4.4: **Generating 1D qubit cluster in the time domain**. (a) On the left, a 'GBS factory' comprising multiplexed GBS devices is used to generate the sequence of pulses, where each pulse contains either a GKP $|+\rangle$ state (red dot) or a momentum-squeezed state (blue dot). Each input interacts with the previous input (which is in the loop) via a CZ gate, enters the loop mode via the swap, interacts with the next mode, and then is swapped into the output mode by the same swap. (b) Simplified diagram for 1D time-domain cluster state source.

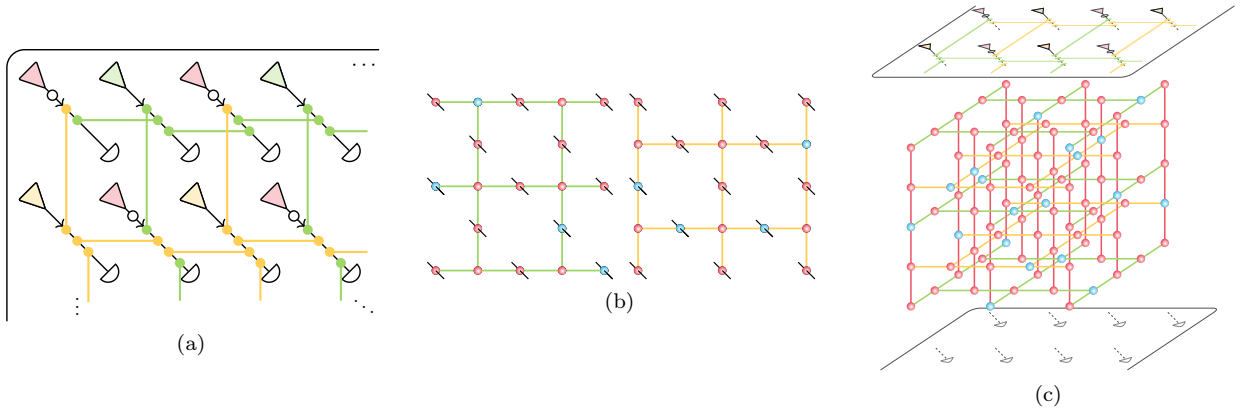


Figure 4.5: **Generating the RHG lattice.** (a) Chip layout to generate the hybrid RHG lattice. Light from three kinds of sources is incident on the chip. The red triangles with circles are sources of hybrid 1D cluster states in the time domain, with a sequence of entangled qubits being emitted at a time delay of τ. The remaining yellow and green triangles are simply qubit sources, but these sources fire only at a time interval of $T = 2\tau$, with yellow sources firing only at the $(2n-1)\tau$ times and the green sources firing at the $(2n)\tau$ times. The lines on the chip represent CZ gates. The yellow CZ gates are turned on only at odd times and the green ones at even times. Together, these qubits and gates generate the different layers of the RHG lattice and the connections between them. (b) A representation of two layers of the RHG lattice. Recall that the dots represent individual computational qubits and the connections between them show the entanglement. Here the two sub-figures represent the even and odd layers of the RHG lattice. (c) A spatio-temporal representation of the hybrid RHG lattice generated by the chip of (a).

Section 4.2.1 that both GKP qubit cluster states and CV cluster states require $CZ = e^{i\hat{q}\otimes\hat{q}}$ gates. Given a single physical CZ gate (implemented via beam-splitters and squeezers as shown in Fig. C.1) and a swap gate, a linear cluster state can be generated in the time domain using optical delay lines [193–195], as depicted in Fig. 4.4a.

Next, additional CZ gates are implemented in the two spatial dimensions to generate the 3D structure of the RHG lattice. Consider a 2D spatial array of 1D time-domain cluster state sources, interspersed by additional state-preparation modules and connected in the spatial domain by a nearest-neighbor array of optical CZ gates, as shown in Fig. 4.5a. These extra state-preparation modules are broken into two sets, indicated by the green and yellow coloring in Fig. 4.5a. Half emit states at even clock cycles, and the other half emit at odd. The CZ gates are also divided into two sets—indicated by green and yellow coloring in Fig. 4.5a—and are applied during even and odd clock cycles, respectively. Thus, the additional spatial connectivity of the lattice for even and odd clock cycles is as shown in Fig. 4.5b. The resulting cluster state has a lattice structure, as shown in Fig. 4.5c in (2+1)-dimensions. After traversing through the CZ gates, all modes are sent to homodyne detectors.

Photonic quantum processing unit (QPU)

The photonic QPU module performs homodyne measurements on the generated resource state in order to perform the required computation. By merely changing the phases of the local oscillator, i.e. the phase space quadrature of the homodyne measurement, the QPU can perform the logical computation, as we reviewed briefly in Section 4.2.2. Output of the homodyne detectors, along with the input state record from the state generation module (whether GKP or squeezed state), and the program instructions (encoding the user's

compiled quantum program), are all fed to a classical controller which determines which quadrature angles to measure, as well as returns decoded data to the user.

Having reviewed the necessary theoretical components for the architecture, along with a feasible method for its implementation, we now construct a noise model for the output state that we can use to determine fault-tolerant thresholds.

4.3 Error Correction for a Quantum Memory

Having laid out the details of our architecture in the previous sections, we move on to describing its operation for quantum computation on logical qubits. The simplest logical computation is the identity logical operation or the *quantum memory*. Here we elucidate the steps required to implement quantum error correction for a quantum memory.

Algorithm 2 Quantum error correction procedure for a quantum memory

1. *Initialization.* Prepare a resource state on N quantum modes corresponding to the nodes of the RHG lattice. With probability $1 - p_0$ and p_0, the state of each node is either a noisy GKP state or a finitely-squeezed momentum eigenstate, respectively. Both node states are characterized by a noise variance parameter δ.
2. *Measurement.* Obtain a list of real-valued outcomes corresponding to p-homodyne measurements on all the modes.
3. *Inner decoder.* Map the real-valued homodyne outcomes to binary qubit measurement outcomes using local and global information via Algorithm 4.
4. *Outer decoder.* Apply qubit decoding techniques for the RHG lattice such as those in Algorithm 5 to obtain a recovery operation which has a corresponding CV implementation.
5. *Error correction.* Perform CV feed-forward operations based on the outcomes obtained and processed in steps 2 and 3. These, combined with the qubit recovery operation obtained in step 4, return the complete CV recovery operation, which can be tracked in software.

At a high level, quantum error correction in the architecture consists of performing homodyne measurements on a subset of nodes of the RHG lattice, followed by processing of the measurement data to output a recovery operation to be applied on the remaining active nodes of the lattice. In our case, the data processing procedure consists of two decoders, the first of which is an inner (CV) decoder that converts the real-valued homodyne measurements into qubit outcomes and probabilities of Z-type qubit-level errors. This information, in turn, is fed into an outer (qubit-level) decoder, which returns an outer recovery operation. As described below, our outer recovery operation can exploit analog information from the inner decoder, resulting in suitable inner recovery operation to be applied on the physical modes of the system.

Thus, the full error correction procedure is specified by the choice of inner decoder (applied to the GKP code) and the outer qubit code (applied to the RHG lattice). We first introduce a noise model for our hybrid lattice in the next subsection. The inner and the outer decoders are tailored to both the noise model and the hybrid GKP/squeezed-state structure of our architecture. The step-wise procedure for implementing the quantum error correction procedure on a quantum memory is overviewed in Algorithm 2.

4.3.1 Error Model

In order to motivate our choice of inner decoder and check its efficacy, we first construct and analyze a simple noise model for our hybrid RHG lattice. This section summarizes the noise model and main conclusions that we draw from it, with full details available in Appendix C.1.

A reasonable model, which is standard in the CV literature, for capturing part of the noise effect of finite-energy GKP states is obtained by the application of a Gaussian noise channel [196]

$$\mathcal{N}_{\boldsymbol{Y}}(\hat{\rho}) = \int_{\mathbb{R}^2} \frac{d^2\boldsymbol{\xi}}{\pi\sqrt{\det \boldsymbol{Y}}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{Y}^{-1}\boldsymbol{\xi}\right] \hat{\mathcal{D}}(\boldsymbol{\xi})\hat{\rho}\hat{\mathcal{D}}(\boldsymbol{\xi})^\dagger \tag{4.3}$$

to the ideal GKP states with noise of variance $\frac{\delta}{2}$ [34, 47] in both quadratures, with $\delta = \Delta^2$ from Eq. (3.20). While this noise model does not capture the peak-damping envelope of Eq. (3.20), it captures the finite width added to each delta-function in phase space. In our case, we find that the same noise model framework can be used to model the replacement of $|+\rangle_{\text{gkp}}$ states with p-squeezed states, setting $\epsilon = \Delta^2 = \delta$ from Eq. (4.2). In particular, we notice that adding Gaussian noise of variance $\delta/2$ ($\frac{1}{2\delta}$) in p (q) quadrature makes the $|+\rangle_{\text{gkp}}$ mimic the Wigner function of a mixture of p-squeezed states. In the context of Eq. (4.3), the noise matrices for GKP and p-squeezed states are given by:

$$\boldsymbol{Y}_{\text{gkp}} = \frac{1}{2}\begin{pmatrix}\delta & 0\\ 0 & \delta\end{pmatrix}, \quad \boldsymbol{Y}_{\text{p}} = \frac{1}{2}\begin{pmatrix}\delta^{-1} & 0\\ 0 & \delta\end{pmatrix}. \tag{4.4}$$

More concretely, the $|+\rangle_{\text{gkp}}$ Wigner function has rows of positive peaks periodically arranged in phase space along even integer multiples of $\sqrt{\pi}$ in the p quadrature, and alternating positive and negative peaks for odd multiples (see Fig. 1a of [36]). The broad distribution in q from $\boldsymbol{Y}_{\text{p}}$ causes the rows of positive and negative peaks to cancel, and the rows of positive-only peaks to add, washing away the Wigner negativity and yielding a distribution mimicking a mixture of p squeezed states spaced by even multiples of $\sqrt{\pi}$ in p. While this is not a true p-squeezed state, we do not expect it to provide an underestimation of the error probability of the quantum memory, especially since a mixture of states (as opposed to a pure p-squeezed state) would only add more noise and hence make the decoding problem more difficult.

Given these two types of initial states, both modelled as GKP states having undergone independent and different Gaussian noise channels, we then model the encoding into the RHG lattice, which simply consists of repeated applications of CZ gates. Propagating the initial state noise through the CZ gates results in a *correlated* Gaussian noise channel, where the correlations depend on the locations of p-squeezed states and on the lattice-dependent pattern of CZ gates applied to the nodes. We assume that the dominant source of noise is the noise in the input states. Additional noise sources include photon loss and noise introduced in CZ gates, which we leave to analyze or improve in future work.

From our model, we can formally write down the distribution of p-homodyne data. Since all the modes are measured in the p-quadrature when the computer is operating as a quantum memory, we can use this model for the distribution to inform our choice of inner decoder. In the case of no initial-state noise, sampling from the distribution of p-homodyne outcomes would simply correspond to sampling a lattice point $\boldsymbol{n}\sqrt{\pi}$ in p-space, where $\boldsymbol{n}$ is dictated by the qubit state of the RHG lattice. However, under the correlated Gaussian noise channel in our model, we find that each lattice point in p-space is converted into a correlated Gaussian distribution centered at the same point with covariance matrix $\widetilde{\boldsymbol{\Sigma}}_p$. Here, $\widetilde{\boldsymbol{\Sigma}}_p$ is the momentum part of the covariance matrix for the Gaussian peaks of the Wigner function in the phase space for the state of our

hybrid lattice, as we show in Appendix C.1. $\widetilde{\mathbf{\Sigma}}_p$ contains the aforementioned correlations and can be used to our advantage in the inner decoder as we show in the next section.

4.3.2 Inner Decoder

As described above, an inner decoder $\mathcal{T}$ is a function that takes real-valued homodyne data and outputs binary data interpreted as qubit measurement outcomes, i.e.,

$$\mathcal{T} : \mathbb{R}^n \to \{0,1\}^n. \tag{4.5}$$

These qubit outcomes can then be combined into stabilizer measurement outcomes and used in the subsequent decoding procedure of the outer code [163]. Additionally, we use our model for noise and the inner decoder strategy to calculate (marginal or correlated) probabilities of qubit error in our readout, which in turn can then be used to inform our outer decoder strategy that we outline in the following subsection. The standard map from homodyne measurement outcomes to qubit measurement outcomes is a binning function derived from the translational symmetry of the original GKP state, i.e., the perfect periodicity in the q and p directions. The $|+\rangle_{\text{gkp}}$ and $|-\rangle_{\text{gkp}}$ states are each $2\sqrt{\pi}$-periodic in momentum but shifted relative to each other by $\sqrt{\pi}$. Therefore we can place the homodyne outcomes into bins of width $\sqrt{\pi}$ that are centred at integer multiples of $\sqrt{\pi}$, associating with $|+\rangle_{\text{gkp}}$ ($|-\rangle_{\text{gkp}}$) the outcomes that fell in bins centered about even (odd) integer multiples of $\sqrt{\pi}$. We refer to this procedure as "standard binning". While this binning procedure uses the original symmetry of the GKP states, it does not account for the correlations in the covariance matrix introduced by the CZ gates and the presence of p-squeezed states, as described in the error model.

As a key proof-of-concept improvement to illustrate the importance of taking correlations into account, consider the example of a momentum-squeezed state at the centre of a primal face of the RHG lattice, which we denote as node 0, surrounded by four neighboring GKP states on nodes 1–4. For simplicity, in this example we assume that all the continuous-variable CZ gates are the same, but this trivially generalizes if the signs of the CZ gates change. The joint quadrature $p_0 + \sum_{j=1}^{4} q_j$ has a large variance on the order of $\frac{1}{2\delta}$. Without using the correlations, the naïve inner decoder described above would result in a high-strength dephasing channel on the four neighboring GKP qubits, since the marginal distributions along p_j would be broadened by $\frac{1}{2\delta}$ and standard binning does not leverage correlations between nodes. On the other hand, by taking correlations into account, the high covariance along the joint quadrature will result in either the identity gate, or a correlated four-body ring of Z operators on the neighboring qubits, which acts trivially on the code space.

More explicitly, consider the binning strategy that makes use of the correlations between optical modes. The momentum part of the noise matrix resulting from the application of the CZ gates is

$$\begin{aligned}\widetilde{\mathbf{\Sigma}}_p &= \frac{1}{2}\begin{pmatrix} 5\delta & 0 & 0 & 0 & 0 \\ 0 & \delta+\delta^{-1} & \delta^{-1} & \delta^{-1} & \delta^{-1} \\ 0 & \delta^{-1} & \delta+\delta^{-1} & \delta^{-1} & \delta^{-1} \\ 0 & \delta^{-1} & \delta^{-1} & \delta+\delta^{-1} & \delta^{-1} \\ 0 & \delta^{-1} & \delta^{-1} & \delta^{-1} & \delta+\delta^{-1} \end{pmatrix}, \\ &= \frac{1}{2}\left[5\delta \oplus \left(\delta \mathbb{1}_4 + \delta^{-1} |\boldsymbol{\nu}\rangle\langle\boldsymbol{\nu}|\right)\right], \boldsymbol{\nu} = (1,1,1,1)^T, \end{aligned} \tag{4.6}$$

where we label the modes $0, 1, \ldots, 4$ with the momentum state corresponding to mode 0. We see that the noise matrix is non-diagonal, i.e., the CV noise is correlated, but it has a specific structure that can be

exploited. Two immediate observations are that mode 0 is uncorrelated from the other modes, meaning we can simply apply standard binning to it; and that there is correlated noise along the direction $(0,1,1,1,1)$ in p-space. Algorithm 3 presents a strategy for dealing with this correlated noise, taking into account consistency checks that our guesses for modes 1–4 must respect.

Algorithm 3 Inner decoder applied to 5-modes: a p-squeezed state surrounded by 4 GKP states

Input: Vector $\boldsymbol{p} = (p_0, ..., p_4)$ of homodyne measurement outcomes, with $p_i \in \mathbb{R}$. Mode 0 is the p-squeezed state, rest are GKPs.

1. Apply the following change-of-basis $\boldsymbol{T}\boldsymbol{p} = \boldsymbol{p}'$ where:

$$\boldsymbol{T} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -1 & -1 \\ 0 & 1 & -1 & 1 & -1 \\ 0 & 1 & -1 & -1 & 1 \end{pmatrix}.$$

   We note the column vectors of this transformation are eigenvectors of $\widetilde{\boldsymbol{\Sigma}}_p$ in this case.

2. Bin the first component of $\boldsymbol{p}'$ to the nearest integer multiple of $\sqrt{\pi}$ to return $n_0'\sqrt{\pi}$, since the p quadrature outcome of mode 0 is uncorrelated from the others.

3. Of the last three components of $\boldsymbol{p}'$, find the component i that is closest to an integer multiple n_i' of $\sqrt{\pi}$. Round $\boldsymbol{p}_i'$ to $n_i'\sqrt{\pi}$. We only choose the last three components since we do not trust the second component which corresponds to homodyne results along $(0,1,1,1,1)$ which has excessive noise of order $\frac{1}{2\delta}$.

4. If n_i' is even (odd), round the remaining two components other than $\boldsymbol{p}_0'$, $\boldsymbol{p}_1'$ and $\boldsymbol{p}_i'$ to the nearest even (odd) integer multiples of $\sqrt{\pi}$ for each component. This yields $\sqrt{\pi}\boldsymbol{v}' = \sqrt{\pi}(n_0', \boldsymbol{p}_1'/\sqrt{\pi}, n_2', n_3', n_4')$, because on applying the change of basis $\boldsymbol{T}$ to an integer vector, the last four components of the new vector should either all be even or all odd.

5. If $(n_2' + n_3' + n_4') \bmod 4 = 0, 1, 2, 3$, then round $\boldsymbol{p}_1'$ to the nearest $n_1'\sqrt{\pi}$ with the constraint that $n_1' \bmod 4 = 0, 3, 2, 1$. This yields $\sqrt{\pi}\boldsymbol{n}' = \sqrt{\pi}(n_0', n_1', n_2', n_3', n_4')$. Again, this is because on applying the change of basis $\boldsymbol{T}$ to an integer vector, the second component and the last three components respect this rule, so this guess should respect it too.

6. Undo the change of basis on the integer-valued vector $\boldsymbol{T}^{-1}\boldsymbol{n}' = \boldsymbol{n}$.

7. Take $\boldsymbol{n} \bmod 2 = \boldsymbol{s}$ to be the five-component binary string output.

Output: 5-qubit measurement values $\boldsymbol{s}$.

In general, the problem of finding a better inner decoder for our hybrid architecture is to find a decoder that takes into account the location of GKP and p-squeezed states, and knowledge of the structured CZ gates that have been applied to form the cluster state.

The distribution of p-homodyne outcomes consists of a periodic arrangement of Gaussian distributions all with covariance $\widetilde{\boldsymbol{\Sigma}}_p$ on N modes, each Gaussian centred at a point $\boldsymbol{n}\sqrt{\pi}$ where $\boldsymbol{n}$ are integer valued vectors from a set that corresponds to the ideal state of the qubits. Suppose we obtain the values $\boldsymbol{p}$ after the homodyne measurements. If we assume $\boldsymbol{p}$ could have resulted from a Gaussian distribution centered at any of the lattice points $\boldsymbol{n}$, then the so-called *responsibility* [197] of a given lattice point for the result $\boldsymbol{p}$ is given by:

$$r(\boldsymbol{n}) = \exp\left[-\frac{1}{2}(\boldsymbol{n}\sqrt{\pi} - \boldsymbol{p})^T \widetilde{\boldsymbol{\Sigma}}_p^{-1}(\boldsymbol{n}\sqrt{\pi} - \boldsymbol{p})\right]. \tag{4.7}$$

The responsibility is directly related to the Gaussian distributions at each lattice point and provides a relative way of ordering which lattice points were most likely to have generated $\boldsymbol{p}$. Specifically, the lattice point which was most likely to have produced the point $\boldsymbol{p}$ is:

$$\boldsymbol{n}_{\mathrm{IQP}} = \arg\min_{\boldsymbol{n}\in\mathbb{Z}^N} (\boldsymbol{n}\sqrt{\pi} - \boldsymbol{p})^T \tilde{\boldsymbol{\Sigma}}_p^{-1} (\boldsymbol{n}\sqrt{\pi} - \boldsymbol{p}), \tag{4.8}$$

where we have chosen the subscript IQP to indicate that this is an integer quadratic program, i.e., a minimization of a quadratic function over an integer domain. As mentioned above and for simplicity, we are using the standard approximation that all peaks in the GKP state have equal weight [34]. However, one could also include an envelope that weights peaks differently, in which case this information could also be included in the calculation of the responsibility. In general, integer quadratic programs are NP-hard [198, 199], so we will require a heuristic strategy that is computationally tractable. Our approach for a generalized version of Algorithm 3 is summarized in Algorithm 4, with the case of more complicated configurations of p-squeezed states left to future study.

Algorithm 4 Inner decoder

Input: Vector $\boldsymbol{p} = (p_0, ..., p_N)$ of homodyne measurement outcomes, with $p_i \in \mathbb{R}$, and the noise model.

1. Identify directions that are noisy and those that are not using the noise matrix.
2. Perform a suitable change of basis to the homodyne data to obtain CV results for joint quadratures, a smaller number of which have reduced noise. In particular, an integer-valued transformation would allow for certain consistency checks (e.g. parity) when making a guess for the p-space lattice point $\boldsymbol{n}$.
3. Apply binning along the new directions to round results to nearest ideal peak position, taking into account self-consistency of the results.
4. Undo the change of basis to return a candidate lattice point $\boldsymbol{n}\sqrt{\pi}$.
5. Obtain a binary string by taking $\boldsymbol{n}$ mod 2.

Output: Interpreted qubit measurement outcome

4.3.3 Outer Decoder and Error Correction

After obtaining and binning the outcomes of the homodyne measurement, error correction is performed for the outer qubit code. The details of the error correction problem we solve are summarized in Algorithm 5 for a particular, standard, choice of decoding algorithm: minimum-weight perfect matching (MWPM) [162, 200, 201].

A few comments are in order. The weights of the matching graph edges in Algorithm 5 are derived from the homodyne measurement outcomes, as well as the positions of the p-squeezed states in the lattice. An example of such weights is presented in Section 4.4. Furthermore, using the homodyne measurement outcomes to calculate matching graph weights has been explored in the context of the toric code [39, 47], but the knowledge of the locations of the p-squeezed states gives us additional information that can be used to improve the performance of the decoder. We discuss this point in more detail in Section 4.4.

As mentioned earlier, due to the measurement-based computation model, feed-forward operations based on the outcomes obtained from the homodyne measurements and the inner decoder are combined with the qubit-recovery operation obtained from the outer decoder. Together, these inform the complete CV recovery

operation that needs to be applied to the active computation layers. In practice, the combined recovery operation need not actually be applied on the qubits; instead, we would keep track of the recovery operations in classical control programming by updating the Pauli frame [72, 184].

Algorithm 5 Outer decoder using MWPM

Input: Qubit measurement outcome from Algorithm 4

1. *Syndrome identification.* Construct relevant stabilizer measurement outcomes from the input qubit outcomes.
2. *Matching graph construction.* Construct a complete graph using:
   - Vertices: One vertex for each unsatisfied stabilizer (include additional vertices if needed for specific boundary conditions.)
   - Edges: Connect every pair of vertices
   - Weights: The edges are assigned weights reflecting probability of the most likely error that could have given rise to the pairs of unsatisfied stabilizers. The choice of weights can be informed by the CV noise model.
3. *Matching algorithm.* Find a minimum-weight perfect matching by running Edmonds' algorithm [202] on the matching graph from the previous step.
4. *Qubit recovery operator.* The recovery operator is given by this rule: for each pair (u, v) in the matching, flip the binary outcomes of qubits in the most likely error that could have given rise to u and v.

Output: Qubit recovery operator

4.4 Threshold Estimation for a Quantum Memory

The operation of the architecture as a fault-tolerant quantum computer is underpinned by the concept that the logical error rate of an encoded computation can be arbitrarily lowered by increasing the size of the code. This concept is based on the idea of fault-tolerance thresholds. The existence of such thresholds for qubit-based architectures has been a subject of extensive research for over twenty years [177, 178, 200, 203–212] but the existence of thresholds for CV-based architectures [34] is less well understood. Furthermore, the question of whether hybrid architectures remain fault-tolerant with *probabilistic* sources of GKP qubits is not obvious. Here we provide numerical evidence that our architecture does indeed have a threshold in the presence of errors arising from finite squeezing and for a range of swap-out probabilities. As we detail in this section, in order to calculate the threshold, we simulate the hybrid architecture operating as a quantum memory and run a complete error-correction procedure [213]. We detail the various steps involved in the simulation of the thresholds in Algorithm 6.

We now briefly review the numerical procedure for estimating the error threshold of a quantum memory. Consider a family of codes of growing size, parameterized by d. In the case of the RHG lattice, d is the code distance (the weight of the minimal weight non-trivial logical operator) and the number of qubits is $n = O(d^3)$. Another parameter is the noise channel, which in our case is described by two numerical parameters: the noise variance δ and the swap-out probability p_0. To estimate the error threshold, we run many trials of Monte Carlo simulations to determine the logical error rates as a function of our physical noise parameters. This is done for different lattice sizes d. In each trial, we generate homodyne measurement outcomes according to the noise parameters, then we run our error correction procedure (inner decoder followed by outer decoder

Algorithm 6 Procedure for obtaining thresholds for a quantum memory

1. *Parameters.* Choose lattice size d, swap-out probability p_0, and noise variance δ
2. *Simulated homodyne measurements.* Generate homodyne measurement outcomes consistent with the noise matrix as given in Eq. (C.20) through a suitable sampling method.
3. *Inner decoder.* Apply the inner decoder of Algorithm 4 on the homodyne data to obtain qubit measurement outcomes.
4. *Outer decoder.* Apply the outer decoder to the qubit outcomes to obtain a recovery operation using Algorithm 5
5. *Error correction.* Apply the recovery operation.
6. *Success check.* Note the success/failure of the error correction procedure.
7. *Error rate.* Repeat steps 2-6 sufficient number of times to obtain an error rate.
8. *Thresholds*. Repeat steps 1-7 for different lattice sizes and noise parameters (p_0, δ).

as described in Section 4.3), and finally check if error correction has been successful. Let us assume that we fix p_0 and vary δ. Then if a threshold, δ_c, exists, we expect to see the following behaviour. For $\delta > \delta_c$, increasing the size of the code (increasing d) increases the logical error rate. But for $\delta < \delta_c$, increasing the size of the code exponentially decreases the logical error rate. We note that the largest code sizes we consider involve $n \approx 5000$ qubits. Simulation of such a large number of qubits is possible due to the fact that we use a classical noise channel to model approximate GKPs and the circuits we simulate belong to the Clifford group, which makes them efficiently classically simulable [137].

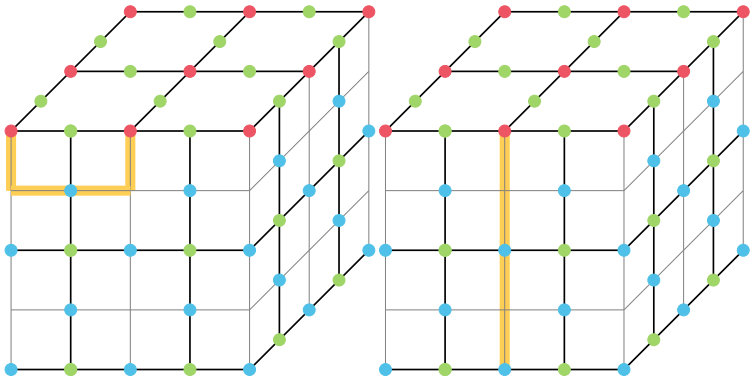
While the other steps of Algorithm 6 are relatively straightforward, we explain the success-check step of Algorithm 6. After applying the recovery operation, all the cluster state stabilizers are guaranteed to be satisfied. Therefore, error correction is successful if the product of the qubit error and the recovery operator is a stabilizer (logical identity operator) and error correction fails if the product of the qubit error and the recovery operator is a non-trivial logical operator. Such operators anti-commute with at least one of the correlation surfaces of the cluster state [161]. Fig. 4.6 shows the X correlation surface (the Z correlation surface is analogous). To summarize, if the product of the qubit error and the recovery operator anti-commutes with either correlation surface, then error correction has failed.

The remainder of this section is structured as follows. In Section 4.4.1, we describe our simulations in detail and compare the performance of different inner and outer decoding strategies. Then, in Section 4.4.2, we present the threshold simulation results for our architecture operating as a quantum memory.

4.4.1 Simulation Details

Here we provide some details on the simulations performed to find the thresholds. First, we note that we only simulate error correction of the primal lattice nodes, as the error correction problem for the dual lattice nodes is the same and each problem can be solved independently. We consider RHG lattices with size parameterized by d, where the left and right boundaries are equivalent to distance d surface codes, and there are d layers of nodes in between these two boundaries (see Fig. 4.6).

We now return to the calculation of the matching graph weights (Step 2 in Algorithm 5). The first step is to construct a decoding graph based on the RHG lattice. For the sake of brevity, we will describe this



(a) (b)

Figure 4.6: X correlation surface in the RHG lattice. The blue circles are the primal qubits, the green circles are the dual qubits, and the pink circles are the primal qubits in the X correlation surface. The yellow highlighted edges represent Z operators, i.e. primal qubits on yellow highlighted edges have a Pauli Z applied to them. In a) we show a logical identity operator that commutes with all the stabilizers and the correlation surface, whereas in b) we show a non-trivial logical operator that commutes with all the stabilizers but does not commute with the correlation surface.

construction for an RHG lattice with periodic boundary conditions (see [214] for the case of lattices with boundaries). We refer to the elements of the decoding graph as vertices and arcs, to avoid confusion with the nodes and edges of the RHG lattice. The decoding graph has a vertex for each six-body X stabilizer acting on the primal qubits of the RHG lattice. These stabilizers are formed from products of cluster state stabilizers surrounding a dual cell. Vertices are connected by arcs if their corresponding stabilizers share a qubit. As each qubit is in the support of two such stabilizers, the arcs of the decoding graph are in a one-to-one correspondence with the primal qubits of the RHG lattice, and hence with a subset of the modes of the cluster state. We assign weights to the arcs of the decoding graph as follows. Consider the mode q corresponding to an arc in the decoding graph. Let m be the number of swapped-out modes neighboring q and let z be the outcome of the homodyne measurement of q. We assign to this mode a heuristic error probability as follows:

$$w(z,m,\tilde{\delta}) = \begin{cases} 2/5 & \text{if } m=4, \\ 1/3 & \text{if } m=3, \\ 1/4 & \text{if } m=2, \\ \frac{\sum_{n\in\mathbb{Z}} \exp[-(z-(2n+1)\sqrt{\pi})^2/\tilde{\delta}]}{\sum_{n\in\mathbb{Z}} \exp[-(z-n\sqrt{\pi})^2/\tilde{\delta}]} & \text{if } m\leq 1. \end{cases} \tag{4.9}$$

If a mode has one swapped-out neighbor, then there are no errors due to swap-outs as the net effect of a single swap-out after applying the CZ gates is a stabilizer. In this case, the error probability is the probability of incorrectly binning the state [47], using the standard binning function and assuming a classical noise channel with parameter $\tilde{\delta}$, which we derive from $\tilde{\Sigma}_p$. For $m \geq 2$, we derive the weights in Eq. (4.9) from simulations which we detail in Appendix C.3. The weight of the corresponding arc in the decoding graph is then $-\log w(z,m,\tilde{\delta})$ [201].

Given the decoding graph, we construct the matching graph weights as follows. For each pair of vertices in the matching graph, we compute the total weight of the minimum weight path between the corresponding vertices in the decoding graph using Dijkstra's algorithm [215].

Many variants are possible for the inner decoder introduced in Section 4.3.2. In this chapter, we considered

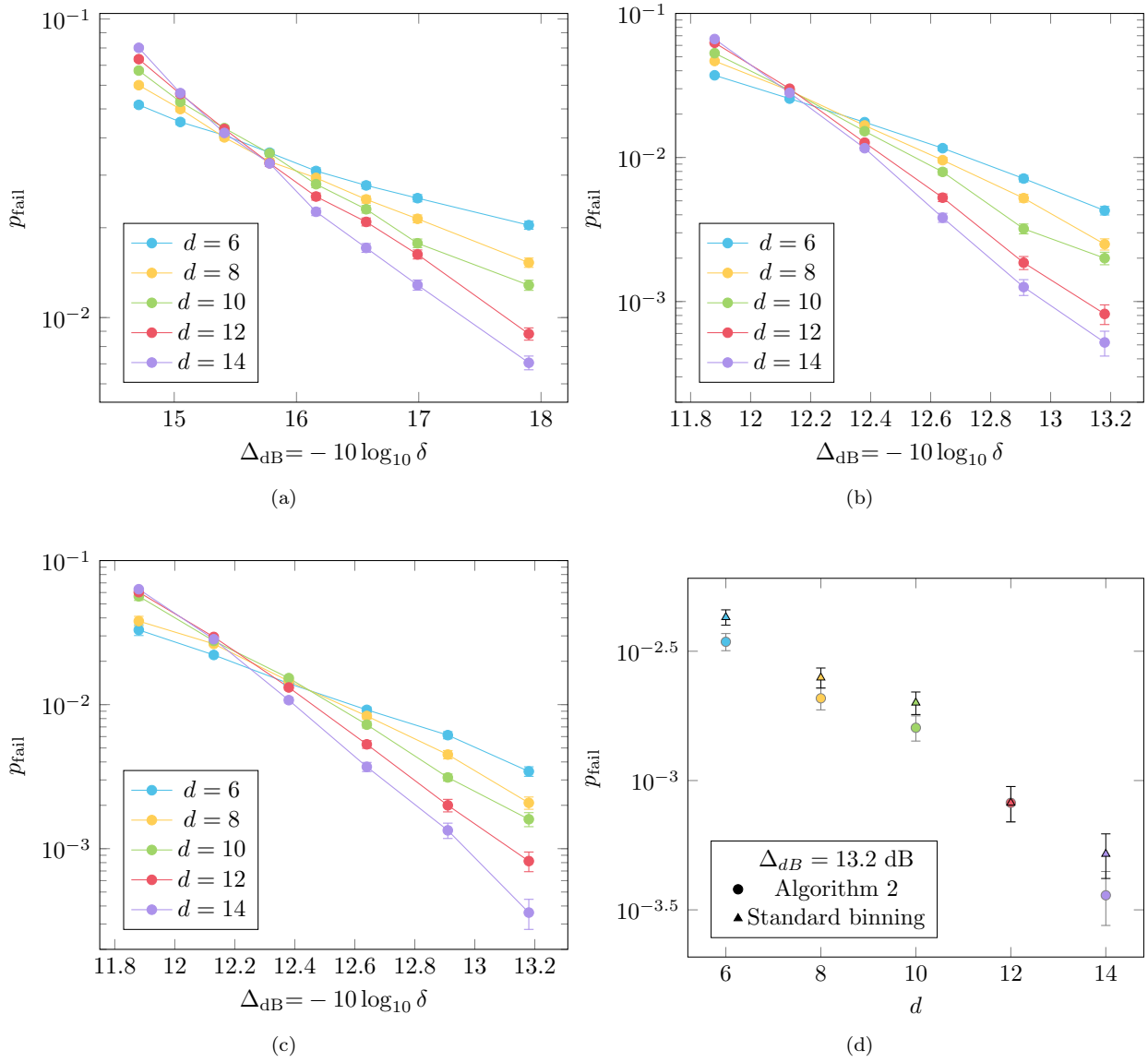


Figure 4.7: Performance comparison for the various inner decoders considered for $p_0 = 0.06$. For (a) and (b), standard binning is used as the inner decoder for every node, and the weights assigned to the edges of the matching graph are either (a) all equal, or (b) assigned following Eq. (4.9). In (c), Algorithm 3 is first used on GKP nodes connected to isolated momentum states, standard binning is used for the remainders, and weights described by Eq. (4.9) are used in the matching graph. (d) A comparison of the failure probability as a function of code distance d, for $\Delta_{\text{dB}} = 13.2$ dB using two different inner decoders. At this point below threshold, we see that Algorithm 3 provides equal or lower error rates than standard binning.

two simple ones. The first is performing standard binning of the homodyne outcomes, irrespective of the presence of momentum-squeezed state in its vicinity. Second, for those momentum-squeezed states which are isolated from others, in the sense that no connected node is also connected to another squeezed state, a variant of Algorithm 3 is used. The modifications are required because of the variable number of neighbors and signs present in the physical application of the CZ gates. We emphasize that, as mentioned in Section 4.3.2, more complex strategies can be devised and are likely to improve the overall decoding performance.

Simulation results for both possibilities are shown in Fig. 4.7, for $p_0 = 0.06$. Fig. 4.7 (a) shows results for naïve binning using uniform weights in the matching graphs, while (b) uses weights as described in Eq. (4.9). In (c), Algorithm 3 is used, and weights are given by Eq. (4.9); we chose $p_0 = 0.06$ as a representative example to test-drive Algorithm 3, as it is best suited to cases of isolated swap-outs, which are common for this value of p_0. Incorporating the analog information into building the matching graph clearly improves the performance, the threshold decreasing from $\sim 15.5\,\mathrm{dB}$ to $\sim 12.2\,\mathrm{dB}$, with both variants of the inner decoder. We note that modifying the inner decoder to leverage Algorithm 3 did not result in any significant differences for the thresholds themselves but the failure rates below threshold are equal or lower using Algorithm 3, as we show in Fig. 4.7 (d). Quantifying and understanding the origin of this effect is left for future work.

4.4.2 Threshold Results

Now we are ready to present the thresholds of our hybrid architecture. Our first result is the error threshold of the RHG-GKP code with approximate GKP states, which we model as ideal states suffering a random displacement with noise variance δ, as discussed in Section 4.3.1. In our noise model, this corresponds to the limit of no swap-outs, i.e., $p_0 = 0$. Similar simulations have been carried out in previous works for the toric-GKP code [39] and the surface-GKP code [47, 48]. We use standard binning and matching graph weights derived from Eq. (4.9). We observe an error threshold of $\Delta_{\mathrm{dB}} = -10\log_{10}(\delta) \approx 10.5$ dB, which is comparable with results for similar noise models in the aforementioned works. The data are shown in Fig. 4.8.

As described above, the full noise model we use involves two noise parameters, the noise variance δ and the swap-out probability p_0; the error threshold is a line in (δ, p_0) parameter space rather than a single point. To estimate this error threshold, we run Monte Carlo simulations as described in Algorithm 6 for different values of δ, p_0 and d (the lattice size). For a particular value of p_0, we can extract the corresponding threshold δ value by plotting the logical error probabilities, p_{fail}, for a range of values of δ and d. The error threshold is then the point where curves for different d intersect. Equivalently, we can instead fix a value of δ and vary p_0 and d. In the inner decoder we use standard binning, and we use matching graph weights derived from Eq. (4.9) in the outer decoder. Fig. 4.9 shows the below-threshold region in (δ, p_0) parameter space, alongside an example threshold plot for $p_0 = 0.1$. We find a high tolerance to swap-outs, with a maximum swap-out threshold of $p_0 \approx 0.236$ (for $\delta = 0$). For $p_0 = 0$, the noise variance error threshold is $\approx 10.5\,\mathrm{dB}$, where the dB value is given by $-10\log_{10}\delta$. As expected, an increase in the swap-out probability leads to an increase in the squeezing thresholds. For an experimentally accessible [87] squeezing value of 15 dB , our simulations suggest a swap-out threshold of $p_0 \approx 0.133$. We note that the noise variance (δ) tolerance of our decoder is markedly better for $p_0 \lesssim 0.19$ than for values nearer the swap-out threshold. Understanding this behaviour is an open problem, with one possible reason being that the inner and outer decoders we are using for the current simulations might be sub-optimal for this regime. Therefore, to investigate this phenomenon further, we should compare our decoding strategy with e.g. maximum-likelihood decoding, in order to ascertain whether the sharp decrease in performance is a fundamental property or an artifact of our decoding strategy. We leave this analysis for future work.

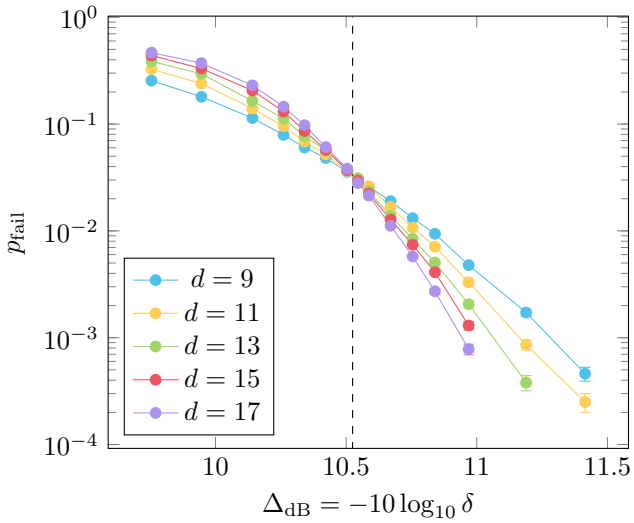


Figure 4.8: Error threshold of the RHG-GKP code with approximate GKP states using standard binning and matching graph weights derived from Eq. (4.9). We estimate the logical error rate p_{fail} using Monte Carlo simulations for different lattice sizes d and noise variance δ. The error threshold is the point where the curve for different values of d intersect. Each data point in the average of $\eta \geq 10^4$ trials and has at least 25 failure events. The error bars show the standard error of the mean $\sqrt{p_{\text{fail}}(1-p_{\text{fail}})/\eta}$. We use the fitting procedure described in [216] to systematically obtain our threshold estimate (dashed vertical line).

Previous works [165–167] have studied the error threshold of the RHG cluster state model when qubits are erased with some probability. This is a natural noise model in optics and bears some resemblance to our model, as one assumes that the locations of the erasures are known. The relationship between the erasure threshold and the Z error threshold was found to be approximately linear [165] and there is a fundamental erasure threshold of 0.249, which is set by percolation theory [217]. It is difficult to directly compare our results with those of [165] because of the differences in the noise models. However, our swap-out threshold is close to the percolation theory erasure threshold, and it is natural to ask whether we can increase the swap-out threshold beyond the erasure threshold by further optimizing our decoder. There are many ways we could improve our current decoder (see Section 4.5), so, unless there is a fundamental limit due to percolation of swap-outs, we are hopeful that we can surpass the erasure threshold. In addition, the question as to whether our decoder has an advantage over the equivalent erasure decoder for finite values of Δ_{dB} remains open and could be a subject of future work.

4.5 Summary and Outlook

We have proposed a concrete and scalable architecture for quantum computing with light. By using a hybrid resource state composed of GKP and squeezed states that can be generated and manipulated using near-future photonic technology, our architecture synthesizes modern techniques in scalable entangled resource state generation and bosonic codes. This "best of both worlds" hybrid approach comes with a novel error structure that arises from the Gaussian model of state imperfections and the use of probabilistic bosonic qubit sources. Numerical results show that such errors can be handled by our tailored two-tier decoder that makes use of continuous- and discrete-variable syndrome data. We find that fault-tolerant quantum computation is possible in the regime where the swap-out probability – the likelihood that any given bosonic qubit source

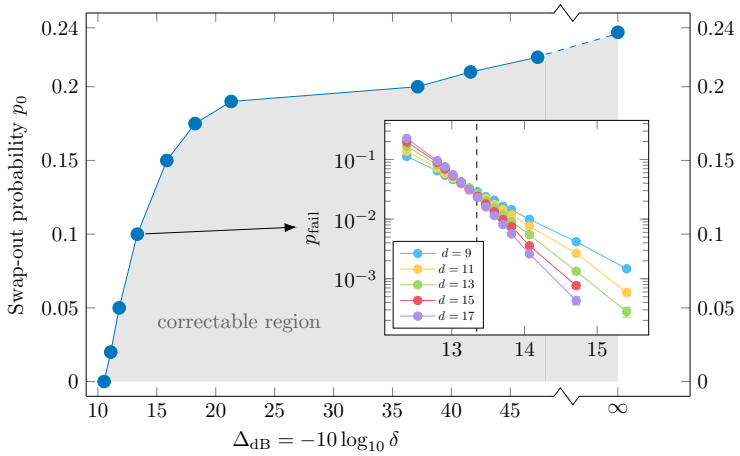


Figure 4.9: Error threshold estimates for a quantum memory using standard binning and matching graph weights derived from Eq. (4.9). Each blue point represents a threshold plot, where we fix either p_0 or δ and use Monte Carlo simulations to find the threshold value of the other noise parameter. The inset shows the threshold plot for $p_0 = 0.1$, with the corresponding threshold $\Delta_{\mathrm{dB}} \approx 13.3$ shown as a dashed line. Pairs of parameters below the blue line lie in the correctable region of parameter space where error correction works.

failed and the input was swapped with a squeezed state – is smaller than $\approx 23.6\%$. In the remainder of this section, we discuss some areas of improvement and open problems for the architecture, which we group into hardware and encoding/decoding improvements.

4.5.1 Hardware improvements

From the hardware perspective, one of the downsides to the current architecture is the CZ gates in the computational module, since these are active transformations, i.e., they require in-line squeezing and displacements, so they can be experimentally challenging to implement. In [46], an updated version of this architecture is presented; there, active transformations are replaced by passive transformations (i.e. beamsplitters), simplifying the computational module and reducing the experimental requirements. This change in the computational module even comes with a *benefit* to the depth of the multiplexing module and to the fault-tolerant thresholds. The simplicity of the updated architecture also allows one to analyze errors due to loss, a source of noise in photonics which was not directly treated in the analysis of the architecture presented here.

Another area of required hardware improvement is increasing the level of squeezing for on-chip sources. A 15 dB level of squeezing has been attained in free-space implementations [87], while the current state-of-the-art level of squeezing demonstrated in integrated sources is 8 dB [218]; it remains an open experimental problem to match the level of squeezing in integrated sources to the free-space record.

Another hardware challenge is that of reliable state preparation. In particular, a key experimental goal highlighted by this work is to gain access to high-quality GKP qubits whose teeth are squeezed at or exceeding the 15 dB level. This motivates the improvement of existing methods for state generation with GBS devices, which previously considered states with up to 10 dB of per-peak squeezing [16, 17, 20, 219]. Reaching 15 dB states with the techniques from Refs. [16, 17, 20, 219] will involve going to higher-order truncations of the Fock basis, which is computationally more demanding. This motivates using the analysis method from [21], or the application of breeding or distillation protocols to lower quality states to reach the 15 dB level.

Progress in these directions is critical to accurate resource estimates for photonic quantum computing.

4.5.2 Encoding/decoding improvements

While the current work focuses on the RHG lattice as a paradigmatic example of an outer qubit code, significant benefit can be expected by moving to other outer encodings. In particular, as we learn more about the structure of noise in realistic GKP-based computation, this opens the possibility of devising better outer encodings that are tailored for the specific noise structure. Furthermore, noise-tailoring along the lines of [48] may provide substantial enhancement to the thresholds obtained in our hybrid architecture.

As for decoding improvements, one question is about the possibility of obtaining a further advantage from accounting for real-valued homodyne outcomes. Although our inner decoder is exploiting the structure of the CV noise, there is still more information that could in principle be exploited, for example, at the level of the outer decoder. It may be possible to use ideas from analog quantum error correction [38] and maximum-likelihood decoding [39] to further reduce our squeezing thresholds.

The question of optimal methods for decoding is also closely related to more fundamental questions related to swap-out-based resource states. Specifically, what is the fundamental swap-out threshold of our architecture? An alternative to swap-outs (i.e., replacing the GKP-node no-shows with squeezed light modes) is to treat the no-shows directly as erasure errors, but for these the threshold is set by percolation theory to be around 24.9%. Is the swap-out threshold higher than the erasure threshold set by percolation theory [165, 217] and in which regions of (δ, p_0) parameter space is it beneficial to have swap-outs rather than erasures? With the current decoders, we obtained around 23.6% swap-out threshold, which likely can be improved substantially as numerous upgrades can be made to our inner and outer decoders. We expect that development of an inner decoder that can treat more complicated arrangements of p-squeezed states in the lattice will provide fewer errors in the readout of qubit outcomes. Furthermore, we expect improvements to the outer decoder by using a more sophisticated method for assigning weights that takes the structure of the inner decoder into account.

Part II

Quantum Key Distribution (QKD)

Chapter 5

Entanglement-Based High-Dimensional QKD and the Measurement Range Problem

This chapter is based on [23], co-authored with Prof. Hoi-Kwong Lo, who initiated and supervised the project. I was the first author for this work. This work was selected as Editor's Pick in the feature issue of *Journal of the Optical Society of America B*, "Quantum Key Distribution and Beyond" [220]. My main contributions were to the literature review, mathematical proofs, numerical simulations, analysis, and manuscript writing. The work benefited from helpful discussions with Prof. Li Qian, Prof. Charles Ci Wen Lim, Dr. Bing Qi, Prof. Zheshen Zhang, Prof. Norbert Lütkenhaus, Aaron Goldberg, and Ilan Tzitrin.

5.1 Introduction

In the quest to encrypt larger amounts of data over untrusted channels, there has been recent interest in creating more efficient QKD strategies that employ high-dimensional photonic degrees of freedom to maximize secret bits per optical signal. In particular, time-energy [221–232], orbital angular momentum [233], and electric field quadrature [234–237] are all candidates for high-dimensional degrees of freedom.

To prove the security of high-dimensional QKD protocols, entropic uncertainty relations have proven to be powerful tools, as they provide a method for quantitatively relating the statistics of measurement outcomes, the operators characterizing those measurements, and the amount of information different parties can have about a system [12, 238–244], allowing Alice and Bob to bound the information held by an eavesdropper, Eve [223, 230, 235, 245–247]. In particular, for an entanglement-based protocol, the optical source and channels may not be trusted, so Alice and Bob must rely on the characterization of their measurements and the statistics of their outcomes to determine security [248].

High-dimensional photonic degrees of freedom are, in principle, unbounded, while any practical detector for measuring them only has a finite range of detection, so a natural question has been whether the potential for the state to fall beyond the range of detection poses any serious consequences for the security of a protocol [222, 223, 249–253]. Qi first noted the potential for a detection range loophole in time-frequency QKD [222], with Nunn *et. al.* outlining a specific strategy for exploiting the loophole [223]. In the context of both time-

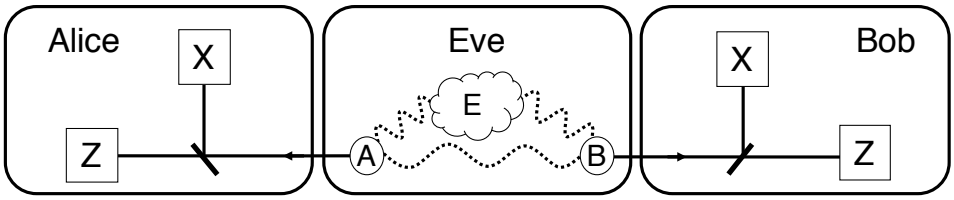


Figure 5.1: A typical scenario for the entropic uncertainty relations with quantum memory. Alice and Bob are given quantum systems, A and B, from a source controlled by Eve, who has quantum memory, E. Alice and Bob independently and randomly either make X-type or Z-type measurements.

frequency QKD and more general continuous variable entanglement verification and quantum cryptographic protocols, Ray and van Enk discussed how data falling outside the measurement range compromises variance-based measures of entanglement, and demonstrated that Rényi entropies without quantum memory provide more optimistic bounds for verifying entanglement [249, 250]. The issues arising from finite measurement ranges have additionally been discussed in [235, 252, 253], in the context of homodyne-based continuous variable (CV) QKD[1], with the main problem being caused by detections above the saturation limit of the detector. A major review of CV QKD is available in Section VI of [106].

In Section 5.2, we review the entropic uncertainty relations with quantum memory, as well as how results outside the measurement range can render the relationship trivial. In Section 5.3, we consider a more general problem: given a measurement outcome one would like to safely ignore, such as the state falling outside the measurement range, we formulate a non-trivial entropic uncertainty relation with quantum memory that depends on the probability of the problematic outcome, rather than on the operators characterizing the problematic outcome. This bound is particularly important for entanglement-based high-dimensional quantum cryptography protocols. In Section 5.4, we discuss the modified bound in the context of time-frequency QKD, and find that additional assumptions are required to deal with practical limitations like loss and vacuum components. Even with some additional assumptions about the source and our modified result, channel loss severely limits the secure key rate of the protocol. Finally, in Section 5.5, we discuss the applicability of our main result to entanglement-based CV QKD using homodyne detection. We find that, since being outside the measurement range corresponds to saturation, not loss, our bound produces the same results as existing protocols with the added benefit that it provides a quantitative way to guard against saturation attacks.

5.2 Entropic Uncertainty Relations and the Measurement Range Problem

We begin this section by summarizing the unmodified entropic uncertainty relations, (a thorough review of entropic uncertainty relations and their applications can be found in [12]). We consider a source, potentially controlled by Eve, that distributes quantum systems A and B to Alice and Bob, respectively, so that the total purified state is ρ_{ABE}. Alice either performs a Z-type measurement, characterized by the positive operator-valued measure (POVM), $Z = \{\mathbb{Z}_A^z\}_z$, or an X-type measurement, characterized by the POVM, $X = \{\mathbb{X}_A^x\}_x$. Bob also alternates between the same two types of measurements. See Fig. 5.1 for an illustration of the set-up.

The most information Eve can have about Alice's Z-type measurement results is quantified by the

[1]CV QKD commonly refers to protocols employing electric field quadratures.

conditional min-entropy of Alice's classical register, Z_A, used for storing results from Z-type measurements, given Eve's quantum memory, E:

$$H_{\min}(Z_A|E) = -\log p_{\text{guess}}(Z_A|E) \tag{5.1}$$

where $p_{\text{guess}}(Z_A|E) = \max_{\mathbb{M}_E} \sum_z p^z_{Z_A} \text{Tr}(\mathbb{M}^z_E \rho^z_E)$ is Eve's guessing probability of Z_A maximized over all POVMs on E, $\mathbb{M}_E$ [12, 240, 245].

The entropic uncertainty relations with quantum memory provide a bound on $H_{\min}(Z_A|E)$ [239, 241, 243]:

$$H_{\min}(Z_A|E) + H_{\max}(X_A|B) \geq -\log c(X, Z) \tag{5.2}$$

where $H_{\max}(X_A|B)$ is the conditional max-entropy of Alice's classical register, X_A, for storing results from X-type measurements, given Bob's system, B. It is given by [243]:

$$H_{\max}(X_A|B) = 2 \log \max_{\sigma_B} F(\rho_{X_A B}, \mathbb{I}_{X_A} \otimes \sigma_B) \tag{5.3}$$

with $F(\rho, \sigma) = Tr(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}})$ denoting the fidelity between operators ρ and σ. $H_{\max}(X_A|B)$ can be upper bounded using $H_{\max}(X_A|X_B)$, in which the max-entropy is conditional on Bob's results from an X-type measurement [254]. Finally, the bound depends on the POVM elements for the two measurements. One bound is the maximum overlap between the two POVMs:

$$c(X, Z) = \max_{x,z} ||\sqrt{\mathbb{X}^x_A}\sqrt{\mathbb{Z}^z_A}||^2_\infty \tag{5.4}$$

where $||\cdot||_\infty$ denotes the maximum singular value [239, 243]. An equal or better bound is provided by [242]:

$$c'(X, Z) = \min\Big\{ \max_x ||\sum_z \mathbb{Z}^z_A \mathbb{X}^x_A \mathbb{Z}^z_A||_\infty, \max_z ||\sum_x \mathbb{X}^x_A \mathbb{Z}^z_A \mathbb{X}^x_A||_\infty \Big\}. \tag{5.5}$$

Given the reliance of the bound on the POVMs, a problem arises when Alice has POVM elements from X and Z that cause $c(X, Z) \approx 1$. The general problem can be summarized as follows: Alice has two POVMs, $Z = \{\mathbb{Z}^z_A\}_{z=1}^{N_Z} \cup \{\mathbb{Z}^\emptyset_A\}$ and $X = \{\mathbb{X}^x_A\}_{x=1}^{N_X} \cup \{\mathbb{X}^\emptyset_A\}$, such that $||\sqrt{\mathbb{Z}^\emptyset_A}\sqrt{\mathbb{X}^\emptyset_A}||^2_\infty \approx 1$, while $||\sqrt{\mathbb{Z}^z_A}\sqrt{\mathbb{X}^x_A}||^2_\infty < 1$ for the other POVM elements. The elements denoted with "$\emptyset$" indicate some measurement outcomes that cause $c(X, Z) \approx 1$. For convenience, we will often refer to these as "null" measurement outcomes, since a common way for $c(X, Z) \approx 1$ is when the detector does not register a result; however, we do not specify a form for these elements, meaning in general they can correspond to any measurement outcomes with POVM elements that are problematic for the maximum overlap. The current entropic uncertainty relation would provide a trivial bound due to these null measurements, but ideally we would like to salvage a bound using the other POVM elements that do not saturate the overlap. In the next section we provide a modified entropic uncertainty relation in which Alice and Bob can still bound Eve's information about the Z-type measurement outcomes, in terms of the X-type measurement outcomes for $1 \leq x \leq N_X$, the POVM elements from $\{\mathbb{Z}^z_A\}_{z=1}^{N_Z}$ and $\{\mathbb{X}^x_A\}_{x=1}^{N_X}$, and the probability Alice's measurements return null outcomes.

Before introducing the bound, we illustrate the relevance of this limitation to current entropic uncertainty relations by considering what occurs when Alice has basis-dependent limitations on her measurement range [222, 223, 235, 249, 250]. For instance, if Alice were measuring the frequency of a single photon, she might have N_Z POVM elements for describing the N_Z frequency bins that the detector can resolve, plus the null

measurement POVM element, $\mathbb{Z}_A^\emptyset$, to characterize the result of the photon falling outside the bandwidth of the detector. If her X-type measurement corresponds to an arrival-time measurement, then she will also have a finite number of bins with good timing resolution, then a null measurement element, $\mathbb{X}_A^\emptyset$, corresponding to the case when the photon arrived before or after her well-resolved time bins. One might think that because frequency and arrival-time are non-commuting observables, Alice and Bob would be able to have a non-trivial bound on Eve's information. Unfortunately, in such a case, because the null measurement POVM elements are measurement-dependent (i.e. $\mathbb{Z}_A^\emptyset \neq \mathbb{X}_A^\emptyset$) and span a semi-infinite region of the Hilbert space, $c(X,Z) = ||\sqrt{\mathbb{Z}_A^\emptyset}\sqrt{\mathbb{X}_A^\emptyset}||_\infty^2 \approx 1$.

To resolve this issue, one might be tempted to discard all the null outcomes from the frequency and arrival-time measurements, and bound the conditional min-entropy of the remaining data from the frequency measurement using the remaining data from the arrival-time measurement. By only maximizing the overlap over the POVM elements corresponding to the well-resolved bins, perhaps one can discount the null measurement POVM elements. However, a conceptually simple strategy, outlined in [223], can be used by Eve to make Alice and Bob overestimate the conditional min-entropy: Eve can make measurements of the photon frequency with very narrow bin widths, such that when Alice and Bob perform arrival-time measurements, the most likely outcome is a null measurement. Thus, Eve gains all information about frequency without introducing any errors to their frequency measurements, and Alice and Bob will almost always discard the arrival-time events for which Eve does not have any information and which would otherwise be used to bound Eve's information.

Note that this problem occurs when the null measurement POVM element depends on the measurement type, i.e. when $\mathbb{X}_A^\emptyset \neq \mathbb{Z}_A^\emptyset$ [247]. In Appendix D.1, we show that, when the null measurement POVM element is the same in both measurements, which can naively lead to $c(X,Z) = 1$, one can always formulate an entropic uncertainty relation in terms of the POVM elements of a related, effective measurement that does not have the problematic null measurement element. We show how to construct the related, effective measurement, and discuss why the approach for solving the problem when the null measurement is basis-independent fails when the null measurement is basis-dependent.

5.3 A Modified Entropic Uncertainty Relation

Our main result is a modified entropic uncertainty relation for a scenario in which Alice has specific outcomes in Z- and X-type measurements for which the POVM elements characterizing those outcomes yield a trivial bound on Eve's information, but those outcomes have low probability of occurring. Our result allows Alice and Bob to achieve a sometimes better bound on Eve's information about the measurement outcomes, with no extra characterization of the state required, simply by including the fact that the problematic outcomes have a low probability of occurring. We do not expect our bound to be tight. Its main utility is for generating non-zero bounds for scenarios where null measurement POVMs yield a large overlap, with the probability of such measurements remaining very low.

Main result : For a tripartite state, ρ_{ABE}, and two POVMs on $\mathcal{H}_A$, $Z = \{\mathbb{Z}_A^z\}_{z=1}^{N_Z} \cup \{\mathbb{Z}_A^\emptyset\}$ and $X = \{\mathbb{X}_A^x\}_{x=1}^{N_X} \cup \{\mathbb{X}_A^\emptyset\}$,

$$H_{\min}(Z_A|E) \geq -2\log\left[\sqrt{p_{Z_A}^\emptyset} + \sqrt{p_{X_A}^\emptyset} + \sqrt{1-p_{X_A}^\emptyset}\sqrt{c^<(X,Z)}\left(\sqrt{2}^{H_{\max}(X_A^<|B)}\right)\right] \tag{5.6}$$

where $p_{Z_A}^\emptyset = \mathrm{Tr}(\rho_A \mathbb{Z}_A^\emptyset)$ and $p_{X_A}^\emptyset = \mathrm{Tr}(\rho_A \mathbb{X}_A^\emptyset)$ are Alice's probabilities of null measurements from Z- and

X-type measurements. $H_{\max}(X_A^{<}|B)$ is the conditional max-entropy of Alice's X-type measurement results, after she has discarded the null measurements, given Bob's system. Finally,

$$c^{<}(X,Z) = \max_{(x,z)\neq\emptyset} ||\sqrt{\mathbb{X}_A^x}\sqrt{\mathbb{Z}_A^z}||_\infty^2 \tag{5.7}$$

is the maximum overlap for the Z and X POVM elements, excluding the null measurement POVM elements that cause the original bound to saturate.

Practically, the quantities in Eq. (5.6) are as experimentally accessible as those in Eq. (5.2), the unmodified entropic uncertainty relation. The new maximum overlap in Eq. (5.7) can be calculated after having characterized the POVMs for the detectors [255]. The null measurement probabilities can be calculated directly from the statistics of the measurement outcomes. $H_{\max}(X_A^{<}|B)$ can be bounded from above by $H_{\max}(X_A^{<}|X_B)$, the conditional max-entropy of Alice's X type measurement, after she has discarded the null measurements, given Bob's X measurement results. As we discuss in Section 5.4, Bob may need to make some modifications to his data, like replacing his null outcomes with bit values fitted to Alice's distribution. $H_{\max}(X_A^{<}|X_B)$ can be further bounded from above using methods from [230].

The result is quite general: we make no assumptions about the dimension of the Hilbert spaces or the commutation relations between the POVM elements. The proof of this result is available in Appendix D.2. We use similar analytical techniques from [243], which provides a proof of Eq. (5.2). Additionally, we use the form of the fidelity function in terms of trace norm, $F(\rho,\sigma) = ||\sqrt{\rho}\sqrt{\sigma}||_{\mathrm{Tr}}$, and exploit the triangle inequality property for norms. In Appendix D.3, we provide a smoothed version of the bound, which could be used to include finite-size effects, as Eq. (5.6) is valid for the asymptotic limit.

As stated above, our result provides a *sometimes* better bound on $H_{\min}(Z_A|E)$, and so a few comments are in order on the applicability of the result. First, note that if $\sqrt{p_{Z_A}^{\emptyset}} + \sqrt{p_{X_A}^{\emptyset}}$ exceeds 1 then the trivial lower bound of $H_{\min}(Z_A|E) \geq 0$ is better. See Fig. 5.2 for a contour plot of necessary values of $p_{Z_A}^{\emptyset}$ and $p_{X_A}^{\emptyset}$ for a non-trivial bound. Having reasonable values of $p_{X_A}^{\emptyset}$ and $p_{Z_A}^{\emptyset}$ depends on what $\mathbb{X}_A^{\emptyset}$ and $\mathbb{Z}_A^{\emptyset}$ correspond to physically. For example, in Section 5.4, in the context of time-frequency QKD, these null measurement probabilities should ideally correspond to the probability that single photons arriving at Alice's detectors fall outside her measurement ranges, resulting in no detector clicks; however, in practice, there are many scenarios that also result in null detection, including loss and vacuum components. While these other scenarios do not compromise security, they certainly cause an overestimation of $p_{X_A}^{\emptyset}$ and $p_{Z_A}^{\emptyset}$ and a pessimistic bound on the secure key rate. We will discuss what *additional* assumptions may need to be made to achieve a good bound. In Section 5.5, when we discuss homodyne-based CV QKD, $\mathbb{X}_A^{\emptyset}$ and $\mathbb{Z}_A^{\emptyset}$ correspond to quadrature intensities above the saturation limit of the detectors; in practice, saturation probabilities can be kept low, so fewer assumptions are needed to achieve a good bound.

Second, to tolerate higher null measurement probabilities, one requires low values of $c^{<}(X,Z)$ and $H_{\max}(X_A^{<}|B)$. The former can be achieved if the degree of freedom being measured is high-dimensional and the Z- and X-type measurements are non-commuting, while the latter can be achieved if Alice and Bob's systems are highly entangled. We will show in the next two sections that our bound provides insight into the security proofs for entanglement-based time-frequency, and continuous variable QKD using homodyne detection, two protocols employing high-dimensional degrees of freedom and Fourier pair measurements.

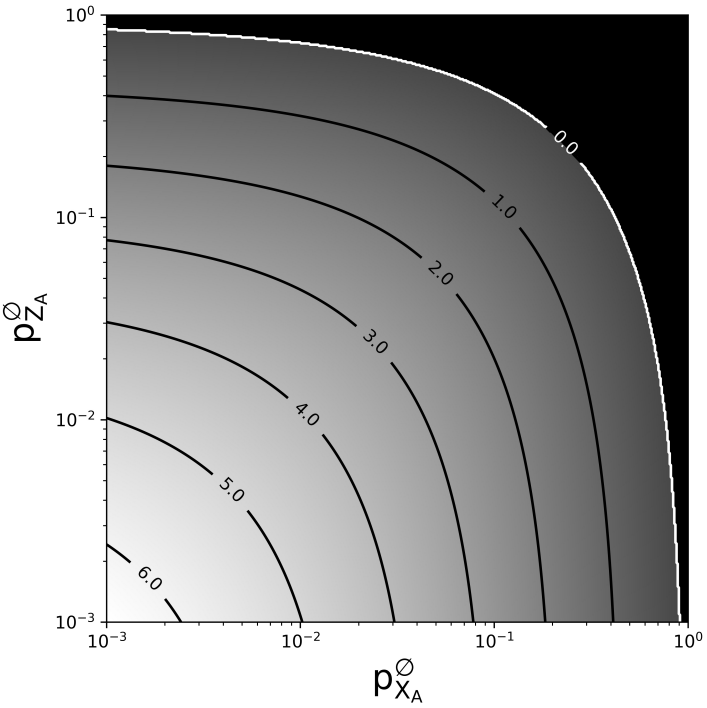


Figure 5.2: Contour plot of the lower bound, in bits, on $H_{\min}(Z_A|E)$ as a function of $p^{\emptyset}_{Z_A}$ and $p^{\emptyset}_{X_A}$ as given by Eq. (5.6). We have fixed $H_{\max}(X^{<}_{A}|B) = 1$, which corresponds to a noiseless scenario, and $c^{<} = 10^{-3}$, which is an experimentally feasible value in time-frequency QKD [230]. This plot corresponds to best-case scenarios for the bound, and provides necessary conditions on the values of $p^{\emptyset}_{Z_A}$ and $p^{\emptyset}_{X_A}$ for protocols employing Eq. (5.6) to prove security, with the black region corresponding to no proven security. Assuming no noise, for $c^{<} = 10^{-3}$, the maximum tolerable equal probability of null measurement is $p^{\emptyset}_{Z_A} = p^{\emptyset}_{X_A} \approx 23.2\%$, while in the limit $c^{<} \to 0$, $p^{\emptyset}_{Z_A} = p^{\emptyset}_{X_A} = 25\%$. Of course, if we make $p^{\emptyset}_{Z_A}$ smaller, we gain tolerance for higher $p^{\emptyset}_{X_A}$, and vice versa. For example, with $p^{\emptyset}_{Z_A} = 10^{-3}$, $p^{\emptyset}_{X_A}$ can be as high as $\sim$92%.

5.4 Application to Time-Frequency QKD

One of the main applications for entropic uncertainty relations is in proofs of the security of QKD protocols. QKD is a method for two spatially separated parties, Alice and Bob, to establish a shared, secret cryptographic key in the presence of an eavesdropper, Eve, where, for instance, the results of one measurement can be used to form the key, and the results of the other measurement can be used to bound Eve's information about the key [8]. A distinction is made between prepare-and-measure protocols, in which Alice uses a fully-characterized source to send signals to Bob, who has fully-characterized measurements, and entanglement-based protocols, in which Alice and Bob both have full characterization of their measurements, and an untrusted source outputs states for them to measure. The prepare-and-measure scenario can usually be framed in terms of the entanglement-based scenario [248], and the latter is suited to be analyzed using entropic uncertainty relations, since the statistics of the measurement outcomes and the POVMs characterizing the measurement can be used to bound Eve's information.

One example of a QKD protocol for which our main result is particularly relevant is time-frequency QKD. Time-frequency QKD employs single photon frequency and arrival-time as non-commuting observables to establish a key [221]. In an entanglement-based protocol, an untrusted source would output two spatial modes, one sent to Alice and one to Bob, who would then randomly alternate between measuring frequency and arrival-time, with each observable having bins of finite width into which results can fall.

Typically, Alice's frequency POVM in the single photon subspace, F_{ideal}, is idealized with elements [223]:

$$\mathbb{F}_A^m = \int_{\omega_m-\delta\omega/2}^{\omega_m+\delta\omega/2} \frac{d\omega}{2\pi} |\omega\rangle\langle\omega| \tag{5.8}$$

where ω_m is the central frequency of the bin; $\delta\omega$ is the bin width; and $|\omega\rangle$ is a frequency eigenstate with normalization $\langle\omega|\omega'\rangle = 2\pi\delta(\omega-\omega')$ [230]. It is often assumed that the central frequencies of the bins are defined over positive and negative frequencies [223, 232], although some have modified the definition to include only positive frequencies [230]. Of course, a full characterization of Alice's measurement will also include POVM elements for vacuum and multiphoton contributions; we will return to this issue later in our discussion.

The arrival-time eigenstates are defined as Fourier pairs of the frequency eigenstates, with Alice's arrival-time POVM in the single photon subspace, T_{ideal}, typically defined by elements [230]:

$$\mathbb{T}_A^k = \int_{t_k-\delta t/2}^{t_k+\delta t/2} dt |t\rangle\langle t|, \tag{5.9}$$

where t_k is the central arrival time of the bin; δt is the bin width; and $|t\rangle = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} e^{i\omega t}|\omega\rangle$ is the Fourier transform of $|\omega\rangle$. As with frequency, the sequence of central times extends over positive and negative arrival times.

Using Eq. (5.4), these POVMs yield a maximum overlap of:

$$c(T_{\text{ideal}}, F_{\text{ideal}}) = \frac{\delta\omega\delta t}{2\pi} S_0^{(1)}(1, \frac{\delta\omega\delta t}{4})^2 \tag{5.10}$$

where $S_0^{(1)}(\cdot,\cdot)$ is the 0th radial prolate spheroidal wave function of the first kind [12, 256]. Eq. (5.10) is roughly proportional to $\delta\omega\delta t$ when their product is small [256]. As expected, $c(T_{\text{ideal}}, F_{\text{ideal}}) \to 1$ as the bin

width product grows.

The obvious problem with this characterization is that realistic detectors cannot have a constant bin width over all frequencies and arrival-times. For instance, a typical source for time-frequency entangled photons may have a repetition rate on the order of 10 MHz [224, 257], which would limit the range of arrival-times that can be employed. Additionally, the best single photon detectors have high efficiency in a finite spectral range of a few tens of nanometers about 1550 nm [258]. We assume that when the photon falls beyond these ranges of the detectors, the detectors do not click. Thus, one step towards a more realistic characterization of the measurements is to include the null measurement elements corresponding to when the photon falls outside the temporal range in the arrival-time measurement or the spectral range in the frequency measurement. If the temporal range is from $[-t_c, t_c]$, then $\mathbb{T}_A^{\emptyset} = \int_{-\infty}^{-t_c} dt|t\rangle\langle t| + \int_{t_c}^{+\infty} dt|t\rangle\langle t|$. If the spectral range is from $[\omega_o - \omega_c, \omega_o + \omega_c]$, then $\mathbb{F}_A^{\emptyset} = \int_{-\infty}^{\omega_o-\omega_c} \frac{d\omega}{2\pi}|\omega\rangle\langle\omega| + \int_{\omega_o+\omega_c}^{+\infty} \frac{d\omega}{2\pi}|\omega\rangle\langle\omega|$. Thus, the refined POVMs are $T = \{\mathbb{T}_A^k\}_{k=1}^{N_T} \cup \{\mathbb{T}_A^{\emptyset}\}$ and $F = \{\mathbb{F}_A^m\}_{m=1}^{N_F} \cup \{\mathbb{F}_A^{\emptyset}\}$, where we assume that t_c (ω_c) or δt ($\delta\omega$) can be chosen so that the N_T (N_F) bins can be defined within $[-t_c, t_c]$ ($[\omega_o - \omega_c, \omega_o + \omega_c]$). Based on (10), $c(T, F) \approx 1$ since the null measurement bins have, in principle, infinite width. As discussed previously, this has been a known issue for time-frequency QKD [222, 223].

A common suggestion for dealing with this problem has been to apply pre-measurement filters to exclude the frequencies or arrival-times that would not be detected [222, 223]. Consider, however, that for the unmodified bound, the maximum overlap in Eq. (5.4) is not sensitive to the probabilities of different measurement outcomes, only to the POVM elements. This means that, as Eq. (5.2) stands, it does not necessarily matter that a filter could keep those probabilities low. Thus, without modifying the entropic uncertainty relations, the only way to ensure that we can safely disregard the null measurement POVM elements with semi-infinite support that saturate Eq. (5.4) is to ensure that the state has no shared support with those POVM elements, effectively reducing the problem to the smaller subspace defined by the support of the state. Unfortunately, it is not possible to construct a filter that could simultaneously ensure compact support in the time domain and frequency domain [259]. In other words, we cannot simultaneously make the state have exactly zero probability of yielding null measurements on the spaces of $\mathbb{T}_A^{\emptyset}$ and $\mathbb{F}_A^{\emptyset}$ without filtering out all of the state. As long as there is some non-zero probability, one would need to consider the POVM elements on that space, which saturate Eq. (5.4) and render the unmodified entropic uncertainty relations trivial.

Alternatively, as discussed in [249, 250], one could assume that there exists a cut-off outside the measurement range, and that the probability of a result beyond the cut-off is negligible, using less coarse-grained measurements with a wider range to estimate this probability. We would then need to ask how small that probability needs to be, and how to quantify its effect within the bound on security.

Our main result, Eq. (5.6), is a step towards dealing with this problem. In [230], an ideal source for time-frequency entangled photons has a biphoton wavefunction modelled as:

$$\psi(\omega_A, \omega_B) = \frac{\exp[-(\omega_A - \omega_B)^2\sigma_{cor}^2/4 - (\omega_A + \omega_B)^2\sigma_{coh}^2]}{\sqrt{\pi/2\sigma_{coh}\sigma_{cor}}} \tag{5.11}$$

where ω_A and ω_B are relative to the central telecom frequency; σ_{coh}, the coherence time, is taken to be 6 ns, and σ_{cor}, the correlation time, is taken to be 2 ps. Assuming the spectral range is between 1520 nm and 1610 nm [258], the state in Eq. (5.11) yields a probability for a null outcome in the frequency measurement of $p_{F_A}^{\emptyset} = 0$ (to within machine precision[2]). Note that this is in the best case scenario when an

[2]We used MATLAB, for which a positive value less than 2^{-52} is treated as 0 [260].

eavesdropper is not tampering with the value; of course, in a real-world entanglement-based QKD scenario this probability value would be measured from the data. Using a repetition rate of 55.6 MHz [230], Eq. (5.11) yields a probability for a null outcome in the arrival-time measurement of $p^{\emptyset}_{T_A} \approx 0.27\%$. We can upper bound $c^{<}(T,F) = \max_{k,m\neq\emptyset} ||\sqrt{\mathbb{T}^k_A}\sqrt{\mathbb{F}^m_A}||^2_\infty$ by $c(T_{\text{ideal}}, F_{\text{ideal}})$ since the sets of POVMs over which the former is maximized are subsets of the sets over which the latter is maximized. Again using values from [230], this yields $c^{<} \leq 10^{-3}$.

The original problem was that the unmodified entropic uncertainty relation, Eq. (5.2), was rendered trivial by the overlap $||\sqrt{\mathbb{Z}^{\emptyset}_A}\sqrt{\mathbb{X}^{\emptyset}_A}||^2_\infty \approx 1$, so it is a clear improvement that the new maximum overlap, $c^{<}$, is no longer saturated by the null measurement POVM elements. Unfortunately, a new problem arises when applying Eq. (5.6) in a practical implementation of entanglement-based, time-frequency QKD: $p^{\emptyset}_{F_A}$, $p^{\emptyset}_{T_A}$, and $H_{\max}(X^{\leq}_A|B)$ will all need to be estimated using the measurement data, a task limited by source, coupling, and channel imperfections.

In a true entanglement-based scenario, only the detectors are trusted. This means that the source, channels and couplings are not. In Eq. (5.6), $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$ should ideally correspond to the probabilities that the single photon portion of Alice's state yields null measurements. However, less than ideal devices may result in an overestimation of these parameters.

First, the source will pose a problem due to vacuum components since vacuum states also yield null measurements. For example, a common choice for creating spectrally entangled photon pairs is a spontaneous parametric down-conversion device, which can have low conversion efficiency [225]. Therefore, to get a better estimate on $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$, Alice would ideally want to determine the probability the untrusted source emits a vacuum state. While passive decoy state methods allow for the characterization of the photon number distribution for untrusted sources in single-mode QKD [261], we are not aware of any methods for characterizing the photon number distribution of an untrusted source for high-dimensional QKD employing many time-frequency modes. Thus, short of such methods, the source will be untrusted and uncharacterized, and we would have to allow null measurements due to vacuum components to be part of the estimation of $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$. However, as shown in Fig. 5.2, this would mean the probability of a vacuum component would need to be less than 25%, and that is additionally on condition that there be perfect correlations in measurement results, and a very low value of $c^{<}$. This may not be possible with current technology, as increasing the number of photon pairs is typically achieved by pumping at higher intensities which also results in increasing the number of multiphoton contributions [262]. Multiphoton detection events are a source of noise [263], and thus the high correlations between Alice and Bob's data will be unattainable.

Second, the problem caused by in-lab coupling from the channels to the detectors is conceptually similar: a lost photon yields a null measurement. Even if the couplings induce a basis-independent loss, if we do not want to trust the couplings, we will need to factor their contribution to null measurements into the estimation of $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$. To have a non-trivial bound, that loss cannot be greater than 25%, on condition that vacuum contributions are negligible, correlations between measurement results are high, and we have a very low value of $c^{<}$.

There is clearly a practicality problem if we do not trust the source since there are so many scenarios that can cause a null measurement. The problem can be slightly improved if we allow partial characterization of the source and in-lab couplings, weakening the full entanglement-based assumption. For instance, Alice need not know the full biphoton wavefunction, but if she can know the probability that the source outputs a vacuum state and characterize the basis-independent coupling losses between the source and her detector, then she can have a much better estimate of $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$. This is because she can first use the process from

Appendix D.1 to discount any negative impact due to basis-independent null measurement POVM elements, and then she can use Eq. (5.6) to treat the remaining, basis-dependent null measurements, whose impact on security is treated via the probabilities $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$. With no tampering of her single photon component, she should expect the very low values for $p^{\emptyset}_{F_A}$ and $p^{\emptyset}_{T_A}$ calculated earlier.

However, in all QKD protocols the channel is untrusted. In Eq. (5.6), $H_{\max}(X^{<}_{A}|B)$ should ideally be the conditional max-entropy of Alice's measurement outcomes from single photon components that fell within her measurement range, given the single photons that arrived at Bob's detector. Unfortunately, Bob will have additional null measurements due to all the photons lost in the channel. Even if Bob can discard null measurements due to the source outputting a vacuum component, or due to coupling from the channel to his detector, he will have to assume all the null measurements due to loss in the channel are instead due to single photons falling outside his detector range. Since there will be cases where Alice did not have a null measurement while Bob did, Bob's strategy in those cases will then be to assign a bit value to these null outcomes based off the publicly known probability distribution of Alice's results, to have a lower chance of error.

Using the state in Eq. (5.11), with a loss of 0.2 dB/km in fiber, assuming Alice has partial characterization of her source so she can safely estimate her null measurement probabilities to be $p^{\emptyset}_{F_A} = 0$ and $p^{\emptyset}_{T_A} = 0.27\%$, and taking $c^{<} = 10^{-3}$, the bound on $H_{\min}(F_A|E) - H(F_A|F_B)$ still saturates at $\sim$2 km, even in an ideal case where Alice and Bob have no dark counts and Eve has not interfered with the results, as shown in Fig. 5.3. Here, $H(F_A|F_B)$ is the conditional Shannon entropy of Alice's frequency results given Bob's, and it is used to quantify the number of bits to correct errors in the key [8, 223]; in this case, it will be non-zero even in noiseless channels because of the finite coherence time of the state. We used methods from [230] to bound $H_{\max}(T^{<}_{A}|B)$. Note that our result significantly differs from the distance of >150 km presented in [230]; however, their analysis did not address the measurement range problem, so it would not provide security given realistic limitations on the measurement range. We have shown the region of security can be expanded, albeit slightly, to more than 2 km.

It appears that bounding the security of entanglement-based time-frequency QKD with a completely untrusted source is still impractical using Eq. (5.6). With knowledge of the source's vacuum component probability, and characterization of their coupling losses, Alice and Bob can achieve a better bound on $H_{\min}(F_A|E)$, but only if the channel distance to Bob is very short. The result may still be useful for applications of time-frequency entanglement other than entanglement-based communication, such as entanglement witnessing [239].

Clearly, time-frequency QKD with arrival-time and frequency treated as Fourier pairs of each other is not currently feasible with an untrusted source and current levels of loss; seeing that the null measurements depend on the type of measurement, loss must be treated as a threat to security. We therefore have the following outlook for time-frequency QKD: if one wants to continue characterizing arrival-time and frequency as Fourier pairs, then one will need to move to a prepare-and-measure or a measurement-device-independent QKD setting. For the former approach, the equivalence between prepare-and-measure and entanglement-based QKD that is often used for security proofs will need to be carefully considered, since an entanglement-based approach is untenable. Alternatively, it may be fruitful to re-examine the characterization of arrival-time and frequency measurements: it is likely that one is not measuring perfect Fourier pairs in the lab, so perhaps the newly-characterized observables that one is measuring will end up having basis-independent null measurements, which we know not to be a problem for security. Additionally, some work has been done to implement measurements in arrival-time-like and frequency-like measurements: rather than measuring

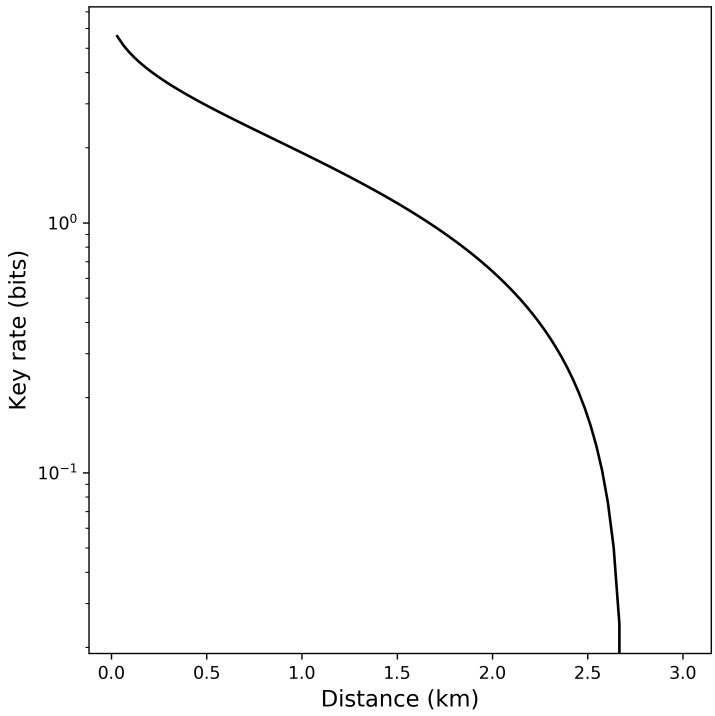


Figure 5.3: Key rate vs. distance for time-frequency QKD, accounting for security repercussions of null measurements on Bob's side. For the state in Eq. (5.11), we have set $\sigma_{coh} = 6$ns, and $\sigma_{cor} = 2$ps [230]. To employ Eq. (5.6) we have used $c^{<} = 10^{-3}$, $p^{\emptyset}_{T_A} = 0.27\%$, and $p^{\emptyset}_{F_A} = 0\%$. We have bounded $H_{\max}(T_A^{<}|B)$ using techniques from [230], employing a bin width of $\delta t = 20$ps. See the text for additional assumptions. Previous security analysis of time-frequency QKD in [230] presented non-zero key rate at distances of >150 km, but the proof does not provide security against the measurement range problem. Our new bound widens the region of security to more than 2 km.

Fourier pairs of observables, one either makes a measurement in a basis of d arrival times, $\{|t_n\rangle\}_{n=0}^{d-1}$, or in a superposition basis, $\{|f_m\rangle = \sum_{n=0}^{d-1} \exp(2\pi i n m/d)|t_n\rangle\}_{m=0}^{d-1}$ [226, 227]. In that case, the measurement range problem is averted because the system of interest is now defined on an effective subspace in which there are lower probabilities of basis-dependent null measurements. Finally, as our bound was not proven to be tight, we do not rule out the possibility of a better bound that is not as quickly saturated by the probability of null measurements.

5.5 Application to Homodyne-Based Continuous Variable QKD

CV QKD exploits the electric field quadratures as the non-commuting observables to establish a secret key. In a typical entanglement-based protocol employing homodyne detection, an untrusted source sends one mode to Alice, another to Bob, and they each randomly perform either X or P quadrature measurements [234, 236]. Entanglement-based CV QKD employing heterodyne detection is also possible [264–267]; however, the entropic uncertainty relations are not applicable for proving security in such schemes [265, 267], so we do not consider them here as our result relates to a modification of the entropic uncertainty relations.

In close analogy to time-frequency QKD, for a homodyne-based CV QKD protocol, Alice sorts her results into bins, yielding the POVMs for X and P quadrature measurements with elements [235]:

$$\mathbb{X}_A^m = \int_{x_m-\delta/2}^{x_m+\delta/2} dx|x\rangle\langle x|, \quad \mathbb{P}_A^k = \int_{p_k-\delta/2}^{p_k+\delta/2} dp|p\rangle\langle p| \tag{5.12}$$

where $|x\rangle$ and $|p\rangle$ are the eigenstates of the quadrature operators. If the constant bin width, δ, can be maintained across the entire measurement range, then we should expect an expression for the maximum overlap just like Eq. (5.10), since that equation is derived for general continuous variables respecting Heisenberg-like uncertainty relations [256].

Of course, as with time-frequency QKD, we cannot expect the same level of coarse-graining over the entire Hilbert space. Above some intensity level, the detectors will become saturated. The measurement operators characterizing saturation in the X and P measurements will be defined over semi-infinite ranges, meaning $c \approx 1$ [235]. The suggestion made in [235] for dealing with this issue is that one needs to trust the source, assume it has a low probability of saturating the detectors, and then incorporate that probability into the smoothing parameter for the entropic uncertainty relations resulting in a failure probability for the protocol. In [237], the solution is to assume that the energy of the source is bounded, so that the probability for saturation can be estimated.

This issue has been discussed in [252, 253], in the context of the Gaussian modulated coherent state protocol, a prepare-and-measure setting. They found that an eavesdropper can shift the mean of the distribution of results into the saturation regime, simultaneously lowering the variance of results, which causes Alice and Bob to overestimate the security of their key. Among countermeasures suggested, they discussed introducing a confidence interval for the results, and if too many results fall beyond the confidence interval, Alice and Bob ought to abort the protocol; however, the range of this confidence interval and the threshold probability for aborting the protocol were left open for future work.

Our main result, Eq. (5.6), addresses this gap: our bound depends explicitly on the probabilities of saturating the detectors, $p_{X_A}^{\emptyset}$ and $p_{P_A}^{\emptyset}$, which can be measured from the data without having to trust the source. This provides a way for Alice and Bob to monitor how many results are beyond their confidence interval, guarding against a saturation attack by Eve. Note that, by depending on $p_{X_A}^{\emptyset}$ and $p_{P_A}^{\emptyset}$, our bound

additionally provides an implicit way to include the choice of the measurement range as an optimization parameter for the protocol.

Note the major difference between the problematic measurements in time-frequency and CV QKD using homodyne detection. In the former, losing a photon had the same consequence for security as falling outside the measurement range, since both resulted in the detector not clicking; the high loss in fiber optic channels compromised security after a short distance. In homodyne-based CV QKD, however, saturation is problematic, and luckily the fiber optic channel will not naturally introduce gain that will convert a low intensity signal into a saturating signal. Other than tampering, the main source for detector saturation is due to the tails of the Gaussian distributions in phase space from the two-mode squeezed vacuum states used for entanglement-based protocols. Luckily, these probabilities are vanishingly small. For example, in [237], the range of measurement is [-61.6, 61.6] in units of vacuum noise, and using an anti-squeezing factor of 19.3 dB [268], this yields $p^{\emptyset}_{X_A} = p^{\emptyset}_{P_A} \approx 0$ to within machine precision. We should not expect the results with the new bound to differ from [237], since Eq. (5.6) reduces to Eq. (5.2) when those probabilities are zero.

5.6 Conclusion

We have presented a modified entropic uncertainty relation to discount unwanted POVM elements that render the unmodified entropic uncertainty relation trivial. This is done at the cost of including the probability of the unwanted measurement in the bound. Our bound offers insight into the measurement range problem, which poses an issue for the characterization of entanglement in high-dimensional systems. We applied the bound to analyze entanglement-based time-frequency QKD, and found that, unlike previous results, we can now guarantee security; however, this is conditional on low loss and high detection efficiency within the measurement range. In a practical setting, this may only be achievable with some characterization of the source, like knowing the probability a vacuum state is emitted, weakening the completely untrusted source assumption of entanglement-based QKD. Under realistic conditions, the key becomes insecure using our bound at $\sim$2 km, mainly due to the high loss in fiber. Since we did not show the bound to be tight, this does not preclude a better bound in the future from extending the secure distance of the protocol. Finally, we discussed our bound as it related to saturation attacks in CV QKD employing homodyne detection schemes. Through the bound's dependence on the probability of detector saturation, we provide a new quantitative way to guard against saturation attacks, without any assumptions about the source but with full characterization of the measurements.

As a consequence of this study, we see the susceptibility of detectors in QKD protocols to security loopholes. Luckily, an alternative strategy is offered by measurement-device-independent (MDI) QKD [24], wherein the channel and detectors are completely untrusted, but the sources are trusted. In the next two chapters, we will examine security proofs in the MDI setting and how they can be adapted to account for imperfections in the optical source.

Chapter 6

Loss-tolerant QKD with a Twist

This chapter is based on [43], co-authored with Ignatius William Primaatmaja, Prof. Charles Ci Wen Lim, and Prof. Hoi-Kwong Lo. The work was initiated by Charles Ci Wen Lim, who supervised the project along with Hoi-Kwong Lo. The project was collaborative in nature, and published in *Physical Review A*. I was the first author for this work. My main contributions were to the literature review, mathematical proofs, numerical simulations, analysis, and manuscript writing. The work benefited from helpful discussions with Ilan Tzitrin, Wenyuan Wang, Thomas van Himbeeck, Emilien Lavie, and Koon Tong Goh.

6.1 Introduction

There has been significant interest in quantum hacking against practical quantum key distribution (QKD) systems [8, 269]. In particular, single photon detectors (SPDs) have been identified as the weakest link in the security of practical QKD. To completely bypass all possible attacks on SPDs, the concept of measurement-device-independent (MDI) QKD has been introduced and widely deployed. MDI QKD allows two distant parties, Alice and Bob, to distribute a shared, secret cryptographic key, even in the presence of an eavesdropper, Eve, who has complete control of their quantum channels and measurement devices [24, 270]. Typically, Alice and Bob prepare a set of signal states, send them to a central measurement node potentially controlled by Eve, which then makes an announcement based on a measurement it may not have faithfully executed. The cost of information-theoretic security in this setting is that Alice and Bob need to trust and characterize the optical sources they employ to send signals. Thus, it is especially valuable to account for the source features and flaws in a security proof when quantifying the key rates offered by an MDI protocol.

In this chapter, we answer a seemingly simple question: how do you construct a security proof for an MDI QKD protocol that employs trusted, yet noisy – i.e. mixed – signal states? To clarify, protocols that employ the decoy state method [271, 272] call for mixed optical states in the form of phase-randomized weak coherent pulses. However, in those protocols, the *signal* states – i.e. the single photon contributions – used for key generation are still often assumed to be pure. Unfortunately, a realistic source will not be able to initialize signal states with perfect purity. Therefore, our task is to build a consistent framework for optimally determining the security of MDI QKD protocols in the case of mixed signal states from a trusted source, using to our advantage that Eve may not hold the purification of the mixture.

There are several leading proof techniques for handling state preparation errors in a QKD protocol. The first major analysis was performed in [273]; however, the authors assumed pessimistically that Eve could

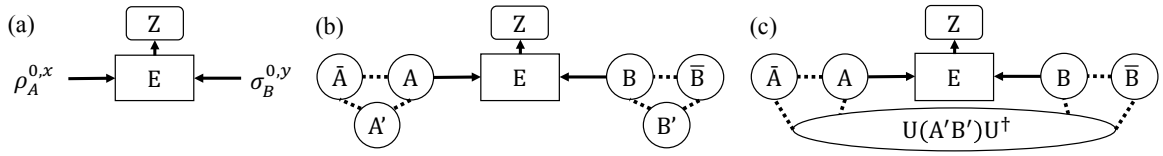


Figure 6.1: (a) A real MDI QKD protocol: Alice and Bob send mixed states associated with bit (x, y) values to a central node controlled by Eve, who announces Z. (b) A virtual (purified) picture of sending the key generation states: Alice and Bob's mixed signal states are entangled with virtual qubits $\bar{A}\bar{B}$ which coherently store the bit values (x, y) until they are revealed by a computational-basis measurement. The AB systems are additionally purified by the $A'B'$ systems to account for trusted noise in the source. Only the A, B systems are sent to Eve. (c) An alternative virtual picture: all purifications are related by unitary operations applied to, in general, a joint purifying ancilla, yielding private states in $\bar{A}\bar{B}A'B'$. These "twisting" operations can optimally boost the secret key rate as they modify the phase error rates which Alice and Bob need to estimate. In (a)–(c), the signal states sent and the observed detection and bit error rates are the same.

amplify such noise to her benefit, so the technique was not robust over long distances against e.g. coherent modulation errors. An improved technique was provided in the loss-tolerant protocol [25], which uses basis mismatch statistics to infer phase error rates that cannot be directly observed when state preparation is non-ideal. However, the technique leaves ambiguous how to treat *mixed* signal states, a gap this work closes. Different extensions of the loss tolerant protocol were considered in [274–276]; however, their focus was primarily on leaky sources, so treatment of mixed states was analogous to [25]. Another notable technique for characterizing security given pure qubit signal states is provided by [277]; however, their technique for generalizing to mixed signal states uses a suboptimal approach of averaging the key rates for each of the pure states in the mixture, which yields an equal or lower key rate than the key rate produced from the true average signal state. Lastly, an approach for finding a numerical lower bound on the Devetak-Winter secret key rate [278] for MDI-QKD protocols is provided by [279, 280]; their technique is in principle extendable to noisy state preparation. In our work, we take a conceptually simpler strategy of directly optimizing the key rate formula from [281], which uses the bit and phase errors of qubits in a virtual picture of the protocol.

In the case state preparation noise can be trusted and characterized, but perhaps not reduced, we provide here a simple analytical and numerical toolbox for calculating an optimal secret key rate. First, we provide a re-framing of the tilted four-state loss-tolerant protocol which provides a method for fixing Eve's degrees of freedom in the secret key rate [25, 282, 283]. However, as the signal states are mixed, the security also depends on how we treat the trusted noise in the signal state generation. Typically, the security of QKD is analyzed in terms of Alice and Bob's ability to virtually distill maximally entangled EPR pairs, since measurement of such pairs yields perfectly correlated keys, and by the monogamy of entanglement, the results cannot be correlated with anyone else, including Eve. However, it is known that a larger class of states known as private states [284–287] are fundamentally what is required to produce secret key. Formally, private states can be constructed from an EPR pair if Alice and Bob take ancillary shield systems they control, and apply a "twisting" unitary operation between the EPR pair and the shields, the condition being that this twisting leave unaffected the measurement results that generate secret key. Since twisting does not change the key, private states can then be understood as deflecting some of Eve's attack on the systems that generate key to the shield systems. See Fig. 6.1 for a diagram of this concept.

In our technique, we show that the mixing noise of the signal states can be treated in a virtual picture as being equivalent to Alice and Bob employing shield systems. Completely within this virtual picture, we can apply unitary twisting operations to the shields to decrease the phase errors of the protocol, increasing the

secret key rate. We provide simple semi-definite programs to find the optimal twisting operations, yielding the optimal key rate under this framework. Semi-definite programming [288] has recently become a powerful tool for quantifying the asymptotic security of QKD protocols [27, 279, 280, 289–293]. While private states have been of significant conceptual interest, as far as we are aware, this is the first application of private states in a practical QKD setting. Finally, we apply our technique to calculating fundamentally achievable key rates in an MDI QKD protocol with randomized modulation error in the state preparation procedure. We note that our technique is applicable to a wide class of MDI QKD protocols in which Alice and Bob each employ four qubit signal states that must not fall in the same plane of the Bloch sphere (which is easy to impose in practice), but which can be subject to general asymmetric preparation noise. Moreover, these signal states can be the single photon components of phase randomized coherent states in a decoy state protocol.

6.2 Characterizing Eve's State

We consider an MDI QKD protocol in which Alice and Bob each prepare four mixed qubit signal states $\{\rho_A^{i,x}\}$ and $\{\sigma_B^{j,y}\}$, that they will send respectively with probabilities $p^{i,x}$ and $q^{j,y}$ to the central measurement node controlled by Eve. Alice and Bob can characterize their initial states by e.g. performing tomography on their sources before the protocol, as in [282]. As implied by the notation, here we assume that the signal states are separable and uncorrelated from round to round of the protocol. As mixed states, there can be random fluctuations of which state is sent from round to round, but the overall average state must be given by the density matrix, and fluctuations between rounds are assumed to be uncorrelated. An interesting future problem would be to combine the technique presented in this chapter with the approach for treating correlated sources in [294].

When Alice and Bob choose $(i, j) = (0, 0)$ these are the key generation states with (x, y) corresponding to their key bit values. All other combinations (i, j, x, y) correspond to test states used to constrain the phase errors. Following the security proof of the loss tolerant protocol [25], we require that the sets of states $\{\rho_A^{i,x}\}$ and $\{\sigma_B^{j,y}\}$ each form a tetrahedron on the Bloch sphere, meaning the Bloch vectors cannot all lie in the same plane. In Appendix E.1, we provide steps for how to embed our technique within a decoy state protocol in the asymptotic limit of an infinite number of decoys. Note we are also assuming collective attacks, with an extension to coherent attacks available in [25].

As qubits, our signal states can be fully characterized with two orthonormal basis vectors, which we take to be the polarization states $|H\rangle, |V\rangle$:

$$\rho_A^{i,x}\sigma_B^{j,y} = \sum_{\substack{m,m',\\ n,n'=H}}^{V} c_{m,m'}^{i,x} d_{n,n'}^{j,y} |m,n\rangle\langle m',n'|_{A,B} \tag{6.1}$$

Under unitary evolution, each of these basis vectors evolves to a (subnormalized) state in Eve's possession as well as a classical announcement, z, which we take to be pass or fail: $|m,n\rangle_{A,B} \to \sum_{z=P}^{F} |e_{m,n}^z\rangle_E |z\rangle_Z$. This process generalizes simply to multiple announcement events, such as which Bell state Eve claims to have detected.

The probabilities that Eve announces a round successfully passed conditioned on the signal states sent

$p_{det}^{i,j,x,y} = p(z = P|i,j,x,y)$, provide constraints on the inner product of Eve's vectors $\langle e_{m',n'}^P | e_{m,n}^P \rangle_E$:

$$p_{det}^{i,j,x,y} = p^{i,x} q^{j,y} \sum_{\substack{m,m',\\ n,n'=H}}^{V} c_{m,m'}^{i,x} d_{n,n'}^{j,y} \langle e_{m',n'}^P | e_{m,n}^P \rangle_E \tag{6.2}$$

$p_{det}^{i,j,x,y}$ are observable satistics in the protocol, and they can also be used to directly calculate some quantities required for the secret key rate formula, such as the detection probability in the key basis, $p_{det}^{0,0} = \sum_{x,y} p_{det}^{0,0,x,y}$, and the bit error rate $e_Z = (p_{det}^{0,0,0,1} + p_{det}^{0,0,1,0})/p_{det}^{0,0}$, where we have taken $|\Phi^+\rangle$ to be the target Bell state that Alice and Bob wish to distill in a virtual picture we describe in the next Section.

We see that Eq. 6.2 can be written compactly as:

$$\vec{p}_{det} = \hat{\gamma}\vec{e} \implies \vec{e} = \hat{\gamma}^{-1}\vec{p}_{det} \tag{6.3}$$

where $\vec{e}_s = \langle e_{m',n'}^P | e_{m,n}^P \rangle_E$, $s = 1, ..., 16$, is the vectorized form of the Gramian matrix of Eve's states associated with a passing announcement; $(\vec{p}_{det})_t = p_{det}^{i,j,x,y}$, $t = 1, ..., 16$, is a vector containing all the successful detection probabilities; and $\hat{\gamma}_{ts} = p^{i,x} q^{j,y} c_{m,m'}^{i,x} d_{n,n'}^{j,y}$ is a matrix dependent on the initial states from Eq. 6.1 used in the protocol. As long as $\hat{\gamma}^{-1}$ exists, then we can exactly solve for $\vec{e}$, which can then be used to calculate any objective function of $\langle e_{m',n'}^P | e_{m,n}^P \rangle_E$, including all the phase error rates in the six-state protocol key rate formula [8, 25, 245, 281], even though we are only using four states, which were chosen to provide complete characterization of Eve's strategy. In Appendix E.2, we show that the invertibility of $\hat{\gamma}$ is equivalent to sending four states that form a tetrahedron on the Bloch sphere, as found in the original security proof of the tilted four-state loss tolerant protocol [25]. Additionally, we provide a generalization of the proof technique to high-dimensional MDI QKD [233, 291, 295–297].

6.3 Optimal Choice of Virtual Protocol

Having characterized Eve's Gramian matrix entirely from observable parameters in the protocol, we now move to a virtual picture for the key generation signal states to calculate the remaining parameters of the secret key rate. In this virtual picture, which is depicted in Fig. 6.1, systems A, B from Eq. 6.1 are entangled with virtual qubits $\bar{A}, \bar{B}$ that Alice and Bob keep in their laboratory [25]. Importantly, since these signal states are mixed, we require additional purifying ancillary systems $A'B'$. We assume that the sources of noise are confined to Alice and Bob's labs, meaning Eve does not have access to manipulate $A'B'$. The mixedness of the signal states then results in an effective virtual shield Alice and Bob can use to minimize Eve's knowledge of the secret key.

The key generation states, $p^{0,x} q^{0,y} \rho_A^{0,x} \sigma_B^{0,y}$ can be purified to:

$$|\zeta\rangle = \sum_{x,y} |x,y\rangle_{\bar{A}\bar{B}} \sum_{m,n=H}^{V} |\gamma_{m,n}^{x,y}\rangle_{A'B'} |m,n\rangle_{AB} \tag{6.4}$$

where we have constraints from the states in Eq. 6.1:

$$\langle \gamma_{m',n'}^{x,y} | \gamma_{m,n}^{x,y} \rangle_{A'B'} = p^{0,x} q^{0,y} c_{m,m'}^{0,x} d_{n,n'}^{0,y} \tag{6.5}$$

since to generate key, Alice and Bob measure $\bar{A}\bar{B}$ in the computational basis. The crucial point is that this

purification is not unique [2], and so we have freedom to choose the virtual picture that yields the optimal key rate. Since Eve does not have access to $A'B'$, any purification will yield a suitable lower bound on the key rate, but we will show how to choose the optimal purification with simple semidefinite programs.

We can parametrize all purifications using twisting unitary operations [284–287] applied to the virtual ancillary systems in $|\zeta\rangle$:

$$U_{\bar{A}\bar{B}A'B'} = \sum_{x,y=0}^{1} |x,y\rangle\langle x,y|_{\bar{A}\bar{B}} \otimes U^{x,y}_{A'B'} \tag{6.6}$$

Such an operation is entirely virtual, so it can be nonlocal in general and never needs to be executed in the real protocol. Twisting does not affect any of the real observed detection probabilities, which correspond to Alice and Bob first projecting $\bar{A}\bar{B}$ in the computational basis, as we show in Appendix E.3. Moreover, since only the A, B portion of $|\zeta\rangle$ evolves unitarily to E, Z, the twisting operation need not be fixed from the beginning of the protocol, and its choice can and should be informed by the statistics of the protocol. Such twisting operations applied to Bell states yield private states. We next show exactly how these twisting operations affect the secret key formula.

To quantify the security of the protocol, we employ the key rate formula from the six-state protocol [8, 245, 281], noting, however, that our protocol employs only four states:

$$R = p^{0,0}_{det}\left(1 - h_2(e_Z) - e_Z h_2\left[\frac{1+(e_X-e_Y)/e_Z}{2}\right] - (1-e_Z)h_2\left[\frac{1-(e_X+e_Y+e_Z)/2}{1-e_Z}\right]\right) \tag{6.7}$$

where $h_2(\cdot)$ is the binary entropy function, and e_X and e_Y are the phase error rates of the virtual qubits $\bar{A}\bar{B}$ in the X and Y Pauli bases. These can be understood as the probability of the virtual qubits being projected into the incorrect Bell states given a passing announcement from Eve:

$$\begin{aligned} e_X &= \frac{\langle\Gamma|\left[\,(|\Psi^-\rangle\langle\Psi^-| + |\Phi^-\rangle\langle\Phi^-|)_{\bar{A}\bar{B}} \otimes |P\rangle\langle P|_Z\,\right]|\Gamma\rangle}{\langle\Gamma|(|P\rangle\langle P|_Z)|\Gamma\rangle} \\ e_Y &= \frac{\langle\Gamma|\left[\,(|\Psi^+\rangle\langle\Psi^+| + |\Phi^-\rangle\langle\Phi^-|)_{\bar{A}\bar{B}} \otimes |P\rangle\langle P|_Z\,\right]|\Gamma\rangle}{\langle\Gamma|(|P\rangle\langle P|_Z)|\Gamma\rangle} \end{aligned} \tag{6.8}$$

Here, $|\Gamma\rangle$ denotes the joint state between $\bar{A}\bar{B}A'B'EZ$ after the AB portion of twisted purified state $U_{\bar{A}\bar{B}A'B'}|\zeta\rangle$ is sent to Eve. Note $\langle\Gamma|(|P\rangle\langle P|_Z)|\Gamma\rangle = p^{0,0}_{det}$. The six-state protocol key rate provides generally higher key rates than the Shor-Preskill key rate [298] because it takes into account correlations between the bit and phase error patterns.

Employing Eq. 6.4, the twisting operation in Eq. 6.6, and the unitary evolution $|m,n\rangle_{A,B} \to \sum_{z=P}^{F} |e^z_{m,n}\rangle_E |z\rangle_Z$, we find that $e_\pm = e_X \pm e_Y$ are linear functions with respect to the elements of Eve's Gramian matrix $\langle e^P_{m',n'}|e^P_{m,n}\rangle_E$, which are already known from Eq. 6.3. Additionally, these phase errors are linear with respect to matrix elements $\langle\gamma^{x',y'}_{m',n'}|U^{x',y'\,\dagger}_{A'B'}U^{x,y}_{A'B'}|\gamma^{x,y}_{m,n}\rangle_{A'B'}$, which are functions of the twisting operation we control. Since our task is to modify the twisting operation to boost the key rate, these elements form the optimization variables of our problem. We note these elements form the Gramian matrix of the twisted ancillary system states, which is a positive semidefinite (PSD) matrix by construction. They are constrained linearly by Eq. 6.5, since by construction when $(x,y) = (x',y')$, the twisting operations cancel to not affect the form of the real protocol signal states.

The additional benefit of choosing $e_\pm$ as our objective functions is that we overcome any nonlinear optimization introduced by $h_2(\cdot)$. We find that e_+ (e_-) only depends on $U_+ = U^{0,0\,\dagger}_{A'B'}U^{1,1}_{A'B'}$ ($U_- = U^{0,1\,\dagger}_{A'B'}U^{1,0}_{A'B'}$).

Intuitively, we have such a dependence since e_- involves the Bell states that underwent a bit flip, so Alice and Bob's bit values will be (0,1) and (1,0), and only those twisting unitaries will be used. Similarly e_+ reflects Bell states that did not undergo a bit flip, so twisting will only involve (0,0) and (1,1). Since the unitaries $\{U^{x,y}_{A'B'}\}$ can be defined independently of each other, the optimizations of $e_\pm$ can be decoupled. Finally, since $h_2(x \leq 1/2)$ is monotonic, optimization of the arguments $e_\pm$ is sufficient.

Taking stock, we have two independent objective functions $e_\pm$, which are linear with respect to our optimization variables $\langle \gamma^{x',y'}_{m',n'} | U^{x',y'\,\dagger}_{A'B'} U^{x,y}_{A'B'} | \gamma^{x,y}_{m,n} \rangle_{A'B'}$, the elements of a PSD matrix subject to linear constraints. Thus, these optimization problems take the form of semidefinite programs which can be solved numerically on a standard laptop in a few seconds using packages for Python [299, 300] or Matlab [301]. While previous literature on twisting operations had noted the opportunity for optimizing U [287], no explicit procedure was constructed. Here, we have closed this gap, increasing the practicality of utilizing a virtual twisting operation as a step in the security proof. In Appendix E.3, we provide complete details for framing the problem in terms of semidefinite programs.

A comment is in order regarding our ability to choose an optimistic and optimal purification to increase the key rate, as it is common in other QKD security proofs to assume Eve holds the purification and thus one might assume we need to choose the most pessimistic purification. Recall from Fig. 6.1 that the qubits from which secret key is extracted are $\bar{A}\bar{B}$. These qubits are entangled with the signal states AB, and with $A'B'$, since the signal states are mixed. We are assuming that the sources are imperfect, but not malicious, so Eve does not have access to $A'B'$. This means that we are able to choose the virtual state of $A'B'$ in the most optimistic manner, which is equivalent to using $A'B'$ as a shield system upon which we can apply twisting operations to yield private states with $\bar{A}\bar{B}$. After the initial states in AB are sent to Eve, who also holds system E, Eve then holds a partial purification of the key systems $\bar{A}\bar{B}$ as well, but she still does not hold the entire purification since she does not have $A'B'$. Indeed, for a general protocol we would need to determine the most pessimistic state Eve could hold, which would correspond to finding the most pessimistic form of her Gramian with respect to the secret key rate; however, as we are using the tilted four state protocol, we have from Eq. 6.3 that Eve's Gramian is fixed, so there is no room to modify the parameters of the secret key rate that depend on Eve. That is, we do not need to take the most pessimistic partial purification that Eve can hold, because we have already uniquely specified her Gramian. Thus, the only remaining free parameters come from the state of the shield system, which we have the benefit of treating optimally by applying the twisting operation in Eq. 6.6. Picking the optimal purification of virtual systems Eve cannot access has been used to advantage in QKD security proofs before, as in choosing the state for the fictitious quantum coin in [273].

6.4 Key Rate Results

The only requirement for applying our technique is that Alice and Bob's initial qubit signal states cannot fall in the same plane of the Bloch sphere, which is easy to satisfy in practice. Otherwise, our technique can handle quite general noisy state preparation: Alice and Bob need not prepare the same sets of states; they can send states with different probabilities; and, the noise channel applied to each state can be dependent on the state.

As a study of fundamentally achievable key rates, we consider the following two-parameter (δ, p)-model for the initial states. We suppose Alice and Bob attempt to prepare the states $\{|H\rangle, |V\rangle, |H\rangle+|V\rangle/\sqrt{2}, |H\rangle-i|V\rangle/\sqrt{2}\}$; however, each state is subject to a modulation error which we treat as a random variable. Given

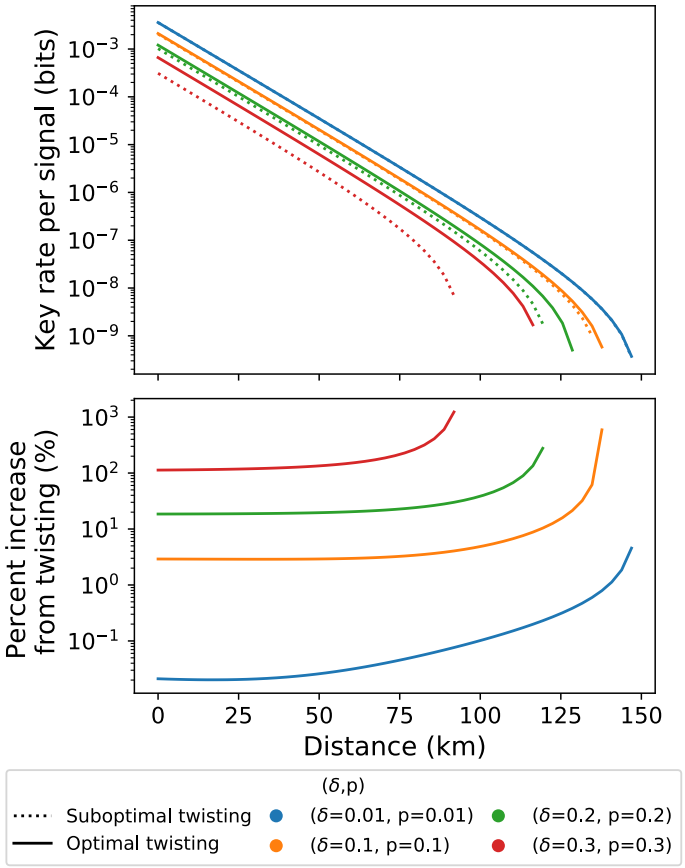


Figure 6.2: Top: Key rate vs. Alice-Charlie distance for various values of modulation error and depolarizing noise (δ, p) (same color indicates same model parameters). The dotted lines are the results from a suboptimal purification, while solid lines indicate our optimized key rates over all virtual twisting operations. Bottom: For each pair (δ, p), the percentage increase offered by optimizing over twisting operations. We assume a single photon source, symmetric distances from Alice and Bob to Charlie, and a Bell state detection scheme similar to [24], with overall detection efficiency of 50%, a dark count probability of 10^{-5} per pulse per detector, loss in fiber of 0.2 dB/km, and error correction efficiency of 1.

a state-dependent distribution for the modulation error on the surface of the Bloch sphere, the resulting average states can be treated as having a coherent modulation error, i.e. a constant offset angle from the ideal state parametrized by δ, as well as a depolarization noise parametrized by p, which introduces incoherent mixing to the states, shortening the Bloch vector. For exact definitions of the signal states, see Appendix E.5. For the case $p = 0$, we expect no improvement in our key rate over the standard loss tolerant protocol, since no mixing implies no virtual ancillary shield system.

In Fig. 6.2, we plot the asymptotic key rate found using our technique as a function of distance for various pairs (δ, p). For comparison with the key rate produced with our optimization, we plot the key rate calculated using a suboptimal purification, which was constructed by diagonalizing Alice and Bob's signal states and having $A'B'$ index the eigenvalues in decreasing order. We find that our technique provides a modest increase over the "naive" purification, our technique's advantages being most significant as the depolarizing noise gets stronger (making the initial states more mixed), and at longer distances when the untrusted channel noises (loss and dark counts) accrue. Additionally, we see a better key rate can be produced by reducing state preparation noise; however, once one has improved as best possible, our technique provides an optimized

key rate given that level of noise. That is, our technique provides confidence that one has optimized over all possible ancillary states of the purification that are consistent with the protocol statistics without worry that one has chosen a pessimistic virtual picture.

6.5 Conclusion

We have presented an extension of the proof technique from [25] to quantify the security of MDI QKD protocols that employ general noisy qubit signal states. We first reframed the analytical technique used for characterizing the parameters in the secret key rate that depend on Eve's system, noting that this new approach lends itself clearly to a generalization for higher-dimensional signal states. Next, we observed that employing trusted but mixed signal states means Alice and Bob have not a single but a set of virtual pictures they can use to analyze security in their protocol; we observed this was equivalent to Alice and Bob employing a *virtual* shield system onto which they can apply *virtual* twisting operations to minimize Eve's knowledge of the key [284, 285]. Finally, we provided a simple numerical technique leveraging semidefinite programming to optimize over all twisting operations to optimize the six-state protocol secret key rate formula, examining the implications for state preparation subject to random modulation error.

While this chapter focused on noisy signal states, we note that we assumed the noise kept the signals within their original Hilbert space; that is, the qubit became a mixed state, but remained within a two-dimensional subspace. In the next chapter, we examine the issue of source side-channels—scenarios in which extra information is sent along with each signal, thereby breaking the assumption that the states remain entirely within a qubit space.

Chapter 7

Measurement Device-Independent QKD with Time-Dependent Source Side-Channels

This chapter is based on [26], co-authored with Amita Gnanapandithan, Li Qian and Hoi-Kwong Lo. The work was collaborative, and I share first-authorship with Amita Gnanapandithan, who initiated the project by identifying the side-channel we use as a case example. Li Qian and Hoi-Kwong Lo supervised the project. My main contributions were to the quantum optical modelling in Section 7.3.2, application of the security proof technique, numerical key rate simulations, and writing various sections of the manuscript. The work benefited from helpful discussions with Thomas Van Himbeeck, Ignatius William Primaatmaja, Emilien Lavie, and Wenyuan Wang.

7.1 Introduction

Although measurement-device-independent (MDI) quantum key distribution (QKD) closes all side-channels in the detectors [24], imperfections of the quantum state source continue to threaten security. To partly address this challenge, the loss tolerant protocol [25] provides a proof technique for dealing with state preparation flaws, with extensions of the proof available to account for the decoy state method [282], and mixed states [43]. The method from [277] can also treat flawed sources, under the condition the encoded signals remain confined to a qubit space. Unfortunately, these methods to deal with state preparation flaws only account for systematic errors in the two-dimensional degree of freedom that Alice and Bob intentionally encode, meaning these techniques are not sufficient to account for source side-channels. On this front, recent security proof techniques have been developed to deal with the source leaking decoy state parameters [275, 302], and encoding information, with analytic approaches given in [276, 294] and numerical techniques in [27, 274].

With such security proof techniques now available, it is time they be applied to develop practical strategies for MDI QKD protocols employing realistic sources, bringing closer together the gap between idealized security proofs and experimental realities. In this chapter, we study a common optical source for polarization-based MDI QKD which relies on a Faraday mirror for polarization stabilization [282, 303–307]. We determine that this experimental setup introduces a passive optical side-channel due to leakage light between optical pulses

being unintentionally modulated in a time-dependent manner, a loophole that has not been identified in the literature to the best of our knowledge. Previous works focus on active side channels (side channels introduced by Eve), such as Trojan horse attacks [308, 309]. Hence, there exists minimal work on computing secure key rates in the presence of such passive side channels, despite these side channels being harder to avoid.

In this particular passive side channel case, we are faced with the seemingly daunting task of incorporating optical states distributed over a continuum of temporal modes into a security proof. However, we find that the versatile proof technique from [27] can be employed even in this scenario, a modest extension of its already wide applicability. As a numerical approach, the technique from [27] is particularly well-suited to our task, as it allows one to integrate detailed information about the initial states sent by Alice and Bob (including the time-dependent side-channels) and all the observed detection statistics in the protocol as constraints in the security proof.

Using the Faraday mirror source as a representative case example of sources with time-varying side-channels, we calculate the secret key rate under various assumptions and scenarios to better determine strategies for mitigating the information leaked via the side-channel. In particular, we investigate how the model for the state of the side-channel can have a significant impact on the amount of key generated, reinforcing the importance of carefully characterizing the optical output of the source. We present a few practical strategies for increasing the key rate, such as using all available detection statistics, sending more states than what would be required in the ideal protocol, and optimizing the choice of which test state to send from the Bloch sphere. As part of this analysis, we determine that while the MDI three-state protocol [25, 310] yields the same key rates as the MDI BB84 protocol [24, 311] in the ideal case of no side-channel, in the presence of leakage light these two protocols diverge, with BB84 being the more advantageous choice.

We briefly summarize why we will employ the numerical approach of [27] based on semi-definite programming to address the side-channel problem. Note that some other approaches such as the loss-tolerant protocol approach [25] and uncharacterized qubit approach [277] cannot be applied to the side-channel problem because those approaches assume the optical source sends out a qubit and such a qubit assumption is violated by side channels.

On the other hand, approaches such as [274] and the reference state approach [276, 294] do work for side channels. Nonetheless, we find that technique from [274] relaxes the task of bounding the phase error to a linear program and, therefore, it gives a less strict result than using the approach in [27]. As for the reference state technique [276, 294], we find that it gives a worse key rate for the MDI version of BB84 protocol in the presence of side-channels than the approach in [27]. For these reasons, we find that the approach in [27] is highly suitable for addressing the side channel problem.

The structure of this chapter is as follows: in Section 7.2, we review two necessary components for calculating key rates—the decoy state method and security proof technique from [27]—and compare [27] to competing proof techniques [25, 274, 276, 277, 294] to justify our choice of approach. Then, in Section 7.3, we study the case example of a polarization-based MDI QKD that employs Faraday mirrors for phase stabilization; here, we identify a side-channel due to leakage light, provide quantum optical modelling of the side-channel, and link this initial state information to the security proof technique. Finally, in Section 7.4, we provide key rate results for various protocol scenarios, determining the impact of the source model, and finding a divergence between the three-state and BB84 protocols in the presence of a side-channel. The takeaway from our work is that identification and careful characterization of side channels can bring secret key rate calculations more in line with experiment, and that the proof technique should be used to develop tailored strategies, such as sending a seemingly redundant state, to mitigate the information leakage of the

side channel.

7.2 Background

To understand the dependence of the secret key rate on the side-channel, we first provide some background on the components that are required for the key rate calculation. In Section 7.2.1, we review the decoy state method, and in Section 7.2.2 we review the proof technique from [27]. In Section 7.2.3, we compare our choice of proof technique to other potential options we could have chosen.

7.2.1 Decoy State Method

When Alice and Bob encode their secret key in a single photon degree of freedom, photon number splitting attacks are a method for an eavesdropper to exploit multiphoton output of the optical source to learn the secret key [273]. As a consequence, only single photon detection events are usable to characterize the amount of information the eavesdropper has about the key; however, multiphoton events can still be used to characterize the correctness of the key. The decoy state method allows Alice and Bob to characterize the photon number statistics of the eavesdropper-controlled channel, and in turn bound security based on the detection events that arose from the single photon components of the source's optical output [272].

Practically, the decoy state method for MDI QKD with polarization encoding consists of Alice and Bob each preparing phase-randomized weak coherent pulses (WCPs) with varying intensities [312]. Each pulse is polarized according to the protocol, e.g. BB84 [24] or three-state [25]. In this case, we can write the photon number distribution of Alice and Bob's states as:

$$p(m,n|k,l) = \frac{e^{-(\mu_k+\nu_l)}\mu_k^m \nu_l^n}{m!n!}, \tag{7.1}$$

where k (l) refers to Alice's (Bob's) optical intensity setting μ_k (ν_l). This assumes that the intensity settings are completely independent of the polarization basis and bit setting choices. If they each use N intensity settings (typically three is sufficient), then for given basis (i,j) and bit (x,y) choices, they have N^2 linear equations for the detection probabilities as a function of photon number:

$$Q_{k,l}^{i,j,x,y} = \sum_{m,n} \frac{e^{-(\mu_k+\nu_l)}\mu_k^m \nu_l^n}{m!n!} p_{\mathsf{pass},m,n}^{i,j,x,y}. \tag{7.2}$$

Here, $Q_{k,l}^{i,j,x,y}$ is the observed probability of Eve announcing that a round passed given that Alice and Bob chose intensity settings (k,l) along with basis and bit choices (i,j,x,y). $p_{\mathsf{pass},m,n}^{i,j,x,y}$ is the probability that the round passes due Alice (Bob) sending m (n) photons, and given basis and bit choices (i,j,x,y). Since we assume that the intensity setting choices are independent of the basis and bit choices, and that phase randomization of the WCP is perfect, $p_{\mathsf{pass},m,n}^{i,j,x,y}$ is independent of (k,l). As we review in Appendix F.1, these N^2 linear equations can be used in a linear program to determine upper and lower bounds on all the detection probabilities due to single photon components of the optical output $p_{\mathsf{pass},1,1}^{i,j,x,y}$. This allows us to bound the relevant detection statistics for computing security.

Using the decoy method, a lower bound on the secret key rate in an MDI QKD protocol is provided by [312]:

$$R \geq p_{\mathsf{pass},1,1}^{0,0}[1 - h_2(e_{ph,1,1})] - Q_{N,N}^{0,0} h_2(E_{bit}), \tag{7.3}$$

where $h_2(\cdot)$ is the binary entropy function. $Q^{0,0}_{N,N}$ is the detection probability of outcomes that generate raw key, and is given by:

$$Q^{0,0}_{N,N} = \sum_{x,y} Q^{0,0,x,y}_{N,N} \tag{7.4}$$

where we choose, without loss of generality, $(i,j) = (0,0)$ to be the key generation basis and $(k,l) = (N,N)$ to be the key generation intensities. E_{bit} is the bit error rate of the raw key, given by:

$$E_{bit} = \frac{\sum_{x \neq y} Q^{0,0,x,y}_{N,N}}{Q^{0,0}_{N,N}} \tag{7.5}$$

$p^{0,0}_{\texttt{pass},1,1}$ is the detection probability due to the single photon components of Alice and Bob's optical output:

$$p^{0,0}_{\texttt{pass},1,1} = \sum_{x,y} p^{0,0,x,y}_{\texttt{pass},1,1}. \tag{7.6}$$

Finally, $e_{ph,1,1}$ is the phase error rate of the protocol which we will more precisely define in the next section. Briefly, were we to consider a virtual picture of the protocol in which the single photon components of the optical output are entangled with qubits kept in Alice and Bob's labs, the phase error is the probability those qubits end up in a target Bell state up to a phase error. Like the single photon detection probabilities, it is not a directly observable quantity of the protocol and must be bounded. Were Alice and Bob to be able to perfectly prepare eigenstates of the conjugate basis to the key generation basis, then $e_{ph,1,1}$ can also be bounded with a simple linear program [312]. When the sources have preparation flaws amounting to constant polarization offsets, a series of linear programs are required (see Appendix A of [282]). In the case that the sources have preparation flaws and have a side-channel, we can employ a more recent technique [27] for bounding the phase error rate that employs semidefinite programming; we review that technique in the next section.

7.2.2 Security Proof Technique Based on Semidefinite Programming

Semidefinite programs (SDPs) are a class of convex optimization problems that can be written in the form:

$$\begin{aligned} &\texttt{maximize} && f_0(G) = \mathrm{Tr}(A_0 G) \\ &\texttt{s.t.} && f_i(G) = \mathrm{Tr}(A_i G) \geq b_i, \;\; i = 1, \ldots, m \\ & && G \succeq 0 \end{aligned}$$

where G is a positive semidefinite (PSD) matrix (i.e. has non-negative eigenvalues) whose elements form the optimization variables of the problem. $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function we seek to maximize. $f_i : \mathbb{R}^n \to \mathbb{R}$ are the constraint functions, and b_i are the constraint bounds. Importantly, the objective and the constraint functions are all linear functions of the elements of G, with the coefficients of the linear functions contained in the matrices A_i. SDPs are increasingly being used for QKD security proofs [27, 279, 280, 289–291, 293, 313, 314], due in part to the availability of fast and mature numerical implementations of solvers [299, 300].

In [27], the authors present a versatile numerical proof technique based on SDPs for MDI QKD protocols. The objective function of the SDP is the phase error of the key rate formula in Eq. (7.3), meaning an optimal solution provides a secure lower bound. The constraints are provided by the detection statistics of

the protocol, as well as initial state information, which is especially useful since it allows more experimental information to be used to quantify security. Here, we review the proof technique from [27] so that later we can apply it to the case of an MDI QKD protocol with a source side-channel.

To begin, we distinguish between the full optical states that Alice and Bob send to Charlie from the components of those states from which we will derive security. We will refer to the components of the optical states used to derive security as the *signal* states. As an example, in an ideal decoy state protocol, the full optical states are phase-randomized WCPs with different intensities, while the signal states used to derive security are the single-photon components. Alternatively, in an ideal phase-encoding protocol that does not use decoy states, the optical state and the signal state are one and the same.

Let the signal states that Alice prepares be denoted by $|\psi_x^i\rangle_A$, where (i, x) indicate her choice of basis and bit value. Analogously, we can write Bob's states as $|\varphi_y^j\rangle_B$, with his basis and bit choice given by (j, y). Like in [27], we will assume that the signal states are pure states; however, a path to treating mixed states is available via the technique from [43]. Note that, even though phase-randomized WCPs in an ideal decoy state protocol are mixed states, the signal states (single photon components) are pure states. At a high level, in an MDI QKD protocol, Alice and Bob send their signal states to Eve, who in turn makes an announcement z conditioned on a measurement she may or may not execute faithfully. For the sake of simplicity, we assume $z = P, F$, corresponding to a binary `pass` or `fail` outcome, but this can be generalized to account for more announcements.

Since quantum mechanics obeys unitary evolution, we can write the evolution of the joint state as:

$$|\psi_x^i \varphi_y^j\rangle_{AB} \to \sum_{z=P,F} |e_{x,y,z}^{i,j}\rangle_E |z\rangle_Z \tag{7.7}$$

where $|e_{x,y,z}^{i,j}\rangle_E$ are sub-normalized vectors that can be used to completely characterize Eve's state. Were we to know the states $|e_{x,y,z}^{i,j}\rangle_E$, we would have full knowledge of Eve's information about the key, and in turn could calculate the key rate exactly. Alas, the exact state of Eve's system is generally unknown; however, we can impose constraints on the state vectors $|e_{x,y,z}^{i,j}\rangle_E$. As we will review now, one can frame the secret key rate calculation as a semidefinite program, where the PSD matrix used as an optimization variable is the Gram matrix of Eve's vectors which we denote G_E. We recall that the elements of a Gram matrix for a set of vectors are all the pairwise inner products of the vectors, meaning G_E has elements $\langle e_{x',y',z'}^{i',j'} | e_{x,y,z}^{i,j}\rangle_E$. Gram matrices are always PSD.

The first type of constraint comes from the unitary evolution of states in quantum mechanics; namely, the inner product structure of the initial states must be preserved in the final states [27]. If Alice and Bob each have n_A and n_B basis choice settings, each basis choice associated with two bit choices, then the inner product constraint yields $(n_A \times n_B \times 2 \times 2)^2$ constraints of the form:

$$\langle \psi_{x'}^{i'} \varphi_{y'}^{j'} | \psi_x^i \varphi_y^j\rangle_{AB} = \sum_z \langle e_{x',y',z}^{i',j'} | e_{x,y,z}^{i,j}\rangle_E \,, \tag{7.8}$$

where the fact that the announcements are classical means $\langle z|z'\rangle_Z = \delta_{z,z'}$. Note that these constraints are linear in the elements of G_E

The next type of constraint on G_E comes from the observed detection statistics [27]. Let the probability of Eve announcing a successful detection event, conditioned on Alice and Bob having chosen basis and bit choices (i, j, x, y) be denoted by $p_{\texttt{pass}}^{i,j,x,y}$. In an ideal decoy state protocol, $p_{\texttt{pass}}^{i,j,x,y} \equiv p_{\texttt{pass},1,1}^{i,j,x,y}$, since the single

photon components are the signal states. This can be related to the elements of G_E as follows:

$$p_{\text{pass}}^{i,j,x,y} = \langle e_{x,y,P}^{i,j} | e_{x,y,P}^{i,j} \rangle_E \tag{7.9}$$

If the signal states are the same as the full optical states, then $p_{\text{pass}}^{i,j,x,y}$ would be directly observable in practice. Alternatively, if one is performing a decoy state MDI QKD protocol, then, as we saw in Section 7.2.1, one first uses the statistics of the full optical states in a linear program to establish upper and lower bounds on $p_{\text{pass}}^{i,j,x,y}$:

$$p_{\text{pass},L}^{i,j,x,y} \leq p_{\text{pass}}^{i,j,x,y} \leq p_{\text{pass},U}^{i,j,x,y}. \tag{7.10}$$

Thus, depending on the protocol, one either obtains $4n_A n_B$ equality constraints, or $8n_A n_B$ inequality constraints on G_E. Like the previous set of constraints, these are also linear in elements of G_E.

So far, we have identified the Gram for Eve's system G_E as a PSD matrix, as well as various linear constraints on its elements. We now review how to write the phase error as the objective function of an SDP. To start, we will assume that the basis choice $(i, j) = (0, 0)$ corresponds to the key generation basis. Moving to a virtual picture, we can think of Alice and Bob's signal states being entangled with virtual qubits $\bar{A}$ and $\bar{B}$ that they keep in their lab:

$$|\Psi_{virt}\rangle_{\bar{A}\bar{B}AB} = \sum_{x,y=0}^{1} |x\rangle_{\bar{A}} |y\rangle_{\bar{B}} |\psi_x^0 \varphi_y^0\rangle_{AB}, \tag{7.11}$$

where measurement of $\bar{A}\bar{B}$ in the computational basis yields the bit values of the secret key. Let the virtual state evolve to $|\Psi_{virt}\rangle_{\bar{A}\bar{B}AB} \to |\Gamma\rangle_{\bar{A}\bar{B}EZ}$ with:

$$|\Gamma\rangle_{\bar{A}\bar{B}EZ} = \sum_{x,y=0}^{1} |x,y\rangle_{\bar{A}\bar{B}} \sum_{z=P,F} |e_{x,y,z}^{0,0}\rangle_E |z\rangle_Z, \tag{7.12}$$

since we used Eq. (7.7).

Through the process of sending their key generation signal states to Eve (who conducts a measurement), as well as postselecting on $z = P$, Alice and Bob end up with a mixture of Bell states in the $\bar{A}\bar{B}$ virtual qubits. The virtual picture therefore allows us to formally define the phase error rate that characterizes security in the key rate in Eq. (7.3). Assuming, without loss of generality, that the target Bell state of the protocol is $|\Phi^+\rangle_{\bar{A}\bar{B}} = \frac{1}{\sqrt{2}}(|00\rangle_{\bar{A}\bar{B}} + |11\rangle_{\bar{A}\bar{B}})$, then the phase error rate is defined to be the probability that the $\bar{A}\bar{B}$ virtual qubits held by Alice and Bob end up in Bell states with the incorrect phase:

$$\begin{aligned} e_{ph} &= \frac{\langle\Gamma| \left(\mathbb{M}_{\bar{A}\bar{B}}^{-} \otimes |P\rangle\langle P|_Z\right) |\Gamma\rangle_{\bar{A}\bar{B}EZ}}{\langle\Gamma| \left(|P\rangle\langle P|_Z\right) |\Gamma\rangle_{\bar{A}\bar{B}EZ}} \\ &= \frac{1}{2} - \frac{Re\left(\langle e_{0,0,P}^{0,0} | e_{1,1,P}^{0,0}\rangle_E + \langle e_{0,1,P}^{0,0} | e_{1,0,P}^{0,0}\rangle_E\right)}{\sum_{x,y} p_{\text{pass}}^{0,0,x,y}}, \end{aligned} \tag{7.13}$$

with

$$\mathbb{M}_{\bar{A}\bar{B}}^{-} = (|\Phi^-\rangle\langle\Phi^-| + |\Psi^-\rangle\langle\Psi^-|)_{\bar{A}\bar{B}}. \tag{7.14}$$

Note that e_{ph} is a linear function of the elements of G_E as required for an SDP. With the additional constraint that $0 \leq e_{ph} \leq 1/2$, such that we are within the region where the binary entropy function increases

monotonically, then we can maximize e_{ph} via an SDP and determine a secure lower bound on the key rate using the Shor-Preskill formula [298] or the key rate from (7.3). Note that in a decoy state protocol without leakage light, $e_{ph} = e_{ph,1,1}$, since the signal states correspond to the single photons components.

In summary, [27] provides a method for obtaining a secure lower bound on the key rate using an SDP of the form:

$$\begin{aligned}
&\texttt{maximize} && e_{ph} \\
&\texttt{s.t.} && \langle \psi^{i'}_{x'} \varphi^{j'}_{y'} | \psi^{i}_{x} \varphi^{j}_{y} \rangle_{AB} = \textstyle\sum_z \langle e^{i',j'}_{x',y',z} | e^{i,j}_{x,y,z} \rangle_E \\
& && p^{i,j,x,y}_{\texttt{pass},L} \leq \langle e^{i,j}_{x,y,P} | e^{i,j}_{x,y,P} \rangle_E \leq p^{i,j,x,y}_{\texttt{pass},U} \\
& && 0 \leq e_{ph} \leq 1/2 \\
& && G_E \succeq 0.
\end{aligned}$$

With this technique in hand, we apply it to the case example of an MDI QKD source with a side-channel, which we will describe in the next section. First, however, we justify our choice of proof technique by comparing it with competing methods.

7.2.3 Comparison to Competing Proof Techniques

Given our review of the numerical approach from [27], it is worth comparing with competing proof techniques for MDI QKD to see why the approach we have chosen is well-suited to the problem we wish to study. The loss tolerant protocol [25] and the proof technique from [277] are both leading techniques for quantifying security in the presence of state-preparation flaws, the former requiring knowledge of the initial states, while the latter can simply use the detection statistics. However, in both techniques, one needs to assume that the optical source is outputting a qubit state, meaning they are insufficient to treat scenarios involving source side-channels, since the extra state sent with the encoded qubit, e.g. an optical coherent state, can easily break the assumption that the source only outputs states from a two-dimensional Hilbert space.

The numerical proof technique developed in [279, 280] also uses SDPs to compute the secret key rate; however, they work directly with the Devetak-Winter key rate formula [278], as opposed to the Shor-Preskill key rate [298]. Working with the Devetak-Winter key rate requires solving a series of SDPs, which is more cumbersome and numerically slower than a direct calculation of the phase error.

Two generalizations of the loss tolerant protocol have been developed to deal with non-qubit sources, in part for the purpose of studying source side-channels: the technique from [274] and the reference state technique [276, 294]. These techniques are directly comparable to [27] as they all use the Shor-Preskill key rate, with the core task of the proof being to find an upper bound on the phase error rate. While shown to perform more poorly than the reference state technique [294], the technique from [274] may be the most directly comparable to [27]. Both methods allow one to consider an arbitrary number of initial states that are not confined to a qubit space, and involve using all observed detection statistics. As we show in Appendix F.2, however, the approach from [274] relaxes the task of bounding the phase error to a linear program. The constraints of the linear program are provided by the detection probabilities and the initial states; however, fewer constraints are provided by the initial states than the approach from [27], since the overlaps between states with different (i, j, x, y) are not considered. Moreover, linear programs are a class of convex optimization problems contained within SDPs, meaning the constraints on the optimization variables used to compute the phase error are less strict than solving the full SDP as done in [27]. Thus, we expect the key

rates provided by the SDP numerical approach to be greater or equal to those calculated using the technique from [274] in general. In Appendix F.2, we provide the illustrative example of the three state protocol with a side channel to explicitly demonstrate the key rates provided by the SDP method outperform those from [274].

Finally, while the reference state technique is a purely analytic approach, a disadvantage is that it can currently only treat the case when Alice and Bob each send three states. The crux of the technique is to consider hypothetical detection statistics and phase error stemming from a fictitious set of reference states, and then bound the actual phase error of the protocol using the real detection statistics and the deviation between the reference and real states [276, 294]. In particular, in [294], the strategy for treating protocols involving four states, such as BB84, is to consider random alternation between two three-state protocols, where each of the X-basis BB84 states act as the third state. While the three-state protocol and BB84 yield the same key rates when the initial states are qubits [25], one of the observations that this work will provide is that information from the seemingly redundant fourth state of BB84 can provide extra constraints to boost the key rate in the presence of a source side-channel (i.e. when the qubit assumption is broken). Thus, a downside of using the reference state technique for protocols involving more than three states is that the key rate calculation will only ever be constrained by the statistics and initial states of three out of the four states, which leaves valuable information on the table, at the cost of higher key rate.

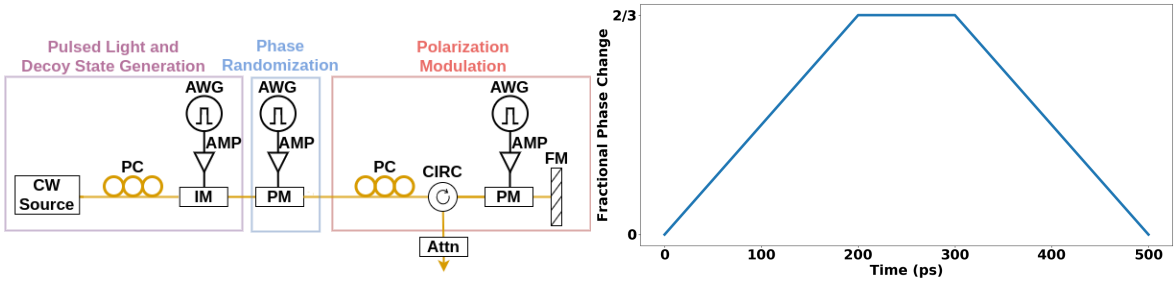
7.3 Source Side-Channels: A Case Example

Having reviewed the necessary components for the security proof, we now study a case example of an MDI QKD source with a side-channel. In Section 7.3.1, we identify a novel, time-dependent passive source side-channel which occurs when using a Faraday mirror for stable electro-optic bit modulation [282, 303–307]. We provide quantum optical modelling of the side-channel in Section 7.3.2, and in Section 7.3.3, we link the model to the security proof technique described in Section 7.2 to assess its impact on security while taking its time-dependent nature into account.

7.3.1 Origin of the Side-Channel

Several polarization and phase encoding implementations of MDI, prepare-and-measure, and plug-and-play QKD make use of an electro-optic phase modulator and Faraday mirror for optical bit modulation [304, 315]. The Faraday mirror is added, as shown in Figure 7.1a, to remove the temperature dependence of the phase modulator [282, 303–307]. Optical pulses first travel forward through the PM, co-propagating with a voltage pulse. During this first trip, they experience both voltage and unintentional temperature induced phase modulation. After reflection from the Faraday mirror, the pulses travel back through the PM, such that they do not collide with any counter-propagating voltage pulses. Hence, during this second trip, they only experience temperature induced phase modulation. Although the usage of a Faraday mirror drastically reduces state preparation flaws, we found that it creates a source side-channel.

This side-channel occurs due to the presence of weak light leakage between the optical pulses into which bits are encoded. These optical pulses are carved out from continuous wave light using an electro-optic intensity modulator (IM). Due to the finite extinction ratio of pulses that can be created with an IM, the presence of weak light leakage is inevitable. Of course, the phase of this leakage light is not intentionally modulated. In other words, no voltage is applied to the phase modulator as this light travels through it for the first time. However, after reflection from the Faraday mirror, some of this leakage light would inevitably



(a) (b)

Figure 7.1: (a) Experimental setup for polarization encoding MDI-QKD transmitter. An intensity modulator is used to create pulsed light (including decoy states) from a continuous wave light source. Then, the pulses go through a phase randomization unit, followed by a polarization modulation/encoding unit. PC - polarization controller, IM - intensity modulator, AMP - voltage amplifier, AWG - arbitrary waveform generator, PM - phase modulator, CIRC - optical circulator, FM - Faraday mirror, Attn - optical attenuator. (b) Time dependent fractional phase change applied to the leakage light. Fraction is with respect to the phase change applied to the corresponding encoded light.

collide with a counter-propagating voltage pulse within the phase modulator. Therefore, this leakage light would experience unintentional voltage-induced phase modulation, forming a source side-channel whose impact on security must be quantified.

We find the modulation of the leakage light to be time-dependent, as it travels in the opposite direction of the voltage pulse. It is also dependent on the voltage pulse shape used for phase modulation and the phase modulator electrode length. For our particular setup, the time dependence is shown in Figure 7.1b. Refer to Appendix F.3 for further details on how Figure 7.1b was derived.

Several simplifying assumptions regarding the optical and leakage signals are made when performing the security analysis.

1. In our particular setup (see Figure 7.1a), polarization encoding occurs after setting the decoy state intensity and performing phase randomization. Hence, we assume that the decoy state intensity and phase randomization of optical pulses (encoded signals) are independent of their polarization encoding.

2. We assume that the voltage pulses delivered to the polarization modulation PM are square pulses, such that the polarization of encoded pulses are time-independent.

3. In our experimental setup, there is a minimal correlation between the decoy state intensity and leakage light intensity. This correlation stems from the tails of the pulse shaping voltage pulses, which are small compared to the full duration of leakage light between pulses. Therefore, we assume that leakage signals carry no information about the decoy state intensity setting of encoded signals.

4. We assume that the voltage pulses strictly overlap with the optical pulses within the phase randomizing PM. Hence, the leakage signals carry no information about the phase randomization of encoded signals.

5. We assume that the leakage intensity is uniform in time, but the method could also easily treat time-varying intensity.

Note that these assumptions could be broken and could be incorporated into the security analysis using techniques from [275]. However, in our particular experimental setup, they are not the leading-order source of

information leakage, which we take to be the unintentional polarization modulation of the leakage light after reflection in the Faraday mirror. We can now proceed to model the quantum state of the source's output light.

7.3.2 Quantum Optical Modelling

To proceed with the security proof technique from Section 7.2.2, we must first model the states transmitted by the source so that we can compute the inner products of the signal states. The full optical state can be written as a separable state of Alice's and Bob's transmitted states: $\rho_A^{k,i,x} \otimes \sigma_B^{l,j,y}$. The parameters (k,l) refer to their decoy state intensity settings, (i,j) refer to their basis choice, and (x,y) refer to their bit choices. The basis and bit choices impact the polarization set with the polarization phase modulator, and are independent from the choice of intensity setting.

Alice's state can be further broken down into the side-channel state which represents the leakage light, and the encoding state which represents the optical pulses into which the basis and bit information are intentionally encoded:

$$\rho_A^{k,i,x} = \rho_{enc}^{k,i,x} \otimes \rho_{leak}^{i,x}. \tag{7.15}$$

We assume that the intensity modulation and phase randomization are timed with the optical pulses, and that in between pulses when the leakage light is passing through these modules, no phase randomization is applied, and the intensity modulation attenuates the light as much as possible to constant minimum but non-zero intensity. This has several consequences: first, this means that only $\rho_{enc}^{k,i,x}$ carries information about the intensity setting, and that $\rho_{leak}^{i,x}$ carries information about neither the random phase nor the intensity. This means we can treat $\rho_{enc}^{k,i,x}$ as a perfectly phase-randomized WCP just as in an ideal MDI decoy state protocol without leakage light. Second, our assumptions mean that we can treat $\rho_{leak}^{i,x}$ as a pure state:

$$\rho_{leak}^{i,x} = |\chi_x^i\rangle \langle \chi_x^i|_{leak}\,. \tag{7.16}$$

We now move to model the polarization module of the source and the time-dependent nature of the side-channel state. When $\rho_{enc}^{k,i,x}$ passes through the phase modulator, the controlling voltage pulse is timed with the optical signal such that the resulting polarization, determined by settings (i,x), is time-independent across the length of the optical pulse. By contrast, because $\rho_{leak}^{i,x}$ is travelling in the opposite direction, it acquires a time-dependent polarization. Let the creation operator for a photon at time t with polarization angles $\theta_x^i(t)$ and $\phi_x^i(t)$ be given by:

$$a_{t,i,x}^\dagger = \cos[\theta_x^i(t)] a_{t,H}^\dagger + \sin[\theta_x^i(t)] e^{i\phi_x^i(t)} a_{t,V}^\dagger, \tag{7.17}$$

where H and V denote the horizontal and vertical polarization modes, with the raising and lowering operators satisfying $[a_{t,m}, a_{t',m'}^\dagger] = \delta(t-t')\delta_{m,m'}$, $m = H, V$. As non-phase-randomized laser light, the leakage light at a given instant in time t can be treated as a coherent state with amplitude $\alpha_x^i(t)$, meaning over multiple times, the state can be written in general as:

$$|\alpha_x^i(t)\rangle = \exp\left\{\int dt[\alpha_x^i(t) a_{t,i,x}^\dagger - \alpha_x^{i\,*}(t) a_{t,i,x}]\right\} |vac\rangle\,. \tag{7.18}$$

In the source we are studying, we assume the leakage light has an effectively constant intensity $|\alpha_0|^2$ and polar angle θ, while the azimuthal angle ϕ changes with time based on the interaction with the phase modulator,

as shown in Fig. 7.1b; however, the technique we will apply could easily be used to study time-dependent intensity and polar angles as this would just modify the integral over time used to calculate the inner product between two side-channel states. The state of the leakage light associated with a given pulse is spread over multiple temporal modes, and is given by

$$|\chi_x^i\rangle_{leak} = \exp\left[\int_{-\Delta/2}^{\Delta/2} dt\, \alpha_0 a_{t,i,x}^\dagger - \alpha_0^* a_{t,i,x}\right]|vac\rangle\,, \tag{7.19}$$

where $a_{t,i,x}^\dagger$ here denotes the creation operator for a polarized photon with time-independent polar angle θ_x^i and time-dependent azimuthal angle $\phi_x^i(t)$ as in Fig. 7.1b, where the angles depend on Alice's basis and bit choices (i,x). Δ is the duration of the leakage light that contains encoding information, which from Fig. 7.1b is 500 ps.

In summary, we have that the output state of Alice's source can be treated as a tensor product of a perfectly phase-randomized WCP in the encoded mode with a time-varying pure coherent state in the side-channel mode. The time-varying polarization of the side-channel state depends on the basis and bit values chosen, but not the intensity setting or random phase used for the decoy state method. We can model Bob's source in the same way, denoting his encoded and leakage states by $\sigma_{enc}^{l,j,y}$ and $|\zeta_y^j\rangle_{leak}$, respectively. In the next section, we use these assumptions about the source, in connection with the decoy state method and proof technique reviewed in Section 7.2, to build up the security proof for this MDI QKD source.

7.3.3 Applying the Proof Technique

Given the model of the optical source, we can now connect it with the decoy state method and security proof technique from Section 7.2. To start, since the intentionally encoded states $\rho_{enc}^{k,i,x} \otimes \sigma_{enc}^{l,j,y}$ can still be treated as phase-randomized WCPs, and since the side-channel states $|\chi_x^i\rangle_{leak} \otimes |\zeta_y^j\rangle_{leak}$ carry no information about the decoy state intensity or random phase, we are able to use the decoy state method with only modifications to how we interpret the detection probabilities obtained by solving the linear programs.

The photon number distribution of the states $\rho_{enc}^{k,i,x} \otimes \sigma_{enc}^{l,j,y}$ follows the form from Eq. (7.1); however, when considering the linear equations provided by the detection probabilities in Eq. (7.2), $p_{\mathtt{pass},m,n}^{i,j,x,y}$ now refers to the probability a round passes given that Alice sent the m-photon component of the state $\rho_{enc}^{k,i,x}$, that Bob sent the n-photon component of the state $\sigma_{enc}^{l,j,y}$, *and* that they together sent the leakage states $|\chi_x^i\rangle_{leak} \otimes |\zeta_y^j\rangle_{leak}$. Note that the state of the leakage light does not depend on the number of photons Alice and Bob sent, just on the polarization encoding choice, so we can still use m,n to label the variable, even though it does not strictly refer to Fock states anymore. Moreover, since the leakage states are independent of the intensity choice settings, $p_{\mathtt{pass},m,n}^{i,j,x,y}$ remains independent of (k,l). Solving the linear program in Appendix F.1, Alice and Bob retrieve, for each basis and bit setting (i,j,x,y), bounds on the probabilities $p_{\mathtt{pass},1,1}^{i,j,x,y}$ that Eve will announce a successful detection event given that they each sent a single photon in the encoded mode *along with* the associated side-channel state.

Interpreting $p_{\mathtt{pass},1,1}^{i,j,x,y}$ as coming from both the single photon components of the encoded mode *and* from the leakage light, means that the definition of the signal states in this protocol no longer refers just to the single photon components of the encoded mode, as is the case for an ideal decoy state protocol. Connecting to Eq. (7.7), Alice and Bob's signal states are given by:

$$|\psi_x^i \varphi_x^j\rangle_{AB} = |\psi_x^i \varphi_y^j\rangle_{enc} \otimes |\chi_x^i \zeta_y^j\rangle_{leak}\,, \tag{7.20}$$

where $|\psi_x^i \varphi_x^j\rangle_{enc}$ refers to the single photon components of the phase-randomized WCPs $\rho_{enc}^{k,i,x} \otimes \sigma_{enc}^{l,j,y}$. In our description of the signal states, we have implicitly chosen that they be separable and uncorrelated from round to round, as we did not identify a mechanism by which such correlations would arise in the case example source considered here. A path to including correlations between pulses may be provided by [294], where the authors use the reference state technique to treat correlated sources by treating the correlations between pulses as a side channel state, and determining the deviation of that state from the ideal state of the protocol. With the signal states in (7.20), the detection probability constraints from Eq. (7.9) employ the probabilities $p_{\texttt{pass},1,1}^{i,j,x,y}$ which come from the signal states, i.e. the leakage light and the single photon component of the encoded mode. Additionally, the inner product constraints from Eq. (7.8) now include the inner products of the states of the leakage light; this reaffirms the versatility of the proof technique we are employing, since the constraints coming from the continuous-variable, time-dependent leakage light states can be coarse-grained down to their inner products, which form a finite-dimensional Gram matrix. The optimization variables (the elements of Eve's Gram G_E) and the objective function (the phase error rate) do not change; however, since the constraints will be affected by the presence of leakage light, the resulting key rate will certainly change.

With a model for the source, and how it connects to the decoy state method and the proof technique, we can now move to calculate the secret key rate for various scenarios and protocols.

7.4 Key Rate Results

Having reviewed the main components required for the security proof technique in Section 7.2, and having studied a case example of an MDI QKD source with a side-channel in Section 7.3, we now calculate key rates under different conditions. In Section 7.4.1, we provide the details of our how our simulations were performed. Then, in Section 7.4.2, we see how the key rate can change depending on the model chosen for the side-channel, including the time-dependent state we derived in the previous section, highlighting the need for careful side-channel characterization. In Sections 7.4.3 through 7.4.5, we explore various strategies that can be used to extract higher key rates, including using mismatch statistics, sending states that would be redundant under ideal conditions but which help in the presence of a side-channel, and choosing which test states to send. From these sections, we conclude that the three-state protocol and BB84, which yield equivalent key rates in the ideal case of no leakage light, have different key rates when a side-channel is present. These simulations should be used to better inform the choice of protocol when working with realistic sources like the one we are studying. While having only considered a specific source with a non-trivial side-channel, we expect the conclusions drawn to be broadly applicable to any leaky source; namely, we emphasize the importance of side-channel characterization, and determine which protocol parameters (e.g. number of states sent) can lead to key rate improvement.

7.4.1 Simulation Details

Recall we have two types of constraints to calculate the key rate: the inner product of the initial states, and the detection probabilities. Before detailing how we simulate these, we first comment on how different aspects of the source model affect these constraints.

To start, we observe that the single photon component of the encoded mode is the only quantity that affects both the inner product and detection probability constraints. However, whether that single photon state is directly sent from a single photon source, or is a component of a phase-randomized WCP is irrelevant to the inner product constraint; the type of source is only relevant to the detection statistics constraint,

since single photons undergoing a lossy channel will provide a different result at a threshold detector than phase-randomized WCPs.

Next, we assume the usage of gated detectors that would be timed to receive the encoded optical pulses. Hence, in our simulations of the detection events, we assume that the state of the leakage light has no impact on the overall observed detection statistics, since they are in temporal modes that are not picked up by the detectors and the already weak leakage light would be even less bright after the lossy channel. As a consequence, in the calculations we present, the side-channel state only affects the inner-product constraint. We note that the proof technique could easily accommodate the case of detection statistics being affected by the leakage light, as this would just be simulating different values of $Q_{k,l}^{i,j,x,y}$ in Eq. (7.2).

Since different aspects of the source model affect constraints in non-trivial ways, to better understand the key rate resulting from the source described in Section 7.3, we provide comparisons to other optical source models. Specifically, we consider:

- Single-photon vs. Phase-randomized WCP sources: when calculating the key rate for a given single photon component and side-channel state, i.e. for fixed inner product constraints in Eq. (7.8), how much is the key rate affected by those signal states being used directly vs. in a decoy state method?
- Sensitivity to the side-channel model: we assume in the detection simulations that the side-channel state has no impact on the observed outcomes, so the detection probability constraint in Eq. (7.9) remains fixed even if we change the model for the leakage light. In Section 7.3.2, we provided a model for the source which resulted in a time-dependent coherent state. Were we to change this model, how much does the key rate change?

In the sections that follow, we will consider these high-level choices of the model, in addition to varying more practically-rooted parameters like the intensity of the leakage light, plus the number and choice of encoded states sent.

For the choice of side-channel model, we compare three different approaches to treating the state of the leakage light:

- Model 1: full encoding information leaked. In this model, we assume the leakage light state is of the form:
$$|\chi_x^i\rangle_{leak} = \sqrt{\epsilon}\,|vac\rangle + \sqrt{1-\epsilon}\,|i,x\rangle \tag{7.21}$$
with $\langle i,x|i',x'\rangle = \delta_{i,i'}\delta_{x,x'}$. This model has been used in previous studies of QKD source side-channels [274, 276, 294]. This model makes a relatively pessimistic assumption, since any non-vacuum component of leakage light provides full-information, while we know, for example, that the single photon component would not be able to unambiguously encode all possible basis and bit choices (i,x).
- Model 2: time-independent coherent state. In this model, we assume the leakage light state is of the form:
$$|\chi_x^i\rangle_{leak} = |\beta\cos\theta_x^i\rangle_H \otimes |\beta\sin\theta_x^i e^{i\phi_x^i}\rangle_V\,. \tag{7.22}$$
The angles (θ_x^i, ϕ_x^i) are chosen to coincide with the polarization angles of the encoded mode. This model is more realistic in that we know the leakage light, as laser light, is in a coherent state; however, it does not account for the time-dependent nature of the polarization encoding in the leakage light, which has the opportunity to act to our advantage since not every instant provides Eve with full encoding information. Time-independent coherent state leakage light was considered in the context of Trojan horse attacks in [27].

- Model 3: time-dependent coherent state (multiple temporal modes). This model assumes the state from Eq. (7.19). It is our most accurate model of the source side-channel we introduced in Section 7.3. The inner product between two general, time-dependent coherent states is given by:

$$\langle\beta(t)|\alpha(t)\rangle = e^{-\frac{1}{2}\int dt\left[|\beta(t)|^2+|\alpha(t)|^2-2\beta^*(t)\alpha(t)\right]}, \tag{7.23}$$

  which we use to calculate the inner product of the side-channel states in Eq. (7.19).

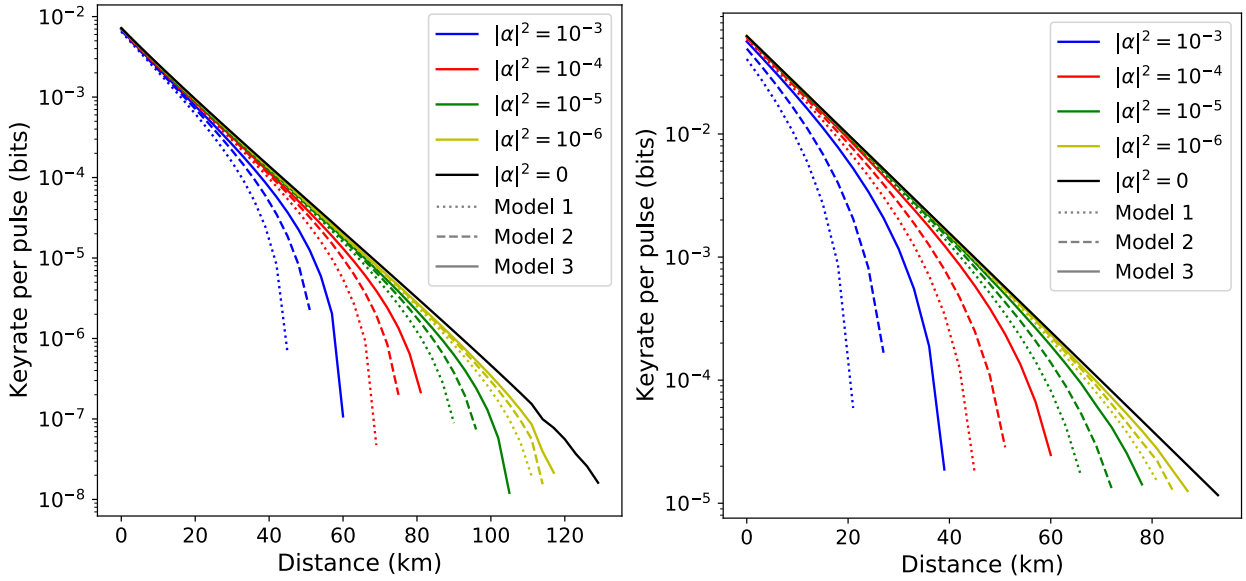
While we have three different models for the leakage light, we can still associate each of them to a fixed leakage light intensity, $|\alpha|^2$. For Model 1, we can set $\epsilon = e^{-|\alpha|^2}$. In Model 2, we can set $\beta = \alpha$, and in Model 3 we can set $\int dt|\alpha_0|^2 = |\alpha|^2$. This means all the models have the same vacuum probability, i.e. chance of sending no information to Eve, while the non-vacuum components carry varying amounts of information about the basis and bit values.

For the simulation of the detection statistics, in all our simulations we assume detection of a single Bell state using the detector setup from [24], with detector efficiency of 50%, dark count rates of 10^{-6} per pulse, and loss in fibre of 0.2 dB/km, with symmetric channel lengths from Alice and Bob to Charlie. To isolate the effect coming from the side-channel, we do not assume any misalignment in the source, but this could easily be added to the detection simulations. When simulating the decoy state method, we have Alice and Bob employ constant intensities of 0.05, 0.1 and 0.6; however, an additional layer of optimization for the decoy state intensities is possible, using our phase error calculation as a subroutine. The detection outcome probabilities $Q_{k,l}^{i,j,x,y}$ for the phase-randomized WCPs were simulated using the method from [312]. All of our calculations are in the asymptotic limit of infinite key length, and in the limit as the sifting rate goes to 1.

7.4.2 The Benefits of Side-Channel Characterization

Our main interest is determining how the key rate is affected by the presence of the side-channel. Here we investigate how the key rate changes depending on the model for the side-channel light, the intensity of the light, and on whether the encoded modes are sent as perfect single-photons source or as phase-randomized WCPs. In these simulations, we assume that Alice and Bob prepare BB84 states $\frac{|H\rangle\pm|V\rangle}{\sqrt{2}}$ and $\frac{|H\rangle\pm i|V\rangle}{\sqrt{2}}$, i.e. there are no encoding flaws.

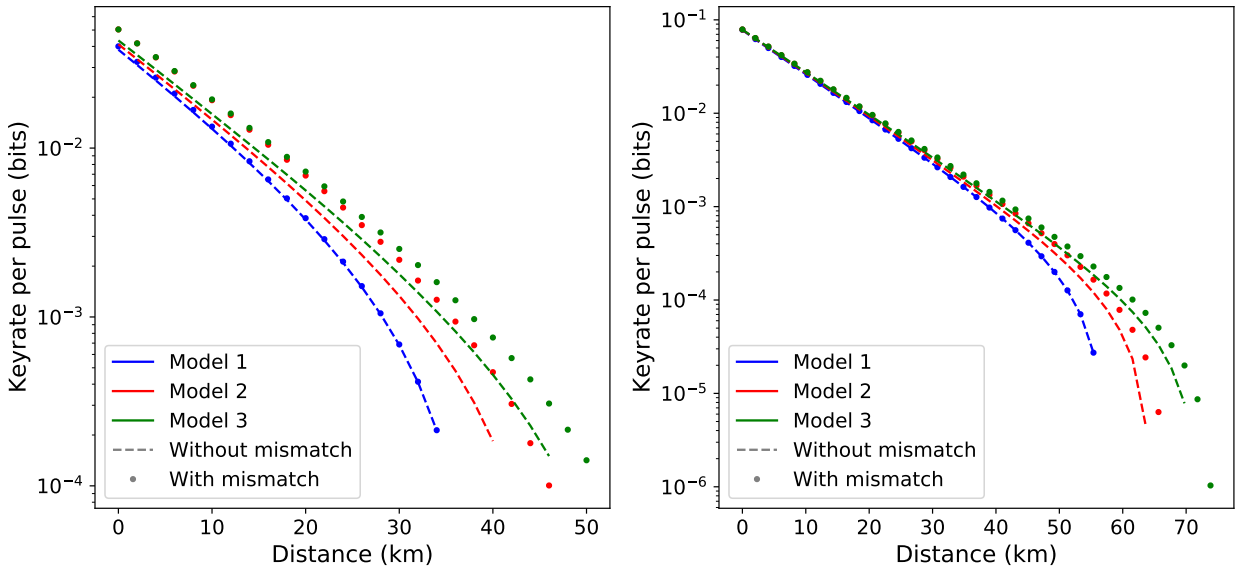
In Fig. 7.2 (a), we plot the key rate as a function of Alice-Charlie distance, assuming a decoy state protocol, for the three models of leakage light. Additionally, we vary the intensity of the leakage light across several orders of magnitude; using the lowest intensity signals from [316] as an order-of-magnitude reference for highly attenuated light, this places realistic leakage light intensity somewhere on the order of 10^{-6} to 10^{-4}. From this plot, we see that the most significant boosts in key rate come from a hardware solution of minimizing the intensity of the light in the side-channel; an order of magnitude reduction in intensity provides a greater improvement than refining the model of the leakage light state. However, there will likely always be some level of leakage light present between pulses. To mitigate this effect, it can be beneficial to carefully characterize the state of the side-channel. We see an improvement in the key rate when moving from Models 1 through 3, in that order. This reflects the intuition that the non-vacuum components of the states in these models carry diminishing levels of information about the basis and bit choices. In MDI QKD, we require that Alice and Bob have complete characterization of their sources but no characterization of the detectors; thus, if they know the state of the side-channel (or at the very least the pairwise inner products of all the initial states), it is straightforward to incorporate more information about the state by modifying the inner product constraints (a simple software solution), rather than making pessimistic assumptions about the



(a) Decoy method

(b) Single-photon source

Figure 7.2: Secret key rate as a function of Alice-Charlie distance for three different models of the leakage light. Model 1 corresponds to treating the leakage light as a superposition of vacuum (with amplitude $e^{-|\alpha|^2/2}$, same as a coherent state) and a state which leaks full encoding information. Model 2 corresponds to treating the leakage light as a coherent state $|\alpha_0 \cos\theta_x^i\rangle_H \otimes |\alpha_0 \sin\theta_x^i e^{i\phi_x^i}\rangle_V$ with the same polarization encoding parametrized by θ and ϕ as the signal state. Model 3 corresponds to treating the leakage light as a coherent state with total intensity $|\alpha|^2$, but with a time-dependent polarization, as given in Eq. (7.19). Across all models, $|\alpha|^2$ can be interpreted as the intensity of the leakage signal. For each model, we plot the key rate for various values of $|\alpha|^2$ which we indicate with different colours. In (a) we assume a decoy state protocol is used to characterize the single photon detection events, while in (b) we assume that the encoded modes of the signal state are perfect single photons.



(a) Single-photon source

(b) Decoy method

Figure 7.3: Secret key rate as a function of Alice-Charlie distance for Models 1-3 of the leakage light. Here, we investigate whether using all initial state inner products and detection statistics, as opposed to just the cases when Alice and Bob choose the same basis, benefit the key rate. We consider the case of BB84 with a preparation flaw, and a side-channel with $|\alpha|^2 = 10^{-4}$, for both (a) a single-photon source, and (b) the decoy state method. For (b), Alice and Bob each use four decoy intensities. We observe that for Models 2 and 3, the key rate benefits from considering all inner products and detection statistics available.

leakage light, as in Model 1, resulting in lower key rates.

Since the model for and intensity of the side-channel light has no bearing on the observed detection statistics, the detection constraints used to produce all the key rate curves in the figure are the same. As extra confirmation that the observed improvements in the key rate due to changing the model of the side-channel and the intensity of the side-channel light are independent of the observed detection statistics, in Fig. 7.2 (b), we provide the same key rate calculations, but assume a single-photon source for the encoded mode. We notice qualitatively the exact same trends as when using the decoy method, as expected.

7.4.3 Basis Mismatch Constraints

For our next investigation, we examine the role of the basis mismatch constraints in the security proof. Here, we are interested in knowing whether any advantage can be gained by using all the detection statistics and all the initial state inner products, including when Alice and Bob's bases do not match, as opposed to simply using the cases when the basis choices match ($i = j$). We observed that when Alice and Bob prepare the BB84 states perfectly, using the basis mismatch statistics and inner products did not produce an increase in the key rate, even in the presence of a side-channel. With perfect state preparation, we know that the conjugate basis statistics alone are strict enough constraints to provide the phase error when there is no leakage light, and we confirm numerically that this extends to the case when leakage light is present.

However, we know that when Alice and Bob have a preparation flaw for their states, i.e. a constant offset angle on the Bloch sphere, the mismatch statistics can help better characterize the key rate [25]. Since the Bloch sphere angle affects the associated side-channel state, the inner products, and the detection statistics,

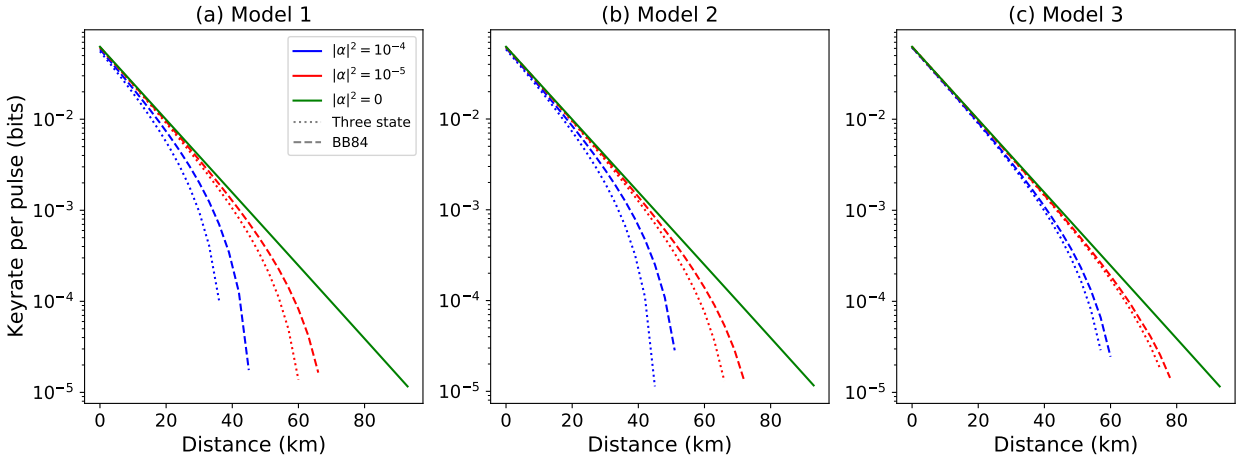


Figure 7.4: Secret key rate as a function of Alice-Charlie distance assuming a single photon source for the encoded mode states. All dotted (dashed) lines correspond to three state (BB84) type protocols, while different colors of lines correspond to different values of intensity $|\alpha|^2$; for $|\alpha|^2 = 0$, a solid line is plotted since the three state and BB84 protocols coincide. Here we see the advantage of sending the four BB84 states instead of using the three-state protocol. This trend is true for different orders of magnitude of $|\alpha|^2$ and across all leakage state models, depicted in (a)-(c). This in contrast to the ideal case of $|\alpha|^2 = 0$, where BB84 and the three-state protocol yield the same key rates.

it is more difficult to predict how the key rate will respond to a preparation flaw, and whether using full or partial detection statistics in the SDP constraints benefits the key rate. For these simulations, we use the preparation flaw model from Appendix D of [282], with the Bloch sphere offset angle parameter $\delta = 0.1$. In this case, we fix the leakage light intensity to $|\alpha|^2 = 10^{-4}$.

In Fig. 7.3 (a), we plot the key rate assuming a single-photon source for the encoded mode. For Model 1, we barely see any increase in the key rate when using full vs. partial detection statistics; this makes sense, since the non-vacuum component of the side-channel state leaks full encoding information, independent of Bloch sphere angle. For Models 2 and 3, we observe a boost in the key rate when using full detection statistics and inner products as constraints. This indicates that when one has a preparation flaw, and the side-channel state depends on the preparation flaw, it is best to use all information available from the detection statistics and initial state inner products.

In Fig. 7.3 (b), we consider the same situation but with a decoy state protocol. For this scenario, we add a fourth decoy with vacuum intensity, and observe an increase in the key rate when using full detection statistics and initial inner products as constraints in Models 2 and 3. Like before, we do not observe an increase in the key rate for Model 1. When we only considered three decoy intensities, we did not observe a meaningful increase in the key rate, likely because the three decoy intensities did not allow for tight enough constraints on the single photon detection statistics, so adding more detection statistics as constraints did not help since the constraints were too loose.

7.4.4 Sending Seemingly Redundant States Helps

It is known that the with only three out of the four BB84 states and using all the detection statistics, that one can produce the same key rate as using all the BB84 states [25]. Here, we are interested in whether the

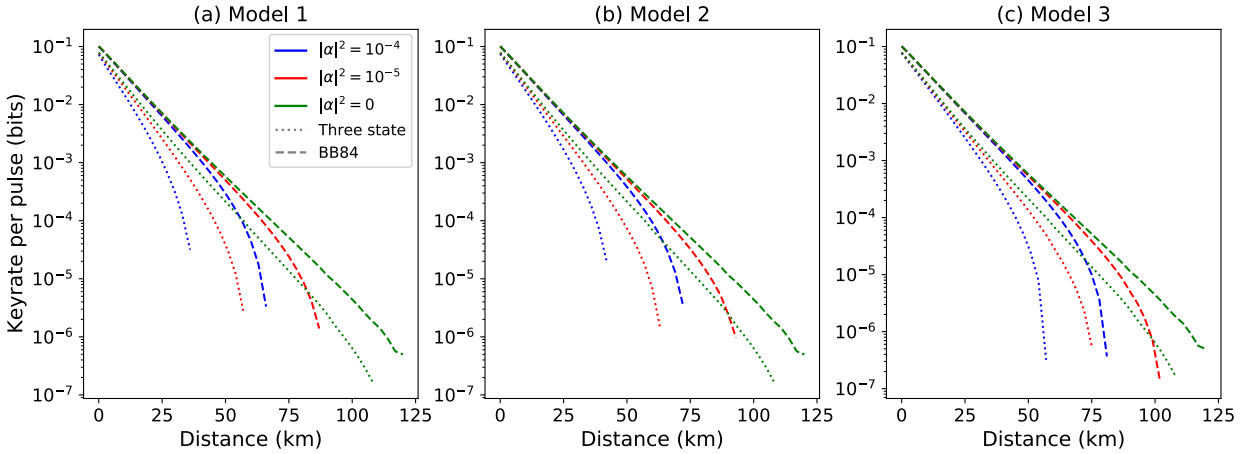


Figure 7.5: Secret key rate as a function of Alice-Charlie distance assuming a phase-randomized WCP and decoy state method. All dotted (dashed) lines correspond to three state (BB84) protocols, while different colors of lines correspond to different values of side-channel intensity $|\alpha|^2$. Like in Fig. 7.4, we see the advantage of sending the four BB84 states instead of using the three-state protocol. Even in the ideal case of $|\alpha|^2 = 0$, BB84 outperforms the three-state protocol, since the detection probabilities only offer inequality constraints, meaning the extra fourth state does offer extra constraint to increase the key rate. When the side-channel is present, we also see that adding an extra state in BB84 can go so far as to achieve a higher key rate than using the three-state protocol with an order of magnitude lower leakage light intensity.

same conclusion extends to the case of when the source has a side-channel. While the fourth BB84 state is redundant in the case that there is no leakage light, when a side-channel is present, the extra state can help characterize Eve's attack on the leakage light. We certainly would not expect the key rate to decrease by sending an extra state, as the extra state will only provide additional inner product and detection constraints to those already provided by the other three states.

In Fig. 7.4, we plot the key rate for a single photon source, examining all three models of leakage light, and a couple different intensities. Across all models and intensities (except for $|\alpha|^2 = 0$) there is an increase in the key rate when all four BB84 states are used as opposed to only three. In Fig. 7.5, we plot the key rates again, this time assuming a decoy state protocol. In this case, the divergence between using three or four states is even more pronounced. Even the $|\alpha|^2 = 0$ case observes a boost in the key rate, since the detection statistic constraints in Eq. (7.10) are inequalities when using the decoy state method, so the extra detection statistics from the fourth state are useful in this case. The key rate boost achieved from switching from three to four states is so pronounced that it can even do better than decreasing the intensity of the leakage light: using four states with a leakage light intensity of 10^{-4} provides a higher key rate than using three states with a leakage intensity of 10^{-5}.

The takeaway message from this analysis is that the three-state protocol is not as practical as BB84 in the presence of source side-channels. The extra resource savings of only having to use three states is undone by the loss of useful constraints that increase the key rate. We also simulated sending five and six states in the same plane of the Bloch sphere as the BB84 states to see if this provided even better key rates, but the key rate appeared to saturate with sending four states. Certainly sending additional states outside of this plane would increase the key rate, as expected from the six-state [281] or tilted four state protocols [25], but this would require additional polarization modulation in the source, whereas it is easier to only vary the

angle along one great circle of the Bloch sphere.

7.4.5 Choice of Test States Matters

Another example of divergence between ideal sources and sources with side-channels occurs in the choice of which test states to send. In the ideal case, if Alice and Bob prepare two orthogonal polarization states, they need only send one other state to achieve the same key rate as BB84 [25]; the location of that state on the Bloch sphere does not matter (as long as it is not the same state as the first two). Here we are interested to see whether this changes in the presence of a source side-channel.

To study this problem, we fix the channel distance, a leakage light intensity of $|\alpha|^2 = 10^{-4}$, fix Alice and Bob to send encoded single photon components $\frac{|H\rangle \pm |V\rangle}{\sqrt{2}}$ as two of their states, then vary the azimuthal angle of the other two states sent $\frac{|H\rangle \pm e^{i\phi}|V\rangle}{\sqrt{2}}$, and observe how the key rate changes. By symmetry, we need only vary $\phi \in [0, \pi]$.

In top of Fig. 7.6, we plot the results assuming a single photon source and a distance of 10 km. As expected, the key rate is independent of ϕ when there is no leakage light. In the presence of leakage light, $\phi = \pi/2$ still remains as the optimum test state to send, but the key rate drops off away from that point, most dramatically for Model 1. Interestingly, there is even a region for which Model 2 outperforms Model 3. To explain this, we can go to the Gram matrix formed by the initial states which form the constraints on the RHS of Eq. 7.8. If we calculate the trace distance between the Gram matrix of Model 2 and the Gram matrix created by the ideal qubit states $\{\frac{|H\rangle \pm |V\rangle}{\sqrt{2}}, \frac{|H\rangle \pm e^{i\phi}|V\rangle}{\sqrt{2}}\}$ as a function of ϕ, we find that it is symmetric about $\phi = \pi/2$; however, doing the same for the Gram matrix of Model 3, we find that the trace distance is not symmetric about that point due to the time-dependent nature of the underlying states and the way the inner product is calculated in Eq. 7.23. As ϕ increases, the Gram matrix of Model 3 eventually becomes a further distance from ideal than the Gram matrix of Model 2 for $\phi \gtrsim 0.8\pi$, so it is conceivable the key rate for Model 3 can perform worse in that region. Of course, the key rate depends on much more than just this trace distance, since the angle also changes the constraints provided by the detection statistics, but this provides some intuition as to why Model 2 can outperform Model 3 in certain regimes.

In bottom of Fig. 7.6, we plot the key rates assuming a decoy state method and a distance of 50 km. Here, even the case of zero leakage light has some sensitivity to the angle of the test state. $\phi = \pi/2$ is still the optimal test state across all models. Like before, there is a limited range of ϕ that provides a positive key rate in the presence of leakage light, with the range being narrowest for Model 1. We observed for both types of sources that the range of ϕ that yields positive key rate narrows as the channel distance is increased; this means that source preparation flaws, especially in the test state, become a greater problem at further distances, unlike in the case of no leakage light where there is greater stability of the key rate with respect to ϕ.

The main point of these simulations is to demonstrate that while the choice of test state is not so important when the source is ideal without side-channels, in the presence of leakage light, we must be careful to choose a test state that provides both good constraints on the encoded mode and on the leakage mode. While $\phi = \pi/2$ seemed to be the best choice for these models—coinciding with the BB84 states—we also observed cases when other values of ϕ produced the maximum key rate at a given distance. In a typical protocol, Alice and Bob simply choose the BB84 states and optimize the decoy state intensities as a function of distance; here, we see that in the presence of leakage light, there is additional benefit to optimizing over the polarization of the test states sent.

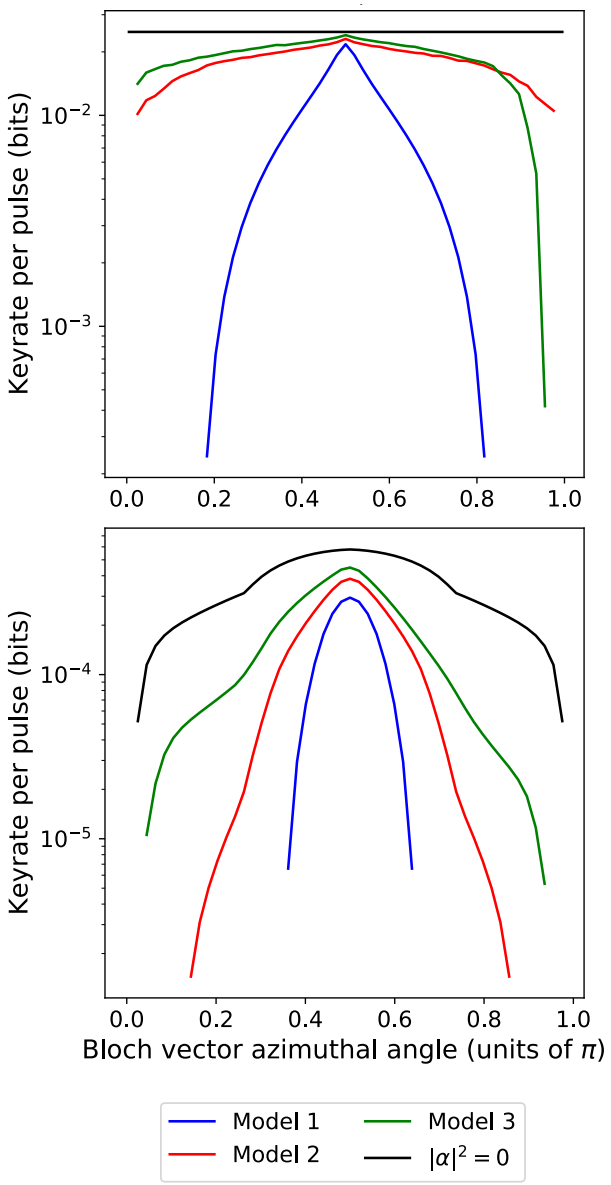


Figure 7.6: Key rate vs. azimuthal angle of the test states, for the case of no leakage light, and for leakage light with intensity $|\alpha|^2 = 10^{-4}$ treated with Models 1-3. The top (bottom) figure provides results for a single-photon source (a decoy state method) at a distance of 10 km (50 km). While the choice of test state is less relevant for the case of no leakage light, it can significantly decrease the key rate in the presence of leakage light, prompting the need to optimize which test states are used at a given distance.

7.5 Conclusion

In this chapter, we have examined the problem of source side-channels in MDI QKD. We reviewed the decoy state method and a recent, versatile proof technique based on semidefinite programming which allows for information about the state of the side-channel to be incorporated into the key rate calculation. With this in hand, we examined a case example of a common MDI QKD source that employs a Faraday mirror for polarization stabilization. For this source, we identified a non-trivial, time-dependent side-channel due to leakage light between encoded optical pulses, provided a quantum optical model of the output, and linked the components of the source model to the security proof techniques. We then examined multiple protocol scenarios to understand strategies for improving the secret key rate under practical circumstances.

We identified how the key rate calculation is affected by the information provided as constraints to the security proof. Most importantly, we saw how the model for the state of the leakage light can significantly impact the key rate, reaffirming that in MDI QKD security is derived in part from knowledge of the initial states sent by Alice and Bob, including any side-channel states. It is clear from our results that in practice one must carefully characterize side-channels, the reward of this work being higher key rates that come from not needing to take overly pessimistic assumptions of how much information is being leaked to the eavesdropper. On top of the importance of using the best available model for the side-channel, we found that in the presence of state-preparation flaws, Alice and Bob benefit from using all information at their disposal for the key rate calculation, i.e. all detection statistics, and all initial state information, rather than discarding cases when their basis choices do not match.

Having models for the state of the leakage light allowed us to develop concrete hardware strategies for mitigating the presence of the side-channel. Besides the obvious hardware improvement of simply suppressing the leakage light, two other physically implementable strategies emerged for when leakage light is present: first, although the three-state protocol promises the same key rates as BB84, when leakage light is present, Alice and Bob can get better key rates by sending all four BB84 states, as the statistics from the normally redundant fourth state actually help to better constrain Eve's attack on the side-channel. Second, while the choice of which test state to send from the Bloch sphere typically does not matter, here we find that in the presence of leakage light, some test states provide better key rates than others, indicating the advantage of optimizing which states to send as a function of distance. Even though we considered a representative case example, we expect that the strategies we developed to mitigate the side channel to be widely applicable to other leaky sources.

While this work examined strategies for treating source side-channels in MDI QKD, the source we considered had the advantage of not leaking information about the intensity setting choice and random phase value of the decoy state protocol, meaning we were able to use the decoy state method with only modifications to how we interpret the output of the linear programs in the security proof. An open problem is how to simultaneously mitigate more general source side-channels that leak information about both the encoding information, as we investigated, and the decoy state method intensity and phase parameters. It would be worthwhile to investigate merging the analysis presented here with the proof technique from [275] for treating intensity and phase information leakage.

Chapter 8

Conclusion

In this thesis we examined two photonic quantum technologies: quantum computers and quantum key distribution. In both parts of the thesis, an overriding theme was how to analyze and deal with ever-present noise. For photonic quantum computers, we saw a promising encoding was the Gottesman-Kitaev-Preskill (GKP) qubit, in large part due to its error correcting properties and ability to perform computation with the Gaussian operations that can be performed in photonic platforms. This thesis served to provide a feasible technique for preparing GKP states in optics, a framework for their analysis in the presence of realistic noise, and a blueprint for their utilization to perform fault-tolerant quantum computation.

In Chapter 2, we reviewed the formalism of GKP states and various figures of merit for assessing the quality of their realistic, finite-energy forms. From there, the chapter provided an exhaustive analysis of a promising photonic state preparation technique for GKP states based on Gaussian Boson Sampling (GBS). This work opens a path to producing GKP states in optics using current experimental resources, a valuable step forward given the many advantages of these bosonic qubits that we have discussed. Next steps for this work could be to add small feedforward operations conditioned on the observed photon number pattern to boost the probability of generation and fidelity, and to determine optimal decompositions of the GBS device in the presence of loss. Further into the future, research should be done towards deterministic generation of optical GKP states, as this would allow for greater error correction capabilities; some initial work is beginning in this area [317].

In Chapter 3, we developed a new mathematical formalism for studying and simulating bosonic qubits in the presence of realistic noise, leveraging the mature Gaussian CV phase space framework. This work opens the possibility to perform fine-grained, component-wise analysis of noise in near-term preparation, gates and measurements for bosonic qubits, aiding in device design. An open question for this work is whether the same computational speed-up observed for simulating true qubits, i.e. states living in a two-dimensional Hilbert space, under Clifford operations [136–140] can be replicated in the case of noisy bosonic qubits; such a result would be helpful for incorporating more realistic noise models for the fault-tolerance analysis of codes containing a few thousand bosonic qubits.

In Chapter 4 we reviewed an architecture for photonic quantum computing that employs GKP states and Gaussian resources, providing a tractable noise analysis, decoding strategy, and fault-tolerance thresholds. Putting together such a blueprint allowed us to identify areas where gaps still exist between current experimental capabilities and what would be required to implement the architecture, as was discussed in detail in the conclusion to that chapter. Already, [46] provides a substantial improvement to the hardware

requirements by removing the need for inline squeezing, an operation whose realistic implementation—measurement-based squeezing—we observed to be susceptible to noise in Chapter 3. Other desirable hardware improvements include boosting the level of achievable squeezing for on-chip squeezed light sources, and devising higher-probability and fidelity sources of GKP states. On the error-correction side, we note that the Raussendorf-Harrington-Goyal encoding was a starting point for choice of code, with our decoder relying on several heuristics, so we expect more tailored codes and decoding procedures to lower the thresholds required for state quality and loss.

The second part of this thesis examined quantum key distribution (QKD). Here, we focused on how idealized security proof techniques could be adapted to account for noisy devices used to implement the protocol. In Chapter 5, we looked at how the entropic uncertainty relations used to prove the security of high-dimensional, entanglement-based QKD were compromised by the measurement-range problem. Since the bounds were compromised by their dependence on the measurement operators associated with falling beyond the range of the detector, we instead provided an improved bound that replaced this dependence with a dependence on the probability of signals going undetected. While this essentially eliminated the problem for QKD schemes reliant on electric field quadrature encoding, for time/frequency schemes, even the refined bound was quickly insufficient due to channel loss increasing the probabilities of undetected signals. While it seems unlikely that entanglement-based protocols (i.e. untrusted sources but trusted detectors) will be tenable for time-frequency encoding, it could be worthwhile to devise measurement-device-independent (MDI) schemes (i.e. untrusted detectors but trusted sources) that can still leverage modern sources of time-frequency entangled photons.

The susceptibility of detectors in QKD schemes seen in Chapter 5 motivated the work on MDI QKD in the two chapters that followed. The chapters brought idealized security proofs closer to experimental reality by tackling two imperfections in the source: incoherent noise and information leakage. In Chapter 6, we examined the loss-tolerant protocol, generalizing it to account for Alice and Bob sending noisy, mixed signal states. As the technique was applied to qubits, e.g. polarization encoding, one area to investigate next would be to generalize the process for phase-encoded protocols, such as the popular twin-field QKD protocol [318]. Finally, in Chapter 7, we applied a recent numerical proof technique [27] to study a time-dependent, passive source side channel discovered in a common polarization-based MDI QKD implementation. After comparing various proof techniques to determine which one was best-suited to our task, we showed how incorporating the time-dependent nature of the side-channel into the security proof was beneficial to the key rate. As we found the technique from [27] to be so useful in this context, a worthwhile theoretical pursuit would be to explore extensions of the method to account for finite-size effects, and so that it is more directly applicable to CV QKD, prepare-and-measure and entanglement-based protocols. Further in the future, it would be worthwhile to create a full-stack numerical toolkit based on the method that can take as input the data characterizing the output states and/or detectors (depending on the protocol assumptions), along with the measurement data, and return a key rate, or, even better, suggest feasible modifications to the protocol—such as preparing an extra state or adjusting intensity—that would have the potential to boost the key rate.

With photonics as a platform, quantum technologies are an exciting prospect. I hope this thesis has served to bring their realization closer to fruition.

Appendices

Appendix A

Supplementary material for A Photonic State Preparation Method for Approximate GKP States

A.1 Notation, Nomenclature, Convention, and Units

We use this section to disambiguate some of the notation and conventions we use in the chapter.

Grid states. See Table A.1 for the notation we use for the different GKP states we refer to in the chapter.

Logical and physical gates. Gates with a bar, as in $\bar{U}$, always refer to logical gates acting on qubits, and gates without a bar, as in U, are always physical gates acting on the oscillator space.

Squeezing. The convention for the squeezing parameter r we use is such that the effect of the squeezing gate on the quadrature operators is

$$S^\dagger(r)\hat{q}S(r) = e^{-r}\hat{q}, \qquad S^\dagger(r)\hat{p}S(r) = e^{r}\hat{p}, \tag{A.1}$$

Category	Symbol	Description	Eq.
Ideal	$\|\psi_I\rangle$	Infinite energy square-lattice GKP states	(2.6)
Normalized	$\|\psi_G\rangle$	$G\|\psi_I\rangle$; ideal GKP states normalized with operator G	(2.13)
	$\|\psi_\epsilon\rangle$	$\|\psi_G\rangle$ with $G = E(\epsilon) \equiv e^{-\epsilon\hat{n}}$	(2.11)
	$\|\psi_{\Delta,\kappa}\rangle$	Gaussians of width Δ enveloped by a Gaussian of width κ^{-1}	(2.8)
	$\|\psi_\Delta\rangle$	$\|\psi_{\Delta,\kappa}\rangle$ with $\Delta = \kappa$; equals $\|\psi_\epsilon\rangle$ for small Δ	(2.10)
Approximate	$\|\psi_A\rangle$	Approximations to a given choice of normalizable states	(2.66)

Table A.1: Notation for various kinds of GKP states referred to in the chapter.

Δ_{dB}	ϵ
0	1
2	0.631
4	0.398
6	0.251
8	0.158
10	0.100
12	0.0630
15	0.0316
20	0.0100

Table A.2: Conversion between selected squeezing values expressed in decibels, corresponding to the width of the peaks in normalizable GKP states $|\psi_\Delta\rangle$ and the epsilon parameter in $|\psi_\epsilon\rangle$ from Eq. (2.11). The final column contains the error in the approximation between Δ and ϵ, expressed as the magnitude of the third-order term in the expansion of $\tanh\frac{\Delta}{2}$ (see Eq. (2.12)).

meaning that, in the position representation,

$$S(r)\psi(q) = e^{r/2}\psi(e^r q). \tag{A.2}$$

Squeezing values are often expressed in decibels in this chapter. The conversion we use for the parameter in the squeezing gate is

$$r = \frac{\ln 10}{20} r_{\text{dB}}, \tag{A.3}$$

and for the width Δ of the normalizable GKP state peaks, it is

$$\Delta = 10^{-\frac{\Delta_{\text{dB}}}{20}}. \tag{A.4}$$

Although we often work in the regime Delta $= \kappa$, it is κ^{-1} that is the width of the overall envelope in the $|\psi_{\Delta,\kappa}\rangle$ states. Therefore we are interested in low values of both Δ and κ (large positive in dB).

Lastly, we often use the envelope operator $E(\epsilon)$ from Eq. (2.11) to express our normalizable states. In the regimes we consider, it is true that

$$\epsilon \approx \Delta^2, \tag{A.5}$$

meaning

$$\epsilon \approx 10^{-\frac{\Delta_{\text{dB}}}{10}}. \tag{A.6}$$

In Table A.2 we list the conversion between a few squeezing values expressed in Δ_{dB} and ϵ.

A.2 Derivations

A.2.1 Dilation for the envelope operator $E(\epsilon)$

As stated in Sec. 2.2.2, a convenient method for constructing normalizable GKP states is by applying the operator $e^{-\epsilon\hat{n}}$ to the ideal GKP states. According to [1], this can be constructed as follows: first pass an

ideal GKP state through a beamsplitter of transmissivity t with a vacuum state:

$$\begin{aligned} |\psi_I\rangle|vac\rangle &= \int \frac{d^2\beta}{\pi}\langle\beta|\psi_I\rangle|\beta\rangle|vac\rangle \\ &\xrightarrow{\text{BS}} \int \frac{d^2\beta}{\pi}\langle\beta|\psi_I\rangle|t\beta\rangle|r\beta\rangle \end{aligned} \tag{A.7}$$

Then, post-select on measuring the second mode in the vacuum state to obtain

$$\begin{aligned} &\int \frac{d^2\beta}{\pi} e^{-|r\beta|^2/2}\langle\beta|\psi_I\rangle|t\beta\rangle \\ =&\int \frac{d^2\beta}{\pi}\langle\beta|\psi_I\rangle \sum_n e^{-|\beta|^2/2}\frac{\beta^n t^n}{n!}|n\rangle \\ =&t^{\hat{n}}|\psi_I\rangle, \end{aligned} \tag{A.8}$$

Finally, make the association $e^{-\epsilon} \to t$.

A.2.2 Readout from the logical subsystem

A measurement in the computational basis of a qubit can be effected with a binned homodyne measurement of the q quadrature of the corresponding GKP state. Here we show that this translates easily to the subsystem picture, meaning binned homodyne measurements allow direct access to the logical subsystem.

Just as with a measurement in the computational basis, a binned homodyne measurement must yield a binary output. Naturally, for the square lattice, the union of bins of width $\sqrt{\pi}$ centered at $2n\sqrt{\pi}$ ($(2n+1)\sqrt{\pi}$) corresponds to binary output 0 (1). We note that the POVMs corresponding to these bins can naturally be written in terms of displacements of the GKP states:

$$M_\mu = \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} d\beta X(\beta)|\mu_I\rangle\langle\mu_I|X^\dagger(\beta);\ \mu = 0,1; \beta \in \mathbb{R}. \tag{A.9}$$

Recall the decomposition from Eq. (2.22), and note that

$$X(\beta)|\mu_I\rangle = \begin{cases} |\mu\rangle_{\mathcal{L}} \otimes X(\beta)\,|+_I\rangle_{\mathcal{G}}\,, & \beta \in 0 \text{ bins}, \\ \bar{X}\,|\mu\rangle_{\mathcal{L}} \otimes X(\beta)\,|+_I\rangle_{\mathcal{G}}\,, & \beta \in 1 \text{ bins}, \end{cases} \tag{A.10}$$

taking care to realize that this is not a decomposition of $X(\beta)$ but only a statement about its action on ideal GKP states. Here, the 0 (1) bins refer to the q quadrature region of $[(2n-\frac{1}{2})\sqrt{\pi}, (2n+\frac{1}{2})\sqrt{\pi}]$ ($[(2n+\frac{1}{2})\sqrt{\pi}, (2n+\frac{3}{2})\sqrt{\pi}]$).

Thus, Eq. (A.9) can be rewritten as:

$$\begin{aligned} M_\mu &= |\mu\rangle\langle\mu|_{\mathcal{L}} \otimes \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} d\beta X(\beta)|+_I\rangle\langle+_I|_{\mathcal{G}}X^\dagger(\beta) \\ &= |\mu\rangle\langle\mu|_{\mathcal{L}} \otimes \mathbb{1}_{\mathcal{G}}, \end{aligned} \tag{A.11}$$

which is exactly a measurement on the logical subsystem in the computational basis.

U	$\tilde{E}(\epsilon) = UE(\epsilon)U^\dagger$
$X(\alpha)$	$e^{-\frac{\epsilon}{2}\left[(\hat{q}-\alpha)^2+\hat{p}^2\right]}$
$Z(\alpha)$	$e^{-\frac{\epsilon}{2}\left[\hat{q}^2+(\hat{p}-\alpha)^2\right]}$
$P(s)$	$e^{-\frac{\epsilon}{2}\left[\hat{q}^2+(\hat{p}-s\hat{q})^2\right]}$
$R(\phi)$	$E(\epsilon)$
SUM(g)	$e^{-\frac{\epsilon}{2}\left[\hat{q}_1^2+(\hat{p}_1-g\hat{p}_2)^2\right]}e^{-\frac{\epsilon}{2}\left[(g\hat{q}_1+\hat{q}_2)^2+\hat{p}_2^2\right]}$
$S(r)$	$e^{-\frac{\epsilon}{2}\left[e^{2r}\hat{q}^2+e^{-2r}\hat{p}^2\right]}$
$B(\theta,\phi)$	$e^{-\epsilon\hat{n}_1}e^{-\epsilon\hat{n}_2}$

Table A.3: Physical operators U and conjugated envelope operator $\tilde{E}(\epsilon) = UE(\epsilon)U^\dagger$ introduced in Eq. 2.11.

A.2.3 Conjugation and commutation with $E(\epsilon)$

In Table A.3, we display how the envelope operator $E(\epsilon)$ from (2.11) transforms under conjugation with Gaussian operations through $E(\epsilon) \to UE(\epsilon)U^\dagger$. In addition to this we might wish to see how the operations themselves change: $U \to E(-\epsilon)UE(\epsilon)$. Note that there are some mathematical difficulties in working with $E(\epsilon)$ that one should be wary of, since it is an exponential of an unbounded operator. For a discussion and a rigorous treatment of some the issues that arise, see, for example, [319].

To derive the commutation relations, we first show that, for any $k \in \mathbb{N}$,

$$\left[\hat{n}, \hat{a}^k\right] = -k\hat{a}^k \tag{A.12}$$

$$\left[\hat{n}, \hat{a}^{\dagger k}\right] = k\hat{a}^{\dagger k}. \tag{A.13}$$

For this, proceed by induction. First, using the identity $[AB, C] = A[B, C] + [A, C]B$, we can see that

$$[\hat{n}, \hat{a}] = \left[\hat{a}^\dagger\hat{a}, \hat{a}\right] = \hat{a}^\dagger\underbrace{[\hat{a}, \hat{a}]}_{0} + \underbrace{\left[\hat{a}^\dagger, \hat{a}\right]}_{-1}\hat{a} = -\hat{a}. \tag{A.14}$$

Now choose $j \in \mathbb{N}$, and assume $\left[\hat{n}, \hat{a}^{j-1}\right] = -a^{j-1}$. In light of the related identity $[A, BC] = [A, B]C + B[A, C]$, we have

$$\left[\hat{n}, \hat{a}^j\right] = \left[\hat{n}, \hat{a}^{j-1}\hat{a}\right] \tag{A.15}$$

$$= \left[\hat{n}, \hat{a}^{j-1}\right]\hat{a} + \hat{a}^{j-1}[\hat{n}, \hat{a}] \tag{A.16}$$

$$= -(j-1)\hat{a}^{j-1}\hat{a} + \hat{a}^{j-1}(-\hat{a}) \tag{A.17}$$

$$= -j\hat{a}^j. \tag{A.18}$$

To show $\left[\hat{n}, \hat{a}^{\dagger k}\right] = \hat{a}^{\dagger k}$, simply take the Hermitian conjugate of both sides of (A.12).

Let us now rewrite these commutation relations in a more helpful form:

$$\left[r\hat{n}, s\hat{a}^k\right] = (-kr)\left(s\hat{a}^k\right) \tag{A.19}$$

$$\left[r\hat{n}, t\hat{a}^{\dagger k}\right] = (kr)\left(t\hat{a}^{\dagger k}\right). \tag{A.20}$$

For relations that look like this, that is $[X, Y] = bY$, there is a braiding identity

$$e^X e^Y = e^{e^b Y}e^X. \tag{A.21}$$

Thus we may write

$$e^{r\hat{n}}e^{s\hat{a}^k} = e^{e^{-kr}s\hat{a}^k}e^{r\hat{n}} \tag{A.22}$$

$$\implies e^{r\hat{n}}e^{s\hat{a}^k}e^{-r\hat{n}} = e^{s\left(e^{-r}\hat{a}\right)^k} \tag{A.23}$$

$$e^{r\hat{n}}e^{t\hat{a}^{\dagger k}} = e^{e^{kr}t\hat{a}^{\dagger k}}e^{r\hat{n}} \tag{A.24}$$

$$\implies e^{r\hat{n}}e^{t\hat{a}^{\dagger k}}e^{-r\hat{n}} = e^{t\left(e^{r}\hat{a}^{\dagger}\right)^k}. \tag{A.25}$$

Now consider an arbitrary single-mode operator of the form

$$U = e^{p_k\left(\hat{a}^\dagger,\hat{a}\right)},$$

where p_k is a k-th degree polynomial. We can insert $E(-\epsilon)E(\epsilon) = \mathbb{1}$ between any two operators in this polynomial, implying that

$$E(-\epsilon)p_k\left(\hat{a}^\dagger,\hat{a}\right)E(\epsilon) \tag{A.26}$$

$$=p_k\left[E(-\epsilon)\hat{a}^\dagger E(\epsilon), E(-\epsilon)\hat{a}E(\epsilon)\right] \tag{A.27}$$

$$=p_k\left(e^{\epsilon}\hat{a}, e^{-\epsilon}\hat{a}^\dagger\right). \tag{A.28}$$

Again using the fact $E(-\epsilon)E(\epsilon) = \mathbb{1}$, we see that

$$E\left(-\epsilon\right)e^{p_k\left(\hat{a}^\dagger,\hat{a}\right)}E\left(\epsilon\right) = e^{E(-\epsilon)p_k\left(\hat{a}^\dagger,\hat{a}\right)E(\epsilon)} \tag{A.29}$$

$$= e^{p_k\left(e^{\epsilon}\hat{a}^\dagger, e^{-\epsilon}\hat{a}\right)}. \tag{A.30}$$

Therefore we conclude that the envelope conjugates across any operator in the form (A.30) – an exponential of an arbitrary polynomial of the creation and anihilation operators – at the expense of changing the inputs via

$$\hat{a} \to e^{-\epsilon}\hat{a} \tag{A.31}$$

$$\hat{a}^\dagger \to e^{\epsilon}\hat{a}^\dagger. \tag{A.32}$$

A.2.4 Measures of non-Gaussianity

One way to characterize genuine non-Gaussianity of a quantum state is through the negativity of the Wigner function [133, 320]. The Wigner function is a phase space quasiprobability distribution defined through

$$W_\rho\left(q,p\right) = \frac{1}{\pi}\int_{-\infty}^{\infty}\left\langle q+x\right|\rho\left|q-x\right\rangle e^{-2ipx}dx, \tag{A.33}$$

and the Wigner negativity is the area of the negative part of the Wigner function:

$$W\left(\rho\right) = \int dqdp\left|W_\rho\left(q,p\right)\right| - 1. \tag{A.34}$$

The Wigner logarithmic negativity is then

$$W_N\left(\rho\right) = \log\left[W\left(\rho\right)+1\right]. \tag{A.35}$$

Although pure states with a vanishing Wigner negativity must be Gaussian, there exist non-Gaussian mixed states with a positive Wigner function [321].

A.2.5 Glancy-Knill probability of no error

In [33], the authors find that GKP error correction can be performed perfectly with ideal GKP states displaced by less than $\sqrt{\pi}/6$. Thus, to determine the probability of successful error correction with normalizable GKP states, they first expand an arbitrary oscillator state in a basis of displaced ideal GKP states:

$$|\psi\rangle = \int_{-\sqrt{\pi}}^{\sqrt{\pi}} du \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} dv \langle u, v\,|\psi\rangle\,|u, v\rangle \tag{A.36}$$

where $|u, v\rangle = \pi^{-1/4} e^{-iu\hat{p}} e^{-iv\hat{x}} |0_I\rangle$ is reminiscent of the modular basis later explored in [322] and $\pi^{-1/4}$ is a normalization factor that ensures completeness:

$$\int_{-\sqrt{\pi}}^{\sqrt{\pi}} du \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} dv\,|u, v\rangle\,\langle u, v| = \mathbb{1}. \tag{A.37}$$

Then, the probability of finding $|\psi\rangle$ within a displaced region of less than $\sqrt{\pi}/6$, so that the errors introduced by the ancilla are small enough for the error correction to go through, is given by

$$P_{\text{no error}} = \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv |\langle u, v\,|\psi\rangle\,|^2. \tag{A.38}$$

Expanding, we find

$$\begin{aligned} P_{\text{no error}} = \pi^{-1/2} \sum_{s,t=-\infty}^{\infty} \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} dv\, e^{2iv(s-t)\sqrt{\pi}} \\ \times \int_{-\sqrt{\pi}/6}^{\sqrt{\pi}/6} du\, \psi^*(2t\sqrt{\pi} + u)\psi(2s\sqrt{\pi} + u). \end{aligned} \tag{A.39}$$

Performing the integral over v and taking $t \to t + s$ and $u \to u - 2s\sqrt{\pi}$, we recover Eq. (2.65). Eq. (A.39) can be interpreted as follows: For $s = t$, one gets the probability that $\psi(x)$ lies within $\frac{\sqrt{\pi}}{6}$ of all integer multiples of $2\sqrt{\pi}$ in position space, that is, where $|0_I\rangle$ has support. For $s \neq t$, one gets the probability that $\psi(x)$ lies within $\frac{\sqrt{\pi}}{6}$ of integer multiples of $2\sqrt{\pi}$ in momentum, without overcounting the correctable position regions.

It is straightforward to extend the formula to error correction with arbitrary mixed states:

$$\begin{aligned} P_{\text{no error}} = \frac{\pi}{3} \sum_{s,t} \operatorname{sinc}\left(\frac{\pi t}{3}\right) \times \\ \int_{\sqrt{\pi}(2s-\frac{1}{6})}^{\sqrt{\pi}(2s+\frac{1}{6})} du \rho(u, 2t\sqrt{\pi} + u). \end{aligned} \tag{A.40}$$

where $\rho(x, x')$ is the density matrix in the position basis.

A.3 Numerical Techniques

In addition to the strawberryfields and the walrus libraries, we employed packages from scipy libraries for special functions, numerical integration, and optimization algorithm implementations. Here we present details of our implementation of the algorithms.

A.3.1 Algorithm 7 details

Algorithm 7 Optimal Approximate States

function `cost`(r, $\boldsymbol{c}$, Δ, μ)
  initialize $|\mu_\Delta\rangle$
  $|\psi\rangle \leftarrow S(r)\sum_n c_n|n\rangle$
  return $|\langle\psi\,|\mu_\Delta\rangle\,|^2$
end function
procedure `optimization`(Δ, μ, $n_{\max}$)
  initialize r #squeezing within desired range
  initialize $\boldsymbol{c}$ #normalized vector of dim $n_{\max}$
  # basinhopping is a global search algorithm
  $r_{\text{opt}}, \boldsymbol{c}_{\text{opt}} \leftarrow$ basinhopping$[(r, \boldsymbol{c}), \texttt{cost}, args = (\Delta, \mu)]$
end procedure

To initialize the normalizable state $|\mu_\Delta\rangle$, we build a numerical wavefunction on a discretized position space as follows:

1. To be safe, we take 7 standard deviations of $1/\Delta$ as the range in q space so that the heights of the peaks at the edge of the range will be negligible.

2. We solve for the integer number of peaks separated by $\alpha = \sqrt{\pi}$ that fit into that range, and choose the number of points in the discretization of the range of q space to be 100 times the number of peaks, so that each peak is well-resolved.

3. With the array of q values in hand, we build the $|\mu_\Delta\rangle$ wavefunction by summing together Gaussians of width Δ centred at each of the $n\sqrt{\pi}$ within the range of q, and weighted according to the value of $\mu \in \{0, 1, +, H_+\}$.

4. We numerically integrate the function using the numpy trapz function to determine the normalization factor.

We define the `cost` function to be the fidelity between the target state $|\mu_\Delta\rangle$ and the approximate state of the form of a squeezed finite superposition of Fock states. We construct a numerical wavefunction for $|\psi\rangle = S(r)\sum_n c_n|n\rangle$. Using the same q array as the $|\mu_\Delta\rangle$ wavefunction, we sum the weighted q space wavefunctions for the Fock states (built using the scipy eval_hermite function). Then, we apply the squeezing using the convention from Eq. (A.2). The fidelity, $|\langle\mu_\Delta\,|\,\psi\rangle\,|^2$, is evaluated as a numerical integration with the numpy trapz function.

In the optimization algorithm, we employ the basinhopping algorithm available from the scipy optimize library. An overview of this algorithm is available in [17], with a more detailed description available in the library documentation and references therein [82, 83]. Broadly, basinhopping consists of two phases:

1. A local minimization (we minimize the negative of the `cost` function) over the optimization parameters $(r, \boldsymbol{c})$ given an initial guess, performed using the sequential least-squares programming method. We impose the constraint that $\boldsymbol{c}$ needs to be normalized and only has even components up to $n_{\max}$. This step produces a candidate for the global optimum (not a true global optimum).

2. A stochastic "hop" is performed in parameter space once a local minimum is found. After the hop is made, the local minimization phase is repeated with the new point in parameter space as the initial guess. The candidate global optimum is updated based on an acceptance test. We found 40 hops to be a suitable number.

When the algorithm terminates we are provided with an optimal squeezing parameter, r_{opt}, and a vector of coefficients, $\boldsymbol{c}_{\text{opt}}$, consistent with our choice of $n_{\max}$.

Importantly, we find that basinhopping is helpful: in some cases the optimal squeezing parameter and the truncated coefficients of the core state do not vary smoothly as we change Δ, even though the fidelity appears to be changing smoothly. That is, two or more different regions of parameter space can yield comparable fidelities, so if one region becomes better than another there is an apparent jump in parameter space. The global properties of the approximate states, such as the fidelity and the average photon number, still vary smoothly because the interplay between the squeezing and the core state coefficients combine to yield similar final states even if the optimal squeezing parameters and coefficients are undergoing discontinuous jumps. Put differently, even though the stellar representation [79] of a state is unique, when we apply a truncation to the core state, there can be many choices for states that meet a fidelity threshold to the target state. Thus, when one chooses the state with the highest fidelity among a collection of states that meet a fidelity threshold, the states can come from very different regions of parameter space even if Δ is only changing slightly.

To speed up the search for optimal parameters, for a given μ and $n_{\max}$, we pass the optimal results from the previous value of Δ as the initial guess for the next value. Since the basinhopping global search algorithm employs random jumps in parameter space, the initial guess does not exclude finding better local optima in other regions of parameter space.

A.3.2 Algorithm 8 details

Algorithm 8 Optimal Circuits for Approximate States

function `cost`($\boldsymbol{c}$, $\mathbf{x}$, $\bar{\mathbf{n}}$, handle)
  $|\psi_{\text{target}}\rangle \leftarrow \sum_n c_n |n\rangle$
  # $\mathbf{x} = (\boldsymbol{z}, \bar{\boldsymbol{\Theta}})$
  $|\psi_{\text{out}}\rangle \leftarrow \langle \bar{\mathbf{n}}|U(\bar{\boldsymbol{\Theta}})S(\boldsymbol{z})|\mathbf{0}\rangle$
  prob $\leftarrow \langle\psi_{\text{out}}|\psi_{\text{out}}\rangle$
  fid $\leftarrow |\langle\psi_{\text{out}}|\psi_{\text{target}}\rangle|^2$/prob
 if handle = 'global' **then**
  return fid + 0.1 × prob
 else if handle = 'local' **then**
  return fid + prob
 end if
end function
procedure `optimization`($\boldsymbol{c}$, $\bar{\mathbf{n}}$, N)
  # for an N-mode circuit
  # random squeezing and BS angles within ranges
  initialize $\mathbf{x}_0 = (\boldsymbol{\zeta}, \bar{\boldsymbol{\Theta}})$
  # basinhopping is a global search algorithm
  $\mathbf{x}_1 \leftarrow$ basinhopping[$\mathbf{x}_0$, `cost`, $args = (\bar{\mathbf{n}}$, 'global')]
  # `local_search` is a local maximization algorithm
  $\mathbf{x}_2 \leftarrow$ `local_search`[$\mathbf{x}_1$, `cost`, $args = (\bar{\mathbf{n}}$, 'local')]
end procedure
procedure `redecompose_circuit`(r,$\mathbf{x}_2$)
  # $|\psi_{\text{approx}}\rangle = S(r)\sum_n c_n|n\rangle$
  # squeezing only applied to mode with $|\psi_{\text{out}}\rangle$
  $\boldsymbol{r} \leftarrow (r, 0, ..., 0)$
  $\tilde{U} \leftarrow S(\boldsymbol{r})U(\bar{\boldsymbol{\Theta}})S(\boldsymbol{z})$
  # given a Gaussian unitary applied to vacuum, `euler` returns circuit parameters in the form $U(\bar{\Theta}')S(\boldsymbol{\zeta'})$
  $(\boldsymbol{z}', \bar{\boldsymbol{\Theta}}') \leftarrow$ `euler`($\tilde{U}$)
end procedure

To run Algorithm 8 for finding the optimal circuit-produced state relative to an approximate state $|\mu_A\rangle$, we supply the squeezing parameter, r; the core state Fock coefficients, $\boldsymbol{c}$, of $|\mu_A\rangle$; the number of modes for the GBS device; and a post-selection pattern, $\bar{\mathbf{n}}$, consistent with the need for the number of photodetections in $\bar{\mathbf{n}}$ to total the n_{max} of the core state. We loop Algorithm 8 over all choices of $\bar{\mathbf{n}}$ up to permutations of modes, as permutations can be absorbed into the interferometer.

There are three essential parts of Algorithm 8. First, we define a `cost` function which is called in the `optimization` procedure. Second, the `optimization` procedure returns the optimal circuit parameters (relative to the `cost` function) for producing the core state of $|\mu_A\rangle$, $\boldsymbol{c}$. Third, the `redecompose_circuit` procedure uses the optimal circuit parameters from `optimization` as well as the squeezing parameter, r, associated with $|\mu_A\rangle$ to find new GBS device parameters, where r is offloaded to the beginning of the circuit

instead of having to be applied in-line after the interferometer to the core state. The algorithm outputs are the parameters for a GBS device that produces a state given a post-selection pattern, $\bar{\mathbf{n}}$, along with the fidelity of that state to $|\mu_A\rangle$ and the success probability of obtaining that photon detection pattern.

The `optimization` procedure runs in two steps:

1. After initializing a random set of circuit parameters, $\mathbf{x}_0$, consistent with the number of modes (constraining the initial squeezed light source to within 12 dB to maintain realistically achievable values), we employ the basinhopping algorithm already described in App. A.3.1 to find a global candidate solution that maximizes the `cost` function labelled by the "global" handle. This global `cost` function is the fidelity between the core state of $|\mu_A\rangle$, $\boldsymbol{c}$, and the state output by the circuit given the post-selection pattern, $\bar{\mathbf{n}}$. Additionally, we add to the cost function the probability of the postselection pattern weighted by a factor of 0.1; if we solely optimized fidelity, we sometimes found solutions with probabilities smaller than 10^{-6}. We found the L-BFGS-B algorithm within basinhopping to be the most efficient local minimization algorithm in our case. We employed 50 hops to search the parameter space.

2. The optimal circuit parameters, $\mathbf{x}_1$, from basinhopping are passed to `local_search`, which employs the scipy minimize function to find the optimal solution to the `cost` function with the "local" handle. This local `cost` function is almost the same as the global `cost` function, but now the weight for the probability of the state is increased to 1 so that a solution in the neighbourhood of $\mathbf{x}_1$ with comparable fidelity but higher probability is returned. The scipy minimize function stops running when the first local minimum is found instead of the iterative process of random hops implemented in basinhopping. This change ensures only the neighbourhood of $\mathbf{x}_1$ is searched instead of the entire parameter space.

We now move on to how we calculate the `cost` function, including how to calculate the output of the GBS circuit. The `cost` function is given the $|\mu_A\rangle$ core state coefficients, $\boldsymbol{c}$, initialized as a vector in the Fock basis that will be used to compute fidelity; the squeezing and interferometer parameters; and the postselection pattern, $\bar{\mathbf{n}}$, for the GBS device. The algorithm proceeds as follows:

1. Using a strawberryfields engine we apply the squeezing to each mode followed by the interferometer in the rectangular decomposition. As the state of the modes at this point is still Gaussian, we use the gaussian backend for the engine, which returns the mean and covariance matrix for the resulting state.

2. The mean and covariance matrix, along with $\bar{\mathbf{n}}$, is passed to the state_vector function from the walrus library to get $|\psi_{\text{out}}\rangle$, the coefficients of the output state in the Fock basis (up to the dimension of $\boldsymbol{c}$), and before they have been normalized by the square root of the probability of obtaining $\bar{\mathbf{n}}$.

3. The probability is computed in the following way:

   (a) The mean and covariance matrix of all the modes before PNR measurements, as well as the indices of the modes we intend to measure, are passed to the strawberryfields reduced_gaussian function, which returns the mean and covariance matrix of ρ_{n-1}, the reduced $(n-1)$-mode state of all but the output mode. ρ_{n-1} is still Gaussian, as it is the partial trace of a Gaussian state.

   (b) Using the mean and covariance matrix of ρ_{n-1}, we can use the density_matrix_element function of the walrus and find the diagonal element $\langle\bar{\mathbf{n}}|\rho_{n-1}|\bar{\mathbf{n}}\rangle$ which is exactly the probability of finding the post-selection pattern $\bar{\mathbf{n}}$.

4. We normalize $|\psi_{\text{out}}\rangle$ with the probability obtained in the previous step.

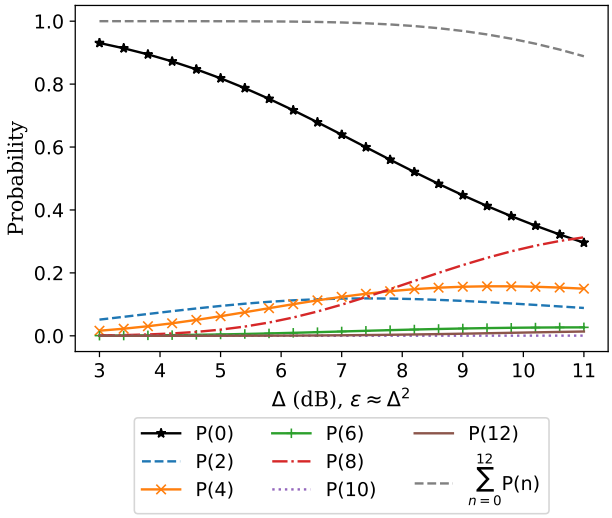


Figure A.1: Probabilities of different even Fock states in $|0_\epsilon\rangle$ up to $n = 12$ as a function of $\epsilon \approx \Delta^2$. We see that the probability of measuring 8 photons is higher than 6; thus, when constructing the GKP states from squeezed core states, increasing $n_{\max}$ of the core state from 4 to 6 provides little to no advantage compared to increasing $n_{\max}$ to 8.

5. We take the inner product between $|\psi_{\text{out}}\rangle$ and $\boldsymbol{c}$ in the Fock basis to compute the fidelity, and use the already calculated probability to compute the relevant `cost` function being called.

Finally, to implement `redecompose_circuit`, we find the Gaussian unitary associated with $S(r)U(\bar{\boldsymbol{\Theta}})D(\boldsymbol{\alpha})S(\boldsymbol{z})$ and use the strawberryfields GaussianTransform function to find $U(\bar{\boldsymbol{\Theta}}')D(\boldsymbol{\alpha}')S(\boldsymbol{z}')$.

A.4 Approximate GKP Z, X, and H Eigenstates

A.4.1 Further comments on $|0_A\rangle$

Our results for constructing the $|0_A\rangle$ state using $n_{\max} = 4$ and 6 were identical until $\Delta \approx 8$ dB, meaning the $n = 6$ component was not required; after that point, there was a jump in parameter space and the $n = 6$ state was used to attain better fidelity than $n_{\max} = 4$, albeit only slightly. Here we explain why it makes sense that the results for $n_{\max} = 4$ and 6 are so close, and why we see a small jump in parameter space for $n_{\max} = 6$.

First, we can use the $|0_\epsilon\rangle$ state to understand the distribution of the Fock coefficients for $|0_\Delta\rangle$, since the states are almost indistinguishable for $\epsilon = \Delta^2$. In Fig. A.1, we plot the probabilities of detecting the even Fock states in $|0_\epsilon\rangle$ up to $n = 12$ as a function of Δ, for $\epsilon = \Delta^2$. We see that, up to $\Delta \approx 8$ dB, the $n = 2$ and 4 states are the most probable states after the vacuum, and beyond $\Delta \approx 8$ dB, the $n = 8$ state becomes more probable. The probabilities for $n = 6, 10$, and 12 remain significantly smaller over the whole range. Thus, if one wants to increase the $n = 8$ component using only squeezing applied to a core state of $n_{\max} = 6$, one will also end up increasing the $n = 6$ component in the distribution. If one wants to keep the $n = 6$ component small, one option is to truncate the core state at $n = 4$; as it turns out, this is the optimal option found numerically up to $\Delta \approx 8$ dB, leading to similar results between $n_{\max} = 4$ and 6 in that regime.

We can also use the wavefunctions of the Fock states to understand why the $n = 2, 4$, and 8 components are weighted higher than $n = 6$. In Fig. A.2, we plot the wavefunction produced by the sum of the Fock

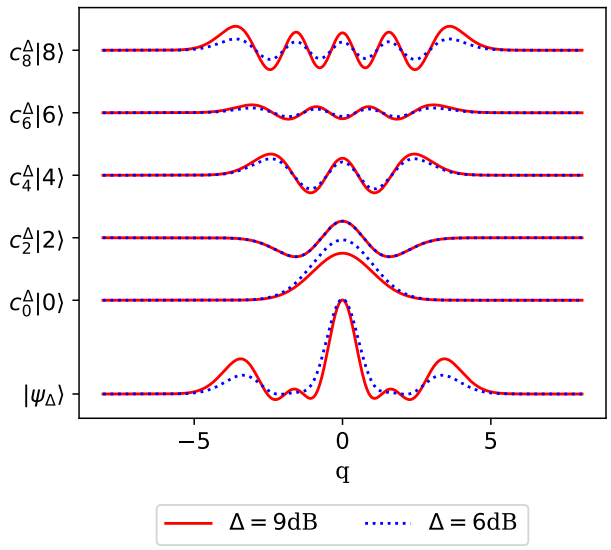


Figure A.2: Wavefunctions produced from summing the first eight Fock states in the distribution for $|0_\epsilon\rangle$ (using $\epsilon \approx \Delta^2$), as well as the wavefunctions for the weighted Fock states in the superposition. We see the important contribution of the 8 photon state to the outer set of peaks for the normalizable GKP state. Thus, without the 8-photon state in the core of $|0_A\rangle$, the $n = 4$ or 6 states are needed to build the outer peaks, leading to worse fidelities.

state components in $|0_\Delta\rangle$ up to $n = 8$, as well as the wavefunctions for the weighted Fock states in the superposition. We start to see the peak structure of the GKP wavefunctions – a central peak and two outer peaks at the correct locations – in addition to some wiggles in between due to the truncation. Importantly, we see that the outer set of peaks is created almost exclusively by the $n = 8$ Fock state, with the $n = 2$ and 4 states working to narrow the central peak. Thus, were one to construct $|0_\Delta\rangle$ by only using a core state up to $n = 6$ and then stretching the Fock wavefunctions through squeezing, building the outer peaks of the $|0_\Delta\rangle$ state would require stretching the $n = 6$ or the $n = 4$ state to align with the GKP grid. But using the $n = 6$ state and weighting it heavily would shrink the central peak, since the $n = 6$ state has a dip at $q = 0$. Thus, the better option for matching the symmetry of the $|0_\Delta\rangle$ state is to stretch the $n = 4$ state, which has outer peaks and a central peak. This is further evidence for why we see no difference between the results for $|0_A\rangle$ using $n_{\max} = 4$ and 6 for $\Delta < 8$ dB.

For $\Delta > 8$ dB, we see a small jump in the $(r, \boldsymbol{c})$ parameter space for the $n_{\max} = 6$ state, although the resulting wavefunction does not change much, leading to only slight jumps in the fidelity and average photon numbers. In that regime, all coefficients in the core state are now nonzero, and the squeezing applied to the core state as compared to $\Delta < 8$ dB is slightly greater. We understand this jump to come from the fact that the achievable fidelity to the target state is already relatively low ($\approx 90\%$) for this range of Δ; in other words, since the core state resource available from $n_{\max}$ is not too high, we are simply trying to find the best fidelity among several poorly contending regions of parameter space. There are two basins in parameter space that provide very similar fidelities when $\Delta = 8$ dB; as Δ increases, one basin overtakes the other to provide marginally higher fidelity, leading to a jump from the first basin to the second.

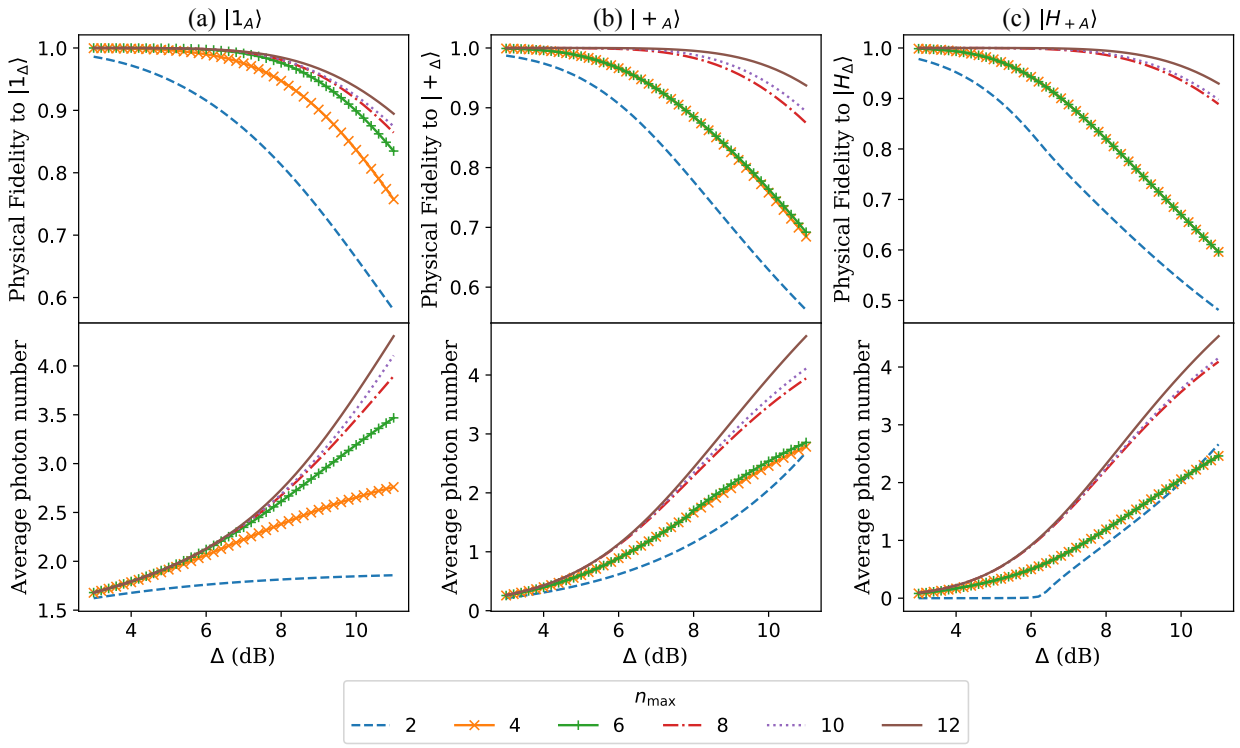


Figure A.3: Fidelity of approximate states $|\mu_A\rangle$ to normalizable states $|\mu_\Delta\rangle$; and photon number vs. Δ for (by column) $\mu =$ (a) 1, (b) +, and (c) H_+. Line colours/styles in each plot represent the results for different $n_{\max}$ in the core states of $|\mu_A\rangle$. The trend in all cases is that a greater $n_{\max}$ leads to better fidelity to the target state and higher energy; if the quality of the target state is improved by increasing the dB value of Δ, for a fixed $n_{\max}$, the fidelity gets worse, and the required energy gets higher. We expect comparable circuit resources (see Sec. 2.3.4) are required to construct approximate states for any point on the Bloch sphere.

A.4.2 Results for $|1_A\rangle, |+_A\rangle$, and $|H_{+A}\rangle$

In Fig. A.3, we provide the numerical results for the construction of the $|1_A\rangle$, $|+_A\rangle$, and $|H_{+A}\rangle$ states using Algorithm 7. All the trends for fidelity to the target states are comparable to those of the $|0_A\rangle$ states. The $|+_A\rangle$ state is effectively as resource-intensive to create as the $|0_A\rangle$ state given their relation by a Fourier transform. Finally, the $|H_{+A}\rangle$ state has squeezing and average photon numbers on the same order as $|0_A\rangle$ and $|+_A\rangle$, so it should also be comparatively demanding to prepare.

Notice that the results for the $|0_A\rangle$ in the main text and $|+_A\rangle$ states are nearly indistinguishable. Since the states are closely relatable by a Fourier transform [1], r for the $|0_A\rangle$ state is approximately $-r$ for the $|+_A\rangle$ state: one is squeezed in q and the other in p. Additionally, since the state $|n\rangle$ is an eigenfunction of the Fourier transform with eigenvalue $(-i)^n$, we expect the coefficients of the Fock states in the core state superposition to be related by phases of the form $(-i)^n$. In fact, we do see in our results for $|0_A\rangle$ and $|+_A\rangle$ that the core Fock states with $n \bmod 4 = 0$, for which $(-i)^n = 1$, have almost the same coefficients, while all the states wth $n \bmod 4 = 2$, for which $(-i)^n = -1$, have coefficients that differ effectively by a phase of -1. As a result of the states having such similar properties, their global properties end up being very close.

[1] They would be exactly relatable if we chose the $E(\epsilon)$ envelope since $E(\epsilon)$ commutes with phase space rotations.

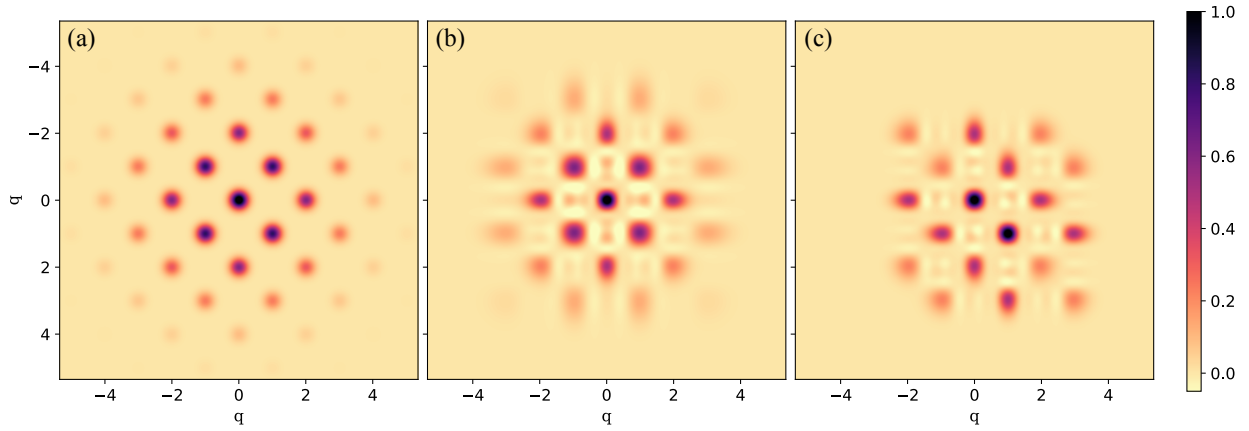


Figure A.4: Single-qubit codespace projectors $P(q, q')$ in q basis (in units of $\sqrt{\pi}$) for three cases: (a) The normalizable GKP states $|0_\Delta\rangle\langle 0_\Delta| + |1_\Delta\rangle\langle 1_\Delta|$, with $\Delta = 10$ dB. (b) Our approximate GKP states $|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|$, designed to target the normalizable states with $\Delta = 10$ dB, constructed using squeezing applied to a core state with $n_{\max} = 12$. Here we assume we prepare both $|0_A\rangle$ and $|1_A\rangle$. (c) Our approximate GKP states $|0_A\rangle\langle 0_A| + X(\sqrt{\pi})|0_A\rangle\langle 0_A|X^\dagger(\sqrt{\pi})$ with the same parameters as (b), except here we assume we prepare only $|0_A\rangle$ and create the logical 1 state by displacing $|0_A\rangle$. We see that the symmetries of the target projector are best preserved when preparing both $|0_A\rangle$ and $|1_A\rangle$, although that symmetry may not be required for encoding.

A.5 Additional Characterization of Approximate GKP States

A.5.1 Projectors

Projectors provide a valuable understanding as well as some visual intuition for how the code subspace behaves. Here we consider projectors defined by

1. The normalizable states $|\psi_\Delta\rangle$: $|0_\Delta\rangle\langle 0_\Delta| + |1_\Delta\rangle\langle 1_\Delta|$;

2. The approximate states $|\psi_A\rangle$: $|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|$;

3. $|0_A\rangle$ and displaced $|0_A\rangle$:
$$|0_A\rangle\langle 0_A| + X(\sqrt{\pi})|0_A\rangle\langle 0_A|X^\dagger(\sqrt{\pi}).$$

We consider the third kind of projector since one may want to prepare just one GKP Z eigenstate and generate the other through gate application. In Fig. A.4, we provide some examples of these projectors plotted in the q basis. We take $\Delta = 10$ dB, and choose the approximate states that are meant to target the $\Delta = 10$ dB states with core states of $n_{\max} = 12$. Note that the projector $|0_A\rangle\langle 0_A| + |1_A\rangle\langle 1_A|$ maintains the symmetry of the target state projector, while the displacement in $|0_A\rangle\langle 0_A| + X(\sqrt{\pi})|0_A\rangle\langle 0_A|X^\dagger(\sqrt{\pi})$ breaks it. As we know from the modular subsystem decomposition, a loss of such symmetry does not affect the logical encoding. Of course, because the approximate states only have $\sim 95\%$ fidelity to the target states, the sharp peaks get blurrier further away from the origin in q space.

A.5.2 Quantum error correction matrix

One object that encompasses the error-correcting properties of general codes is the quantum error correction (QEC) matrix [2, 15]. For an error map with Kraus operators $\{E_k\}$ and code states $|0_G\rangle$ and $|1_G\rangle$, the QEC matrix elements are defined by

$$\gamma_{kk'} \equiv P_G E_k^\dagger E_{k'} P_G, \tag{A.41}$$

where P_G is the projector constructed from the code states, as defined in App. A.5.

The 2×2 matrix γ can then be expanded in a basis of Pauli matrices:

$$\gamma_{kk'} = \varepsilon_0 \mathbb{1} + \varepsilon_x \sigma_x + \varepsilon_y \sigma_y + \varepsilon_z \sigma_z. \tag{A.42}$$

If γ is proportional to the identity, then there exists a recovery which can perfectly correct the error [2]. When this is not the case, errors can only be approximately corrected. However, the magnitude of the coefficients ε_i offer a convenient interpretation of the error's effects; namely, ε_x, ε_y, and ε_z are the probability of bit, bit-phase, and phase flip errors [15].

Ideal square-lattice GKP states were designed to correct displacement errors in phase space of up to $\sqrt{\pi}/2$. Thus, we examine the QEC matrix for displacement errors applied to the normalizable and approximate states. In Fig. A.5, we calculate $P_G D(\beta) P_G$ as a function of displacement, β, expand it in terms of the Pauli matrices, and then plot the magnitudes of ε_i as functions of β. We do this for four choices of states (that is, four different P_G): the normalizable GKP states with $\Delta = 10$ dB and the approximate GKP states for the same Δ but with $n_{\max} = 4, 8$, and 12.

The first row of plots in Fig. A.5 corresponds to the normalizable states $|0_\Delta\rangle$ with $\Delta = 10$ dB. We see that, for small displacements, the largest magnitude is ε_0, meaning the matrix is basically proportional to the identity, meaning these displacement are correctable. A displacement by $\sqrt{\pi}$ in q (p) leads to the dominant terms becoming ε_x (ε_z), corresponding to a bit (phase) flip. Further away from the origin, the blobs become more smeared, indicating that a displacement of that magnitude is more likely to lead to an uncorrectable error.

As we progress down the column, we see how the approximate states perform. The radii of the blobs become bigger as $n_{\max}$ decreases, meaning the probability of error increases. Moreover, the blobs further from the origin become quite smeared in phase space, meaning only the first multiples of $\sqrt{\pi}/2$ will be correctable.

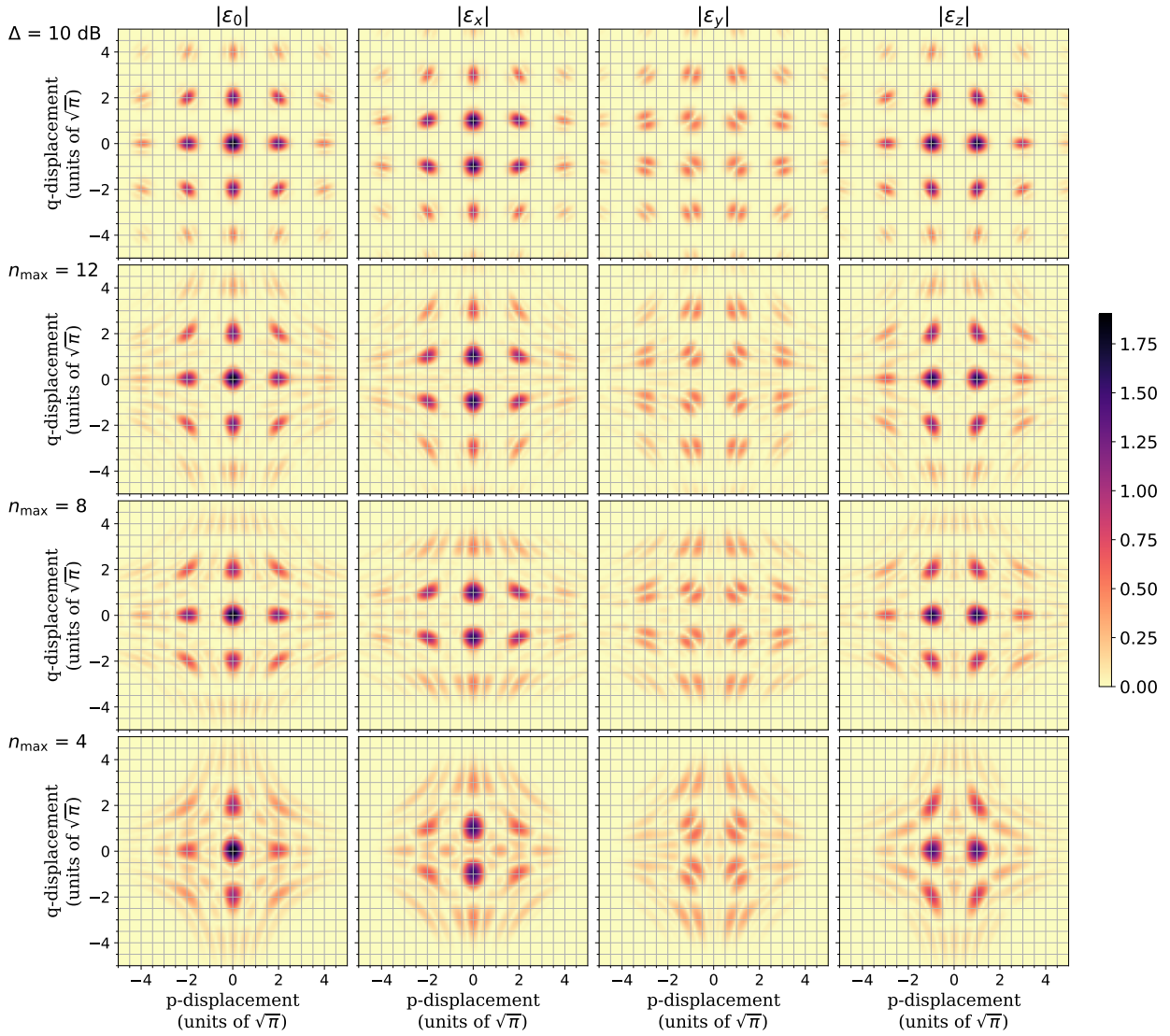


Figure A.5: Visualization of the quantum error correction (QEC) matrices (Eq. (A.41)) of normalizable and approximate GKP states, $|\psi_\Delta\rangle$ and $|\psi_A\rangle$, subject to displacement errors. Each row of subfigures corresponds to a different choice of code states. In each subfigure, the y (x) axis corresponds to a value of displacement along q (p) in phase space, so that a single point corresponds to a net displacement error. The QEC matrix associated with this displacement error can then be decomposed in terms of the identity and Pauli operators. Each column corresponds to the magnitude of the coefficients in the decomposition, which can be interpretted as no error, bit flip, bit-phase flip, and phase flip probabilities. The first row corresponds to $|\psi_\Delta\rangle$ with $\Delta = 10$ dB. The second through fourth rows correspond to $|\psi_A\rangle$ meant to approximate $|\psi_\Delta\rangle$ with $\Delta = 10$ dB with core states of $n_{\max} = 12$, 8, and 4 photons.

Appendix B

Supplemental material for Fast Simulation of Bosonic Qubits via Gaussian Functions in Phase Space

B.1 Coefficients of Ideal GKP

For ideal GKP qubits in the state $|\psi\rangle = \cos\frac{\theta}{2}|0\rangle_{\text{gkp}} + e^{-i\phi}\sin\frac{\theta}{2}|1\rangle_{\text{gkp}}$, the coefficients for the Wigner function are provided by [53]:

$$c_{k,\ell}(\theta,\phi) = \begin{cases} \frac{1}{4\sqrt{\pi}} & \text{for } k \bmod 2 = 0,\ \ell \bmod 2 = 0, \\ \frac{1}{4\sqrt{\pi}}\cos\theta & \text{for } k \bmod 4 = 0,\ \ell \bmod 2 = 1, \\ -\frac{1}{4\sqrt{\pi}}\cos\theta & \text{for } k \bmod 4 = 2,\ \ell \bmod 2 = 1, \\ \frac{1}{4\sqrt{\pi}}\sin\theta\cos\phi & \text{for } \begin{cases} k \bmod 4 = 3,\ \ell \bmod 4 = 0, \\ k \bmod 4 = 1,\ \ell \bmod 4 = 0, \end{cases} \\ \frac{-1}{4\sqrt{\pi}}\sin\theta\cos\phi & \text{for } \begin{cases} k \bmod 4 = 3,\ \ell \bmod 4 = 2, \\ k \bmod 4 = 1,\ \ell \bmod 4 = 2, \end{cases} \\ \frac{-1}{4\sqrt{\pi}}\sin\theta\sin\phi & \text{for } \begin{cases} k \bmod 4 = 3,\ \ell \bmod 4 = 3, \\ k \bmod 4 = 1,\ \ell \bmod 4 = 1, \end{cases} \\ \frac{1}{4\sqrt{\pi}}\sin\theta\sin\phi & \text{for } \begin{cases} k \bmod 4 = 3,\ \ell \bmod 4 = 1, \\ k \bmod 4 = 1,\ \ell \bmod 4 = 3. \end{cases} \end{cases} \tag{B.1}$$

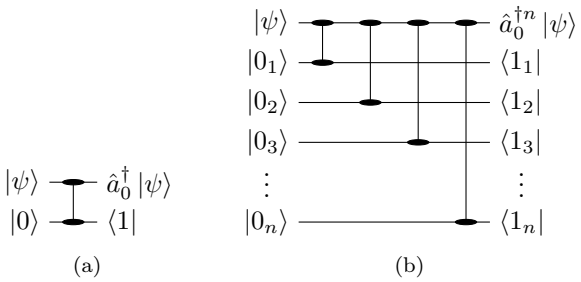


Figure B.1: Heralding scheme to do (a) single- and (b) many-photon addition. The vertical line with two ellipses at the end is used to denote a two-mode squeezing operation. Note that we always postselect on a single click for all but the first mode.

B.2 Cat and Fock States

B.2.1 Cat States

As is the case with finite-energy GKP states, cat states are linear superpositions of pure Gaussian states. We can write the density matrix of these states as

$$|k^\alpha\rangle\langle k^\alpha|_{\text{cat}} = \mathcal{N}\Big(|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle-\alpha| + |\alpha\rangle\langle-\alpha| e^{-i\pi k} + |-\alpha\rangle\langle\alpha| e^{i\pi k} \Big).$$

Again by linearity of the mapping between density operators and Wigner functions, we see that the Wigner functions of the first two terms $|\pm\alpha\rangle\langle\pm\alpha|$ are Gaussian functions with the covariance matrix of the vacuum and displacement $\boldsymbol{\mu}_\pm = \pm\sqrt{2\hbar}(\Re(\alpha), \Im(\alpha))$. Using the results from Appendix A of [21], we find the Wigner function of $|\alpha\rangle\langle-\alpha|$ and $|-\alpha\rangle\langle\alpha|$ to be complex-valued Gaussians with prefactor $e^{-2|\alpha|^2}$, vacuum covariance matrix, and complex-valued vectors of means $\boldsymbol{\mu}_z = \sqrt{2\hbar}(i\Im(\alpha), -i\Re(\alpha))$ with $\boldsymbol{\mu}_{\bar{z}} = \boldsymbol{\mu}_z^*$. From these results, we have that

$$\begin{aligned}
\mathcal{M} &= \{+, -, z, \bar{z}\}, \\
c_\pm &= \mathcal{N}, \quad c_z = (c_{\bar{z}})^* = e^{-i\pi k - 2|\alpha|^2}\mathcal{N} \\
\boldsymbol{\mu}_\pm &= \pm\sqrt{2\hbar}(\Re(\alpha), \Im(\alpha)), \\
\boldsymbol{\mu}_z &= (\boldsymbol{\mu}_{\bar{z}})^* = \sqrt{2\hbar}(i\Im(\alpha), -i\Re(\alpha)),
\end{aligned} \tag{B.2}$$

and where all the Gaussian functions have the same vacuum covariance matrix given in (3.11).

B.2.2 Fock States

In this section we consider the representation of Fock states as linear combinations of Gaussians. Compared with GKP and cat states, it is less clear whether Fock states can be can be written this way, since they cannot be represented as discrete coherent superpositions of Gaussian states. To obtain a representation for Fock states we use the idea of photon addition; that is, we study a quantum-optical circuit in which postselected heralded outcomes in certain ancillary modes allow us to apply the creation operation of a given mode to an arbitrary input state.

We first consider photon addition applied to the vacuum state for the generation of a single photon,

as considered in Ref. [74, 130, 323–326]. The circuit for the addition of a single photon is shown in Fig. B.1 (a), where the vertical line with ellipses on the ends corresponds to a two-mode squeezing operation $\hat{S}_{0,1}(r) = \exp\left(r\left[\hat{a}_0^\dagger \hat{a}_1^\dagger - \hat{a}_0 \hat{a}_1\right]\right)$ with squeezing parameter $r \ll 1$. To see how it works, we simply need to calculate

$$\begin{aligned}
&\left(\hat{\mathbb{I}}_0 \otimes |1_1\rangle\langle 1_1|\right) \hat{S}_{0,1}^{(2)}(r)\left(|\psi_0\rangle \otimes |0_1\rangle\right) \\
&\approx \left(\hat{\mathbb{I}}_0 \otimes |1_1\rangle\langle 1_1|\right)\left(\hat{\mathbb{I}} + r\hat{a}_0^\dagger \hat{a}_1^\dagger\right)\left(|\psi_0\rangle \otimes |0_1\rangle\right) &\text{(B.3)}\\
&= r\left(\hat{a}_0^\dagger |\psi_0\rangle\right) \otimes |1_1\rangle\,, &\text{(B.4)}
\end{aligned}$$

where $\hat{\mathbb{I}}_j$ is the identity operation in the Hilbert space of mode j. If $|\psi\rangle = |0_0\rangle$ is the vacuum state of mode 0, then the state at the output is a single photon in this mode. Since we are working in the regime where $r \ll 1$, to make progress we replace the photon-number-resolving detection $|1\rangle\langle 1|$ by its poor-man's version, $\hat{\mathbb{I}} - |0\rangle\langle 0|$, the threshold detection.

We can now write the probability of successful and failed heralding more formally as

$$p_1 = \mathrm{tr}\left(\left[\hat{\mathbb{I}}_1 - |0_1\rangle\langle 0_1|\right]|\Psi\rangle\langle\Psi|\right) = 1 - p_0, \tag{B.5}$$

$$p_0 = \frac{1}{1+\bar{n}}, \tag{B.6}$$

where $|\Psi\rangle = \hat{S}_{0,1}(r)\,|0_0 0_1\rangle$ and $\bar{n} = \sinh^2 r$ is the mean photon number in either of the two modes, 0 or 1. The state conditioned on successful heralding with threshold detectors in mode 0 is

$$|1\rangle\langle 1| \approx \hat{\rho}_1 = \frac{\mathrm{tr}_1\left(\left[\hat{\mathbb{I}}_1 - |0_1\rangle\langle 0_1|\right]|\Psi\rangle\langle\Psi|\right)}{p_1} \tag{B.7}$$

$$= \frac{\hat{\rho}^{\mathrm{th}} - p_0\,|0\rangle\langle 0|}{p_1}, \tag{B.8}$$

which is a linear combination of density matrices for two Gaussian states, namely, a thermal state with mean photon number $\bar{n}$ and the single-mode vacuum. A thermal state with mean photon number $\langle \hat{a}^\dagger \hat{a}\rangle = \langle \hat{n}\rangle = \bar{n}$ is a mixed Gaussian state with zero mean and covariance matrix $\mathbf{\Sigma} = \hbar(\bar{n} + \frac{1}{2})\mathbb{1}$ and can be expressed in the Fock basis as [106]

$$\hat{\rho}^{\mathrm{th}} = \frac{1}{1+\bar{n}} e^{-\epsilon \hat{n}} = \frac{1}{1+\bar{n}} \sum_{m=0}^{\infty} e^{-\epsilon m}\,|m\rangle\langle m|\,, \tag{B.9}$$

where $\bar{n} = [e^\epsilon - 1]^{-1}$ or, equivalently, $e^\epsilon = 1 + 1/\bar{n}$. With this expression one can easily confirm that $\hat{\rho}_1$ in Eq. (B.7) approaches a single photon in the limit $r \to 0$.

We generalize this scheme to an n-photon addition that, as schematically shown in Fig. B.1 (b), gives a Fock state with n photons when applied to the vacuum. The mathematical details of the derivation are provided in Appendix D of [21]; the final result is that we can approximate any n-particle Fock state using

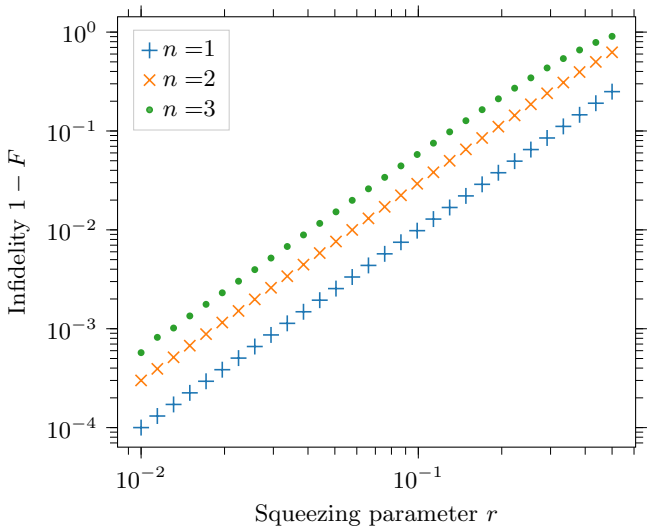


Figure B.2: Infidelity between the Gaussian approximation of a Fock state $\hat{\rho}_n$ with Wigner function described in Eq. (B.10) and an exact Fock state $|n\rangle$ as a function of the squeezing parameter r. The fidelity is defined as $F = \langle n|\hat{\rho}_n|n\rangle$.

the general notation of Eq. (3.23) with

$$\begin{aligned}
\mathcal{M} &= \{0, \ldots, n\}, \\
c_m &= \frac{(-1)^{n-m}}{\mathcal{N}_n}\binom{n}{m}\left[\frac{1-nr^2}{1-(n-m)r^2}\right], \\
\boldsymbol{\Sigma}_m &= \frac{\hbar}{2}\frac{1+(n-m)r^2}{1-(n-m)r^2}\mathbb{1}, \\
\boldsymbol{\mu}_m &= \mathbf{0}, \\
\mathcal{N}_n &= \frac{n!\left(n\left(\frac{-r^2-1}{r^2}\right)! + \left(-\frac{1}{r^2}\right)!\right)}{\left(\frac{nr^2-1}{r^2}\right)!}.
\end{aligned} \tag{B.10}$$

Although the equations above were derived with the assumption that the squeezing parameter $r \ll 1$, and they depend explicitly on this parameter, as long as $r < 1/\sqrt{n}$, the expressions correspond to a physical state. Formally, it is only in the limit $r \to 0$ that one recovers with perfect fidelity a Fock state. In Fig. B.2 we study the behaviour of the (in)fidelity between an ideal Fock state and our approximation. For $r \sim 10^{-2}$ one gets a fidelity of at least 99.9% for $n = 1, 2, 3$.

B.3 Derivation of the Wigner function of a Finite-Energy GKP State

B.3.1 Real Weights, Means, and Covariances *Via* Optical Circuits

This Appendix explicitly computes the Wigner function shown in eq. (3.43) using the physical application of the Fock damping operator using an optical circuit. For a derivation involving the applying $\hat{E}(\epsilon)$ directly in the phase space picture via the Moyal product, see Appendix C1 of [21].

The index set, weights, means and covariances matrices for the ideal GKP state are provided in Eq. (3.41). Finite energy GKP states can be recovered by applying the Fock damping channel $\hat{E}(\epsilon) = e^{-\epsilon\hat{n}}$ to the state, a channel which neatly fits into the Gaussian-inspired transformation framework introduced in Section 3.3.3. For the Fock damping channel with $e^{-\epsilon} = \cos\theta$, we derive the exact update rules for states in Appendix B.4. Thus, using Eq. (B.26), the covariances of the ideal GKP under Fock damping become:

$$\begin{aligned}
\Sigma_m &= \lim_{\delta\to 0+} \delta\mathbb{1} \\
&\overset{\hat{E}(\epsilon)}{\to} \lim_{\delta\to 0+} \left[\cos^2\theta\delta\mathbb{1} + \frac{\hbar}{2}\sin^2\theta\mathbb{1} - \cos^2\theta\sin^2\theta\left(\frac{\hbar}{2}\mathbb{1} - \delta\mathbb{1}\right)\left[\sin^2\theta\delta\mathbb{1} + \frac{\hbar}{2}(\cos^2\theta+1)\mathbb{1}\right]^{-1}\left(\frac{\hbar}{2}\mathbb{1} - \delta\mathbb{1}\right)\right] \\
&= \frac{\hbar}{2}\sin^2\theta\mathbb{1} - \frac{\hbar^2}{4}\cos^2\theta\sin^2\theta\mathbb{1}\left[(\cos^2\theta+1)\frac{\hbar}{2}\mathbb{1}\right]^{-1}\mathbb{1} \\
&= \frac{\hbar}{2}\frac{1-e^{-2\epsilon}}{1+e^{-2\epsilon}}\mathbb{1} = \boldsymbol{\Sigma}_m(\epsilon),
\end{aligned} \tag{B.11}$$

while the means become:

$$\begin{aligned}
\boldsymbol{\mu}_m &\overset{\hat{E}(\epsilon)}{\to} \lim_{\delta\to 0+}\left[\cos\theta\boldsymbol{\mu}_m - \cos\theta\sin^2\theta\left(\frac{\hbar}{2}\mathbb{1} - \delta\mathbb{1}\right)\left[\sin^2\theta\delta\mathbb{1} + \frac{\hbar}{2}(\cos^2\theta+1)\mathbb{1}\right]^{-1}\boldsymbol{\mu}_m\right] \\
&= \frac{2e^{-\epsilon}}{1+e^{-2\epsilon}}\boldsymbol{\mu}_m = \boldsymbol{\mu}_m(\epsilon).
\end{aligned} \tag{B.12}$$

The re-weighting of the peak is provided by Eq. (B.27):

$$\begin{aligned}
c_m(\epsilon;\theta,\phi) &= \lim_{\delta\to 0+}\frac{w(\mathbf{0}|\sin\theta\boldsymbol{\mu}_m,\ \sin^2\theta\delta\mathbb{1} + \hbar\cos^2\theta\mathbb{1}/2,\ \hbar\mathbb{1}/2)}{p(\mathbf{0};|\psi(\theta,\phi)\rangle_{\text{gkp}},\hbar\mathbb{1}/2)} \\
&= \frac{c_m(\theta,\phi)G_{\sin\theta\boldsymbol{\mu}_m,\hbar(1+\cos^2\theta)\mathbb{1}/2}(\mathbf{0})}{p(\mathbf{0};|\psi(\theta,\phi)\rangle_{\text{gkp}},\hbar\mathbb{1}/2)} \\
&= \frac{c_m(\theta,\phi)}{\mathcal{N}_\epsilon}\exp\left[-\frac{1-e^{-2\epsilon}}{\hbar(1+e^{-2\epsilon})}\boldsymbol{\mu}_m^T\boldsymbol{\mu}_m\right]
\end{aligned} \tag{B.13}$$

where $\mathcal{N}_\epsilon$ is an overall normalization. This completes the derivation.

B.3.2 Complex Weights and Means, and Real Covariances *Via* the Wavefunction to Wigner Transformation

The ideal GKP state for a single qubit is:

$$|\psi(\boldsymbol{a})\rangle = a_0|0\rangle_{\text{gkp}} + a_1|1\rangle_{\text{gkp}} = \sum_{t=0}^{1} a_t \sum_{k=-\infty}^{\infty}|\sqrt{\pi\hbar}(2k+t)\rangle_q, |a_0|^2 + |a_1|^2 = 1 \tag{B.14}$$

Next consider application of $\hat{E}(\epsilon) = e^{-\epsilon\hat{n}}$:

$$\hat{E}(\epsilon)|\psi\rangle = \sum_{t=0}^{1} a_t \sum_{k=-\infty}^{\infty}\int ds\left[\langle s|\hat{E}(\epsilon)|\sqrt{\pi\hbar}(2k+t)\rangle_q\right]|s\rangle_q \tag{B.15}$$

Next we use Mehler's kernel [327]:

$$\begin{aligned}\langle s_1|\hat{E}(\epsilon)|s_2\rangle_q &= \sum_{n=0}^{\infty} e^{-\epsilon n}\langle s_1|n\rangle_q\langle n|s_2\rangle_q \\ &= \sum_{n=0}^{\infty}\frac{e^{-\epsilon n}}{2^n n!\sqrt{\pi\hbar}}e^{-(s_1^2+s_2^2)/2\hbar}H_n(s_1)H_n(s_2) \\ &= \frac{e^{\epsilon/2}}{\sqrt{2\pi\sinh(\epsilon)}}\exp\left[-\coth(\epsilon)(s_1^2+s_2^2)/2\hbar+\operatorname{csch}(\epsilon)s_1s_2/\hbar\right] \\ &= \frac{e^{\epsilon/2}}{\sqrt{2\pi\hbar\sinh(\epsilon)}}\exp[-\frac{1}{2\hbar}\boldsymbol{s}^T\boldsymbol{\sigma}_{\boldsymbol{\epsilon}}^{-1}\boldsymbol{s}]\end{aligned} \tag{B.16}$$

where the second last line was obtained with Mehler's formula. In the last line:

$$\begin{aligned}\boldsymbol{s} &= (s_1,s_2)^T \\ \boldsymbol{\sigma}_{\boldsymbol{\epsilon}}^{-1} &= \begin{pmatrix}\alpha & -\beta\\ -\beta & \alpha\end{pmatrix} = \begin{pmatrix}\coth(\epsilon) & \operatorname{csch}(\epsilon)\\ \operatorname{csch}(\epsilon) & \coth(\epsilon)\end{pmatrix}\end{aligned} \tag{B.17}$$

Therefore, we have that the (subnormalized) wavefunction in position quadrature is given by:

$$\begin{aligned}\psi_{\boldsymbol{a},\epsilon}(x) &= \sum_{t=0}^{1}a_t\sum_{k=-\infty}^{\infty}\langle x|\,\hat{E}(\epsilon)\,|\sqrt{\pi\hbar}(2k+t)\rangle_q \\ &= \sum_{t=0}^{1}a_t\sum_{k=-\infty}^{\infty}\exp[-\alpha\pi(2k+t)^2/2]\exp[-\alpha x^2/2\hbar-\beta\sqrt{\pi}x(2k+t)/\sqrt{\hbar}].\end{aligned} \tag{B.18}$$

Note that we can write the amplitude inside the sum as

$$\exp[-\alpha x^2/2\hbar-\beta\sqrt{\pi}x(2k+t)/\sqrt{\hbar}] = \sqrt[4]{\frac{\pi\hbar}{\alpha}}e^{\frac{\pi\beta^2(2k+t)^2}{2\alpha}}\langle x|\,\hat{D}\left(-\sqrt{\tfrac{\pi}{2}}(2k+t)\tfrac{\beta}{\alpha}\right)\hat{S}\left(\tfrac{1}{2}\log_{\mathrm{e}}\alpha\right)|0\rangle\,, \tag{B.19}$$

where $\hat{D}$ and $\hat{S}$ are the single mode displacement and squeezing operators (cf. Eq. (3.13)) and $|0\rangle$ is the single mode vacuum. We can use these results to write the Hilbert space vector of a finite energy GKP state as

$$|\psi(\boldsymbol{a},\epsilon)\rangle = \sum_{t=0}^{1}a_t\sum_{k=-\infty}^{\infty}\kappa_{k,t}\;\hat{D}\left(\sqrt{\tfrac{\pi}{2}}(2k+t)\operatorname{sech}(\epsilon)\right)\hat{S}\left(-\tfrac{1}{2}\log_{\mathrm{e}}\tanh(\epsilon)\right)|0\rangle\,, \tag{B.20}$$

$$\kappa_{k,t} = \sqrt[4]{\pi\hbar\tanh(\epsilon)}e^{-\frac{1}{2}\pi(2k+t)^2\tanh(\epsilon)}, \tag{B.21}$$

showing that finite-energy GKP states can be written as linear combinations of single-mode pure Gaussian states like the ones discussed in Eq. (3.26).

Now we move on to the (subnormalized) Wigner function calculation:

$$\begin{aligned}W(\boldsymbol{\xi};|\psi(\boldsymbol{a},\epsilon)\rangle) &= \frac{1}{2\pi\hbar}\int dy\exp(-ipy/\hbar)\psi^*_{\boldsymbol{a},\epsilon}(x-y/2)\psi_{\boldsymbol{a},\epsilon}(x+y/2) \\ &= \frac{1}{2\pi\hbar}\sum_{s,t=0}^{1}a_ta_s^*\sum_{k,\ell=-\infty}^{\infty}\exp[-\alpha\pi(2k+t)^2/2]\exp[-\alpha\pi(2\ell+s)^2/2]I(x,p;2k+t,2\ell+s)\end{aligned} \tag{B.22}$$

where:

$$\begin{aligned} I(x,p;a,b) = &\int dy \exp(-ipy/\hbar) \exp[-\alpha(x+y/2)^2/2\hbar - a\beta\sqrt{\pi}(x+y/2)/\sqrt{\hbar}] \\ &\times \exp[-\alpha(x-y/2)^2/2\hbar - b\beta\sqrt{\pi}(x-y/2)/\sqrt{\hbar}] \\ =&\sqrt{\frac{4\pi\hbar}{\alpha}} \exp\left[-\frac{\alpha}{\hbar}x^2 - x\frac{\beta\sqrt{\pi}}{\sqrt{\hbar}}(a+b)\right] \exp\left[-\frac{1}{\alpha\hbar}\left(p + \frac{i\beta\sqrt{\pi\hbar}}{2}(b-a)\right)^2\right] \end{aligned} \tag{B.23}$$

Here, the result was obtained by performing the Gaussian integral.

Putting it all together, we find that the normalized Wigner function is a linear combination of Gaussian functions in phase space:

$$W(\boldsymbol{\xi}; |\psi(\boldsymbol{a},\epsilon)\rangle) = \sum_{m\in\mathcal{M}} c_m G_{\boldsymbol{\mu}_m,\boldsymbol{\Sigma}_m}(\boldsymbol{\xi}) \tag{B.24}$$

with:

$$\begin{aligned} &\mathcal{M} = \{m \equiv (k,\ell,s,t) \,|\, s,t \in \{0,1\} \,\&\, k,\ell \in \mathbb{Z}\} \\ &\boldsymbol{\mu}_m^T = \left(\frac{-\beta\sqrt{\pi\hbar}(2\ell+2k+t+s)}{2\alpha}, \frac{-i\beta\sqrt{\pi\hbar}(2\ell+s-2k-t)}{2}\right) \\ &\boldsymbol{\Sigma}_m = \frac{\hbar}{2}\begin{pmatrix} 1/\alpha & 0 \\ 0 & \alpha \end{pmatrix} \\ &(\alpha,\beta) = (\coth(\epsilon), -\operatorname{csch}(\epsilon)) \\ &c_m = \frac{1}{\mathcal{N}(\boldsymbol{a},\epsilon)} a_t a_s^* \exp[-\alpha\pi(2k+t)^2/2] \exp[-\alpha\pi(2\ell+s)^2/2] \exp\left[\frac{\beta^2\pi(2k+2\ell+t+s)^2}{4\alpha}\right] \end{aligned} \tag{B.25}$$

where $\mathcal{N}(\boldsymbol{a},\epsilon)$ is an overall normalization chosen so that $\sum_{m\in\mathcal{M}} c_m = 1$, since the Gaussian functions are already normalized over phase space.

B.4 Fock Damping

Here we show how a state with a Wigner function of the form in Eq. (3.23) is transformed by the Fock damping operator in Eq. (3.42). Recall that the $\hat{E}(\epsilon)$ operator can be viewed as arising from passing a state through a beam-splitter of transmissivity $\cos\theta = e^{-\epsilon}$ with an ancillary mode in a vacuum state, then measuring vacuum on the ancillary mode [16, 28].

Let the state have means and covariance $(\boldsymbol{\mu}_{m,0}, \boldsymbol{\Sigma}_{m,0})$. First we pass the state and an ancillary vacuum mode (covariance matrix of $\hbar\mathbb{1}/2$) through a beam-splitter represented by symplectic matrix $\boldsymbol{S}_\theta = \begin{pmatrix} \cos\theta\mathbb{1} & \sin\theta\mathbb{1} \\ -\sin\theta\mathbb{1} & \cos\theta\mathbb{1} \end{pmatrix}$ where we have chosen the mode-wise ordering (q_1, p_1, q_2, p_2):

$$\boldsymbol{\Sigma}_m = \boldsymbol{S}_\theta \begin{pmatrix} \boldsymbol{\Sigma}_{m,0} & \boldsymbol{0} \\ \boldsymbol{0} & \hbar\mathbb{1}/2 \end{pmatrix} \boldsymbol{S}_\theta^T = \begin{pmatrix} \cos^2\theta\boldsymbol{\Sigma}_{m,0} + \hbar\sin^2\theta\mathbb{1}/2 & \cos\theta\sin\theta(\hbar\mathbb{1}/2 - \boldsymbol{\Sigma}_{m,0}) \\ \cos\theta\sin\theta(\hbar\mathbb{1}/2 - \boldsymbol{\Sigma}_{m,0}) & \sin^2\theta\boldsymbol{\Sigma}_{m,0} + \hbar\cos^2\theta\mathbb{1}/2 \end{pmatrix} \tag{B.26}$$

$$\boldsymbol{\mu}_m = \boldsymbol{S}_\theta \begin{pmatrix} \boldsymbol{\mu}_{m,0} \\ \boldsymbol{0} \end{pmatrix} = \begin{pmatrix} \cos\theta\boldsymbol{\mu}_{m,0} \\ \sin\theta\boldsymbol{\mu}_{m,0} \end{pmatrix} \tag{B.27}$$

where $\boldsymbol{\Sigma}_m$ and $\boldsymbol{\mu}_m$ now have the same form as Eq. (3.34).

Next, we project the second mode onto vacuum. Here we can use Eq. (3.35), setting $\boldsymbol{\Sigma}_j = \hbar\mathbb{1}/2$ and $\boldsymbol{r}_j = \mathbf{0}$, so that the full mapping becomes:

$$\begin{aligned}
\boldsymbol{\Sigma}_{m,0} &\to \cos^2\theta\boldsymbol{\Sigma}_{m,0} + \frac{\hbar}{2}\sin^2\theta\mathbb{1} - \cos^2\theta\sin^2\theta(\frac{\hbar}{2}\mathbb{1} - \boldsymbol{\Sigma}_{m,0})\left[\sin^2\theta\boldsymbol{\Sigma}_{m,0} + \frac{\hbar}{2}(\cos^2\theta+1)\mathbb{1}\right]^{-1}(\frac{\hbar}{2}\mathbb{1} - \boldsymbol{\Sigma}_{m,0}),\\
\boldsymbol{\mu}_{m,0} &\to \cos\theta\boldsymbol{\mu}_{m,0} - \cos\theta\sin^2\theta(\frac{\hbar}{2}\mathbb{1} - \boldsymbol{\Sigma}_{m,0})\left[\sin^2\theta\boldsymbol{\Sigma}_{m,0} + \frac{\hbar}{2}(\cos^2\theta+1)\mathbb{1}\right]^{-1}\boldsymbol{\mu}_{m,0}.
\end{aligned} \tag{B.28}$$

Finally, as in Eq. (3.37) we find a per-peak re-weighting of:

$$c_m \to \frac{w(\mathbf{0}|\sin\theta\boldsymbol{\mu}_{m,0},\ \sin^2\theta\boldsymbol{\Sigma}_{m,A} + \hbar\cos^2\theta\mathbb{1}/2,\ \hbar\mathbb{1}/2)}{p(\mathbf{0};\hat{\rho},\hbar\mathbb{1}/2)}, \tag{B.29}$$

where $d_j = 1$ as the measurement is onto only a single Gaussian in phase space.

B.5 Measurement-Based Gates that Employ Squeezed Ancillae

In Table B.1, we provide a summary of the Gaussian CPTP maps corresponding to measurement-based squeezing, implemented using an ancillary squeezed state, a beam-splitter and homodyne measurement followed by a feedforward displacement.

Gate	**Average Map**
q-squeezing	$\boldsymbol{X}_s^{(q)} = \begin{pmatrix}\cos\theta & 0\\ 0 & \cos\theta^{-1}\end{pmatrix}$, $\cos\theta = e^{-s}$, $\boldsymbol{Y}_{s,r,\eta}^{(q)} = \frac{\hbar}{2}\begin{pmatrix}\sin^2\theta e^{-2r} & 0\\ 0 & \eta^{-1}\tan^2\theta(1-\eta)\end{pmatrix}$
p-squeezing	$\boldsymbol{X}_s^{(p)} = \begin{pmatrix}\cos\theta^{-1} & 0\\ 0 & \cos\theta\end{pmatrix}$, $\boldsymbol{Y}_{s,r,\eta}^{(p)} = \frac{\hbar}{2}\begin{pmatrix}\eta^{-1}\tan^2\theta(1-\eta) & 0\\ 0 & \sin^2\theta e^{-2r}\end{pmatrix}$

Table B.1: Average Gaussian CPTP maps for inline squeezing, as implemented with an ancillary squeezed vacuum state. Here, $\cos\theta = e^{-s}$ is the squeezing parameter, r is the squeezing parameter of the ancillary state, and η is the loss parameter of the inefficient homodyne measurement on the ancilla.

In Table B.2, we provide a summary of the map for a single-shot run of ancilla-assisted squeezing.

B.5.1 Inline Squeezing Gate

Here we derive the map effected by measurement-based squeezing, as described in [101], and shown in Fig. B.3 (a).

Initial states Let the covariance matrix and mean of the initial state be of the form:

$$\boldsymbol{\Sigma}_{m,A,0} = \frac{\hbar}{2}\begin{pmatrix}a_{m,0} & b_{m,0}\\ b_{m,0} & d_{m,0}\end{pmatrix},\ \boldsymbol{\mu}_{m,A,0} = \begin{pmatrix}\bar{x}_{m,0}\\ \bar{p}_{m,0}\end{pmatrix}. \tag{B.30}$$

Covariance	*Initial*: $\boldsymbol{\Sigma}_{m,A,0} = \frac{\hbar}{2}\begin{pmatrix} a_{m,0} & b_{m,0} \\ b_{m,0} & d_{m,0} \end{pmatrix}$, $\boldsymbol{\Sigma}_r = \frac{\hbar}{2}\begin{pmatrix} e^{-2r} & 0 \\ 0 & e^{2r} \end{pmatrix}$ *Final*: $\cos^2\theta\boldsymbol{\Sigma}_{m,A,0} + \sin^2\theta\boldsymbol{\Sigma}_r - \frac{\hbar\eta(\sin\theta\cos\theta)^2}{2F_m}\begin{pmatrix} b_{m,0}^2 & d_{m,r}b_{m,0} \\ d_{m,r}b_{m,0} & d_{m,r}^2 \end{pmatrix}$ $F_m = \left[\eta(s_\theta^2 d_{m,0} + c_\theta^2 e^{2r}) + 1 - \eta\right]$, $d_{m,r} = d_{m,0} - e^{2r}$
Mean	*Initial*: $\boldsymbol{\mu}_{m,A,0} = \begin{pmatrix} \bar{x}_{m,0} \\ \bar{p}_{m,0} \end{pmatrix}$ *Final*: $\cos\theta\boldsymbol{\mu}_{m,A,0} - \frac{\sqrt{\eta}\sin\theta\cos\theta(p_M - \sqrt{\eta}\sin\theta\bar{p}_{m,0})}{F_m}\begin{pmatrix} b_{m,0} \\ d_{m,r} \end{pmatrix} + \begin{pmatrix} f_x(p_M) \\ f_p(p_M) \end{pmatrix}$
Weight	*Initial*: c_m *Final*: $c_m \frac{\exp[-2(p_M - \sqrt{\eta}s_\theta\bar{p}_{m,0})^2/\hbar F_m]}{\sum_{n\in\mathbb{M}} c_n \exp[-2(p_M - \sqrt{\eta}s_\theta\bar{p}_{n,0})^2/\hbar F_n]}$

Table B.2: Update rules for each Gaussian peak in a linear combination from a single-shot run of ancilla-assisted squeezing in q quadrature. r is the squeezing level of the ancilla state, θ is the angle of the beam-splitter with $\cos\theta = e^{-s}$ defining the squeezing parameter s, η is the loss parameter of the homodyne detector, and p_M is the p-homodyne outcome on the ancilla mode. To match the average case, one sets the feedforward operations to $f_x(p_M) = 0$ and $f_p(p_M) = p_M \tan\theta/\sqrt{\eta}$, and integrate over p_M.

The covariance matrix and mean of the ancillary squeezed state is:

$$\boldsymbol{\Sigma}_{B,0} = \frac{\hbar}{2}\begin{pmatrix} e^{-2r} & 0 \\ 0 & e^{2r} \end{pmatrix}, \ \boldsymbol{\mu}_{B,0} = \mathbf{0}. \tag{B.31}$$

The initial separable form of the joint covariance and mean is thus $\boldsymbol{\Sigma}_{m,0} = \boldsymbol{\Sigma}_{m,A,0} \oplus \boldsymbol{\Sigma}_{B,0}$ and $\boldsymbol{\mu}_{m,0} = \boldsymbol{\mu}_{m,A,0} \oplus \boldsymbol{\mu}_{B,0}$.

Combine at a beam-splitter When combined on a beam-splitter we evolve the two-mode covariance matrix and means to

$$\boldsymbol{\Sigma}_{m,1} = \begin{pmatrix} \cos^2\theta\boldsymbol{\Sigma}_{m,A,0} + \sin^2\theta\boldsymbol{\Sigma}_{B,0} & -\sin\theta\cos\theta(\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0}) \\ -\sin\theta\cos\theta(\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0}) & \sin^2\theta\boldsymbol{\Sigma}_{m,A,0} + \cos^2\theta\boldsymbol{\Sigma}_{B,0} \end{pmatrix}, \boldsymbol{\mu}_{m,1} = \begin{pmatrix} \cos\theta\boldsymbol{\mu}_{m,A,0} \\ \sin\theta\boldsymbol{\mu}_{m,A,0} \end{pmatrix} \tag{B.32}$$

where symplectic matrix for a beam-splitter is given after Eq. (3.46). Note we are working with basis ordering (q_1, p_1, q_2, p_2). For convenience we now move to $c_\theta \equiv \cos\theta$, $s_\theta \equiv \sin\theta$.

Inefficient homodyne measurement We now apply loss to the second mode to model inefficient homodyne measurement. The Gaussian CPTP map for the loss channel is provided in Eq. (3.45), so loss on the second mode only is simply $(\boldsymbol{X}, \boldsymbol{Y}) = (\mathbb{1} \oplus \sqrt{\eta}\mathbb{1}, \mathbf{0} \oplus [1-\eta]\hbar\mathbb{1}/2)$. Applying this channel yields the updated covariance and means:

$$\boldsymbol{\Sigma}_m = \begin{pmatrix} c_\theta^2\boldsymbol{\Sigma}_{m,A,0} + s_\theta^2\boldsymbol{\Sigma}_{B,0} & -\sqrt{\eta}s_\theta c_\theta(\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0}) \\ -\sqrt{\eta}s_\theta c_\theta(\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0}) & \eta(s_\theta^2\boldsymbol{\Sigma}_{m,A,0} + c_\theta^2\boldsymbol{\Sigma}_{B,0}) + (1-\eta)\hbar\mathbb{1}/2 \end{pmatrix}, \boldsymbol{\mu}_m = \begin{pmatrix} c_\theta\boldsymbol{\mu}_{m,A,0} \\ \sqrt{\eta}s_\theta\boldsymbol{\mu}_{m,A,0} \end{pmatrix} \tag{B.33}$$

At this stage, $\boldsymbol{\Sigma}_m$ and $\boldsymbol{\mu}_m$ are in the form required for using Eq. (3.34).

Measurement and conditional dynamics with feedforward We now perform a ideal p-homodyne measurement on mode B which yields outcome $\boldsymbol{r}_j^T = (0, p_M)$. Employing now Eq. (3.35), we find that the updated covariance and means for mode A are provided by:

$$\begin{aligned}\boldsymbol{\Sigma}_{m,A} = c_\theta^2 \boldsymbol{\Sigma}_{m,A,0} + s_\theta^2 \boldsymbol{\Sigma}_{B,0} \\ - \eta (s_\theta c_\theta)^2 (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^T [\eta(s_\theta^2 \boldsymbol{\Sigma}_{m,A,0} + c_\theta^2 \boldsymbol{\Sigma}_{B,0}) + (1-\eta)\hbar \mathbb{1}/2 + \boldsymbol{\Sigma}_j]^{-1} (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0}),\end{aligned} \tag{B.34}$$

$$\boldsymbol{\mu}_{m,A} = c_\theta \boldsymbol{\mu}_{m,A,0} - \sqrt{\eta} s_\theta c_\theta (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^T [\eta(s_\theta^2 \boldsymbol{\Sigma}_A + c_\theta^2 \boldsymbol{\Sigma}_{B,0}) + (1-\eta)\hbar \mathbb{1}/2 + \boldsymbol{\Sigma}_j]^{-1} (\boldsymbol{r}_j - \sqrt{\eta} s_\theta \boldsymbol{\mu}_{m,A,0}) \tag{B.35}$$

Here, $\boldsymbol{\Sigma}_j$ is the covariance matrix of the Gaussian state which is detected. In this case, we are considering ideal p-homodyne measurement, so it is given by $\boldsymbol{\Sigma}_j = \lim_{\epsilon \to 0} \frac{\hbar}{2} \begin{pmatrix} \epsilon^{-1} & 0 \\ 0 & \epsilon \end{pmatrix}$. We note that there is only one weight with $d_j = 1$, since the measurement is onto a Gaussian state.

Because of the form of $\boldsymbol{\Sigma}_M$, the inverse matrix in both the expressions for the covariance matrix and mean reduces to

$$\left[\eta(s_\theta^2 \boldsymbol{\Sigma}_{m,A,0} + c_\theta^2 \boldsymbol{\Sigma}_{B,0}) + (1-\eta)\hbar \mathbb{1}/2 + \boldsymbol{\Sigma}_M\right]^{-1} = \frac{2}{\hbar} \begin{pmatrix} 0 & 0 \\ 0 & 1/F_m \end{pmatrix}, \tag{B.36}$$

$$F_m = \eta(s_\theta^2 d_{m,0} + c_\theta^2 e^{2r}) + 1 - \eta. \tag{B.37}$$

Let $\boldsymbol{T}^{(i,j)}$ denote the i^{th} row and j^{th} column of matrix $\boldsymbol{T}$. We find that after the feedforward displacement, the covariance and mean update as

$$\begin{aligned}\boldsymbol{\Sigma}_{m,A} = c_\theta^2 \boldsymbol{\Sigma}_{m,A,0} + s_\theta^2 \boldsymbol{\Sigma}_{B,0} - \frac{\eta (s_\theta c_\theta)^2}{F_m} \\ \times \begin{pmatrix} [(\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^{(1,2)}]^2 & (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^{(2,2)} (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^{(1,2)} \\ (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^{(2,2)} (\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^{(1,2)} & [(\boldsymbol{\Sigma}_{m,A,0} - \boldsymbol{\Sigma}_{B,0})^{(2,2)}]^2 \end{pmatrix}\end{aligned} \tag{B.38}$$

$$= c_\theta^2 \boldsymbol{\Sigma}_{m,A,0} + s_\theta^2 \boldsymbol{\Sigma}_{B,0} - \frac{\hbar \eta (s_\theta c_\theta)^2}{2F_m} \begin{pmatrix} b_{m,0}^2 & (d_{m,0} - e^{2r}) b_{m,0} \\ (d_{m,0} - e^{2r}) b_{m,0} & (d_{m,0} - e^{2r})^2 \end{pmatrix} \tag{B.39}$$

$$\boldsymbol{\mu}'_{m,A} = c_\theta \boldsymbol{\mu}_{m,A,0} - \frac{\sqrt{\eta} s_\theta c_\theta (p_M - \sqrt{\eta} s_\theta \bar{p}_{m,0})}{F_m} \begin{pmatrix} b_{m,0} \\ d_{m,0} - e^{2r} \end{pmatrix} + \begin{pmatrix} f_x(p_M) \\ f_p(p_M) \end{pmatrix}, \tag{B.40}$$

Note that the outcome covariance matrix does not depend on the measurement outcome, while the mean does. Here, $f_x(p_M), f_p(p_M)$ correspond to the value of the feedforward displacement that is applied after the homodyne measurement.

Single run peak reweighting While we have found the covariance matrix and mean of mode A after detecting outcome p_M on mode B, we must also take into account the probability of detecting p_M, which is

provided by Eq. (3.27):

$$p(p_M;\hat{\rho},\boldsymbol{\Sigma}_M)=\sum_{m\in\mathcal{M}}c_m\frac{\exp[-2(p_M-\sqrt{\eta}s_\theta\bar{p}_{m,0})^2/\hbar F_m]}{\sqrt{\pi\hbar F_m/2}} \tag{B.41}$$

Note that each term in the sum depends on $\bar{p}_{m,0}$ and F_m. This means that if mode A is a linear combination of Gaussian functions, the coefficients update as in Eq. (3.37):

$$c_m(p_M)=c_m\frac{\exp[-2(p_M-\sqrt{\eta}s_\theta\bar{p}_{m,0})^2/\hbar F_m]}{p(p_M;\hat{\rho},\boldsymbol{\Sigma}_M)\sqrt{\pi\hbar F_m/2}} \tag{B.42}$$

At this stage, we have all the information required to update a single-shot run of measurement-based squeezing.

Average map However, we can also use the probability to determine the average map applied to the state. Importantly, we find that the average map for specific $f_x(p_M), f_p(p_M)$ simply corresponds to a Gaussian CPTP map applied to mode A, which means we do not need to worry about per-peak reweighting, but can simply update the covariance and means of each peak as in Eq. (3.32). Consider the characteristic function of the output state $\hat{\rho}_M$, which is a function of the input state and the measurement outcome p_M:

$$\chi(\hat{\rho}_M;\boldsymbol{\xi})=\sum_{m\in\mathcal{M}}c_m(p_M)\exp[-\frac{1}{4}\boldsymbol{\xi}^T\boldsymbol{\Omega}^T\boldsymbol{\Sigma}_{m,A}\boldsymbol{\Omega}\boldsymbol{\xi}-i\boldsymbol{\xi}^T\boldsymbol{\Omega}\boldsymbol{\mu}'_{m,A}],\ \text{with}\ \boldsymbol{\Omega}=\begin{pmatrix}0&1\\-1&0\end{pmatrix} \tag{B.43}$$

The measurement-averaged output state is then given by

$$\hat{\rho}=\int dp_M p(p_M;\hat{\rho},\boldsymbol{\Sigma}_M)\hat{\rho}_M$$

The characteristic function of the output state is then given by

$$\chi(\hat{\rho};\boldsymbol{\xi})=\int dp_M p(p_M;\hat{\rho},\boldsymbol{\Sigma}_M)\chi(\hat{\rho}_M;\boldsymbol{\xi}) \tag{B.44}$$

Since the covariance matrix is independent of p_M:

$$\chi(\hat{\rho};\boldsymbol{\xi})=\sum_{m\in\mathcal{M}}\frac{c_m}{\sqrt{\pi\hbar F_m/2}}\exp[-\frac{1}{4}\boldsymbol{\xi}^T\boldsymbol{\Omega}^T\boldsymbol{\Sigma}_{m,A}\boldsymbol{\Omega}\boldsymbol{\xi}]\int dp_M\exp\left[-\frac{2(p_M-\sqrt{\eta}s_\theta\bar{p}_{m,0})^2}{\hbar F_m}\right]\exp[-i\boldsymbol{\xi}^T\boldsymbol{\Omega}\boldsymbol{\mu}'_{m,A}] \tag{B.45}$$

Let us rewrite $\boldsymbol{\mu}'_{m,A}$ as:

$$\boldsymbol{\mu}'_{m,A}=\tilde{\boldsymbol{r}}(p_M)+(p_M-\sqrt{\eta}s_\theta\bar{p}_{m,A})\boldsymbol{\nu}_m,\quad \boldsymbol{\nu}_m=-\frac{\sqrt{\eta}s_\theta c_\theta}{F_m}\begin{pmatrix}b_{m,0}\\d_{m,0}-e^{2r}\end{pmatrix},\quad \tilde{\boldsymbol{r}}(p_M)=c_\theta\boldsymbol{\mu}_{m,0}+\begin{pmatrix}f_x[p_M]\\f_p[p_M]\end{pmatrix} \tag{B.46}$$

Furthermore, we relabel $\tilde{p}_{M,m}=p_M-\sqrt{\eta}s_\theta\bar{p}_{m,A}$. The characteristic function then becomes:

$$\chi(\hat{\rho};\boldsymbol{\xi})=\sum_{m\in\mathcal{M}}\frac{c_m\exp[-\frac{1}{4}\boldsymbol{\xi}^T\boldsymbol{\Omega}^T\boldsymbol{\Sigma}_{m,A}\boldsymbol{\Omega}\boldsymbol{\xi}]}{\sqrt{\pi\hbar F_m/2}}\int dp_M\exp\left[-i\boldsymbol{\xi}^T\boldsymbol{\Omega}^T\tilde{\boldsymbol{r}}(p_M)-\frac{2\tilde{p}_{M,m}^2}{\hbar F_m}-i\tilde{p}_{M,m}\boldsymbol{\xi}^T\boldsymbol{\Omega}\boldsymbol{\nu}_m\right] \tag{B.47}$$

Displacing mode A by linear and quadratic functions of p_M, the above integral reduces to a Gaussian integral. Let us assume that

$$f_x[p_M] = g_0 + g_1 p_M + g_2 p_M^2, \quad f_p[p_M] = h_0 + h_1 p_M + h_2 p_M^2. \tag{B.48}$$

Sanity check As a sanity check, let us check what happens if we apply no feedforward. Performing the Gaussian integral, we get that:

$$\chi(\hat{\rho}; \boldsymbol{\xi}) = \sum_{m \in \mathcal{M}} c_m \exp[-\frac{1}{4} \boldsymbol{\xi}^T \boldsymbol{\Omega}^T \boldsymbol{\Sigma}_{m,A} \boldsymbol{\Omega} \boldsymbol{\xi}] \exp[-i \boldsymbol{\xi}^T \boldsymbol{\Omega} c_\theta \boldsymbol{\mu}_{m,0}] \exp[-\frac{\hbar F_m}{8} (\boldsymbol{\xi}^T \boldsymbol{\Omega} \boldsymbol{\nu}_m)^2] \tag{B.49}$$

This can be simplified into

$$\chi(\hat{\rho}; \boldsymbol{\xi}) = \sum_{m \in \mathcal{M}} c_m \exp[-\frac{1}{4} \boldsymbol{\xi}^T \boldsymbol{\Omega}^T \tilde{\boldsymbol{\Sigma}}_{m,A} \boldsymbol{\Omega} \boldsymbol{\xi}] \exp[-i \boldsymbol{\xi}^T \boldsymbol{\Omega} c_\theta \boldsymbol{\mu}_{m,0}], \quad \tilde{\boldsymbol{\Sigma}}_{m,A} = \boldsymbol{\Sigma}_{m,A} + \frac{\hbar F_m}{2} \boldsymbol{\nu}_m \boldsymbol{\nu}_m^T. \tag{B.50}$$

We see that we recover the partial trace condition, i.e. that $\boldsymbol{\Sigma}_{m,0} \to c_\theta^2 \boldsymbol{\Sigma}_{m,0} + s_\theta^2 \boldsymbol{\Sigma}_{B,0}$ and $\boldsymbol{\mu}_{m,0} \to c_\theta \boldsymbol{\mu}_{m,0}$, which is exactly what happens if we pass through a beam-splitter and trace out mode B.

Feedforward linear in m Consider linear feedforward, by setting $\tilde{\boldsymbol{r}}_2 = \tilde{\boldsymbol{r}}_0 = 0$. Instead, we have that

$$\begin{aligned} \chi(\hat{\rho}; \boldsymbol{\xi}) = \sum_{m \in \mathcal{M}} \frac{c_m}{\sqrt{\pi \hbar F_m / 2}} \exp\left[-\frac{1}{4} \boldsymbol{\xi}^T \boldsymbol{\Omega}^T \boldsymbol{\Sigma}_{m,A} \boldsymbol{\Omega} \boldsymbol{\xi} - i \boldsymbol{\xi}^T \boldsymbol{\Omega} (c_\theta \boldsymbol{\mu}_{m,0}) - \frac{2\eta s_\theta^2 \bar{p}_{m,A}^2}{\hbar F_m} + i \sqrt{\eta} s_\theta \bar{p}_{m,A} \boldsymbol{\xi}^T \boldsymbol{\Omega} \boldsymbol{\nu}_m \right] \\ \times \int dp_M \exp\left[p_M \left[\frac{4\sqrt{\eta} s_\theta \bar{p}_{m,A}}{\hbar F_m} - i \boldsymbol{\xi}^T \boldsymbol{\Omega} (\tilde{\boldsymbol{r}}_1 + \boldsymbol{\nu}_m) \right] \exp[-p_M^2 \left(\frac{2}{\hbar F_m} \right) \right] \end{aligned} \tag{B.51}$$

Let us further choose the linear feedforward to be $g_1 = 0$ and $h_1 = \sqrt{\eta^{-1}} \tan\theta$, which is based on the suggestion in [101]. Then the output covariance matrix can be written as

$$\boldsymbol{\Sigma}_{A,out} = \boldsymbol{\Sigma}_{m,A} + \frac{\hbar F_m}{2} (\boldsymbol{\nu}_m + \tilde{\boldsymbol{r}}_1)(\boldsymbol{\nu}_m + \tilde{\boldsymbol{r}}_1)^T = \boldsymbol{X} \boldsymbol{\Sigma}_{m,A,0} \boldsymbol{X}^T + \boldsymbol{Y}, \quad \boldsymbol{\mu}_{A,\text{out}} = \boldsymbol{X} \boldsymbol{\mu}_{m,A,0}$$

with

$$\boldsymbol{X} = \begin{pmatrix} c_\theta & 0 \\ 0 & c_\theta^{-1} \end{pmatrix}, \quad \boldsymbol{Y} = \frac{\hbar}{2} \begin{pmatrix} s_\theta^2 e^{-2r} & 0 \\ 0 & \eta^{-1} \tan^2\theta (1 - \eta) \end{pmatrix} \tag{B.52}$$

This is a Gaussian CPTP map, with the $\boldsymbol{X}$ part of the map corresponding to squeezing mode A in q by a factor $\cos\theta$. An analogous derivation is possible for squeezing in p by using a p-squeezed state, performing homodyne in q and feedforward displacement in p. Moreover, we can achieve complex squeezing value $re^{i\phi}$ by placing the squeezing circuit between phase shifters.

B.5.2 Gates that Employ Inline Squeezing

Given our examination of inline squeezing using squeezed vacuum ancillae, we now summarize valuable CV gates which employ inline squeezing.

Phase gate The CV quadratic phase gate $\hat{P}(s)$ is just a squeezing gate $\hat{S}(re^{i\phi})$ composed with a rotation gate $\hat{R}(\theta)$. The ideal decomposition is given by [84, 328]

$$\hat{P}(s) = \hat{R}(\theta)\hat{S}(re^{i\phi}), \quad \theta = \tan^{-1}\left(\tfrac{s}{2}\right), \ \phi = -\text{sign}(s)\frac{\pi}{2} - \theta, \ r = \cosh^{-1}\left(\sqrt{1+\tfrac{s^2}{4}}\right). \tag{B.53}$$

as shown in Fig. B.3 (b), with values provided for the GKP qubit phase gate. Further noise could be assumed by using a lossy rotation gate $\hat{\mathcal{L}}(\eta)\hat{R}(\theta)$. If we assume perfect rotation operations, then the performance of the phase gate is limited by the squeezing gate.

Entangling gates The CV SUM gate can be decomposed as in terms of beam-splitters $\hat{BS}(\theta)$ and squeezing [84], where the control mode is the first mode:

$$\hat{CX}(s) = \hat{BS}(\theta+\pi/2)[\hat{S}(r)\otimes\hat{S}(-r)]\hat{BS}(\theta), \ r = \sinh^{-1}\left(-\frac{s}{2}\right), \ \theta = \frac{1}{2}\tan^{-1}\left(-\frac{2}{s}\right). \tag{B.54}$$

A simple noise model is to add single-mode losses to the beam-splitter outputs $\hat{\mathcal{L}}(\eta)\otimes\hat{\mathcal{L}}(\eta)\hat{BS}(\theta)$. A CZ gate can be achieved by applied Fourier rotations on mode B both before and after the CX gate. The CZ gate is depicted in Fig. B.3 (c) with specific values provided for the GKP qubit CZ gate.

Amplifier channel Amplification could be a useful in tackling photon loss effects [15]. In the ideal average case, an amplifier modifies the covariance matrices and means in the following manner [109]:

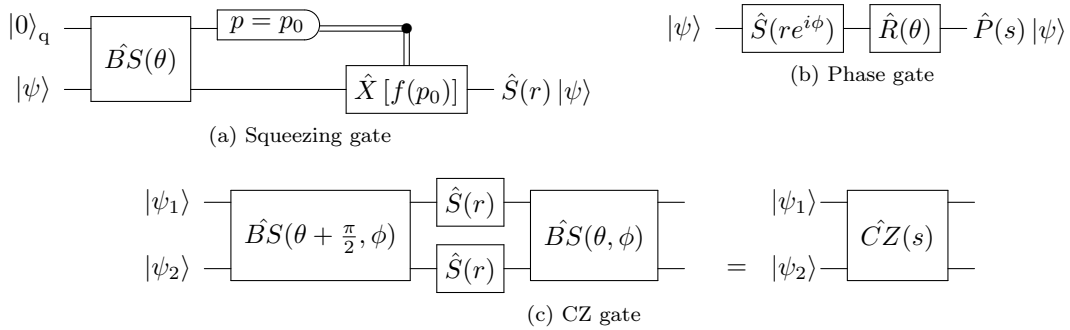
$$\boldsymbol{\Sigma} \to \kappa^2\boldsymbol{\Sigma} + (\kappa^2-1)\hbar\mathbb{1}/2, \quad \kappa > 1, \quad \boldsymbol{\mu} \to \kappa\boldsymbol{\mu}. \tag{B.55}$$

The ideal amplifier amplifies the input signal and adds a symmetric noise. To construct a realistic noise channel with $\kappa = \cosh r$, we use its Stinespring dilation to get [125]:

$$\hat{\mathcal{A}}(\kappa)[\hat{\rho}] = \text{tr}_B[\hat{S}_{0,1}(r)(\hat{\rho}\otimes|0\rangle\langle 0|_B)\hat{S}_{0,1}(r)^\dagger], \tag{B.56}$$

$$\hat{S}_{0,1}(r) = \hat{BS}(\pi/4)^\dagger[\hat{S}(r)\otimes\hat{S}(-r)]\hat{BS}(\pi/4), \tag{B.57}$$

where $\hat{S}_{0,1}(r)$ is a two-mode squeezing operation. Here, lossy beam-splitters and the noise we examined for inline squeezing form the dominant sources of error.



(a) Squeezing gate

(b) Phase gate

(c) CZ gate

Figure B.3: Review of CV gates and their realistic optical implementations. (a) The circuit for measurement-based squeezing [101]. An ancillary position eigenstate is combined at a beam-splitter of angle θ with the target mode. The ancilla is then measured in p quadrature with efficiency η, and informing a feedforward q displacement of $p_0 \tan\theta/\sqrt{\eta}$. The level of q-quadrature squeezing applied to the target mode is $r = \ln\cos\theta$. In practice, the quality of measurement-based squeezing is dependent on the level of squeezing of the ancillary state, and loss in the beam-splitter. Squeezing in p can be achieved with a momentum eigenstate, a q quadrature measurement, and a p displacement. Complex squeezing parameters can be implemented by sandwiching the circuit between two phase shifters. (b) A CV phase gate applied using rotations and inline squeezing [84]. Here, $r = \operatorname{arccosh}\sqrt{1+\frac{s^2}{4}}$, $\theta = \arctan\frac{s}{2}$, and $\phi = -\frac{\pi}{2}\mathrm{sign}(s) - \theta$. (c) A CV CZ gate [84], with $r = \operatorname{arcsinh}-\frac{s}{2}$, $\phi = \frac{\pi}{2}$, $\sin 2\theta = (\cosh r)^{-1}$, and $\cos 2\theta = \tanh r$. For the GKP encoding, the qubit phase and CZ gates corresponds to $s = 1$. In practice, the inline squeezing in the phase and CZ gates can be implemented using measurement-based squeezing, and additional noise in the CZ gate can be modelled by lossy beam-splitters.

Appendix C

Supplementary Material for Noise Analysis for a Fault-Tolerant Photonic Quantum Computer

C.1 Noise Model for a Hybrid RHG Lattice Operating as a Memory

In this Appendix, we provide the full details of the error model summarized in Section 4.3.1, that we in turn used to justify our choice of inner decoder.

C.1.1 Noisy Initial States

Additive Gaussian Noise Channel

The cluster state that the hardware generates will be populated by two kinds of states as mentioned in the main text: the $|+\rangle_{\text{gkp}}$ state and the momentum-squeezed state. Since the computer is operating in memory mode, we do not need to consider magic states as one of the possible states prepared. The position wavefunction of the ideal GKP state is [13]

$$|+\rangle_{\text{gkp}} = \sum_{n=-\infty}^{\infty} |n\sqrt{\pi}\rangle_q, \tag{C.1}$$

where $|\cdot\rangle_q$ corresponds to a position eigenstate. To model the state initialization error, we apply the single-mode additive Gaussian noise channel given by [196]

$$\mathcal{N}_{\boldsymbol{Y}}(\hat{\rho}) = \int_{\mathbb{R}^2} \frac{d^2\boldsymbol{\xi}}{\pi\sqrt{\det \boldsymbol{Y}}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{Y}^{-1}\boldsymbol{\xi}\right] \hat{\mathcal{D}}(\boldsymbol{\xi})\hat{\rho}\hat{\mathcal{D}}^\dagger(\boldsymbol{\xi}), \tag{C.2}$$

where $\boldsymbol{Y} \geq 0$ is the noise matrix, applied independently on each mode depending on the state that populates it. The Weyl-Heisenberg displacement operator is defined as $\hat{\mathcal{D}}(\boldsymbol{\xi}) = \exp[i\boldsymbol{\xi}^T\boldsymbol{\Omega}\hat{\boldsymbol{r}}]$, where $\boldsymbol{\xi} = (\xi_q, \xi_p)^T \in \mathbb{R}^2$ for a single mode, $\boldsymbol{\Omega}$ is the anti-symmetric symplectic metric, and $\hat{\boldsymbol{r}} = (\hat{q}, \hat{p})^T$. For the GKP states and the

momentum states, the corresponding noise matrices are chosen as

$$\boldsymbol{Y}_{\text{gkp}} = \frac{1}{2}\begin{pmatrix} \delta & 0 \\ 0 & \delta \end{pmatrix}, \quad \boldsymbol{Y}_{\text{p}} = \frac{1}{2}\begin{pmatrix} \delta^{-1} & 0 \\ 0 & \delta \end{pmatrix}. \tag{C.3}$$

In other words, we start with ideal $|+\rangle_{\text{gkp}}$ states and either apply the noise channel $\mathcal{N}_{\boldsymbol{Y}_{\text{gkp}}}$ or $\mathcal{N}_{\boldsymbol{Y}_{\text{p}}}$ with probabilities $1-p_0$ and p_0, respectively. Note that we know what the state is at each site, so we know which noise channel was applied. For GKP states, although this noise model does not capture the damping of peaks due to finite energy seen in Eq. (3.20), it captures the broadening of peaks that results from finite-energy effects [34, 47]. Similarly, this method approximates the realistic momentum state well in the position basis but gives it a periodic structure in momentum space. These points can be viewed transparently through the Wigner picture. In both cases, the application of a noise channel renders the output states mixed.

The Wigner Picture

The Wigner function for ideal GKP states consists of a linear combination of two-dimensional δ-functions in phase space [13]:

$$\begin{aligned} W_{|+\rangle_{\text{gkp}}}(\boldsymbol{r}) &= \frac{\sqrt{\pi}}{2} \sum_{s,t=-\infty}^{\infty} (-1)^{st} \delta(\boldsymbol{r} - \boldsymbol{\mu}_{s,t}), \\ \boldsymbol{\mu}_{s,t}^{T} &= \frac{\sqrt{\pi}}{2}(s, 2t). \end{aligned} \tag{C.4}$$

Note that the lattice spacing of the Dirac-delta peaks in the momentum direction is twice that of the position direction in the Wigner picture. For a clear diagram of the phase space unit cell distribution, see Fig. 1a of [36]. Treating each δ-function as a Gaussian of infinitely small width in phase space, we see that the effect of the noise channel is to replace the δ-functions with Gaussian distributions with covariance $\boldsymbol{Y}_{\text{gkp}}$, by using Eq. (C.2). Thus, the linear combination of δ-functions is mapped to a linear combination of Gaussian functions centred at the same points in phase space and with the same weights in the linear combination of Eq. (C.4).

With regard to the momentum states in the RHG lattice, consider the same noise model of Eq. (C.2), now instead with covariance

$$\boldsymbol{Y}_{\text{p}}(\epsilon, \delta) = \frac{1}{2}\begin{pmatrix} \epsilon^{-1} & 0 \\ 0 & \delta \end{pmatrix}. \tag{C.5}$$

Returning again to the Wigner function picture for the $|+\rangle_{\text{gkp}}$ state, this noise replaces the δ-functions with Gaussians of covariance $\boldsymbol{Y}_{\text{p}}(\epsilon, \delta)$. In the limit that $\epsilon \to 0$, we see that for odd values of t, these Gaussians in phase space cancel each other out, while for even values of t, the Gaussians add together. This is due to the phase factor $(-1)^{st}$ in Eq. (C.4). The resulting phase space distribution is a periodic *mixture* of p-quadrature eigenstates, each separated by $2\sqrt{\pi}$, and each with a Gaussian noise of variance $\delta/2$ applied in the p-quadrature. That is, the noise channel $\boldsymbol{Y}_{\text{p}}(\epsilon, \delta)$ turns $|+\rangle_{\text{gkp}}$ into a classical mixture of noisy p-squeezed states. Since we are examining the regime where δ is small, we set $\delta = \epsilon$, which leads to ϵ^{-1} being large as needed for the states to approach p-squeezed states. While it is the case that this noise model for momentum squeezed states returns a Wigner function which still, in principle, has a periodic structure, we do not expect it to positively bias the decoding procedure. The periodic structure in position space is essentially washed away

by the broadness of the envelope of order δ^{-1}, while the discrete $2\sqrt{\pi}$ translational symmetry in momentum introduces a mixture of momentum eigenstates, and hence more noise than a pure momentum squeezed state.

The initialization step for all N nodes can therefore be written compactly in one equation as:

$$\begin{aligned}\hat{\rho}_0 &= \mathcal{N}_{\boldsymbol{\Sigma}_0}(|+\rangle_{\text{gkp}}\langle+|^{\otimes N}) \\ &= \int_{\mathbb{R}^{2N}} \frac{d^{2N}\boldsymbol{\xi}}{\pi^N\sqrt{\det\boldsymbol{\Sigma}_0}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}^T\boldsymbol{\Sigma}_0^{-1}\boldsymbol{\xi}\right] \\ &\quad \times \hat{\mathcal{D}}(\boldsymbol{\xi})(|+\rangle\langle+|_{\text{gkp}}^{\otimes N})\hat{\mathcal{D}}^\dagger(\boldsymbol{\xi}),\end{aligned} \tag{C.6}$$

where $\boldsymbol{\xi} = (\boldsymbol{\xi_q}, \boldsymbol{\xi_p})^T = (\xi_{q_1}, \cdots, \xi_{q_N}, \xi_{p_1}, \cdots, \xi_{p_N})^T \in \mathbb{R}^{2N}$, and $\hat{\boldsymbol{r}} = (\hat{q}_1, ..., \hat{q}_N, \hat{p}_1, ..., \hat{p}_N)^T$, corresponding to N-modes. Here, $\boldsymbol{\Sigma}_0$ is a direct sum of matrices, where the i^{th} matrix in the direct sum is either of the form $\boldsymbol{Y}_{\text{gkp}}$ or $\boldsymbol{Y}_{\text{p}}$ depending on whether the i^{th} mode is a GKP or a p-squeezed state. In other words,

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} \boldsymbol{\Sigma}_x & \boldsymbol{0} \\ \boldsymbol{0} & \frac{\delta}{2}\mathbb{1} \end{pmatrix}, \tag{C.7}$$

where $\boldsymbol{\Sigma}_x$ is a diagonal matrix with elements $\frac{1}{2\delta}$ or $\frac{\delta}{2}$ depending on if the mode is p-squeezed or GKP, respectively.

There are several reasons to model the state preparation error with the noise channel described in Eqs. (C.2) and (C.3). For one, there are many physical gates that use a measurement-based squeezing operation [101, 329, 330] that naturally leads to imperfections modeled as the classical noise channel. Furthermore, this type of noise is closely related to pure loss—following a pure loss channel by an amplifier of the inverse strength leads to a classical noise channel [122, 123, 125, 331]. In settings where loss can be treated this way, such as in measurement imperfections, this relationship would play an important role.

The classical noise channel is easily described in the Heisenberg picture, so we use this representation in our simulations. Let us consider the quadrature operators $\hat{r}$ of the N-modes. The noise channel on each mode can be described as

$$\hat{\boldsymbol{r}} \to \hat{\boldsymbol{r}} + \boldsymbol{\xi}, \tag{C.8}$$

where $\boldsymbol{\xi}$ is a vector of random variables drawn from the corresponding normal distribution $\boldsymbol{\Sigma}_0$ associated with the state initialization errors.

A final note is that we assume that the CZ gates and the measurement procedure are noiseless. Imperfections in both these modules are likely to reduce the error threshold. Inefficiencies in the measurement outcomes can be modeled as a lossy channel and can be converted into a classical noise channel by virtually applying an amplifier that would affect the measurement readout. Similarly, classical noise channels in the CZ can also be tracked due to the Gaussian nature of the noise. However, for simplicity of the presentation, we leave the analysis of this case and possibly more complicated noise models to future work.

C.1.2 Propagation of Noise in the Cluster State Preparation

For each node, with probability p_0 prepare a momentum-squeezed state, and with probability $(1-p_0)$ prepare a $|+\rangle_{\text{gkp}}$. Next in our model, we apply CZ gates perfectly according to the structure of the cluster state, i.e., apply CZ gates to each pair of qubits connected by an edge. We invert some of the CZ gates to match the CV toric code convention [332]. Since we are operating in memory mode, no further gates are applied before

the p-homodyne measurements.

The symplectic transformation for a CZ gate defined as $\exp[i\hat{q_1}\hat{q_2}]$ in the (q_1, q_2, p_1, p_2) basis ordering is given by

$$\boldsymbol{S}_{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \boldsymbol{A} & \mathbb{1} \end{pmatrix}, \tag{C.9}$$

where $\boldsymbol{A}$ is a 2×2 adjacency matrix. This motivates the symplectic matrix that links all N optical modes into an RHG lattice as

$$\boldsymbol{S}_{RL} = \begin{pmatrix} \mathbb{1} & \mathbf{0} \\ \boldsymbol{A}_{RL} & \mathbb{1} \end{pmatrix}, \tag{C.10}$$

where $\boldsymbol{A}_{RL}$ is the $N \times N$ matrix with 1 at position (i, j) if two nodes are entangled with a CZ gate and 0 otherwise, with a suitable parity function dictated by the toric code convention. $\boldsymbol{A}_{RL}$ corresponds to the links depicted in Figs. 4.1 and 4.5.

It is also instructive to look at the effect of the cluster state preparation on the noise matrix. Under the action of all the CZ gates, the full noise matrix evolves under the symplectic transformation to

$$\boldsymbol{\Sigma}_0 \to \widetilde{\boldsymbol{\Sigma}}_0 = \boldsymbol{S}_{RL} \boldsymbol{\Sigma}_0 \boldsymbol{S}_{RL}^T \tag{C.11}$$

$$= \begin{pmatrix} \boldsymbol{\Sigma}_x & \boldsymbol{\Sigma}_x \boldsymbol{A}_{RL}^T \\ \boldsymbol{A}_{RL} \boldsymbol{\Sigma}_x & \frac{\delta}{2}\mathbb{1} + \boldsymbol{A}_{RL} \Sigma_x \boldsymbol{A}_{RL}^T \end{pmatrix}. \tag{C.12}$$

Since we are mainly concerned with the momentum homodyne measurement values, it turns out that the momentum component of the covariance matrix is useful to write down for subsequent sections. To achieve this, we trace out the position degrees of freedom of the covariance matrix of the noise channel to obtain

$$\widetilde{\boldsymbol{\Sigma}}_p = \frac{\delta}{2}\mathbb{1} + \boldsymbol{A}_{RL} \boldsymbol{\Sigma}_x \boldsymbol{A}_{RL}^T. \tag{C.13}$$

C.1.3 Probability Distribution in Momentum Space

So far we have only focused on the noise model and the correlated noise matrix that one obtains once all the CZ gates have been applied to the initial states in each mode. We now detail the connection between the noise matrix obtained in Eq. (C.2) and the homodyne distribution.

Let us define the unitary corresponding to the symplectic transformation in Eq. (C.10) as $\hat{U}_{RL}$. Since the preparation of the RHG cluster state and the noise channel on the initial states are both Gaussian, we can conjugate $\hat{U}_{RL}$ through the noise matrix to obtain

$$\begin{aligned} \hat{\rho}_{RL} &= \hat{U}_{RL} \left[\mathcal{N}_{\boldsymbol{\Sigma}_0} (|+\rangle \langle +|_{\text{gkp}}^{\otimes N}) \right] \hat{U}_{RL}^\dagger \\ &= \mathcal{N}_{\widetilde{\boldsymbol{\Sigma}}_0} \left[(\hat{U}_{RL} |+\rangle \langle +|_{\text{gkp}}^{\otimes N} \hat{U}_{RL}^\dagger) \right]. \end{aligned} \tag{C.14}$$

This corresponds to taking the ideal state of the RHG lattice had all the GKP states been initialized perfectly without noise and then applying a correlated multimode Gaussian noise channel with covariance $\widetilde{\boldsymbol{\Sigma}}_0$.

To understand the probability distribution produced by conjugating the unitary through the Gaussian noise channel, we first show that the probability distribution in p-quadrature of a state under a Gaussian

noise channel is given by the convolution of the noiseless probability distribution with the marginal Gaussian distribution of the noise channel along the p-quadrature.

Consider a Gaussian random displacement channel applied to a state:

$$\mathcal{N}_{\boldsymbol{\Sigma}}(\hat{\rho}) = \int_{\mathbb{R}^{2N}} \frac{d^{2N}\boldsymbol{\xi}}{\pi^N \sqrt{\det \boldsymbol{\Sigma}}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}\right] \hat{\mathcal{D}}(\boldsymbol{\xi}) \hat{\rho} \hat{\mathcal{D}}^\dagger(\boldsymbol{\xi}). \tag{C.15}$$

Next we note we can break the displacement into displacements $\hat{X}(\cdot)$ and $\hat{Z}(\cdot)$ along $\boldsymbol{q}$ and $\boldsymbol{p}$ in phase space, respectively, along with a phase factor:

$$\hat{\mathcal{D}}(\boldsymbol{\xi}) |\boldsymbol{p}\rangle = e^{i\boldsymbol{\xi}_{\boldsymbol{x}} \cdot \boldsymbol{\xi}_{\boldsymbol{p}}/2} \hat{X}(\boldsymbol{\xi}_{\boldsymbol{x}}) \hat{Z}(\boldsymbol{\xi}_{\boldsymbol{p}}) |\boldsymbol{p}\rangle \tag{C.16}$$

$$= e^{-i\boldsymbol{\xi}_{\boldsymbol{x}} \cdot \boldsymbol{\xi}_{\boldsymbol{p}}/2} e^{-i\boldsymbol{\xi}_{\boldsymbol{x}} \cdot \boldsymbol{p}} |\boldsymbol{p} + \boldsymbol{\xi}_{\boldsymbol{p}}\rangle . \tag{C.17}$$

Thus:

$$\begin{aligned} &\langle \boldsymbol{p}| \hat{\mathcal{D}}(\boldsymbol{\xi}) |\boldsymbol{p}'\rangle \langle \boldsymbol{p}''| \hat{\mathcal{D}}^\dagger(\boldsymbol{\xi}) |\boldsymbol{p}\rangle \\ =&\delta(\boldsymbol{p} - \boldsymbol{p}' - \boldsymbol{\xi}_{\boldsymbol{p}}) \delta(\boldsymbol{p} - \boldsymbol{p}'' - \boldsymbol{\xi}_{\boldsymbol{p}}) e^{i\boldsymbol{\xi}_{\boldsymbol{x}} \cdot (\boldsymbol{p}'' - \boldsymbol{p}')} \end{aligned} \tag{C.18}$$

Putting these equations together, we find:

$$\begin{aligned} &\langle \boldsymbol{p}| \mathcal{N}_{\boldsymbol{\Sigma}}(\hat{\rho}) |\boldsymbol{p}\rangle \\ &= \int_{\mathbb{R}^{2N}} \frac{d^{2N}\boldsymbol{\xi}}{\pi^N \sqrt{\det \boldsymbol{\Sigma}}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\xi}\right] \rho(\boldsymbol{p} - \boldsymbol{\xi}_{\boldsymbol{p}}, \boldsymbol{p} - \boldsymbol{\xi}_{\boldsymbol{p}}) \\ &= \int_{\mathbb{R}^N} \frac{d^N \boldsymbol{\xi}_{\boldsymbol{p}}}{\sqrt{\pi^N \det \boldsymbol{\Sigma}_p}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{p}}^T \boldsymbol{\Sigma}_p^{-1} \boldsymbol{\xi}_{\boldsymbol{p}}\right] \rho(\boldsymbol{p} - \boldsymbol{\xi}_{\boldsymbol{p}}, \boldsymbol{p} - \boldsymbol{\xi}_{\boldsymbol{p}}), \end{aligned} \tag{C.19}$$

where in the last step we performed the Gaussian integral over $\boldsymbol{\xi}_{\boldsymbol{x}}$.

Therefore, returning to the probability distribution in p-space of our hybrid lattice, we find:

$$\begin{aligned} &\langle \boldsymbol{p}| \mathcal{N}_{\widetilde{\boldsymbol{\Sigma}}_0}(\hat{U}_{RL} |+\rangle \langle +|_{\text{gkp}}^{\otimes N} \hat{U}_{RL}^\dagger) |\boldsymbol{p}\rangle \\ &= \int_{\mathbb{R}^N} \frac{d^N \boldsymbol{\xi}_{\boldsymbol{p}}}{\sqrt{\pi^N \det \widetilde{\boldsymbol{\Sigma}}_p}} \exp\left[-\frac{1}{2}\boldsymbol{\xi}_{\boldsymbol{p}}^T \widetilde{\boldsymbol{\Sigma}}_p^{-1} \boldsymbol{\xi}_{\boldsymbol{p}}\right] |\psi_{RL}(\boldsymbol{p} - \boldsymbol{\xi}_{\boldsymbol{p}})|^2, \end{aligned} \tag{C.20}$$

where $\psi_{RL}(\boldsymbol{p})$ is the wavefunction in p-space of the ideal RHG cluster state and $\widetilde{\boldsymbol{\Sigma}}_p$ was defined in Eq. (C.13). We know that $|\psi_{RL}(\boldsymbol{p})|^2$ consists of a lattice in p-space, where each point of the lattice is located at $\boldsymbol{n}\sqrt{\pi}$, where $\boldsymbol{n}$ is an integer-valued N-component vector chosen from a set dictated by the ideal qubit state of the RHG lattice. The addition of the Gaussian noise channel broadens each of these lattice points into Gaussian functions with covariance $\widetilde{\boldsymbol{\Sigma}}_p$. Therefore, we see that we can interpret homodyne momentum outcomes as being sampled from the noise matrix $\widetilde{\boldsymbol{\Sigma}}_p$ using Eq. (C.20). Given the measurement outcomes, we then apply a classical decoder to these values to yield us the net recovery operation.

C.2 Optical Components for GKP Qubit Operations

A primary advantage of the GKP qubit encoding is the fact that Clifford gates and measurements correspond to CV Gaussian gates and measurements. In Fig. C.1, we provide optical circuits for the application of

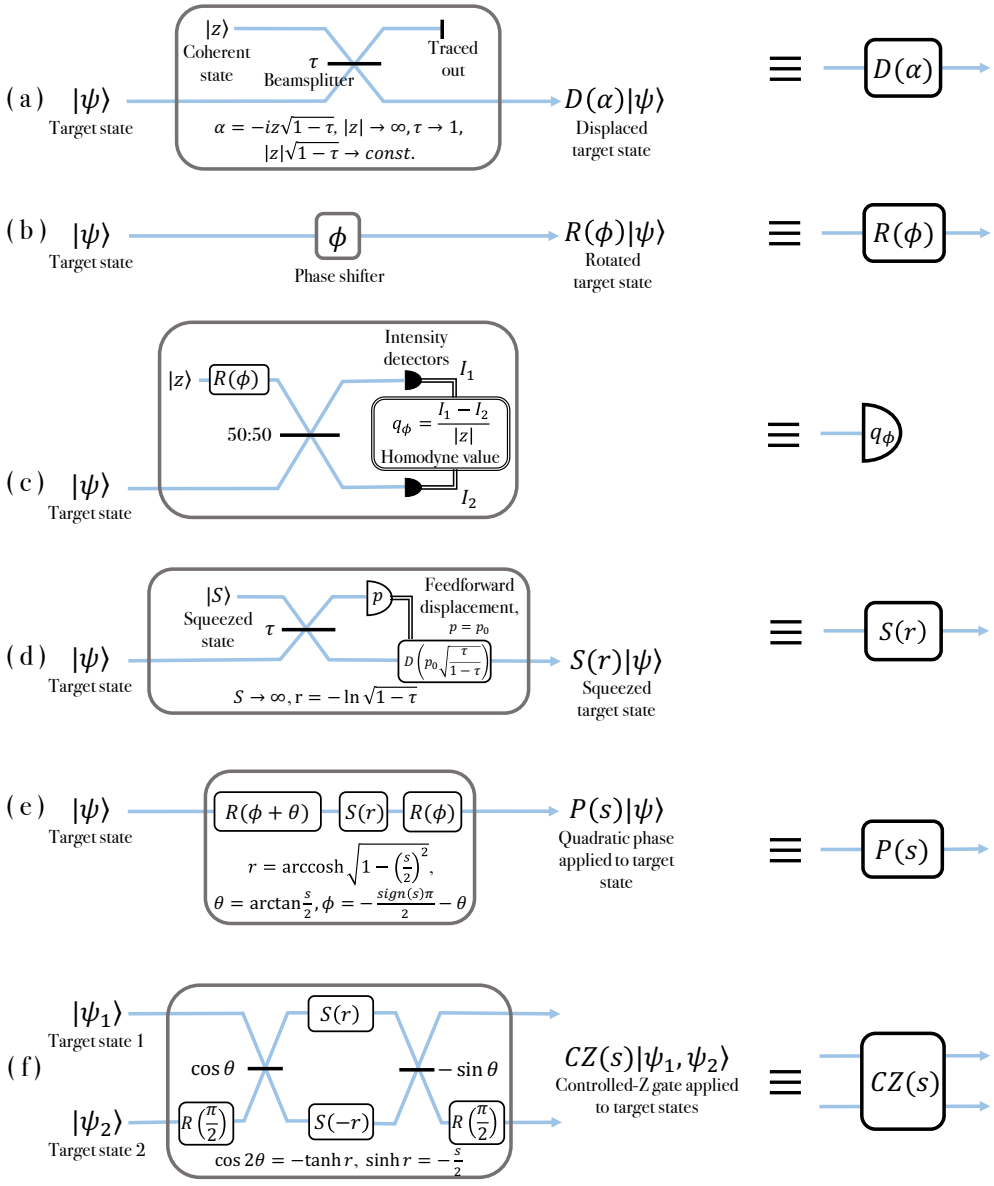


Figure C.1: A review of optical implementations of the gates and measurements required for Clifford operations in the GKP encoding, including limits required to achieve ideal, perfect CV gate application. (a) A general displacement module [333]. Displacement by $\sqrt{\pi}$ in q (p) corresponds to a GKP qubit Pauli X (Z) gate. (b) Rotation module as performed by e.g. an optical thermoelectric heating element. $\phi = \pi/2$ corresponds to the CV Fourier transform as well as the GKP-qubit Hadamard gate. (c) Homodyne measurement module. Changing the rotation ϕ changes the axis in phase space along which the measurement is performed. $\phi = 0$ ($\pi/2$) corresponds to q (p) homodyne measurement, which is the GKP qubit Pauli Z (X) measurement. (d) Measurement-based squeezing module [101]. On-demand, in-line squeezing is in general required for implementing CV quadratic phase and Controlled-X/phase gates, and a measurement-based approach allows for offline preparation of squeezed resource state. (e) Quadratic phase gate module [84]. $s = \pm 1$ corresponds to the GKP qubit phase gate. (f) CV CZ gate module. $s = \pm 1$ corresponds to the GKP qubit CZ gate. Application of $\pi/2$ rotations on the second mode before and after the CZ gate implements a CV CX gate [84] with Target state 1 becoming the control and Target state 2 becoming the target, and thus a GKP qubit $CNOT$ gate.

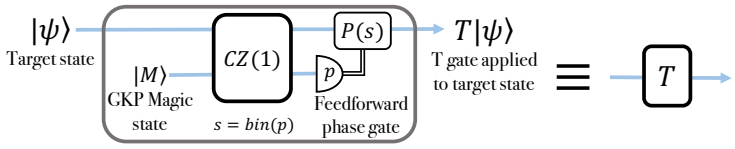


Figure C.2: Optical implementation of the GKP qubit T gate up to global phase, following the method from [13]. Here, in the ideal limit, $|M\rangle = e^{-i\pi/8}|+\rangle_{\text{gkp}} + e^{i\pi/8}|-\rangle_{\text{gkp}}$, and the feedforward phase gate is applied if the ancillary mode detects $|-\rangle_{\text{gkp}}$ via a qubit X measurement (CV p homodyne).

GKP Clifford gates and measurements in an optical setting. These circuits present how the gates would be implemented in a circuit-based setting. In contrast, the actual gates on our physical qubits are implemented in a measurement-based manner, and hence their implementation would only involve performing homodyne measurements on the computational resource state. Thus, this section is included for completeness of the background material and to demonstrate the accessibility of Gaussian resources in optics, rather than as the actual implementation of how the gates would be performed in our architecture.

For the non-Clifford T gate, non-Gaussian CV gates are required. In Fig. C.2, we provide an optical circuit for T gate application via gate teleportation using a GKP magic state as a resource. In [16], it was observed that magic states and Pauli basis states are comparably resource-intensive to produce using GBS state preparation.

C.3 Heuristic Weights for the Outer Decoder

In the outer decoder algorithm detailed in Section 4.3.3 and applied in Section 4.4.1, we have the opportunity to assign different weights in the outer decoder for the RHG lattice depending on the expected error at a site. In the most naÃ¯ve approach, one could simply use the marginal probability that a node undergoes a phase flip to determine the weight; however, this does not take into account correlated phase flips that we expect to see from replacing some nodes with p-squeezed states, as we have discussed previously. For instance, in either the case of applying a ring of four Z gates or of applying the identity, neither result in any error in the decoding, but if we simply looked at the marginal probabilities of the sites in the ring, we might pessimistically assume each site has an independent probability of incurring a phase flip, which would result in pessimistic weight assignment.

While a full analysis of correlated errors and weight assignments for general configurations of p-squeezed states and GKP states is left to future work, the heuristic choice for weight assignments given by Eq. (4.9) can be found by considering a simple configuration. Amazingly, this choice of weights already provides a significant improvement over the use of marginal probability of error at each site. Here, we detail the motivation for this choice of heuristic weights.

Consider a single node e_0 in the RHG lattice surrounded by four neighbors e_1, e_2, e_3, e_4, which can be either GKP or p-squeezed, so that e_0 can have anywhere from 0 to 4 p-squeezed states as neighbors. For simplicity, we will assume that the next layers of neighbors of e_1, e_2, e_3, e_4 are GKP states. See Figs. 4.1 and 4.5 for a visualization of the lattice configuration. Whether e_0 is GKP or p-squeezed does not impact the following argument. Additionally, we will assume the limit of infinite squeezing ($\delta \to 0$) for all the sites. Note that these assumptions are used to select a choice of weights, not to run the actual simulation, so

discrepancies between the assumptions for choosing weights and the actual parameters of the simulation will nonetheless result in a perfectly usable, albeit suboptimal, set of weights.

In this scenario, for each node that is a momentum eigenstate, it imparts a random displacement – sampled from the uniform distribution of its q-quadrature since we assumed $\delta \to 0$ – onto the p-quadrature of its neighbors via the action of the CV CZ gates, while each node that is a perfect GKP state does not impart any random displacements onto its neighbors. Let the displacements from e_1, e_2, e_3, e_4 be d_1, d_2, d_3, d_4, where we specifically mean the excess displacement beyond $n\sqrt{\pi}$ [47]. Thus, the displacement on e_0 is given by $d_1 + d_2 + d_3 + d_4$, assuming the CZ gates are all $+1$; changing the sign of the CZ gates does not change the argument since d_1, d_2, d_3, d_4 are sampled from symmetric distributions. Moreover, let b_i, $i = 1, 2, 3, 4$ be the binary value returned by performing standard GKP binning on the homodyne output of the neighbors of node e_i other than e_0; note that the only shared neighbor of e_1, e_2, e_3, e_4 is e_0, and since we assumed the nodes beyond e_1, e_2, e_3, e_4 were GKP, then we know we will get the same singular outcome b_i on all neighbors of e_i other than e_0. Let b_0 be the binary value returned by performing standard GKP binning on the homodyne output of e_0.

Consider now the scenarios that would cause an error at the level of the qubit decoder. We know that a closed ring of Z gates commutes with the stabilizers of the RHG lattice. If only one of e_1, e_2, e_3, e_4 is a momentum eigenstate, then we have already shown that the resulting effect on the binary outcomes b_i is a ring of four Z gates around e_i, which we know causes no problem. If two or more of e_1, e_2, e_3, e_4 are momentum eigenstates, then we now have potential for error when using standard binning. In particular, for the readout to correspond to a closed ring of Z gates, we require that $(b_0 + b_1 + b_2 + b_3 + b_4) \bmod 2 = 0$. This condition will always be true if only one of d_1, d_2, d_3, d_4 is sampled from uniform (since we assumed $\delta \to 0$) while the others are zero, since $d_1 + d_2 + d_3 + d_4$ will then be equal to the only non-zero displacement. We find that if two, three or four of d_1, d_2, d_3, d_4 are nonzero, then the condition $(b_0 + b_1 + b_2 + b_3 + b_4) \bmod 2 = 0$ is violated with probabilities of approximately $0.25, 0.33$, and 0.40, respectively. This means that even with multiple nodes replaced by p-squeezed states, the probability of the resulting readout indicating a series of gates that do not commute with the stabilizer is less than 50%, which would be the marginal probability of phase flip at each site. Finally, we use these probabilities of error to assign heuristic, relative weighting in the outer decoder.

Appendix D

Supplementary material for Entanglement-Based High-Dimensional QKD and the Measurement Range Problem

D.1 Basis-Independent Null Measurements Pose No Problem for Entropic Uncertainty Relations

We show that if the null measurement is independent of the measurement type, then we can always reformulate an entropic uncertainty relation that does not depend on the null measurement operator common to both measurement types. However, it depends on new, effective POVMs related to the original POVMs. We show how those can be constructed.

Let $\mathcal{H}$ be the total Hilbert space. Suppose we have two measurements, Z and X, characterized by POVMs $Z = \{\mathbb{Z}_n\}_{n=0}^{N_Z} \cup \mathbb{N}$ and $X = \{\mathbb{X}_m\}_{m=0}^{N_X} \cup \mathbb{N}$, where $\mathbb{N} = \mathbb{I} - \sum_{n=0}^{N_Z} \mathbb{Z}_n = \mathbb{I} - \sum_{m=0}^{N_X} \mathbb{X}_m = \mathbb{I} - \mathbb{M}$ represents the basis-independent null measurement.

We can define $\mathcal{H}_{\mathbb{N}}$ ($\mathcal{H}_{\mathbb{M}}$) to be the space on which $\mathbb{N}$ ($\mathbb{M}$) has support. In general, $\mathcal{H}_{\mathbb{N}}$ and $\mathcal{H}_{\mathbb{M}}$ will have overlap, with $\mathcal{H}_{\mathbb{N}} \cup \mathcal{H}_{\mathbb{M}} = \mathcal{H}$.

With the definition of those spaces in hand, we can treat some operators in the problem as the direct sum of operators with support only on $\mathcal{H}_{\mathbb{M}}$, and of the zero operator on $\mathcal{H}/\mathcal{H}_{\mathbb{M}}$:

$$\begin{aligned} \mathbb{M} &= (\mathbb{M})_{\mathcal{H}_{\mathbb{M}}} \oplus 0_{\mathcal{H}/\mathcal{H}_{\mathbb{M}}} \\ \mathbb{Z}_n &= (\mathbb{Z}_n)_{\mathcal{H}_{\mathbb{M}}} \oplus 0_{\mathcal{H}/\mathcal{H}_{\mathbb{M}}} \\ \mathbb{X}_m &= (\mathbb{X}_m)_{\mathcal{H}_{\mathbb{M}}} \oplus 0_{\mathcal{H}/\mathcal{H}_{\mathbb{M}}} \end{aligned} \tag{D.1}$$

where $(\cdot)_{\mathcal{H}_{\mathbb{M}}}$ denotes an operator with only support on $\mathcal{H}_{\mathbb{M}}$.

Claim: $[\sqrt{(\mathbb{M})_{\mathcal{H}_{\mathbb{M}}}}]^{-1}$ is well-defined.

Proof: As a POVM element, $\mathbb{M}$ is Hermitian and non-negative, so all its eigenvalues are real and non-

negative. Moreover, when we consider the operator $(\mathbb{M})_{\mathcal{H}_\mathbb{M}}$, we have now restricted to only the eigenvalues that are positive. Thus, the square root of $(\mathbb{M})_{\mathcal{H}_\mathbb{M}}$ is well-defined, and invertible. ■

We can now define new operators:

$$\begin{aligned}\tilde{\mathbb{Z}}_n &= \big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}(\mathbb{Z}_n)_{\mathcal{H}_\mathbb{M}}\big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}\\ \tilde{\mathbb{X}}_m &= \big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}(\mathbb{X}_m)_{\mathcal{H}_\mathbb{M}}\big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}.\end{aligned} \tag{D.2}$$

Claim: $\tilde{Z} = \{\tilde{\mathbb{Z}}_n\}_{n=0}^{N_Z}$ and $\tilde{X} = \{\tilde{\mathbb{X}}_m\}_{m=0}^{N_X}$ each form a POVM on $\mathcal{H}_\mathbb{M}$.

Proof: To show $\tilde{Z}$ is a POVM we need to show its elements are Hermitian and non-negative, and that they sum to $\mathbb{I}_\mathbb{M}$. Because $(\mathbb{Z}_n)_{\mathcal{H}_\mathbb{M}}$ and $[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}]^{-1}$ are Hermitian and non-negative, then $\tilde{\mathbb{Z}}_n = \big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}(\mathbb{Z}_n)_{\mathcal{H}_\mathbb{M}}\big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}$ is also Hermitian and non-negative. Moreover, we can check:

$$\begin{aligned}\sum_{n=0}^{N_Z}\tilde{\mathbb{Z}}_n =& \big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}\sum_{n=0}^{N_Z}(\mathbb{Z}_n)_{\mathcal{H}_\mathbb{M}}\big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}\\ =& \big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1}(\mathbb{M})_{\mathcal{H}_\mathbb{M}}\big[\sqrt{(\mathbb{M})_{\mathcal{H}_\mathbb{M}}}\big]^{-1} = \mathbb{I}_\mathbb{M}.\end{aligned} \tag{D.3}$$

The same process follows for $\tilde{X}$. ■

Now that we have two new effective POVMs, we will use them to understand an equivalence between two different scenarios:

Scenario 1 (a "realistic" model): An untrusted source outputs to Alice a state ρ that could be entangled with Eve and/or Bob. With probability q, Alice performs a Z measurement, and with probability $(1-q)$ Alice performs an X measurement. Alice knows what type of measurement she is performing. Both Z and X measurements include the null measurement element, $\mathbb{N}$. Depending on if she performs a Z measurement or an X measurement, Alice records the result in separate registers, R_Z or R_X, respectively. After many iterations, Alice considers two probability distributions: $p(m|\rho, X, \neg\mathbb{N})$ ($p(n|\rho, Z, \neg\mathbb{N})$) is the distribution of measurement results given she performed an X (Z) measurement on the incoming state, ρ, and post-selected to discard the null measurement. This distribution represents the state of the register, R_X (R_Z), having discarded null measurements. The POVMs characterizing the Z and X measurements have been outlined in the previous section.

Scenario 2 (a fictitious model): An untrusted source outputs the same state ρ, but before it is sent to Alice, Eve passes it through a filter that blocks the state if it returns the null measurement, and lets it pass if it does not return the null measurement. The state Alice then receives is:

$$\rho' = \frac{\sqrt{\mathbb{M}}\rho\sqrt{\mathbb{M}}}{\mathrm{Tr}(\rho\mathbb{M})}. \tag{D.4}$$

Note that this state is only non-zero on $\mathcal{H}_\mathbb{M}$ because it filters out the portion of the state that has support on $\mathcal{H}/\mathcal{H}_\mathbb{M}$.

Now Alice performs either one of the two new measurements we defined, $\tilde{Z}$ or $\tilde{X}$, which are related to, but different from, Z and X. Similarly to before, depending on if she performs a $\tilde{Z}$ measurement or an $\tilde{X}$ measurement, Alice records the result in separate registers, $\tilde{R}_Z$ or $\tilde{R}_X$, respectively. After many iterations, Alice considers two probability distributions: $p(m|\rho', \tilde{X})$ ($p(n|\rho', \tilde{Z})$) is the distribution of measurement results given that an $\tilde{X}$ ($\tilde{Z}$) measurement was performed on the incoming state, ρ'. This distribution represents the state of the register, $\tilde{R}_X$ ($\tilde{R}_Z$).

One can check that:

$$p(m|\rho', \tilde{X}) = p(m|\rho, X, \neg\mathbb{N}) = \frac{\mathrm{Tr}(\rho\mathbb{X}_m)}{\mathrm{Tr}(\rho\mathbb{M})}$$
$$p(n|\rho', \tilde{Z}) = p(n|\rho, Z, \neg\mathbb{N}) = \frac{\mathrm{Tr}(\rho\mathbb{Z}_n)}{\mathrm{Tr}(\rho\mathbb{M})}. \tag{D.5}$$

Thus, because Alice cannot distinguish between these two distributions, we can assume Eve can have control of when a null measurement outcome will occur.

We can now examine an entropic uncertainty relation for Scenario 2. Note that an entropic uncertainty relation for Scenario 1 would lead to the trivial lower bound of 0 on $H_{\min}(Z|E)$ because of the overlap between the common null measurement element in both POVM sets: $c(X, Z) = ||\mathbb{N}||_\infty^2 = 1$ because the eigenvalues of $\mathbb{N}$ on $\mathcal{H}/\mathcal{H}_\mathbb{M}$ are all 1.

Because $\tilde{X}$ and $\tilde{Z}$ are both POVMs on $\mathcal{H}_\mathbb{M}$, the only space on which ρ' can have support, then:

$$H_{\min}(\tilde{Z}|E)_{\rho'} + H_{\max}(\tilde{X}|B)_{\rho'} \geq -\log \max_{m,n} ||\sqrt{\tilde{\mathbb{Z}}_n}\sqrt{\tilde{\mathbb{X}}_m}||^2.$$

Hence, we have provided a proof that basis-independent null measurements do not pose a problem for entanglement-based QKD protocols.

Note that the same process cannot be done with basis-dependent null measurements since there would not necessarily be a common POVM element $\mathbb{N}$. Thus, defining the subspace, $\mathcal{H}_\mathbb{M}$, would not be possible, precluding the definition of new effective POVMs, $\tilde{Z}$ and $\tilde{X}$, and the reduction to an entropic uncertainty relation in terms of just those operators.

D.2 Proof of Main Result, Eq. 5.6

We assume a tripartite state, ρ_{ABE}, and two POVMs on $\mathcal{H}_A$, $Z = \{\mathbb{Z}_A^z\}_{z=1}^{N_Z} \cup \{\mathbb{Z}_A^\emptyset\}$ and $X = \{\mathbb{X}_A^x\}_{x=1}^{N_X} \cup \{\mathbb{X}_A^\emptyset\}$. We begin by following the procedure from [243]. Define $\rho_{ZZ'ABE} = V_{ZZ'}\rho_{ABE}V_{ZZ'}^\dagger$, where $V_{ZZ'} = \sum_{z=1}^{N_Z} |z\rangle_Z|z\rangle_{Z'}\sqrt{\mathbb{Z}_A^z} + |\emptyset\rangle_Z|\emptyset\rangle_{Z'}\sqrt{\mathbb{Z}_A^\emptyset}$ is an isometry that maps the input state to a state entangled with the Z register. As in [243], we start from the duality relation:

$$H_{\min}(Z|E) + H_{\max}(Z|Z'AB) \geq 0. \tag{D.6}$$

Based on the definition for conditional max entropy in Eq. (5.3), we see our task is to upper bound $\max_{\sigma_{Z'AB}} F(\rho_{ZZ'AB}, \mathbb{I}_Z \otimes \sigma_{Z'AB})$. One of our objectives is to remove dependence on $\mathbb{Z}_A^\emptyset$, so we first prove a useful lemma.

Lemma: For a composite Hilbert space, $\mathcal{H}_C \otimes \mathcal{H}_D$, if $\sqrt{\Gamma_C}\sqrt{\mathbb{I}_C - \Gamma_C} = 0$ for some positive operator Γ_C on $\mathcal{H}_C$ such that $\mathbb{I}_C - \Gamma_C$ is also positive, then $F(\rho_{CD}, \mathbb{I}_C \otimes \sigma_D) \leq F(\rho_{CD}, \Gamma_C \otimes \sigma_D) + F(\rho_{CD}, (\mathbb{I}_C - \Gamma_C) \otimes \sigma_D)$.

Proof: Consider the trace norm formulation of the fidelity, and recall the triangle inequality for norms. Thus,

$$F(\rho_{CD}, \mathbb{I}_C \otimes \sigma_D) = ||\sqrt{\rho_{CD}}\sqrt{\mathbb{I}_C \otimes \sigma_D}||_{\mathrm{Tr}} \leq ||\sqrt{\rho_{CD}}\sqrt{\Gamma_C \otimes \sigma_D}||_{\mathrm{Tr}} + ||\sqrt{\rho_{CD}}\sqrt{(\mathbb{I}_C - \Gamma_C) \otimes \sigma_D}||_{\mathrm{Tr}} \tag{D.7}$$

on condition we can write $\sqrt{\mathbb{I}_C \otimes \sigma_D} = \sqrt{\Gamma_C \otimes \sigma_D} + \sqrt{(\mathbb{I}_C - \Gamma_C) \otimes \sigma_D}$. Consider the square of both sides:

$$\begin{aligned} \mathbb{I}_C \otimes \sigma_D &= \mathbb{I}_C \otimes \sigma_D + \sqrt{\Gamma_C \otimes \sigma_D}\sqrt{(\mathbb{I}_C - \Gamma_C) \otimes \sigma_D} + \sqrt{(\mathbb{I}_C - \Gamma_C) \otimes \sigma_D}\sqrt{\Gamma_C \otimes \sigma_D} \\ &= \mathbb{I}_C \otimes \sigma_D + \sqrt{\Gamma_C}\sqrt{(\mathbb{I}_C - \Gamma_C)} \otimes \sigma_D + \sqrt{(\mathbb{I}_C - \Gamma_C)}\sqrt{\Gamma_C} \otimes \sigma_D. \end{aligned} \tag{D.8}$$

We see the equality holds if $\sqrt{\Gamma_C}\sqrt{\mathbb{I}_C - \Gamma_C} = 0$. ■

Note that $\mathbb{I}_Z \otimes \sigma_{Z'AB} = |\emptyset\rangle\langle\emptyset|_Z \otimes \sigma_{Z'AB} + \sum_{z=1}^{N_Z} |z\rangle\langle z|_Z \otimes \sigma_{Z'AB}$, and $\sqrt{|\emptyset\rangle\langle\emptyset|_Z}\sqrt{\sum_{z=1}^{N_Z} |z\rangle\langle z|_Z} = 0$, so we use the lemma to find:

$$F(\rho_{ZZ'AB}, \mathbb{I}_Z \otimes \sigma_{Z'AB}) \leq F(\rho_{ZZ'AB}, |\emptyset\rangle\langle\emptyset|_Z \otimes \sigma_{Z'AB}) + F(\rho_{ZZ'AB}, \sum_{z=1}^{N_Z} |z\rangle\langle z|_Z \otimes \sigma_{Z'AB}). \tag{D.9}$$

Using the data-processing inequality for fidelities [242], $F(\rho, \sigma) \leq F[\mathcal{E}(\rho), \mathcal{E}(\sigma)]$, where $\mathcal{E}(\cdot)$ is a trace-preserving, completely positive map, we find:

$$F(\rho_{ZZ'AB}, |\emptyset\rangle\langle\emptyset|_Z \otimes \sigma_{Z'AB}) \leq F(\rho_Z, |\emptyset\rangle\langle\emptyset|_Z) = \sqrt{p^{\emptyset}_{Z_A}}. \tag{D.10}$$

Next, we use the fact from [243] that relative entropies are invariant under isometries. Since the max relative entropy is simply proportional to the logarithm of the fidelity, this means fidelity is also invariant under isometries. Following the process done in equation (6) of the supplementary material of [243], we get:

$$\begin{aligned} F(\rho_{ZZ'AB}, \sum_{z=1}^{N_Z} |z\rangle\langle z|_Z \otimes \sigma_{Z'AB}) &\leq F(\rho_{AB}, V^{\dagger}_{ZZ'} \sum_{z=1}^{N_Z} |z\rangle\langle z|_Z \otimes \sigma_{Z'AB} V_{ZZ'}) \\ &= F\big[\rho_{AB}, \mathrm{Tr}_{Z'}(\sum_{z=1}^{N_Z} |z\rangle\langle z| \sqrt{\mathbb{Z}^z_A}\sigma_{Z'AB}\sqrt{\mathbb{Z}^z_A})\big]. \end{aligned} \tag{D.11}$$

Note that $\mathrm{Tr}_{Z'}(\sum_{z=1}^{N_Z} |z\rangle\langle z| \sqrt{\mathbb{Z}^z_A}\sigma_{Z'AB}\sqrt{\mathbb{Z}^z_A})$ may not be a normalized density matrix, but we will fix this later.

Now, we define the isometry associated with the X-type measurement, $V_{XX'} = \sum_{x=1}^{N_X} |x\rangle_X |x\rangle_{X'} \sqrt{\mathbb{X}^x_A} + |\emptyset\rangle_X |\emptyset\rangle_{X'} \sqrt{\mathbb{X}^{\emptyset}_A}$, use again the fact that fidelity is invariant under isometries, and again use the data-processing inequality to trace over subsystems A and X':

$$\begin{aligned} &F\big[\rho_{AB}, \mathrm{Tr}_{Z'}(\sum_{z=1}^{N_Z} |z\rangle\langle z| \sqrt{\mathbb{Z}^z_A}\sigma_{Z'AB}\sqrt{\mathbb{Z}^z_A})\big] \\ &\leq F\Big\{\rho_{XB}, \sum_x |x\rangle\langle x|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z} \Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}^z_A}\mathbb{X}^x_A\sqrt{\mathbb{Z}^z_A}\Big)\sigma_{Z'AB}\Big]\Big\} \end{aligned} \tag{D.12}$$

where $\rho_{XB} = \sum_x |x\rangle\langle x|_X \otimes \mathrm{Tr}_A\big(\rho_{AB}\mathbb{X}_A^x\big)$. We now note:

$$\begin{aligned} & F\Big\{\rho_{XB}, \sum_x |x\rangle\langle x|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^x\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\} \\ &= F\Big\{|\emptyset\rangle\langle\emptyset|_X \otimes \mathrm{Tr}_A\big(\rho_{AB}\mathbb{X}_A^\emptyset\big), |\emptyset\rangle\langle\emptyset|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^\emptyset\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\} \\ &+F\Big\{\sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_A\big(\rho_{AB}\mathbb{X}_A^x\big), \sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^x\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\}. \end{aligned} \tag{D.13}$$

Applying the data-processing inequality to the first term to trace over B:

$$\begin{aligned} & F\Big\{|\emptyset\rangle\langle\emptyset|_X \otimes \mathrm{Tr}_A\big(\rho_{AB}\mathbb{X}_A^\emptyset\big), |\emptyset\rangle\langle\emptyset|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^\emptyset\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\} \\ &\leq F(p_{X_A}^\emptyset|\emptyset\rangle\langle\emptyset|_X, |\emptyset\rangle\langle\emptyset|_X) = \sqrt{p_{X_A}^\emptyset}. \end{aligned} \tag{D.14}$$

Additionally, we can define $\rho_{X<B} = \sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_A\big(\rho_{AB}\mathbb{X}_A^x\big)/(1-p_{X_A}^\emptyset)$ to be a normalized density operator, so that:

$$\begin{aligned} & F\Big\{\sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_A\big(\rho_{AB}\mathbb{X}_A^x\big), \sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^x\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\} \\ &= \sqrt{1-p_{X_A}^\emptyset}F\Big\{\rho_{X<B}, \sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^x\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\}. \end{aligned} \tag{D.15}$$

Finally, we know from [243] that if $\tilde{\sigma} \geq \sigma$, then $F(\rho,\tilde{\sigma}) \geq F(\rho,\sigma)$, so:

$$\begin{aligned} & F\Big\{\rho_{X<B}, \sum_{x=1}^{N_X}|x\rangle\langle x|_X \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}\Big(|z\rangle\langle z|_{Z'} \otimes \sqrt{\mathbb{Z}_A^z}\mathbb{X}_A^x\sqrt{\mathbb{Z}_A^z}\Big)\sigma_{Z'AB}\Big]\Big\} \\ &\leq \max_{(x,z)\neq\emptyset}||\sqrt{\mathbb{Z}_A^z}\sqrt{\mathbb{X}_A^x}||_\infty F\Big\{\rho_{X<B}, \mathbb{I}_{X<} \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}(|z\rangle\langle z|_{Z'})\sigma_{Z'AB}\Big]\Big\}. \end{aligned} \tag{D.16}$$

Applying this property again for $\sigma_B \geq \mathrm{Tr}_{Z'A}[\sum_{z=1}^{N_Z}(|z\rangle\langle z|_{Z'})\sigma_{Z'AB}]$, we get:

$$\begin{aligned} \max_{\sigma_{Z'AB}} F\Big\{\rho_{X<B}, \mathbb{I}_{X<} \otimes \mathrm{Tr}_{Z'A}\Big[\sum_{z=1}^{N_Z}(|z\rangle\langle z|_{Z'})\sigma_{Z'AB}\Big]\Big\} &\leq \max_{\sigma_B} F(\rho_{X<B}, \mathbb{I}_{X<} \otimes \sigma_B) \\ &= \sqrt{2}^{H_{\max}(X_A^<|B)}. \end{aligned} \tag{D.17}$$

Putting together Eq. (D.6), and (D.9)-(D.17), we get the result in Eq. (5.6). ■

D.3 Smooth Version of Main Result

Smooth min- and max- entropies are useful quantities for incorporating finite key size effects [12, 230, 240–242, 247]. ε-smooth conditional min- and max-entropies are defined as:

$$H_{\min}^{\varepsilon}(A|B)_{\rho} = \max_{\rho' \in \mathcal{B}^{\varepsilon}(\rho)} H_{\min}(A|B)_{\rho'}, \; H_{\max}^{\varepsilon}(A|B)_{\rho} = \min_{\rho' \in \mathcal{B}^{\varepsilon}(\rho)} H_{\max}(A|B)_{\rho'} \tag{D.18}$$

where $\mathcal{B}^{\varepsilon}(\rho) = \{\rho' | \; \frac{1}{2}||\rho - \rho'||_{\text{Tr}} \leq \varepsilon\}$ denotes the set of operators within an ε-distance of ρ.

Proceeding similarly to the proof from [241], we first note that for some $\tau \in \mathcal{B}^{\varepsilon}(\rho)$, $H_{\max}^{\varepsilon}(X_A^{\leq}|B)_{\rho} = H_{\max}(X_A^{\leq}|B)_{\tau}$. Thus, we can write down our main result, Eq. (5.6) for that state:

$$H_{\min}(Z_A|E)_{\tau} \geq -2\log\left[\sqrt{p_{Z_A}^{\emptyset}(\tau)} + \sqrt{p_{X_A}^{\emptyset}(\tau)} + \sqrt{1 - p_{X_A}^{\emptyset}(\tau)}\sqrt{c^{<}(X,Z)}\left(\sqrt{2}^{H_{\max}^{\varepsilon}(X_A^{\leq}|B)_{\rho}}\right)\right] \tag{D.19}$$

where the $p_i^{\emptyset}(\tau)$ denote the null measurement probabilities given the state τ.

Next, we would like to express $p_i^{\emptyset}(\tau)$ in terms of $p_i^{\emptyset}(\rho)$ and ε to determine how much the probabilities from the two states can differ. Using $1 - F(\rho,\tau) \leq \frac{1}{2}||\rho - \tau||_{\text{Tr}}$ [240], and the data-processing inequality [242], we find:

$$f_{-}[p_i^{\emptyset}(\rho), \varepsilon] \leq p_i^{\emptyset}(\tau) \leq f_{+}[p_i^{\emptyset}(\rho), \varepsilon] \tag{D.20}$$

where

$$f_{\pm}[p_i^{\emptyset}(\rho), \varepsilon] = 2\varepsilon + p_i^{\emptyset}(\rho) + 2p_i^{\emptyset}(\rho)\varepsilon^2 - 4p_i^{\emptyset}(\rho)\varepsilon - \varepsilon^2 \pm 2(1-\varepsilon)\sqrt{p_i^{\emptyset}(\rho)\varepsilon[1 - p_i^{\emptyset}(\rho)][2-\varepsilon]}. \tag{D.21}$$

Finally, following [241], knowing that $H_{\min}^{\varepsilon}(Z_A|E)_{\rho} = \max_{\rho' \in \mathcal{B}^{\varepsilon}(\rho)} H_{\min}(Z_A|E)_{\rho'} \geq H_{\min}(Z_A|E)_{\tau}$, we get the smooth version of Eq. (5.6):

$$\begin{aligned} H_{\min}^{\varepsilon}(Z_A|E)_{\rho} \geq -2\log\Bigg[&\sqrt{f_{+}[p_{Z_A}^{\emptyset}(\rho), \varepsilon]} + \sqrt{f_{+}[p_{X_A}^{\emptyset}(\rho), \varepsilon]} + \\ &\sqrt{1 - f_{-}[p_{X_A}^{\emptyset}(\rho), \varepsilon]}\sqrt{c^{<}(X,Z)}\left(\sqrt{2}^{H_{\max}^{\varepsilon}(X_A^{\leq}|B)_{\rho}}\right)\Bigg]. \end{aligned} \tag{D.22}$$

Appendix E

Supplemental material for Loss-tolerant QKD with a Twist

E.1 Embedding our Technique Within a Decoy State Protocol

In the main text, we showed that, using two independent semidefinite programs (SDP), we can optimize the secret key rate of a loss-tolerant protocol where the signal states are two-dimensional mixed states. In the context of quantum key distribution, those states are normally encoded in the mode (such as polarization or time-bin) of single photons. However, in practice, many protocols employ weak coherent pulses with decoy states [271, 272]. In this section, we outline how we can embed our technique in a loss-tolerant protocol which uses decoy states. In this chapter, we will consider a protocol with infinite number of decoy states as in [25]. We leave the case where a finite number of decoy states are used for future work. Like in the main text, we work in the asymptotic limit where we can ignore finite key effects.

Recall from the main text that to use our technique, we need two pieces of information: before sending optical signals to Eve, we need the density matrices of the qubit signal states (the single photon component), and then once the optical signals are sent, we need the probability of successful Bell state measurement given the states that Alice and Bob chose. As such, to apply our technique when phase-randomised weak coherent pulses are used together with a decoy state protocol, we need to calculate the state of the single photon component of the optical signal as well as the probability of successful Bell state measurement given that Alice and Bob send the corresponding single photon signals. In the literature, that conditional probability is often referred to as the single photon yield, denoted by $Y_{11}^{i,j,x,y}$.

To obtain the single photon component of the signals, we can simply project the coherent signal states to their single photon components. Suppose that Alice and Bob prepare the phase-randomised coherent state $\tilde{\rho}_A^{i,x,\mu_A}$ and $\tilde{\sigma}_B^{j,y,\mu_B}$ with intensity μ_A and μ_B respectively, the single photon components of those states can

be easily obtained by performing the following projection and then normalizing the resulting states

$$\rho_A^{i,x} = \frac{\Big(|0\rangle\langle 0|_{H_A} \otimes |1\rangle\langle 1|_{V_A} + |1\rangle\langle 1|_{H_A} \otimes |0\rangle\langle 0|_{V_A}\Big)\tilde{\rho}_A^{i,x,\mu_A}\Big(|0\rangle\langle 0|_{H_A} \otimes |1\rangle\langle 1|_{V_A} + |1\rangle\langle 1|_{H_A} \otimes |0\rangle\langle 0|_{V_A}\Big)}{e^{-\mu_A}\mu_A}$$
$$\sigma_B^{j,y} = \frac{\Big(|0\rangle\langle 0|_{H_B} \otimes |1\rangle\langle 1|_{V_B} + |1\rangle\langle 1|_{H_B} \otimes |0\rangle\langle 0|_{V_B}\Big)\tilde{\sigma}_B^{j,y,\mu_B}\Big(|0\rangle\langle 0|_{H_B} \otimes |1\rangle\langle 1|_{V_B} + |1\rangle\langle 1|_{H_B} \otimes |0\rangle\langle 0|_{V_B}\Big)}{e^{-\mu_B}\mu_B} \tag{E.1}$$

where $|0\rangle_m$ and $|1\rangle_m$ are the vacuum and single photon states in mode m respectively.

Hence, it is important that we characterize the sources of each legitimate party before performing the protocol. Ideally, this should be done by performing tomography on the signal states $\tilde{\rho}_A^{i,x,\mu_A}$ and $\tilde{\sigma}_B^{j,y,\mu_B}$. Alternatively, one can have a model for the source, taking into account the finite precision and randomness in the modulation of the signal states. Once we have the single photon component of the signal states (i.e. $\rho_A^{i,x}$ and $\sigma_B^{j,y}$ in Eq. (1) of the main text), we can use them to construct the $\hat{\gamma}$ matrix in Eq. (3) of the main text, and to impose the constraints on the ancillary systems $A'B'$ as described in Eq. (5) of the main text .

On the other hand, from the parameter estimation step of the protocol, we can estimate the gain $Q_{\mu_A,\mu_B}^{i,j,x,y}$ for each choice of states (i,x) and (j,y) and intensities μ_A, μ_B. When using infinite number of decoy states, Alice and Bob can determine the values of the single photon yield $Y_{11}^{i,j,x,y}$ exactly for all i,j,x,y. Once the values of $Y_{11}^{i,j,x,y}$ are obtained, we can replace the $p_{det}^{i,j,x,y}$ with $Y_{11}^{i,j,x,y}$ in Eq. (3) of the main text and then proceed with our method.

E.2 Relationship Between the Invertibility of $\hat{\gamma}$ and the States in the Bloch Sphere Forming a Tetrahedron

In the main text, we demonstrated that we could solve for the elements of Eve's Gramian matrix using the equation:

$$\vec{p}_{det} = \hat{\gamma}\vec{e} \implies \vec{e} = \hat{\gamma}^{-1}\vec{p}_{det} \tag{E.2}$$

where $\vec{e}_s = \langle e_{m',n'}^P | e_{m,n}^P \rangle_E$ are the elements of the vectorized form of the Gramian matrix of Eve's states associated with a passing announcement. $(\vec{p}_{det})_t = p_{det}^{i,j,x,y}$ form a vector containing all the successful detection probabilities, and $\hat{\gamma}_{ts} = p^{i,x}q^{j,y}c_{m,m'}^{i,x}d_{n,n'}^{j,y}$ form a matrix dependent on the initial states used in the protocol which were taken to be:

$$\rho_A^{i,x}\sigma_B^{j,y} = \sum_{\substack{m,m',\\ n,n'=H}}^{V} c_{m,m'}^{i,x}d_{n,n'}^{j,y}\,|m,n\rangle\langle m',n'|_{A,B} \tag{E.3}$$

Here we show that the invertibility of $\hat{\gamma}$ is equivalent to the condition in the loss tolerant protocol [25] that Alice and Bob each need to choose four signal states that form a tetrahedron in the Bloch sphere, as shown in Fig. E.1.

We begin by noting that:

$$\begin{aligned}\hat{\gamma}_{ts} &= p^{i,x}q^{j,y}c_{m,m'}^{i,x}d_{n,n'}^{j,y} \\ &= p^{i,x}q^{j,y}\,\langle m,n|\rho_A^{i,x}\sigma_B^{j,y}|m',n'\rangle_{A,B}\end{aligned} \tag{E.4}$$

meaning we can always choose the basis ordering of $\hat{\gamma}$ so that its rows are $\mathrm{vec}(p^{i,x}\rho_A^{i,x})^T \otimes \mathrm{vec}(q^{j,y}\sigma_B^{j,y})^T$, the

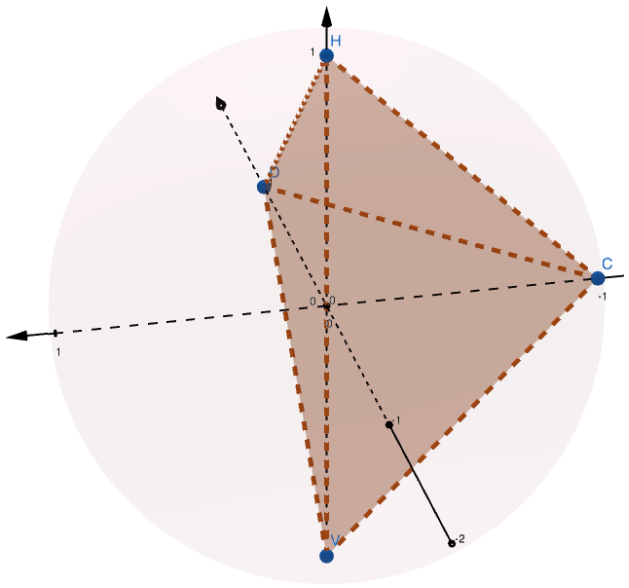


Figure E.1: A tetrahedron in the Bloch sphere representing qubits encoded in horizontal, vertical, diagonal and clockwise circularly polarized single photons. A tetrahedron is formed so long as the states don't all fall in the same plane.

tensor product of the vectorized forms of the probability-weighted signal states. Invertibility of $\hat{\gamma}$ is equivalent to showing its rows are linearly independent. Since all the row vectors have tensor product form, we just need to show that the $\{\text{vec}(p^{i,x}\rho_A^{i,x})\}$ and the $\{\text{vec}(q^{j,y}\sigma_B^{j,y})\}$ each form sets of linearly independent vectors.

Next, we recall that four states forming a tetrahedron in the Bloch sphere is equivalent to them having linearly independent Stokes vectors. Let's focus on Alice's states, since the states are qubits, they can be expressed in terms of Stokes parameters:

$$p^{i,x}\rho_A^{i,x} = \frac{1}{2}\sum_{r=0}^{3} P_r^{i,x}\sigma_r \tag{E.5}$$

where σ_0 is the identity and σ_r, $r = 1, 2, 3$ are the Pauli matrices, while $P_r^{i,x} = p^{i,x}Tr(\sigma_r\rho_A^{i,x})$ form the elements of the Stokes vector $\vec{P}^{i,x}$ for that state. Thus:

$$\text{vec}(p^{i,x}\rho_A^{i,x}) = \frac{1}{2}\sum_{r=0}^{3} P_r^{i,x}\text{vec}(\sigma_r) \tag{E.6}$$

It is easy to show that $\text{vec}(\sigma_r)^T\text{vec}(\sigma_{r'}) = \delta_{rr'}$, which means the inner product of any two $\text{vec}(p^{i,x}\rho_A^{i,x})$ is related to the inner product of the Stokes vectors by a constant factor:

$$\begin{aligned}\text{vec}^T(p^{i,x}\rho_A^{i,x})\text{vec}(p^{i',x'}\rho_A^{x'i'}) =& \frac{1}{2}\sum_{r=0}^{3} P_r^{i,x}P_r^{x',i'} \\ =& \frac{1}{2}\vec{P}^{i,x}\cdot\vec{P}^{x',i'}\end{aligned} \tag{E.7}$$

Thus, since the inner product structure of the rows of $\hat{\gamma}$ is identical to that of the Stokes vectors up to a factor, the linear independence of the Stokes vectors is equivalent to the invertibility of $\hat{\gamma}$.

Generalization to high-dimensional protocols: Having reframed the security proof from the loss-tolerant protocol in this form, we observe that the matrix inversion Eq. E.2 can also generalize straightforwardly to MDI QKD protocols employing discrete-variable high-dimensional degrees of freedom, such as those employing orbital angular momentum [233, 295, 296] or timebin encodings [291, 297].

For Alice and Bob each sending d-dimensional systems, Eve's Gramian matrix will have d^4 elements that we can flatten to a vector $\vec{e}$. Thus, Alice and Bob can each prepare d^2 states within the d-dimensional space, that will in turn yield d^4 observable detection probabilities that we can store in the vector $\vec{p}_{det}$. The basis ordering of $\hat{\gamma}$ can be chosen so that its rows are the tensor product of the vectorized forms of the probability-weighted signal states $\text{vec}(p^{i,x}\rho_A^{i,x})^T \otimes \text{vec}(q^{j,y}\sigma_B^{j,y})^T$. Since invertibility of $\hat{\gamma}$ is equivalent to its rows being linearly independent, this provides a clear minimum requirement for state preparation, namely that the vectorized forms of the states which Alice and Bob use in the protocol must be linearly independent.

If they can prepare their states so that the d^4 rows of $\hat{\gamma}$ are linearly independent, then they will be able to satisfy Eq. E.2. This condition of linear independence for the vectorized density matrices is a much less stringent condition to satisfy for high-dimensional protocols than e.g. employing mutually unbiased bases [334, 335], while still allowing for complete characterization of the parameters in the high-dimensional secret key rate formula that are dependent on Eve's system [336].

E.3 The Virtual Picture and Optimization of the Key Rate with Semidefinite Programming

In the main text, we provided an overview of how to determine the optimal virtual picture for our protocol using a virtual twisting operation [284] and semidefinite programming. Here we provide the full mathematical details of our analytical and numerical techniques.

E.3.1 Moving to a Virtual Picture

Given states of the form in Eq. 6.1, we can define a virtual purified picture for the key generation states:

$$\{p^{0,x}\rho_A^{0,x}q^{0,y}\sigma_B^{0,y}\} \to |\zeta\rangle_{\bar{A}\bar{B}A'B'AB} = \sum_{x,y} |x,y\rangle_{\bar{A}\bar{B}} \sum_{m,n=H}^{V} |\gamma_{m,n}^{x,y}\rangle_{A'B'} |m,n\rangle_{AB} \tag{E.8}$$

where we know from the initial states that $Tr_{\bar{A}\bar{B}A'B'}(|x,y\rangle\langle x,y|_{\bar{A}\bar{B}} |\zeta\rangle\langle\zeta|_{\bar{A}\bar{B}A'B'AB}) = p^{0,x}\rho_A^{0,x}q^{0,y}\sigma_B^{0,y}$, which fixes the constraint:

$$\langle\gamma_{m',n'}^{x,y}|\gamma_{m,n}^{x,y}\rangle_{A'B'} \equiv p^{0,x}q^{0,y}c_{m,m'}^{0,x}d_{n,n'}^{0,y} \tag{E.9}$$

where these also correspond to a subset of the matrix elements from Eq. E.4 for the key generation states.

The constraint in Eq. E.9 on the ancillary systems is not sufficient to identically fix the purification. Rather, any unitary twisting operation of the form:

$$U_{\bar{A}\bar{B}A'B'} = \sum_{x,y=0}^{1} |x,y\rangle\langle x,y|_{\bar{A}\bar{B}} \otimes U_{A'B'}^{x,y} \tag{E.10}$$

preserves the real signal states, since:

$$\begin{aligned} &Tr_{\bar{A}\bar{B}A'B'}(|x,y\rangle\langle x,y|_{\bar{A}\bar{B}} U_{\bar{A}\bar{B}A'B'} |\zeta\rangle\langle\zeta|_{\bar{A}\bar{B}A'B'AB} U^{\dagger}_{\bar{A}\bar{B}A'B'}) \\ =&Tr_{\bar{A}\bar{B}A'B'}(|x,y\rangle\langle x,y|_{\bar{A}\bar{B}} U^{x,y}_{A'B'} |\zeta\rangle\langle\zeta|_{\bar{A}\bar{B}A'B'AB} U^{x,y\,\dagger}_{A'B'}) \\ =&p^{0,x}\rho^{0,x}_{A} q^{0,y}\sigma^{0,y}_{B} \end{aligned} \tag{E.11}$$

Since the produced signal states are independent of this twisting operation, it cannot affect any of the real detection probabilities observed in the execution of the protocol which depend only on the A, B systems. Thus, the characterization of Eve's Gramian matrix elements $\langle e^{P}_{m',n'}|e^{P}_{m,n}\rangle_{E}$ is independent of the twisting operation. Moreover, since the produced signal states are independent of the twisting operation it never needs to actually be implemented in the real protocol, and can remain a useful virtual analytical tool in the characterization of security after the real signal exchange. Finally, the twisting operation can be chosen after the detection statistics are produced, which gives Alice and Bob the power to adjust their virtual strategy based on what they observe.

The major change when adding the twisting operation is the *definition* of the phase error rates, i.e. how we use the information about Eve's state from $\vec{e}_s$. Consider the joint state between Alice, Bob, their ancillae, and Eve in the purified picture after they apply a twisting operation and send the systems A, B to Eve:

$$U_{\bar{A}\bar{B}A'B'}|\zeta\rangle_{\bar{A}\bar{B}A'B'AB} \to |\Gamma(U)\rangle_{\bar{A}\bar{B}A'B'EZ} = \sum_{x,y}|x,y\rangle_{\bar{A}\bar{B}} \sum_{m,n=H}^{V} U^{x,y}_{A'B'}|\gamma^{x,y}_{m,n}\rangle_{A'B'} \sum_{z=P}^{F} |e^{z}_{m,n}\rangle_{E}|z\rangle_{Z} \tag{E.12}$$

Thus, taking the target virtual Bell state to be $|\Phi^{+}\rangle\langle\Phi^{+}|_{\bar{A}\bar{B}}$, the phase error rates now become the *twisted phase error rates*:

$$\begin{aligned} e_X(U) &= \frac{1}{p^{0,0}_{det}}\langle\Gamma(U)|\,[\,(|\Psi^{-}\rangle\langle\Psi^{-}| + |\Phi^{-}\rangle\langle\Phi^{-}|)_{\bar{A}\bar{B}} \otimes |P\rangle\langle P|_{Z}\,]\,|\Gamma(U)\rangle_{\bar{A}\bar{B}A'B'EZ} \\ &= \frac{1}{2} - \frac{1}{p^{0,0}_{det}} \sum_{m,n,m',n'} Re\Big[\big(\,\langle\gamma^{0,0}_{m',n'}|U^{0,0}_{A'B'}{}^{\dagger}U^{1,1}_{A'B'}|\gamma^{1,1}_{m,n}\rangle_{A'B'} + \langle\gamma^{0,1}_{m',n'}|U^{0,1}_{A'B'}{}^{\dagger}U^{1,0}_{A'B'}|\gamma^{1,0}_{m,n}\rangle_{A'B'}\,\big)\,\langle e^{P}_{m',n'}|e^{P}_{m,n}\rangle_{E}\Big] \end{aligned}$$

$$\begin{aligned} e_Y(U) &= \frac{1}{p^{0,0}_{det}}\langle\Gamma(U)|\,[\,(|\Phi^{-}\rangle\langle\Phi^{-}| + |\Psi^{+}\rangle\langle\Psi^{+}|)_{\bar{A}\bar{B}} \otimes |P\rangle\langle P|_{Z}\,]\,|\Gamma(U)\rangle_{\bar{A}\bar{B}A'B'EZ} \\ &= \frac{1}{2} - \frac{1}{p^{0,0}_{det}} \sum_{m,n,m',n'} Re\Big[\big(\,\langle\gamma^{0,0}_{m',n'}|U^{0,0}_{A'B'}{}^{\dagger}U^{1,1}_{A'B'}|\gamma^{1,1}_{m,n}\rangle_{A'B'} - \langle\gamma^{0,1}_{m',n'}|U^{0,1}_{A'B'}{}^{\dagger}U^{1,0}_{A'B'}|\gamma^{1,0}_{m,n}\rangle_{A'B'}\,\big)\,\langle e^{P}_{m',n'}|e^{P}_{m,n}\rangle_{E}\Big] \end{aligned} \tag{E.13}$$

In general, there exists an optimal purification provided by some $U_{\bar{A}\bar{B}A'B'}$ such that the key rate is maximized. As it is unlikely to choose this optimal purification at random when constructing the problem, we expect the calculation of $e_X(U)$ and $e_Y(U)$ to benefit from an optimization over $U_{\bar{A}\bar{B}A'B'}$, and thus, in general, for the key rate to increase by employing an optimal twisted phase error rate.

Remark: employing noisy, i.e. mixed, signal states in the protocol begs the question of how best to purify the states in the virtual protocol when defining the phase error rate. This necessarily leads to the definition of a twisted phase error rate and a search for an optimal twisting operation with respect to the key rate. The twisting operation is entirely in the virtual picture, so the optimal $U_{\bar{A}\bar{B}A'B'}$ can and should be computed after the exchange of signals and observation of detection probabilities and bit error rates.

Although the form of the optimal twisting operation can even be nonlocal across $\bar{A}\bar{B}$ in practice, it does not need to be implemented in the real protocol, so its locality does not matter.

It is worth emphasizing that employing a twisting operation is an optional step in the security proof. Indeed, any phase error rates of the form Eq. (E.13) will supply a suitable lower bound on the key rate, since Eve does not have control over the $A'B'$ systems. Nonetheless, optimizing over the twisting operation will in general boost the key rate, safeguarding against a poor, overly pessimistic initial choice of purification.

We next demonstrate how optimizing the phase error rates over twisting operations can be framed as two semidefinite programs, closing a gap in the previous literature on twisting operations [287].

E.3.2 Semidefinite Programs for Evaluating the Six State Key Rate

A general optimization problem is of the form [288]:

$$\begin{aligned} &\texttt{minimize} && f_0(\mathbf{x}) \\ &\texttt{s.t.} && f_i(\mathbf{x}) \geq b_i,\ i = 1, \ldots, m \end{aligned}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ are the variables over which we optimize; $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the objective function; $f_i : \mathbb{R}^n \to \mathbb{R}$ are the constraint functions; and, b_i are the constraint bounds. An optimal solution, $\mathbf{x}^\star$ would mean that for all $\mathbf{z}$ such that $f_i(\mathbf{z}) \geq b_i$ then $f_0(\mathbf{z}) \geq f_0(\mathbf{x}^\star)$.

Semidefinite programs (SDPs) are a class of convex optimization problems with linear objective and constraint functions over a cone of positive semidefinite (PSD) matrices [288]. That is, the optimization variables, $\mathbf{x}$, form the elements of a matrix with non-negative eigenvalues, and $f_i(x) = \mathbf{c}_i \cdot \mathbf{x}$. They have become an incredibly versatile tool for QKD security proofs in recent years [27, 279, 280, 289–291].

Objective Functions

At first glance, the optimization problem required for the six-state key rate in Eq. 7 of the main text looks daunting. It appears we have two quantities to optimize with the twisting operation, $e_X(U)$ and $e_Y(U)$, appearing in a nonlinear function due to the binary entropy. However, consider a simple change of variable so that the two unknowns are given by:

$$e_-(U) = (e_X - e_Y)(U),\ e_+(U) = (e_X + e_Y)(U) \tag{E.14}$$

These remain linear objective functions of the only free variables in the problem, which we recall are $\left\{ \langle \gamma^{x',y'}_{m',n'} | U^{x',y'\,\dagger}_{A'B'} U^{x,y}_{A'B'} | \gamma^{x,y}_{m,n} \rangle_{A'B'} \right\}$:

$$\begin{aligned} e_+(U) &= 1 - \frac{2}{p^{0,0}_{det}} \sum_{m,n,m',n'} Re\left(\langle \gamma^{0,0}_{m',n'} | U^{0,0\,\dagger}_{A'B'} U^{1,1}_{A'B'} | \gamma^{1,1}_{m,n} \rangle_{A'B'} \langle e^P_{m',n'} | e^P_{m,n} \rangle_E \right) \\ &= e_+(U^{0,0\,\dagger}_{A'B'} U^{1,1}_{A'B'}) \end{aligned} \tag{E.15}$$

and:

$$\begin{aligned} e_-(U) &= -\frac{2}{p^{0,0}_{det}} \sum_{m,n,m',n'} Re\left(\langle \gamma^{0,1}_{m',n'} | U^{0,1\,\dagger}_{A'B'} U^{1,0}_{A'B'} | \gamma^{1,0}_{m,n} \rangle_{A'B'} \langle e^P_{m',n'} | e^P_{m,n} \rangle_E \right) \\ &= e_-(U^{0,1\,\dagger}_{A'B'} U^{1,0}_{A'B'}) \end{aligned} \tag{E.16}$$

The only free parameters over which we can optimize are the twisting unitaries, $\{U^{0,1}_{A'B'}, U^{1,0}_{A'B'}, U^{0,0}_{A'B'},$

$U^{1,1}_{A'B'}\}$, where each of the four unitaries can be defined independently of the others. Here, we have found that the two objective functions in the key rate $e_\pm(U)$, are functions of independent variables: $e_+(U)$ only depends on $U_+ = U^{0,0\,\dagger}_{A'B'}U^{1,1}_{A'B'}$ and $e_-(U)$ only depends on $U_- = U^{0,1\,\dagger}_{A'B'}U^{1,0}_{A'B'}$. This is very good, since it means the difficult task of nonlinear optimization of the six state key rate formula can be avoided. Using the monotonicity of the binary entropy, we can directly optimize $e_\pm(U_\pm)$ within the binary entropy functions:

$$\begin{aligned} R &= \max_U p^{0,0}_{det}\left(1 - h_2(e_Z) - e_Z h_2\left[\frac{1+e_-(U)/e_Z}{2}\right] - (1-e_Z)h_2\left[\frac{1-[e_+(U)+e_Z]/2}{1-e_Z}\right]\right) \\ &= p^{0,0}_{det}\left(1 - h_2(e_Z) - e_Z h_2\left[\frac{1+\max_{U_-} e_-(U_-)/e_Z}{2}\right] - (1-e_Z)h_2\left[\frac{1-[\min_{U_+} e_+(U_+)+e_Z]/2}{1-e_Z}\right]\right) \end{aligned} \tag{E.17}$$

with the extra conditions $0 \leq e_-(U_-) \leq e_Z$ and $e_Z \leq e_+(U_+) \leq 1$ so that the arguments of the binary entropy functions remain between 0 and 1.

Two Independent Semidefinite Programs

We recall that $\langle e^P_{m',n'}|e^P_{m,n}\rangle_E$ are known from Eq. 6.3, while $\langle \gamma^{x',y'}_{m',n'}|{U^{x',y'}_{A'B'}}^\dagger U^{x,y}_{A'B'}|\gamma^{x,y}_{m,n}\rangle_{A'B'}$ are the optimization variables. This leads to two independent semidefinite programs.

- For the linear objective function $e_-(U_-)$, we note the optimization variables

$$\langle \gamma^{x,(x+1 \bmod 2)}_{m',n'}|{U^{x,(x+1 \bmod 2)}_{A'B'}}^\dagger U^{y,(y+1 \bmod 2)}_{A'B'}|\gamma^{y,(y+1 \bmod 2)}_{m,n}\rangle_{A'B'} \tag{E.18}$$

  form the 8×8 positive semidefinite Gram matrix for the vectors $\{U^{0,1}_{A'B'}|\gamma^{0,1}_{m,n}\rangle_{A'B'}, U^{1,0}_{A'B'}|\gamma^{1,0}_{m,n}\rangle_{A'B'}\}$, subject to the eight linear constraints from Eq. E.9:

$$\langle \gamma^{x,(x+1 \bmod 2)}_{m',n'}|{U^{x,(x+1 \bmod 2)}_{A'B'}}^\dagger U^{x,(x+1 \bmod 2)}_{A'B'}|\gamma^{x(x+1 \bmod 2)}_{m,n}\rangle_{A'B'} = \langle \gamma^{x,(x+1 \bmod 2)}_{m',n'}|\gamma^{x,(x+1 \bmod 2)}_{m,n}\rangle_{A'B'} \tag{E.19}$$

  The optimization is additionally constrained by $0 \leq e_-(U_-) \leq e_Z$.

- For the linear objective function $e_+(U_+)$, we note the optimization variables

$$\langle \gamma^{x,x}_{m',n'}|{U^{x,x}_{A'B'}}^\dagger U^{y,y}_{A'B'}|\gamma^{y,y}_{m,n}\rangle_{A'B'} \tag{E.20}$$

  form the 8×8 PSD Gram matrix for the vectors $\{U^{0,0}_{A'B'}|\gamma^{0,0}_{m,n}\rangle_{A'B'}, U^{1,1}_{A'B'}|\gamma^{1,1}_{m,n}\rangle_{A'B'}\}$, subject to the eight linear constraints from Eq. E.9:

$$\langle \gamma^{x,x}_{m',n'}|{U^{x,x}_{A'B'}}^\dagger U^{x,x}_{A'B'}|\gamma^{x,x}_{m,n}\rangle_{A'B'} = \langle \gamma^{x,x}_{m',n'}|\gamma^{x,x}_{m,n}\rangle_{A'B'} \tag{E.21}$$

  The optimization is additionally constrained by $e_Z \leq e_+(U_+) \leq 1$.

With that, we have two independent semidefinite programs which can be used to optimize the six state key rate formula. In section E.4, we provide a pseudocode overview of our numerical technique for calculating the key rate.

E.4 Pseudocode for Key Rate Calculation

Here we present a sketch of our numerical implementation for calculating key rates. For the semidefinite programs we employed CVXPY [299, 300], a convex optimization library for Python. All codes are available upon request.

Algorithm 9 Key rate function

function `keyrate`$(\rho_A, \sigma_B, p_A, q_B, p_{dark}, \eta, l)$

ρ_A and σ_B are arrays containing Alice and Bob's four density matrices, s.t. $\rho_A[i,x] = \rho_A^{i,x}$, $\sigma_B[j,y] = \sigma_B^{j,y}$

p_A and q_B are lists of the probabilities for sending their four states, s.t. $p_A[2i+x] = p_A^{i,x}$, $q_B[2j+y] = q_B^{j,y}$

p_{dark} is the dark count probability per detector

η is the overall transmissivity

l is the Alice-Charlie distance (same for Bob)

#

Probability of losing a photon

$p_0 = 1 - \eta 10^{-0.2l/20}$

Extract protocol statistics

$\vec{p}_{det}$,$\hat{\gamma}$= stats(ρ_A,σ_B,p_A,q_B,p_0,p_{dark})

Key generation detection probability

$p_{det}^{0,0} = \sum_{i=0}^{3} \vec{p}_{det}[i]$

Bit error rate

$e_Z = \vec{p}_{det}[1] + \vec{p}_{det}[2]$

Solving for Eve's Gramian matrix (Eq. 6.3 of main text)

$\vec{e} = \hat{\gamma}^{-1}\vec{p}_{det}$

Phase error rates

$e_- = \text{emin}(\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p_{det}^{0,0})$

$e_+ = \text{eplus}(\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p_{det}^{0,0})$

Key rate

$R = p_{det}^{0,0}\left[1 - h_2(e_Z) - e_Z h_2\left(\frac{1+e_-/e_Z}{2}\right) - (1-ez)h_2\left(\frac{1-(e_++e_Z)/2}{1-e_Z}\right)\right]$

return R

end function

Algorithm 10 Protocol statistics function

function `stats`(ρ_A,σ_B,p_A,q_B,p_0,p_{dark})

Loop over all 16 combinations of states in lists: i,j=0,1; x,y=0,1

Probability of passing if both photons arrive

$p_{pass}[8i+4j+2x+y] = p_A[2i+x]q_B[2j+y]Tr(\rho_A[i,x]\sigma_B[j,y]\left|\Phi^+\right\rangle\left\langle\Phi^+\right|_{AB})$

Detection probability including dark counts and loss

$p_{det}[8i+4j+2x+y] = (1-p_0)^2(1-p_{dark})^2 p_{pass}[8i+4j+2x+y]$

$p_{det}[8i+4j+2x+y] \mathrel{+}= 2p_A[2i+x]q_B[2j+y]\left[p_0^2 p_{dark}^2(1-p_{dark})^2 + p_0(1-p_0)p_{dark}(1-p_{dark})^2\right]$

Filling the 16 rows of the $\hat{\gamma}$ matrix

$\hat{\gamma}[8i+4j+2x+y] = p_A[2i+x]q_B[2j+y]\mathrm{vec}(\rho_A[i,x]\sigma_B[j,y])$

return $\vec{p}_{det}, \hat{\gamma}$

end function

Algorithm 11 Phase errors

function `emin`$(\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p^{0,0}_{det})$

Reshape $\vec{e}$ into a matrix

$\hat{e} = \text{reshape}(\vec{e})$

We use the CVXPY and Mosek packages for solving semidefinite programs

import cvxpy

Define the 8x8 Gramian matrix from Eq. E.18 for the $A'B'$ systems as the optimization variables of the system

G = cvxpy.Variable((8,8))

Define a list of constraints on G, such as PSD and constraint from Eq. E.19

constraints = $[G \succeq 0]$

#For $x = 0, 1; m, m', n, n' = 0, 1$

constraints += $[G[4x + 2m + n, 4x + 2m' + n'] = p_A[x]q_B[(x + 1) \bmod 2]\rho_A[0, x][m, m']\sigma_B[0, (x + 1) \bmod 2][n, n']]$

Define the objective function

$e_- = -\frac{2}{p^{0,0}_{det}} \sum_{m,m',n,n'} Re(\hat{e}[2m + n, 2m' + n']G[2m + n, 4 + 2m' + n'])$

constraints += $[e_- \geq 0, e_- \leq e_Z]$

Use cvxpy to solve problem

prob = cvxpy.Problem(cvxpy.Maximize(e_-),constraints)

prob.solve(solver = cvxpy.MOSEK)

e_- = prob.value

return e_-

end function

function `eplus`$(\rho_A, \sigma_B, p_A, q_B, \vec{e}, e_Z, p^{0,0}_{det})$

Reshape $\vec{e}$ into a matrix

$\hat{e} = \text{reshape}(\vec{e})$

We use the CVXPY and Mosek packages for solving semidefinite programs

import cvxpy

Define the 8x8 Gramian matrix from Eq. E.20 for the $A'B'$ systems as the Variable of the system

G = cvxpy.Variable((8,8))

Define a list of constraints on G, such as PSD and constraint from Eq. E.21

constraints = $[G \succeq 0]$

#For $x = 0, 1; m, m', n, n' = 0, 1$

constraints += $[G[4x + 2m + n, 4x + 2m' + n'] = p_A[x]q_B[x]\rho_A[0, x][m, m']\sigma_B[0, x][n, n']]$

Define the objective function

$e_+ = 1 - \frac{2}{p^{0,0}_{det}} \sum_{m,m',n,n'} Re(\hat{e}[2m + n, 2m' + n']G[2m + n, 4 + 2m' + n'])$

constraints += $[e_+ \geq e_Z, e_+ \leq 1]$

Use cvxpy to solve problem

prob = cvxpy.Problem(cvxpy.Minimize(e_+),constraints)

prob.solve(solver = cvxpy.MOSEK)

e_+ = prob.value

return e_+

end function

E.5 (δ, p)-Model for Signal States

We consider the following two-parameter (δ, p)-model for the initial states which Alice and Bob prepare:

$$\begin{aligned}
\rho_A^{0,0} &= \sigma_B^{0,0} = (1-p)\left|\xi_{00}^{\delta}\right\rangle\left\langle\xi_{00}^{\delta}\right| + p/2\mathbb{1} \\
\rho_A^{0,1} &= \sigma_B^{0,1} = (1-p)\left|\xi_{01}^{\delta}\right\rangle\left\langle\xi_{01}^{\delta}\right| + p/2\mathbb{1} \\
\rho_A^{1,0} &= \sigma_B^{1,0} = (1-p)\left|\xi_{10}^{\delta}\right\rangle\left\langle\xi_{10}^{\delta}\right| + p/2\mathbb{1} \\
\rho_A^{1,1} &= \sigma_B^{1,1} = (1-p)\left|\xi_{11}^{\delta}\right\rangle\left\langle\xi_{11}^{\delta}\right| + p/2\mathbb{1}
\end{aligned} \tag{E.22}$$

where the states $|\xi^{\delta}\rangle$ are of the form:

$$\begin{aligned}
\left|\xi_{00}^{\delta}\right\rangle &= |H\rangle \\
\left|\xi_{01}^{\delta}\right\rangle &= -\sin\frac{\delta}{2}|H\rangle + \cos\frac{\delta}{2}|V\rangle \\
\left|\xi_{10}^{\delta}\right\rangle &= \cos\frac{\pi+\delta}{4}|H\rangle + \sin\frac{\pi+\delta}{4}|V\rangle \\
\left|\xi_{11}^{\delta}\right\rangle &= \cos\frac{-\pi+\delta}{4}|H\rangle + i\sin\frac{-\pi+\delta}{4}|V\rangle
\end{aligned} \tag{E.23}$$

The states $|\xi^{\delta}\rangle$ parametrized by δ are a model for Alice and Bob attempting to prepare $\{|H\rangle, |V\rangle, {}^{|H\rangle+|V\rangle}/\sqrt{2}, {}^{|H\rangle-i|V\rangle}/\sqrt{2}\}$, but each state is subject to a different, constant state-dependent modulation error. The pure $|\xi^{\delta}\rangle$ states and the resulting key rates were considered in the loss-tolerant protocol [25]. Additionally, were the modulation error a random variable subject to a distribution on the Bloch sphere, we expect the average state to be mixed with a shorter than unit Bloch vector. This effect is accounted for with the depolarizing channel parametrized by p, which indicates with some probability the maximally mixed state is sent, shortening the Bloch vector. The depolarizing channel can also be used to model any thermal photons that are accidentally produced during state preparation.

Appendix F

Supplementary Material for Measurement Device-Independent QKD with Time-Dependent Source Side-Channels

F.1 Linear Programming for Decoy States

Here we review how to construct a linear program to calculate upper and lower bounds on the single photon detection probabilities of a decoy state MDI QKD protocol [312]. Starting from Eq. (7.2), we find that, for a fixed basis and bit choice (i, j, x, y) and N choices of intensity settings each, Alice and Bob have N^2 linear equality constraints on the single photon detection probability $p^{i,j,x,y}_{\mathtt{pass},1,1}$:

$$Q^{i,j,x,y}_{k,l} = \sum_{m,n} \frac{e^{-(\mu_k+\nu_l)}\mu_k^m \nu_l^n}{m!n!} p^{i,j,x,y}_{\mathtt{pass},m,n}. \tag{F.1}$$

Here, the variables of the optimization are $p^{i,j,x,y}_{\mathtt{pass},m,n}$. To establish a lower (upper) bound on $p^{i,j,x,y}_{\mathtt{pass},1,1}$ given these constraints, we solve the linear program to find the minimum (maximum) possible value of $p^{i,j,x,y}_{\mathtt{pass},1,1}$ consistent with the constraints. If Alice and Bob each have n_A and n_B basis choice settings, each basis choice associated with two bit choices, we repeat the process of finding lower and upper bounds for all $n_A \times n_B \times 2 \times 2$ combinations of (i, j, x, y).

Since there are in principle infinitely many $p^{i,j,x,y}_{\mathtt{pass},m,n}$, for a practical linear program, we impose a cutoff photon number $N_{\max}$. In that case, the N^2 linear equality constraints become $2N^2$ linear inequality constraints. The first N^2 constraints are:

$$Q^{i,j,x,y}_{k,l} \geq \sum_{m,n=0}^{N_{\max}} \frac{e^{-(\mu_k+\nu_l)}\mu_k^m \nu_l^n}{m!n!} p^{i,j,x,y}_{\mathtt{pass},m,n}, \tag{F.2}$$

stemming from the fact that summing up to the cutoff will yield a value less than the total detection

probability. For the next N^2 constraints, we find:

$$1-\sum_{m,n=0}^{N_{\max}}\frac{e^{-(\mu_k+\nu_l)}\mu_k^m\nu_l^n}{m!n!}\geq\sum_{m,n=N_{\max}+1}^{\infty}\frac{e^{-(\mu_k+\nu_l)}\mu_k^m\nu_l^n}{m!n!}p^{i,j,x,y}_{\mathtt{pass},m,n}. \tag{F.3}$$

which means we can provide the constraints:

$$Q^{i,j,x,y}_{k,l}+\sum_{m,n=0}^{N_{\max}}\frac{e^{-(\mu_k+\nu_l)}\mu_k^m\nu_l^n}{m!n!}-1\leq\sum_{m,n=0}^{N_{\max}}\frac{e^{-(\mu_k+\nu_l)}\mu_k^m\nu_l^n}{m!n!}p^{i,j,x,y}_{\mathtt{pass},m,n}. \tag{F.4}$$

In practice, we found an N_{max} of 10 photons was sufficient to provide good upper and lower bounds on $p^{i,j,x,y}_{\mathtt{pass},1,1}$ while not taking too long to compute.

F.2 Comparison to the Proof Technique From Pereira et. al.

Here we will compare the proof technique we are using to the technique from [274]. We will show that [274] relaxes the SDP inherent to optimizing the phase error rate to a linear program, which we would expect to give equal or lower bounds on the key rate than computing the full SDP. For simplicity, we will consider a protocol where when Alice and Bob choose the Z basis, they perfectly prepare qubit states $|0\rangle, |1\rangle$, but we allow for their test states to have leakage components outside of the qubit subspace spanned by $\{|0\rangle, |1\rangle\}$. The following comparison can be generalized in a straightforward manner to arbitrary initial states.

Let U be the unitary that takes $|\psi^i_x\phi^j_y\rangle_{A,B}\to\sum_z|e^{i,j}_{x,y,z}\rangle$. Thus, for this case of initial states, the phase error rate would be given by:

$$e_{ph}=\frac{1}{\sum_{x,y}p^{0,0,x,y}_{\mathtt{pass}}}Tr\left[|P\rangle\langle P|_Z U\left(\frac{\mathbb{1}+\sigma_X\otimes\sigma_X}{2}\right)_{A,B}U^\dagger\right], \tag{F.5}$$

where σ_m, $m=I,X,Y,Z$ refer to Pauli operators in the qubit space spanned by $\{|0\rangle, |1\rangle\}$. Following [25, 274], e_{ph} can be decomposed in terms of the transmission rates of the Pauli operators $q_{\mathtt{pass}|i,j}=Tr\left(|P\rangle\langle P|_Z U\sigma_i\otimes\sigma_j U^\dagger\right)$, since the Pauli operators form a basis for any operator. Were $\epsilon=1$, then we could use the states Alice and Bob send to exactly solve for $q_{\mathtt{pass}|i,j}$ (assuming their test states are some superposition of $\{|0\rangle, |1\rangle\}$). However, when their signal states have a leakage component, we cannot exactly constrain these quantities; in [274] $q_{\mathtt{pass}|i,j}$ form the variables of a linear program that are optimized to determine a lower bound on e_{ph}.

First, we write Alice and Bob's signal states as linear combinations of states in a two-qubit space, and a space orthogonal to it (the leakage space), just as in Eq. (1) of [274]:

$$|\psi^i_x\phi^j_y\rangle_{A,B}=a^{i,j}_{x,y}|\tilde{\psi}^i_x\tilde{\phi}^j_y\rangle_{A,B}+b^{i,j}_{x,y}|\tilde{\psi}^i_x\tilde{\phi}^{j\perp}_y\rangle_{A,B} \tag{F.6}$$

where the orthogonality of the two-qubit and leakage space means $\langle\cdot|\cdot^\perp\rangle_{A,B}=0$.

Next, the detection probabilities provide the constraint:

$$p^{i,j,x,y}_{\mathtt{pass}}=Tr\left(|P\rangle\langle P|_Z U|\psi^i_x\phi^j_y\rangle\langle\psi^i_x\phi^j_y|_{A,B}U^\dagger\right). \tag{F.7}$$

Using the decomposition from Eq. (F.6), we find this is also equal to:

$$p_{\text{pass}}^{i,j,x,y} = |a_{x,y}^{i,j}|^2 Tr\left[|P\rangle\langle P|_Z U |\tilde{\psi}_x^i\tilde{\phi}_y^j\rangle\langle\tilde{\psi}_x^i\tilde{\phi}_y^j|_{A,B} U^\dagger\right] + Tr\left[|P\rangle\langle P|_Z U \left(a_{x,y}^{i,j} b_{x,y}^{i,j\,*} |\tilde{\psi}_x^i\tilde{\phi}_y^j\rangle\langle\tilde{\psi}_x^i\tilde{\phi}_y^{j\perp}| + a_{x,y}^{i,j\,*} b_{x,y}^{i,j} |\tilde{\psi}_x^i\tilde{\phi}_y^{j\perp}\rangle\langle\tilde{\psi}_x^i\tilde{\phi}_y^j| + |b_{x,y}^{i,j}|^2 |\tilde{\psi}_x^i\tilde{\phi}_y^{j\perp}\rangle\langle\tilde{\psi}_x^i\tilde{\phi}_y^{j\perp}|\right)_{A,B} U^\dagger\right] \tag{F.8}$$

which coincides with the MDI QKD version of Eq. (19) from [274]. Since the operator $|\tilde{\psi}_x^i\tilde{\phi}_y^j\rangle\langle\tilde{\psi}_x^i\tilde{\phi}_y^j|$ lives in the two-qubit subspace, it too can be written in terms of the Pauli operators $\sigma_m \otimes \sigma_n$, meaning the first trace term in Eq. (F.8) can be written in terms of $q_{\text{pass}|i,j}$. Because of the second trace term, we cannot solve for them exactly as is done in the loss-tolerant proof technique [25].

Rather than solving the semidefinite program for e_{ph} using the linear equality constraints provided by the detection probabilities, [274] considers a relaxation to a linear program. Specifically, the second trace term in Eq. (F.8) can be bounded above and below by the maximum and minimum eigenvalues of the matrix:

$$M_{x,y}^{i,j} = \begin{pmatrix} 0 & a_{x,y}^{i,j} b_{x,y}^{i,j\,*} \\ a_{x,y}^{i,j\,*} b_{x,y}^{i,j} & |b_{x,y}^{i,j}|^2 \end{pmatrix}. \tag{F.9}$$

Thus, F.8 leads to inequalities linear in $q_{\text{pass}|i,j}$, which can be used as constraints in a linear program to find an upper bound for e_{ph}. However, because the exact equality constraints coming from the detection probabilities have been relaxed using the maximum and minimum eigenvalues of $M_{x,y}^{i,j}$, we would expect that this would lead to a greater upper bound on e_{ph} (and hence a weaker lower bound on the key rate) than if the exact constraints were kept, as they would be in the numerical approach from [27] that we reviewed in Section 7.2.2.

Note that the proof approach we have used in this paper does away with needing to frame the calculation of e_{ph} in terms of $q_{\text{pass}|i,j}$ (even though we could, in principle, do so since they are linear functions of the elements of Eve's Gram matrix); after all, since the signal states are no longer qubits, we need not make the distinction between a qubit subspace and the leakage space, since Eve's operation can blend these two spaces. Instead, given that the phase error can be expressed in terms of the elements of a positive semidefinite matrix associated with Eve's information, and given that we have linear equality constraints on this matrix, the phase error can be maximized directly with a simple SDP, rather than relaxing to a linear program.

As an example to demonstrate the superiority of the SDP method over the method from [274], we consider a toy example of the three state protocol with a single photon source, for which Alice and Bob prepare a leaky third state,

$$|+\rangle_{enc}(\sqrt{\epsilon}\,|vac\rangle_{leak} + \sqrt{1-\epsilon}_{enc}\,|1\rangle_{leak}), \tag{F.10}$$

as opposed to the ideal $|+\rangle_{enc}$. We assume a detection efficiency of 1 and a dark count rate of 10^{-6}. In Fig. F.1, we plot the key rates calculated using the SDP method we reviewed in Section 7.2.2 and using the method from [274]. We find that across values of ϵ, the SDP method performs much better.

F.3 Derivation of Figure 7.1b

Here, we derive the fractional phase change applied to the leakage light as a function of time, as shown in Figure 7.1b.

Refer to the experimental setup shown in Figure 7.1a, specifically the polarization modulation unit. First,

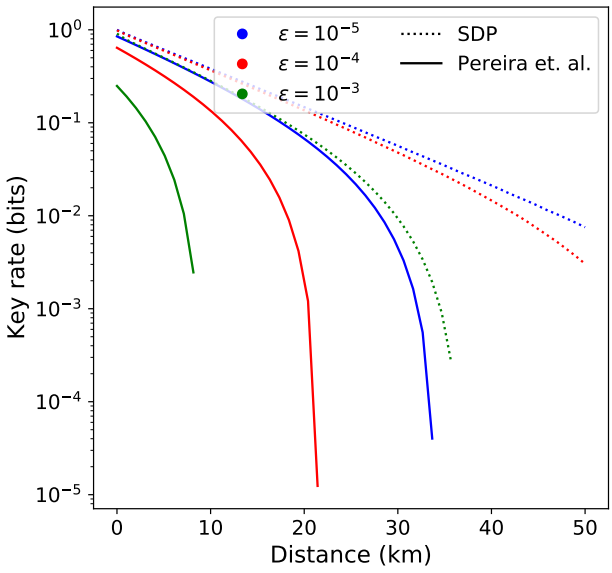


Figure F.1: Key rate vs. distance for various values of ϵ, calculated using the SDP method reviewed in 7.2.2 and the method from [274]. We see the advantage of the SDP approach over the relaxation to a linear program, as done in [274].

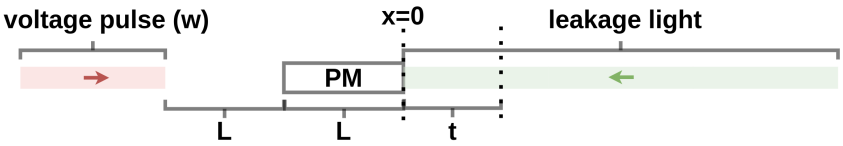


Figure F.2: An illustration depicting a moment in time as a voltage pulse approaches the phase modulator (PM).

optical pulses travel forward through the PM for the purpose of polarization modulation. Simultaneously, voltage pulses overlapping in time with the optical pulses are sent into the PM, propagating in the same direction as the optical pulses. The PM is designed such that the optical and voltage pulses travel through the PM at the same speed. The voltage is what enables a phase change and therefore a polarization change. Since the leakage light is not meant to encode information, voltage is not sent through the PM as this light travels through the first time.

However, when traveling back through the PM after reflection from the Faraday mirror, the leakage light will inevitably collide temporally with a voltage pulse that is travelling in the opposite direction along with an optical pulse it is intended to modulate.

The overall phase modulation experienced by a temporal slice of light after traveling through the PM can be expressed as:

$$\phi = K \int_0^L V(z) dz. \tag{F.11}$$

Here, L represents the length of the PM and $V(z)$ represents the applied voltage overlapping with the slice of light. K is simply a proportionality constant. When light is travelling in the same direction as the voltage wave through the phase modulator, Eq. (F.11) reduces to $K \times V \times L$. This occurs due to the voltage, which is moving at the same speed as the light, being a constant along the length of the PM.

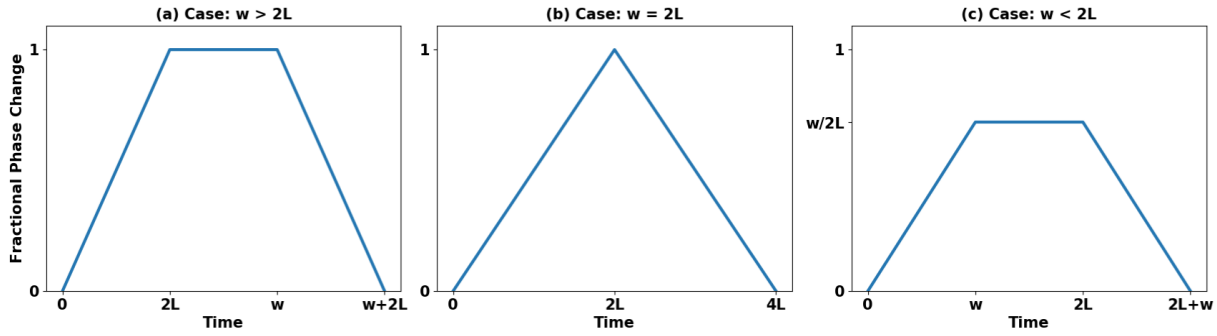


Figure F.3: The various solutions for the integral in Eq. F.13.

In our case, we are also dealing with leakage light that is travelling in the opposite direction. We will use L to refer to the PM length and w to refer to the width of the square voltage pulses sent to the PM. Given these parameters, we can determine the phase change experienced by the leakage light as follows:

1. Refer to Figure F.2. We will use this moment in time as our starting point. First we will create a coordinate system by defining $x = 0$ to be the right hand edge of the phase modulator. We can parametrize a slice of leakage light with t, the time it crosses the point $x = 0$.

2. At the point in time shown in Figure F.2, we can define the voltage pulse as $A(x) = H(x + 2L + w) - H(x + 2L)$ (a square pulse) and the leakage light as $B(x) = H(x)$ where H refers to the Heaviside step function.

3. Now, notice that the movement in time of the voltage pulse and leakage light can also be incorporated into these functions. After τ ps, the function defining the voltage pulse will become $A(x - \tau)$ while the function defining the leakage light will become $B(x + \tau)$.

4. Notice that $A(x-\tau) \times B(x+\tau)$ represents the overlap between the voltage and leakage light at position x and time τ. It has a value of 1 if there is an overlap and a value of 0 if there is no overlap.

5. Now, suppose we want to calculate the amount of time for which the slice t experiences an overlap within the phase modulator. We need to integrate $A(x - \tau) \times B(x + \tau)$ from $\tau = t$ to $\tau = t + L$. In other words, we need to integrate the overlap function over the values of τ for which the slice at position t is inside the phase modulator.

6. The slice at position t has an x position of $t - \tau$ at time τ. Substitute this into the integral for x.

7. The resulting integral is as follows:

$$\int_{t}^{t+L} A(t - 2\tau) \times B(t)\, d\tau. \tag{F.12}$$

This integral represents the amount of time the slice t is in contact with a voltage pulse within the phase modulator. Recall that the optical pulse which is travelling along with (in the same direction as) this voltage pulse would be in contact along the entire length of the phase modulator (L). Therefore,

the phase change experienced by slice t is

$$\frac{1}{L}\int_{t}^{t+L} A(t-2\tau)\times B(t)\,d\tau \tag{F.13}$$

when written as a fraction of the phase change experienced by the pulse. The solution to this integral is shown in Figure F.3. In our particular experimental setup, $L = 150$ ps and $w = 200$ ps. The value of the integral for these parameter values is plotted in Figure 7.1b. The maximal fractional phase change is $\frac{2}{3}$.

Bibliography

[1] Kyungjoo Noh, SM Girvin, and Liang Jiang. Encoding an oscillator into many oscillators. *Physical Review Letters*, 125(8):080503, 2020.

[2] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011. ISBN 1107002176, 9781107002173.

[3] Lov K Grover. Quantum mechanics helps in searching for a needle in a haystack. *Physical review letters*, 79(2):325, 1997.

[4] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[5] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[6] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.

[7] Ronald L Rivest, Adi Shamir, and Leonard Adleman. A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2):120–126, 1978.

[8] Valerio Scarani, Helle Bechmann-Pasquinucci, Nicolas J. Cerf, Miloslav Dušek, Norbert Lütkenhaus, and Momtchil Peev. The security of practical quantum key distribution. *Rev. Mod. Phys.*, 81:1301–1350, Sep 2009.

[9] William K Wootters and Wojciech H Zurek. A single quantum cannot be cloned. *Nature*, 299(5886): 802–803, 1982.

[10] Artur K Ekert. Quantum cryptography based on bell's theorem. *Physical review letters*, 67(6):661, 1991.

[11] Ryszard Horodecki, Pawel Horodecki, Michal Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of modern physics*, 81(2):865, 2009.

[12] Patrick J. Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. *Reviews of Modern Physics*, 89(1):1–65, 2017.

[13] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. *Phys. Rev. A*, 64:012310, Jun 2001.

[14] Emanuel Knill, Raymond Laflamme, and Gerald J Milburn. A scheme for efficient quantum computation with linear optics. *Nature*, 409(6816):46–52, 2001.

[15] Victor V Albert, Kyungjoo Noh, Kasper Duivenvoorden, Dylan J Young, RT Brierley, Philip Reinhold, Christophe Vuillot, Linshu Li, Chao Shen, SM Girvin, et al. Performance and structure of single-mode bosonic codes. *Phys. Rev. A*, 97(3):032346, 2018.

[16] Ilan Tzitrin, J. Eli Bourassa, Nicolas C. Menicucci, and Krishna Kumar Sabapathy. Progress towards practical qubit computation using approximate gottesman-kitaev-preskill codes. *Phys. Rev. A*, 101: 032315, 3 2020.

[17] Krishna Kumar Sabapathy, Haoyu Qi, Josh Izaac, and Christian Weedbrook. Production of photonic universal quantum gates enhanced by machine learning. *Phys. Rev. A*, 100:012326, Jul 2019.

[18] Daiqin Su, Casey R. Myers, and Krishna Kumar Sabapathy. Generation of photonic non-gaussian states by measuring multimode gaussian states, 2019.

[19] Daiqin Su, Casey R. Myers, and Krishna Kumar Sabapathy. Conversion of gaussian states to non-gaussian states using photon-number-resolving detectors. *Phys. Rev. A*, 100:052301, Nov 2019.

[20] N. Quesada, L. G. Helt, J. Izaac, J. M. Arrazola, R. Shahrokhshahi, C. R. Myers, and K. K. Sabapathy. Simulating realistic non-gaussian state preparation. *Phys. Rev. A*, 100:022341, Aug 2019.

[21] J Eli Bourassa, Nicolás Quesada, Ilan Tzitrin, Antal Száva, Theodor Isacsson, Josh Izaac, Krishna Kumar Sabapathy, Guillaume Dauphinais, and Ish Dhand. Fast simulation of bosonic qubits via gaussian functions in phase space. *arXiv preprint arXiv:2103.05530*, 2021.

[22] J Eli Bourassa, Rafael N Alexander, Michael Vasmer, Ashlesha Patil, Ilan Tzitrin, Takaya Matsuura, Daiqin Su, Ben Q Baragiola, Saikat Guha, Guillaume Dauphinais, et al. Blueprint for a scalable photonic fault-tolerant quantum computer. *Quantum*, 5:392, 2021.

[23] J Eli Bourassa and Hoi-Kwong Lo. Entropic uncertainty relations and the measurement range problem, with consequences for high-dimensional quantum key distribution. *JOSA B*, 36(3):B65–B76, 2019.

[24] Hoi-Kwong Lo, Marcos Curty, and Bing Qi. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.*, 108:130503, Mar 2012.

[25] Kiyoshi Tamaki, Marcos Curty, Go Kato, Hoi-Kwong Lo, and Koji Azuma. Loss-tolerant quantum cryptography with imperfect sources. *Phys. Rev. A*, 90:052314, Nov 2014.

[26] J Eli Bourassa, Amita Gnanapandithan, Li Qian, and Hoi-Kwong Lo. Measurement device-independent quantum key distribution with passive, time-dependent source side-channels. *arXiv preprint arXiv:2108.08698*, 2021.

[27] Ignatius William Primaatmaja, Emilien Lavie, Koon Tong Goh, Chao Wang, and Charles Ci Wen Lim. Versatile security analysis of measurement-device-independent quantum key distribution. *Phys. Rev. A*, 99:062332, Jun 2019.

[28] Kyungjoo Noh, Victor V. Albert, and Liang Jiang. Quantum capacity bounds of gaussian thermal loss channels and achievable rates with gottesman-kitaev-preskill codes. *IEEE Trans. Inf. Theory*, 65(4): 2563–2582, 4 2019.

[29] Kosuke Fukui, Rafael N Alexander, and Peter van Loock. All-optical long-distance quantum communication with gottesman-kitaev-preskill qubits. *arXiv preprint arXiv:2011.14876*, 2020.

[30] Filip Rozpędek, Kyungjoo Noh, Qian Xu, Saikat Guha, and Liang Jiang. Quantum repeaters based on concatenated bosonic and discrete-variable quantum codes. *arXiv preprint arXiv:2011.15076*, 2020.

[31] Kasper Duivenvoorden, Barbara M Terhal, and Daniel Weigand. Single-mode displacement sensor. *Phys. Rev. A*, 95(1):012305, 2017.

[32] Quntao Zhuang, John Preskill, and Liang Jiang. Distributed quantum sensing enhanced by continuous-variable error correction. *New Journal of Physics*, 22(2):022001, 2020.

[33] S. Glancy and E. Knill. Error analysis for encoding a qubit in an oscillator. *Phys. Rev. A*, 73:012325, Jan 2006.

[34] Nicolas C. Menicucci. Fault-tolerant measurement-based quantum computing with continuous-variable cluster states. *Phys. Rev. Lett.*, 112:120504, 3 2014.

[35] Kosuke Fukui, Akihisa Tomita, Atsushi Okamoto, and Keisuke Fujii. High-threshold fault-tolerant quantum computation with analog quantum error correction. *Phys. Rev. X*, 8(2):021054, 2018.

[36] Ben Q. Baragiola, Giacomo Pantaleoni, Rafael N. Alexander, Angela Karanjai, and Nicolas C. Menicucci. All-Gaussian universality and fault tolerance with the Gottesman-Kitaev-Preskill code. *Phys. Rev. Lett.*, 123:200502, Nov 2019.

[37] Blayney W. Walshe, Lucas J. Mensen, Ben Q. Baragiola, and Nicolas C. Menicucci. Robust fault tolerance for continuous-variable cluster states with excess antisqueezing. *Phys. Rev. A*, 100:010301, Jul 2019.

[38] Kosuke Fukui. High-threshold fault-tolerant quantum computation with the GKP qubit and realistically noisy devices. *arXiv preprint arXiv:1906.09767*, 2019.

[39] Christophe Vuillot, Hamed Asasi, Yang Wang, Leonid P Pryadko, and Barbara M Terhal. Quantum error correction with the toric gottesman-kitaev-preskill code. *Phys. Rev. A*, 99(3):032344, 2019.

[40] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf, L. Frunzio, M. Mirrahimi, and M. H. Devoret. Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature*, 584(7821):368–372, August 2020.

[41] B M Terhal, J Conrad, and C Vuillot. Towards scalable bosonic quantum error correction. *Quantum Science and Technology*, 5(4):043001, Jul 2020.

[42] Hayata Yamasaki, Kosuke Fukui, Yuki Takeuchi, Seiichiro Tani, and Masato Koashi. Polylog-overhead highly fault-tolerant measurement-based quantum computation: all-Gaussian implementation with Gottesman-Kitaev-Preskill code. *arXiv preprint arXiv:2006.05416*, 2020.

[43] J Eli Bourassa, Ignatius William Primaatmaja, Charles Ci Wen Lim, and Hoi-Kwong Lo. Loss-tolerant quantum key distribution with mixed signal states. *Physical Review A*, 102(6):062607, 2020.

[44] Mikkel V Larsen, Christopher Chamberland, Kyungjoo Noh, Jonas S Neergaard-Nielsen, and Ulrik L Andersen. A fault-tolerant continuous-variable measurement-based quantum computation architecture. *arXiv preprint arXiv:2101.03014*, 2021.

[45] Kyungjoo Noh, Christopher Chamberland, and Fernando GSL Brandão. Low overhead fault-tolerant quantum error correction with the surface-gkp code. *arXiv preprint arXiv:2103.06994*, 2021.

[46] Ilan Tzitrin, Takaya Matsuura, Rafael N Alexander, Guillaume Dauphinais, J Eli Bourassa, Krishna K Sabapathy, Nicolas C Menicucci, and Ish Dhand. Fault-tolerant quantum computation with static linear optics. *arXiv preprint arXiv:2104.03241*, 2021.

[47] Kyungjoo Noh and Christopher Chamberland. Fault-tolerant bosonic quantum error correction with the surface–Gottesman-Kitaev-Preskill code. *Phys. Rev. A*, 101:012316, Jan 2020.

[48] Lisa Hänggli, Margret Heinze, and Robert König. Enhanced noise resilience of the surface–gottesman-kitaev-preskill code via designed bias. *Phys. Rev. A*, 102:052408, Nov 2020.

[49] Hayata Yamasaki, Takaya Matsuura, and Masato Koashi. Cost-reduced all-gaussian universality with the gottesman-kitaev-preskill code: Resource-theoretic approach to cost analysis. *Phys. Rev. Research*, 2:023270, Jun 2020.

[50] Takaya Matsuura, Hayata Yamasaki, and Masato Koashi. Equivalence of approximate gottesman-kitaev-preskill codes. *Phys. Rev. A*, 102:032408, Sep 2020.

[51] Lucas J Mensen, Ben Q Baragiola, and Nicolas C Menicucci. Phase-space methods for representing, manipulating, and correcting gottesman-kitaev-preskill qubits. *arXiv preprint arXiv:2012.12488*, 2020.

[52] Laura García-Álvarez, Cameron Calcluth, Alessandro Ferraro, and Giulia Ferrini. Efficient simulatability of continuous-variable circuits with large wigner negativity. *Physical Review Research*, 2(4):043322, 2020.

[53] L. García-Álvarez, A. Ferraro, and G. Ferrini. From the bloch sphere to phase-space representations with the gottesman–kitaev–preskill encoding. In Tsuyoshi Takagi, Masato Wakayama, Keisuke Tanaka, Noboru Kunihiro, Kazufumi Kimoto, and Yasuhiko Ikematsu, editors, *International Symposium on Mathematics, Quantum Theory, and Cryptography*, pages 79–92, Singapore, 2021. Springer Singapore.

[54] Christa Flühmann, Thanh Long Nguyen, Matteo Marinelli, Vlad Negnevitsky, Karan Mehta, and JP Home. Encoding a qubit in a trapped-ion mechanical oscillator. *Nature*, 566(7745):513, 2019.

[55] C. Flühmann and J. P. Home. Direct characteristic-function tomography of quantum states of the trapped-ion motional oscillator. *Phys. Rev. Lett.*, 125:043602, Jul 2020.

[56] Brennan de Neeve, Thanh Long Nguyen, Tanja Behrle, and Jonathan Home. Error correction of a logical grid state qubit by dissipative pumping. *arXiv preprint arXiv:2010.09681*, 2020.

[57] Adriana E. Lita, Aaron J. Miller, and Sae Woo Nam. Counting near-infrared single-photons with 95% efficiency. *Opt. Express*, 16(5):3032, 2008.

[58] Hilma M Vasconcelos, Liliana Sanz, and Scott Glancy. All-optical generation of states for ëncoding a qubit in an oscillator. *Optics letters*, 35(19):3261–3263, 2010.

[59] Miller Eaton, Rajveer Nehra, and Olivier Pfister. Non-Gaussian and Gottesman–Kitaev–Preskill state preparation by photon catalysis. *New J. Phys.*, 21(11):113034, nov 2019.

[60] Keith R Motes, Ben Q Baragiola, Alexei Gilchrist, and Nicolas C Menicucci. Encoding qubits into oscillators with atomic ensembles and squeezed light. *Phys. Rev. A*, 95(5):053819, 2017.

[61] Giacomo Pantaleoni, Ben Q Baragiola, and Nicolas C Menicucci. Modular bosonic subsystem codes. *Phys. Rev. Lett.*, 125(4):040501, 2020.

[62] Jacob Hastrup, Mikkel V Larsen, Jonas S Neergaard-Nielsen, Nicolas C Menicucci, and Ulrik L Andersen. Unsuitability of cubic phase gates for non-clifford operations on gottesman-kitaev-preskill states. *Physical Review A*, 103(3):032409, 2021.

[63] Giacomo Pantaleoni, Ben Q Baragiola, and Nicolas C Menicucci. Subsystem analysis of continuous-variable resource states. *arXiv preprint arXiv:2102.10500*, 2021.

[64] Giacomo Pantaleoni, Ben Q Baragiola, and Nicolas C Menicucci. Hidden qubit cluster states. *arXiv preprint arXiv:2103.11556*, 2021.

[65] Krishna Kumar Sabapathy. Some aspects of the interplay between bipartite correlations and quantum channels, 2016.

[66] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Phys. Rev. A*, 71:022316, Feb 2005.

[67] Alison L. Gibbs and Francis Edward Su. On choosing and bounding probability metrics. *Int. Stat. Rev.*, 70(3):419–435, dec 2002.

[68] Michael A. Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. *Phys. Lett. Sect. A Gen. At. Solid State Phys.*, 303(4):249–252, oct 2002.

[69] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, Aug 1996.

[70] A. M. Steane. Error correcting codes in quantum theory. *Phys. Rev. Lett.*, 77:793–797, Jul 1996.

[71] E. Knill. Scalable quantum computing in the presence of large detected-error rates. *Phys. Rev. A*, 71: 042322, Apr 2005.

[72] Emanuel Knill. Quantum computing with realistically noisy devices. *Nature*, 434(7029):39, 2005.

[73] Yang Wang. Quantum error correction with the gkp code and concatenation with stabilizer codes, 2019.

[74] Jaromír Fiurášek, Raúl García-Patrón, and Nicolas J Cerf. Conditional generation of arbitrary single-mode quantum states of light by repeated photon subtractions. *Phys. Rev. A*, 72(3):033822, 2005.

[75] BC Travaglione and Gerard J Milburn. Preparing encoded states in an oscillator. *Phys. Rev. A*, 66(5): 052322, 2002.

[76] Stefano Pirandola, Stefano Mancini, David Vitali, and Paolo Tombesi. Continuous variable encoding by ponderomotive interaction. *The European Physical Journal D-Atomic, Molecular, Optical and Plasma Physics*, 37(2):283–290, 2006.

[77] C. Flühmann, V. Negnevitsky, M. Marinelli, and J. P. Home. Sequential modular position and momentum measurements of a trapped ion mechanical oscillator. *Phys. Rev. X*, 8:021001, Apr 2018.

[78] David Menzies and Radim Filip. Gaussian-optimized preparation of non-gaussian pure states. *Phys. Rev. A*, 79:012313, Jan 2009.

[79] Ulysse Chabaud, Damian Markham, and Frédéric Grosshans. The stellar representation of non-gaussian quantum states. *Phys. Rev. Lett.*, 124:063605, 2020.

[80] Biswadeb Dutta, N Mukunda, R Simon, et al. The real symplectic groups in quantum mechanics and optics. *Pramana*, 45(6):471–497, 1995.

[81] Yoshichika Miwa, Jun-ichi Yoshikawa, Noriaki Iwata, Mamoru Endo, Petr Marek, Radim Filip, Peter van Loock, and Akira Furusawa. Exploring a new regime for processing optical qubits: Squeezing and unsqueezing single photons. *Phys. Rev. Lett.*, 113:013601, Jul 2014.

[82] Pauli Virtanen, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature methods*, 17(3):261–272, 2020.

[83] David J. Wales and Jonathan P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, July 1997.

[84] Nathan Killoran, Josh Izaac, Nicolás Quesada, Ville Bergholm, Matthew Amy, and Christian Weedbrook. Strawberry fields: A software platform for photonic quantum computing. *Quantum*, 3:129, March 2019.

[85] Andreas Björklund, Brajesh Gupt, and Nicolás Quesada. A faster hafnian formula for complex matrices and its benchmarking on a supercomputer. *Journal of Experimental Algorithmics (JEA)*, 24:1–17, 2019.

[86] Brajesh Gupt, Josh Izaac, and Nicolás Quesada. The walrus: a library for the calculation of hafnians, hermite polynomials and gaussian boson sampling. *J. Open Source Softw.*, 4(44):1705, 2019.

[87] Henning Vahlbruch, Moritz Mehmet, Karsten Danzmann, and Roman Schnabel. Detection of 15 db squeezed states of light and their application for the absolute calibration of photoelectric quantum efficiency. *Phys. Rev. Lett.*, 117:110801, Sep 2016.

[88] https://github.com/XanaduAI/approximate-GKP-prep.

[89] Kazunori Miyata, Hisashi Ogawa, Petr Marek, Radim Filip, Hidehiro Yonezawa, Jun-ichi Yoshikawa, and Akira Furusawa. Experimental realization of a dynamic squeezing gate. *Phys. Rev. A*, 90:060302, Dec 2014.

[90] Christopher Chamberland, Kyungjoo Noh, Patricio Arrangoiz-Arriola, Earl T Campbell, Connor T Hann, Joseph Iverson, Harald Putterman, Thomas C Bohdanowicz, Steven T Flammia, Andrew Keller, et al. Building a fault-tolerant quantum computer using concatenated cat codes. *arXiv preprint arXiv:2012.04108*, 2020.

[91] Sara Bartolucci, Patrick Birchall, Hector Bombin, Hugo Cable, Chris Dawson, Mercedes Gimeno-Segovia, Eric Johnston, Konrad Kieling, Naomi Nickerson, Mihir Pant, et al. Fusion-based quantum computation. *arXiv preprint arXiv:2101.09310*, 2021.

[92] Shruti Puri, Lucas St-Jean, Jonathan A. Gross, Alexander Grimm, Nicholas E. Frattini, Pavithran S. Iyer, Anirudh Krishna, Steven Touzard, Liang Jiang, Alexandre Blais, Steven T. Flammia, and S. M. Girvin. Bias-preserving gates with stabilized cat qubits. *Science Advances*, 6(34):eaay5901, August 2020.

[93] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret. Stabilization and operation of a kerr-cat qubit. *Nature*, 584(7820):205–209, August 2020.

[94] C R Myers and T C Ralph. Coherent state topological cluster state production. *New J. Phys.*, 13(11): 115015, November 2011.

[95] Kosuke Fukui, Warit Asavanant, and Akira Furusawa. Temporal-mode continuous-variable three-dimensional cluster state for topologically protected measurement-based quantum computation. *Phys. Rev. A*, 102:032614, Sep 2020.

[96] Timothy C. Ralph, Alexei Gilchrist, Gerard J Milburn, William J Munro, and Scott Glancy. Quantum computation with optical coherent states. *Phys. Rev. A*, 68(4):042319, 2003.

[97] Isaac L Chuang and Yoshihisa Yamamoto. Simple quantum computer. *Phys. Rev. A*, 52(5):3489, 1995.

[98] Arne L Grimsmo, Joshua Combes, and Ben Q Baragiola. Quantum computing with rotation-symmetric bosonic codes. *Phys. Rev. X*, 10(1):011058, 2020.

[99] Filippo M. Miatto and Nicolás Quesada. Fast optimization of parametrized quantum optical circuits. *Quantum*, 4:366, 2020.

[100] A. Mari and J. Eisert. Positive wigner functions render classical simulation of quantum computation efficient. *Phys. Rev. Lett.*, 109:230503, Dec 2012.

[101] Radim Filip, Petr Marek, and Ulrik L. Andersen. Measurement-induced continuous-variable quantum interactions. *Phys. Rev. A*, 71:042308, Apr 2005.

[102] https://github.com/XanaduAI/strawberryfields.

[103] An introduction to the bosonic backend. https://strawberryfields.ai/photonics/demos/run_intro_bosonic.html, Mar 2021.

[104] Sampling quadratures of non-gaussian states using the bosonic backend. https://strawberryfields.ai/photonics/demos/run_sampling_bosonic.html, Mar 2021.

[105] Studying realistic bosonic qubits using the bosonic backend. https://strawberryfields.ai/photonics/demos/run_GKP_bosonic.html, Mar 2021.

[106] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84:621–669, May 2012.

[107] R. Simon, N. Mukunda, and Biswadeb Dutta. Quantum-noise matrix for multimode systems: U(n) invariance, squeezing, and normal forms. *Phys. Rev. A*, 49:1567–1583, Mar 1994.

[108] R. Simon, E. C. G. Sudarshan, and N. Mukunda. Gaussian-wigner distributions in quantum mechanics and optics. *Phys. Rev. A*, 36:3868–3880, Oct 1987.

[109] Alessio Serafini. *Quantum Continuous Variables*. CRC Press, 2017.

[110] Thomas R Bromley, Juan Miguel Arrazola, Soran Jahangiri, Josh Izaac, Nicolás Quesada, Alain Delgado Gran, Maria Schuld, Jeremy Swinarton, Zeid Zabaneh, and Nathan Killoran. Applications of near-term photonic quantum computers: Software and algorithms. *Quantum Sci. Technol.*, 5(3):034010, 2020.

[111] J. M. Arrazola, V. Bergholm, K. Brádler, T. R. Bromley, M. J. Collins, I. Dhand, A. Fumagalli, T. Gerrits, A. Goussev, L. G. Helt, J. Hundal, T. Isacsson, R. B. Israel, J. Izaac, S. Jahangiri, R. Janik, N. Killoran, S. P. Kumar, J. Lavoie, A. E. Lita, D. H. Mahler, M. Menotti, B. Morrison, S. W. Nam, L. Neuhaus, H. Y. Qi, N. Quesada, A. Repingon, K. K. Sabapathy, M. Schuld, D. Su, J. Swinarton, A. Száva, K. Tan, P. Tan, V. D. Vaidya, Z. Vernon, Z. Zabaneh, and Y. Zhang. Quantum circuits with many photons on a programmable nanophotonic chip. *Nature*, 591:54–60, 2021.

[112] Nicolás Quesada and Juan Miguel Arrazola. Exact simulation of gaussian boson sampling in polynomial space and exponential time. *Phys. Rev. Research*, 2(2):023005, 2020.

[113] Charles R. Harris, K. Jarrod Millman, St'efan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fern'andez del R'ıo, Mark Wiebe, Pearu Peterson, Pierre G'erard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. *Nature*, 585(7825):357–362, September 2020. doi: 10.1038/s41586-020-2649-2.

[114] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, pages 265–283, 2016.

[115] Kwok Ho Wan, Alex Neville, and Steve Kolthammer. Memory-assisted decoder for approximate gottesman-kitaev-preskill codes. *Phys. Rev. Research*, 2:043280, Nov 2020.

[116] Namrata Shukla, Stefan Nimmrichter, and Barry C Sanders. Squeezed comb states. *Phys. Rev. A*, 103 (1):012408, 2021.

[117] Shohini Ghose and Barry C. Sanders. Non-gaussian ancilla states for continuous variable quantum computation via gaussian maps. *J. Mod. Opt.*, 54(6):855–869, March 2007.

[118] Seckin Sefi, Vishal Vaibhav, and Peter van Loock. Measurement-induced optical kerr interaction. *Phys. Rev. A*, 88:012303, Jul 2013.

[119] Kevin Marshall, Raphael Pooser, George Siopsis, and Christian Weedbrook. Repeat-until-success cubic phase gate for universal continuous-variable quantum computation. *Phys. Rev. A*, 91:032321, Mar 2015.

[120] Kazunori Miyata, Hisashi Ogawa, Petr Marek, Radim Filip, Hidehiro Yonezawa, Jun-ichi Yoshikawa, and Akira Furusawa. Implementation of a quantum cubic gate by an adaptive non-Gaussian measurement. *Phys. Rev. A*, 93(2):022301, 2016.

[121] Francesco Arzani, Nicolas Treps, and Giulia Ferrini. Polynomial approximation of non-gaussian unitaries by counting one photon at a time. *Phys. Rev. A*, 95:052352, May 2017.

[122] Krishna Kumar Sabapathy, J Solomon Ivan, and R Simon. Robustness of non-Gaussian entanglement against noisy amplifier and attenuator environments. *Phys. Rev. Lett.*, 107(13):130501, 2011.

[123] Raul Garcia-Patron, Carlos Navarrete-Benlloch, Seth Lloyd, Jeffrey H Shapiro, and Nicolas J. Cerf. Majorization theory approach to the Gaussian channel minimum entropy conjecture. *Phys. Rev. Lett.*, 108(11):110505, 2012.

[124] F Caruso, V Giovannetti, and A S Holevo. One-mode bosonic gaussian channels: a full weak-degradability classification. *New J. Phys.*, 8(12):310–310, December 2006.

[125] J. Solomon Ivan, Krishna Kumar Sabapathy, and R. Simon. Operator-sum representation for bosonic Gaussian channels. *Phys. Rev. A*, 84:042311, Oct 2011.

[126] Z Vernon, N Quesada, M Liscidini, B Morrison, M Menotti, K Tan, and JE Sipe. Scalable squeezed-light source for continuous-variable quantum sampling. *Phys. Rev. Applied*, 12(6):064024, 2019.

[127] L G Helt and N Quesada. Degenerate squeezing in waveguides: a unified theoretical approach. *J. Phys.: Photonics*, 2(3):035001, 2020.

[128] Blayney W Walshe, Ben Q Baragiola, Rafael N Alexander, and Nicolas C Menicucci. Continuous-variable gate teleportation and bosonic-code error correction. *Phys. Rev. A*, 102(6):062411, 2020.

[129] Victor Veitch, Nathan Wiebe, Christopher Ferrie, and Joseph Emerson. Efficient simulation scheme for a class of quantum optics experiments with non-negative wigner representation. *New Journal of Physics*, 15(1):013037, 2013.

[130] Nicolás Quesada, Juan Miguel Arrazola, and Nathan Killoran. Gaussian boson sampling using threshold detectors. *Phys. Rev. A*, 98:062322, 9 2018.

[131] Art B. Owen. *Monte Carlo theory, methods and examples.* 2013.

[132] Guillaume Rabusseau and François Denis. Learning negative mixture models by tensor decompositions. *arXiv preprint arXiv:1403.4224*, 2014.

[133] Anatole Kenfack and Karol Zyczkowski. Negativity of the Wigner function as an indicator of non-classicality. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(10):396–404, jun 2004.

[134] N. M. Temme. Asymptotic estimates for laguerre polynomials. *Z. Angew. Math. Phys.*, 41(1):114–126, January 1990.

[135] Stephen D Bartlett, Barry C Sanders, Samuel L Braunstein, and Kae Nemoto. Efficient classical simulation of continuous variable quantum information processes. *Phys. Rev. Lett.*, 88(9):097904, 2002.

[136] Daniel Gottesman. The heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998.

[137] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Physical Review A*, 70(5):052328, 2004.

[138] Sergey Bravyi, Graeme Smith, and John A Smolin. Trading classical and quantum computational resources. *Physical Review X*, 6(2):021043, 2016.

[139] Sergey Bravyi and David Gosset. Improved classical simulation of quantum circuits dominated by clifford gates. *Physical review letters*, 116(25):250501, 2016.

[140] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, David Gosset, and Mark Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum*, 3:181, 2019.

[141] Daniel J Weigand and Barbara M Terhal. Generating grid states from schrödinger-cat states without postselection. *Phys. Rev. A*, 97(2):022341, 2018.

[142] Ulrik L Andersen, Jonas S Neergaard-Nielsen, Peter van Loock, and Akira Furusawa. Hybrid discrete- and continuous-variable quantum information. *Nat. Phys.*, 11(9):713–719, 2015.

[143] Nicolas C. Menicucci, Peter van Loock, Mile Gu, Christian Weedbrook, Timothy C. Ralph, and Michael A. Nielsen. Universal quantum computation with continuous-variable cluster states. *Phys. Rev. Lett.*, 97:110501, Sep 2006.

[144] Shota Yokoyama, Ryuji Ukai, Seiji C Armstrong, Chanond Sornphiphatphong, Toshiyuki Kaji, Shigenari Suzuki, Jun-ichi Yoshikawa, Hidehiro Yonezawa, Nicolas C. Menicucci, and Akira Furusawa. Ultra-large-scale continuous-variable cluster states multiplexed in the time domain. *Nat. Photonics*, 7(12): 982–986, 2013.

[145] Moran Chen, Nicolas C. Menicucci, and Olivier Pfister. Experimental realization of multipartite entanglement of 60 modes of a quantum optical frequency comb. *Phys. Rev. Lett.*, 112(12):120505, 2014.

[146] Jun-ichi Yoshikawa, Shota Yokoyama, Toshiyuki Kaji, Chanond Sornphiphatphong, Yu Shiozawa, Kenzo Makino, and Akira Furusawa. Invited article: Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing. *APL Photonics*, 1(6):060801, 2016.

[147] Rafael N. Alexander, Pei Wang, Niranjan Sridhar, Moran Chen, Olivier Pfister, and Nicolas C. Menicucci. One-way quantum computing with arbitrarily large time-frequency continuous-variable cluster states from a single optical parametric oscillator. *Phys. Rev. A*, 94:032327, Sep 2016.

[148] Mikkel V. Larsen, Xueshi Guo, Casper R. Breum, Jonas S. Neergaard-Nielsen, and Ulrik L. Andersen. Deterministic generation of a two-dimensional cluster state. *Science*, 366(6463):369–372, Oct 2019.

[149] Mikkel V. Larsen, Jonas S. Neergaard-Nielsen, and Ulrik L. Andersen. Architecture and noise analysis of continuous-variable quantum gates using two-dimensional cluster states. *Phys. Rev. A*, 102:042608, Oct 2020.

[150] Rafael N. Alexander, Shota Yokoyama, Akira Furusawa, and Nicolas C. Menicucci. Universal quantum computation with temporal-mode bilayer square lattices. *Phys. Rev. A*, 97:032302, Mar 2018.

[151] Warit Asavanant, Yu Shiozawa, Shota Yokoyama, Baramee Charoensombutamon, Hiroki Emura, Rafael N. Alexander, Shuntaro Takeda, Jun-ichi Yoshikawa, Nicolas C. Menicucci, Hidehiro Yonezawa, and et al. Generation of time-domain-multiplexed two-dimensional cluster state. *Science*, 366(6463): 373–376, Oct 2019.

[152] Pei Wang, Moran Chen, Nicolas C. Menicucci, and Olivier Pfister. Weaving quantum optical frequency combs into continuous-variable hypercubic cluster states. *Phys. Rev. A*, 90(3):032325, 2014.

[153] Bo-Han Wu, Rafael N Alexander, Shuai Liu, and Zheshen Zhang. Quantum computing with multi-dimensional continuous-variable cluster states in a scalable photonic platform. *Phys. Rev. Res.*, 2(2): 023138, 2020.

[154] Austin P. Lund, Timothy C. Ralph, and Henry L. Haselgrove. Fault-tolerant linear optical quantum computing with small-amplitude coherent states. *Phys. Rev. Lett.*, 100:030503, Jan 2008.

[155] Terry Rudolph. Why I am optimistic about the silicon-photonic route to quantum computing. *APL Photonics*, 2(3):030901, 2017.

[156] M. Dakna, J. Clausen, L. Knöll, and D.-G. Welsch. Generation of arbitrary quantum states of traveling fields. *Phys. Rev. A*, 59:1658–1661, Feb 1999.

[157] Stefano Pirandola, Stefano Mancini, David Vitali, and Paolo Tombesi. Generating continuous variable quantum codewords in the near-field atomic lithography. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 39(4):997, 2006.

[158] J. Sperling, D. S. Phillips, J. F. F Bulmer, G. S. Thekkadath, A. Eckstein, T. A. W. Wolterink, J. Lugani, S. W. Nam, A. Lita, T. Gerrits, W. Vogel, G. S. Agarwal, C. Silberhorn, and I. A. Walmsley. Detector-agnostic phase-space distributions. *Phys. Rev. Lett.*, 124:013605, Jan 2020.

[159] G. S. Thekkadath, D. S. Phillips, J. F. F. Bulmer, W. R. Clements, A. Eckstein, B. A. Bell, J. Lugani, T. A. W. Wolterink, A. Lita, S. W. Nam, T. Gerrits, C. G. Wade, and I. A. Walmsley. Tuning between photon-number and quadrature measurements with weak-field homodyne detection. *Phys. Rev. A*, 101: 031801, Mar 2020.

[160] V. D. Vaidya, B. Morrison, L. G. Helt, R. Shahrokshahi, D. H. Mahler, M. J. Collins, K. Tan, J. Lavoie, A. Repingon, M. Menotti, N. Quesada, R. C. Pooser, A. E. Lita, T. Gerrits, S. W. Nam, and Z. Vernon. Broadband quadrature-squeezed vacuum and nonclassical photon number correlations from a nanophotonic device. *Sci. Adv.*, 6(39):eaba9186, September 2020.

[161] Robert Raussendorf, Sergey Bravyi, and Jim Harrington. Long-range quantum entanglement in noisy cluster states. *Phys. Rev. A*, 71:062313, Jun 2005.

[162] Robert Raussendorf, Jim Harrington, and Kovid Goyal. A fault-tolerant one-way quantum computer. *Ann. Phys. (N. Y.)*, 321(9):2242–2270, 2006.

[163] Robert Raussendorf, Jim Harrington, and Kovid Goyal. Topological fault-tolerance in cluster state quantum computation. *New J. Phys.*, 9(6):199, 2007.

[164] Thomas M Stace, Sean D Barrett, and Andrew C Doherty. Thresholds for topological codes in the presence of loss. *Phys. Rev. Lett.*, 102(20):200501, 2009.

[165] Sean D Barrett and Thomas M Stace. Fault tolerant quantum computation with very high threshold for loss errors. *Phys. Rev. Lett.*, 105(20):200502, 2010.

[166] James M. Auger, Hussain Anwar, Mercedes Gimeno-Segovia, Thomas M. Stace, and Dan E. Browne. Fault-tolerant quantum computation with nondeterministic entangling gates. *Phys. Rev. A*, 97(3):5–9, 2018.

[167] Adam C. Whiteside and Austin G. Fowler. Upper bound for loss in practical topological-cluster-state quantum computing. *Phys. Rev. A*, 90:052316, Nov 2014.

[168] P. T. Cochrane, G. J. Milburn, and W. J. Munro. Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping. *Phys. Rev. A*, 59:2631–2634, Apr 1999.

[169] Zaki Leghtas, Gerhard Kirchmair, Brian Vlastakis, Robert J. Schoelkopf, Michel H. Devoret, and Mazyar Mirrahimi. Hardware-efficient autonomous quantum memory protection. *Phys. Rev. Lett.*, 111: 120501, Sep 2013.

[170] Mazyar Mirrahimi, Zaki Leghtas, Victor V Albert, Steven Touzard, Robert J Schoelkopf, Liang Jiang, and Michel H Devoret. Dynamically protected cat-qubits: a new paradigm for universal quantum computation. *New J. Phys.*, 16(4):045014, April 2014.

[171] Marios H. Michael, Matti Silveri, R. T. Brierley, Victor V. Albert, Juha Salmilehto, Liang Jiang, and S. M. Girvin. New class of quantum error-correcting codes for a bosonic mode. *Phys. Rev. X*, 6:031006, Jul 2016.

[172] A. I. Lvovsky and M. G. Raymer. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.*, 81:299–332, Mar 2009.

[173] Hans J Briegel and Robert Raussendorf. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.*, 86(5):910, 2001.

[174] Robert Raussendorf and Hans J Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86(22):5188, 2001.

[175] Mile Gu, Christian Weedbrook, Nicolas C. Menicucci, Timothy C. Ralph, and Peter van Loock. Quantum computing with continuous-variable clusters. *Phys. Rev. A*, 79:062318, Jun 2009.

[176] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Phys. Rev. A*, 52(4): 2493–2496, 1995.

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

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

[179] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. Roads towards fault-tolerant universal quantum computation. *Nature*, 549(7671):172–179, Sep 2017.

[180] A.Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Ann. Phys. (N. Y.)*, 303(1):2 – 30, 2003.

[181] Sergey B Bravyi and A Yu Kitaev. Quantum codes on a lattice with boundary. *arXiv preprint quant-ph/9811052*, 1998.

[182] Robert Raussendorf and Jim Harrington. Fault-tolerant quantum computation with high threshold in two dimensions. *Phys. Rev. Lett.*, 98(19), May 2007.

[183] Austin G. Fowler, Ashley M. Stephens, and Peter Groszkowski. High-threshold universal quantum computation on the surface code. *Phys. Rev. A*, 80:052312, Nov 2009.

[184] Austin G. Fowler, Matteo Mariantoni, John M. Martinis, and Andrew N. Cleland. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A*, 86(3), Sep 2012.

[185] Austin G Fowler and Kovid Goyal. Topological cluster state quantum computing. *Quantum Inf. Comput.*, 9(9-10):0721–0738, 2009.

[186] Clare Horsman, Austin G Fowler, Simon Devitt, and Rodney Van Meter. Surface code quantum computing by lattice surgery. *New J. Phys.*, 14(12):123011, dec 2012.

[187] Benjamin J. Brown, Katharina Laubscher, Markus S. Kesselring, and James R. Wootton. Poking holes and cutting corners to achieve Clifford gates with the surface code. *Phys. Rev. X*, 7:021029, May 2017.

[188] Daniel Litinski and Felix von Oppen. Lattice surgery with a twist: Simplifying Clifford gates of surface codes. *Quantum*, 2:62, May 2018.

[189] Daniel Litinski. A Game of Surface Codes: Large-Scale Quantum Computing with Lattice Surgery. *Quantum*, 3:128, March 2019.

[190] Daniel Herr, Alexandru Paler, Simon J Devitt, and Franco Nori. Lattice surgery on the Raussendorf lattice. *Quantum Sci. Technol.*, 3(3):035011, jun 2018.

[191] Benjamin J. Brown and Sam Roberts. Universal fault-tolerant measurement-based quantum computation. *Phys. Rev. Research*, 2:033305, Aug 2020.

[192] Damien Bonneau, Gabriel J Mendoza, Jeremy L O'Brien, and Mark G Thompson. Effect of loss on multiplexed single-photon sources. *New J. Phys.*, 17(4):043057, April 2015.

[193] Nicolas C. Menicucci. Temporal-mode continuous-variable cluster states using linear optics. *Phys. Rev. A*, 83:062314, Jun 2011.

[194] Ish Dhand, Melanie. Engelkemeier, Linda. Sansoni, Sonja Barkhofen, Christine Silberhorn, and Martin B. Plenio. Proposal for quantum simulation via all-optically-generated tensor network states. *Phys. Rev. Lett.*, 120:130501, 2018.

[195] Michael Lubasch, Antonio A. Valido, Jelmer J. Renema, W. Steven Kolthammer, Dieter Jaksch, M. S. Kim, Ian Walmsley, and Raúl García-Patrón. Tensor network states in time-bin quantum optics. *Phys. Rev. A*, 97:062304, Jun 2018.

[196] Michael JW Hall. Gaussian noise and quantum-optical communication. *Phys. Rev. A*, 50(4):3295, 1994.

[197] Christopher Bishop. *Pattern recognition and machine learning*. Springer, New York, 2006. ISBN 0-387-31073-8.

[198] Christoph Buchheim, Ruth Hübner, and Anita Schöbel. Ellipsoid bounds for convex quadratic integer programming. *SIAM J. Optim.*, 25(2):741–769, January 2015.

[199] Jaehyun Park and Stephen Boyd. A semidefinite programming method for integer convex quadratic minimization. *Optim. Lett.*, 12(3):499–518, March 2017.

[200] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. Topological quantum memory. *J. Math. Phys.*, 43(9):4452–4505, Sep 2002.

[201] David S. Wang, Austin G. Fowler, and Lloyd C. L. Hollenberg. Surface code quantum computing with error rates over 1%. *Phys. Rev. A*, 83:020302, Feb 2011.

[202] Jack Edmonds. Optimum branchings. *J. Res. Natl. Bur. Stand. Sect. B Math. Math. Phys.*, 71B(4): 233, oct 1967.

[203] John Preskill. Reliable Quantum Computers. *Proc. R. Soc. Lond. A*, 454:385–410, 1997.

[204] Daniel Gottesman. Stabilizer codes and quantum error correction. *arXiv preprint quant-ph/9705052*, 1997.

[205] A Yu Kitaev. Quantum computations: algorithms and error correction. *Russ. Math. Surv.*, 52(6): 1191–1249, dec 1997.

[206] Panos Aliferis, Daniel Gottesman, and John Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Inf. Comput.*, 6(2):097–165, 2006.

[207] David P. Divincenzo and Panos Aliferis. Effective fault-tolerant quantum computation with slow measurements. *Phys. Rev. Lett.*, 98(2):020501, 2007.

[208] Panos Aliferis and John Preskill. Fault-tolerant quantum computation against biased noise. *Phys. Rev. A*, 78(5):052331, 2008.

[209] Panos Aliferis, Daniel Gottesman, and John Preskill. Accuracy threshold for postselected quantum computation. *Quantum Inf. Comput.*, 8(3-4):181–244, mar 2008.

[210] Alexey A. Kovalev and Leonid P. Pryadko. Fault tolerance of quantum low-density parity check codes with sublinear distance scaling. *Phys. Rev. A*, 87(2):020304, feb 2013.

[211] Daniel Gottesman. Fault-tolerant quantum computation with constant overhead. *Quantum Inf. Comput.*, 14(15-16):1338–1372, oct 2014.

[212] Omar Fawzi, Antoine Grospellier, and Anthony Leverrier. Constant overhead quantum fault-tolerance with quantum expander codes. In *Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS*, volume 2018-Octob, pages 743–754. IEEE, oct 2018.

[213] Barbara M. Terhal. Quantum error correction for quantum memories. *Rev. Mod. Phys.*, 87:307–346, Apr 2015.

[214] D. S. Wang, A. G. Fowler, A. M. Stephens, and L. C.L. Hollenberg. Threshold error rates for the toric and planar codes. *Quantum Inf. Comput.*, 10(5-6):456–469, may 2010.

[215] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numer. Math.*, 1(1):269–271, December 1959.

[216] J Harrington. *Analysis of quantum error-correcting codes: symplectic lattice codes and toric codes*. PhD thesis, Caltech, 2004.

[217] Christian D. Lorenz and Robert M. Ziff. Precise determination of the bond percolation thresholds and finite-size scaling corrections for the sc, fcc, and bcc lattices. *Phys. Rev. E*, 57:230–236, Jan 1998.

[218] Yong Zhang, M Menotti, K Tan, VD Vaidya, DH Mahler, L Zatti, M Liscidini, B Morrison, and Z Vernon. Single-mode quadrature squeezing using dual-pump four-wave mixing in an integrated nanophotonic device. *arXiv preprint arXiv:2001.09474*, 2020.

[219] Daiqin Su, Krishna Kumar Sabapathy, Casey R. Myers, Haoyu Qi, Christian Weedbrook, and Kamil Brádler. Implementing quantum algorithms on temporal photonic cluster states. *Phys. Rev. A*, 98: 032316, Sep 2018.

[220] Mohsen Razavi, Anthony Leverrier, Xiongfeng Ma, Bing Qi, and Zhiliang Yuan. Quantum key distribution and beyond: introduction. *Journal of the Optical Society of America B*, 36(3):QKD1, March 2019.

[221] Bing Qi. Single-photon continuous-variable quantum key distribution based on the energy-time uncertainty relation. *Optics letters*, 31(18):2795–2797, 2006.

[222] Bing Qi. Quantum key distribution based on frequency-time coding: security and feasibility. *arXiv preprint arXiv:1101.5995*, 2011.

[223] J. Nunn, L. Wright, C. Söller, L. Zhang, I. A. Walmsley, and B. J. Smith. Large-Alphabet Time-Frequency Entangled Quantum Key Distribution by means of Time-to-Frequency Conversion. *Optics Express*, 21(13):15959, jul 2013.

[224] Catherine Lee, Zheshen Zhang, Gregory R. Steinbrecher, Hongchao Zhou, Jacob Mower, Tian Zhong, Ligong Wang, Xiaolong Hu, Robert D. Horansky, Varun B. Verma, Adriana E. Lita, Richard P. Mirin, Francesco Marsili, Matthew D. Shaw, Sae Woo Nam, Gregory W. Wornell, Franco N.C. Wong, Jeffrey H. Shapiro, and Dirk Englund. Entanglement-based quantum communication secured by nonlocal dispersion cancellation. *Physical Review A - Atomic, Molecular, and Optical Physics*, 90(6):1–6, 2014.

[225] Tian Zhong, Hongchao Zhou, Robert D Horansky, Catherine Lee, Varun B Verma, Adriana E Lita, Alessandro Restelli, Joshua C Bienfang, Richard P Mirin, Thomas Gerrits, et al. Photon-efficient quantum key distribution using time–energy entanglement with high-dimensional encoding. *New Journal of Physics*, 17(2):022002, 2015.

[226] Nurul T. Islam, Charles Ci Wen Lim, Clinton Cahall, Jungsang Kim, and Daniel J. Gauthier. Provably secure and high-rate quantum key distribution with time-bin qudits. *Science Advances*, 3(11), 2017.

[227] Nurul T. Islam, Clinton Cahall, Andrés Aragoneses, A. Lezama, Jungsang Kim, and Daniel J. Gauthier. Robust and Stable Delay Interferometers with Application to d -Dimensional Time-Frequency Quantum Key Distribution. *Physical Review Applied*, 7(4):1–12, 2017.

[228] Thomas Brougham, Stephen M. Barnett, Kevin T. McCusker, Paul G. Kwiat, and Daniel J. Gauthier. Security of high-dimensional quantum key distribution protocols using Franson interferometers. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 46(10), 2013.

[229] Darius Bunandar, Zheshen Zhang, Jeffrey H. Shapiro, and Dirk R. Englund. Practical high-dimensional quantum key distribution with decoy states. *Physical Review A - Atomic, Molecular, and Optical Physics*, 91(2), 2015.

[230] Murphy Yuezhen Niu, Feihu Xu, Jeffrey H. Shapiro, and Fabian Furrer. Finite-key analysis for time-energy high-dimensional quantum key distribution. *Physical Review A*, 94(5), 2016.

[231] Jacob Mower, Zheshen Zhang, Pierre Desjardins, Catherine Lee, Jeffrey H. Shapiro, and Dirk Englund. High-dimensional quantum key distribution using dispersive optics. *Physical Review A - Atomic, Molecular, and Optical Physics*, 87(6), 2013.

[232] Zheshen Zhang, Jacob Mower, Dirk Englund, Franco N C Wong, and Jeffrey H. Shapiro. Unconditional security of time-energy entanglement quantum key distribution using dual-basis interferometry. *Physical Review Letters*, 112(12), 2013.

[233] Alicia Sit, Frédéric Bouchard, Robert Fickler, Jérémie Gagnon-Bischoff, Hugo Larocque, Khabat Heshami, Dominique Elser, Christian Peuntinger, Kevin Günthner, Bettina Heim, Christoph Marquardt, Gerd Leuchs, Robert W. Boyd, and Ebrahim Karimi. High-dimensional intracity quantum cryptography with structured photons. *Optica*, 4(9):1006, August 2017.

[234] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and Ph. Grangier. Virtual Entanglement and Reconciliation Protocols for Quantum Cryptography with Continuous Variables. *Quantum Information and Computation*, 3(7):535–552, 2003.

[235] F. Furrer, T. Franz, M. Berta, A. Leverrier, V. B. Scholz, M. Tomamichel, and R. F. Werner. Continuous variable quantum key distribution: Finite-key analysis of composable security against coherent attacks. *Physical Review Letters*, 109(10):1–5, 2012.

[236] Fabian Laudenbach, Christoph Pacher, Chi-Hang Fred Fung, Andreas Poppe, Momtchil Peev, Bernhard Schrenk, Michael Hentschel, Philip Walther, and Hannes Hübel. Continuous-Variable Quantum Key Distribution with Gaussian Modulation – The Theory of Practical Implementations. *Adv. Quantum Technol.*, 1(1):1–37, 2018.

[237] Tobias Gehring, Vitus Händchen, Jörg Duhme, Fabian Furrer, Torsten Franz, Christoph Pacher, Reinhard F. Werner, and Roman Schnabel. Implementation of continuous-variable quantum key distribution with composable and one-sided-device-independent security against coherent attacks. *Nature Communications*, 6:1–7, 2015.

[238] Hans Maassen and J. B M Uffink. Generalized entropic uncertainty relations. *Physical Review Letters*, 60(12):1103–1106, 1988.

[239] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M. Renes, and Renato Renner. The uncertainty principle in the presence of quantum memory. *Nature Physics*, 6(9):659–662, 2010.

[240] Robert König, Renato Renner, and Christian Schaffner. The operational meaning of min-and max-entropy. *IEEE Transactions on Information Theory*, 55(9):11, 2009.

[241] Marco Tomamichel and Renato Renner. Uncertainty relation for smooth entropies. *Physical Review Letters*, 106(11):1–4, 2011.

[242] Marco Tomamichel. *A Framework for Non-Asymptotic Quantum Information Theory*. PhD thesis, ETH Zurich, 2012.

[243] Patrick J. Coles, Roger Colbeck, Li Yu, and Michael Zwolak. Uncertainty relations from simple entropic properties. *Physical Review Letters*, 108(21):1–5, 2012.

[244] Patrick J. Coles and Marco Piani. Improved entropic uncertainty relations and information exclusion relations. *Physical Review A - Atomic, Molecular, and Optical Physics*, 89(2):1–11, 2014.

[245] Renato Renner. *Security of Quantum Key Distribution*. PhD thesis, ETH Zurich, 2005.

[246] Anthony Leverrier, Frédéric Grosshans, and Philippe Grangier. Finite-size analysis of a continuous-variable quantum key distribution. *Physical Review A - Atomic, Molecular, and Optical Physics*, 81(6): 1–11, 2010.

[247] Marco Tomamichel, Charles Ci Wen Lim, Nicolas Gisin, and Renato Renner. Tight finite-key analysis for quantum cryptography. *Nature Communications*, 3(May 2011):634–636, 2012.

[248] Charles H. Bennett, Gilles Brassard, and N. David Mermin. Quantum cryptography without Bell's theorem. *Physical Review Letters*, 68(5):557–559, 1992.

[249] Megan R. Ray and S. J. Van Enk. Missing data outside the detector range: Continuous-variable entanglement verification and quantum cryptography. *Physical Review A - Atomic, Molecular, and Optical Physics*, 88(4):1–6, 2013.

[250] Megan R. Ray and S. J. Van Enk. Missing data outside the detector range. II. Application to time-frequency entanglement. *Physical Review A - Atomic, Molecular, and Optical Physics*, 88(6):2–7, 2013.

[251] Fabricio Toscano, Daniel S. Tasca, Łukasz Rudnicki, and Stephen P. Walborn. Uncertainty relations for coarse-grained measurements: An overview. *Entropy*, 20(6):1–36, 2018.

[252] Hao Qin, Rupesh Kumar, and Romain Alléaume. Saturation attack on continuous-variable quantum key distribution system. *Proc. SPIE*, 8899:88990N, 2013.

[253] Hao Qin, Rupesh Kumar, Vadim Makarov, and Romain Alléaume. Homodyne-detector-blinding attack in continuous-variable quantum key distribution. *Physical Review A*, 98(1):1–13, 2018.

[254] Rupert L. Frank and Elliott H. Lieb. Monotonicity of a relative Rényi entropy. *Journal of Mathematical Physics*, 54(12), 2013.

[255] S J Van Enk. Photodetector figures of merit in terms of POVMs. *Journal of Physics Communications*, 1(4):045001, Aug 2017.

[256] Aukasz Rudnicki, Stephen P. Walborn, and Fabricio Toscano. Optimal uncertainty relations for extremely coarse-grained measurements. *Physical Review A - Atomic, Molecular, and Optical Physics*, 85(4):1–9, 2012.

[257] Changjia Chen, Eric Y. Zhu, Arash Riazi, Alexey V. Gladyshev, Costantino Corbari, Morten Ibsen, Peter G. Kazansky, and Li Qian. Compensation-free broadband entangled photon pair sources. *Optics Express*, 25(19):22667, 2017.

[258] F. Marsili, V. B. Verma, J. A. Stern, S. Harrington, A. E. Lita, T. Gerrits, I. Vayshenker, B. Baek, M. D. Shaw, R. P. Mirin, and S. W. Nam. Detecting single infrared photons with 93% system efficiency. *Nature Photonics*, 7(3):210–214, 2013.

[259] Walter Rudin. *Real and Complex Analysis, International Edition*. McGraw-Hill, New York, 1987.

[260] Cleve Moler. Floating points: Ieee standard unifies arithmetic model. *Cleve's Corner, The MathWorks, Inc*, 1996.

[261] Yi Zhao, Bing Qi, Hoi-Kwong Lo, and Li Qian. Security analysis of an untrusted source for quantum key distribution: passive approach. *New Journal of Physics*, 12(2):023024, feb 2010.

[262] Tian Zhong, Franco N. C. Wong, Alessandro Restelli, and Joshua C. Bienfang. Efficient single-spatial-mode periodically-poled $KTiOPO_4$ waveguide source for high-dimensional entanglement-based quantum key distribution. *Optics Express*, 20(24):26868, Nov 2012.

[263] Norbert Lütkenhaus. Security against individual attacks for realistic quantum key distribution. *Phys. Rev. A*, 61:052304, Apr 2000.

[264] Christian Weedbrook, Andrew M. Lance, Warwick P. Bowen, Thomas Symul, Timothy C. Ralph, and Ping Koy Lam. Quantum cryptography without switching. *Phys. Rev. Lett.*, 93:170504, Oct 2004.

[265] Anthony Leverrier, Raúl García-Patrón, Renato Renner, and Nicolas J. Cerf. Security of continuous-variable quantum key distribution against general attacks. *Phys. Rev. Lett.*, 110:030502, Jan 2013.

[266] Anthony Leverrier. Composable security proof for continuous-variable quantum key distribution with coherent states. *Phys. Rev. Lett.*, 114:070501, Feb 2015.

[267] Anthony Leverrier. Security of continuous-variable quantum key distribution via a gaussian de finetti reduction. *Phys. Rev. Lett.*, 118:200501, May 2017.

[268] Moritz Mehmet, Stefan Ast, Tobias Eberle, Sebastian Steinlechner, Henning Vahlbruch, and Roman Schnabel. Squeezed light at 1550 nm with a quantum noise reduction of 12.3 dB. *Optics Express*, 19 (25):25763–25772, 2011.

[269] Feihu Xu, Xiongfeng Ma, Qiang Zhang, Hoi-Kwong Lo, and Jian-Wei Pan. Secure quantum key distribution with realistic devices. *Rev. Mod. Phys.*, 92:025002, May 2020.

[270] Samuel L. Braunstein and Stefano Pirandola. Side-channel-free quantum key distribution. *Phys. Rev. Lett.*, 108:130502, Mar 2012.

[271] Won-Young Hwang. Quantum key distribution with high loss: Toward global secure communication. *Phys. Rev. Lett.*, 91:057901, Aug 2003.

[272] Hoi-Kwong Lo, Xiongfeng Ma, and Kai Chen. Decoy state quantum key distribution. *Phys. Rev. Lett.*, 94:230504, Jun 2005.

[273] D Gottesman, Hoi-Kwong Lo, Norbert Lütkenhaus, and John Preskill. Security of quantum key distribution with imperfect devices. *Quantum Information and Computation*, 4:136, 01 2003.

[274] Margarida Pereira, Marcos Curty, and Kiyoshi Tamaki. Quantum key distribution with flawed and leaky sources. *npj Quantum Information*, 5(1), July 2019.

[275] Weilong Wang, Kiyoshi Tamaki, and Marcos Curty. Measurement-device-independent quantum key distribution with leaky sources. *Scientific reports*, 11(1):1–11, 2021.

[276] Álvaro Navarrete, Margarida Pereira, Marcos Curty, and Kiyoshi Tamaki. Practical quantum key distribution that is secure against side channels. *Phys. Rev. Applied*, 15:034072, Mar 2021.

[277] Zhen-Qiang Yin, Chi-Hang Fred Fung, Xiongfeng Ma, Chun-Mei Zhang, Hong-Wei Li, Wei Chen, Shuang Wang, Guang-Can Guo, and Zheng-Fu Han. Measurement-device-independent quantum key distribution with uncharacterized qubit sources. *Phys. Rev. A*, 88:062322, Dec 2013.

[278] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053): 207–235, 2005.

[279] Patrick J. Coles, Eric M. Metodiev, and Norbert Lütkenhaus. Numerical approach for unstructured quantum key distribution. *Nature Communications*, 7(1), May 2016.

[280] Adam Winick, Norbert Lütkenhaus, and Patrick J. Coles. Reliable numerical key rates for quantum key distribution. *Quantum*, 2:77, July 2018.

[281] Hoi-Kwong Lo. Proof of unconditional security of six-state quantum key distribution scheme. *Quantum Info. Comput.*, 1(2):81–94, August 2001.

[282] Zhiyuan Tang, Kejin Wei, Olinka Bedroya, Li Qian, and Hoi-Kwong Lo. Experimental measurement-device-independent quantum key distribution with imperfect sources. *Phys. Rev. A*, 93:042308, Apr 2016.

[283] Feihu Xu, Kejin Wei, Shihan Sajeed, Sarah Kaiser, Shihai Sun, Zhiyuan Tang, Li Qian, Vadim Makarov, and Hoi-Kwong Lo. Experimental quantum key distribution with source flaws. *Phys. Rev. A*, 92:032305, Sep 2015.

[284] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim. Unconditional privacy over channels which cannot convey quantum information. *Phys. Rev. Lett.*, 100:110502, Mar 2008.

[285] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim. General paradigm for distilling classical key from quantum states. *IEEE Transactions on Information Theory*, 55(4):1898–1929, 2009.

[286] Joseph M. Renes and Graeme Smith. Noisy processing and distillation of private quantum states. *Phys. Rev. Lett.*, 98:020502, Jan 2007.

[287] K. Horodecki, M. Horodecki, P. Horodecki, D. Leung, and J. Oppenheim. Quantum key distribution based on private states: Unconditional security over untrusted channels with zero quantum capacity. *IEEE Transactions on Information Theory*, 54(6):2604–2620, 2008.

[288] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004. ISBN 0521833787.

[289] Nurul T. Islam, Charles Ci Wen Lim, Clinton Cahall, Jungsang Kim, and Daniel J. Gauthier. Securing quantum key distribution systems using fewer states. *Phys. Rev. A*, 97:042347, Apr 2018.

[290] Yukun Wang, Ignatius William Primaatmaja, Emilien Lavie, Antonios Varvitsiotis, and Charles Ci Wen Lim. Characterising the correlations of prepare-and-measure quantum networks. *npj Quantum Information*, 5(1), February 2019.

[291] Nurul T Islam, Charles Ci Wen Lim, Clinton Cahall, Bing Qi, Jungsang Kim, and Daniel J Gauthier. Scalable high-rate, high-dimensional time-bin encoding quantum key distribution. *Quantum Science and Technology*, 4(3):035008, June 2019.

[292] Li Liu, Yukun Wang, Emilien Lavie, Chao Wang, Arno Ricou, Fen Zhuo Guo, and Charles Ci Wen Lim. Practical quantum key distribution with non-phase-randomized coherent states. *Phys. Rev. Applied*, 12: 024048, Aug 2019.

[293] Ernest Y.-Z. Tan, René Schwonnek, Koon Tong Goh, Ignatius William Primaatmaja, and Charles C-W Lim. Computing secure key rates for quantum key distribution with untrusted devices. *arXiv preprint arXiv:1908.11372*, 2019.

[294] Margarida Pereira, Go Kato, Akihiro Mizutani, Marcos Curty, and Kiyoshi Tamaki. Quantum key distribution with correlated sources. *Science Advances*, 6(37):eaaz4487, 2020.

[295] Le Wang, Sheng-Mei Zhao, Long-Yan Gong, and Wei-Wen Cheng. Free-space measurement-device-independent quantum-key-distribution protocol using decoy states with orbital angular momentum. *Chinese Physics B*, 24(12):120307, December 2015.

[296] Xing-Yu Wang, Shang-Hong Zhao, Chen Dong, Zhuo-Dan Zhu, and Wen-Yuan Gu. Orbital angular momentum-encoded measurement device independent quantum key distribution under atmospheric turbulence. *Quantum Information Processing*, 18(10), August 2019.

[297] Yan-Lin Tang, Hua-Lei Yin, Si-Jing Chen, Yang Liu, Wei-Jun Zhang, Xiao Jiang, Lu Zhang, Jian Wang, Li-Xing You, Jian-Yu Guan, Dong-Xu Yang, Zhen Wang, Hao Liang, Zhen Zhang, Nan Zhou, Xiongfeng Ma, Teng-Yun Chen, Qiang Zhang, and Jian-Wei Pan. Measurement-device-independent quantum key distribution over 200 km. *Physical Review Letters*, 113(19), November 2014.

[298] Peter W. Shor and John Preskill. Simple proof of security of the bb84 quantum key distribution protocol. *Phys. Rev. Lett.*, 85:441–444, Jul 2000.

[299] Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016.

[300] Akshay Agrawal, Robin Verschueren, Steven Diamond, and Stephen Boyd. A rewriting system for convex optimization problems. *Journal of Control and Decision*, 5(1):42–60, 2018.

[301] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014.

[302] Kiyoshi Tamaki, Marcos Curty, and Marco Lucamarini. Decoy-state quantum key distribution with a leaky source. *New Journal of Physics*, 18(6):065008, 2016.

[303] I Lucio-Martinez, Philip Chan, Xiaofan Mo, Steve Hosier, and Wolfgang Tittel. Proof-of-concept of real-world quantum key distribution with quantum frames. *New Journal of Physics*, 11(9):095001, 2009.

[304] Zhiyuan Tang, Zhongfa Liao, Feihu Xu, Bing Qi, Li Qian, and Hoi-Kwong Lo. Experimental demonstration of polarization encoding measurement-device-independent quantum key distribution. *Physical review letters*, 112(19):190503, 2014.

[305] Jindong Wang, Xiaojuan Qin, Yinzhu Jiang, Xiaojing Wang, Liwei Chen, Feng Zhao, Zhengjun Wei, and Zhiming Zhang. Experimental demonstration of polarization encoding quantum key distribution system based on intrinsically stable polarization-modulated units. *Optics express*, 24(8):8302–8309, 2016.

[306] Chenyang Li, Marcos Curty, Feihu Xu, Olinka Bedroya, and Hoi-Kwong Lo. Secure quantum communication in the presence of phase-and polarization-dependent loss. *Physical Review A*, 98(4):042324, 2018.

[307] Eleftherios Moschandreou, Brian J Rollick, Bing Qi, and George Siopsis. Experimental decoy-state bennett-brassard 1984 quantum key distribution through a turbulent channel. *Physical Review A*, 103 (3):032614, 2021.

[308] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy. Trojan-horse attacks on quantum-key-distribution systems. *Physical Review A*, 73(2), February 2006.

[309] M. Lucamarini, I. Choi, M. B. Ward, J. F. Dynes, Z. L. Yuan, and A. J. Shields. Practical security bounds against the trojan-horse attack in quantum key distribution. *Phys. Rev. X*, 5:031030, Sep 2015.

[310] Chi-Hang Fred Fung and Hoi-Kwong Lo. Security proof of a three-state quantum-key-distribution protocol without rotational symmetry. *Physical Review A*, 74(4):042342, 2006.

[311] Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and con tos5. In *Proceedings of the International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.

[312] Xiongfeng Ma, Chi-Hang Fred Fung, and Mohsen Razavi. Statistical fluctuation analysis for measurement-device-independent quantum key distribution. *Phys. Rev. A*, 86:052305, Nov 2012.

[313] Gong Zhang, Ignatius William Primaatmaja, Jing Yan Haw, Xiao Gong, Chao Wang, and Charles Ci Wen Lim. Securing practical quantum cryptosystems with optical power limiters. *arXiv preprint arXiv:2012.08702*, 2020.

[314] Ian George, Jie Lin, and Norbert Lütkenhaus. Numerical calculations of the finite key rate for general quantum key distribution protocols. *Physical Review Research*, 3(1):013274, 2021.

[315] Chunyuan Zhou, Guang Wu, Xiuliang Chen, and Heping Zeng. "plug and play" quantum key distribution system with differential phase shift. *Applied Physics Letters*, 83(9):1692–1694, September 2003.

[316] Xiaoqing Zhong, Jianyong Hu, Marcos Curty, Li Qian, and Hoi-Kwong Lo. Proof-of-principle experimental demonstration of twin-field type quantum key distribution. *Physical Review Letters*, 123(10), September 2019.

[317] Jacob Hastrup and Ulrik L Andersen. Generation of optical gottesman-kitaev-preskil states with cavity qed. *arXiv preprint arXiv:2104.07981*, 2021.

[318] Marco Lucamarini, Zhiliang L Yuan, James F Dynes, and Andrew J Shields. Overcoming the rate–distance limit of quantum key distribution without quantum repeaters. *Nature*, 557(7705):400–403, 2018.

[319] Stefano Biagi, Andrea Bonfiglioli, and Marco Matone. On the baker-campbell-hausdorff theorem: non-convergence and prolongation issues. *Linear and Multilinear Algebra*, 0(0):1–19, 2018.

[320] Francesco Albarelli, Marco G. Genoni, Matteo G.A. Paris, and Alessandro Ferraro. Resource theory of quantum non-Gaussianity and Wigner negativity. *Phys. Rev. A*, 98(5):052350, nov 2018.

[321] T Bröcker and RF Werner. Mixed states with positive wigner functions. *Journal of Mathematical Physics*, 36(1):62–75, 1995.

[322] A Ketterer, A Keller, SP Walborn, T Coudreau, and P Milman. Quantum information processing in phase space: A modular variables approach. *Phys. Rev. A*, 94(2):022325, 2016.

[323] Jens Eisert. *Discrete Quantum States versus Continuous Variables*, chapter 3, pages 39–54. John Wiley & Sons, Ltd, 2016. ISBN 9783527805785.

[324] DB Horoshko, Stephan De Bièvre, Giuseppe Patera, and MI Kolobov. Thermal-difference states of light: Quantum states of heralded photons. *Phys. Rev. A*, 100(5):053831, 2019.

[325] Oliver F Thomas, Will McCutcheon, and Dara PS McCutcheon. A general framework for multimode gaussian quantum optics and photo-detection: Application to hong–ou–mandel interference with filtered heralded single photon sources. *APL Photonics*, 6(4):040801, 2021.

[326] Brajesh Gupt, Juan Miguel Arrazola, Nicolás Quesada, and Thomas R Bromley. Classical benchmarking of gaussian boson sampling on the titan supercomputer. *Quantum Inf. Process.*, 19(8):1–14, 2020.

[327] F Gustav Mehler. Ueber die entwicklung einer function von beliebig vielen variablen nach laplaceschen functionen höherer ordnung. *J. Reine Angew. Math.*, 1866(66):161–176, 1866.

[328] Timjan Kalajdzievski and Nicolás Quesada. Exact and approximate continuous-variable gate decompositions. *Quantum*, 5:394, 2021.

[329] Ryuji Ukai, Shota Yokoyama, Jun-ichi Yoshikawa, Peter van Loock, and Akira Furusawa. Demonstration of a controlled-phase gate for continuous-variable one-way quantum computation. *Phys. Rev. Lett.*, 107: 250501, Dec 2011.

[330] Jun-ichi Yoshikawa, Yoshichika Miwa, Alexander Huck, Ulrik L. Andersen, Peter van Loock, and Akira Furusawa. Demonstration of a quantum nondemolition sum gate. *Phys. Rev. Lett.*, 101:250501, Dec 2008.

[331] Filippo Caruso, Vittorio Giovannetti, and Alexander S Holevo. One-mode bosonic Gaussian channels: a full weak-degradability classification. *New J. Phys.*, 8(12):310, 2006.

[332] Tommaso F. Demarie, Trond Linjordet, Nicolas C. Menicucci, and Gavin K. Brennen. Detecting topological entanglement entropy in a lattice of quantum harmonic oscillators. *New J. Phys.*, 16(8): 085011, 2014.

[333] Matteo G.A. Paris. Displacement operator by beam splitter. *Phys. Lett. A*, 217(2-3):78–80, July 1996.

[334] Mhlambululi Mafu, Angela Dudley, Sandeep Goyal, Daniel Giovannini, Melanie McLaren, Miles J. Padgett, Thomas Konrad, Francesco Petruccione, Norbert Lütkenhaus, and Andrew Forbes. Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases. *Physical Review A*, 88(3), September 2013.

[335] Daniele Cozzolino, Davide Bacco, Beatrice Da Lio, Kasper Ingerslev, Yunhong Ding, Kjeld Dalgaard, Poul Kristensen, Michael Galili, Karsten Rottwitt, Siddharth Ramachandran, and Leif Katsuo Oxenløwe. Orbital angular momentum states enabling fiber-based high-dimensional quantum communication. *Physical Review Applied*, 11(6), June 2019.

[336] Lana Sheridan and Valerio Scarani. Security proof for quantum key distribution using qudit systems. *Phys. Rev. A*, 82:030301, Sep 2010.